CAFRAL
CENTRE FOR ADVANCED FINANCIAL RESEARCH AND LEARNING
Promoted by Reserve Bank of India

CAFRAL RESEARCH ECONOMIC INSIGHTS: WORKING PAPER

2024

CAFRAL
CENTRE FOR ADVANCED FINANCIAL RESEARCH AND LEARNING
Promoted by Reserve Bank of India

CAFRAL RESEARCH ECONOMIC INSIGHTS: WORKING PAPERS

2024

DIRECTOR'S NOTE

I am pleased to introduce the inaugural series of research working papers by the Centre for Advanced Financial Research and Learning (CAFRAL), an esteemed institution established by the Reserve Bank of India (RBI). CAFRAL's inception is rooted in the recognition of the need to conduct high-quality research that can help shape policy debates in India. With a mandate to become a world-class global institution for research and learning in macroeconomics, finance, and public policy, CAFRAL stands as a beacon of excellence and innovation.

This collection of working papers reflects CAFRAL's commitment to advancing knowledge and fostering a deeper understanding of India's economy. As an independent body under the auspices of the RBI, CAFRAL is uniquely positioned to contribute to scholarly research that is not only relevant to the Indian context but also resonates globally. These working papers showcase CAFRAL's research agenda, covering topics relevant to economic policy. The contributions in this series signify CAFRAL's dedication to becoming a hub for cutting-edge research, bridging the gap between theory and practice, and providing valuable insights for policymakers, practitioners, and researchers alike.

As we embark on this journey with the first installment of the working paper series, we anticipate that this initiative will lay the groundwork for a continuous stream of scholarly contributions. CAFRAL's vision is to establish a dynamic platform for research and learning, fostering collaboration and dialogue among experts in the field. We aspire for this series to be a catalyst for future endeavors, promoting a culture of inquiry and academic excellence. I hope this series becomes a testament to CAFRAL's commitment to excellence in research and learning, and inspires and guides future research on solving India's economic challenges.

Bibhu Prasad Kanungo
Director, CAFRAL
May 2, 2024

FOREWORD

It is a privilege to introduce the inaugural series of research working papers by the Centre for Advanced Financial Research and Learning (CAFRAL), an institution promoted by the Reserve Bank of India as a house of intellectual excellence. This collection signifies a milestone in our commitment to advancing knowledge in banking, finance, macroeconomics, trade, and public policy. It aims to address critical knowledge gaps essential for informed decision-making in the context of India's unique economic landscape.

In the ever-changing global economic scenario, a nuanced understanding of the forces shaping our macroeconomy and financial systems is paramount. CAFRAL seeks to be a torchbearer in this pursuit, contributing to scholarly discourse through rigorous research that is topical, relevant and tailored to the specific socio-economic characteristics of our nation so that the formulation and implementation of public policies is well informed. These working papers exemplify CAFRAL's dedication to advancing knowledge in areas crucial to India's economic development.

The contributions in this series reflect research undertaken by CAFRAL scholars and cover a rich set of topics such as the transmission of global price shocks to India, firm financing, sovereign defaults and their implications, trade policies, and firm productivity. This volume commences a journey that shines light on CAFRAL's endeavour to establish itself as a hub for interdisciplinary research, bridging critical gaps in policy-oriented research and facilitating evidence-based policymaking.

This working paper series is envisioned as a platform for continual dissemination of applied academic knowledge that deepens understanding of India's economic systems. The richness and diversity of these papers lay the foundation for future research endeavour aimed at fostering collaborations and inspiring scholars to delve deeper into the complexities of India's economy.

I commend the authors for their insightful contributions which is inclined to engage the wider academic community in a vibrant intellectual exchange, fuelling a continuous dialogue on the economic challenges and opportunities that define our nation's developmental trajectory.

Michael Debabrata Patra
Deputy Governor, Reserve Bank of India
May 2, 2024

CONTENTS

FOOD, FUEL, AND FACTS: DISTRIBUTIONAL EFFECTS OF GLOBAL PRICE SHOCKS

Saroj Bhattarai* **Arpita Chatterjee†** **Gautham Udupa‡**

January 4, 2024

We show that exogenous global commodity price shocks lead to a significant decline over time in Indian household consumption. These negative effects are heterogeneous along the income distribution: households in lower income groups experience more adverse consumption effects following an exogenous rise in food prices, whereas households in the lowest and the two highest income groups are affected similarly following an exogenous rise in oil prices. We investigate how income and relative price changes contribute to generating these heterogeneous consumption effects. Global food price shocks lead to significant negative wage income effects that mirror the pattern of negative consumption effects along the income distribution. Both global oil and food price shocks pass-through to local consumer prices in India and increase the relative prices of fuel and food respectively. Expenditure share of food increases with such a rise in relative prices, which provides unambiguous evidence for non-homothetic preferences. Specifically, we show that food, compared to fuel, is an essential consumption good for all income groups in India.

JEL Codes: F41, F62, O11

Keywords: Global Price shocks; Food prices; Oil prices; Inequality; Household heterogeneity; Household consumption; Essential consumption good; Non-homotheticity; India

*Department of Economics, University of Texas at Austin, U.S.A. Email: saroj.bhattarai@austin.utexas.edu.

†Department of Economics, IIM-Bangalore, India and University of New South Wales, Australia. Email: chatterjee.econ@gmail.com.

‡CAFRAL, Reserve Bank of India, India. Email: gautham.udupa@cafral.org.in.

Non-Technical Summary

The global economy has gone through turbulent times in the last few years. The fiscal and monetary stimulation during the pandemic, as well as the Russia-Ukraine war that followed it, has put considerable pressure on commodity prices globally. For example, the composite food price index tracked by United Nation's Food and Agriculture Organization rose by 8 per cent year-on-year in August 2022, whereas the West Texas Industrial crude oil price rose 23 percent in the same month compared to a year ago. The inflation in consumer prices that accompanied the commodity price inflation was not seen in the developed economies since the global financial crisis in 2008.

From the developing economies' perspective, there have been concerns about their susceptibility to the global shocks. Different factors such as exchange rate volatility, as well as impact on household consumption due to lack of adequate savings and access to credit, have been flagged as reasons why developing economies are particularly susceptible to global commodity price inflation. In particular poorer households for whom food and fuel could be essentials with very few substitution possibilities, may face more adverse consequences. And because in emerging economies poorer households are a larger share of the population, these distributional effects may have larger aggregate demand consequences. Such adverse consequences are cited as the reason for urgent monetary policy actions to curb inflation.

This paper is the first to measure the distributional consequences of global commodity prices on household consumption in a major developing economy – India. It documents differential responses to global food and fuel price fluctuations across the rich and poorer segments of the population. A rise in global food prices leads to a fall in consumption for all households, and the effects are larger for poorer households. Effects are also economically meaningful: average food price rise of 3 per cent leads to more than 4 per cent fall in non-durable consumption for the poorest households, more than 3 per cent for the middle income groups and more than 2.5 per cent for richer households. The food inflation following the Russia-Ukraine war lead to a consumption loss of 12 per cent for the poorer households while for the middle income groups itâs close to 10 per cent.

A rise in global oil prices have a more nuanced effect. Both ends of the income distribution, i.e., the poorest and the two richest income groups, are vulnerable to oil price increases. A 10 per cent inflation leads to a nearly 2 per cent fall in non-durable consumption loss for both these groups.

The paper identifies local wage earnings and local prices as the key channels through which the global commodity price inflation impact local consumption. The pattern of effects on wage earnings mirror that of consumption: while all households show a fall in wage earnings due to rise in global food prices, the effects are monotonically larger on poorer households. The rise in oil prices erode the earnings on the poor and the rich, with no such statistically discernible effects on the middle income groups.

Global price shocks pass-through to local prices. Global food price inflation raises the food component of CPI, which in turn affects headline CPI, and makes food and its various subcomponents relatively more expensive compared to fuel. Global oil prices raise local fuel prices. It also raises core CPI especially in urban areas and both of these lead to a rise in overall CPI. The pass-through magnitudes are larger for global food inflation, but are more broad-based for global fuel inflation.

The rest of the paper documents that standard demand structures used in the business cycles literature are at odds with the estimated responses. Importantly, relative expenditure on food shows strong patterns on non-homotheticity, wherein consumption share of food are is a constant fraction of total consumption. A rise in global food prices raises relative prices of various subcomponents of food relative to fuel. Standard expenditure switching effects under demand implies a decline in relative expenditure on food. Both the overall food to fuel expenditure ratios, and particularly some sub-components such as sugar and pulses, contrary to predictions of the standard demand structure, register a rise in relative expenditure following an increase in global food prices. It suggests that these items are 'essentials' in consumption not just for the poor, but often even for the rich income groups.

1 Introduction

There has been a rapid increase in *global* oil and food prices recently. These large external shocks have raised major concerns worldwide, but especially so in emerging markets, whose economies tend to be more vulnerable to global shocks. For emerging markets, food and fuel price shocks affect livelihoods of a very large part of the population, which further increases their salience. Strong effects on inflation and cost-of-living are expected to hit these countries as a result of these external price increases, thereby exacerbating existing inequality.

Effects of such global price shocks driven inflation on macroeconomic outcomes and inequality have therefore been at the forefront of policy makers' agenda. For instance, in the April 2022 issue of the World Economic Outlook (WEO), the International Monetary Fund (IMF) states: *Fuel and food prices have increased rapidly, with vulnerable populations–particularly in low-income countries–most affected. ... Higher food prices will hurt consumers' purchasing power–particularly among low-income households–and weigh on domestic demand.* Moreover, with deteriorating conditions in food and energy markets, the IMF's stance is more grave in the July 2022 issue of the WEO: *Rising food and energy prices cause widespread hardship, famine, and unrest. Because energy and food are essential goods with few substitutes, higher prices are particularly painful for households. When the price of other items, such as electronics, furniture, or entertainment, increases, families can simply reduce or even eliminate spending on them. For food, heating, and transportation–often essential to earn a living, this is much harder.*

Despite previous research on the impact of such global shocks, particularly oil prices, on the overall macroeconomy, there exists limited rigorous evidence on the *distributional* consequences of such shocks, especially in emerging markets.This paper seeks to address this gap by examining the causal connection between rises in global food and fuel prices and consumption inequality in India, a major emerging economy that has experienced significant inflationary pressures in these sectors recently. To this end, we begin by presenting motivating evidence that shows a positive correlation between global food and oil price fluctuations and aggregate measures of consumption inequality in India. While this correlation is intriguing and suggestive, it alone does not establish causality or demonstrate that the effects of external price shocks are sustained over time. Moreover, it cannot help identify which segments of the population are more sensitive to food and fuel price shocks, which components of consumption constitute essential consumption goods for different income groups, or elucidate the economic mechanisms through which such global price shocks lead to consumption inequality in India.

To make headway on these questions, we utilize a comprehensive monthly household panel dataset from India that spans 2014-2019. Leveraging the panel dimension of the data, in a local projection framework at the household level, we investigate whether the dynamic effects on consumption of global oil and food price fluctuations differ along the income distribution. Our analysis involves categorizing households into five income brackets and estimating interaction effects between these groups and the global price shocks.

Furthermore, to ensure a more accurate and causal interpretation of our findings, we devise an instrumental variable (IV) strategy. Since we assume India to be a small open economy, we can regard changes in global oil and food prices as an external shock in our panel local projection exercise. However, as the literature on the macroeconomic impact of oil shocks suggests, separating out the effects of global demand (and commodity-specific demand) shocks from those of global supply shocks is essential for a clear interpretation of the results. Using changes in world oil and food prices as a measure of shock would thus produce OLS estimates that conflate the effects of both types of shocks. Moreover, global demand shocks can lead to omitted variable bias if they directly affect Indian household consumption through their exposure to the global business cycle.

To tackle this challenge, we employ an IV approach, using supply-side instruments for the change in global oil and food prices.[1] For the global oil price change, we use the oil supply shock estimated in Baumeister and Hamilton [2019] as an IV, while for the global food price change, we construct our own IV. The latter is based on residuals of food commodity prices, after extracting two common factors, a food-specific and an aggregate common factor. These factors are estimated by imposing sign restrictions in a dynamic factor model that uses panel data on commodity prices.

Our principal findings are as follows. Effects on consumption of the global price shocks are clearly adverse, and heterogeneously so, along the income distribution. Households in the lower income deciles bear a greater burden from an exogenous rise in food prices and the negative consumption effects become progressively less severe as we move up the income distribution. Exogenous rise in fuel prices, in contrast, affects consumption of both the lowest and the two higher income groups similarly.

Our IV estimates reveal substantive differences from the corresponding OLS estimates in

[1]Using an IV helps address any reverse causality concerns as well since India is growing fast and may have an impact on global demand.

some cases. For instance, there are discernible differences in the consumption responses of the highest income group when comparing the OLS and IV findings. While OLS results indicate an increase in consumption for this income group following an oil price increase, IV results exhibit a decrease instead. This finding is intuitive, as positive global demand shocks, which are a part of the OLS results, and which increase oil prices, are likely to benefit high-income households in India.[2]

We utilize the IV framework to further investigate the mechanisms that cause heterogeneous effects in consumption. First, we estimate heterogeneous wage earnings effects of the global price shocks. Our analysis reveals that food price shocks lead to a negative effect on real wage incomes of the poorest households. Moreover, this shock also affects non-trivially the labor earnings of other households, with the effects monotonically decreasing as we move up the income distribution. This suggests that food price shocks affect consumption heterogeneously through their differential effects on wage income. For oil price shocks, there are consistently negative labor income effects on the lowest and highest income groups, which is aligned with the negative consumption effects for these groups.

Second, state-level panel local projection IV results show that both global shocks exhibit a "pass-through" effect on local prices in India, affecting not just own-category prices but also overall (headline) prices. Both these shocks also affect relative prices: Global food price shocks elevate the relative price of food, while global oil price shocks drive up the relative price of fuel in India.[3]

Conventional homothetic demand functions would predict expenditure-switching effects due to such relative price effects. We, however, find strong evidence to the contrary. Specifically, we show that in response to the global food price shock, the food expenditure ratio clearly increases for the lower income groups. Given that the relative price of food increases with the global food price shock, these consumption share responses unambiguously suggest a role for income effects in relative demand. Moreover, using the responses of relative food prices, relative food expenditures, and real non-durable consumption expenditure together, we infer econometrically that food is an essential consumption good,

[2]We show direct evidence of this phenomenon by examining the effects on labor earnings. We observe positive impacts on labor income for the high-income group following an oil price increase in OLS estimates, which are absent in IV estimates. This is reminiscent of the key conclusion in Kilian [2009]. Generally, in OLS estimates of effects of oil price shocks, we find very little negative consumption effects irrespective of the income group.

[3]Recognising the composite nature of the global food price index, we also go deeper in our analysis to uncover particularly pronounced relative price effects resulting from an increase in global food prices on local relative prices of pulses, sugar, oils and fats, and vegetables.

compared to fuel, for all income groups in India.[4]

Our paper is related to several strands of the literature. The two-way relationship between global oil prices and the U.S. macroeconomy, as well as the implications for US monetary policy, has been studied extensively in Hamilton [1983], Hamilton [2003], Barsky and Kilian [2004], and Kilian [2009]. We extend this body of work by estimating the distributional effects of global oil prices, an area that has only recently garnered empirical attention (see, for example, Gelman, Gorodnichenko, Kariv, Koustas, Shapiro, Silverman, and Tadelis [2023], Peersman and Wauters [2022], and Känzig [2021]). Moreover, we use the oil supply shock estimated in Baumeister and Hamilton [2019] as an instrument in our panel IV specifications.

Although there has been some recent research on global food price shocks, such as De Winne and Peersman [2016] and Peersman [2022], this area of study is less developed than the oil shock literature. Previous studies have mainly examined the aggregate or sectoral effects of food price shocks, such as their impact on sectoral inflation. We contribute to this literature by estimating the distributional effects, at the household level, of global food prices. In our IV specification, the instrument we develop is novel as we use a statistical factor-based method with data from a broad cross-section of commodity prices to isolate food-specific and aggregate demand factors from global food price dynamics.

Our paper also contributes to two other strands of the literature that have examined the distributional effects of domestic monetary policy shocks. On the theoretical front, Auclert [2019] develops a general model that encompasses various redistribution based channels for monetary policy transmission. On the empirical front, Coibion, Gorodnichenko, Kueng, and Silvia [2017] study the effects of US monetary policy shocks on inequality, while Holm, Paul, and Tischbirek [2021] estimate the heterogeneous household effects of Norwegian monetary policy shocks along the liquid asset distribution. In building on this body of work, our paper focuses on the distributional implications of an external shock in the context of an emerging market. Additionally, we use detailed household panel consumption and income data at a monthly frequency to investigate these implications and transmission mechanisms.

Our paper shares a common theme with the literature that highlights the significant impact of external shocks on emerging market economies and their role in driving busi-

[4]As we discuss later, the class of preferences that align well with our results are iso-elastic non-homothetic constant elasticity of substitution preferences between food and fuel. Our econometric analysis uses this class of preferences. See Matsuyama [2022] for a discussion of such preferences. We also do the econometric analysis at a disaggregated food category level and find evidence that many food categories constitute an essential consumption good.

ness cycle dynamics. For example, Neumeyer and Perri [2005] have emphasized the role of global interest rate shocks, while Fernandez-Villaverde, Guerron-Quintana, Rubio-Ramirez, and Uribe [2011] and Bhattarai, Chatterjee, and Park [2020] have highlighted the importance of global volatility or uncertainty shocks. Our paper focuses on a different type of external shock, namely global food and oil price shocks, and estimates their distributional implications in India.[5]

Furthermore, our findings regarding distributional effects on consumption imply that monetary policy in emerging markets might need to respond to such external shocks, even though they arise in sectors with flexible prices, in order to decrease consumption dispersion in the economy. The sticky price monetary policy literature highlights that optimal policy should not put any weight on inflation of flexible price sectors as they do not cause relative price distortions. In canonical open economy sticky price models, optimal monetary policy only targets domestic price inflation if import prices are determined flexibly ([Clarida, Galı, and Gertler, 2002]). Such insights do not account for effects of such shocks on consumption inequality. If such effects are present, optimal policy will need to respond to mitigate consumption disparities in the economy because with incomplete markets, consumption dispersion appears in the objective of the benevolent central bank ([Bhattarai, Lee, and Park, 2015] and [Acharya, Challe, and Dogra, 2020]).

2 Data, Stylized Facts, and Instrumental Variables

We now discuss our data, present some stylized facts that serve as motivation for the econometric exercise, and describe in detail our instrumental variables.

2.1 Data Description

Our household data is from the Consumer Pyramid Household Survey (CPHS) dataset, a survey conducted by the Centre for Monitoring the Indian Economy (CMIE). CPHS has surveyed over 236,000 unique households since 2014 and is the most comprehensive longitudinal consumption data available for India. CPHS is unique in including both income data and detailed consumption data in a single longitudinal dataset. Moreover, it is available at the monthly frequency, which allows an analysis of the dynamic effects of

[5] Lakdawala and Singh [2019] study aggregate output and price effects in India of oil supply shocks using the Baumeister and Hamilton [2019] supply shocks directly as a measure of external shocks.

global food and oil prices in a straight-forward way, without having to impute data due to frequency mismatch between the shock series and the consumption/income data. The time period of our analysis is Jan 2014-Dec 2019.

We construct consumption, income, and earnings measures following closely the method of Coibion et al. [2017]. Consumption expenditure comprises of 153 categories. Total consumption measure we construct is the sum of non-durable consumption (food, cooking fuel, electricity and transport, intoxicants), durable consumption (appliances, furniture, jewelry, clothing, electronics, toys, cosmetics), and service consumption (entertainment, beauty services, fitness services, restaurants, etc). We present results on total consumption and non-durable consumption separately in all our analysis.[6] We also present results on food, fuel (including cooking fuel and electricity and transport fuel), and detailed food sub-components consumption.

Total consumption is deflated using monthly state-region level Consumer Price Index (CPI) - Combined series (2012 base) available from the Ministry of Statistics and Program Implementation (MoSPI), Government of India. The remaining consumption categories are deflated using their respective CPIs as follows. Food consumption is deflated by the index available from the MoSPI. Fuel consumption, where we include not just the cooking fuel expenditure given directly by the MoSPI but also fuel expenditure on transportation, is deflated using a weighted average of the two categories with the weights provided from the MoSPI. Non-durable consumption is deflated using a weighted average of food, cooking fuel, and transport price indices with the weights provided from the MoSPI.[7]

We construct income as the sum of household income from rent, wages, self-production, private transfers, government transfers, business profits, sale of assets, lotteries and gambling, pension, dividends, interest and deposit provident fund and insurance. These categories are an exhaustive list of all income sources collected in the CPHS. Our earnings measure is constructed using income from wages and overtime bonus. To construct real values of these nominal income and earnings variables we use the state-region level CPI - Combined series (2012 base). Using these data, measures of inequality we construct for stylized facts are: Gini coefficients, cross-sectional standard deviations, and differences

[6] The average share of non-durable consumption in total consumption is 75.1% in our dataset.

[7] We use the most detailed state and region (urban or rural) level monthly deflator available for India at a monthly frequency for our time period, following the suggestions in Deaton [2019]. There are 35 states and union territories (regions administered by the central government) in our dataset. While headline and food CPI is available for each state-region, nondurable CPI has to be constructed. We overcome this challenge by constructing state-region (urban and rural) level non-durable CPI using state-region level headline CPI as well as state-region level food and energy consumption shares in the CPI basket. We provide further details on the data in the Appendix in Section A.

between individual percentiles (90th-10th and 75th-25th) on log levels.

Finally, we use IMF's Global Price of Food Index (Nominal, US Dollar) and the West Texas Intermediate crude oil prices (WTI, US Dollar per barrel) as our measure of global food and oil prices respectively.

2.2 Summary Statistics Along the Income Distribution

Several summary statistics from our household panel data, along the income distribution, are in Appendix B. Most importantly, we present in Table A1 summary statistics on average (across households and months) monthly income, monthly consumption, share of non-durable consumption, and share of food consumption by various income deciles, where the deciles are on the basis of the initial period (2014) real household income. The poorest income group is definitely below poverty line and are net borrowers with a high share of non-durable and food in consumption. Savings rate rises while non-durable and food shares decline with income. The top income decile has nearly a 75% savings rate and a relatively low share of food in total consumption.[8]

These statistics motivate us to divide households in five broad income groups when we estimate heterogeneous consumption effects of global commodity price shocks. In these regressions where we estimate interaction effects, we consider five income groups: very low income (decile 1), low income (deciles 2 and 3), lower middle income (deciles 4, 5 and 6), upper middle income (deciles 7, 8 and 9), and high income (decile 10). We determine the cut-offs for deciles based on real income in 2014, and assign each household to a group based on those cutoffs.[9]

2.3 Global Commodity Prices and Aggregate Inequality

We now present some stylized facts on global commodity prices and aggregate inequality in India.

[8]We also show in Table A2 that earnings constitute the most important category of income for lower income groups. For higher income groups, capital income constitute a significant fraction of total income whereas for lower income groups transfers is an important component.

[9]Note that the same household may belong to different income groups at different points in time depending on their current real income. This is an issue we address in the sensitivity analysis in Section 5.

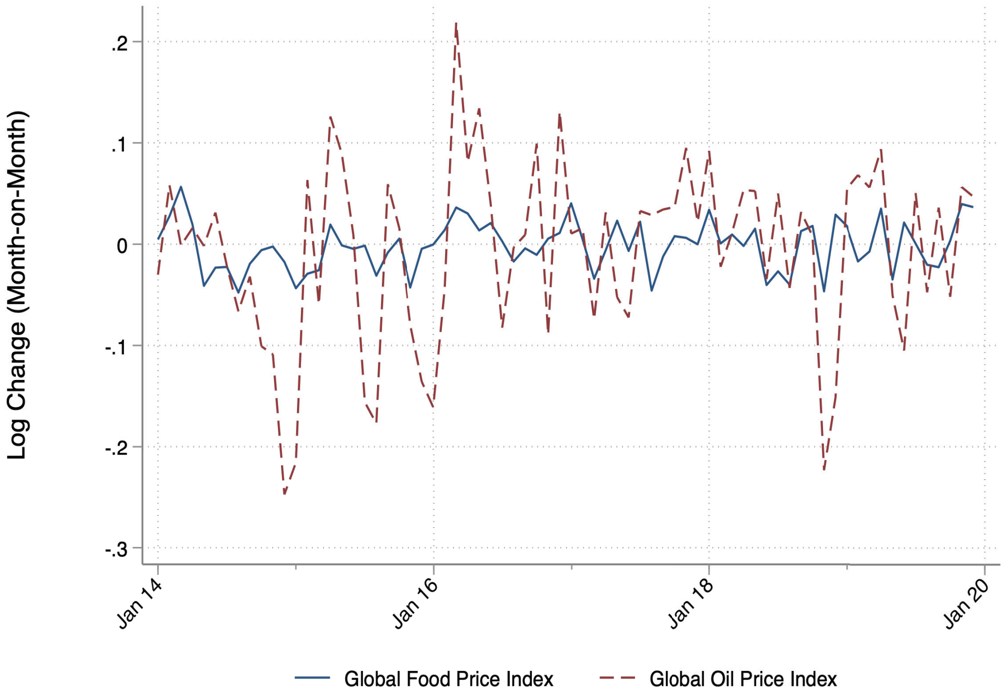


Figure 1: Changes in Global Food and Fuel Prices

Notes: This figure plots the log change in IMF's Global Price of Food Index (Nominal US Dollar) and West Texas Intermediate Crude Oil Prices (US Dollar per barrel).

We first plot the log changes in global food and oil prices in Figure 1. As expected, average of the changes is close to zero while the standard deviation is approximately 2.4% for food price and nearly 8.9% for oil price, confirming a higher oil price volatility. AR(1) coefficients of the estimated processes for these change in prices are very low, indicating that the changes are largely transitory in nature. Finally, changes in the two series are positively correlated but not very highly so, hence implying independent sources of variation.[10]

We construct various measures of inequality and compute the conditional correlation of them with one period lagged values of global food and oil price changes. As shown in Table 1, the correlations are mostly positive. The largest positive correlations are observed between consumption inequality and the two commodity price changes. Response of consumption to global price shocks is the most relevant metric to evaluate welfare effects of such shocks. Our raw data reveals a strong correlation between aggregate consumption

[10]Correlation between the two series is 0.4 in our sample. Some co-movement such as this to be expected given the role of energy as input in production of food as well as the possible role of global demand in driving both commodity prices. See, for example, Peersman, Rüth, and Van der Veken [2021].

inequality and the external commodity price changes. Does this "smell test" pass an econometric examination? This is the key focus of our paper.

Table 1: Correlations of Aggregate Inequality with Global Food and Oil Price Changes

Panel A: Correlations With Global Food Price Change				
	Gini	SD	90th-10th	75th-25th
Income Inequality	0.108	0.097	0.153	0.063
Earnings Inequality	0.058	0.098	0.077	-0.062
Consumption Inequality	**0.234**	**0.232**	**0.215**	**0.229**
Panel B: Correlations With WTI Crude Oil Price Change				
	Gini	SD	90th-10th	75th-25th
Income Inequality	0.137	0.134	0.118	0.134
Earnings Inequality	0.047	0.098	0.090	-0.023
Consumption Inequality	**0.285**	**0.296**	**0.269**	**0.230**

Notes: This table shows correlations of one-period lag of global food and oil price changes with various inequality measures for income, earnings, and consumption that are constructed using the micro household panel data.

2.4 Instrumental Variables for Global Commodity Prices

In our stylized facts above we use changes in world oil and food prices (in logs) as an external shock, motivated by the small open economy assumption for India. These results can be considered as OLS versions of our estimation framework as they conflate the effects of various underlying shocks that lead to changes in world oil and food prices. As has been shown in the oil shock literature however, for a cleaner interpretation of these results, it is instructive to separate out such global oil/food price changes as coming from global demand or commodity-specific demand or global supply shocks. In addition to allowing a cleaner interpretation, isolating supply side variation also guards against omitted variable bias (OVB) problems. For instance, OVB can arise as Indian households could have *direct* exposure to the global business cycle, which is well-known to drive global commodity prices. To address this issue, we take an Instrumental Variable (IV) approach where we use supply side instruments for the change in global oil and food prices.

For the oil price change, our IV is the oil supply shock estimated in Baumeister and Hamilton [2019]. For the food price change, we construct an IV based on residuals of global food commodity prices after extracting two common factors from a cross-section of commodity prices. We residualize the global food price index with a common factor and a food specific factor estimated using sign restrictions on a panel of 37 non-energy commodity prices (13 industrial metals and 24 food prices, available from FRED and Bloomberg in the time period 1990-2022) in a dynamic factor model. Our estimation method for this dynamic factor model is outlined in Appendix C. We plot the changes in global commodity prices and their IVs, the respective supply shocks, in Appendix D in Figure A1.[11]

3 Distributional Effects of Global Commodity Price Shocks

To econometrically estimate the distributional effects of global shocks in a dynamic setting, we use a household panel local projection regression framework where we estimate heterogeneous dynamic effects on consumption of global oil and food price shocks. In particular, these consumption effects will be allowed to differ along the income distribution.

[11] It is challenging to estimate a supply shock for the food sector in a way analogous to the oil supply shock due to two main reasons. Unlike oil, food is not a single commodity–it is a composite of several commodities. Also, while monthly price data is available for various components of food, monthly production data is generally not available. There are two approaches that one can take to circumvent these problems. The first is to use a large cross-section of non-energy commodity prices and a combination of statistical and theory based identification to disentangle supply and demand shocks (e.g., as in Alquist, Bhattarai, and Coibion [2020]) and this is the approach we take. The second is to use a limited cross-section of price and a proxy for monthly production data, as outlined in De Winne and Peersman [2016]. However, the major crops of India are subject to various price regulations both on the supply and demand side in the domestic market due to minimum support prices for farmers and public distribution system of basis cereals for consumers. Hence, we prefer to rely on an approach that uses a broad cross-section of prices.

3.1 Panel Local Projection Framework

To capture such dynamic heterogeneous effects, we estimate a household level panel local projection model with interaction effects. Our estimation equation is:

$$\begin{aligned} c_{i,t+h} - c_{i,t-1} = {} & \beta^{g,h}_{0,\text{food}} ext^{\text{food}}_t \times \mathbb{1}_{i\in g(t)} + \beta^{g,h}_{0,\text{oil}} ext^{\text{oil}}_t \times \mathbb{1}_{i\in g(t)} + \sum_{j=1}^{J} \alpha^h (c_{i,t-j} - c_{i,t-j-1}) \\ & + \sum_{k=1}^{K} \beta^h_{k,\text{food}} ext^{\text{food}}_{t-k} + \sum_{k=1}^{K} \beta^h_{k,\text{oil}} ext^{\text{oil}}_{t-k} + \sum_{d=0}^{D} \delta^h D_{t-d} + \gamma^{g,h} X_t \times \mathbb{1}_{i\in g(t)} \\ & + \delta_{\text{c},t} + \delta_{\text{l},t} + \delta_{\text{e},t} + \delta_{\text{city},t} + \delta_{\text{age},t} + \mathbb{1}_{i\in \text{s}} \times \mathbb{1}_{\text{year}} + \mathbb{1}_{i\in \text{s}} \times \mathbb{1}_{\text{month}} + \epsilon_{i,t+h} \end{aligned} \tag{3.1}$$

Here, c_i is the log of consumption for household i for various measures of consumption; ext^{food} and ext^{oil} stand for measures of the global food and oil price shocks respectively; $\delta_{\text{c},t}, \delta_{\text{l},t}, \delta_{\text{e},t}, \delta_{\text{city},t},$ and $\delta_{\text{age},t}$ are respectively the fixed effects for household's caste, religion, education, residence in a big city, and age; $\mathbb{1}_{i\in \text{s}} \times \mathbb{1}_{\text{year}}$ and $\mathbb{1}_{i\in \text{s}} \times \mathbb{1}_{\text{month}}$ are a set of household i's residence state by calendar year and residence state by calendar month fixed effects to account for state-specific trends and seasonality respectively; and D is the dummy for the Indian government's demonetization policy, which is allowed to have lagged effects up-to three lags.[12] For the AR and MA coefficients, we choose $J = 3, K = 3$.

We estimate the above specification separately for each horizon ranging from $h = 0$ to $h = 12$. In all the regressions, the observations are weighted using sampling weights provided by CMIE which takes into account the non-response factor. The standard errors are clustered at the state level.

The most important aspect of this specification is that we allow the consumption effects to differ by income of the household. That is, $g(t)$ denotes income group of household i at time t constructed using cutoffs from 2014 real income data. The effects of external shocks are thus, allowed to vary by income groups. We consider five income groups: very low income (decile 1), low income (deciles 2 and 3), low middle income (deciles 4, 5 and 6), upper middle income (deciles 7, 8 and 9), and high income (decile 10). $\beta^{g,h}_{0,\text{food}}$ is the coefficient of interest that captures the impact of global food price shock at time t on

[12] There are a total of eleven age groups defined based on the age of the household head. The youngest and the oldest groups consist of households below twenty years and above 65 years respectively. Households between these two ages, which roughly corresponds to working age, are classified in to groups of five years each. Education groups are defined similarly based on the education level of the household head. We consider three groups – below high school, high school educated but less than college educated, and college graduate and above. Summary statistics for different household characteristics are presented in Appendix B in Table A3.

households of group g at horizon h; $\beta_{0,\text{oil}}^{g,h}$ is similarly the estimate for the global oil price shock. We report cumulative impulse responses below.

X denotes controls for aggregate world conditions: world industrial production as a proxy for aggregate demand ([Kilian, 2009]); US monetary policy stance as captured by the federal funds rate; and global financial volatility as captured by the VIX index. These aggregate global controls are interacted with household income group dummies.

As we discussed above, in addition to estimating OLS results, based on using changes in global food and oil prices as the shock measure ext, we will also estimate IV results where we instrument the changes in global food and oil prices. These IV results will isolate variation coming from global supply shocks and also guard against omitted variable bias problems, as we discussed previously.

Table A5 in Appendix E lists all the control and instrumental variables in our household panel local projection estimation.

3.2 Heterogeneous Consumption Effects: IV Results

For the local projection household panel exercise based on estimation of equation (3.1), where the effects on consumption are allowed to vary along the income distributions, our key IV results are in Figures 2 and 3 for food price shocks and oil price shocks respectively.[13] We present results for total consumption, non-durable consumption, and own-category consumption (fuel consumption for oil price shocks and food consumption for food price shocks). The results show that there are adverse effects on consumption, for both total and non-durable consumption, for both the shocks.

On heterogeneous effects along the income distribution, the two shocks show different patterns. For the food price shock, the lower income groups are hit harder and the negative effects become progressively less pronounced as we move to higher income groups. For the oil price shocks, the effects are more nuanced and more symmetric along the income distribution. For instance, both upper-middle and high-income households' consumption drops as much as the lower income groups. Taking statistical uncertainty into account, for oil price shocks, the consumption effects are more muted for the low-income and lower-middle income groups. That is, the adverse effects of oil price shocks are most precisely estimated and clearly negative for the tails of the income distribution.

[13] The first-stage F-statistics for these IV regressions are reported in Appendix E in Table A6.

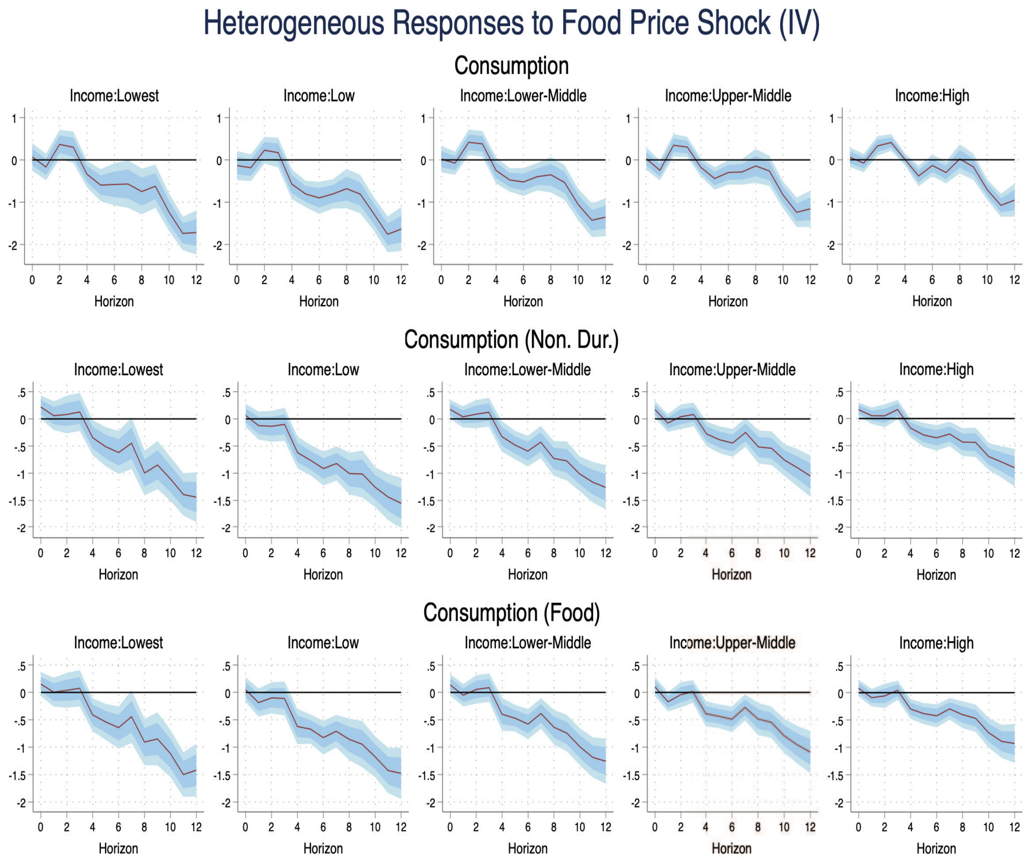


Figure 2: Response of Consumption to External Food Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global food price, which is instrumented by a global food supply shock and the dependent variable is log changes in household consumption. The light blue region is the 90% confidence interval and the dark blue region is the 68% confidence interval.

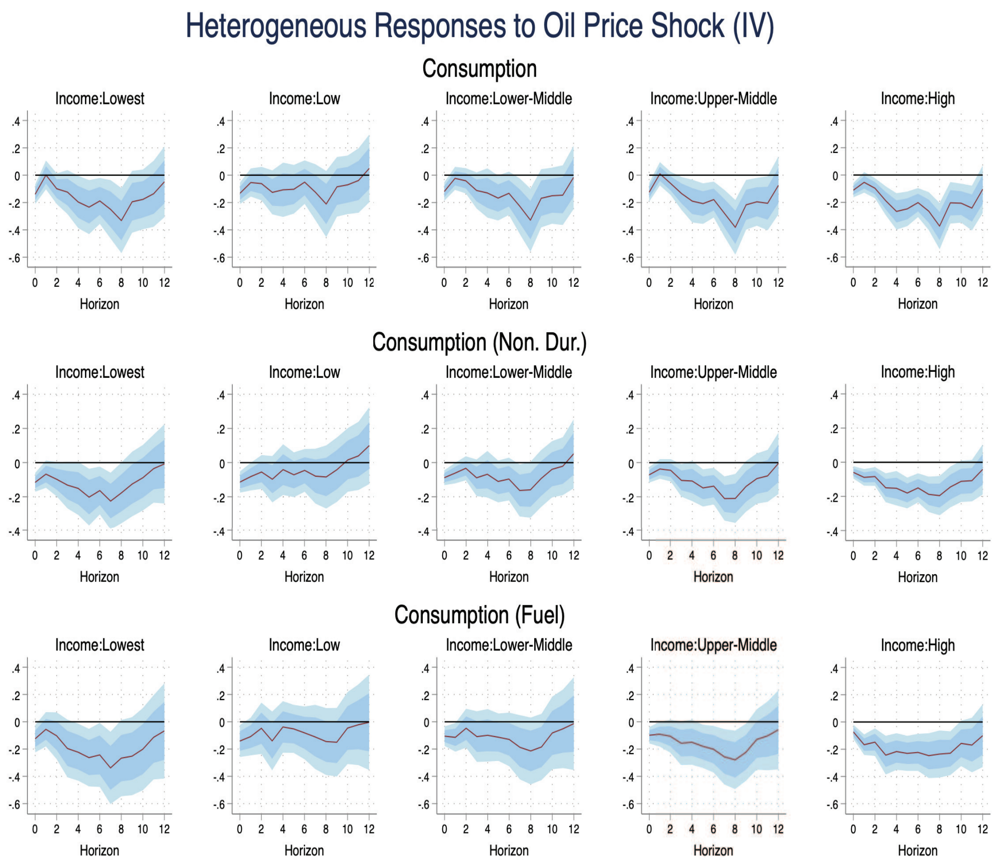


Figure 3: Response of Consumption to External Oil Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global oil price, which is instrumented by a global oil supply shock and the dependent variable is log changes in household consumption. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

To conclude, Figure 4 provides summary statistics of the results, using a box and whisker plot, for non-durable consumption. As we mentioned before, we observe that poorer income groups suffer a substantially larger consumption loss across all horizons following an exogenous rise in global food prices, whereas the poorest, the upper-middle, and the high income groups are equally vulnerable to an exogenous rise in oil prices. Moreover, the low income group (deciles 2 and 3) is shielded most from oil price increases but suffers most from food price increases.[14]

[14] The low income (deciles 2 and 3) group is even more vulnerable than the poorest to rising food prices possibly due to the public distribution system shielding the consumption loss of the very poor. To assess the statistical significance of the difference in effects across income groups, we show in Appendix E in Figure A2 the effects for all groups relative to the low income group. They shows that the differences are statistically significant and also go in the direction emphasized above.

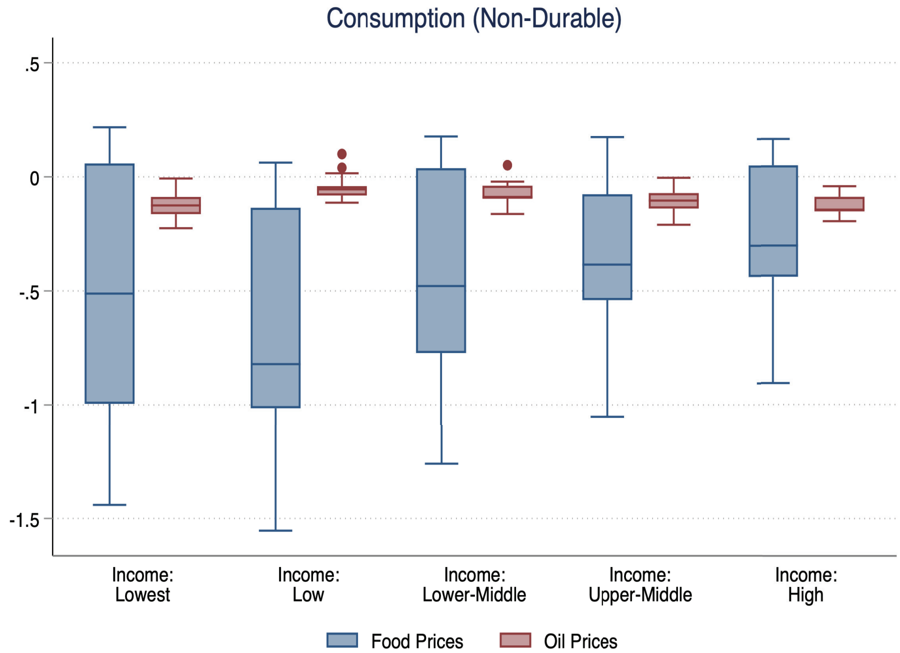


Figure 4: Summary Statistics of Response of Non-durable Consumption to Food and Fuel Price Shocks by Income Quintiles (IV)

Notes: This figures is a box and whisker plot that summarizes the responses of non-durable consumption to the two external shocks that are presented in Figures 2 and 3. The line in the center of each box represents the median impulse response estimate across thirteen horizons; the top and the bottom edges of each box represent the 75th and the 25th percentiles respectively; and the lines above and below each box represent respectively the upper and lower adjacent values calculated as in Tukey et al. [1977].

Table 2: Magnitude of Real Non-durable Consumption Loss (in %)

	1 SD Price Shock		**2022 External Price Shock**	
Income Group	Max Impact 1 sd food shock	Max Impact 1 sd oil shock	Max Impact 2022 food shock (8% in Aug 2022)	Max Impact 2022 oil shock (26% in Aug 2022)
Lowest	-3.54	-2.01	-11.53	-5.85
Low	-3.82	-1.01	-12.43	-2.95
Lower middle	-3.10	-1.46	-10.08	-4.24
Upper middle	-2.58	-1.88	-8.41	-5.46
High	-2.22	-1.75	-7.24	-5.08

Notes: This table shows the loss in real non-durable consumption (in % terms) for the five income groups based on the estimates of elasticities presented in Figure 4. Columns (2)-(3) refer to a 1 standard deviation shock to food prices (2.4%) and oil prices (8.9%). Columns (4)-(5) refer to the August 2022 massive rise in food prices (8 %) and oil prices (26 %).

To appreciate the magnitude of consumption loss due to the external shocks and the pattern of heterogeneity along the income distribution, in Table 2, we translate the elasticity estimates presented in Figure 4 to consumption loss in % terms. The first two columns of Table 2 capture the maximum negative impact of a 1 standard deviation shock in global food and oil prices (2.4% rise in the food price index and 8.9% rise in the WTI crude oil price index, respectively, as presented earlier in Figure 1). This clearly shows the pattern of heterogeneity we have emphasized: for an exogenous food price increase, the poorest two groups clearly suffer the most in consumption loss in % terms and there is a clear pattern of monotonicity along the income distribution, while for an exogenous oil price increase, the lowest and the two highest income groups suffer similarly. Moreover, the low income group is shielded most from oil price increases but suffers most from food price increases.

To make this analysis salient for recent events, a similar pattern is observed when we evaluate the consumption loss for the massive rise in food and oil prices in 2022 in the last two columns of Table 2. In August 2022, the IMF Food Price Index was higher by 8 percent while the WTI Crude Oil Prices was higher by 26 percent, compared to a year ago. For such a large rise in external food prices, the poorest two groups suffer around 12% loss in non-durable consumption and the effect declines monotonically with income. The oil price increase leads to a smaller, but still sizeable, negative effects of around 5 % for the poorest and the two highest income groups.

4 Channels for Heterogeneous Consumption Effects

After establishing these baseline results on heterogeneous effects on consumption, we delve further into interpretation and possible transmission mechanisms. In particular, we aim to assess the channels that work via real income, through relative price effects (say across sectors), and those that reflect non-homotheticity in preferences.

4.1 Transmission Mechanisms in Theory

We start by developing a theoretical framework that will inform how we explore and test for various transmission mechanisms in the data.

4.1.1 Dynamic Consumption-Saving Problem

As in Auclert [2019], we consider a infinite horizon consumption-savings problem in a perfect foresight environment with unexpected shocks, where the household can trade various nominal and real assets of different maturities. The household chooses $\left\{C_t, \frac{B_t}{P_t}, \frac{B_{2,t}}{P_t}, E_t, L_t\right\}$ to maximize lifetime utility

$$\sum_{t=0}^{\infty}\beta^t\left[\frac{C_t^{1-\sigma}}{1-\sigma}-\frac{L_t^{1+\phi}}{1+\phi}\right]$$

subject to a sequence of flow budget constraints

$$C_t+Q_tb_t+Q_{2,t}b_{2,t}+S_tE_t=b_{t-1}\frac{1}{\Pi_t}+Q_tb_{2,t-1}\frac{1}{\Pi_t}+\left(S_t+D_t\right)E_{t-1}+w_tL_t+T_t, \tag{4.1}$$

where C_t is aggregate consumption, L_t is labor, B_t is holdings of one-period risk-free nominal bonds, $B_{2,t}$ is holdings of two-period risk-free nominal bonds, and E_t is holdings of stocks.[15] Q_t, $Q_{2,t}$, and S_t are prices of the one-period bond, the two-period bond, and the stock respectively. The stock yields dividends D_t, P_t is the aggregate nominal price level, $\Pi_t=\frac{P_t}{P_{t-1}}$ is gross inflation, w_t is real wages, and T_t is lump-sum transfers. Thus, $b_t=\frac{B_t}{P_t}$ is the real holdings of the one-period nominal bonds and $b_{2,t}=\frac{B_{2,t}}{P_t}$ is the real holdings of the two-period nominal bonds. Finally, $\beta\in(0,1)$ is the discount factor, σ^{-1} is the intertemporal elasticity of substitution, and ϕ^{-1} is the Frisch elasticity of labor supply.

The flow budget constraint, Equation (4.1), makes clear how shocks in period t affect consumption-savings decisions through their effects not only on labor earnings w_tL_t, but also through revaluations of financial positions by affecting inflation and asset prices Π_t, Q_t, and S_t. In this perfect foresight environment, the asset pricing conditions imply equal interest rates across the various assets. Using these no-arbitrage conditions and the Transversality condition together with the flow budget constraints yields the intertemporal budget constraint

$$\sum_{s=0}^{\infty}\rho_{t,t+s}C_{t+s}=\left[\frac{1}{\Pi_t}\left(b_{t-1}+Q_tb_{2,t-1}\right)+\left(S_t+D_t\right)E_{t-1}\right]+\sum_{s=0}^{\infty}\rho_{t,t+s}\left(w_{t+s}L_{t+s}+T_{t+s}\right) \tag{4.2}$$

where

$$\rho_{t,t}=1;\,\rho_{t,t+s+1}=\prod_{j=0}^{s}R_{t+j+1}^{-1};\,R_{t+j+1}=\frac{1}{Q_{t+j}\Pi_{t+j+1}}.$$

The intertemporal budget constraint, Equation (4.2), states that the present discounted

[15] The household also faces an appropriate no-Ponzi game constraint.

value of consumption, using time-varying interest rates for discounting, equals the present discounted value of labor income and transfers as well as the real value of payoffs from ex-ante financial positions. It also shows that unexpected shocks can affect consumption through (a) wage earnings by affecting current or future wages or labor supply; (b) discount factors by affecting current or future real interest rates; and (c) real value of payoffs on ex-ante financial holdings by affecting current inflation, short-term nominal interest rate, or stock prices. Heterogeneity in how such unexpected shocks affect wage earnings or heterogeneity in ex-ante financial positions in terms of nominal bonds, maturity of nominal bonds, and stocks in turn can then generate heterogeneity in consumption effects.

Going further, if we impose a unit intertemporal elasticity of substitution (σ^{-1}=1), since

$$\rho_{t,t+s+1} = \prod_{j=0}^{s} \frac{\beta C_{t+j}}{C_{t+j+1}},$$

by manipulating equation (4.2), we get the solution for current consumption as

$$C_t = (1-\beta)\left[\frac{1}{\Pi_t}\left(b_{t-1} + Q_t b_{2,t-1}\right) + \left(S_t + D_t\right) E_{t-1} + \sum_{s=0}^{\infty} \rho_{t,t+s}\left(w_{t+s} L_{t+s} + T_{t+s}\right)\right]. \quad (4.3)$$

Equation (4.3) makes clear how the various transmission mechanisms discussed above, (a)-(c), govern the effect of unexpected shocks on current consumption. Perhaps even more importantly, it shows that heterogeneity in the response of wage income as well as heterogeneity in ex-ante positions in nominal bonds, maturity of nominal bonds, and stocks will lead to heterogeneity in consumption. Given our data, we state a key prediction that we can test below.

Testable Prediction 1: From Equation (4.3), heterogeneous response of wage income to the external price shocks leads to a heterogeneous response of consumption.

4.1.2 Static Expenditure Allocation Problems

Given the dynamic consumption-saving problem and solution in the previous section that determines the level of aggregate consumption, we now present the static expenditure allocation problem across various consumption categories, given a level of total consumption, C_t.

C_t is a standard constant elasticity of substitution (CES) aggregator of non-durable consumption goods and the rest of consumption goods, for example, durable and services, denoted by $C_{N,t}$ and $C_{S,t}$ respectively:

$$C_t = \left[(1-\alpha)^{\frac{1}{\eta}} C_{N,t}^{\frac{\eta-1}{\eta}} + \alpha^{\frac{1}{\eta}} C_{S,t}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}. \tag{4.4}$$

Here, $\eta > 0$ is the elasticity of substitution and $\alpha > 0$ governs the share of the two types of consumption goods. Standard expenditure minimization problem gives the corresponding optimal price indices and relative expenditure shares as:

$$P_t = \left[(1-\alpha) P_{N,t}^{1-\eta} + \alpha P_{S,t}^{1-\eta} \right]^{\frac{1}{1-\eta}},$$

$$\frac{P_{N,t}C_{N,t}}{P_t C_t} = (1-\alpha) \left(\frac{P_{N,t}}{P_t} \right)^{1-\eta}; \frac{P_{N,t}C_{N,t}}{P_{S,t}C_{S,t}} = \frac{1-\alpha}{\alpha} \left(\frac{P_{N,t}}{P_{S,t}} \right)^{1-\eta},$$

where $P_{N,t}$ and $P_{S,t}$ are prices of the non-durable and rest of consumption goods respectively. Hence, relative expenditure shares between non-durable and total consumption are completely governed by relative prices. In particular, expenditure share of non-durable consumption ($M_i, i = N$) in total consumption is given in logs as

$$\log M_{N,t} = \log(1-\alpha) - (\eta - 1) \log \left(\frac{P_{N,t}}{P_t} \right). \tag{4.5}$$

Given our data, we state a key prediction that we can test below.

Testable Prediction 2: Assuming $\eta > 1$, from Equation (4.5), if relative price of non-durable consumption increases in response to the external price shocks, it leads to a decrease in the expenditure share of non-durable consumption.

Next, we model non-durable consumption as an iso-elastic non-homothetic CES aggregator of food and fuel.[16] Hence, expenditure shares of food and fuel ($M_i, i = F, O$ for food

[16]For this class of utility function, for a consumption bundle $\mathbf{x}$, $U(\mathbf{x})$ is given implicitly as:

$$[\sum_{i=1}^{n} \gamma_i^{\frac{1}{\sigma}} U(\mathbf{x})^{\frac{\epsilon_i - \sigma_\epsilon}{\sigma_\epsilon}} x_i^{1-\frac{1}{\sigma_\epsilon}}]^{\frac{\sigma_\epsilon}{\sigma_\epsilon - 1}} \equiv 1,$$

where $\sigma_\epsilon > 0$ ensures global quasi-concavity, and $\frac{\epsilon_i - \sigma_\epsilon}{1-\sigma_\epsilon} > 0$ ensures global monotonicity. Given total expenditure on this bundle of consumption, E, the cost of living index (P) is implicitly given by:

$$[\sum_{i=1}^{n} \gamma_i (\frac{E}{P})^{\epsilon_i - 1} (\frac{p_i}{P})^{1-\sigma_\epsilon}]^{\frac{1}{1-\sigma_\epsilon}} \equiv 1.$$

See Matsuyama [2022] for further details and references.

and fuel respectively) in non-durable consumption are given by

$$M_{it} \equiv \frac{P_{it}C_{it}}{P_{N,t}C_{N,t}} = \gamma_i \frac{E_{Nt}}{P_{Nt}}^{\epsilon_i - 1} \frac{P_{it}}{P_{Nt}}^{1-\sigma_\epsilon}, \tag{4.6}$$

where $P_{i,t}$ are prices of the food and fuel ($i = F, O$ for food and fuel respectively), $E_{Nt} = P_{N,t}C_{N,t}$ is nominal expenditure on non-durable goods, ϵ_i is the slope of the Engel curve, $\sigma_\epsilon > 0$ is the price elasticity, and γ_i is the expenditure share in non-durable.

In the presence of such non-homotheticity, the expenditure shares depend on the level of real non-durable consumption. More precisely, relative expenditure shares are

$$\underbrace{\log\left(\frac{M_{it}}{M_{jt}}\right)}_{\text{Relative expenditure}} = \log\left(\frac{\gamma_i}{\gamma_j}\right) - \underbrace{(\sigma_\epsilon - 1)\log\left(\frac{P_{it}}{P_{jt}}\right)}_{\text{Relative price effect}} + \underbrace{(\epsilon_i - \epsilon_j)\log\left(\frac{E_{Nt}}{P_{Nt}}\right)}_{\text{Total expenditure effect}}. \tag{4.7}$$

In Equation (4.7) above, a good i is a necessity if and only if $\epsilon_i < \bar{\epsilon}$ and a luxury if and only $\epsilon_i > \bar{\epsilon}$, where $\bar{\epsilon}$ is the budget-share weighted average of ϵ_i. This means that iso-elastic non-homothetic CES can allow the same good to be luxury or necessity depending on the level of real expenditure.

As long as (overall, and various sub-components of) food is a substitute with fuel (implying an elasticity of substitution, $\sigma_\epsilon > 1$), an increase in relative prices reduces relative expenditure via the standard expenditure switching effect. If real non-durable expenditure ($\frac{E_{Nt}}{P_{Nt}}$) falls, the only way the relative expenditure may increase with rising relative prices is if $\epsilon_i < \epsilon_j$. This condition, $\epsilon_i < \epsilon_j$, in a two-good framework implies that consumption of good i is a necessity. Given our data, we state a key prediction that we can test below.

Testable Prediction 3: Assuming $\sigma_\epsilon > 1$, from Equation (4.7), a *sufficient* condition for a good to be essential is that relative expenditure in the good rises following a rise in relative prices and a fall in total real expenditure.

4.2 Empirical Evidence on Transmission Mechanisms

We now present empirical evidence on the various transmission mechanisms we developed theoretically above. In particular, our analysis here will be guided by the testable predictions we developed above.

4.2.1 Labor Earnings Channel

We start by assessing heterogeneous real labor earnings effects of these shocks in the household panel IV local projection framework. That is, we estimate Equation (3.1), but with real labor earnings as the dependent variable. In our theoretical framework, the intertemporal budget constraint, Equation (4.2), and the solution for consumption, Equation (4.3)), have shown how heterogeneity in effects on labor income can lead to heterogeneity in effects on consumption. We had summarized this channel in Testable Prediction 1 in Section 4.1.1.

Figure 5 shows the response of labor income to oil price shocks (top panel) and food price shocks (bottom panel). It is clear that food price shocks have a significant negative effect on labor earnings through out the income distribution. Moreover, these negative earnings effects of food price shocks are monotonically decreasing along the income distribution, analogous to their negative consumption effects that we showed in Figure 2. This suggests that heterogeneity in labor income effects are one source of heterogeneity in consumption response for food price shocks. For oil price shocks, the negative effects are more limited and are significant for the poorest group initially and for the rich over time. Nevertheless, these are still consistent with the negative consumption effects of oil price shocks for these two income groups that we showed in Figure 3.[17]

[17]For oil price shocks, effects on (non-durable) consumption are also significant for the upper middle income group in Figure 3. We however do not find negative labor income effects for this group in Figure 5.

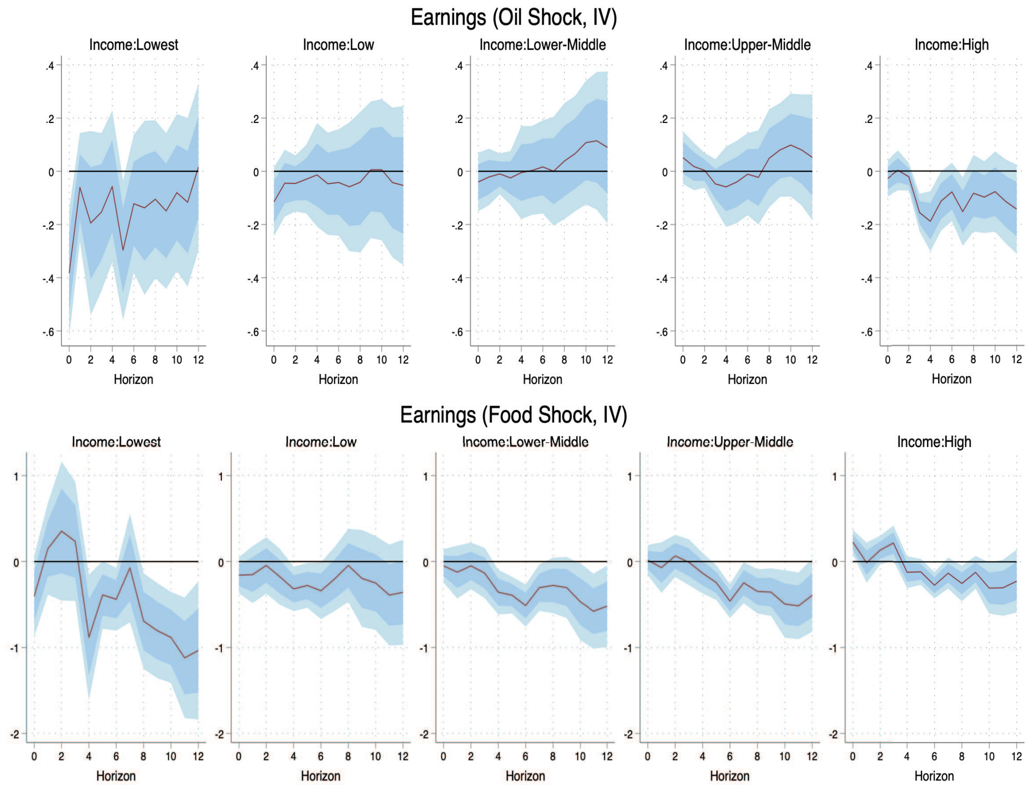


Figure 5: Response of Earnings to External Food and Fuel Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global food price, which is instrumented by a global food supply shock, and log changes in global oil price, which is instruments by a global oil supply shock. The dependent variable is log changes in household labor earnings. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

4.2.2 Aggregate Price and Relative Price Channels

We now assess the various price channels through which external commodity price shocks can affect consumption inequality. To do so, we investigate whether external price shocks pass-through to domestic prices that Indian consumers face. We use, for various components, state-region level monthly CPI data from MoSPI as measures of domestic prices. We then estimate the dynamic responses of domestic prices, that is, how these various components of state-region level monthly CPI respond to global price shocks in a panel local projection framework.

In our theoretical framework, the intertemporal budget constraint, Equation (4.2), and the solution for consumption, Equation (4.3), have shown how aggregate inflation can affect consumption by affecting the real value of pay-offs of nominal assets and how heterogeneity in ex-ante positions on such assets can lead to heterogeneous effects on consumption. Moreover, assessing the effects of these external shocks on relative prices is critical to understanding relative consumption responses across various categories, as given in Equation (4.5) and Equation (4.7).

In addition, while not incorporated in our modelling framework, in standard sticky-price models, if external commodity price shocks lead to aggregate inflation, then by acting like "cost-push" shocks, they can cause a recession and lead to a fall in real income and wage earnings domestically. Empirically, there is evidence for such effects. For instance, De Winne and Peersman [2021] have established how global food price shocks, driven by adverse weather shocks, can negatively impact real economic activity in emerging economies.

The specification for the state-region level panel local projection regression to estimate dynamic effects on regional prices (and relative prices) of the external commodity price shocks is:

$$p_{s,r,t+h} - p_{s,r,t-1} = \beta_{0,u}^{h,\text{food}} ext_t^{\text{food}} \times \mathbb{1}_{r=urban} + \beta_{0,u}^{h,\text{oil}} ext_t^{\text{oil}} \times \mathbb{1}_{r=urban} + \gamma_h X_t + \sum_{d=0}^{D} \delta^h D_{t-d}$$
$$+ \sum_{j=1}^{J} \alpha_j^h (p_{s,r,t-j} - p_{s,r,t-j-1}) + \sum_{k=1}^{K} ext_{t-k}^{\text{food}} + \sum_{k=1}^{K} ext_{t-k}^{\text{oil}} + \theta_s + \zeta_r + \epsilon_{s,r,t+h} \tag{4.8}$$

where $p_{s,r,t}$ denotes (log) prices or relative prices in period t for state s and region r, h denotes the projection horizon, ext denotes different measures of the external commodity price shock, and $J = 1, K = 1$ are respectively the AR and MA coefficients in the specification.

Moreover, D is the dummy for the Indian government's demonetization policy. X denotes controls for aggregate world conditions: world industrial production as a proxy for aggregate demand (

We will present IV results where we instrument the changes in global food and oil prices by the corresponding supply shocks. These IV results will isolate variation coming from supply shocks to global food and oil prices as we discussed previously. We report cumulative impulse responses. Table A7 in Appendix E lists our control and instrumental variables.

We present results based on the IV specification.[18] Figure 6 shows that there is pass-

[18] The OLS results for comparison are in Appendix E in Figure A3. The first-stage F-statistics for these IV regressions are in Appendix E in Table A8.

through to consumer prices, both to the direct category prices (right column) as well as to overall prices (left column). Dynamic effects of global food price change on overall CPI very closely follows its effects on the food component of CPI.[19] Finally, global oil price shock passes through strongly to domestic energy prices (comprising of fuel, light and transportation cost) in India as well as to headline prices. These results show that increases in food prices can be expected to hurt the poor more because food features more prominently in their consumption basket.

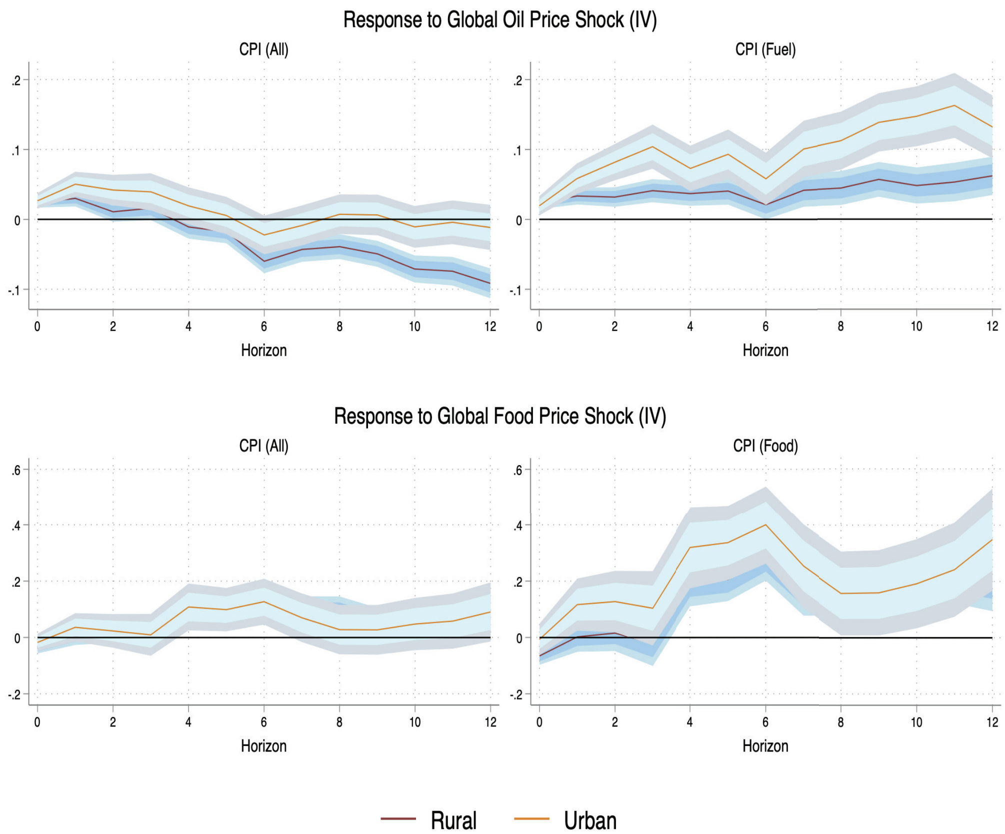


Figure 6: Response of State Level Prices to External Oil and Food Price Shocks (IV)

Notes: Cumulative IRFs on the basis of equation (4.8) where external shock is log changes in global oil price in the top panel and log changes in global food price in the bottom panel. These external price changes are instrumented by global supply shocks. The dependent variable is log changes in state level prices. The light blue region is the 90% confidence interval and the dark blue region is the 68% confidence interval.

Next, to illustrate the effects of these shocks on relative prices, in Figure 7 we show the responses of three relative prices: the ratio of food price to fuel price (left column), and

[19] This is to be expected given that on average, food constitutes around 45 percent of the CPI index share in India.

the ratio of food or fuel price to the non-durables consumption price (right column).[20] As is clear, global food price shocks clearly increase the relative price of food in India while global oil price shocks increase the relative price of fuel in India. These results hold across both urban and rural India.

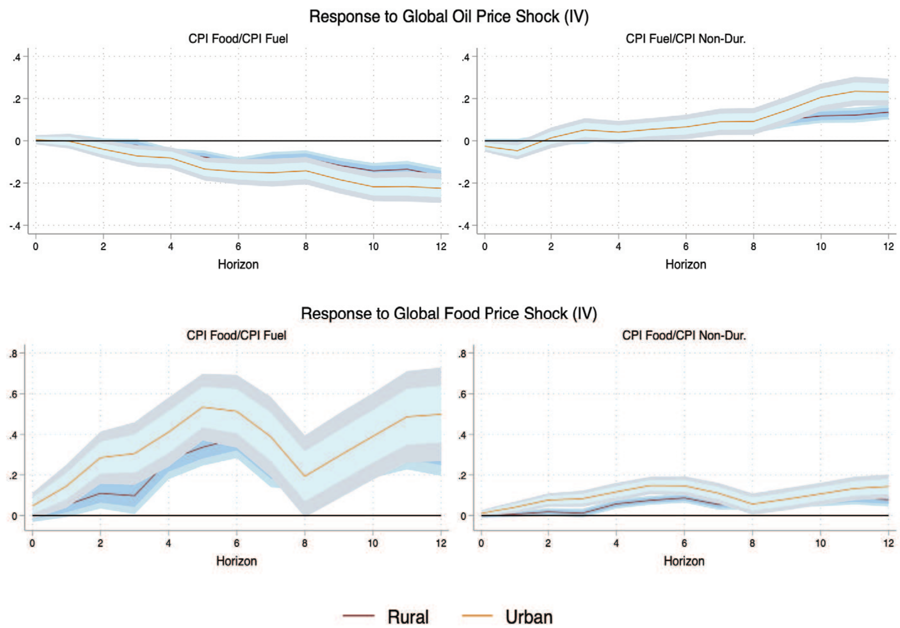


Figure 7: Response of State Level Relative Prices to External Oil and Food Price Shocks (IV)

Notes: Cumulative IRFs on the basis of equation (4.8) where external shock is log changes in global oil price in the top panel and log changes in global food price in the bottom panel. These external price changes are instrumented by global supply shocks. The dependent variable is log changes in state level relative prices. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

Finally, we look at relative price effects using more dis-aggregated food categories to investigate in more detail how the external price shocks pass-through to local Indian prices as well as to understand later the results on expenditure share effects. In Figure 8 we present results for relative price responses of various food components, as ratio to fuel prices, for the case of food price shocks. That is, Figure 8 presents in a more dis-aggregated form the results that we presented above in the second row of Figure 7. It shows that in response to an exogenous increase in global food prices, relative prices of many food categories (compared to fuel prices) increase. While the increase in relative price of food categories is broad-based, quantitatively, they appear particularly salient for certain food types, such as pulses, sugar, oil and fats, and vegetables. In addition, the increase in relative prices of food categories occurs in both rural and urban India.

[20]Note that the sum of food and fuel prices cover most of non-durable prices in our data.

Overall, these results confirm that external commodity price shocks have a strong impact on different components of regional inflation in India, changing both the general cost of living (for example, as captured by overall CPI) as well as relative prices (for example, as captured by food and fuel CPI ratios). This, certainly, has implications for consumer behaviour. Moreover, the effects on relative prices of both shocks suggests that relative consumption expenditures will be affected non-trivially by them.

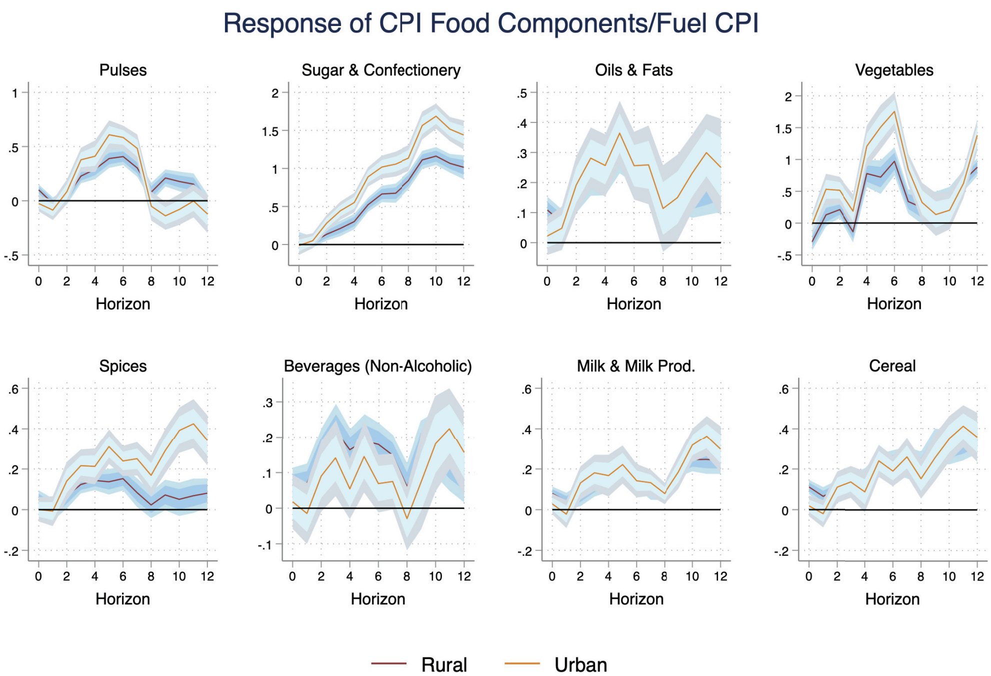


Figure 8: Response of State Level Relative Prices to External Food Price Shocks (IV)

Notes: Cumulative IRFs on the basis of equation (4.8) where external shock is log changes in global food price. The external food price changes are instrumented by global supply shocks. The dependent variable is log changes in state level relative prices, the ratio of various food category CPI to fuel CPI. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

4.2.3 Non-Homotheticity: Consumption Share Effects Within Non-Durables

Responses of food to fuel expenditure ratios We next investigate the effects on nominal consumption expenditure ratios within non-durables. Note that in Figure 7 we showed that relative food prices increase in response to global food price shocks while in Figure 2 we showed that real non-durable consumption expenditure falls in response to global food price shocks. Estimating effects on ratios of nominal consumption expenditures of food, with respect to non-durables and fuel, now allows us to assess if there are any effects that suggest non-homotheticity or if these share responses are simply consistent with expenditure switching due to relative price movements. Our theoretical framework showed how both of these channels can be captured by Equation (4.6) and Equation (4.7). In particular, Testable Prediction 3 in Section 4.1.2 had summarized that relative expenditure on food increasing in response to the food price shock, given the relative price of food and non-durable consumption expenditure responses, would be sufficient evidence for non-homotheticity.

Figure 9 presents results for food expenditure ratios with respect to non-durable consumption expenditure (top row) and fuel expenditure (bottom row). As is clear, in response to the global food price shock, food expenditure ratio increases for the lower income groups while it decreases for the two higher income groups. Given the relative price responses in Figure 7 and the decline in real non-druable consumption in Figure 2, as pointed out in Testable Prediction 3, these consumption share responses show a role for income effects in relative demand. In particular, we infer that food is unambiguously a necessity for the lower income groups. For the richer households, the response we find could be consistent with standard expenditure switching as relative expenditure on food falls. We can however still not rule out non-homotheticity as relative expenditure rising is only a sufficient condition not a necessary one, as stated in Testable Prediction 3.[21]

[21] We will address this issue below using some structural assumptions on the elasticity parameter that governs expenditure switching.

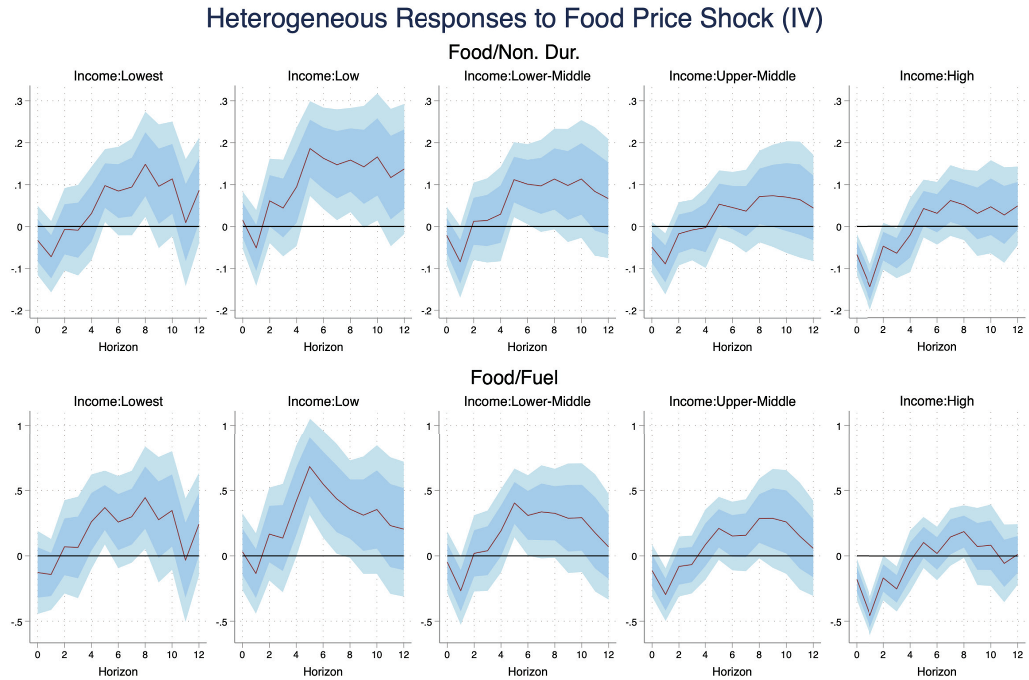


Figure 9: Response of Food Consumption Shares to External Food Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global food price, which is instrumented by a global food supply shock. In the top panel, the dependent variable is the ratio of household nominal food to non-durable consumption expenditures and in the bottom panel the dependent variable is the ratio of household nominal food to fuel consumption expenditures. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

Responses of detailed food categories to fuel expenditure ratios Given this evidence for food expenditure ratios, we next delve into expenditure ratio results for various food components. This is warranted as food expenditure is a composite of different food categories and so the average response might not be indicative of non-homotheticity for all types of food expenditures. Again, Testable Prediction 3 in Section 4.1.2 had summarized that relative expenditure on food components increasing in response to the food price shock would be a sufficient proof for non-homotheticity, as relative price of these food components increases (Figure 8) and total real non-durable consumption expenditure falls (Figure 2).

Figure 10 presents results for responses of various food components to fuel expenditure ratios. For reference, we also plot the relevant relative price response in the left column. It shows that the evidence for non-homotheticity in preferences of the poor (including the two lowest income groups) with respect to various food categories is quite clear for sugar, oil and fats, and vegetables as the expenditure ratios for these categories increase.

In addition, in these three food categories, for the rich, the expenditure ratio goes down. This might suggest expenditure switching, but as we emphasized above, it does not necessarily have to be inconsistent with non-homotheticity as expenditure ratios increasing is a sufficient condition not a necessary one.

Moreover, Figure 10 shows that the clear pattern of lack of expenditure switching by the poor (again including both the low income groups) is prominent also for other food categories, such as pulses and spices. Interestingly, for pulses and spices, the results suggest no expenditure switching even by the rich. Thus, pulses and spices are clearly an essential good for all income groups in India as the results for them satisfy the condition laid out in Testable Prediction 3 in Section 4.1.2.

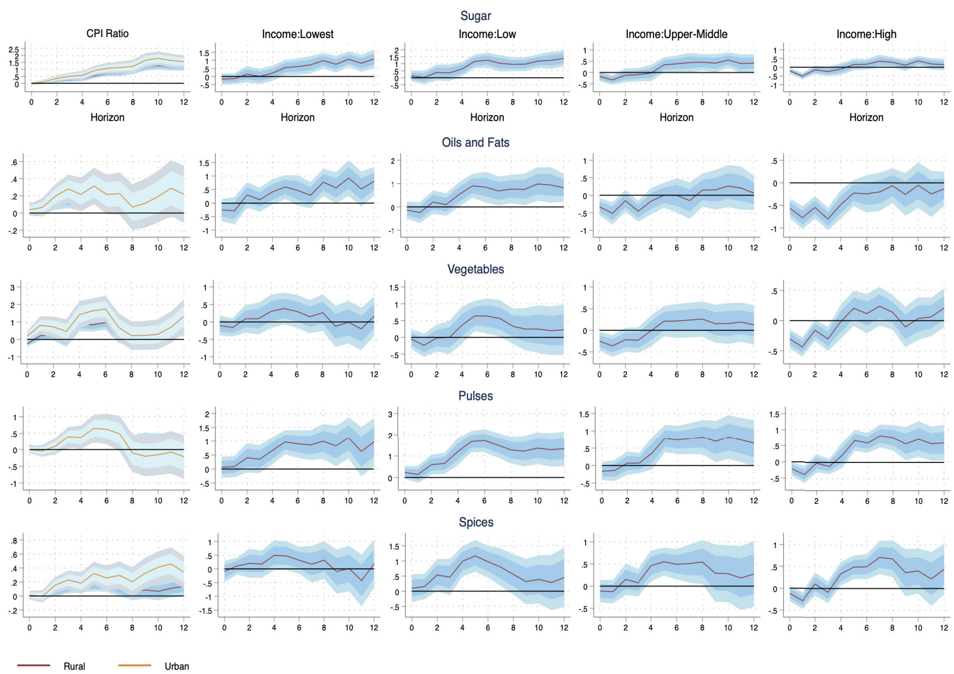


Figure 10: Response of Various Food Categories Consumption Shares to External Food Price Shocks by Income Quintiles(IV)

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global food price (bottom panel), which is instrumented by a global food supply shock. The dependent variable is the ratio of household nominal food categories to fuel consumption expenditures. The left column shows the responses of the relevant relative prices. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

Estimation of demand function parameters We now directly estimate the key demand function parameters in Equation (4.7). In particular, we fix $\sigma_\epsilon = 2$, allowing for some expenditure switching when relative prices change, and then estimate $\epsilon_i - \epsilon_j$ by matching the responses on relative expenditures and total real non-durable consumption expenditures to global price shocks.

The results are in Table 3. In Panel A we use the results on global good price shocks. First, we present the estimates based on total food to fuel expenditure ratio. That is, fixing $\sigma_\epsilon = 2$, we use the impulse responses of the food to fuel relative price from Figure 7, the food to fuel expenditure ratio from Figure 9, and real non-durable consumption expenditure from Figure 2, to estimate $\epsilon_i - \epsilon_j$ in Equation (4.7). As is clear, $\epsilon_i - \epsilon_j$ is estimated to be negative, showing that food is an essential good for all income groups. Next, we repeat the same exercise for various food components, using the appropriate food components to fuel relative prices and relative expenditure shares. It shows that various food components are an essential good for all income groups as $\epsilon_i - \epsilon_j$ is negative.

Some of the $\epsilon_i - \epsilon_j$ in Panel A are not estimated very precisely. One reason for this is that in each estimation exercise, we are only relying on 13 observations (the horizon of the impulse responses). In Panel B, we report estimates if we repeat the same exercise as in Panel A, but now use the impulse responses for both the global food and oil price shocks. As is clear, then $\epsilon_i - \epsilon_j$ are estimated more precisely to be negative.

Table 3: Estimates of Demand Function Parameters ($\epsilon_i - \epsilon_j$)

	Lowest	Low	Low-middle	Upper-middle	High
	Panel A: Estimated Using IRFs From Only Food Shocks				
Food (All)	-0.524**	-0.613**	-0.640**	-0.852**	-0.917**
	(0.156)	(0.186)	(0.184)	(0.212)	(0.224)
Sugar	-2.422***	-2.656***	-2.702***	-3.112***	-3.271***
	(0.257)	(0.223)	(0.297)	(0.477)	(0.440)
Oils and Fats	-0.595**	-0.801***	-0.581**	-0.640**	-0.702**
	(0.146)	(0.158)	(0.135)	(0.153)	(0.165)
Vegetables	-0.648	-1.109	-0.995	-1.585*	-1.613
	(0.520)	(0.544)	(0.579)	(0.689)	(0.780)
Pulses	-0.232	-0.595	-0.357	-0.414	-0.555
	(0.317)	(0.431)	(0.414)	(0.498)	(0.544)
Spices	-0.096	-0.333	-0.382	-0.569*	-0.871**
	(0.148)	(0.219)	(0.177)	(0.222)	(0.248)
	Panel B: Estimated Using IRFs From Food and Oil Shocks				
Food (All)	-0.548***	-0.654***	-0.674***	-0.807***	-0.770***
	(0.106)	(0.091)	(0.123)	(0.156)	(0.168)
Sugar	-1.928***	-2.069***	-2.080***	-2.287***	-2.252***
	(0.170)	(0.102)	(0.180)	(0.270)	(0.288)
Oils and Fats	-0.726***	-0.896***	-0.605***	-0.484***	-0.417*
	(0.097)	(0.071)	(0.070)	(0.097)	(0.166)
Vegetables	-0.806**	-1.207***	-1.184***	-1.576***	-1.524**
	(0.272)	(0.209)	(0.308)	(0.408)	(0.481)
Pulses	-0.852**	-1.391***	-1.015***	-1.116**	-1.156**
	(0.235)	(0.218)	(0.257)	(0.327)	(0.385)
Spices	-0.129	-0.566***	-0.398***	-0.721***	-0.958***
	(0.095)	(0.123)	(0.103)	(0.185)	(0.227)

Notes: This Table reports the estimates of $\epsilon_i - \epsilon_j$ obtained by estimating Equation (4.7) in a regression framework after fixing $\sigma_\epsilon = 2$ and using as data the impulse responses of relative prices, relative expenditures, and real non-durable consumption expenditure. Each row represents estimates from a separate regression, with 13 observations used in Panel A and 26 observations used in Panel B. The columns represent the various income groups. $\epsilon_i - \epsilon_j < 0$ indicates that good i (food and various food categories here) is an essential consumption good.

5 Discussion, Sensitivity Analyses, and Extensions

In this section, we discuss some features of the empirical results that we have presented so far that demand further attention and context. We also discuss some key sensitivity analyses to show that our key conclusions regarding heterogeneous consumption impact of global price shocks is robust. Finally, we discuss some extensions that provide complementary evidence.

5.1 OLS Results on Consumption and Income

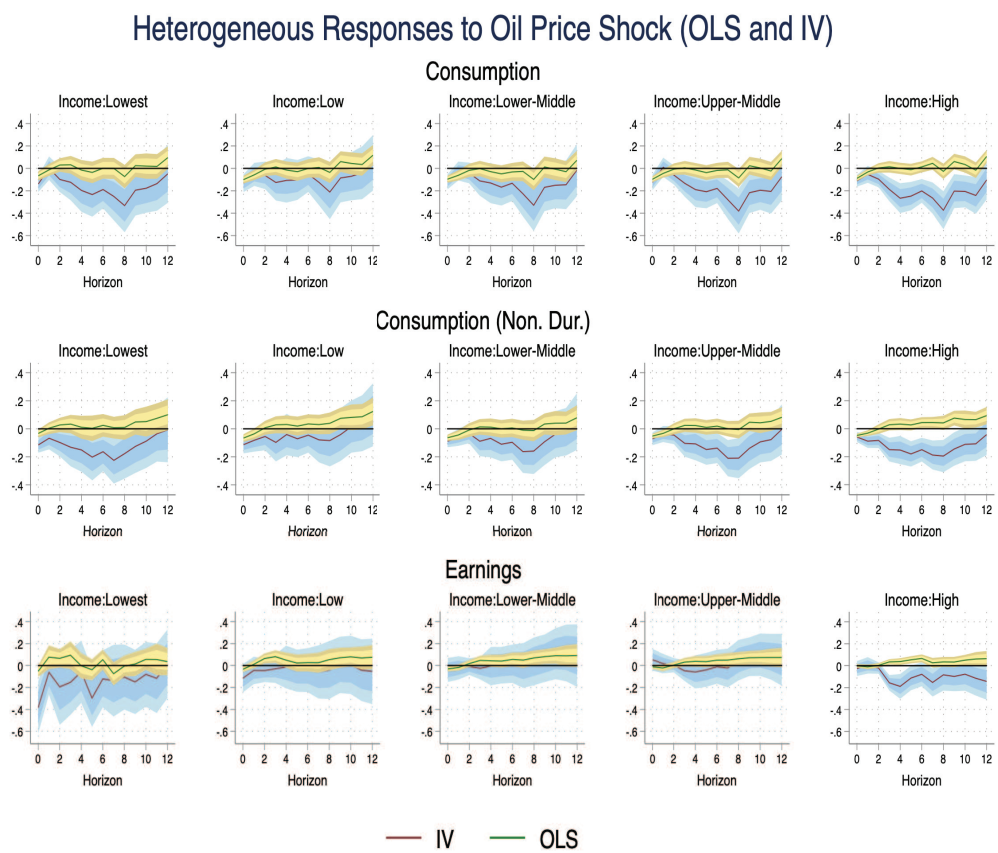


Figure 11: Response of Consumption and Earnings to External Oil Price Shocks

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global oil price. In the IV version, the log changes in global oil price is instrumented by a global supply shock. The dependent variable is log changes in household consumption, non-durable consumption, and labor earnings. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

As we discussed before, our IV estimates isolate global supply side variation in global commodity prices, which is important in order to both avoid mixing in the effects of supply and global demand shocks as well as guarding against omitted variable bias problems. We now further show the importance of the IV estimates by presenting the OLS results for oil price shocks. Based on household data and estimation of equation (3.1), Figure 11 present the OLS results to a key question of the paper: How does the dynamic response of consumption and income to global oil price shocks vary by ex-ante income quintiles? We show results for total and non-durable consumption as well as wage income. For comparison, we also plot the IV results we presented earlier.

As is clear, in the OLS results, as predicted, the consumption effects are not consistently and persistently negative for any income group. Moreover, the earnings effects are even positive for the higher income groups. Thus, as expected, richer Indian households benefit with relatively higher income/earnings if higher oil prices are caused by an increase in global demand, which in turn affects the Indian economy positively. Thus, it is imperative to isolate demand and supply side variation while studying distributional consequences of global price shocks.[22]

5.2 Expenditure Switching: Non-Durable Consumption Share Effects

We investigate the effects on nominal consumption expenditure ratio of non-durable to total consumption to show that expenditure switching is a reasonable model assumption. In Figure A5 we show that global food price shocks increase the relative price of non-durables while global oil price shocks, after a delay, decrease the relative price of non-durables. Estimating effects on ratios of nominal consumption expenditures on non-durables to total consumption now allows us to assess if these share responses are consistent with expenditure switching due to relative price movements, as suggested by the theoretical framework in equation (4.5) and given in Testable Prediction 2 in Section 4.1.2
.

Thus, in the household panel IV local projection framework we estimate equation (3.1), but with the nominal expenditure share of non-durable consumption to total consumption as the dependent variable. Figure A5 presents results for the response of non-durable to total consumption expenditure ratio for both external price shocks. The results show

[22]As Kilian [2009] originally noted for aggregate effects, here, for distributional effects too, not all oil shocks are alike! Similar, albeit not statistically significant, differences between OLS and IV results are observed in the case of food price changes in Figure A4.

that consistent with expenditure switching that comes about due to relative price changes, this ratio increases for the global oil shock, while it decreases for the global food shock.[23] In addition, response of the non-durable consumption shares are very similar across various income groups, suggesting that relative price movements are the main determinant, as appropriately captured by a homothetic CES aggregator in equation (4.5).

5.3 Discussion

We start by discussing some aspects of our results that need further elaboration and also put them in context with related literature. First, in Figure A6 we compute summary statistics to compare earnings and non-durable consumption responses from Figures 2, 3, and 5. For all income groups other than the poorest, effects on consumption are larger in magnitude than that on labor earnings. What can explain such a large consumption response? From the point of view of the theoretical model presented in Section 4.1.1, and in particular, the solution for consumption in equation (4.3), a relatively larger response of consumption (compared to the present discounted value of labor earnings using a constant discount factor) may arise either due to time-varying interest rates and wealth effects or due to asset price changes that alter the real value of payoffs from ex-ante asset positions. We, in fact, do find evidence of such effects in Indian macro data, as presented in Figure A7. A rise in global food and oil prices, instrumented by corresponding supply shocks in an aggregate time series local projection framework, leads to a rise in India's short term interest rate spread (compared to the U.S. benchmark) and lowers stock prices. Such an effect on interest rates is further confirmed using detailed annual bank-branch level (weighted) average lending rate data in a panel regression in Table A9.[24]

We also note that our results showing a strong response of consumption to these shocks (even non-durable consumption), compared to wage income, connect to the unconditional stylized facts from the emerging market business cycle literature. For instance, for a large sample of countries, Uribe and Schmitt-Grohe [2017] documents that consumption growth is clearly more volatile than income growth for emerging market economies and that consumption-smoothing in this sense appears to be limited. Using Indian annual data (1965-2010), Uribe and Schmitt-Grohe [2017] finds that relative volatility of consumption is higher than income (relative volatility $\frac{\sigma_c}{\sigma_y}$: 1.07) in India, while we reach the

[23]The results for the food price shock are comparatively noisier towards the end of the impulse response horizon.

[24]Data source of the detailed interest rate data is Basic Statistical Return of Reserve Bank of India, 1998-2016.

same conclusion using first-differenced (relative volatility $\frac{\sigma_c}{\sigma_y}$: 1.94) or HP filtered (relative volatility $\frac{\sigma_c}{\sigma_y}$: 1.35) quarterly national income accounts data (1996:Q2 to 2019:Q4).

Another important insight emerges from Figure A6 while comparing the effects on earnings with the effects on non-durable consumption for food price shocks. Earnings effects of the exogenous rise in food price are clearly much larger for the poorest than for the poor (Figure 5), but non-durable consumption effects are comparable (Figure 4). This suggests a role for the public distribution system as an insurance mechanism for those below the poverty line.

Second, as seen in Figure 5, what might be an economic mechanism at work that gives a clear role for the earnings channel for food price shocks? This can be rationalized based on the fact that most of the poorest work in the informal sector in India, as documented in Table A10, where there is no inflation adjustment in nominal wage income.[25] In other words, inflation is very likely to outstrip nominal wage growth for the poor. With a rise in oil prices and its broad-based price impact, there is more likely an immediate impact on the cost of living of the middle class, a large fraction of whom work in the more organized, formal sector. Consequently, there is a nominal wage adjustment of the middle class which then percolates to the nominal income of the poor.[26] However, with a rise in global food price and its direct effect on domestic food prices, cost of living of the poor rises disproportionately due to the higher share of food in the consumption basket of the poor.[27] This however, does not set off the same nominal wage adjustment process in the informal sector and as a result, we observe a decline in the real earnings of the poor.

Third, in a summary statistics format, in Figure A8 we compare the non-durable and total consumption responses to global food and oil price shocks from the IV panel results presented in Figures 2 and 3. Remarkably, the elasticity of total consumption is uniformly larger than that of non-durable consumption, pointing towards larger response of durable consumption to external price shocks.[28]

[25]Note that the classification of formal and informal occupations for CMIE data is from IIM Bangalore doctoral student Shweta Shogani's ongoing PhD thesis work. We are extremely grateful to her for sharing this classification.

[26]We do observe a initial decline in the real earnings of the poor immediately after the oil supply shock as well.

[27]Summary statistics for these various income groups, including share of food consumption, are presented in Appendix B, which show this pattern.

[28]Note that total consumption includes non-durable, durable, and service consumption. The excess volatility of durable consumption is a well-known empirical fact in other contexts (see, for example, Alvarez-Parra, Brandao-Marques, and Toledo [2013]) and can be rationalized in canonical models of consumption smoothing.

5.4 Sensitivity Analyses

The answer to our key research question: does household consumption response in India to global price shocks differ by income levels, is a clear yes. Moreover, different global shocks lead to different patterns of consumption heterogeneity: global food prices have a monotonically larger impact on poorer segments of the population, whereas global oil prices affect most clearly the lowest and the highest income groups.

Naturally, our answer may be sensitive to how we assign individuals to different income groups. In our baseline results summarized in Figure 4, households are grouped according to cut-offs based on total household real income in the initial period, 2014. Moreover, while the definition of the groups is on the basis of the initial income distribution, depending on current income, households can and do transition to a different income group over time. In this section, we report two important sensitivity analyses of our baseline results where we change the definition of income groups.

In the first exercise, instead of total household real income, we group households according to *per capita* household real income in the initial period. Because average household sizes differ by income groups, per capita household income may more accurately capture the resources available to household members, as argued in Deaton [2019]. Indeed, characteristics of households by per capita income deciles, as reflected in Table A11, are somewhat different from those reported in our baseline summary statistics in Table A1. In order to account for this, we group households into five income groups according to per capita income deciles and estimate the heterogeneous consumption responses according to equation (3.1). Summary statistics of non-durable consumption response, estimated using panel IV regressions, are in Figure A9. The results in Figure A9 reflect the same pattern of heterogeneous consumption response as in Figure 4.

In the second exercise, we retain the grouping according to total household real income in the initial period, but we restrict the transition matrix. The baseline transition matrix across income groups is presented in Table A12. While more than 80% of households remain in the same income group over time (as captured by the diagonal entries of Table A12), there are some households who transition from the highest to lowest income groups. Such a transition can potentially reflect measurement error. In order to restrict such unusual movements, we estimate the baseline panel IV local projection framework of equation 3.1 while restricting the transition matrix such that no household is allowed to move more than two (absolute) steps in the transition matrix. The resulting summary statistics of non-durable consumer response is in Figure A10. These results again are similar in nature to the baseline results of Figure 4.

Thus, alternate definitions of income groups leave our key conclusions regarding heterogeneous household consumption response to global price shocks unchanged. While everyone suffers consumption losses due to rising food prices, poorer income groups are far more vulnerable to such food price shocks. In contrast, the lowest and highest income groups suffer equally from an increase in global oil prices.

5.5 Extensions

In this section, motivated by the vast regional heterogeneity of India, we conduct two extensions of our baseline empirical framework.

In the first exercise, in our baseline household regression as given in equation (3.1), we allow the impact of global price shocks to differ not just by income groups, but also by the location (rural or urban) of the household. We may reasonably expect these responses to differ by location because, among other stark differences, households who work in agriculture primarily reside in rural areas and may be differentially impacted by food price shocks. Our extended specification is as follows:

$$
\begin{aligned}
c_{i,t+h} - c_{i,t-1} = {} & \beta^{g,h,r}_{0,\text{food}} ext^{\text{food}}_t \times \mathbb{1}_{i \in g(t)} \times \mathbb{1}_{i \in \text{urban}} + \beta^{g,h,r}_{0,\text{oil}} ext^{\text{oil}}_t \times \mathbb{1}_{i \in g(t)} \times \mathbb{1}_{i \in \text{urban}} \\
& + \sum_{j=1}^{J} \alpha^h (c_{i,t-j} - c_{i,t-j-1}) + \sum_{k=1}^{K} \beta^h_{k,\text{food}} ext^{\text{food}}_{t-k} + \sum_{k=1}^{K} \beta^h_{k,\text{oil}} ext^{\text{oil}}_{t-k} + \sum_{d=0}^{D} \delta^h D_{t-d} \\
& + \gamma^{g,h} X_t \times \mathbb{1}_{i \in g(t)} + \delta_{\text{c},t} + \delta_{\text{l},t} + \delta_{\text{e},t} + \delta_{\text{city},t} + \delta_{\text{age},t} + \mathbb{1}_{i \in \text{s}} \times \mathbb{1}_{\text{year}} + \mathbb{1}_{i \in \text{s}} \times \mathbb{1}_{\text{month}} + \epsilon_{i,t+h}
\end{aligned}
\tag{5.1}
$$

where $\mathbb{1}_{i \in \text{urban}}$ captures whether household i resides in an urban area, $r =$ rural, urban.

Results for consumption responses are presented in Figures A11 and A12, and the corresponding earnings results are in Figure A13. Low income groups in rural areas seem to be suffer a larger (compared to their urban counterparts) non-durable consumption loss due to a global food price shock, while the high income households in rural areas suffer a larger (again, compared to their urban counterparts) consumption loss due to a global oil price shock. On the earnings side, the sharp drop in earnings for the low income households due to an increase in global food prices seems to be driven by urban households. Apart from these differences, the broad pattern of responses are largely similar across urban and rural India.

In the second exercise, we assess the effects on regional inequality of the global commodity price shocks, which we make operational by constructing several measures of regional inequality. The specification for the state-level panel local projection regression to estimate dynamic effects on regional consumption inequality of the external commodity price shocks is:

$$Cineq_{s,t+h} - Cineq_{s,t-1} = c + \sum_{j=1}^{J} \alpha_j^h (Cineq_{s,t-j} - Cineq_{s,t-j-1}) + \sum_{k=0}^{K} \beta_k^h ext_{t-k} + \sum_{d=0}^{D} \delta^h D_{t-d} + \gamma_h X_t + \theta_s + \delta_t + \epsilon_{s,t+h} \tag{5.2}$$

where $Cineq_{s,t}$ denotes various measures of state-level inequality (in log) for total consumption and non durable consumption in period t, h denotes the projection horizon, ext denotes different measures of the external commodity price shock, and $J = 1, K = 1$ are respectively the AR and MA coefficients in the specification. Finally, our specification includes state and time fixed-effects. Standard errors are clustered at the state level. This specification is similar to the state price regression specification outlined in equation (4.8).

Figure A14 presents the IV results for the food price shock. It shows that consumption inequality increases robustly following an exogenous change in global food prices. Figure A15 presents the IV results for the oil price shock. It shows that an increase in global oil prices does not have as clear of an effect on consumption inequality as does global food prices, suggesting that traditional inequality measures might not capture the subtle ways in which household get differentially affected along the income distribution by exogenous oil price shocks.

6 Conclusion

In this paper, we explore the distributional implications of the increasing global food and oil prices by utilizing rich consumption and income panel data from India. Our results show robust evidence that lower income deciles are more affected by an exogenous increase in food prices, while both tails of the income distribution are affected similarly by an exogenous increase in fuel prices. We also find that these heterogeneous consumption responses largely mirror the pattern of heterogeneity in earnings response to these global price shocks.

Examining relative expenditure responses, in light of relative price effects, allows us to uncover very interesting patterns of non-homotheticity in non-durable consumption. We find food, compared to fuel, is an essential consumption good for all income groups in India. Our analysis provides a novel way of identifying essential consumption components by relying on a non-homothetic isoelastic CES demand structure and impulse response matching using external instruments.

Our findings have significant implications for monetary policy. The substantial distributional effects on consumption that we have documented suggest that in emerging markets, monetary policy may need to react to external shocks in the food and oil sectors, despite the flexibility of prices in these sectors. This response would be advantageous in terms of reducing consumption inequality in the economy. In our future research, we intend to further investigate this matter by studying optimal monetary policy in a heterogeneous agent open economy model.

References

Sushant Acharya, Edouard Challe, and Keshav Dogra. Optimal monetary policy according to hank. 2020.

Ron Alquist, Saroj Bhattarai, and Olivier Coibion. Commodity-price comovement and global economic activity. *Journal of Monetary Economics*, 112:41–56, 2020.

Fernando Alvarez-Parra, Luis Brandao-Marques, and Manuel Toledo. Durable goods, financial frictions, and business cycles in emerging economies. *Journal of Monetary Economics*, 60(6):720–736, 2013.

Adrien Auclert. Monetary policy and the redistribution channel. *American Economic Review*, 109(6):2333–67, June 2019. doi: 10.1257/aer.20160137. URL https://www.aeaweb.org/articles?id=10.1257/aer.20160137.

Robert B. Barsky and Lutz Kilian. Oil and the macroeconomy since the 1970s. *Journal of Economic Perspectives*, 18(4):115–134, 2004. URL https://ideas.repec.org/a/aea/jecper/v18y2004i4p115-134.html.

Christiane Baumeister and James D. Hamilton. Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review*, 109(5):1873–1910, 2019. doi: 10.1257/aer.20151569. URL https://www.aeaweb.org/articles?id=10.1257/aer.20151569.

Saroj Bhattarai, Jae Won Lee, and Woong Yong Park. Optimal monetary policy in a currency union with interest rate spreads. *Journal of International Economics*, 96(2):375–397, 2015. URL https://EconPapers.repec.org/RePEc:eee:inecon:v:96:y:2015:i:2:p:375-397.

Saroj Bhattarai, Arpita Chatterjee, and Woong Yong Park. Global spillover effects of US uncertainty. *Journal of Monetary Economics*, 114(C):71–89, 2020. doi: 10.1016/j.jmoneco.2019.05. URL https://ideas.repec.org/a/eee/moneco/v114y2020icp71-89.html.

Arpita Chatterjee. Globalization and monetary policy comovement: International evidence. *Journal of International Money and Finance*, 68:181–202, 2016.

Richard Clarida, Jordi Galı, and Mark Gertler. A simple framework for international monetary policy analysis. *Journal of monetary economics*, 49(5):879–904, 2002.

Olivier Coibion, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. Innocent bystanders? Monetary policy and inequality. *Journal of Monetary Economics*, 88(C):70–89,

2017. doi: 10.1016/j.jmoneco.2017.05. URL https://ideas.repec.org/a/eee/moneco/v88y2017icp70-89.html.

Jasmien De Winne and Gert Peersman. Macroeconomic effects of disruptions in global food commodity markets: Evidence for the United States. *Brookings Papers on Economic Activity*, 47(Fall):183–286, 2016.

Jasmien De Winne and Gert Peersman. The adverse consequences of global harvest and weather disruptions on economic activity. *Nature Climate Change*, 11(8):665–672, 2021.

Angus Deaton. *The analysis of household surveys (reissue edition with a new preface): A microeconometric approach to development policy*. World Bank Publications, 2019.

Jesus Fernandez-Villaverde, Pablo Guerron-Quintana, Juan F. Rubio-Ramirez, and Martin Uribe. Risk matters: The real effects of volatility shocks. *American Economic Review*, 101 (6):2530–61, 2011. doi: 10.1257/aer.101.6.2530. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.6.2530.

Michael Gelman, Yuriy Gorodnichenko, Shachar Kariv, Dmitri Koustas, Matthew D Shapiro, Dan Silverman, and Steven Tadelis. The response of consumer spending to changes in gasoline prices. *American Economic Journal: Macroeconomics*, 15(2):129–160, 2023.

James D Hamilton. Oil and the macroeconomy since World War II. *Journal of Political Economy*, 91(2):228–248, 1983. doi: 10.1086/261140. URL https://ideas.repec.org/a/ucp/jpolec/v91y1983i2p228-48.html.

James D. Hamilton. What is an oil shock? *Journal of Econometrics*, 113(2):363–398, 2003. URL https://ideas.repec.org/a/eee/econom/v113y2003i2p363-398.html.

Martin Blomhoff Holm, Pascal Paul, and Andreas Tischbirek. The transmission of monetary policy under the microscope. *Journal of Political Economy*, 129(10):2861–2904, 2021. doi: 10.1086/715416. URL https://doi.org/10.1086/715416.

Diego R Känzig. The unequal economic consequences of carbon pricing. *Available at SSRN 3786030*, 2021.

Lutz Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–1069, 2009. URL https://ideas.repec.org/a/aea/aecrev/v99y2009i3p1053-69.html.

M Ayhan Kose, Christopher Otrok, and Charles H Whiteman. Understanding the evolution of world business cycles. *Journal of international Economics*, 75(1):110–130, 2008.

Aeimit Lakdawala and Sanjay Singh. The effects of foreign shocks in the indian economy. *India Policy Forum*, 6, 2019.

Kiminori Matsuyama. Non-ces aggregators: a guided tour. *Buffett Institute Global Poverty Research Lab Working Paper Forthcoming*, 2022.

Pablo A. Neumeyer and Fabrizio Perri. Business cycles in emerging economies: The role of interest rates. *Journal of Monetary Economics*, 52(2):345–380, 2005. URL https://ideas.repec.org/a/eee/moneco/v52y2005i2p345-380.html.

Gert Peersman. International food commodity prices and missing (dis)inflation in the Euro Area. *The Review of Economics and Statistics*, 104(1):85–100, 01 2022. ISSN 0034-6535. doi: 10.1162/rest_a_00939. URL https://doi.org/10.1162/rest_a_00939.

Gert Peersman and Joris Wauters. Heterogeneous household responses to energy price shocks. 2022.

Gert Peersman, Sebastian K Rüth, and Wouter Van der Veken. The interplay between oil and food commodity prices: has it changed over time? *Journal of International Economics*, 133:103540, 2021.

John W Tukey et al. *Exploratory data analysis*, volume 2. Reading, MA, 1977.

Martin Uribe and Stephanie Schmitt-Grohe. *Open Economy Macroeconomics*. Princeton University Press, Princeton, New Jersey, 2017.

Appendix A Data Description

Survey Data

We use data from the Consumer Pyramid Household Survey (CPHS) dataset, a survey conducted by the Centre for Monitoring the Indian Economy (CMIE). The CPHS has surveyed over 236,000 unique households since it began in 2014, and is the most comprehensive longitudinal consumption data available for India. The CPHS is itself divided into 4 distinct datasets: Consumption Pyramids, Income Pyramids, People of India Survey and Aspirational India survey. We use the data from the Consumption and Income Pyramid surveys to construct our variables, and data from the People of India survey for our control variables about demographics. Our analysis spans data from January 2014 to December 2019.

Level variables

We construct income, earnings, and consumption categories closely following the definitions given by Coibion and Gorodnichenko (2017). We first construct income as the sum of household income from rent, wages, self-production, private transfers, government transfers, business profits, sale of assets, lotteries and gambling, pension, dividends, interest and deposit provident fund and insurance. These categories are an exhaustive list of all income sources collected in the CPHS survey.

Additionally, we construct narrow and broad measures of capital income. The narrow measure of capital income includes income accrued from dividends, interest and business, whereas the broad measure includes the income sources from the narrow measure as well as income accrued from sale of assets and rent.

Our (labor) earnings measure is constructed using only the category of income from wages and overtime bonus.

We construct consumption closely matching the categories constructed by Coibion and Gorodnichenko (2017). The consumption variable we construct is the sum of non-durable consumption (food, fuel, intoxicants), durable consumption (appliances, furniture, jewelry, clothing, electronics, toys, cosmetics), and service consumption (electricity, entertainment, public transport, airfare, highway tolls, beauty services, fitness services, restaurants etc). We then deflate all our income and earnings measures by the Consumer Price Index (CPI) - Combined series (2012 base). We also winsorize our constructed variables at the 1 percent level.

Inequality variables

The measures of inequality we construct using these variables are: Gini coefficients, cross-sectional standard deviations and differences between individual percentiles (90th-10th and 75th-25th) on log levels.

External shock measure

We use IMF's Global Price of Food Index (Nominal USD) and WTI crude oil prices (USD) at monthly frequency as our data for global food and oil prices. We construct shocks by taking differences of the logs of both food and oil prices.

Data on Prices

The Ministry of Statistics and Programme Implementation (MoSPI), Government of India, releases detailed data on prices at the monthly frequency. The series has 2012 as the base year and data is available from January 2011. The data is dis-aggregated by geography as well as by products. Geographically, data is available by urban and rural areas within each state. There is some missing data at the state-geography level, but it is not a major concern (97% of India's consumption is covered in the state-geography data).

On the product side, aggregate CPI is broken down into six broad sub-classifications (national level weights are in parenthesis): i) food and beverages (45.86%); ii) pan, tobacco, and intoxicants (2.38%); iii) clothing and footwear (6.53%); iv) housing (10.07%); v) fuel and light (6.84%); and vi) miscellaneous (28.32%). The coverage (in terms of sub-products) varies across the sub-classifications. The most detailed data is available only for food categories. It comprises of i) cereals and products; ii) meat and fish; iii) egg; iv) milk and milk products; v) oils and fats; vi) fruits; vii) vegetables; viii) pulses and products; ix) sugar and confectionery; x) spices; xi) non-alcoholic beverages; and xii) prepared meals, snacks, sweets, etc.

We construct price indexes for fuel and non-durables. Although the MoSPI provides an index for fuel, it only includes fuel used for cooking and excludes the fuel used in transportation; the index for transportation is available under the "miscellaneous (transportation and communication)" category. While this category has several missing values at the state-geography level, the missing values are concentrated among smaller states (such as Andaman and Nicobar islands) that contribute to under 3% of India's consumption. We use the state-geography level weights of fuel and light (FL) and miscellaneous (transportation and communication, or TC) categories to construct a new composite index:

$$CPI(FL+TC)_{sgt} = \frac{W(FL)_{sg}CPI(FL)_{sgt} + W(TC)_{sg}CPI(TC)_{sgt}}{W(FL)_{sg} + W(TC)_{sg}}$$

where subscript s represents state, $g \in \{Urban, Rural\}$ represents geography, and t represents month. This provides a closer measure of energy consumption.

The non-durable price index includes food and the composite fuel prices. It is calculated as

$$CPI(NonDur.)_{sgt} = \frac{W(Food)_{sg}CPI(Food)_{sgt} + W(FL)_{sg}CPI(FL)_{sgt} + W(TC)_{sg}CPI(TC)_{sgt} + W(Pan)_{sg}CPI(Pan)_{sgt}}{W(Food)_{sg} + W(FL)_{sg} + W(TC)_{sg} + W(Pan)_{sg}}$$

Appendix B Summary Statistics by Income

Table A1: Summary Statistics by Income Decile

	No. of Hhs	Income	Consumption	Non-durable Share	Food Share
1st Decile	43,210.66	476.65	1,221.19	0.78	0.62
2nd Decile	8,034.27	4,365.17	4,168.35	0.78	0.64
3rd Decile	10,832.47	5,579.25	4,708.03	0.77	0.63
4th Decile	12,049.10	6,720.92	5,190.72	0.77	0.61
5th Decile	12,673.92	7,972.92	5,669.16	0.76	0.60
6th Decile	14,400.27	9,613.48	6,093.47	0.76	0.59
7th Decile	15,089.84	11,736.90	6,655.97	0.75	0.58
8th Decile	15,101.95	14,850.60	7,345.38	0.74	0.56
9th Decile	15,345.27	20,166.59	8,302.80	0.73	0.55
10th Decile	19,239.84	37,353.66	10,447.05	0.72	0.51

Notes: This table presents some summary statistics by income deciles. Income and consumption are in real terms where they are deflated by the CPI (all, 2012=100). Non-durable and food share refer to consumption shares of non-durable and food consumption.

Table A2: Income Composition by Income Groups

Income Group	Earnings	Transfers	Capital Income (Broad)	Pensions
Lowest	.63	.16	.03	.03
Low	.88	.04	.05	.02
Lower middle	.84	.02	.09	.04
Upper middle	.75	.02	.15	.08
High	.67	.01	.21	.10

Notes: This table presents some summary statistics by income groups, where it shows shares of various sources of income.

Table A3: Socio-Economic Variables by Income Quintiles

	Q_1	Q_2	Q_3	Q_4	Q_5
	Quintile Shares (%)				
Panel A: Religion					
Buddhist	26	13	16	20	25
Christian	30	10	14	20	26
Hindu	30	14	16	18	21
Jain	21	4	10	17	47
Khasi	32	5	6	20	36
Muslim	30	16	19	19	15
Not Applicable	25	1	5	8	62
Other Religion	33	11	14	17	25
Religion not stated	50	9	9	11	20
Sikh	24	7	12	19	38
Panel B: Caste Category					
Intermediate Caste	32	10	14	17	27
Not Applicable	25	1	5	8	62
Not Stated	35	7	11	18	29
OBC	30	15	18	19	18
SC	29	19	19	19	14
ST	41	17	16	14	12
Upper Caste	27	10	14	19	31
Panel C: Education Category					
Upto 7th Std	34	18	19	18	12
Upto 12th Std	28	13	17	20	22
$\geq$ College Graduate	24	5	9	17	44
Panel D: Region					
Urban	25	13	17	20	25
Rural	39	16	16	15	13

Notes: This table presents summary statistics on some socio-economic variables by income deciles.

Appendix C Estimation of the Food Price Shock IV

Table A4: Non-energy Prices used in estimating the Dynamic Factor Model

Food	Food	Industrial Metals
Rice	Wheat	Iron ore
Bananas	Barley	Aluminium
Beef	Cocoa	Copper
Coffee Arabica	Coffee Robust	Cotton
Fishmeal	Corn	Lead
Poultry	Fish	Soft logs
Shrimp	Sugar	Hard logs
Orange	Tobaco	Nickel
Tea	Olive Oil	Rubber
Palm Oil	Rapseed Oil	Tin
Soybean Oil	Sunflower Oil	Wool coarse
Groundnut oil	Coconut oil	Wool fine
		Zinc

Notes: Commodity prices data are collected from FRED and Bloomberg and are quoted in US Dollar per unit. The units differ by commodity, but in the estimation we only use the log difference in price levels, i.e., returns.

In order to estimate the food shock, we first estimate a dynamic factor model with one common factor and two sector specific factors in a panel of 37 non-energy commodity prices (see, Table A4) which comprise of 13 industrial metals and 24 food prices. The dynamic factor model can be described as:

$$r_{i,t} = B^i f_t + C^i S_{j,t} + \eta_{i,t}$$

where $r_{i,t}$ is the log difference in commodity price, f_t is the common factor and $S_{j,t}, j = 1, 2$ is the sector-specific factor, and $\eta_{i,t}$ is the idiosyncratic component. Both the common factor and the sector-specific factors follow an AR(1) process. In order to identify the common factor as an aggregate demand factor, following the interpretation of Alquist et al. [2020], we impose the sign restriction that the factor loadings of the common factor, $B^i \geq 0$. Similarly, we interpret the food specific factor as the common demand factor for the food sector. Along with identification restriction, we also need to impose a normalization restriction, in order to overcome the well-known problem of unidentified models resulting from rotational indeterminacies of factors and loadings. Following Kose, Otrok, and Whiteman [2008], we normalize the contemporaneous factor loading of the iron ore for the common factor, and the contemporaneous factor loading of poultry for the food factor, to unity.[29]

We cast the dynamic factor model in the state space form and estimate it using Bayesian methods using Markov Chain Monte Carlo (MCMC). Two approaches have become popular for the estimation and identification of factor models: the analysis of principal components and Bayesian methods. Due to its simplicity and the availability of high speed computers, principal component analysis is extensively used for both static and dynamic factor models, extending to models using hundreds of series. As Kose et al. [2008] explain, principal component method is, however, not well suited for estimating models under exclusion restrictions. Model estimation using principal component requires deriving factors from the variance or spectrum of all series simultaneously, and therefore, it becomes inappropriate when a subset of variables is assumed to relate to the factors in a different manner than the rest of the variables. In other words, factors cannot be derived in one step. Therefore, we follow Kose et al. [2008] and use the Bayesian method which easily accommodates restrictions on how the factors affect subsets of series. The following paragraph outlines our estimation technique.

We need to use special techniques to estimate the model as the factors are unobservable. Following Chatterjee [2016], we apply the Bayesian posterior simulation method to estimate the dynamic latent factor model. The estimation procedure is based on the following vital observation: if the factors were observable, under a conjugate prior, the models would be a simple set of regressions with Gaussian autoregressive errors; that simple structure can, in turn, be used to determine the conditional normal distribution of the factors given the data and the parameters of the model. This conditional distribution can, then, easily be used to generate random samples, which can serve as proxy series for the unobserved factors. As the full set of conditional distribution is known − parameters given data and factors and factors given data and parameters − it is possible to generate samples from the joint posterior distribution for the unknown parameters and the unobserved factors using sequential sampling of the full set of conditional distributions in a Gibbs sampling. The process is iterated a large number of times. Under the regularity conditions satisfied here, the Markov chain so produced converges, and yields a sam-

[29] We have tried alternative normalizations of the food factor setting factor loadings for rice, or wheat or beef, to unity. The results are remarkably similar.

ple from the joint posterior distribution of the parameters and the unobserved factors, conditioned on the data.

Once we estimate the common demand factor and the food specific factor from the dynamic factor model, we residualize the log changes in global food price index to construct our food commodity shock that we used as an instrument.

Appendix D Global Price Changes and Instrumental Variables

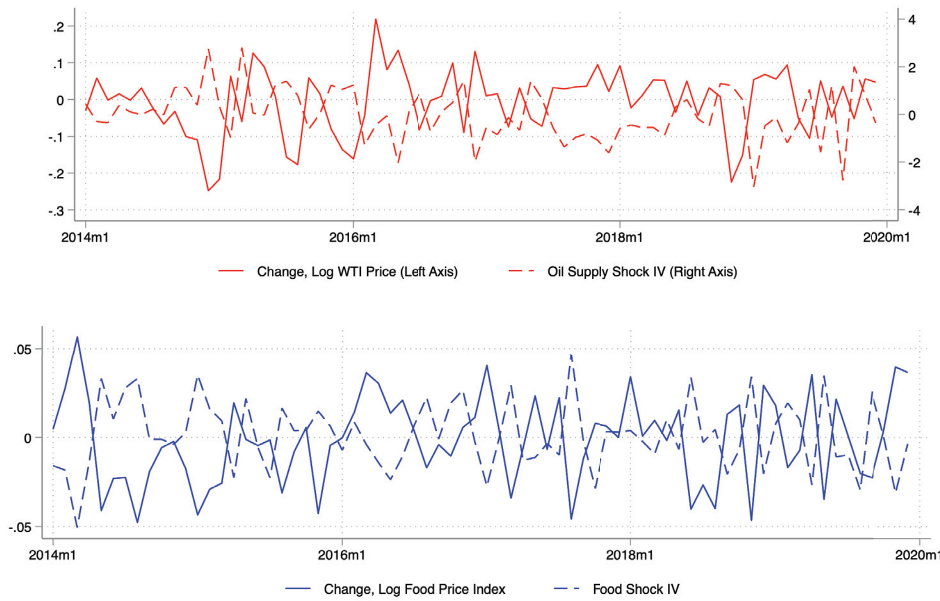


Figure A1: Oil and Food Price Shocks and the Respective Instrumental Variables

Notes: This figure plots the time series of oil (top panel) and food (bottom panel) price shocks and the respective IVs. The oil price shock series is the month-on-month change in log WTI crude oil prices; food price shock is the month-on-month change in log globalfood price index published by the IMF. The supply instrument for oil price shocks is taken from Baumeister and Hamilton [2019]. The food supply instrument is constructed using a dynamic factor model as described in Appendix C.

Appendix E Details on Empirical Specifications

E.1 List of Controls and IVs in Household Regressions

Table A5: Instrumental and Control Variables in Household Panel Local Projection

Panel A. Instrumental Variables

○ Oil supply shock estimated in Baumeister and Hamilton [2019]

○ Food supply shock estimated using a dynamic factor model of non-energy commodity prices

Panel B. Control Variables

○ Lags of outcome variables
– 3 lags

○ Lags of global oil and food price changes
– 3 lags

○ State-by-time-fixed effects
– State-by-calendar month-fixed effects
– State-by-calendar year-fixed effects

○ Socio economic status-fixed effects
– Caste
– Religion
– Education groups
– Big city

○ Demographic controls
– Age fixed effects for 5-year age bins over working life

○ Aggregate world condition controls (interacted with household income group dummies)
– World Industrial Production
– US Federal Funds Rate
– Change in global VIX

○ Demonetization policy dummy

Notes: This table shows our instrumental variables and a set of control variables in our baseline panel household local projection regressions.

E.2 First-Stage F-stats for Household IV Specifications

Table A6: F-statistics for Panel Local Projection IV Regressions of Household Consumption

	(1)	(2)
	Consumption	Non-durable Consumption
	Panel A : Global Food Price Shock & Food Supply IV	
First stage F-stats	5644.9	5684.6
	Panel B : Global Oil Price Shock & Oil Supply IV	
First stage F-stats	810.3	799.3

Notes: This table shows F-statistics from first-stage regressions for our panel IV local projection estimation of effects on household consumption (Column(1)) and non-durable consumption (Column (2)).

E.3 Responses Relative to Group 2

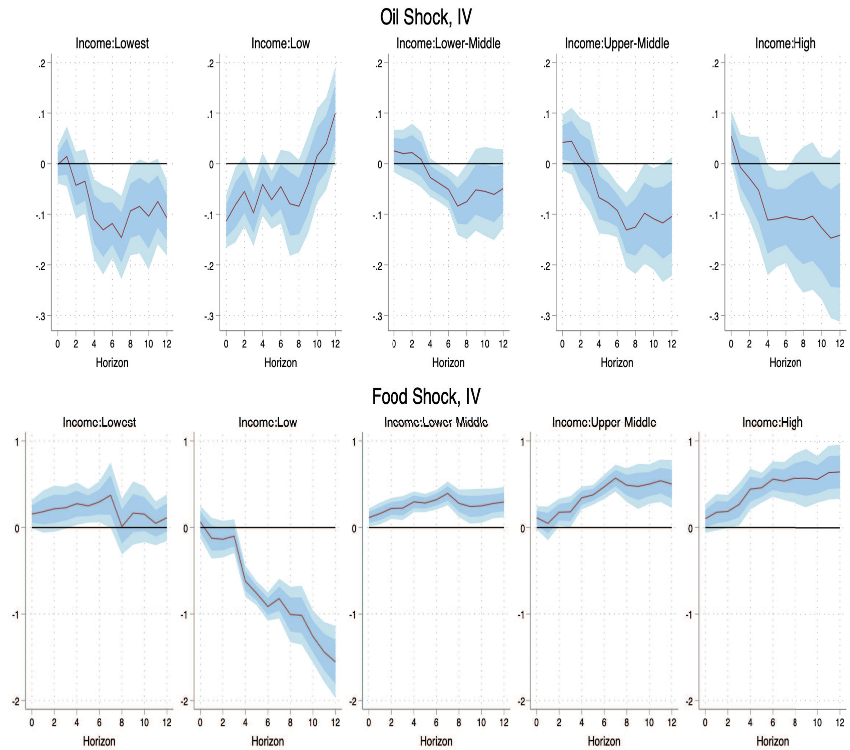


Figure A2: Response of Non-Durable Consumption to External Oil and Food Price Shocks by Income Quintiles

Notes: Cumulative IRFs on the basis of equation (3.1) where the external shock is log changes in global WTI oil and food prices. The IRFs are presented relative to the second (Income: Low) group. The external shocks are instrumented by global supply shocks. The dependent variable is log changes in households' non-durable consumption. The light blue region refers to the 90% confidence interval and the dark blue region is the 68 % confidence interval.

E.4 List of Controls and IVs in Regional Regressions

Table A7: Instrumental and Control Variables in Regional Panel Local Projection

Panel A. Instrumental Variables
○ Oil supply shock estimated in Baumeister and Hamilton [2019]
○ Food supply shock estimated using a dynamic factor model of non-energy commodity prices
Panel B. Control Variables
○ Lags of outcome variables
– 1 lag
○ Lags of global oil and food price changes
– 1 lag
○ State-fixed effects
○ Time-fixed effects
– Calendar month
– Calendar year
○ Aggregate world condition controls
– World Industrial Production
– US federal funds rate
– Change in global VIX
○ Demonetization policy dummy

Notes: This table shows our instrumental variables and a set of control variables in our baseline panel regional local projection regressions.

E.5 First-Stage F-stats for Regional IV Specifications

Table A8: F-statistics for Panel Local Projection IV Regressions of State-Geography Level Prices

	(1)	(2)	(3)
	CPI (All)	CPI (Food)	CPI (Fuel)
	Panel A : Global Food Price Shock & Food Supply IV		
First stage F-stats	3,635.9	3,614.5	2,384.9
	Panel B : Global Oil Price Shock & Oil Supply IV		
First stage F-stats	1410.0	1,515.7	666.2

Notes: This table shows F-statistics from first-stage regressions for our panel IV local projection estimation of effects on regional prices. Columns (1) through (3) show the F-statistics for estimation of effect on CPI (headline), CPI (Food), and CPI (Fuel) respectively.

E.6 OLS Results on Regional Price Effects of External Shocks

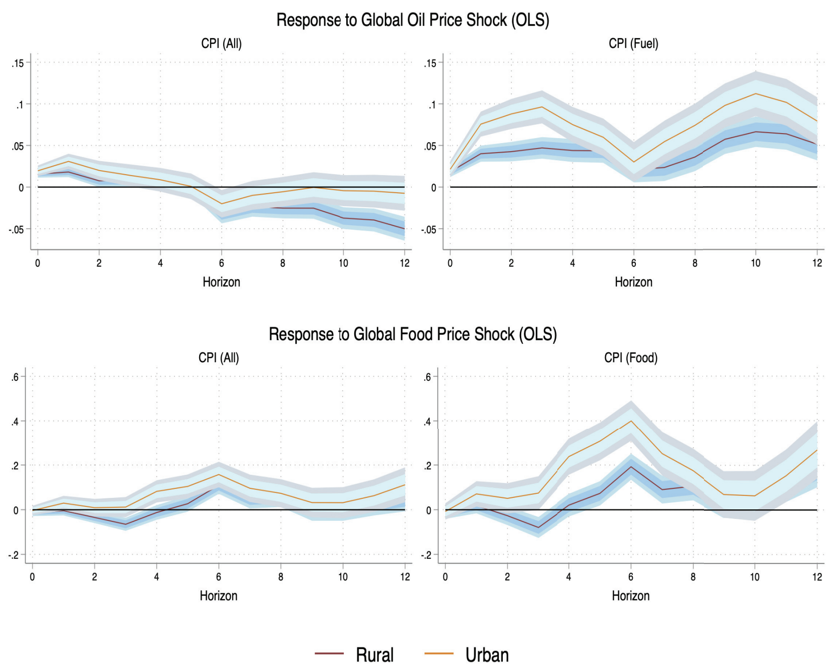


Figure A3: Response of State Level Prices to External Oil and Food Price Shocks (OLS)

Notes: Cumulative IRFs on the basis of equation (4.8) where external shock is log changes in global oil price in the top panel and log changes in global food price in the bottom panel. These are OLS estimates. The dependent variable is log changes in state level prices. The light blue region is the 90% confidence interval and the dark blue region is the 68% confidence interval.

Appendix F Additional Results

F.1 OLS and IV Comparison for Food Price Shocks

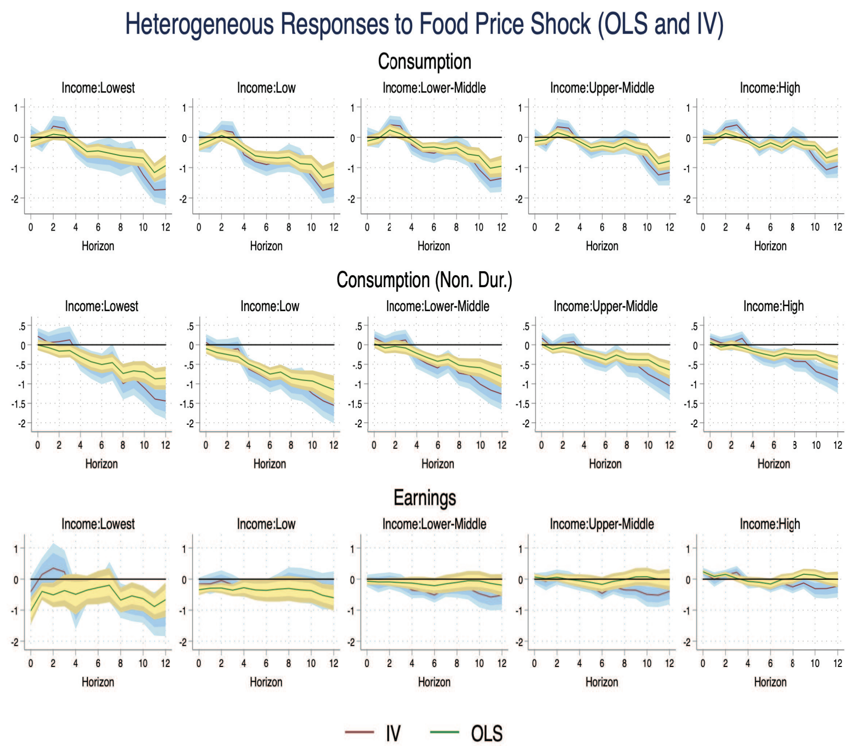


Figure A4: Response of Consumption and Earnings to External Food Price Shocks by Income Quintiles

Notes: Cumulative IRFs on the basis of equation (3.1) where the external shock is log changes in global food price. In the IV version, the log changes in global food price is instrumented by a global supply shock. The dependent variable is log changes in household consumption, non-durable consumption, and labor earnings. The light blue region refers to the 90% confidence interval and the dark blue region is the 68 % confidence interval.

F.2 Expenditure Switching of Non-durables

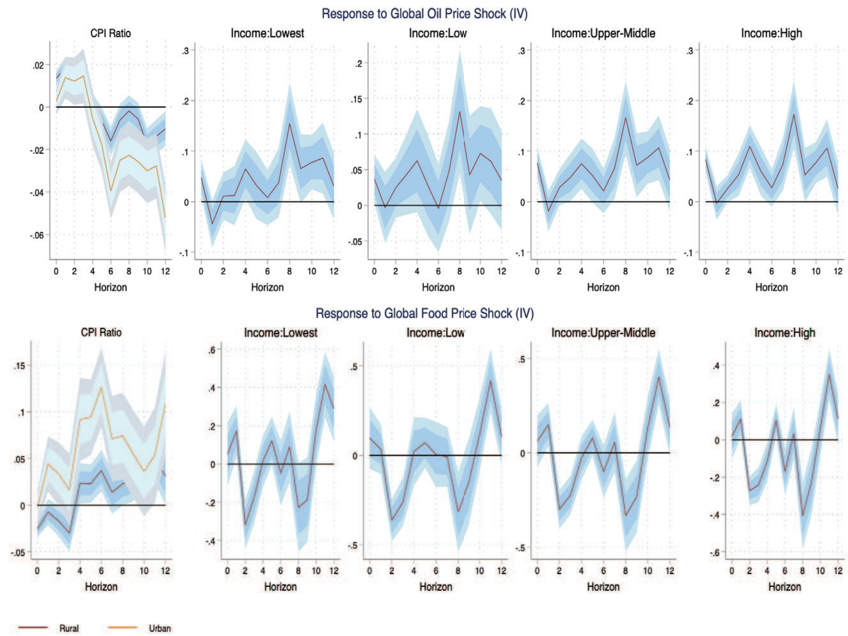


Figure A5: Response of Non-durable to Total Price Ratio and Non-durable Consumption Share to External Oil and Food Price Shocks (IV)

Notes: Cumulative IRFs on the basis of equation (3.1) where external shock is log changes in global food price (top panel), which is instrumented by a global food supply shock and log changes in global oil price (bottom panel), which is instrumented by a global oil supply shock. The dependent variable is non-durable consumption share in total expenditures. The left column plots the response of non-durable price to overall price.

F.3 Discussion Results

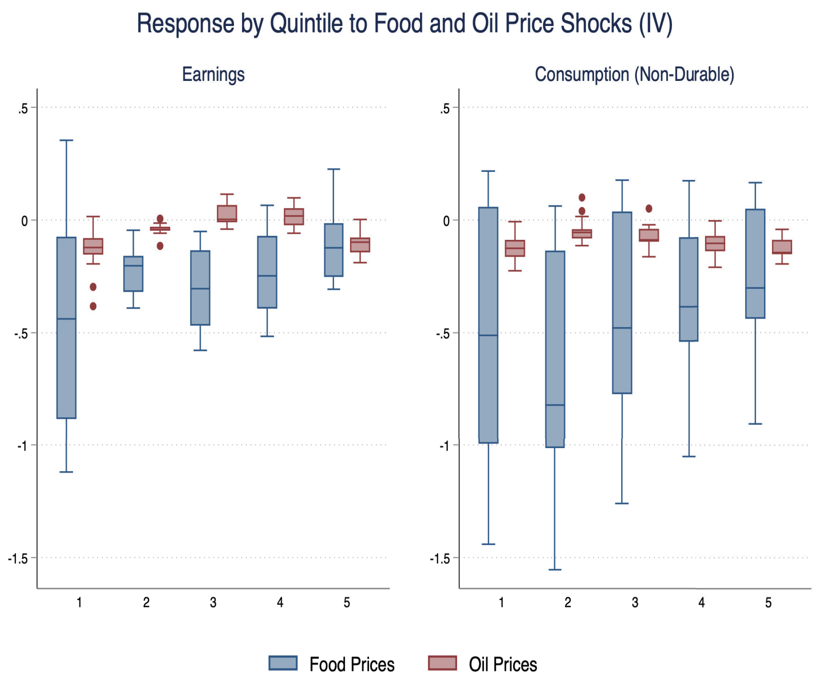


Figure A6: Summary Statistics of Response of Earnings and Non-durable Consumption to Food and Fuel Price Shocks by Income Quintiles (IV)

Notes: This figures is a box and whisker plot that summarizes the responses of labor earnings and non-durable consumption to the two external shocks that is presented in Figures 2, 3, and 5. The line in the center of each box represents the median impulse response estimate across thirteen horizons; the top and the bottom edges of each box represent the 75th and the 25th percentiles respectively; and the lines above and below each box represent respectively the upper and lower adjacent values calculated as in Tukey et al. [1977].

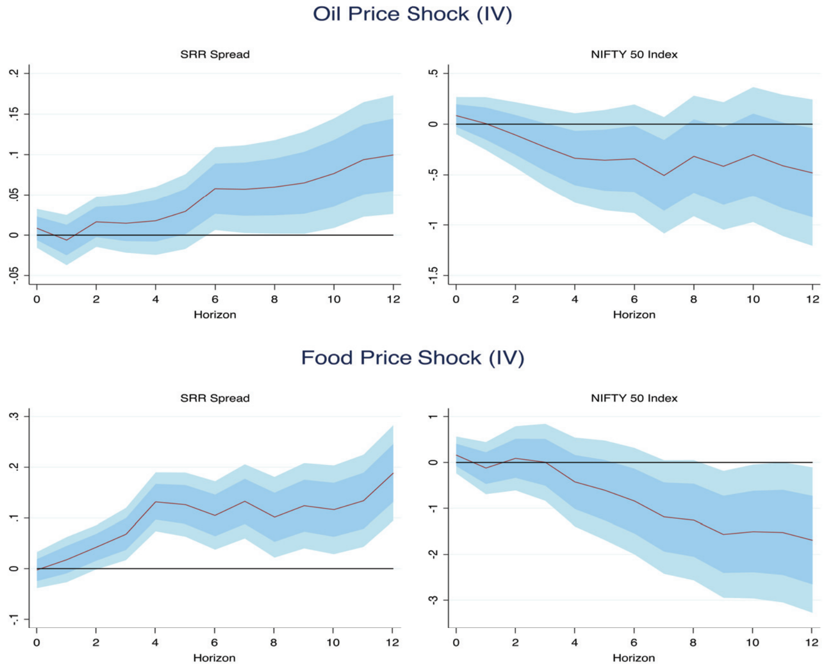


Figure A7: Response of Short Term Interest Rate Spread and Stock Prices to Food and Fuel Price Shocks (IV)

Notes: This figures is the impulse response estimated on the basis of a time series local projection framework.

In Figure A7, in the left columns, we use monthly time series data on short-run (3 month) interest rate spread (relative to the U.S.) as the dependent variable, Y, and Indian industrial production and CPI as well as dummies for global financial crisis, taper tantrum and demonetization as controls, X, in a local projection framework:

$$Y_{t+h} - Y_{t-1} = c + \sum_{j=1}^{J} \alpha_j^h (Y_{t-j}) + \sum_{k=0}^{K} \beta_k^h ext_{t-k} + \gamma_h X_t + \epsilon_{t+h}.$$

Here, ext is our measure of global food and oil price changes instrumented by the corresponding supply shocks, and $J = 3; K = 3$. In the right two columns, our dependent variables are month-to-month changes in the Indian stock price index, Nifty 50, and the estimated impulse responses are cumulative. Our monthly time series data is obtained from Datastream and CEIC, and covers the period January, 2000 to March, 2018.

Table A9: Impact of Global Food and Oil Price on Branch Level Lending Rate (IV)

	(1) Lending Rate (t)	(2) Lending Rate (t)	(3) Lending Rate (t+1)	(4) Lending Rate (t+1)
Global Food Price Change	0.019***		0.026***	
	(0.0001)		(0.0001)	
Global Oil Price Change		0.009***		0.016***
		(0.0000)		(0.0001)
Lending Rate (t-1)	0.540***	0.536***	0.309***	0.307***
	(0.0023)	(0.0023)	(0.0022)	(0.0022)
Observations	1,161,401	1,161,401	1,014,277	1,014,277
R-squared	0.34	0.33	0.17	0.10
Bank FE	Y	Y	Y	Y
District FE	Y	Y	Y	Y

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

For the estimates reported in Table A9, we use the administrative lending rate data at bank-branch level obtained from the Basic Statistical Returns reported to the Reserve Bank of India. Because the data is annual (1998-2016), we need to aggregate our monthly global price changes to the annual level. Because of the short time series, we use a dynamic panel regression to estimate the impact of global price rises on lending rates contemporaneously and one period ahead (instead of a local projection to estimate the entire horizon of dynamic effects as impulse response functions). Our regression specification is:

$$lr_{b,t+1} = c + \alpha(lr_{b,t-1}) + \beta ext_t + \gamma X_t + \delta_{bank} + \delta^{*}_{district} + \epsilon_{t+h}$$

As before, ext is our measure of global food and oil price changes instrumented by the corresponding supply shocks and the standard errors are clustered at the district level.

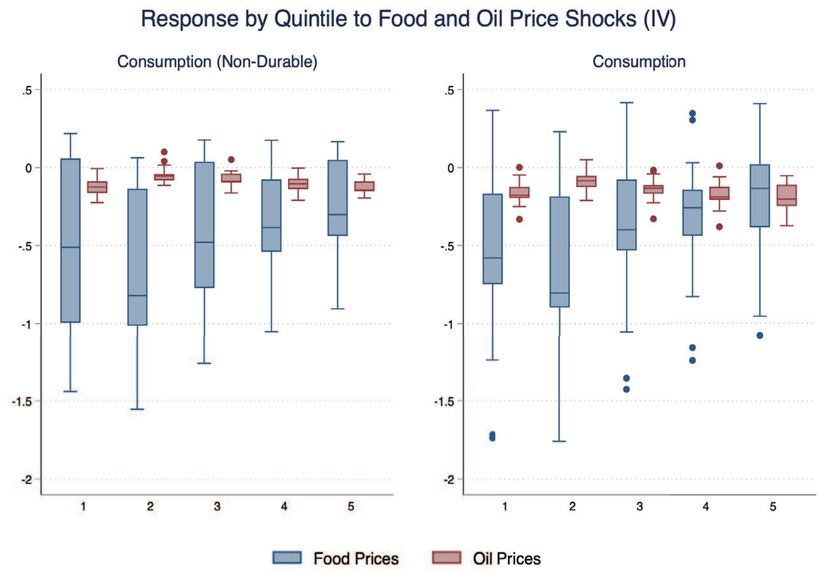


Figure A8: Summary Statistics of Response of Total and Non-durable Consumption to Food and Fuel Price Shocks by Income Quintiles (IV)

Notes: This figures is a box and whisker plot that summarizes the responses of total and non-durable consumption to the two external shocks that is presented in Figures 2 and 3. The line in the center of each box represents the median impulse response estimate across thirteen horizons; the top and the bottom edges of each box represent the 75th and the 25th percentiles respectively; and the lines above and below each box represent respectively the upper and lower adjacent values calculated as in Tukey et al. [1977].

Table A10: Share of Formal and Informal Occupations by Income Groups (in %)

Income Groups	Share of Formal Occupation	Share of Informal Occupation
Lowest	34	66
Low	24	76
Low middle	32	68
Upper middle	46	54
High	75	25

Notes: This table presents the average share of formal and informal occupations among people in labor force for corresponding income groups. Informal occupations include agricultural laborers, home-based worker, small farmer, small trader/ hawker/ businessman without fixed premises, self employed entrepreneurs, legislator/ social workers/ activists and wage laborer. Formal occupations include businessman, industrial workers, managers, non-industrial technical employee, organised farmer, qualified self employed professionals, support staff, white collar clerical employees and White-Collar Professional Employees and Other Employees. The share is calculated on the basis of people who report occupations. While for the top four income groups roughly 8 % do not report an occupation, in the lowest income group, nearly 40 % do not report an occupation.

F.4 Sensitivity Analysis: Groups based on per capita household income

Table A11: Summary Statistics by Income Decile (Based on Per Capita Household Income)

	No. of Hhs	Income		Consumption		Non-durable Share	Food Share
		Household	Per Capita	Household	Per Capita	Household	Household
1st Decile	43,223.65	728.16	111.33	1,413.53	274.08	0.78	0.62
2nd Decile	6,959.47	5,596.87	916.17	4,888.35	855.09	0.78	0.64
3rd Decile	8,060.78	6,453.36	1,148.40	5,138.74	966.81	0.78	0.63
4th Decile	9,017.66	7,319.30	1,373.23	5,400.72	1,068.39	0.77	0.62
5th Decile	10,436.45	8,173.62	1,621.02	5,646.06	1,183.47	0.77	0.61
6th Decile	12,116.28	9,287.45	1,923.05	5,950.28	1,304.91	0.76	0.60
7th Decile	13,258.50	10,657.55	2,315.70	6,255.31	1,445.09	0.75	0.59
8th Decile	14,910.67	12,704.20	2,869.67	6,713.44	1,619.70	0.75	0.58
9th Decile	18,170.81	16,230.00	3,852.63	7,367.53	1,885.03	0.74	0.56
10th Decile	29,823.33	28,500.95	8,177.00	8,928.43	2,665.15	0.72	0.52

Notes: This table presents some summary statistics by income deciles, where the deciles were calculated based on per-capita household income. Income and consumption are in real terms where they are deflated by the CPI (all, 2012=100) Non-durable and food share refer to consumption shares of non-durable and food consumption.

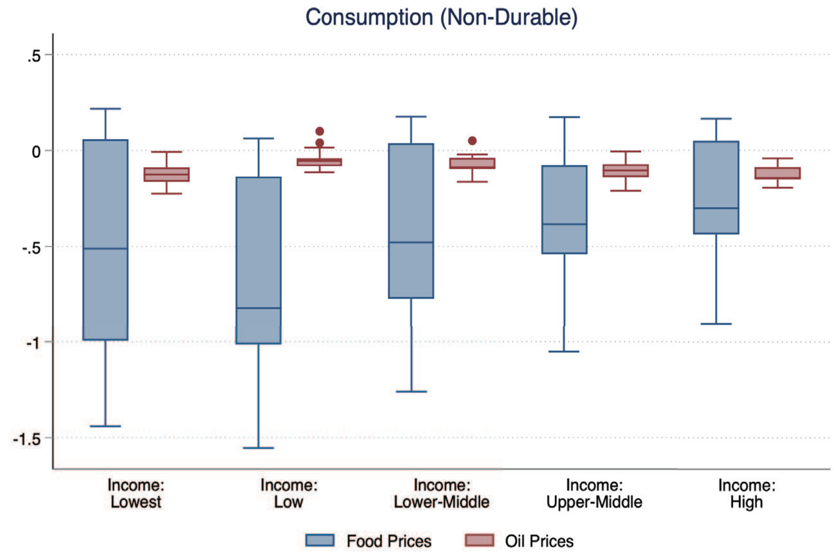


Figure A9: Summary Statistics of Response of Non-durable Consumption to Food and Fuel Price Shocks by Per Capita Income Quintiles (IV)

Notes: This figures is a box and whisker plot that summarizes the responses of non-durable consumption to the two external shocks when households are grouped on the basis of real per capita income in 2014. The line in the center of each box represents the median impulse response estimate across thirteen horizons; the top and the bottom edges of each box represent the 75th and the 25th percentiles respectively; and the lines above and below each box represent respectively the upper and lower adjacent values calculated as in Tukey et al. [1977].

F.5 Sensitivity Analysis: Restricting the income transition matrix

Table A12: Transition Matrix of Real Income

	Q_1	Q_2	Q_3	Q_4	Q_5	Total
Q_1	80.98	4.15	4.40	6.16	4.31	100.00
Q_2	7.26	76.07	13.70	2.21	0.75	100.00
Q_3	4.17	6.35	80.13	8.60	0.75	100.00
Q_4	4.89	0.89	6.87	83.71	3.65	100.00
Q_5	7.70	0.76	1.59	8.15	81.80	100.00

Notes: This table presents the average transition probabilities (in % terms) between different income groups in our sample.

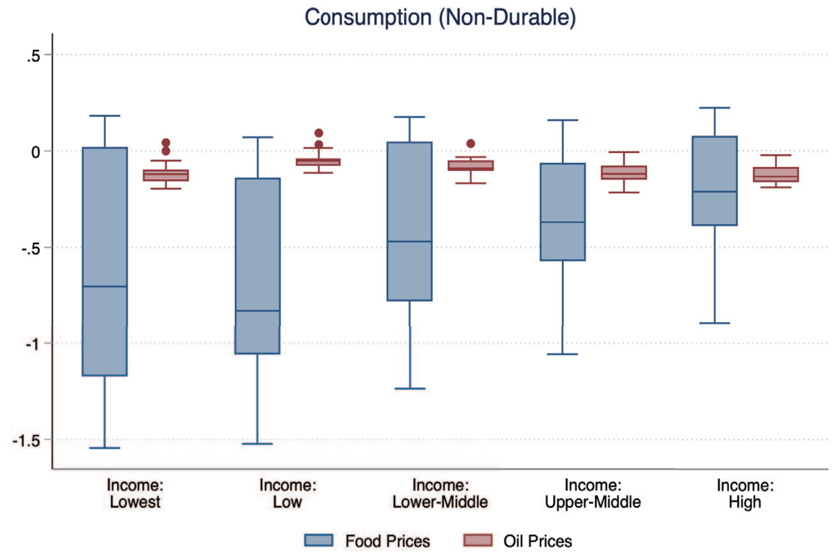


Figure A10: Summary Statistics of Response of Non-durable Consumption to Food and Fuel Price Shocks by Income Quintiles with restricted transition matrix(IV)

Notes: This figure is a box and whisker plot that summarizes the responses of non-durable consumption to the two external shocks while restricting the income transition matrix. The line in the center of each box represents the median impulse response estimate across thirteen horizons; the top and the bottom edges of each box represent the 75th and the 25th percentiles respectively; and the lines above and below each box represent respectively the upper and lower adjacent values calculated as in Tukey et al. [1977].

F.6 Rural Urban Comparison

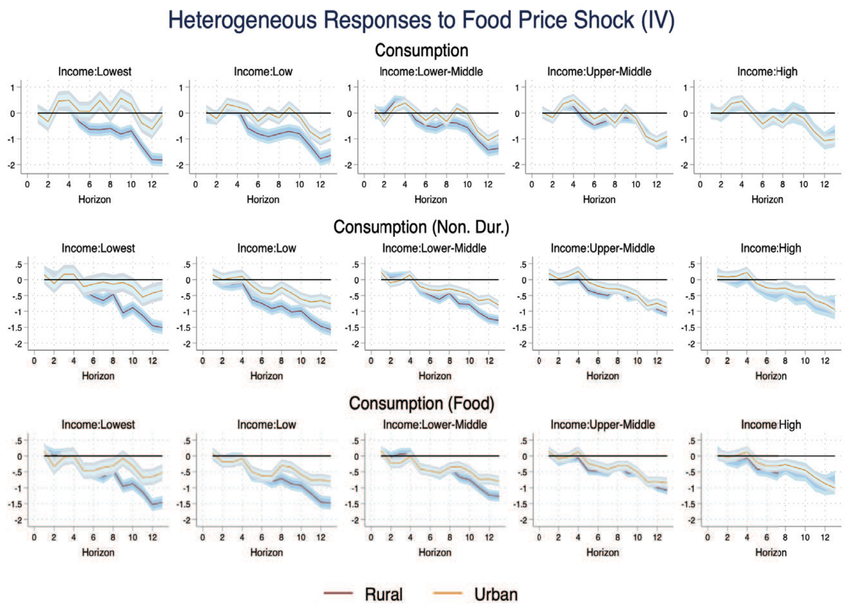


Figure A11: Response of Consumption to External Food Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (5.1) where external shock is log changes in global food price, which is instrumented by a global food supply shock and the dependent variable is log changes in household consumption. The light blue region is the 90% confidence interval and the dark blue region is the 68% confidence interval.

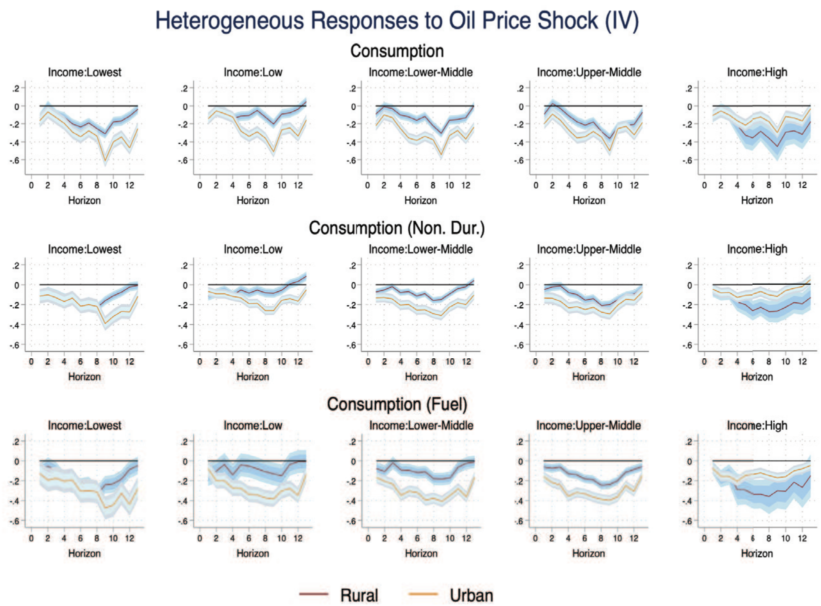


Figure A12: Response of Consumption to External Oil Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (5.1) where external shock is log changes in global oil price, which is instrumented by a global oil supply shock and the dependent variable is log changes in household consumption. The light blue region is the 90% confidence interval and the dark blue region is the 68% confidence interval.

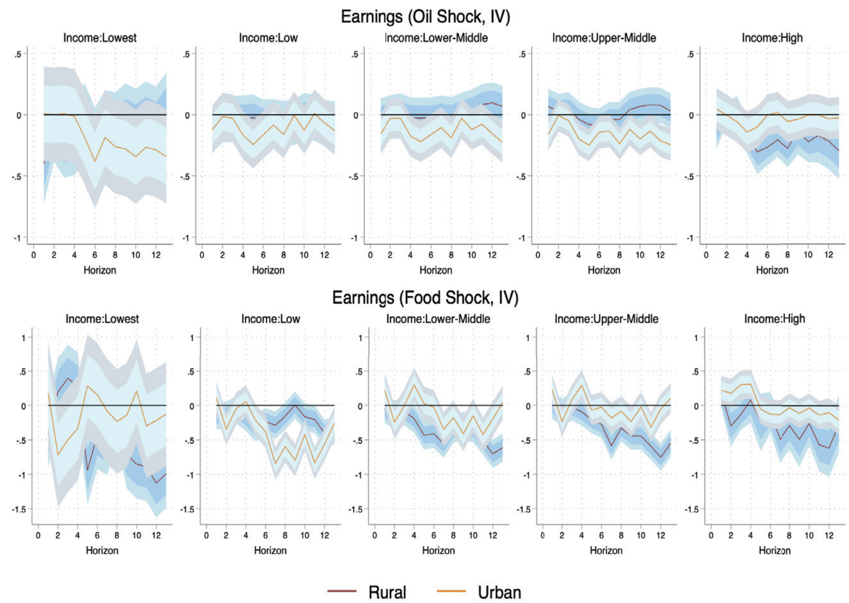


Figure A13: Response of Earnings to External Food and Oil Price Shocks by Income Quintiles (IV)

Notes: Cumulative IRFs on the basis of equation (5.1) where external shock is log changes in global food price, which is instrumented by a global food supply shock, and log changes in global oil price, which is instruments by a global oil supply shock. The dependent variable is log changes in household labor earnings. The light blue region is the 90% confidence interval and the dark blue region is the 68 % confidence interval.

F.7 Effects on Regional Inequality

Here we present IV results regarding effects of global price shocks on regional inequality where we instrument the changes in global food and oil prices. These IV results will isolate variation coming from supply shocks to global food and oil prices as we discussed previously. The specification is given in equation (5.2). We report cumulative impulse responses below. Table A13 lists our control and instrumental variables.

Figures A14 and A15 present answers to the key question of this section based on equation (5.2): How does regional consumption inequality evolve dynamically in response to external food and oil price changes? Broadly speaking, in Figure A14, we observe that an increase in global food prices increases consumption inequality within a state over time, with effects on both total and non-durable consumption inequality statistically significant and persistent.

In Figure A15, we see that an increase in global oil prices does not have as clear of an effect on consumption inequality. In fact, consumption inequality for most of these measures

Table A13: Instrumental and Control Variables in Regional Panel Local Projection

Panel A. Instrumental Variables

◦ Oil supply shock estimated in Baumeister and Hamilton [2019]

◦ Food supply shock estimated using a dynamic factor model of food commodity prices

Panel B. Control Variables

◦ Lags of outcome variables

– 1 lag

◦ Lags of global oil and food price changes

– 1 lag

◦ State-fixed effects

◦ Time-fixed effects

– Calendar month

– Calendar year

◦ Aggregate world condition controls

– World Industrial Production

– US federal funds rate

– Change in global VIX

◦ Demonetization policy dummy

Notes: This table shows our instrumental variables and a set of control variables in our baseline panel regional local projection regressions.

seems to decrease over time. The effects of oil shocks on regional inequality hence appear to be more nuanced, consistent with what we uncover from detailed household level data.

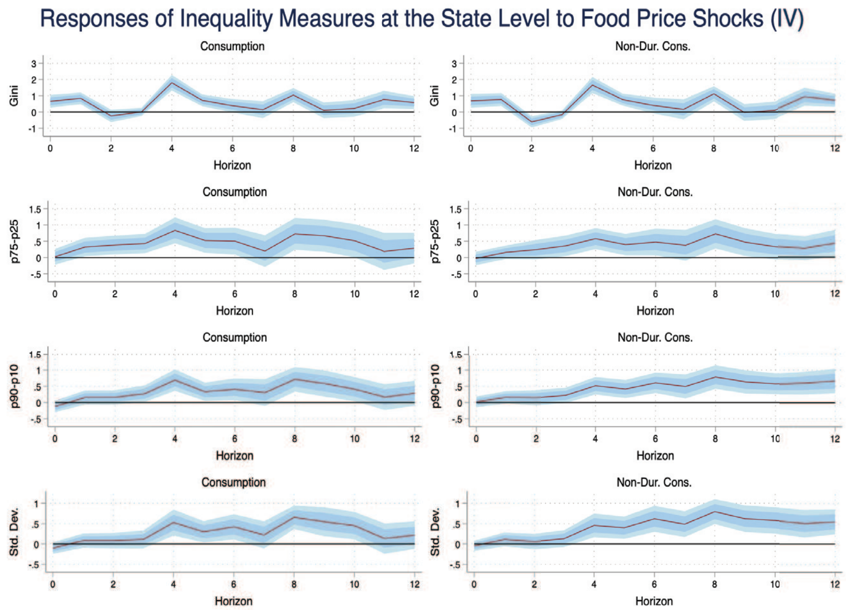


Figure A14: Response of Regional Consumption Inequality to External Food Price Shocks (IV)

Notes: Cumulative IRFs on the basis of equation (5.2) where external shock is log changes in global food price, which is instrumented by a global food supply shock and the dependent variable is log changes in inequality. The light blue region refers to the 90% confidence interval and the dark blue region is the 68 % confidence interval.

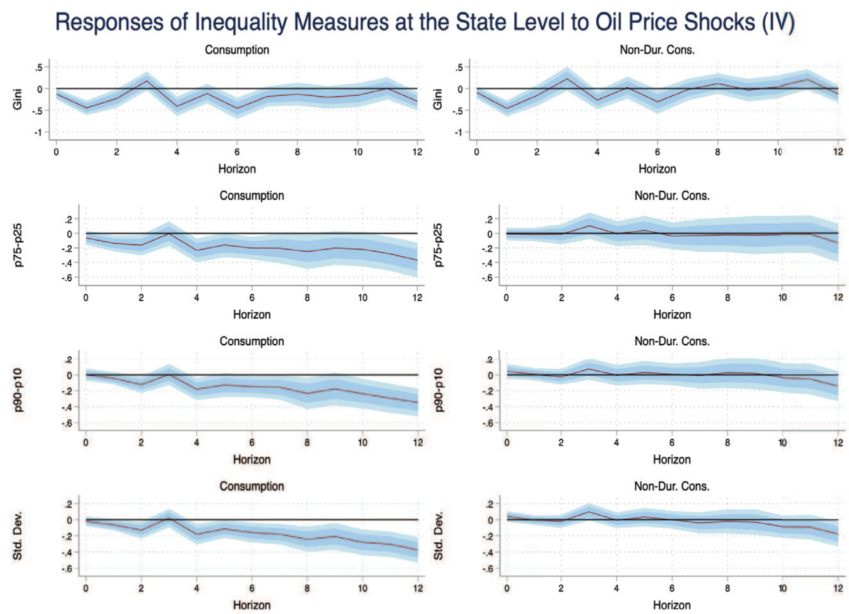


Figure A15: Response of Regional Consumption Inequality to External Oil Price Shocks (IV)

Notes: Cumulative IRFs on the basis of equation (5.2) where external shock is log changes in global oil price, which is instrumented by a global oil supply shock and the dependent variable is log changes in inequality. The light blue region refers to the 90% confidence interval and the dark blue region is the 68 % confidence interval.

PROTECTIONISM IN A GREEN SUIT? MARKET POWER IN CARBON-BASED TRADE POLICY

Sankalp Mathur

December 29, 2023

Carbon tariffs have received widespread support as a second-best policy tool to regulate foreign emissions indirectly. In this paper, I document novel evidence suggesting that carbon-intensive sectors have higher market power and thus charge higher markups. Thus, carbon tariffs lead to sizable profit-shifting across countries. I build a multi-industry structural model of international trade with input-output linkages to analyze the welfare implications of a carbon-based trade policy reform. I study the nature of profit shifting in response to the carbon-embodied tariffs and quantify the aggregate and distributional effects on welfare and emissions. The findings suggest that accounting for market power increases the effectiveness of trade policy in reducing global emissions. However, it generates heterogeneous effects across countries where countries may lose as high as four percentage points or gain more after accounting for profits with the counterfactual trade policy reform. India, for example, experiences a welfare gain of 0.34 percentage points.

JEL Codes: F18, F12, F14, Q50, Q56, H23, F13

Keywords: Carbon tariffs, climate change, Profit shifting, imperfect competition, markups, market power, gains from trade, environmental bias

[1]Current Affiliation: CAFRAL, Reserve Bank of India. Email. sankalpmathur@cafral.org.in. Website www.sankalpmathur.com

Non-Technical Summary

Climate change is one of the most urgent global policy challenges. It is a universal problem that requires worldwide consensus among various nation-states. However, there is a lack of globally coordinated efforts to combat the climate crisis. The Paris Agreement does not prescribe legally binding emission caps for individual countries. Many countries are committed to reducing their greenhouse gas (GHG) emissions through nationally determined contributions (NDCs). The contributions depend on the national government's commitment and public support toward climate change and the trade-off it faces with its economic growth. There is no legal penalty if a government fails to reach its targets. As a result, different countries have pledged to different emission reduction targets. Moreover, countries with stringent environmental policies face the risk of production leakage of emission-intensive and trade-exposed industries. Given the lack of stringent and uniform emission pricing, international trade policies, through import carbon tariffs, have garnered policy attention to increase the effectiveness of unilateral emissions pricing and indirectly regulate foreign emissions. It is crucial to understand the effectiveness of trade policy in reducing global emissions and its economic impact on various countries.

Carbon-based tariff adjustments make it expensive to trade carbon-intensive goods from other countries. This results in production reallocation to domestic countries or countries with comparative advantage. Interestingly, this paper documents a novel finding suggesting that carbon-intensive industries have higher market power than cleaner industries. Thus, they charge higher markups and earn higher profits. As a result, carbon-based trade policy leads to sizable profit-shifting across countries. This paper aims to quantify the role of market power in understanding the impact of carbon-based trade policy adjustments on global emissions and economic welfare.

The emission intensity (tons of CO_2 per \$ of output) of net importers of carbon-intensive goods drives the impact on global emissions. As tariffs are raised on carbon-intensive products, production reallocates to countries otherwise importing these goods. If these importers have lower emission intensity, it will lead to a reduction in global emissions. On average, advanced economies are net importers with lower emission intensity. As production reallocates to countries with lower emission intensity, the policy change reduces global emissions. Notably, the production reallocates high-markup and high-profit industries in home countries. Thus, market power exacerbates the reallocation of carbon-intensive industries in countries with lower emission intensity. As a result, introducing markups increases the effectiveness of reducing global emissions through trade policy.

Unpacking the impact of market power on economic welfare is more challenging. The effect depends on whether a country is a net importer of carbon-intensive goods and which of the two channels, price or profits, is stronger. Carbon-based tariffs may be costlier for countries that net import upstream carbon-intensive goods, particularly when countries do not have the comparative advantage in producing these goods. Higher importing costs, through input-output linkages, spill over to higher prices, reducing welfare. Contrastingly, the profit channel may be stronger for countries if the production of high-markup, high-profit industries reallocate to home, increasing welfare. On the other hand, net exporters of carbon-intensive goods lose market share in high-markup and high-profit markets, reducing welfare.

To study these channels and the impact of market power, I run a counterfactual experiment where tariffs are effectively raised in carbon-intensive industries and lowered in cleaner industries. I then compare the results with a setup with no market power. First, I find that incorporating market power increases the effectiveness of trade policy in reducing global emissions. Secondly, I find that countries experience a heterogenous impact on economic welfare. Net exporters of carbon-intensive goods lose more when market power is accounted. The effect on net importers depends on which of the two channels is stronger. Overall, the paper studies a novel channel of market power and reassesses the environmental and distributional consequences of a carbon-based trade policy.

1 Introduction

The lack of global consensus on a uniform carbon tax has led to the increased acceptance of carbon tariffs by advanced economies to regulate foreign emissions indirectly. (Böhringer et al. (2012); Branger and Quirion (2014); Carbone and Rivers (2017)). However, carbon tariffs are controversial as they may be a protectionism tool under the guise of climate security.[2]Advanced economies may levy carbon tariffs to shelter domestic industries, undermining the pro-competitive effects of global trade. Carbon tariffs may come at the expense of developing countries which are net exporters of carbon-intensive goods, exacerbating the burden-shifting effect as profits shift from developing to advanced economies. Most research assumes perfect competition to analyze carbon-based tariffs masking the potential reshuffling of profits across countries. In this paper, I break the assumption of perfectly competitive markets. I explore the role of market power, quantifying the international profit-shifting in response to carbon-based tariff adjustments using a structural trade model.

The profit-shifting channel is crucial in studying carbon-based tariffs because I find that incorporating market power increases the effectiveness of trade policy in reducing global emissions. However, it generates heterogeneous effects across countries. I document novel, suggestive evidence exploring the relationship between an industry's market power and carbon intensity. I find that carbon-intensive goods are, on average, high markup goods, showing that market power is more widespread in energy-intensive sectors. As a result, carbon-based tariff adjustments relocate high markup industries to the domestic country or the countries with comparative advantage. The production reallocation results in significant profit-shifting and revision of aggregate welfare across countries. Net exporters of carbon-intensive goods are most likely to lose, while net importers are most likely to gain, relative to a case of perfectly competitive markets. I then build a quantitative structural model of international trade to estimate the effect of this relationship on welfare and global emissions, accounting for the revision of aggregate profits. The model evaluates the distributional outcome across countries to examine the burden-shifting effects and assess if carbon-based tariffs contradict the principles of Common But Differentiated Responsibility[3] in which industrialized countries commit to avoiding negative

[2]see National Post Opinion: Carbon tariffs as protectionism tool

[3]The UNFCCC describes CBDR in Article 3.1: "The Parties should protect the climate system for the benefit of present and future generations of humankind, on the basis of equity and in accordance with their common but differentiated responsibilities and respective capabilities. Accordingly, the developed country Parties should take the lead in combating climate change and the adverse effects thereof."

economic spillover of their environmental policy to developing countries.

I use two data estimation techniques to explore the relationship between an industry's carbon intensity and market power. First, to measure carbon intensity, I use environmentally-extended multi-region input-output tables to measure the direct carbon emissions (through the combustion of fossil fuels) and the indirect carbon emissions (through the use of carbon-intensive intermediates). Second, to measure the market power, I estimate the import demand elasticity of an industry using the limited information maximum likelihood (LIML) technique detailed in Soderbery (2015).

I present a stylized fact that reveals a positive correlation between the two variables. Thus, the evidence suggests that carbon-intensive goods are high markup goods. Carbon-based tariff adjustments aim to increase the cost of importing a carbon-intensive good from a given source country. A higher import cost might be welfare-reducing for net importers of carbon-intensive goods as higher import cost spillovers to a higher cost of living. However, if carbon-intensive goods are high markup goods, carbon tariffs might relocate the production of the high-markup industries to the domestic country, reaping profits and enhancing welfare. In contrast, a tariff increase is potentially costlier for exporters of high-markup goods, as it shifts the profits away from the exporter to the destination country or other competing countries with comparative advantage.

I build a multi-country, multi-industry structural model of international trade with global input-output linkages to account for this profit-shifting channel and estimate the aggregate effects on welfare and global emissions. The model departs from the assumption of perfect competition and tractably accounts for the profit-shifting across countries in response to carbon-based tariff adjustments. Heterogeneity in the demand elasticity drives the responsiveness of the import demand to the higher import cost. It thus affects the relative market shares of the source country and its aggregate markups. Given the dispersion in the demand elasticity across industries, carbon-based tariff adjustments can generate sizable gains or losses for the countries.

I study a counterfactual trade policy reform where each country implements a single tariff per trading partner (equal to the mean baseline bilateral tariff). As documented in Shapiro (2021), the baseline tariffs are lower for carbon-intensive imports. This fact is referred to as the environmental bias of trade policy and is widespread across all countries. Thus, the counterfactual policy reform effectively increases import tariffs for carbon-intensive industries and lowers them for cleaner industries. The counterfactual is similar to the central quantitative assessment in Shapiro (2021), allowing a straightforward comparison with the literature. More generally, low protection in carbon-intensive upstream industries stems from a strong industry lobby for lower import costs for these upstream goods. WTO has sought to decrease the protection of downstream industries relative to upstream industries since such policy reform would let developing countries sell more advanced goods to rich countries. This counterfactual is similar to WTO's policy objective and imposes uniform bilateral tariffs across all goods while pricing the carbon emissions embodied in its imports. Governments have a sub-optimal baseline trade policy where tariffs are lower for carbon-intensive sectors. As a result, the policy reform effectively raises tariffs on carbon-intensive sectors and lowers them for cleaner sectors.

Accounting for profit shifting increases the effectiveness of trade policy to reduce global emissions. The counterfactual policy reform reduce global emissions by 5.7% in the case of imperfect competition and by 3.7% in the case of perfect competition. The extent of domestic reallocation explains the difference between these two cases. When we account for profit-shifting, there is a higher domestic reallocation of carbon-intensive sectors in countries that are net importers, as these sectors are high markup sectors. For example, domestic production reallocation is considerably higher under imperfect competition for many advanced economies, including the United States. Since advanced economies, on average, have lower emission intensity and are net importers of carbon-intensive goods, the strength of domestic production reallocation exacerbates the reduction in total global emissions under imperfect competition. The results highlight the potency of trade policy as a second-best policy tool to reduce global emissions.

Quantifying the markup channel has important implications for a country's aggregate welfare in response to carbon-based trade policy reform. Under perfect competition, all countries gain real income. While increasing tariffs on carbon-intensive sectors, carbon-based tariffs also reduce tariffs on cleaner sectors. The clean industries are downstream industries and thus have a higher share of the final consumption of the consumers (Shapiro (2021), Antràs et al. (2012)). As a result, the policy reform lowers the price and encourages international trade of cleaner industries. It leads to heterogeneous positive distributional effects across countries depending on each country's consumption basket and international input-output linkages. The aggregate welfare effect varies widely from 10% in Norway and 3.65% in Sweden to less than 1% in India, Indonesia, and Australia. These results are similar to Shapiro (2021), which reports a positive effect on real income across different regions under perfect competition.

The quantitative assessment in this paper importantly accounts for profit-shifting across countries as the production of high-markup carbon-intensive industries reallocates in response to carbon-based tariff adjustments. The results highlight that unaccounted profits overestimate the welfare effects of countries like China as high as 4.48 percentage points (pp) and Russia (2.69 pp). Unlike perfect competition, a few countries lose real income: China loses by 1.40%, Indonesia by 0.23%, Mexico by 0.17%, and the Netherlands by 1.18%. Russia only gains by 0.82% relative to 3.51% in case of perfect competition. The heterogeneous effects across countries are explained by their trade portfolio and their role in the global input-output linkages. Countries that are net exporters of carbon-intensive goods lose the most from this adjustment. The trade reform is strikingly costly for net exporters of carbon-intensive goods as their export profits erode. In contrast, the net importers of carbon-intensive goods gain real income as the profits and production of high markup industries likely reallocate to their own country.

Literature

Climate change is one of the most urgent global policy challenges. Due to industrialization, deforestation, and large-scale agriculture, human-generated carbon emissions and other greenhouse gases have increased global temperatures and thus are an undisputed driver of climate change. A changing climate can disastrously impact our ecological, physical, and health systems, including extreme weather events, rising sea levels, hampered agricultural growth, and polluted water systems (IPCC (2014)). It is a universal problem that requires worldwide consensus among various nation-states. However, there need to be more globally coordinated efforts to combat the climate crisis. The Paris Agreement of 2015 does not prescribe legally binding emission caps for individual countries. Many countries are committed to reducing their greenhouse gas (GHG) emissions through nationally determined contributions (NDCs). The contributions depend on the national government's commitment and public support toward climate change and the trade-off it faces with its economic growth. As a result, different countries have pledged to different emission reduction targets.[4] If a government fails to reach its targets, there is no legal penalty. Given the lack of stringent and uniform emission pricing, international trade policies, through import carbon tariffs, can play a crucial role in increasing the effectiveness of unilateral emissions pricing.

The differences in the stringency of environmental regulations across countries contribute to the leakage risk whereby rich countries offshore the production of trade-exposed and carbon-intensive goods to countries where the environmental regulation is relatively relaxed. (Hoel (1991); Felder and Rutherford (1993); Markusen (1975); Fischer and Fox

[4]To view updated NDCs for different countries, visit here

(2012); Fowlie and Reguant (2022)). Leakage reduces the cost-effectiveness of unilateral climate policy and comes at the cost of competition loss for emission-intensive and trade-exposed (EITE) industries in stringent countries. Furthermore, emission regulation in stringent countries lowers the demand for fossil fuels and thus reduces their international price. Lower prices incentivize non-stringent countries to increase the use of fossil fuels. The empirical evidence on the leakage risk is relatively modest. Most papers find leakage rates of less than a third using numerical methods and partial equilibrium quantitative models (Babiker (2005); Droege and Panezi (2022)). Some papers find potential leakage ranging from 20 to 73% in emission-intensive industries. (Demailly and Quirion (2006); Ponssard and Walker (2008); Fowlie et al. (2018)). This leakage risk jeopardizes any gains in environmental quality by trading off emission reduction in stringent countries for emission increases in non-stringent countries, defeating the goal of domestic regulation.

These issues have highlighted the significance of using carbon tariffs or carbon border adjustments to price the behavior of foreign firms. For instance, European Union passed a proposal for Carbon Border Adjustment Mechanism whose transitional phase is set to begin in October 2023.[5] Nordhaus (2015) shows how an international climate treaty that combines target carbon pricing and trade sanctions can induce substantial abatement. The modeling results in Nordhaus (2015) indicate that modest trade penalties on relaxed countries by a ”climate club” of stringent countries can induce a coalition that approaches the optimal level of abatement.

Trade policy is important because carbon emissions create a pollution externality that is inconsequential for location. Unlike flow pollutants, CO_2 is a stock pollutant. With flow pollutants, environmental regulations can result in production leaving the country and hurting people in other countries.[6] However, with carbon emissions, if a country regulates emissions and the production moves to another country, the regulating country will still suffer from carbon emissions since the concern is with the stock of CO_2, regardless of its origin. Thus, the ”optimal” policy would impose a uniform Pigovian tax (or a quantity mechanism like cap-and-trade) in all countries and industries.

Most studies estimate the tax to be around 40-80$ per ton of CO_2, depending on the dis-

[5]Under the adjustment, EU importers will buy carbon certificates corresponding to the carbon price that would have been paid had the goods been produced under the EU's carbon pricing rules. The mechanism creates a level playing field between home and foreign producers to avoid the risk of leakage, effectively increasing tariffs on carbon-intensive imports.

[6]For example, when California banned spray finishes, production moved to Mexico. The negative externality from the pollutant in spray finishes (like the health effects when spraying on the furniture - a flow) also moved to Mexico.

count rate.[7] However, due to the lack of globally coordinated efforts to impose a uniform tax on carbon, trade policy has been proposed as a 'second-best' policy tool to regulate foreign emissions indirectly. Redesigning tariff policy in conjunction with their carbon content raises the global prices of carbon-intensive goods, reducing their market shares; and lowering the costs of cleaner goods, encouraging international exchange. Empirical analysis on carbon tariffs has found the potency of carbon tariffs in reducing emission leakage and lowering output losses for EITE industries in stringent countries (Böhringer et al. (2012); Branger and Quirion (2014); Carbone and Rivers (2017)). However, the literature has not explored the quantitative importance of reshuffling of aggregate profits across countries in response to carbon-based tariff adjustments.

In addition to the leakage risk, the current international trade structure and tariff policy contribute to global emissions. Specific facts established in the literature highlight the consequences of our international trade policy on global carbon emissions. Internationally traded goods account for about 22 to 35 percent of global pollution emissions (Copeland et al. (2021)). Pollution-intensive industries are more exposed to trade and relish lower tariff and non-tariff barriers (Shapiro (2021)). This pattern exists globally and within almost all countries. The difference in trade barriers between carbon-intensive (or dirty) and clean industries creates an implicit subsidy for carbon emissions in internationally traded goods. Shapiro (2021) estimates that this global implicit subsidy is between 85$ to 120$ per ton of CO_2, which is higher than the estimated optimal tax on CO_2 at $40 per ton. He explains the existence of environmental bias in our trade policy. Trade barriers are generally lower on upstream goods because industries lobby for lower tariffs on intermediate goods.[8] He finds that upstream goods are also carbon-intensive. The downstream goods are cleaner but face higher tariffs as final consumers are poorly organized. Given this unintentional design, international trade encourages the exchange of dirtier carbon-intensive goods, contributing to global emissions and climate change. This paper contributes to this strand of literature by exploring a novel relationship between carbon intensity and market power. In particular, carbon-intensive sectors are not only subjected to lower tariffs but also, as this research suggests, enjoy a higher market power and thus charge higher markups.

Shapiro (2021) builds a quantitative world economy model to understand the effectiveness of redesigning tariffs. The paper studies a single tariff per trading partner (equal to mean baseline tariffs), raising tariffs in carbon-intensive sectors and lowering them in

[7]Social cost of carbon under different discount rates

[8]Examples of industries with low upstreamness is automobiles, household furniture, etc. while industries with high upstreamness include petrochemicals, copper smelting/refining, etc.

cleaner sectors. Shapiro (2021) finds that such a policy change will lead to global emission reduction by 3.6% and a modest increase in real income in all countries. This striking result fills a critical knowledge gap in the literature. The quantitative assessment shows that we can correct the unintentional environmental bias of our trade policy without hampering the global real income by redesigning our tariff structure.

This paper builds on Shapiro (2021) but targets quantifying the role of profit-shifting under the same trade policy reform. It breaks the assumption that the markets are perfectly competitive and focuses on an increasingly relevant trend in economic literature: markups and market power (De Loecker and Eeckhout (2018); De Loecker et al. (2021)). To capture market power, the model allows for demand elasticity to vary not only across sectors but also across destination countries.[9] Heterogeneity in the destination-specific demand elasticity generates differences in import and export markups, which augments the fact that some countries are net exporters of high-markup goods. At the same time, some are net importers of high-markup goods (Firooz and Heins (2020)). Consequently, a tariff increase is potentially costlier for exporters of high-markup goods, as it shifts the profits away from the exporter to the destination country or other competing countries with comparative advantage. Specifically, carbon-based tariff adjustments aim to increase the cost of importing a carbon-intensive good from a given source country. Heterogeneity in the destination elasticity drives the responsiveness of the import demand to the higher import cost. It thus affects the relative market shares of the source country and its aggregate markups. Given the motivating evidence that carbon-intensive sectors are high markup sectors, tariff adjustments can generate sizable gains or losses for the countries. The paper aims to quantify profit-shifting and its role in evaluating aggregate welfare and emission changes in response to carbon-based tariff adjustments.

In addition to quantifying profits, the paper runs the analysis for a more granular sectoral detail and with the recent data for 2016. This is relevant because Shapiro runs the quantitative assessment for the year 2007. However, as documented in Böhringer et al. (2021), the global economy has undergone a substantial structural change regarding production, consumption, and trade since the financial crisis. Trade in carbon embodied in goods increased until 2007-2008 but modestly decreased afterward. They find that the effectiveness of carbon tariffs is higher when there is a higher trade in carbon. This paper analyzes 2016 to understand the quantitative importance of the profit channel for more recent data. Another reason more current data is useful is that carbon intensity changes due to decarbonization over time (by a change in the mix of fuels, output, or efficiency).

[9] Evidence of heterogeneous demand elasticity has been expressed in various papers, e.g., Handbury (2021), Faber and Fally (2017), Adao et al. (2017), or Heins (2019)

Furthermore, the paper analyzes a sample of 24 countries and a Rest of the World (ROW) aggregate rather than combining them with specific regions (OECD, Non-OECD, Eurozone, Asia). Understanding if the carbon-based tariff adjustments spur global inequalities by shifting export-based profits from developing countries to rich countries is imperative. This paper identifies specific countries in these groups that may gain or lose due to these adjustments.

The remainder of the paper is structured as follows. In Section 2, I detail the data and estimation technique to measure an industry's carbon intensity and market power (through the estimated import demand elasticity). In Section 3, I provide suggestive evidence about the positive relationship between carbon intensity and market power. In Section 4, I describe a quantitative structural model that enables us to account for profit-shifting tractably and estimates the aggregate effects on welfare and emissions. In Section 5, I present the results of a counterfactual exercise and evaluate the aggregate and distributional effect of a trade policy reform. Section 6 concludes.

2 Data and Estimation

The paper uses three main variables to present a stylized fact and to run the quantitative analysis: carbon emissions, market power, and trade policy. To measure carbon emissions, I use the Exiobase Version 3.8 Database (Stadler et al. (2021)). I measure the market power of an industry by estimating the import demand elasticity (σ) using the limited information maximum likelihood hybrid estimator (Soderbery (2015)). To measure the baseline trade policy, I use the WITS database[10] that reports the sector-specific bilateral ad-valorem tariffs. I use the TRAINS database[11] to measure the ad valorem equivalent of non-tariff measures.

2.1 Carbon Emissions

I use the environmentally extended multi-regional input-output tables (EE MRIO) from Exiobase Version 3.8. EE-MRIO by Exiobase, supported by the European Union, has emerged as a pivotal database to analyze the environmental impact of production, trade, and global supply chains. Exiobase reports the world input-output details for 200 prod-

[10]WITS database can be accessed here
[11]Non-Tariff Measures can be accessed here

ucts for 44 countries and five country aggregates, which collectively account for 90% of the world GDP. The database is consistent with the recommended accounting systems of the United National System of Environmental-Economic Accounting (UN-SEEA), which makes it the best available resource for international comparisons across countries and regions (Stadler et al. (2018)). Exiobase is built from several primary data sources (Wood et al. (2014)). It measures trade using BACI, based on the UN's Comtrade database and the UN's services trade databases. Much of the global supply chain literature has used either the World Input-Output Database (WIOD) or the Global Trade Analysis Project (GTAP - MRIO). However, Exiobase EE MRIO provides a richer and more consistent level of sector detail, enabling us to assess the carbon sources of the global value chains adequately.

There are two types of carbon emissions: direct and indirect emissions. Carbon emitted by the combustion of fossil fuels (coal, petroleum, and natural gas) accounts for the direct emissions of industry. Indirect emissions are the carbon emitted in the intermediate goods produced as inputs to make that final good. An input-output table reports the dollars of output from a row industry used to produce one dollar of the output of the column industry. If there are S industries, an input-output table is an $[S \times S]$ matrix detailing the dollar value of each input used to produce each good in the economy. This permits the calculation of the dollar amount of fossil fuels used to produce each good. The database combines the emission factor for each fossil fuel (obtained from the TEAM model, Pulles et al. (2007)) with the national price per physical unit to arrive at the tons of CO_2 emitted to infer the carbon emissions for each dollar of fossil fuels used directly in the production. An industry's "direct" emission rate is inferred by multiplying the total expenditures on each fossil fuel with their respective CO_2 emission per dollar of the fossil fuels used. Exiobase reports the direct impact of each industry in terms of CO_2 emissions in 1000 tons[12] per one million dollars of output from fossil fuel combustion. Table 1 summarizes the cleanest and dirtiest five industries by the median value of direct carbon emissions. In most countries, the highest emissions stem from electricity production, which is not traded internationally as much but is used as a primary input in the production process. Therefore, it is crucial to account for the use of electricity (and hence fossil fuels) as intermediates in producing final tradable goods.

The "default" method of CO_2 described above measures the direct emissions from producing a good but ignores the life-cycle measurement of CO_2 emissions in its value chain

[12]Here, tons refers to metric tons. Exiobase reports the carbon emissions from the combustion of fossil fuels in Gigagrams (Gg) per 1 million euros of output. 1Gg = 1000 metric tons. I use the mean annual exchange rate from the IMF's International Financial Statistics to convert Euros to dollars.

Table 1: Cleanest and Dirtiest Industries by Direct Carbon Emissions

Industry	Direct CO_2 emissions
Panel A: Cleanest Industries	
Post and Telecommunication Services	0.08
Education/Financial Services	0.11
Public Administration and Defense Services	0.13
Electricity by Nuclear/Hydro	0.14
Real Estate Services	0.15
Panel B: Dirtiest Industries	
Electricity by Coal	86.23
Wool, Silk-Worm Cocoons	56.12
Natural Gas Extraction	19.63
Electricity by Petroleum/Oil Derivatives	17.75
Coke Oven Coke	15.92

CO_2 combustion measures the 1000 metric tons of CO_2 per million USD from combustion of fossil fuels

(such as the use of electricity as an intermediate good). To measure the total carbon intensity of an industry, we need to account for the CO_2 emitted in the production of the inputs used, the inputs used in that input, and so on. For example, direct CO_2 emissions from Automobile manufacturing account for the fossil fuels used to manufacture the automobile. However, it neglects the emissions from inputs like steel, which in turn has burned fossil fuels for its production. To permit the calculation of indirect emissions, I calculate the Leontief inverse or a matrix of total requirements. Leontief inverse details the dollars of each input (including those required to produce intermediate inputs, inputs to inputs, and so on) needed to make an additional dollar of the final demand of that industry.

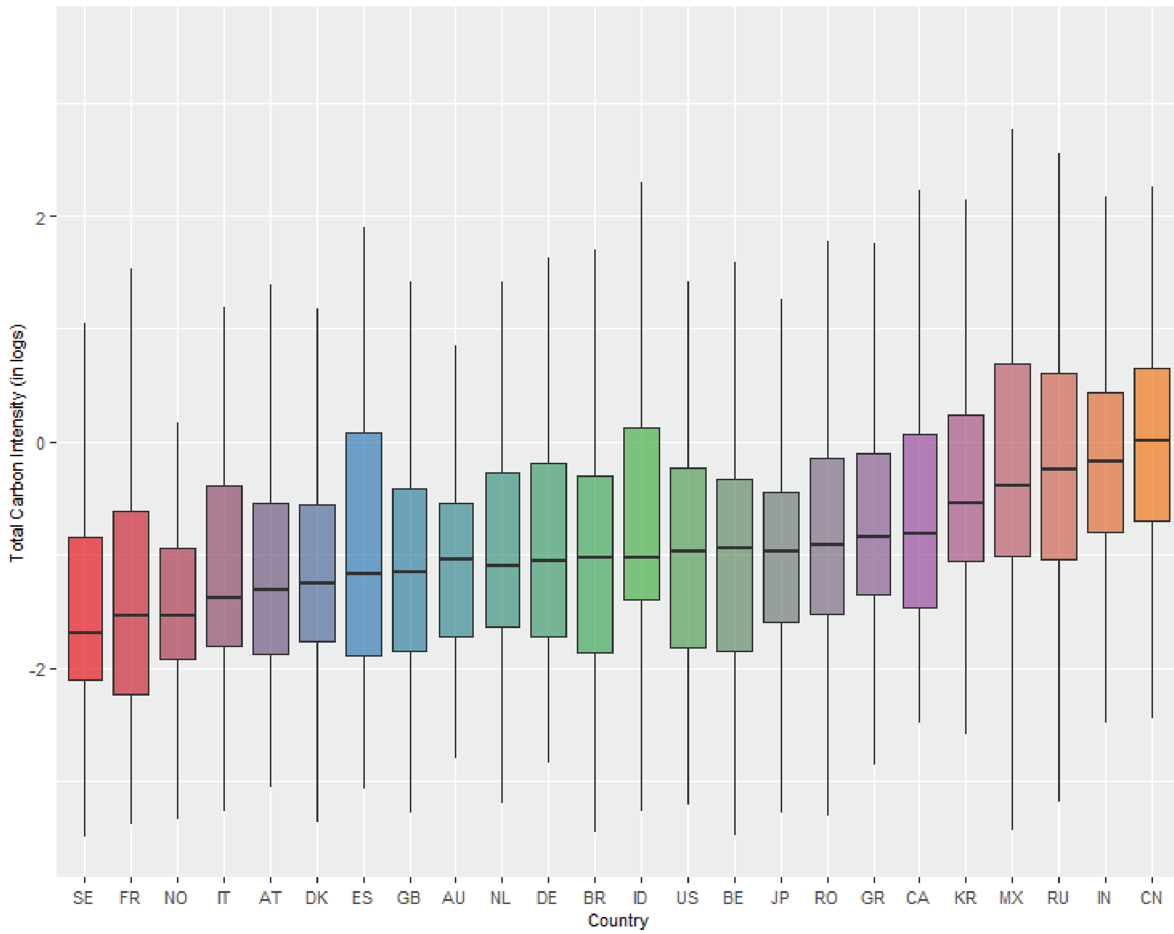
To explain in the context of multiple open countries, a multi-region input-output table (MRIO) extends the national I-O tables to many countries and accounts for the dollar value from one country-industry used in the production of 1$ of output in any other country-industry. If there are N countries and S industries in each country, then A is a MRIO matrix, $[NS \times NS]$, describing this "global" supply chain. Let X be a $[NS \times 1]$ vector of gross output and D be a $[NS \times 1]$ vector of final demand. Then the accounting equation, $X = AX + D$ states that an industry's gross output is equal to its output value used as intermediate goods and as final demand. Here, $X = (I - A)^{-1}D$. This enables us to calculate the Leontief inverse, $L \equiv (I - A)^{-1}$. Let L_{ijst} denote the entry of the multi-region Leontief inverse, that is, the dollar of output from industry s in country i required to produce \$1 of output in industry t in country j.

If E^d_{is} is the direct emissions from combustion of fossil fuels in industry s in country i, then the total carbon intensity of a sector-country, E_{it} is measured by incorporating the

Leontief inverse:

$$E_{jt} = \sum_{i,s} L_{ijst} E_{is}^{d} \tag{1}$$

Figure 1: Summary Statistics: Total Carbon Intensity



The variation in the total carbon intensity of producing a $ of output across countries and products. Each box shows the median and the interquartile range in the carbon intensity across products in each country. Carbon Intensity measure the metric tons of CO_2 emitted per $ of output directly (through the combustion of fossil fuels) as well as indirectly (through the use of intermediates)

Figure 1 summarizes the CO_2 intensity of production in each country. The figure shows the variation within and across countries in total carbon emissions of tradable goods for selected countries, arranged in ascending order.[13] For each country, the boxplot shows the

[13] Countries from left to right: Sweden, France, Norway, Italy, Austria, Denmark, Spain, United Kingdom, Australia, Netherlands, Germany, Brazil, Indonesia, United States, Belgium, Japan, Romania,Greece, Canada, South Korea, Mexico, Russia, India, China.

median and interquartile range (25th-75th percentile) across tradable industries. One of the key takeaways from Figure 1 is that countries with relatively relaxed environmental regulations, like Russia, Mexico, India, and China, are at the right extreme of the graph. Most European countries occupy the left extreme, with the median in China being eight times that of Sweden. This figure complements the finding that differences in the emission intensities vary hugely across countries[14].

2.2 Market Power and Demand Elasticity

To quantify the sources of market power, I estimate the import demand elasticity of various country-sector pairs, using the hybrid estimator method (LIML) proposed in Soderbery (2015). The estimation procedure is based on the approach developed by Feenstra (1994), but addresses potential small sample biases and grid search inefficiencies in previous applications. Elasticity estimates based on the Feenstra method have been frequently used and referred to in various papers in the literature (Broda et al. (2008), Hsieh and Klenow (2009), Khandelwal (2010), and Ossa (2014)). Soderbery (2015) approach is also consistent with the theoretical framework as the demand side in both settings is derived from CES preferences. This section details the estimation methodology used by Soderbery (2015) to estimate demand elasticity.

Specifically, the methodology introduces a time subscript t and time-variety-specific taste shocks $b_{jst}(\omega)$ into the CES aggregator. Here, q_{jst} is the quantity of good s demanded in country j at time period t, which is a CES aggregate over varieties ω.

$$q_{jst} = \left[\int b_{jst}^{j(k)}(\omega)^{\frac{1}{\sigma_{js}}} q_{jst}(\omega)^{\frac{\sigma_{js}-1}{\sigma_{js}}} d\omega \right]^{\frac{\sigma_{js}}{\sigma_{js}-1}} \tag{2}$$

I treat each HS6-country[15] pair that we observe in data as one variety ω who is the winner of the competition in a subset of HS2 products. I follow Soderbery (2015) and Broda and Weinstein (2006) and allow for a potentially upward-sloping export supply curve, in

[14]Stylized Fact 5 in Copeland et al. (2021)

[15]HS6 here refers to the Harmonized Commodity Description and Coding System, generally referred to as "Harmonized System". It comprises more than 5,000 commodity groups; each identified by a six-digit code, arranged in a legal and logical structure, and supported by well-defined rules to achieve uniform classification.

which case this structure implies demand and supply curves of the form

$$\Delta^m ln(s_{jst}(\omega)) = -(\sigma_{js} - 1)\Delta^m ln(p_{jst}(\omega)) - \xi_{jst}(\omega) \quad (3)$$

$$\Delta^m ln(p_{jst}(\omega)) = \left[\frac{\kappa_{js}}{1 + \kappa_{js}}\right]\Delta^m ln(s_{jst}(\omega)) + \delta_{jst}(\omega) \quad (4)$$

where Δ^m denotes double differencing with respect to time and a reference variety m, κ_{js} denotes the inverse export supply elasticity for good s, s_{jst} its expenditure share, and $\xi_{jst}(\omega)$ and $\delta_{jst}(\omega)$ reflect unobservable demand and supply shocks.

Following Feenstra (1994)'s identifying assumption that these demand and supply shocks are orthogonal, i.e., $E[\xi_{jst}(\omega)\delta_{jst}(\omega)] = 0$, one can then multiply the two shocks to convert the structural equations of demand and supply into one estimation equation

$$\left(\Delta^m ln(p_{jst}(\omega))\right)^2 = \lambda_{1,s}\left(\Delta^m ln(s_{jst}(\omega))\right)^2 + \lambda_{2,s}\left(\Delta^m ln(p_{jst}(\omega))\right)\left(\Delta^m ln(s_{jst}(\omega))\right) + u_{jst} \quad (5)$$

where $\lambda_{1,s} = \kappa_{js}/[(\kappa_{js}+1)\cdot(\sigma_{js}-1)]$ and $\lambda_{2,s} = [1-\kappa_{js}(\sigma_{js}-2)]/[(\kappa_{js}+1)\cdot(\sigma_{js}-1)]$, which can be consistently estimated using 2SLS estimation with variety indicators as instruments.[16]

I employ bilateral trade data on the HS6 level for the years between 1995 and 2015 and estimate σ_{js} separately for each HS2 sector and country to allow for the possibility that traded varieties of each good as well as the demand for them may differ across countries. Table 2 provides summary statistics and shows the distribution of the estimated import demand elasticities across countries. I estimate σ to be particularly low for Australia, the U.K., Italy, and Japan. On the other end, I estimate comparably large elasticities of substitution for India, Canada, Belgium, and China.

The parameter estimate σ plays an important role as it quantifies model-implied profits ($\frac{1}{\sigma}$) and markups ($\frac{\sigma}{\sigma-1}$) as is detailed in the quantitative model in Section 4. Essentially, lower demand elasticity in an industry (σ) implies that firms enjoy higher market power, reap higher implied profits and charge higher markups.

[16] Following Soderberry(2015), I weight varieties by their respective estimated residuals to limit the impact of outliers.

Table 2: Distribution of parameter estimates for σ

σ	Median	1st Quartile	3rd Quartile
Australia	1.93	1.47	3.11
Austria	2.76	1.70	6.31
Belgium	3.15	1.94	6.73
Brazil	2.58	1.74	4.28
Canada	4.41	2.09	11.13
China	3.05	1.85	6.52
Denmark	2.40	1.67	4.71
France	2.49	1.64	4.98
Germany	2.65	1.70	5.23
Greece	2.27	1.68	3.59
India	3.48	2.08	7.68
Indonesia	2.37	1.70	3.87
Italy	2.10	1.53	3.71
Japan	2.19	1.61	3.65
Rep. of Korea	2.63	1.70	4.65
Mexico	2.64	1.77	4.89
Netherlands	2.45	1.65	4.71
New Zealand	2.70	1.78	4.77
Norway	2.30	1.72	3.38
Romania	2.48	1.70	4.19
Russia	2.53	1.73	4.38
Spain	2.56	1.76	4.11
Sweden	3.04	1.79	6.73
United Kingdom	1.96	1.50	3.34
USA	2.49	1.61	5.99
ROW	2.72	1.58	7.02

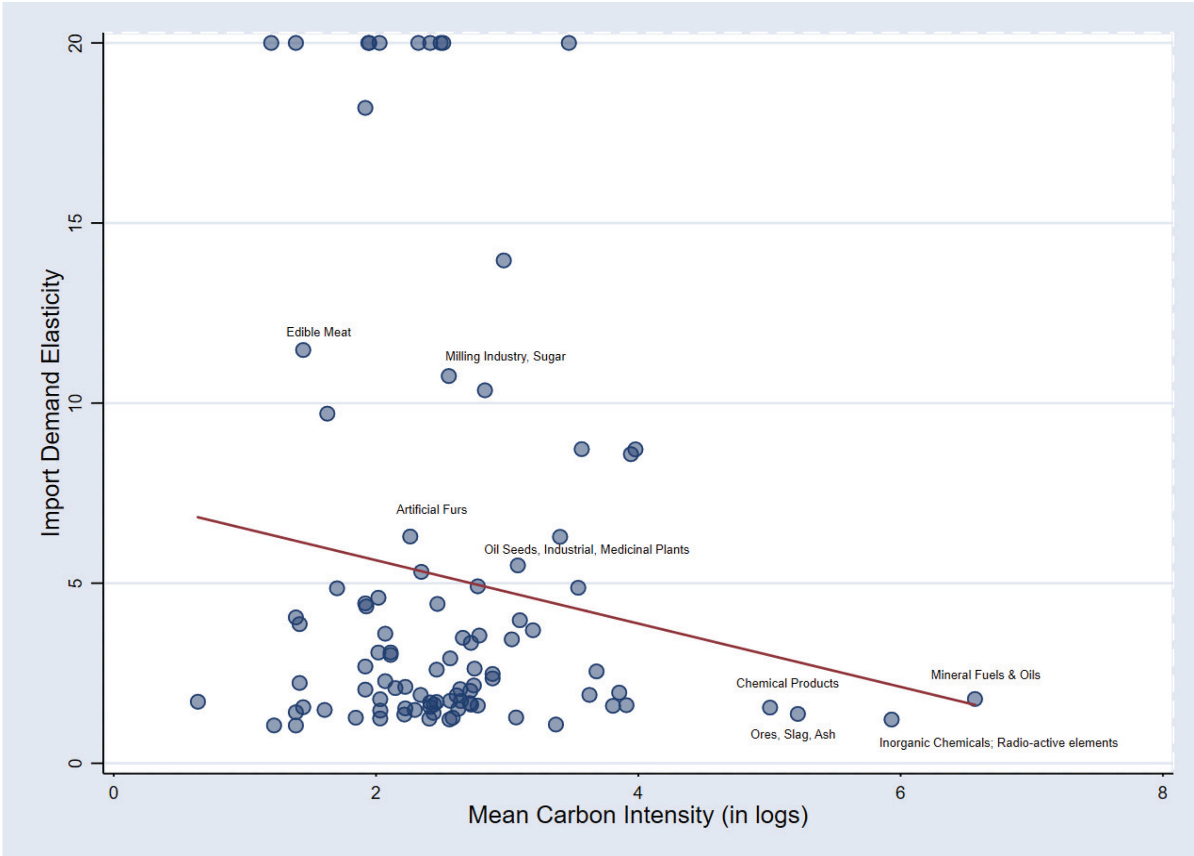
Notes: This table provides summary statistics for the parameter estimates of σ. The median and quartiles are taken over product categories.

3 Relationship between Carbon Intensity and Market Power

In this section, I provide suggestive evidence about the correlation between the carbon intensity of sectors with market power. Figure 2 plots the two variables to motivate the relationship. The figure indicates that carbon-intensive sectors have relatively lower demand elasticity, suggesting that the carbon-intensive sectors have higher market power. To further support the correlation, I run a regression of the per-unit model implied profits ($\frac{1}{\sigma_{js}}$) on the total carbon intensity (E_{js}) of the country-sector pair. As implied by the measured demand elasticity, the results support a positive correlation between the two variables, indicating that the carbon-intensive sectors are, *on average*, high markup sectors.

Exiobase reports the direct emissions from the combustion of fossil fuels for country-

Figure 2: Relationship between Carbon Intensity and Demand Elasticity



Relationship between carbon intensity and demand elasticity. The x axis plot the mean carbon intensity across countries, weighted by the expenditures. The y axis plots the estimated demand elasticity

sector pairs. I use the global input-output tables to infer indirect and total emissions for a particular good using the Leontief matrix. The database uses product codes closely based on the International Standard Industrial Classification. The elasticity estimates are based on the Comtrade database, which uses Harmonized System (HS 2012) classification. I use published concordances to crosswalk each of the Exiobase product codes to an HS2 category using import share as weights[17]. I run the regression with a sample of 24 major economies and the rest of the world aggregate for 96 HS-2 digit industries. I run two sets of regressions to establish the association between carbon intensity and market power. The first uses the US industry-specific elasticity to infer the market power of an industry. The second allows the elasticity to vary across countries.

Table 3 reports the results when US elasticity is used to infer the industry's market power. I use the implied profits in the model, $1/\sigma$, as the dependent variable and regress it on an

[17]Concordances between various industrial classifications can be accessed here

industry's mean total carbon intensity by weighing each country by its import size. The estimates reported in Columns (1) and (2) indicate that carbon-intensive sectors are, on average, associated with higher per-unit implied profits. Measuring total CO_2 emissions from the input-output table may invite measurement errors as it requires harmonization of industrial data across many countries. To address the potential measurement error, I use direct emissions as an instrumental variable for the total emissions. I use the two stages SLS to investigate if the relationship still holds. Columns (4) and (5) report the results of the two-stage IV regression. The results confirm the suggestion that carbon-intensive sectors are correlated with higher profits.

Table 3: Regression of US Elasticity on Mean Carbon Intensity, by Industry

	(1)	(2)	(3)	(4)	(5)
Dependant Variable	$1/\sigma$	$\log(1/\sigma)$	σ	$1/\sigma$	$\log(1/\sigma)$
Mean Carbon Intensity (log)	0.020**	0.066***	-0.379***	0.022**	0.070**
	(0.009)	(0.023)	(0.141)	(0.009)	(0.027)
Constant	0.357***	-1.365***	6.231***	0.353***	-1.376***
	(0.040)	(0.132)	(0.886)	(0.041)	(0.137)
Observations	96	96	96	96	96
IV Regression	No	No	No	Yes	Yes

Standard errors in parentheses
Independant Variable: Mean Carbon Intensity by industry, weighted by the import of each country
Data source: Exiobase and ComTrade
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the second case, I use the estimated country-specific elasticity to allow for the variation in the market power across industries and countries. To measure the difference in the implied profits between clean and dirty industries, I estimate the following equation:

$$\tilde{\pi}_{js} = \alpha_0 + \alpha E_{js} + \mu_j + \epsilon_{js} \quad (6)$$

Here, E_{js} reflects the total carbon intensity of sector s in country j. This means that it incorporates not only the direct emissions by the combustion of fossil fuels, but also the indirect emissions incorporated in the inputs used in the production of the sector s. The parameter α represents the difference in the implied profits $(1/\sigma)$ of carbon-intensive (dirty) industries with respect to cleaner industries; μ_j reflects country fixed effects.

Column (1) in Table 4 reports the value of α. The results indicate that the association of carbon-intensive sectors with higher implied profits still holds. In Column (2), I weigh each country sector by the size of their import share and still find a positive association. The results for the IV regression are reported in Column (4), indicating a more positive and statistically significant relationship between the implied profits and the carbon intensity.

Table 4: Regression of Inverse Elasticity on Carbon Intensity

	(1)	(2)	(3)	(4)
Total carbon intensity (Log)	0.018*	0.013*	0.025**	0.019**
	(0.009)	(0.007)	(0.011)	(0.009)
Constant	-1.802***	-1.819***	-1.851***	-1.856***
	(0.114)	(0.119)	(0.124)	(0.124)
Observations	2399	2398	2399	2398
Country FE	Yes	Yes	Yes	Yes
Weights	No	Yes	No	Yes
IV Regression	No	No	Yes	Yes

Standard errors in parentheses

Dependant variable: Inverse of Demand Elasticity (Sigma)

Data source: Exiobase and ComTrade

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In light of this suggestive evidence, carbon-based tariffs could have sizable welfare implications in reshuffling bilateral profits. Notably, a tariff could be hurtful for countries that export carbon-intensive goods to destinations with high implied profits. This is because their export profits dwindle in response to production potentially shifting to the destination country or other countries with comparative advantages. Consequently, for an importing country with a positive association between carbon intensity and implied profits, a carbon-based tariff might be beneficial as it shifts the production of the high markup sector domestically, reaping additional aggregate profits. A country that exports goods to destinations where carbon-intensive sectors are high markup goods are the ones to lose the most from carbon-based tariffs. On the contrary, a country that imports high-markup carbon-intensive goods is most likely to gain from carbon-based tariffs. The extent to which carbon-based tariffs result in aggregate welfare gains for a country is mired in its complex relationship of carbon intensity with market power, market shares, and its global input-output linkages. In the next section, I build a quantitative structural model of trade which tractably accounts for the profit-shifting channel to quantify the welfare implications of carbon-based tariff changes.

4 Quantitative Model

This section details the paper's theoretical contribution, which enables us to account for markups and profit shifting in the counterfactual analysis of carbon-based tariff changes. Trade between countries creates transboundary pollution, which negatively affects the utility. Countries have pre-existing and sub-optimal trade policy where tariffs are lower on carbon-intensive sectors. In the counterfactual scenario, the tariffs adjust to accommodate the carbon content in its offshore production. Such a policy internalizes the cost of carbon and indirectly regulates the emissions among the trading partners. The heterogeneity in the import demand elasticity across countries and sectors drives the responsiveness of the import demand to these carbon-based tariff adjustments. Production relocates to countries with relative comparative advantage in response to the adjusted import costs. The industries are linked through global input-output linkages. As a result, the tariff adjustments generate general equilibrium effects, which account for the changes in the industry expenditures and, thus, the aggregate revision of the market shares and the bilateral profits. The objective is to measure the effects on social welfare while accounting for profits in response to tariff shocks, considering the global linkages.

4.1 Model

Preferences: The representative agent in country j maximizes national utility which is Cobb-Douglas over various sectors s

$$U_j = \Pi_s Q_{js}^{\beta_{js}} f(Z) \tag{7}$$

Q_{js} is the quantity of output consumed in sector s. β_{js} is the weight given to each sector by the consumer for final consumption. $f(Z)$ is the climate damage function which accounts for the disutility from increase in global emissions Z[18].

[18]Specifically, functional form of climate damage is $f(Z) = [1+\delta(Z-Z_0)]^{-1}$. δ is the damage parameter, Z_0 refers to the baseline global emissions and Z refers to the counterfactual global emissions. δ is estimated such that one unit increase in global emissions ($Z-Z_0$) leads to a utility decline equivalent to the calculated social cost of carbon (approx. \$40 per ton).

Sector s is a CES aggregate over various varieties, depending on the source country (Armington (1969)). The elasticity of substitution is allowed to differ at country-sector level.

$$Q_{js} = \Big(\sum_i q_{ijs}^{\frac{\sigma_{js}-1}{\sigma_{js}}}\Big)^{\frac{\sigma_{js}}{\sigma_{js}-1}} \tag{8}$$

The representative consumer chooses q_{ijs} to maximize Q_{js} such that the total expenditure on good s in country j is $X_{js} = \sum_i p_{ijs}q_{ijs}$. Solving the maximization problem gives us

$$q_{ijs} = \frac{p_{ijs}^{-\sigma_{js}}}{\sum_i p_{ijs}^{1-\sigma_{js}}} X_{js} \tag{9}$$

Total expenditure on good s from country i in country j is

$$X_{ijs} = p_{ijs}q_{ijs} = \frac{p_{ijs}^{1-\sigma_{js}}}{\sum_i p_{ijs}^{1-\sigma_{js}}} X_{js} \tag{10}$$

Thus the market share of country i in country j is $\lambda_{ijs} = \frac{p_{ijs}^{1-\sigma_{js}}}{\sum_i p_{ijs}^{1-\sigma_{js}}}$

Technology: Production of good s in country i requires labor and intermediate inputs. Thus the unit cost is given by

$$c_{is} = w_i^{1-\alpha_{is}} \Pi_k P_{ik}^{\alpha_{iks}} \tag{11}$$

where α is the cost share to produce the goods.

To serve market j, the representative firm in country i incurs an additional cost in terms of tariff and non-tariff barriers given by $\phi_{ijs} = (1+t_{ijs})(1+\eta_{ijs})$. Here t_{ijs} are the ad-valorem tariffs imposed by country j on good s from country i. η_{ijs} are the non-tariff barriers. Thus the marginal cost of exporting good from country i to country j is $c_{ijs} = \phi_{ijs}c_{is}$. Intranational trade costs equal one:$\phi_{jjs} = 1$

Pollution: The pollution emitted to sell goods in a particular destination is

$$Z_{ijs} = \gamma_{is} \frac{X_{ijs}}{p_{ijs}} \tag{12}$$

The coefficient γ_{is} refers to the pollution intensity of producing good s in country i, which is the metric tons of CO_2 emitted per dollar of output in the fossil fuel extraction. A country's total pollution emission is the sum of emission incurred for domestic production and its exports. Thus $Z_i = \sum_{s,j} Z_{ijs}$. The global pollution is $Z = \sum_i Z_i$

Market Structure: We assume that firms in industry s and in country j engage in monopolistic competition and choose prices to maximize profits, given the preference equations. The price charged is a markup over the marginal cost of production

$$\pi_{ijs} = p_{ijs}q_{ijs} - c_{ijs}q_{ijs}$$

Maximizing profits given (8) gives us

$$p_{ijs} = \frac{\sigma_{js}}{\sigma_{js} - 1} c_{ijs} \tag{13}$$

The aggregate price in country j for sector s is

$$P_{js} = \left[\sum_i p_{ijs}^{1-\sigma_{js}}\right]^{\frac{1}{1-\sigma_{js}}} = \frac{\sigma_{js}}{\sigma_{js}-1}\left[\sum_i c_{ijs}^{1-\sigma_{js}}\right]^{\frac{1}{1-\sigma_{js}}} = \frac{\sigma_{js}}{\sigma_{js}-1}\left[\sum_i [w_i^{1-\alpha_{is}}\Pi_k P_{ik}^{\alpha_{iks}}\phi_{ijs}]^{1-\sigma_{js}}\right]^{\frac{1}{1-\sigma_{js}}} \tag{14}$$

Aggregation: Total expenditure incurred on goods from country j in sector s is the expenditure incurred by the representative consumer as final consumption expenditure on that good (which is the share of total income) and the expenditure incurred by other industries k using s as intermediate inputs. The total income, $I_j = w_j L_j + \pi_j + T_j + D_j$

$$X_{js} = \beta_{js} I_j + \sum_k \alpha_{jsk} R_{jk} = \beta_{js}(w_j L_j + \pi_j + T_j + D_j) + \sum_k \alpha_{jsk} \sum_i \frac{\lambda_{jik} X_{ik}}{(1 + t_{ijk})} \tag{15}$$

The total profits π_j can be written as

$$\pi_j = \sum_s \sum_i \frac{1}{\sigma_{is}} \frac{\lambda_{jis} X_{is}}{1 + t_{jis}} \tag{16}$$

The total tariff revenues:

$$T_j = \sum_s \sum_i \frac{t_{ijs}}{1 + t_{ijs}} \lambda_{ijs} X_{js} \tag{17}$$

The expenditures are then given as:

$$X_{js} = \beta_{js}(w_j L_j + \sum_s \sum_i \frac{1}{\sigma_{is}} \frac{\lambda_{jis} X_{is}}{1+t_{jis}} + \sum_s \sum_i \frac{t_{ijs}}{1+t_{ijs}} \lambda_{ijs} X_{js} + D_j) + \sum_k \alpha_{jsk} \sum_i \frac{\lambda_{jik} X_{ik}}{(1+t_{ijk})} \tag{18}$$

The deficits,D_j are given by:

$$\sum_s \sum_i \frac{\lambda_{jis} X_{is}}{1+t_{jis}} = \sum_s \sum_i \frac{\lambda_{ijs} X_{js}}{1+t_{ijs}} - D_j \tag{19}$$

Value Added and Labor Market Clearing: The value added of a sector-country is the difference between the revenues and the total expenditures on intermediate goods. This means,

$$Y_{js} = \sum_i \frac{\lambda_{jis} X_{is}}{1+t_{jis}} - \sum_k \alpha_{jks} \sum_i \lambda_{ijk} X_{jk} = R_{js} - \alpha_{js} R_{js} = (1-\alpha_{js}) R_{js}$$

The value added is the revenues after paying for intermediate inputs. This revenue is then allocated between labor in form of wages ($w_j L_j$) and profits.

$$w_j L_j = \sum_s Y_{js} - \sum_s \sum_i \frac{1}{\sigma_{is}} \frac{\lambda_{jis} X_{is}}{1+t_{jis}}$$

$$w_j L_j = \sum_s \Big[\sum_i \frac{\lambda_{jis} X_{is}}{1+t_{jis}} - \sum_k \alpha_{jks} \sum_i \lambda_{ijk} X_{jk} \Big] - \sum_s \sum_i \frac{1}{\sigma_{is}} \frac{\lambda_{jis} X_{is}}{1+t_{jis}}$$

This gives us

$$w_j L_j = \sum_s \sum_i \frac{\sigma_{is}-1}{\sigma_{is}} \frac{\lambda_{jis} X_{is}}{1+t_{jis}} - \sum_s \sum_k \alpha_{jks} \sum_i \lambda_{ijk} X_{jk} \tag{20}$$

In baseline and counterfactual equilibrium, consumers maximize their utility, firms maximize profits, deficits remain the same and markets clear.

4.2 Counterfactual Analysis

For a counterfactual change in trade costs given by $\hat{\phi}_{ijs}$, I use hat algebra to solve for equilibrium in changes (Dekle et al. (2008); Caliendo and Parro (2015)). If x' denotes a variable in the counterfactual scenario and x denotes the variable in the baseline, then $\hat{x} = \frac{x'}{x}$

$$\hat{P}_{js} = \left[\sum_i \lambda_{ijs} (\hat{c}_{is} \hat{\phi}_{ijs})^{1-\sigma_{js}} \right]^{\frac{1}{1-\sigma_{js}}} \tag{21}$$

$$\hat{\lambda}_{ijs} = \left(\frac{\hat{c}_{is} \hat{\phi}_{ijs}}{\hat{P}_{js}} \right)^{1-\sigma_{js}} \tag{22}$$

$$\hat{c}_{is} = \hat{w}_i^{1-\alpha_{is}} \Pi_k \hat{P}_{ik}^{\alpha_{iks}} \tag{23}$$

Counterfactual changes in trade costs ϕ_{ijs} affect the marginal cost of importing goods from country i to the country j. This change alters the aggregate price of good s in the country j (Equation 21). The price changes affect the marginal cost of producing goods through input-output linkages (Equation 23). This in turn alters the market share of country i in country j (Equation 22). I solve the above system of non-linear equations to get $\hat{\lambda}_{ijs}$ and $\hat{P}_{js}$.

To close the model, I then solve for the counterfactual sector-country expenditures. This is motivated by the fact that in the counterfactual equilibrium, expenditures equal revenues.

$$\hat{X}_{js} X_{js} = \beta_{js} \Big[\hat{w}_j (w_j L_j) + \sum_s \sum_i \frac{1}{\sigma_{is}} \frac{\hat{\lambda}_{jis} \lambda_{jis} \hat{X}_{is} X_{is}}{1 + t'_{jis}} + \sum_s \sum_i \frac{t'_{ijs}}{1 + t'_{ijs}} \hat{\lambda}_{ijs} \lambda_{ijs} \hat{X}_{js} X_{js} + D_j \Big] + \sum_k \alpha_{jsk} \sum_i \frac{\hat{\lambda}_{jik} \lambda_{jik} \hat{X}_{ik} X_{ik}}{(1 + t'_{ijk})} \tag{24}$$

Equation 24 is a system of equations where the expenditures X_{js} is not only dependant on expenditures in other countries X_{is} but also expenditures on different sectors in different countries X_{ik} through input-output linkages.

After solving the above system of equation, I solve for labor market clearing which implies that the country deficits remain the same in the counterfactual.

$$\hat{w}_j w_j L_j = \sum_s \sum_i \frac{\sigma_{is} - 1}{\sigma_{is}} \frac{\hat{\lambda}_{jis} \lambda_{jis} \hat{X}_{is} X_{is}}{1 + t'_{jis}} - \sum_s \sum_k \alpha_{jks} \sum_i \hat{\lambda}_{ijk} \lambda_{ijk} X_{jk} \hat{X}_{jk} \tag{25}$$

To find the changes in the real income (welfare) $\hat{V}_j$, I use the changes in wages, profits, tariff revenues and prices to infer

$$\hat{V}_j = \frac{\hat{w}_j w_j L_j + \hat{\pi}_j + \hat{T}_j + D_j}{\hat{P}_j} \tag{26}$$

where

$$\hat{P}_j = \Pi_s \hat{P}_{js}^{\beta_{js}} \tag{27}$$

The effect on aggregate profits in country j is:

$$\hat{\pi}_j = \frac{\pi_j^{'}}{\pi_j} = \frac{\sum_{s,i} \frac{1}{\sigma_{is}} \frac{\hat{\lambda}_{jis}\lambda_{jis}\hat{X}_{is}X_{is}}{1+t_{jis}^{'}}}{\sum_{s,i} \frac{1}{\sigma_{is}} \frac{\lambda_{jis}X_{is}}{1+t_{jis}}}$$

Total change in pollution in a given country for a change in tariffs are given as

$$\hat{Z}_i = \frac{\sum_s \gamma_{is}\lambda_{ijs}\hat{\lambda}_{ijs}X_{js}\hat{X}_{js}/P'_{ijs}}{\sum_s \gamma_{is}\lambda_{ijs}X_{js}/P_{ijs}} \tag{28}$$

4.3 Estimation Procedure

1. Start with a guess of a vector of ones for the changes in the marginal costs and wage changes: $(\hat{c}_{is}, \hat{w}_{is})$. This is $[N \times S + N]$ vector
2. Use Equations 21-23 to calculate $\hat{P}_{js}$ and $\hat{\lambda}_{ijs}$
3. Calculate the "ERROR 1" by using Equation 23: the unit cost equation. This is $[N \times S]$
4. Use Equation 24 to solve for counterfactual expenditures given $\hat{P}_{js}$ and $\hat{\lambda}_{ijs}$
5. Calculate "ERROR 2" using Equation 25. This is $[N \times 1]$
6. Use an optimizer to find $(\hat{c}_{is}, \hat{w}_i)$ such that the errors $[N \times S + N]$ are minimized.
7. Use Equations 26 and 28 to infer the changes in real income and carbon emissions.

I elaborate on these steps in detail in Appendix Section B.

5 Results

In this section, I report the results of a counterfactual trade policy reform where a country's import tariffs are adjusted to accommodate the carbon embodied in its imports. Particularly, each country's bilateral import tariffs are set equal to the country's weighted mean baseline bilateral tariffs, with weights equal to the baseline bilateral trade value.

$$t_{ijs}^{'} = \frac{\sum_s \frac{t_{ijs}X_{ijs}}{1+t_{ijs}}}{\sum_s \frac{X_{ijs}}{1+t_{ijs}}}$$

As documented in Shapiro (2021), the baseline tariffs are lower for carbon-intensive imports. This fact is referred to as the environmental bias of trade policy and is widespread across all countries. Thus, this counterfactual effectively increases import tariffs for carbon-intensive industries and lowers them for cleaner industries. This trade policy reform is similar to the quantitative assessment in Shapiro (2021). More generally, low protection in carbon-intensive *upstream* industries stems from a strong industry lobby for lower import costs for these upstream goods. Downstream goods, which are relatively less carbon-intensive, have higher tariffs as consumers are poorly organized to lobby for lower tariffs. From the policy perspective, WTO has sought to decrease the protection of downstream industries relative to upstream industries since such reform would let developing countries sell more advanced goods to rich countries. This counterfactual is similar to WTO's policy objective of imposing uniform bilateral tariffs across all goods while pricing the carbon emissions embodied in its imports.

If we eliminate the environmentally-biased price signal in our trade policy, will such a reform lead to a reduction in global emissions? What will be the welfare impact across countries? In particular, since carbon-intensive goods have higher market power, the paper aims to quantify the importance of market power in estimating the effect of the trade policy reform on global emissions and real income across countries.

I run two models to quantitatively assess the significance of the market power channel: one with perfect competition and the other with monopolistic competition. Table 5 reports the results in a perfect competition model. Table 6 reports the results of imperfect competition. In both the tables, I report the changes in the *real* variables, except for emission changes ($\frac{Z_j^{'}}{Z_j}$). For example, Column 1 in Table 5 reports the changes in real income ($\frac{I_j^{'}/P_j^{'}}{I_j/P_j}$). For example, if the table reports the real income change is 1.02, the real income increases by 2% in the counterfactual policy relative to the baseline policy. Global changes in real income, real wages, real tariff revenues, and aggregate prices are calculated by weighing each country by its GDP share in the baseline. In both cases, global emissions decrease with the trade policy reform. In perfect competition, the global emissions reduce by 3.7%, while in imperfect competition, the global emissions reduce by 5.7%. In other words, the results highlight the potency of redesigning trade barriers as an effective second-best policy tool to curtail global emissions.

Perfect Competition: As reflected in Table 5, the policy reform leads to heterogeneous effects across countries in perfect competition. Notably, all countries gain real income. This is augmented by a decline in aggregate prices in many countries, as given in equation (27). Aggregate prices reduce because the counterfactual reform also lowers the baseline tariffs in cleaner industries. Cleaner industries are downstream goods with a higher share in final consumption β_{js}, predicting an overall decline in aggregate prices. The policy reform causes the global emissions to reduce by 3.76% and increase global real income by 1.98%.[19]Eliminating or harmonizing trade policy across goods can increase real income because baseline trade policy encourages consumption and production of dirty goods. Eliminating this price signal decreases the consumption and production of those dirty goods and encourages the consumption and production of downstream cleaner goods.

Market Power and Imperfect Competition: The stylized fact presented in Section 3 highlights that carbon-intensive sector are correlated with higher market power. As a result, carbon-based policy reform leads to a sizable shifting of aggregate profits and markups across countries. How does the markup channel affect the impact of policy change on global emissions? Do all countries benefit from this reform (as is the case with perfect competition)? In Table 6, I present the results of imperfect markets that account for profit shifting across countries.

The most striking result is that accounting for market power increases the effectiveness of trade policy to reduce global emissions. The emissions reduce by 5.7%, relative to 3.7% in the case of perfect competition. When import tariffs increase on carbon-intensive goods, countries find it expensive to trade them and reallocate production to the domestic country. The fact that these sectors reap high profits expedites the extent of the domestic reallocation of these industries. Two facts in the data further support the higher decline in global emissions in imperfect competition. Advanced economies are, on average, net

[19] Shapiro (2021) finds global emissions reduce by 3.6% while a moderate increase in real income by 0.6% for the same policy experiment.

Table 5: Counterfactual Results: Perfect Competition

	Welfare	Emissions	Wages	Tariff Rev.	Prices
Global	1.0198	0.9624	1.0198	1.0317	0.9809
Australia	1.0002	0.9916	1.0004	0.9812	1.0126
Austria	1.0233	0.9877	1.0259	0.9897	0.9853
Belgium	1.0227	0.9669	1.0269	0.9825	0.9742
Brazil	1.0201	1.0016	1.0196	1.0489	0.9708
Canada	1.0256	0.9901	1.0259	1.0074	0.9873
China	1.0298	0.9791	1.0249	1.1077	1.0106
Denmark	1.0147	0.9863	1.0192	0.9536	0.9828
France	1.0129	1.0105	1.0162	0.9746	0.9946
Germany	1.0256	1.0097	1.0277	0.9986	0.9797
Greece	1.0273	0.9116	1.0308	0.9996	0.9602
India	1.0077	1.0223	1.0046	1.1245	0.9785
Indonesia	1.0037	0.9928	1.0039	0.9847	0.9951
Italy	1.0260	0.9620	1.0301	0.9892	0.9741
Japan	1.0185	1.0058	1.0181	1.0612	1.0067
Mexico	1.0070	0.9863	1.0053	1.0960	0.9996
Netherlands	1.0029	0.9365	1.0059	0.9525	1.0061
Norway	1.1085	0.9933	1.1124	1.0666	0.9091
Romania	1.0476	0.9285	1.0547	1.0015	0.9335
Russia	1.0351	0.9262	1.0339	1.1368	1.0263
Spain	1.0164	0.9766	1.0199	0.9774	0.9867
Sweden	1.0365	0.9708	1.0428	0.9731	0.9776
United Kingdom	1.0111	0.9974	1.0128	0.9782	0.9976
United States	1.0031	0.9959	1.0031	1.0033	1.0040
ROW	1.0169	0.8842	1.0211	0.9745	0.9546

Notes: This table reports the $\hat{x} = x'/x$ for the counterfactual changes in tariffs in a perfectly competitive markets. The aggregate real income is decomposed using Equation 26 in terms of real wages and real tariff revenues. Profits are not accounted here. The global values for economic variables are calculated by weighing each country by its relative GDP. The change in emissions, $Z = \sum_i \hat{Z}_i' / \sum_i Z_i$

importers of carbon-intensive goods.[20]Second, the emission intensity is relatively lower in advanced economies.[21] As domestic reallocation speeds up, the production of carbon-intensive goods reallocates to advanced economies with lower emission intensity. Figure 3 plots the change in domestic profits of carbon-intensive sectors and the emission intensity against the net imports of most carbon-intensive sectors. There are three takeaways from the figure. One, net importers of carbon-intensive goods gain a higher increase in domestic profits (blue). Second, The net importers have a lower emission intensity (or-

[20]see Figure 14 and Figure 15
[21]see Figure 1

Table 6: Counterfactual Results: Imperfect Competition, US Elasticity

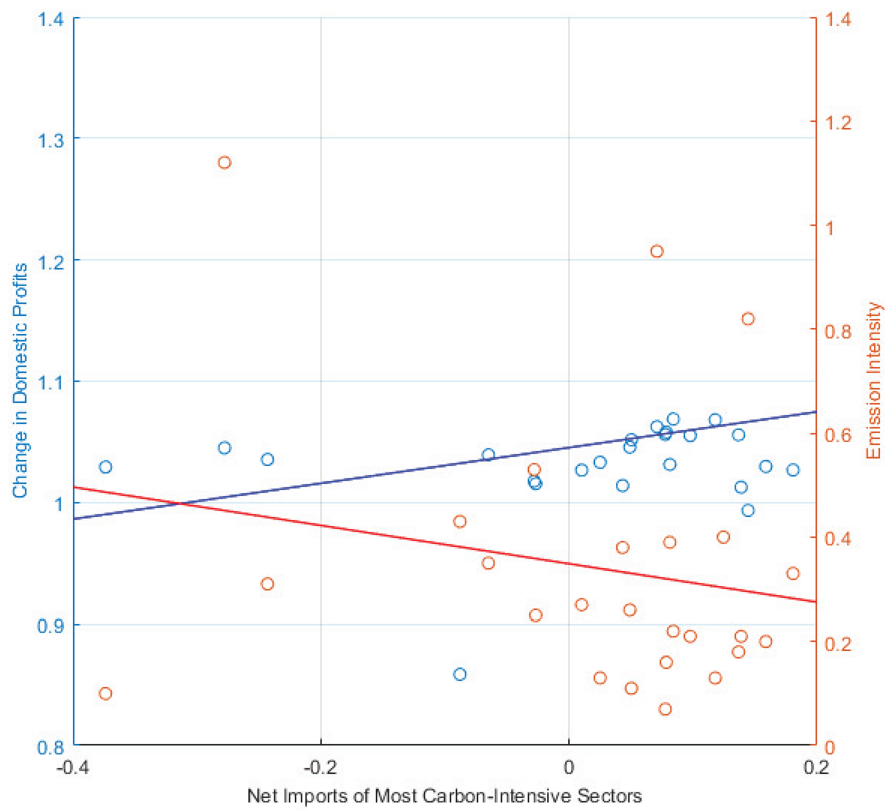
	Welfare	Emissions	Wages	Profits	Tariff Rev.	Prices
Global	1.0076	0.9430	1.0196	0.9969	1.0139	0.9931
Australia	1.0011	0.9962	0.9989	1.0042	0.9784	1.0218
Austria	1.0154	0.9819	1.0211	1.0160	0.9720	0.9984
Belgium	1.0080	0.9552	1.0238	1.0007	0.9625	0.9878
Brazil	1.0148	0.9936	1.0192	1.0063	1.0385	0.9813
Canada	1.0290	0.9940	1.0243	1.0401	1.0058	0.9963
China	0.9860	0.9418	1.0293	0.9636	1.0903	1.0261
Denmark	1.0205	0.9932	1.0116	1.0444	0.9443	0.9943
France	1.0118	1.0122	1.0099	1.0236	0.9599	1.0072
Germany	1.0236	1.0120	1.0207	1.0326	0.9841	0.9921
Greece	1.0036	0.8902	1.0298	0.9807	0.9762	0.9758
India	1.0111	1.0198	1.0026	1.0157	1.1176	0.9880
Indonesia	0.9977	0.9854	1.0031	0.9917	0.9701	1.0058
Italy	1.0125	0.9501	1.0268	1.0074	0.9694	0.9883
Japan	1.0156	1.0054	1.0166	1.0132	1.0527	1.0163
Mexico	0.9983	0.9797	1.0066	0.9873	1.0883	1.0101
Netherlands	0.9882	0.9230	1.0033	0.9746	0.9330	1.0201
Norway	1.0862	0.9756	1.1151	1.0663	1.0443	0.9214
Romania	1.0120	0.9015	1.0554	0.9733	0.9701	0.9492
Russia	1.0082	0.9052	1.0349	0.9675	1.1087	1.0420
Spain	1.0102	0.9717	1.0150	1.0139	0.9609	1.0000
Sweden	1.0277	0.9649	1.0376	1.0318	0.9556	0.9906
United Kingdom	0.9975	0.9889	1.0110	0.9885	0.9587	1.0098
United States	1.0054	0.9976	1.0016	1.0105	0.9945	1.0132
ROW	0.9829	0.8607	1.0255	0.9417	0.9367	0.9694

Notes: This table reports the $\hat{x} = x^{'}/x$ for the counterfactual changes in tariffs in a monopolistically competitive markets. Welfare, wages and tariffs are in real terms. The global values for economic variables are calculated by weighing each country by its relative GDP. The change in emissions, $Z = \sum_i \hat{Z}_i^{'} / \sum_i Z_i$

ange). Third, these two facts together explain the decline in global emissions. Essentially, eliminating the price signal between dirty and clean industries leads to a modest increase in domestic reallocation in countries with lower emission intensity. These mechanisms contribute to the decline in global emissions.

Impact on Real Income: Accounting for market power generates a heterogeneous impact on real income across countries. It creates sizable losses as high as 4.38 percentage points for China and sizable gains for countries like Australia (0.79 pp), India (0.34 pp), and the United States (0.23 pp). Overall, in the case of imperfect competition, I find that China, Mexico, Indonesia, the United Kingdom, etc., lose real income from the trade pol-

Figure 3

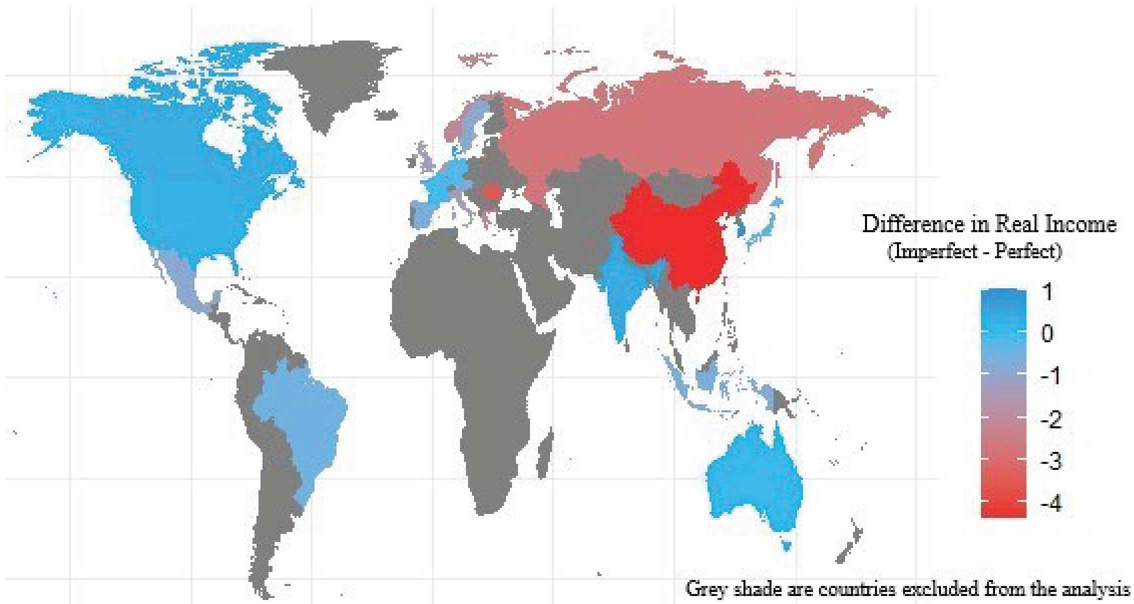


Notes: The left y axis plots the change in domestic profit for carbon-intensive goods, as predicted by the model. The right y axis plots the emission intensity of producing a 1$ worth of output. The x axis plots the net imports of most carbon-intensive goods (mineral fuels, mineral oils, waxes, lime and cement, ores, slag and ash, inorganic chemicals, compound of precious metals, radio-active elements and isotopes) relative to a country's GDP

icy reform. To visualize the difference in the real income in both models, Figure 4 plots the difference in a world map. Countries in red, like China, Russia, Romania, Mexico, Greece, Italy, etc., lose more when we account for market power, while countries like the United States, Canada, India, Australia, etc., gain the most in real income after accounting for the reshuffling of profits. As is evident from the figure, the welfare impact is markedly different in imperfect competition relative to perfect competition.

How important is the quantification of market power in estimating the welfare effects across countries? Figure 5 plots the change in real profits against the change in real income. Countries that lose in real income also experience a decline in their real profits, indicating that market power is an essential channel that explains the decline in real income in these countries under imperfect competition.

Figure 4: Differences in Real Income

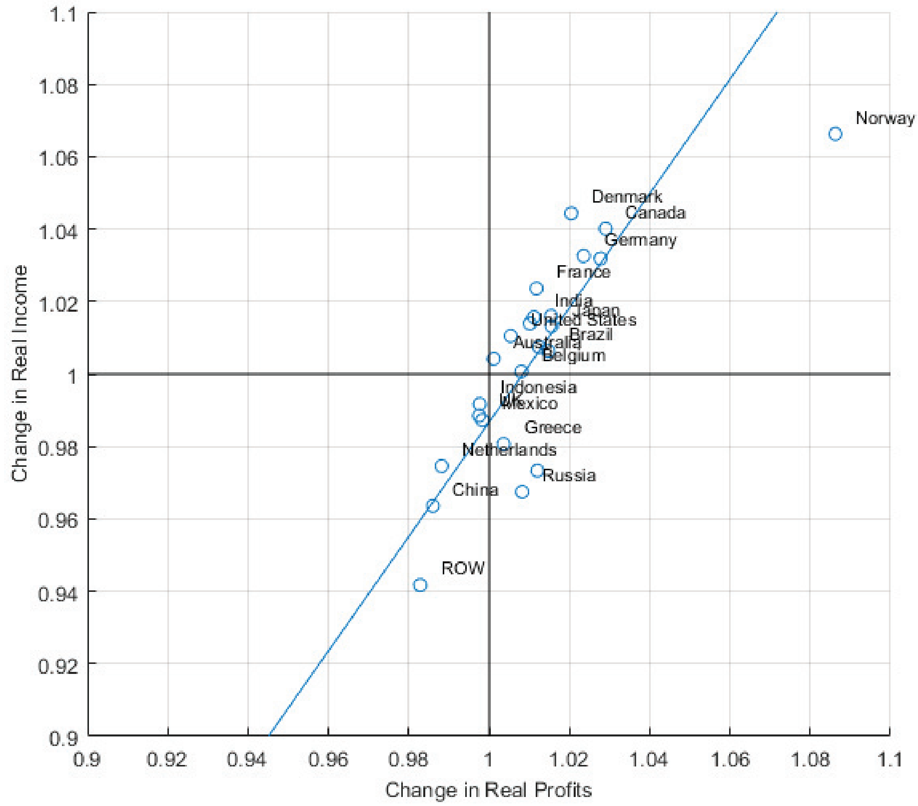


The figure shows the heterogeneity in the welfare estimates across countries. It plots the difference in the real income change between imperfect and perfect competition. Countries in the red lose the most, and countries in the blue gain the most when profits are accounted for. The figure plots the difference in percentage points: *(Change in Real Income in Imperfect Competition - Change in Real Income in Perfect Competition)*100* for each country. The

What explains the heterogeneous effects across countries? The aggregate welfare effects depend on whether a country is a net exporter of the most carbon-intensive goods. As tariffs increase on carbon-intensive sectors, the production reallocation shifts profits to the domestic country. As a result, we expect net importers of the most carbon-intensive goods to gain real profits (if the increase in nominal profits is more significant than the increase in aggregate prices). Figure 6 plots the net imports of the most carbon-intensive goods against the change in real profits predicted by the model. I find that, on average, countries that are net importers of carbon-intensive goods gain real profits from the counterfactual trade policy reform.

Production Reallocation in Developing Countries, including India: Given that the richer countries, on average, have a higher environmental bias and are net importer of carbon-intensive goods, we expect advanced countries to gain the most from the counterfactual trade policy reform. In contrast, as net exporters of high-markup carbon-intensive goods, developing countries would likely lose from the trade policy reform. However, among the developing economies, India gains from the trade policy reform while China, Russia, Mexico, Indonesia, etc, lose from the carbon-based tariff adjustments.

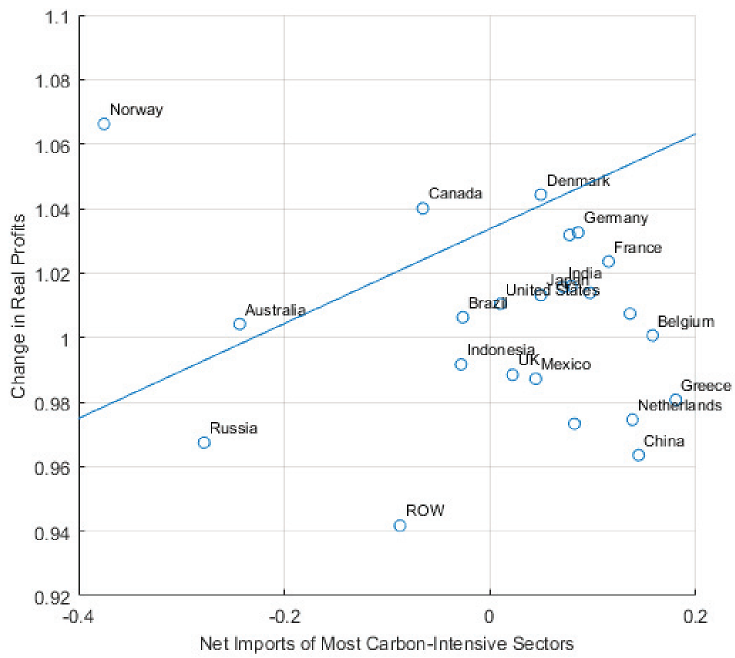
Figure 5: Real Profits vs Real Income



As tariffs increase on high-markup carbon-intensive sectors, the production potentially reallocates to the domestic country, reaping aggregate profits. The extent to which production reallocates to the domestic country explain the magnitude of welfare-enhancing effects. In Figure 7, I compare the size of production reallocation for a select set of developing countries. In the counterfactual, we observe that China, Russia, and Indonesia lose from the counterfactual trade policy reform while India gains. For each country, I plot the change in mean import tariffs for each industry against the change in domestic profits, as predicted by the model. I find that the extent of production reallocation is greatest for India, which potentially explains India's sizable gain from counterfactual trade policy reform relative to other countries.

Further, as documented in Figure 10, the average inverse demand elasticity $1/\sigma_s$ for a country's imports relative to its exports in the carbon-intensive sectors, i.e., the average profit per dollar that countries pay (via imports) minus the average profit per dollar that countries earn (via exports) in carbon-intensive sectors, has a large predictive power for

Figure 6: Net Imports of Carbon-Intensive Goods

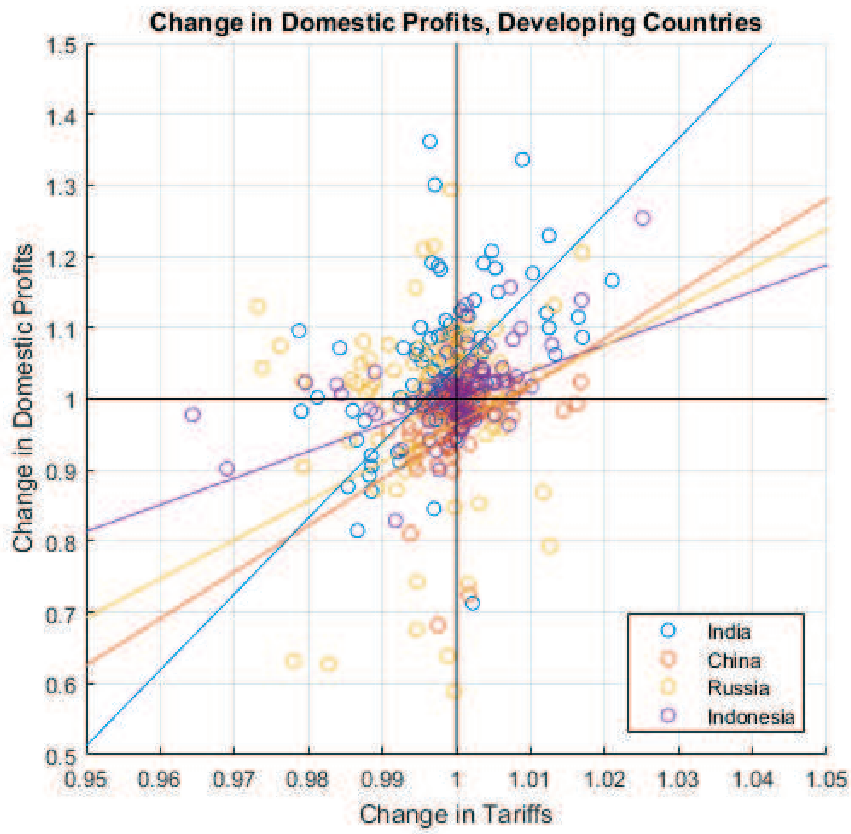


The figure plots the net imports (imports-exports) of the most carbon-intensive sectors (mineral fuels, mineral oils, waxes, lime and cement, ores, slag and ash, inorganic chemicals, compound of precious metals, radio-active elements and isotopes) relative to its GDP against the change in real profits, predicted by the model.

the welfare results. This means that countries that pay more for the imports of carbon-intensive goods relative to earnings in their exports gain real income. This channel is ignored when we are in a perfectly competitive environment.

Country Quadrants: The countries can be classified into four different quadrants after the accounting of market power. The classification depends on whether the country is a net importer of carbon-intensive goods and the relative strength of the markup and price channel. As explained above, a country loses real income if it net exports carbon-intensive goods. This is because its export profits dwindle in response to higher tariffs. These exporting countries include Russia, Brazil, Indonesia, Mexico, etc. Countries that are net importers gain as the domestic reallocation of carbon-intensive sectors, and the resultant gains in profits enhance welfare. These importing countries include the United States, India, and Denmark. A carbon-intensive importer may lose from the policy reform because the increase in import costs spillovers to the aggregate prices for these countries

Figure 7: Production Reallocation in Select countries

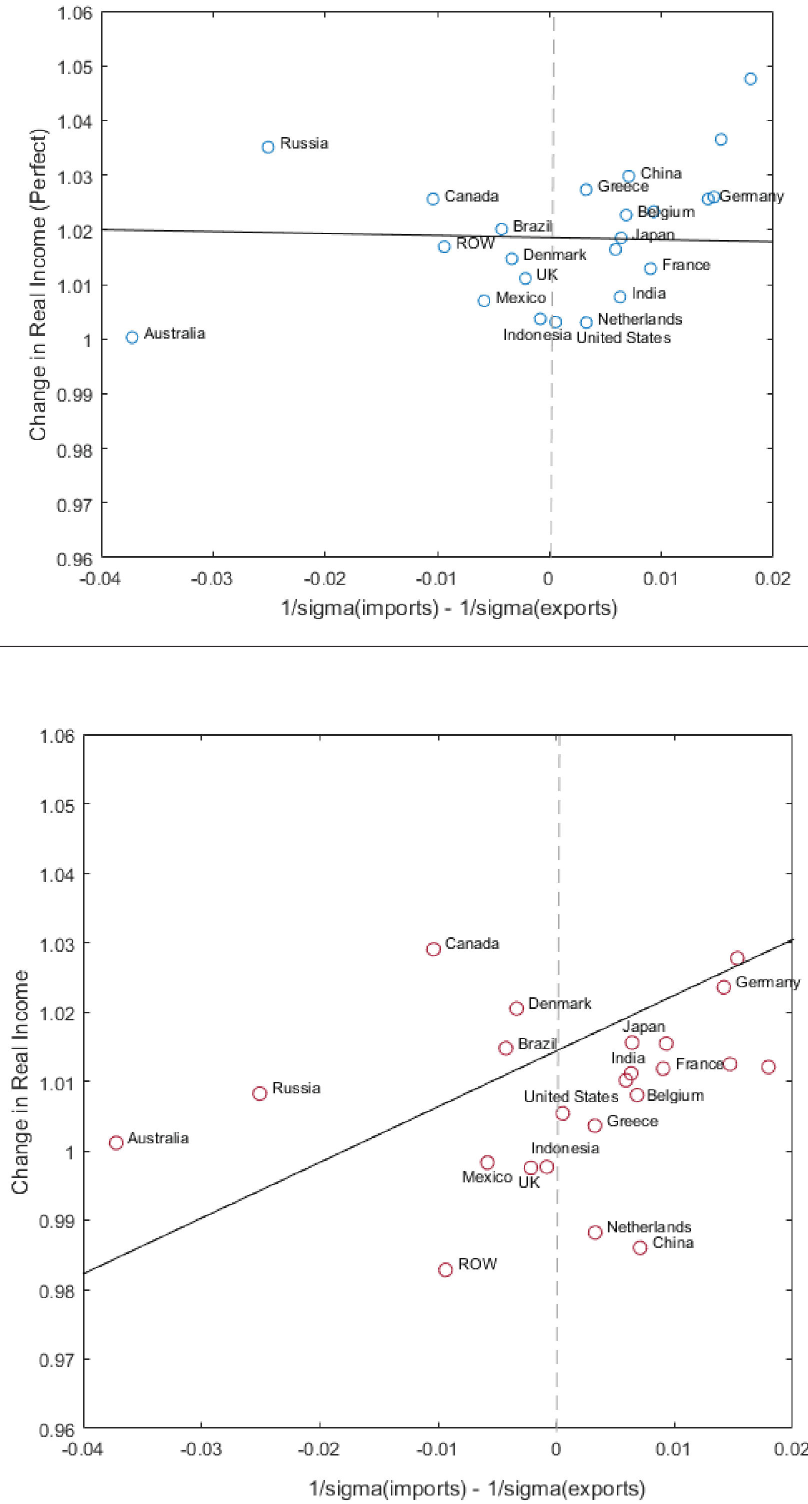


The figure plots the change in mean import tariffs in the counterfactual against the change in the domestic profits, as predicted by the model. For example, If the change in tariffs is greater than 1.01, that means that counterfactual reform increase the mean import tariffs on that industry by 1 percent. If change in domestic profits is 1.1, this means that model predicts a 10% increase in domestic profits.

hampering real income. Conversely, some exporters may gain real income depending on their comparative advantage in cleaner industries (Australia, Canada, etc.)

How does this outcome influence our understanding of carbon-based policy reform? One of the core findings is that not accounting for market power, on average, leads to overestimating the welfare gains of net exporters of carbon-intensive goods. While the counterfactual trade policy reforms lead to heterogeneous effects across countries, the results point that, on average, countries that are a net exporters of carbon-intensive goods lose more. Figure 12 plots the change in the real income against the net imports of most carbon-intensive products (mineral fuels) relative to the GDP for both perfect and imperfect competition. Figure 12 highlights that the welfare changes are relatively lower for net importers of carbon-intensive goods. When we account for markups, we see that the welfare impact is higher. Thus, not accounting for markups underestimates the welfare effect of importers of carbon-intensive goods and overestimates the welfare effect of exporters of carbon-intensive goods.

Figure 10: Change in Real Income against average profits



The horizontal axis measures the average inverse demand elasticity for imports minus that for each country's exports in the carbon-intensive sector. The vertical axis describes the percentage changes in welfare for the counterfactual policy reform under perfect and imperfect competition

Figure 11: Country Quadrants

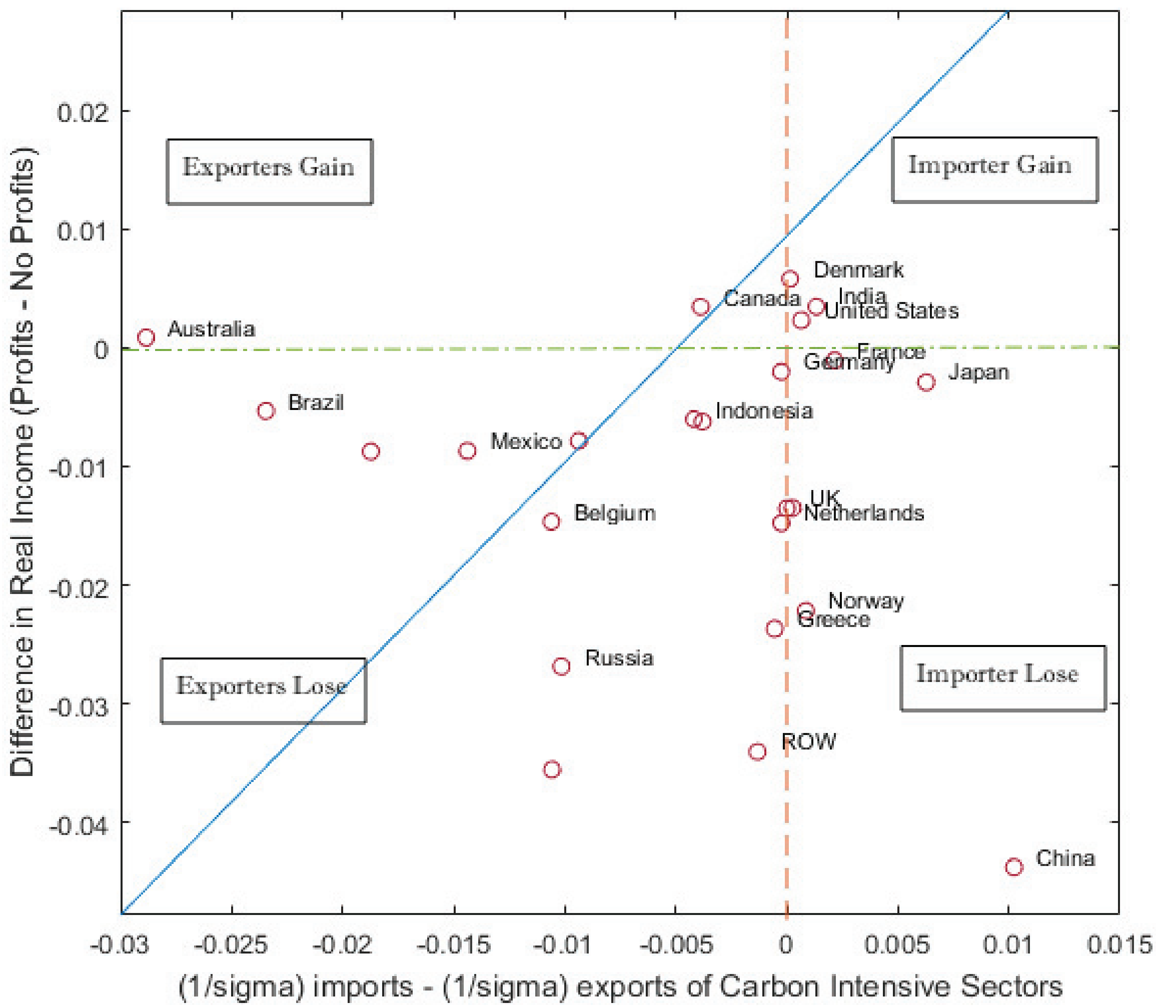
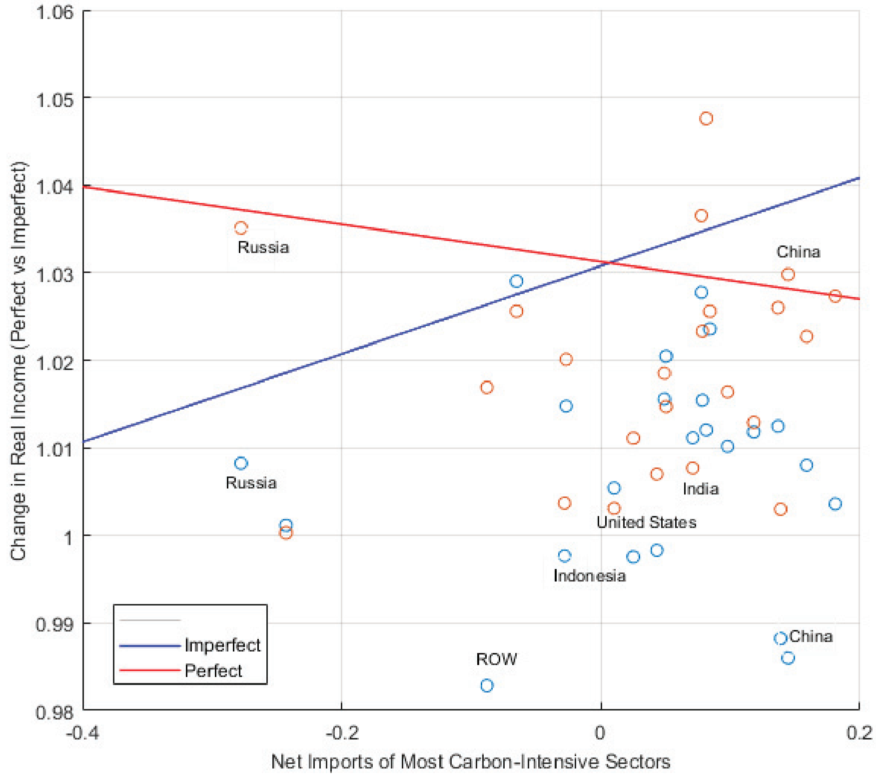


Figure 12: Differences in Real Income: Perfect vs Imperfect



The figure highlights that the welfare changes are relatively lower for net importers of carbon-intensive goods. When we account for markups, we see that the welfare impact is higher. The figure highlights that not accounting for markups underestimates the welfare effect of importers of carbon-intensive goods and overestimates the welfare effect of exporters of carbon-intensive goods.

6 Conclusion

This paper documents novel evidence that suggests a positive correlation between carbon-intensive sectors and market power. In light of this evidence, carbon-based tariffs result in sizable profit-shifting across countries. The paper develops a multi-industry, multi-country structural model of international trade with global input-output linkages. The model accounts for the profit-shifting across countries as tariffs are adjusted to accommodate its carbon content. There are three main takeaways from the paper. First, the findings suggest that accounting for profits increases the effectiveness of trade policy to reduce global emissions. Second, it may result in welfare losses for countries like China, Russia, Indonesia, etc., and may lead to higher welfare gains for many advanced economies, including the United States and Canada, and emerging economies like India. Third, not accounting for market power overestimates the welfare gains for net exporters and underestimates the welfare gains for net importers. The paper discusses the potential forces behind the heterogeneous effects across countries: production reallocation effect and carbon content in its net imports.

References

Rodrigo Adao, Arnaud Costinot, and Dave Donaldson. Nonparametric counterfactual predictions in neoclassical models of international trade. *American Economic Review*, 107(3):633–89, 2017.

Pol Antràs, Davin Chor, Thibault Fally, and Russell Hillberry. Measuring the upstreamness of production and trade flows. *American Economic Review*, 102(3):412–16, 2012.

Paul S Armington. A theory of demand for products distinguished by place of production. *Staff Papers*, 16(1):159–178, 1969.

Mustafa H Babiker. Climate change policy, market structure, and carbon leakage. *Journal of international Economics*, 65(2):421–445, 2005.

Christoph Böhringer, Edward J Balistreri, and Thomas F Rutherford. The role of border carbon adjustment in unilateral climate policy: Overview of an energy modeling forum study (emf 29). *Energy Economics*, 34:S97–S110, 2012.

Christoph Böhringer, Jan Schneider, and Emmanuel Asane-Otoo. Trade in carbon and carbon tariffs. *Environmental and Resource Economics*, 78(4):669–708, 2021.

Frédéric Branger and Philippe Quirion. Would border carbon adjustments prevent carbon leakage and heavy industry competitiveness losses? insights from a meta-analysis of recent economic studies. *Ecological Economics*, 99:29–39, 2014.

Christian Broda and David E Weinstein. Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585, 2006.

Christian Broda, Nuno Limao, and David E Weinstein. Optimal tariffs and market power: the evidence. *American Economic Review*, 98(5):2032–65, 2008.

Lorenzo Caliendo and Fernando Parro. Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, 82(1):1–44, 2015.

Jared C Carbone and Nicholas Rivers. The impacts of unilateral climate policy on competitiveness: evidence from computable general equilibrium models. *Review of Environmental Economics and Policy*, 2017.

Brian R Copeland, Joseph S Shapiro, and M. Scott Taylor. Globalization and the environment. Working Paper 28797, National Bureau of Economic Research, May 2021. URL http://www.nber.org/papers/w28797.

Jan De Loecker and Jan Eeckhout. Global market power. Technical report, National Bureau of Economic Research, 2018.

Jan De Loecker, Jan Eeckhout, and Simon Mongey. Quantifying market power and business dynamism in the macroeconomy. Technical report, National Bureau of Economic Research, 2021.

Robert Dekle, Jonathan Eaton, and Samuel Kortum. Global rebalancing with gravity: Measuring the burden of adjustment. *IMF Staff Papers*, 55(3):511–540, 2008.

Damien Demailly and Philippe Quirion. Co2 abatement, competitiveness and leakage in the european cement industry under the eu ets: grandfathering versus output-based allocation. *Climate Policy*, 6(1):93–113, 2006.

Susanne Droege and Maria Panezi. How to design border carbon adjustments. In *Handbook on Trade Policy and Climate Change*. Edward Elgar Publishing, 2022.

Benjamin Faber and Thibault Fally. Dp12209 firm heterogeneity in consumption baskets: Evidence from home and store scanner data. 2017.

Robert C Feenstra. New product varieties and the measurement of international prices. *The American Economic Review*, pages 157–177, 1994.

Stefan Felder and Thomas F Rutherford. Unilateral co2 reductions and carbon leakage: the consequences of international trade in oil and basic materials. *Journal of Environmental Economics and management*, 25(2):162–176, 1993.

Hamid Firooz and Gunnar Heins. Profit shifting, import and export markups, and the gains from trade. *Import and Export Markups, and the Gains from Trade (November 25, 2020)*, 2020.

Carolyn Fischer and Alan K Fox. Comparing policies to combat emissions leakage: Border carbon adjustments versus rebates. *Journal of Environmental Economics and management*, 64(2):199–216, 2012.

Meredith Fowlie, Michael Greenstone, and Catherine Wolfram. Do energy efficiency investments deliver? evidence from the weatherization assistance program. *The Quarterly Journal of Economics*, 133(3):1597–1644, 2018.

Meredith L Fowlie and Mar Reguant. Mitigating emissions leakage in incomplete carbon markets. *Journal of the Association of Environmental and Resource Economists*, 9(2):307–343, 2022.

Jessie Handbury. Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities. *Econometrica*, 89(6):2679–2715, 2021.

Gunnar Heins. Endogenous vertical differentiation, variety, and the unequal gains from trade. *The Review of Economics and Statistics*, pages 1–45, 2019.

Michael Hoel. Global environmental problems: the effects of unilateral actions taken by one country. *Journal of environmental economics and management*, 20(1):55–70, 1991.

Chang-Tai Hsieh and Peter J Klenow. Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448, 2009.

IPCC. Climate change 2014 synthesis report. *IPCC: Geneva, Szwitzerland*, 2014.

Amit Khandelwal. The long and short (of) quality ladders. *The Review of Economic Studies*, 77(4):1450–1476, 2010.

James R Markusen. International externalities and optimal tax structures. *Journal of international economics*, 5(1):15–29, 1975.

William Nordhaus. Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105(4):1339–1370, 2015.

Ralph Ossa. Trade wars and trade talks with data. *American Economic Review*, 104(12): 4104–46, 2014.

Jean Pierre Ponssard and Neil Walker. Eu emissions trading and the cement sector: a spatial competition analysis. *Climate Policy*, 8(5):467–493, 2008.

Tinus Pulles, Maarten van het Bolscher, Roel Brand, and Antoon Visschedijk. Assessment of global emissions from fuel combustion in the final decades of the 20th century. *TNO Rep. 2007-A-R0132B*, 2007.

Joseph S Shapiro. The environmental bias of trade policy. *The Quarterly Journal of Economics*, 136(2):831–886, 2021.

Anson Soderbery. Estimating import supply and demand elasticities: Analysis and implications. *Journal of International Economics*, 96(1):1–17, 2015.

Konstantin Stadler, Richard Wood, Tatyana Bulavskaya, Carl-Johan Södersten, Moana Simas, Sarah Schmidt, Arkaitz Usubiaga, José Acosta-Fernández, Jeroen Kuenen, Martin Bruckner, et al. Exiobase 3: Developing a time series of detailed environmentally extended multi-regional input-output tables. *Journal of Industrial Ecology*, 22(3):502–515, 2018.

Konstantin Stadler, Richard Wood, Tatyana Bulavskaya, Carl-Johan Södersten, Moana Simas, Sarah Schmidt, Arkaitz Usubiaga, José Acosta-Fernández, Jeroen Kuenen, Martin Bruckner, Stefan Giljum, Stephan Lutter, Stefano Merciai, Jannick H Schmidt, Michaela C Theurl, Christoph Plutzar, Thomas Kastner, Nina Eisenmenger, Karl-Heinz Erb, Arjan Koning, and Arnold Tukker. Exiobase 3, October 2021. URL https://doi.org/10.5281/zenodo.5589597..

Richard Wood, Konstantin Stadler, Tatyana Bulavskaya, Stephan Lutter, Stefan Giljum, Arjan De Koning, Jeroen Kuenen, Helmut Schütz, José Acosta-Fernández, Arkaitz Usubiaga, et al. Global sustainability accounting—developing exiobase for multi-regional footprint analysis. *Sustainability*, 7(1):138–163, 2014.

APPENDIX

A Derivation of Equilibrium Conditions

A.1 Demand

A representative consumer in country j derives utility from consumption of s products, and derives disutility from total global pollution Z

$$U_j = \Pi_s Q_{js}^{\beta_{js}} f(Z) \tag{29}$$

where Q_{js} is the consumption aggregate given by CES aggregate $Q_{js} \equiv (\sum_i q_{ijs}^{\frac{\sigma_{js}-1}{\sigma_{js}}})^{\frac{\sigma_{js}}{\sigma_{js}-1}}$ where elasticity of substitution is allowed to vary by country-sector. β_{js} is the weight given to each sectors by the share of final consumption in country j.

Representative consumer chooses q_{ijs} to maximize Q_{js} subject to the budget constraint

$$\mathcal{L} = (\sum_i q_{ijs}^{\frac{\sigma_{js}-1}{\sigma_{js}}})^{\frac{\sigma_{js}}{\sigma_{js}-1}} + \lambda\Big[X_{js} - \sum_i p_{ijs} q_{ijs}\Big]$$

The solution to the above problem gives us

$$q_{ijs} = \frac{p_{ijs}^{-\sigma_{js}}}{\sum_i p_{ijs}^{1-\sigma_{js}}} X_{js}$$

Thus, the expenditure on goods from country i in country j is

$$X_{ijs} = p_{ijs} q_{ijs} = \underbrace{\frac{p_{ijs}^{1-\sigma_{js}}}{\sum_i p_{ijs}^{1-\sigma_{js}}}}_{L_{ijs}} X_{js}$$

Here, L_{ijs} is the market share of country i in country j. The aggregate price index of good s in country j is $P_{js} = \Big(\sum_i p_{ijs}^{1-\sigma_{js}}\Big)^{\frac{1}{1-\sigma_{js}}}$.

A.2 Firms

Given the demand equation, representative firm in an industry engages in monopolistic competition to choose quantity to maximize profits

$$\max_{q_{ijs}} \pi_{ijs} = p_{ijs}q_{ijs} - c_{ijs}q_{ijs}$$

Maximizing the above problem gives us

$$\frac{\partial \pi_{ijs}}{\partial q_{ijs}} = p_{ijs} + q_{ijs}\frac{\partial p_{ijs}}{\partial q_{ijs}} - c_{ijs} = 0$$

$$\implies p_{ijs}\Big[1 + \underbrace{\frac{q_{ijs}}{p_{ijs}}\frac{\partial p_{ijs}}{\partial q_{ijs}}}_{-1/\sigma_{js}}\Big] = c_{ijs}$$

From the demand equation,

$$\frac{q_{ijs}}{p_{ijs}}\frac{\partial p_{ijs}}{\partial q_{ijs}} = -1/\sigma_{js}$$

This gives us

$$p_{ijs} = \underbrace{\frac{\sigma_{js}}{\sigma_{js} - 1}}_{\mu_{ijs}} c_{ijs}$$

Here, c_{ijs} is the marginal cost which includes the local cost of production $c_{is} = w_i^{1-\alpha_{is}}\Pi_k P_{ik}^{\alpha_{iks}}$ and the trade cost ϕ_{ijs} that includes tariff and non-tariff barriers.

Thus the firm in country i charges a Dixit-Stiglitz markup $\frac{\sigma_{js}}{\sigma_{js}-1}$ over their marginal cost of serving market j. Thus, the profits that the firm in country i earns by serving market j is

$$\pi_{ijs} = \frac{\sigma_{js}}{\sigma_{js} - 1}c_{ijs}q_{ijs} - c_{ijs}q_{ijs} \implies \pi_{ijs} = \frac{1}{\sigma_{js}}X_{ijs} \equiv \frac{1}{\sigma_{js}}L_{ijs}X_{js}$$

B Estimation Procedure

I present step by step description on how to solve the model. We consider a change in policy from t to new policy t', which incorporates carbon-based tariff adjustments, which changes the marginal cost of serving from country i to country j, given by $\hat{\phi}_{ijs}$.

- Step 1: Start with a guess for the changes in the marginal costs for each country-sector $\hat{c}_{is} \equiv (\hat{c}_{11}, \hat{c}_{12}, ...\hat{c}_{1S}, \hat{c}_{21}, \hat{c}_{22}, ..., \hat{c}_{NS})$ and changes in wages $\hat{w}_i \equiv (\hat{w}_1, \hat{w}_2, ...\hat{w}_N)$

- Step 2: For the given trade policy shock (carbon-based tariffs), $p\hat{h}i_{ijs}$ and our guess for $\hat{c}_{is}$ and $\hat{w}_i$, and the market shares in the data λ_{ijs} we use the following equilibrium conditions to calculate $\hat{P}_{js}$ and $\hat{c}_{is}$

$$\hat{P}_{js} = \left[\sum_i \lambda_{ijs}(\hat{c}_{is}\hat{\phi}_{ijs})^{1-\sigma_{js}}\right]^{\frac{1}{1-\sigma_{js}}} \tag{30}$$

$$\hat{c}_{is} = \hat{w}_i^{1-\alpha_{is}} \Pi_k \hat{P}_{ik}^{\alpha_{iks}} \tag{31}$$

$$\hat{\lambda}_{ijs} = \left(\frac{\hat{c}_{is}\hat{\phi}_{ijs}}{\hat{P}_{js}}\right)^{1-\sigma_{js}} \tag{32}$$

- Step 3: Given the changes in market shares $\hat{\lambda}_{ijs}$, sigma elasticities σ_{is} and share of intermediates α_{jsk}, we solve for counterfactual changes in the expenditure $X'_{js} = \hat{X}_{js}X_{js}$

$$\hat{X}_{js}X_{js} = \beta_{js}\Big[\hat{w}_j(w_jL_j) + \sum_s\sum_i \frac{1}{\sigma_{is}}\frac{\hat{\lambda}_{jis}\lambda_{jis}\hat{X}_{is}X_{is}}{1+t'_{jis}} + \sum_s\sum_i \frac{t'_{ijs}}{1+t'_{ijs}}\hat{\lambda}_{ijs}\lambda_{ijs}\hat{X}_{js}X_{js} + D'_j\Big] + \sum_k \alpha_{jsk}\sum_i \frac{\hat{\lambda}_{jik}\lambda_{jik}\hat{X}_{ik}X_{ik}}{(1+t'_{ijk})} \tag{33}$$

The above equation is a system of $N \times S$ non-linear equations in $N \times S$ unknown counterfactual expenditures. If $\phi_{ijs} = \phi'_{ijs}$ then $D'_j = D_j$ and $\hat{w} = 1$ and $X'_{js} = X_{js}$. Re-writing the system of equations in matrix form:

$$\underset{NS\times NS}{\boldsymbol{\Omega}(\hat{\boldsymbol{w}}, \hat{\boldsymbol{c}})} \underset{NS\times 1}{X} = \underset{NS\times 1}{C(\hat{\boldsymbol{w}}, \hat{\boldsymbol{c}})}$$

Here $\mathbf{X}$ is a vector of expenditures for each sector and country and C is a vector of containing the share of labor income and aggregate trade deficits

$$\mathbf{X} = \begin{bmatrix} X'_{11} \\ \vdots \\ X'_{1S} \\ \vdots \\ \vdots \\ X'_{N1} \\ \vdots \\ X'_{NS} \end{bmatrix} \qquad C(\hat{\boldsymbol{w}}, \hat{\boldsymbol{c}}) = \begin{bmatrix} \beta_{11}(\hat{w}_1 w_1 L_1 + D_1) \\ \vdots \\ \beta_{1S}(\hat{w}_1 w_1 L_1 + D_1) \\ \vdots \\ \beta_{N1}(\hat{w}_N w_N L_N + D_N) \\ \vdots \\ \beta_{NS}(\hat{w}_N w_N L_N + D_N) \end{bmatrix}$$

The matrix $\boldsymbol{\Omega}(\hat{\boldsymbol{w}}, \hat{\boldsymbol{c}})$ captures the general equilibrium effects of how carbon-based tariff adjustments in one sector and country can impact the expenditures in all other countries and sectors. $\boldsymbol{\Omega}(\hat{\boldsymbol{w}}, \hat{\boldsymbol{c}})$ is constructed using five matrices: $\boldsymbol{\beta}$, $\boldsymbol{\Pi}$, $\boldsymbol{\zeta}$, $\boldsymbol{\Gamma}$ and $\boldsymbol{\kappa}$

The matrix $\boldsymbol{\beta}$ captures the share of final expenditure of each sector in a country and the matrix.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} \\ \vdots \\ \beta_{1S} \\ \vdots \\ \vdots \\ \beta_{N1} \\ \vdots \\ \beta_{NS} \end{bmatrix}$$

$\boldsymbol{\Pi}$ captures the equilibrium effects of profits that a source sectors earns in a destination country, which is dependant on the expenditures in the destination country.

Here $\Lambda'_{ijs} = \frac{\lambda'_{ijs}}{t'_{ijs}}$

$$
\underset{NS\times NS}{\mathbf{\Pi}} = \begin{bmatrix}
\frac{1}{\sigma_{11}}\Lambda'_{111} & \frac{1}{\sigma_{12}}\Lambda'_{112}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{11S} & \frac{1}{\sigma_{21}}\Lambda'_{121}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{12S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{1N1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{1NS} \\
\frac{1}{\sigma_{11}}\Lambda'_{111} & \frac{1}{\sigma_{12}}\Lambda'_{112}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{11S} & \frac{1}{\sigma_{21}}\Lambda'_{121}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{12S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{1N1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{1NS} \\
\vdots & \ddots & \vdots \\
\frac{1}{\sigma_{11}}\Lambda'_{111} & \frac{1}{\sigma_{12}}\Lambda'_{112}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{11S} & \frac{1}{\sigma_{21}}\Lambda'_{121}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{12S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{1N1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{1NS} \\
\frac{1}{\sigma_{11}}\Lambda'_{211} & \frac{1}{\sigma_{12}}\Lambda'_{212}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{21S} & \frac{1}{\sigma_{21}}\Lambda'_{221}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{22S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{2N1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{2NS} \\
\frac{1}{\sigma_{11}}\Lambda'_{211} & \frac{1}{\sigma_{12}}\Lambda'_{212}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{21S} & \frac{1}{\sigma_{21}}\Lambda'_{221}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{22S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{2N1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{2NS} \\
\vdots & \ddots & \vdots \\
\frac{1}{\sigma_{11}}\Lambda'_{211} & \frac{1}{\sigma_{12}}\Lambda'_{212}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{21S} & \frac{1}{\sigma_{21}}\Lambda'_{221}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{22S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{2N1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{2NS} \\
\vdots & \ddots & \vdots \\
\vdots & \ddots & \vdots \\
\vdots & \ddots & \vdots \\
\frac{1}{\sigma_{11}}\Lambda'_{N11} & \frac{1}{\sigma_{12}}\Lambda'_{N12}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{N1S} & \frac{1}{\sigma_{21}}\Lambda'_{N21}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{N2S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{NN1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{NNS} \\
\frac{1}{\sigma_{11}}\Lambda'_{N11} & \frac{1}{\sigma_{12}}\Lambda'_{N12}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{N1S} & \frac{1}{\sigma_{21}}\Lambda'_{N21}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{N2S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{NN1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{NNS} \\
\vdots & \ddots & \vdots \\
\frac{1}{\sigma_{11}}\Lambda'_{N11} & \frac{1}{\sigma_{12}}\Lambda'_{N12}\cdots\frac{1}{\sigma_{1S}}\Lambda'_{N1S} & \frac{1}{\sigma_{21}}\Lambda'_{N21}\cdots\frac{1}{\sigma_{2S}}\Lambda'_{N2S}\cdots\cdots\frac{1}{\sigma_{N1}}\Lambda'_{NN1}\cdots\frac{1}{\sigma_{NS}}\Lambda'_{NNS}
\end{bmatrix}
$$

The matrix ζ captures the aggregate tariff revenues. Here $\Psi_{js} = \sum_i \frac{t_{ijs}}{1+t_{ijs}}\lambda_{ijs}$

$$\underset{NS\times NS}{\boldsymbol{\zeta}} = \begin{bmatrix} \begin{bmatrix} \Psi_{11} & \Psi_{12}\dots\Psi_{1S} \\ \vdots & \vdots \\ \Psi_{11} & \Psi_{12}\dots\Psi_{1S} \end{bmatrix} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \begin{bmatrix} \Psi_{21} & \Psi_{22}\dots\Psi_{2S} \\ \vdots & \vdots \\ \Psi_{21} & \Psi_{22}\dots\Psi_{2S} \end{bmatrix} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \begin{bmatrix} \ddots \end{bmatrix} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \begin{bmatrix} \Psi_{N1} & \Psi_{N2}\dots\Psi_{NS} \\ \vdots & \vdots \\ \Psi_{N1} & \Psi_{N2}\dots\Psi_{NS} \end{bmatrix} \end{bmatrix}$$

The matrix $\mathbf{\Gamma}$ captures the share of intermediates α_{jsk}, i.e, the share of good s used in the production of good k in country j.

$$\underset{NS\times NS}{\mathbf{\Gamma}} = \begin{bmatrix} \alpha_{111} & \alpha_{112} & \dots & \alpha_{11S} & \alpha_{111} & \alpha_{112} & \dots & \alpha_{11S} & \dots & \dots & \alpha_{111} & \alpha_{112} & \dots & \alpha_{11S} \\ \alpha_{121} & \alpha_{122} & \dots & \alpha_{12S} & \alpha_{121} & \alpha_{122} & \dots & \alpha_{12S} & \dots & \dots & \alpha_{121} & \alpha_{122} & \dots & \alpha_{12S} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ \alpha_{1S1} & \alpha_{1S2} & \dots & \alpha_{1SS} & \alpha_{1S1} & \alpha_{1S2} & \dots & \alpha_{1SS} & \dots & \dots & \alpha_{1S1} & \alpha_{1S2} & \dots & \alpha_{1SS} \\ \alpha_{211} & \alpha_{212} & \dots & \alpha_{21S} & \alpha_{211} & \alpha_{212} & \dots & \alpha_{21S} & \dots & \dots & \alpha_{211} & \alpha_{212} & \dots & \alpha_{21S} \\ \alpha_{221} & \alpha_{222} & \dots & \alpha_{22S} & \alpha_{221} & \alpha_{222} & \dots & \alpha_{22S} & \dots & \dots & \alpha_{221} & \alpha_{222} & \dots & \alpha_{22S} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ \alpha_{2S1} & \alpha_{2S2} & \dots & \alpha_{2SS} & \alpha_{2S1} & \alpha_{2S2} & \dots & \alpha_{2SS} & \dots & \dots & \alpha_{2S1} & \alpha_{2S2} & \dots & \alpha_{2SS} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ \alpha_{N11} & \alpha_{N12} & \dots & \alpha_{N1S} & \alpha_{N11} & \alpha_{N12} & \dots & \alpha_{N1S} & \dots & \dots & \alpha_{N11} & \alpha_{N12} & \dots & \alpha_{N1S} \\ \alpha_{N21} & \alpha_{N22} & \dots & \alpha_{N2S} & \alpha_{N21} & \alpha_{N22} & \dots & \alpha_{N2S} & \dots & \dots & \alpha_{N21} & \alpha_{N22} & \dots & \alpha_{N2S} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \\ \alpha_{NS1} & \alpha_{NS2} & \dots & \alpha_{NSS} & \alpha_{NS1} & \alpha_{NS2} & \dots & \alpha_{NSS} & \dots & \dots & \alpha_{NS1} & \alpha_{NS2} & \dots & \alpha_{NSS} \end{bmatrix}$$

$\boldsymbol{\kappa}$ captures the market share of country i in country j in a sector s

$$\underset{NS\times NS}{\boldsymbol{\kappa}} = \begin{bmatrix} \lambda'_{111} & \lambda'_{112}\dots\lambda'_{11S} & \lambda'_{121}\dots\lambda'_{12S}\dots\dots\lambda'_{1N1}\dots\lambda'_{1NS} \\ \lambda'_{111} & \lambda'_{112}\dots\lambda'_{11S} & \lambda'_{121}\dots\lambda'_{12S}\dots\dots\lambda'_{1N1}\dots\lambda'_{1NS} \\ \vdots & \ddots & \vdots \\ \lambda'_{111} & \lambda'_{112}\dots\lambda'_{11S} & \lambda'_{121}\dots\lambda'_{12S}\dots\dots\lambda'_{1N1}\dots\lambda'_{1NS} \\ \lambda'_{211} & \lambda'_{212}\dots\lambda'_{21S} & \lambda'_{221}\dots\lambda'_{22S}\dots\dots\lambda'_{2N1}\dots\lambda'_{2NS} \\ \lambda'_{211} & \lambda'_{212}\dots\lambda'_{21S} & \lambda'_{221}\dots\lambda'_{22S}\dots\dots\lambda'_{2N1}\dots\lambda'_{2NS} \\ \vdots & \ddots & \vdots \\ \lambda'_{211} & \lambda'_{212}\dots\lambda'_{21S} & \lambda'_{221}\dots\lambda'_{22S}\dots\dots\lambda'_{2N1}\dots\lambda'_{2NS} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ \lambda'_{N11} & \lambda'_{N12}\dots\lambda'_{N1S} & \lambda'_{N21}\dots\lambda'_{N2S}\dots\dots\lambda'_{NN1}\dots\lambda'_{NNS} \\ \lambda'_{N11} & \lambda'_{N12}\dots\lambda'_{N1S} & \lambda'_{N21}\dots\lambda'_{N2S}\dots\dots\lambda'_{NN1}\dots\lambda'_{NNS} \\ \vdots & \ddots & \vdots \\ \lambda'_{N11} & \lambda'_{N12}\dots\lambda'_{N1S} & \lambda'_{N21}\dots\lambda'_{N2S}\dots\dots\lambda'_{NN1}\dots\lambda'_{NNS} \end{bmatrix}$$

The solutions for counterfactual expenditures is given by

$$\underset{NS\times NS}{\boldsymbol{X}} = \underbrace{\left[\underset{NS\times NS}{\boldsymbol{I}} - \Big(\boldsymbol{\beta}\cdot(\boldsymbol{\Pi}+\boldsymbol{\zeta}) + \boldsymbol{\Gamma}\cdot\boldsymbol{\kappa}\Big)\right]^{-1}}_{\Omega(\hat{\boldsymbol{w}})} \cdot \boldsymbol{C}(\hat{\boldsymbol{w}})$$

- Step 4: After solving for the counterfactual expenditures $\boldsymbol{X}(\hat{\boldsymbol{w}}$, substitute the solution with $\hat{\boldsymbol{\lambda}}_{\boldsymbol{ijs}}$ such that the labor market clears in the aggregate and the deficits remain the same in the counterfactual

$$\hat{w}_j w_j L_j = \sum_s \sum_i \frac{\sigma_{is}-1}{\sigma_{is}} \frac{\hat{\lambda}_{jis}\lambda_{jis}\hat{X}_{is}X_{is}}{1+t'_{jis}} - \sum_s \sum_k \alpha_{jks} \sum_i \hat{\lambda}_{ijk}\lambda_{ijk}X_{jk}\hat{X}_{jk} \tag{34}$$

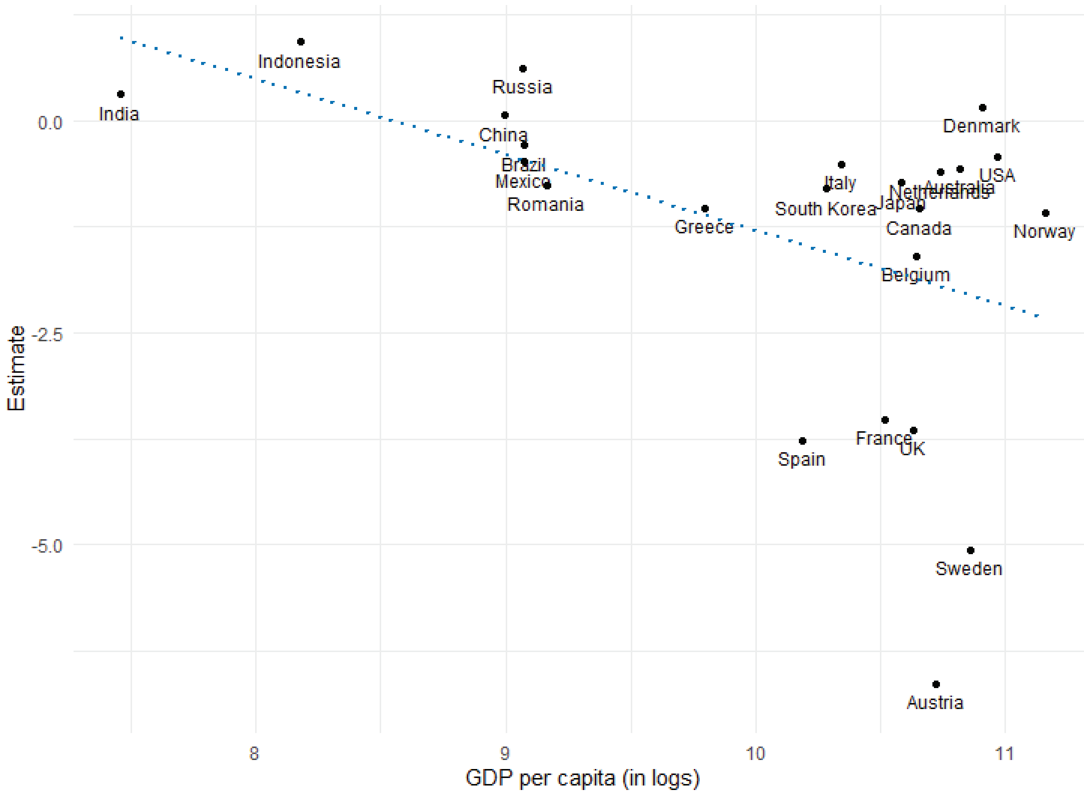
- Step 5: If the above equation does not hold true, we update our guess of $\hat{\boldsymbol{w}}$ until the equilibrium conditions are satisfied.

C Discussion

The exact mechanism that explains these results is not straightforward, as it is hard to assess the extent to which specific endogenous changes in the model influence the aggregate results. I aim to illustrate some essential country characteristics that justify the distributional impact across countries.

Environmental bias of trade policy: Tariff barriers are lower in carbon-intensive sectors for most countries, which is referred to as the environmental bias of trade policy (Shapiro (2021)). However, the extent of this bias differs across countries. The counterfactual tariff adjustments are highest for the countries with the most bias. The country with the lowest trade barrier on carbon-intensive goods experiences the highest increase in import tariffs. This may be welfare-reducing as importing upstream goods become expensive, spilling over to aggregate prices. If there is a significant reallocation of the high-markup carbon-intensive industries, a higher environmental bias might enhance welfare. Contrastingly, trade policy reform lowers tariffs on cleaner industries benefiting the countries with comparative advantage in cleaner sectors. Figure 13 reports the coefficient when a country's weighted mean import tariff is regressed against the carbon embodied in its imports, weighted by the import value. A negative coefficient implies that, on average, sectors with high carbon-embodied in imports have a lower import tariff. Richer countries, on average, have a higher environmental bias. This corroborates that richer economies import upstream carbon-intensive goods from middle-income developing countries, hence imposing a lower import tariff.

Figure 13: Coefficient of environmental bias



The figure reports the coefficient when the weighted mean import tariff of a country is regressed against the weighted carbon-embodied in its imports, weighted by the import value. A negative coefficient implies that on average, sectors with high carbon-embodied in imports have a lower import tariff

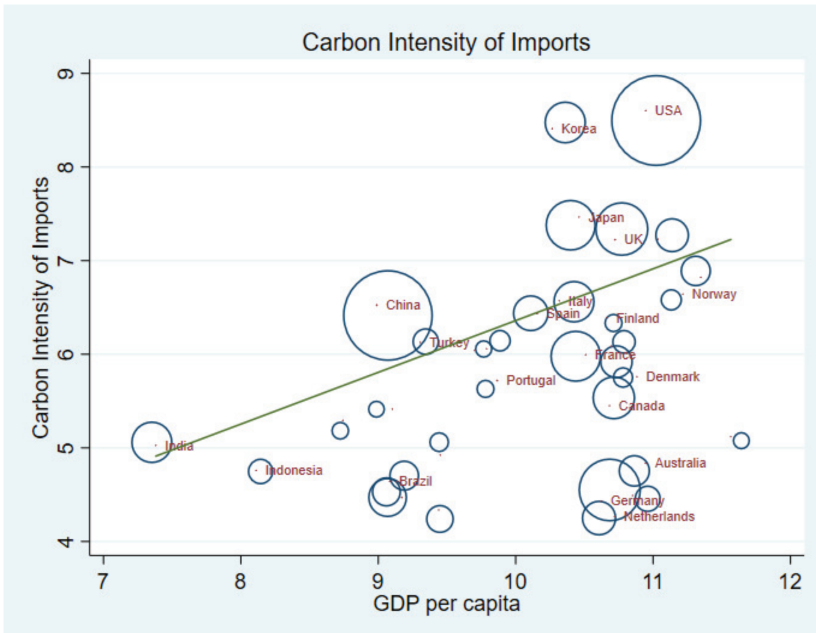
Carbon-Intensity of Trade: The magnitude of the welfare effects of a country in response to the carbon-based tariff adjustments depends on the carbon intensity of its imports vis-a-vis its exports. I define the carbon intensity of imports, ζ_j^{imp} as the total carbon embodied in the imports of a particular country, j

$$\zeta_j^{imp} = \sum_s \theta_{js}^{imp} \sum_{i \neq j} \lambda_{ijs} E_{is} \tag{35}$$

Here, E_{is} is the carbon intensity of producing sector s in country i. The emission intensity is weighted by the market share of country i in country j, given by λ_{ijs}. Each sector s is weighted by their import share in country j, θ_{js}.

Figure 14 plots the carbon intensity of imports of each country with respect to its GDP per capita. Each country is weighted by the size of its total imports relative to global imports. The plot reflects that, on average, the developed countries have a much higher carbon content in their imports than developing countries. This relationship is particularly striking for rich countries like USA, Norway, and Japan and for developing and emerging market economies like Indonesia, China, and Brazil.

Figure 14: Carbon Intensity of Imports



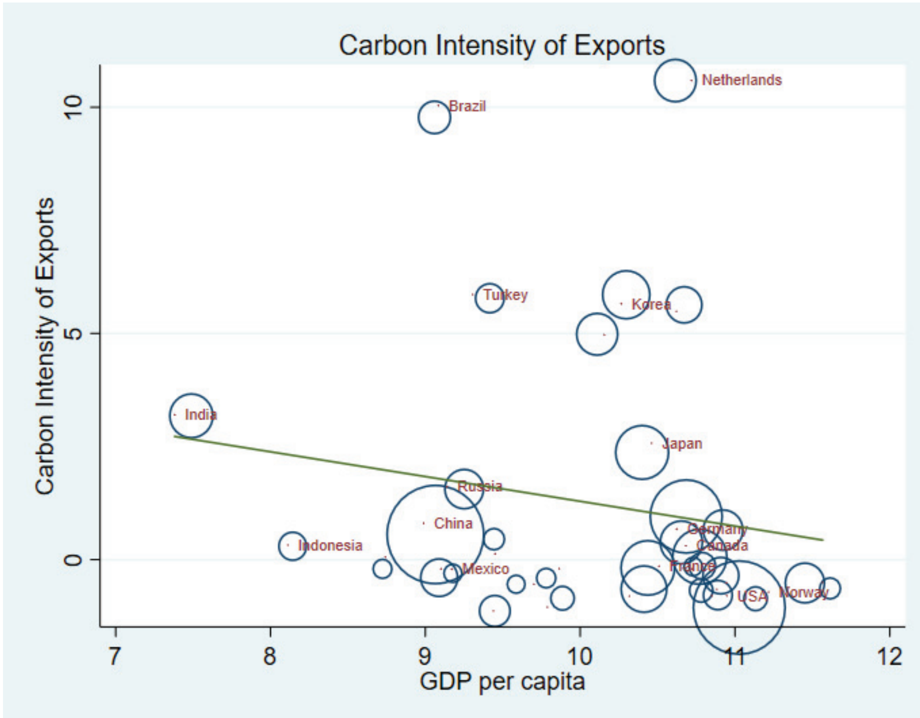
The figure plots the carbon-intensity of imports (in logs) with respect to its GDP per capita (in logs). The circle size represents the share of country's imports in global imports. The fitted line reflects the positive association between the carbon-intensity of imports and GDP per capita, weighted by the country's total imports in global imports

An opposite relationship is observed when we measure the carbon-intensity of exports against the country's GDP per capita. I define the carbon-intensity of exports ζ_j^{exp} as the total carbon-embodied in the exports of a particular country, j

$$\zeta_j^{exp} = \sum_s \kappa_{js} E_{js} \tag{36}$$

Here, κ_{js} is the share of sector s exports in country j's total exports. Figure 15 plots the carbon-intensity of exports of each country relative to its GDP per capita. Each country is weighted by the relative size of its total exports in global exports. We find that, on average, the developing countries export carbon-intensive goods, relative to the developed country.

Figure 15: Carbon Intensity of Exports



Carbon Intensity of Exports. The figure plots the carbon-intensity of exports (in logs) with respect to its GDP per capita (in logs). The circle size represents the share of country's exports in global imports. The fitted line reflects the negative association between the carbon-intensity of exports and GDP per capita, weighted by the country's total exports in global export value.

D Additional Counterfactuals

To ascertain the quantitative importance of the heterogeneity in market power across industries, Table 7 and Table 8 report a counterfactual where σ is the mean value across industries (equal for all industries, $\sigma = 4.75$). The counterfactual eliminates the difference in market power across industries. In the case of perfect competition, global emissions reduce by less than 1%, while all countries still gain real income due to tariff changes. In the case of imperfect competition, emissions increase by almost 6%. We observe that the exporters of carbon-intensive goods do not experience as high loss in real income. Carbon-intensive sectors have higher market power, exacerbating the production reallocation to countries and industries with lower emission intensity. Eliminating the market power shows that carbon-based tariff adjustments are counter-productive and create almost no effect or an increase in global emissions (under different market assumptions). The results highlight the importance of heterogeneity in market power for the policy reform to be effective in reducing emissions.

Table 7: Results: Perfect Competition, sigma = 4.75

	Welfare	Emissions	Wages	Tariff Rev.	Prices
Global Total	1.0240	0.9910	1.0228	1.0624	0.9782
Australia	1.0056	1.0073	1.0055	1.0121	1.0151
Austria	1.0258	1.1113	1.0283	0.9938	0.9750
Belgium	1.0209	1.0900	1.0241	0.9902	0.9585
Brazil	1.0049	1.0564	1.0045	1.0271	0.9809
Canada	1.0374	1.0242	1.0373	1.0437	0.9742
China	1.0316	1.0428	1.0232	1.1649	1.0024
Denmark	1.0188	1.1015	1.0233	0.9585	0.9693
France	1.0177	1.1588	1.0209	0.9809	0.9843
Germany	1.0291	1.1545	1.0317	0.9968	0.9696
Greece	1.0254	1.0351	1.0281	1.0042	0.9446
India	1.0273	1.1164	1.0179	1.3783	1.0095
Indonesia	1.0064	1.0190	1.0061	1.0292	1.0038
Italy	1.0250	1.1213	1.0282	0.9967	0.9604
Japan	1.0178	1.0412	1.0171	1.0936	1.0002
Mexico	1.0080	0.9924	1.0071	1.0550	1.0050
Netherlands	1.0007	0.9733	1.0021	0.9766	0.9945
Norway	1.1640	1.0388	1.1652	1.1511	0.8783
Romania	1.0407	1.1177	1.0427	1.0278	0.9242
Russia	1.0437	0.8751	1.0420	1.1901	1.0264
Spain	1.0185	1.0958	1.0220	0.9780	0.9747
Sweden	1.0443	1.0644	1.0492	0.9944	0.9635
United Kingdom	1.0228	1.0972	1.0217	1.0433	0.9813
United States	1.0038	1.0200	1.0038	1.0123	1.0028
ROW	1.0230	0.8168	1.0284	0.9677	0.9538

Notes: This table reports the $\hat{x} = x^{'}/x$ for the counterfactual changes in tariffs in a perfectly competitive markets. The aggregate real income is decomposed using Equation 26 in terms of real wages and real tariff revenues. Profits are not accounted here. The global values for economic variables are calculated by weighing each country by its relative GDP. The change in emissions, $\hat{Z} = \sum_i Z_i'/\sum_i Z_i$

Table 8: Results: Imperfect Competition, sigma = 4.75

	Welfare	Emissions	Wages	Profits	Tariff Rev.	Prices
Global Total	1.0715	1.0604	1.0248	1.0747	1.0912	0.9415
Australia	0.9938	1.0126	1.0125	0.9642	1.0139	0.9835
Austria	1.0216	1.1218	1.0277	1.0179	0.9958	0.9432
Belgium	1.0260	1.1071	1.0234	1.0341	0.9985	0.9265
Brazil	1.0042	1.0582	1.0050	1.0017	1.0321	0.9428
Canada	1.0288	1.0135	1.0429	1.0050	1.0418	0.9380
China	1.0891	1.1050	1.0275	1.0741	1.3326	0.9577
Denmark	1.0177	1.1116	1.0229	1.0207	0.9605	0.9375
France	1.0118	1.1600	1.0237	0.9985	0.9888	0.9532
Germany	1.0276	1.1677	1.0311	1.0285	0.9981	0.9380
Greece	1.0111	1.0220	1.0325	0.9816	1.0110	0.9171
India	1.0788	1.1673	1.0173	1.1456	1.4195	0.9698
Indonesia	0.9931	1.0263	1.0068	0.9690	1.0065	0.9704
Italy	1.0276	1.1343	1.0285	1.0322	1.0053	0.9298
Japan	1.0301	1.0745	1.0171	1.0400	1.1116	0.9614
Mexico	1.0185	1.0168	1.0127	1.0199	1.1279	0.9658
Netherlands	0.9870	0.9689	1.0062	0.9608	0.9760	0.9626
Norway	1.1558	1.0402	1.1705	1.1400	1.1314	0.8470
Romania	1.0277	1.1105	1.0432	1.0093	1.0234	0.8960
Russia	1.0388	0.8895	1.0443	1.0222	1.1708	0.9914
Spain	1.0142	1.1008	1.0240	1.0071	0.9845	0.9438
Sweden	1.0447	1.0775	1.0480	1.0510	0.9980	0.9313
United Kingdom	1.0245	1.1239	1.0238	1.0226	1.0500	0.9470
United States	1.0112	1.0351	1.0045	1.0198	1.0181	0.9653
ROW	1.0063	0.8212	1.0370	0.9734	0.9350	0.9210

Notes: This table reports the $\hat{x} = x'/x$ for the counterfactual changes in tariffs in a monopolistically competitive markets. Welfare, wages and tariffs are in real terms. The global values for economic variables are calculated by weighing each country by its relative GDP. The change in emissions, $Z = \sum_i \hat{Z}'_i / \sum_i Z_i$

BRANCHING OUT IN BANKING DESERTS: DO CREDIT CONSTRAINED FIRMS GAIN?

Nirupama Kulkarni* **Kanika Mahajan†** **S. K. Ritadhi‡**

January 15, 2024

Does financial deepening increase capital investment by credit-constrained firms? We exploit India's nationwide bank expansion policy in 2005 that incentivized banks to open branches in "underbanked" districts. Using establishment-level data in a regression discontinuity design, we find substantial increases in capital expenditures and credit growth of manufacturing establishments in underbanked districts post intervention. Establishments likely to face credit constraints — small and young establishments, and establishments not owned by publicly listed corporations — drive the effects. The mechanism is increased physical proximity of lenders to small, informationally opaque borrowers. Financial deepening by bringing lenders closer to borrowers can relax credit constraints in economies with imperfect credit markets.

JEL Codes: E22, E58, G21, D24

Keywords: Bank Branch Expansions, Credit Constraints, Small and Micro-Enterprises

*CAFRAL, Reserve Bank of India, India. Email: nirupama.kulkarni@cafral.org.in.

†Ashoka University. Email: kanika.mahajan@ashoka.edu.in.

‡Ashoka University. Email: sk.ritadhi@ashoka.edu.in.

Non-Technical Summary

This research explores the impact of a banking policy in India, known as the Branch Authorisation Policy (BAP) of 2005, on small and medium enterprises (SMEs), particularly those facing credit constraints. The BAP incentivized banks to expand into underbanked regions by making the approval of their annual expansion plans conditional on increased financial intermediation in such areas. The study employs a regression discontinuity design, using an arbitrary threshold to identify underbanked districts, and examines the effects of the policy on credit access and capital investment by SMEs.

The paper focuses on the Indian context, where micro and small enterprises constitute a significant portion of registered manufacturing establishments, with a substantial percentage facing credit constraints. The BAP encouraged banks to open new branches and increase credit to farm activities, rural borrowers, and small and micro-enterprises in underbanked districts. Private banks, in particular, responded disproportionately to the policy, with a notable increase in branches and manufacturing credit.

Using establishment-level data, the study employs a differences-in-differences design to investigate the impact of increased private bank credit on capital investment by manufacturing establishments, especially those in underbanked districts. The findings reveal a significant increase in capital spending for establishments in underbanked districts compared to those in non-underbanked districts. The effects are concentrated among smaller establishments employing under 25 workers, as well as small and young establishments, suggesting a targeted alleviation of credit constraints.

The paper explores potential channels through which the policy affected SMEs, including the physical proximity of banks, increased staffing by banks, and the comparative advantage of lenders in serving small borrowers. The results indicate a reduction in the minimum distance to private banks for smaller establishments in underbanked districts, suggesting improved access. Additionally, the study finds evidence of increased staffing by private banks in these areas, with a focus on officers responsible for loan processing and branch management.

The distributional implications of financial deepening are also examined, showing that the increase in capital investment is primarily driven by smaller and younger establishments, particularly those not publicly listed. The findings highlight the importance of private banks in catering to the financial needs of credit-constrained SMEs and underscore the role of physical proximity in enhancing credit access.

In terms of broader outcomes, the research indicates a positive impact on establishment profitability over the long term, while the effects on output and employment are more nuanced. The study contributes to the literature by offering causal evidence on how bank expansions, especially those led by private banks, can alleviate credit constraints for SMEs and stimulate capital investment.

1 Introduction

Recent research shows that small and medium enterprises (SMEs) constitute a key part of the economy, accounting for over half of formal employment in developing countries (Ayyagari et al., 2014). Large SME presence is also associated with higher per capita GDP growth (Beck et al., 2005a). However, SMEs, along with micro-entrepreneurs, are also more likely to be credit constrained (Banerjee and Duflo, 2014; Beck et al., 2005b; De Mel et al., 2008). Existing studies highlight the role played by high costs of information and monitoring, combined with low collateral, as key barriers restricting these firms' access to credit (Banerjee and Duflo, 2010). There is however limited evidence on policies at-scale which can improve SMEs' credit access through a relaxation of these barriers. We answer this question by exploiting a novel bank expansion policy in India to examine whether greater ease of access to banking institutions affect capital investment by credit-constrained SMEs.

India forms an ideal context to study this question. Micro and small enterprises are ubiquitous, comprising over 80% of registered manufacturing establishments with 74.3% of these facing credit constraints (World Bank Enterprise Data). The financial system, too, is bank-dependent: between 2001 and 2011, less than 15% of registered manufacturing establishments were owned by publicly listed corporations. Against this background, we examine the impacts of the Branch Authorisation Policy (BAP) of 2005, which incentivised banks to expand operations in under-served regions. Specifically, under the BAP, India's central bank – the Reserve Bank of India (RBI) – made its approval of banks' annual expansion plans conditional on their financial intermediation in "underbanked" districts.[1] A district was deemed underbanked if its bank branch density in 2005 was less than the prevailing national average. While no specific targets were provided, banks were encouraged to expand operations in these districts by opening new branches, and increasing credit to farm activities, rural borrowers and small and micro-enterprises.

The use of an arbitrary threshold lends itself to causal identification using a regression discontinuity (RD) design. Importantly, the policy was announced in the latter half of 2005 and used bank branch information from March 2005, and population data from 2001 to classify districts as underbanked. This limited the ability of lenders and local governments to manipulate districts' assignment to underbanked ("treatment") status. We use a McCrary test (McCrary, 2008) to formally rule out selective sorting of districts into

[1]In addition to currency management and monetary policy, the RBI also serves as the banking regulator in India. Districts form the third tier of administration in India, below states.

treatment, and also verify pre-treatment district observables to vary smoothly around the discontinuity threshold. We document an increase in financial intermediation in underbanked districts in response to the policy, driven primarily by private banks. We identify 5 additional private bank branches in underbanked districts in the post-treatment period (a 25% increase), accompanied by a 12% increase in private bank manufacturing credit, with no increase in government-owned banks or credit.[2] The disproportionate response by private banks is not unexpected when considering that over 60% of underbanked districts (and 45% of districts overall) had no private bank presence prior to 2005. These results are broadly consistent with prior findings of Young (2017) and Khanna and Mukherjee (2021).

We then examine whether the expansion in private bank manufacturing credit affected capital investment using data from the Annual Survey of Industries (ASI) – a large nationally representative survey of registered manufacturing establishments in India. The ASI provides annual data on establishment capital, fixed assets, raw materials, output, workers hired, and wages. The ASI also provides district and establishment identifiers between 1998 and 2014. We use these to construct an 11 year establishment-level panel (2001-2011), with treatment status being determined by whether the establishment was located in an underbanked district. Importantly, while the dataset by definition is restricted to registered establishments, the median establishment size in the year 2000 was 15 employees, and two-thirds of establishments hired less than 20 employees, allowing us to identify the distributional effects of banking expansions across small and micro enterprises.

We exploit the time-variation in the onset of the BAP and the panel structure of the establishment data to combine the RD design with a differences-in-difference (DiD) design. Similar to Grembi et al. (2016) and Cingano et al. (2016), our empirical strategy compares establishment outcomes across underbanked and non-underbanked districts, before and after the policy intervention, for establishments located in districts within a narrow neighbourhood of the discontinuity threshold. The inherent discontinuity in districts' assignment to underbanked status ensures the comparability of treatment and control units prior to the policy intervention. We empirically confirm that pre-treatment manufacturing outcomes were statistically indistinguishable across these two sets of districts within a narrow window around the discontinuity threshold.

[2] Government-owned banks dominate the banking landscape in India, accounting for over 60% of the credit disbursed and the majority of bank branches. These banks had also led the initial expansion in branch banking to areas lacking financial infrastructure in the period between 1969 and 1990, after the nationalization of the 14 largest banks (Burgess and Pande, 2005).

Our primary outcome of interest is capital expenditures by manufacturing establishments, defined as investments in plant and machinery. We identify a 6 percentage point increase in capital spending (equivalent to INR 1.8 million) for establishments in underbanked districts, relative to observationally equivalent manufacturing establishments in non-underbanked districts. Our preferred specification uses establishment and industry-year fixed effects, along with establishment and district-level covariates. The use of industry-year fixed effects restricts the comparison of manufacturing investment to establishments in the same broad industry category and year, with the identifying variation stemming from differences in districts' underbanked status. An event-study specification documents the absence of differential pre-treatment trends in capital spending in underbanked districts, but a sharp uptick following the BAP. The increase in manufacturing investment in underbanked districts is also accompanied by higher credit growth. While we cannot distinguish across sources of credit, these findings are consistent with the aggregate increase in manufacturing credit by private banks in underbanked districts.

Our results are robust to alternate functional forms and definitions of the outcome variable. We also verify that our results are unaffected by the choice of bandwidth. We also show robustness to the exclusion of any individual state or industry, making it unlikely that the identified treatment effect is emanating from confounding state or industry-specific policies, the timing of which were also correlated with the BAP. Using a placebo test, we confirm a null effect upon restricting the sample to 2005: the period prior to the BAP.

We next explore the distributional implications of financial deepening and assess whether the increase in manufacturing investment in underbanked districts was driven by enterprises most likely to be credit-constrained. A large literature has shown credit constraints are likely to bind for smaller firms (Beck et al., 2008; Clark et al., 2004; Galindo and Micco, 2007). Consistent with this, we find the increase in capital spending and credit growth to be concentrated amongst establishments employing under 25 workers. Further, motivated by Hadlock and Pierce (2010) and Criscuolo et al. (2019), we find the increase in capital investment and credit growth in underbanked districts to be driven by small *and* young establishments. These establishments witnessed 12 (30) percent higher capital spending (credit growth) following the policy intervention, relative to comparable establishments in control districts. Similar conclusions are drawn if establishment fixed assets are used to measure establishment size. Notably, the positive treatment effects are concentrated amongst establishments satisfying the administrative classification of small-scale industries. Consistent with the inference of Farre-Mensa and Ljungqvist (2016) that being unlisted is a necessary condition for firms to be financially constrained, we also find the positive impact on capital investments and credit growth to be concentrated among establishments owned by corporations which are not publicly listed.

We examine five potential channels that can explain the increase in manufacturing investment in underbanked districts by smaller establishments: namely, the physical proximity of lenders to borrowers; increased staffing by banks to facilitate financial intermediation; comparative advantage of creditors in lending to small borrowers; reductions in the cost of credit due to lender competition; and aggregate demand. In the absence of granular establishment location identifiers in the ASI, we use data from the Economic Census (EC) to explore the first channel. The EC is an enumeration of all business establishments in India, conducted once every 8 years, and provides geographic identifiers at the granular level of Census tracts. As the 2011 population Census provided pincodes for select rural Census tracts, we are able to map a subset of manufacturing establishments in the EC to the nearest private bank branch. Compared to manufacturing establishments in non-underbanked districts, the minimum distance to private banks declined by 5 kilometres (33 percent) in underbanked districts. Importantly, the effects are driven by relatively smaller manufacturing establishments employing between 5 and 25 workers. In contrast, the minimum distance to private banks for establishments employing in excess of 25 workers remained comparable across underbanked and non-underbanked districts.

Aggregating the minimum distance to private banks at the district level, we find the increase in establishments' capital spending and credit growth to be driven by underbanked districts with relatively low (below median) distance to private bank branches. While the flexibility accorded to banks to choose locations within underbanked districts warrants that we interpret the latter finding as suggestive, it supports the overall explanation that the expansion in private bank branches reduced the physical distance between private banks and small manufacturing enterprises, leading to higher credit growth and capital investment. While we cannot ascertain whether physical proximity improved information acquisition or monitoring, these findings nevertheless underline the local nature of credit markets for small borrowers (Aggarwal and Hauswald, 2010; Nguyen, 2019).

Along with the physical proximity of brick-and-mortar branches to manufacturing establishments, we also find the BAP to have affected the staffing decisions of lenders. Using proprietary data on branch-level officers and employees, we identify underbanked districts to have 35 additional private bank officers, and 50 additional private bank employees. Officers are typically responsible for processing loan applications and overall branch management, with the role of loan officers deemed critical to the collection of subjective "soft" information for small borrowers (Aggarwal and Hauswald, 2010; Chen et al., 2022). The fact that two-thirds of the aggregate increase in private bank staffing was driven by bank officers signifies that private lenders were investing in necessary human capital to engage in financial intermediation in underbanked districts.

For the third channel, we exploit proprietary data on banks' lending portfolios. We find the increase in capital investment in treated districts to be driven by districts witnessing entry by a private bank transacting with a relatively high (above median) share of small manufacturing borrowers in the pre-treatment period. As district-specific unobservable factors could have influenced banks' entry decisions, this finding also needs to be interpreted with caution but points to the role of lenders' comparative advantage in transacting with small borrowers in explaining the increase in capital investment.

Using establishment-level data on total interest payments, and proprietary data on branch-level lending rates, we rule out that the expansion in private bank credit came through reductions in the cost of credit due to heightened lender competition. Finally, we draw from Mian and Sufi (2014) and show that the increase in capital investment in underbanked districts was undertaken by establishments in both tradable and non-tradable industries. If financial deepening boosted local demand for manufacturing products and the increase in manufacturing investment arose solely through this channel, we would have expected the treatment effects to be driven exclusively by establishments producing non-tradables.

Lastly, we examine whether the identified increase in capital spending following the BAP affected establishment output, profitability and employment. Surprisingly, we find a muted impact of the BAP on establishment output and profits over the medium term. As capital investments can have a lagged effect on production technologies and profits, we extend our sample to 2014 and identify a positive treatment effect on establishment profitability over the long-run. Manufacturing establishments in underbanked districts were 5 percentage points more likely to report positive profits over the long-term. We also identify a non-significant 7 percent increase in manufacturing output over the long-term. This indicates that the benefits from capital investment in terms of output and profitability might manifest only over the long-term. We however do not find an increase in employment at enterprise level, showing that capital is used to invest in technology that is likely to be labor saving.

In terms of aggregate effects at the district level, we find the BAP to have positively affected the number of establishments operating in underbanked districts, with no corresponding increase in exits. This points to the positive impact of financial deepening on firm entry. Consistent with higher firm entry, we show evidence pointing to an aggregate increase in manufacturing workers in these districts, although the coefficient is imprecisely estimated.

Our study contributes to four broad stands of literature. First, we directly extend the literature studying the impact of financial development on credit allocation. Prior work by Jayaratne and Strahan (1996), Bai et al. (2018), Bruhn and Love (2014), and Fonseca and Matray (2022) documents the impact of banking expansions on economic growth, employment, capital, poverty and entrepreneurship. While Jayaratne and Strahan (1996) and Bai et al. (2018) find improvements in lending quality in response to increased competition, Bruhn and Love (2014) shows how access to credit histories from existing retail partners facilitates banks' selection of creditworthy borrowers. Fonseca and Matray (2022) documents physical distance to banks, proxied by lack of banking services in the pre-treatment period, as the key friction limiting firm expansion and consequently labor demand. In contrast, our paper is the first to causally show that increasing physical proximity to banking institutions, particularly private banks, can spur credit uptake and investment by manufacturing firms. In general, this literature has largely focused on expansion in government banking. Private banks are more aligned with market forces, have profit-maximizing motives, superior corporate governance, and are less susceptible to political capture. Consequently, our paper shows that private banks too can be appropriately incentivized to engage in financial intermediation in under-served areas. Additionally, we show that private bank-led financial deepening alleviated credit constraints through higher credit disbursement to small and young enterprises. This is unaccompanied by higher costs of credit, or a deterioration in credit quality: two typical downsides from excessive risk-taking by private lenders.

Importantly, unlike Bruhn and Love (2014), private banks in our context had negligible operations in underbanked areas prior to the BAP. This further underlines how physical proximity can facilitate financial intermediation to borrowers with high costs of screening and monitoring, even in the absence of long-standing information networks. Here, as we show, investments in human capital by lenders to acquire information through the hiring of officers plays an important role. Consequently, our paper relates to the literature examining how monitoring and screening costs, proxied by distance to bank branches, affect financial intermediation (Greenwood et al., 2010; Ji et al., 2021).[3] Second, the emphasis on physical distance between lenders and small informationally opaque borrowers also re-

[3]Our focus on private banks and credit constraints also serves as the point of distinction from Chakraborty et al. (2021) who study a later reform aimed at expanding bank branches in relative smaller urban centres to show how increased lender competition disciplines government-owned banks.

lates our paper to the literature documenting the local nature of credit markets (Aggarwal and Hauswald, 2010; Berger and Udell, 1995; Chen et al., 2015; Nguyen, 2019; Petersen and Rajan, 1994). Unlike Nguyen (2019), which studies how bank closures affect small business loans, our paper identifies the impact of banking expansions. While we cannot distinguish whether increased physical proximity improves lenders' information collection or ex-post monitoring, our findings resonate with Petersen and Rajan (2002), who suggest a lowering of credit market frictions when borrowers' physical distance from financial institutions reduce. Recent work by Aggarwal and Hauswald (2010), Gilje et al. (2016), Hollander and Verriest (2016), Liberti and Petersen (2019), Chen et al. (2022) and Chernenko et al. (2022) also show the cost of information acquisition to be increasing in distance, and impeding resource allocation for financially constrained firms (Giroud and Mueller, 2015).

Third, with the increase in capital investment and credit growth being driven by small and young establishments, and those not publicly listed, our paper also adds to the literature on how financial deepening can alleviate financial constraints. Related work by Fafchamps and Schündeln (2013) shows that small firms in high potential growth sectors gained from adequate local banking infrastructure, while Larrain and Stumpner (2017) finds that capital account liberalization in Eastern Europe increased investments for firms in sectors dependent on external finance.[4] Our paper differs from these by offering direct causal evidence on how bank expansions affect capital spending by credit-constrained establishments, and uncovering reduction in physical distance to lenders as the principal mechanism.

Lastly, in the Indian context, our paper joins a growing body of work studying the economic impacts of the BAP. We extend the work of Young (2017) by focusing on the distributional aspects of financial deepening, and also identify the mechanisms through which financial deepening affects manufacturing investment.[5] Other studies on financial deepening in India by Burgess and Pande (2005), Kochar (2011) and Gupta and Dehejia (2021) focus on the massive state-directed push by government-owned banks in the 1980s and consider outcomes such as poverty, income distribution and labor reallocation from self-

[4]Qin and Kong (2022) show that increased loan access through branches of a single bank which specialized in small and medium income businesses in China increased entrepreneurship in regions with a larger number of branches. Since the bank in question specialized in lending to smaller firms, the evidence cannot be generalized to general bank expansions.

[5]Young (2017) showed that the BAP-induced increase in private bank branches positively affected farm credit and nightlights-based measures of economic activity. Khanna and Mukherjee (2021) also exploits the same policy intervention to show how bank branches served as a coping mechanism when districts faced an aggregate negative shock to cash supply. Cramer (2022) uses the same policy intervention to show the impact of financial deepening on health outcomes.

employment to salaried work. Our study distinguishes itself through its focus on private bank expansions, its impact on manufacturing investment for credit constrained firms, and highlighting physical distance as the key friction in accessing formal credit markets.

The remainder of the paper is organized as follows: Section 2 outlines the Branch Authorisation Policy; Section 3 formally details our empirical strategy and data sources; Section 4 presents our key findings; Section 5 explores potential mechanisms explaining our results; and Section 6 reports aggregate effects of financial deepening.

2 Background: Policy Intervention

Prior to 1991, India's banking industry was dominated by government-owned banks. In the era of "social banking" between 1970 and 1991, the central bank oversaw a massive expansion of government-owned bank branches, which has been the subject of multiple studies (Burgess and Pande, 2005; Gupta and Dehejia, 2021; Kochar, 2011).[6] With the onset of economic liberalization in 1991, the central bank formally abandoned the rule-based branching policy in 1993 and allowed commercial banks to open branches as determined by market forces. In 2005, the RBI revisited its stance and unveiled the Branch Authorisation Policy (BAP) in an effort to expand financial infrastructure in under-served regions.

2.1 Branch Authorisation Policy, 2005

The "liberalised" Branch Authorisation Policy of 2005 simplified the process for opening new branches, but also accorded greater weightage to branches opened in hitherto "underbanked" areas (RBI, 2005). Unlike the social banking era, the policy did not provide an explicit rule on the number of new branches to be opened; neither was it targeted to establish brick-and-mortar branches in banking deserts. Instead, the RBI used an incentive-based design, where banks were directed to submit annual branch expansion plans, detailing the number of bank branches it intended to open in the coming year. For every branch proposed, the bank had to furnish details on the specific location, the number of bank branches presently operating in that area, and the volume of deposits and loans expected in the first year of operation. The RBI committed to meeting bank officials within 4 weeks of the bank's submission of the annual expansion plan. The final approval

[6]The branch expansion was enabled by large-scale nationalization of existing private banks, such as the one in 1969. Under social banking, banks were required to open 4 additional branches in "underbanked" areas, for every branch opened in a "banked" area.

of the plan would occur after the meeting with central bank regulators, and the bank had one year in which to establish the proposed branches (RBI, 2005).

The implicit nudge provided in the BAP was that the overall expansion plan of banks would be favorably received, conditional on banks expanding their operations in underbanked areas. Specifically, the policy document noted that while evaluating proposals for bank expansion, weightage would be accorded to "the nature and scope of banking facilities provided by banks to common persons, particularly in underbanked areas (districts), actual credit flow to the priority sector, pricing of products and overall efforts for promoting financial inclusion." (RBI, 2007) Furthermore, the RBI noted that it would assess "compliance with not only the letter of the regulations but also whether the bank's activities are in compliance with the spirit and underlying principles of the regulations." (RBI, 2007) [7]

To classify regions as "underbanked", the RBI followed a simple rule based on districts' bank branch density in 2005. For each district, the RBI computed branches per capita using the district's population from the 2001 Census, and the number of commercial bank branches in operation on March 31, 2005. This was compared to "national" branches per capita for the country; districts were classified as "underbanked" if their branches per capita was smaller than the national branches per capita. Formally,

where $BranchPC_d$ is the number of bank branches in district d, scaled by district population in millions and $\overline{BranchPC}$ is the national bank branches per million persons. Using this rule, the RBI published in September 2005 a list of 386 "underbanked" districts, which remained fixed over subsequent years.[8]

As data prior to 2005 was used to determine whether a district was underbanked, districts could not plausibly select into "underbanked" status. There is no anecdotal evidence of advance intimation of the policy, which could have influenced banks to open branches prior to 2005 in underbanked districts. Empirically, Figure 1 confirms using the McCrary test (McCrary, 2008) the absence of any selective sorting of districts into treatment and control status around the national average threshold. This allows us to use the national

[7]For instance, annual branch expansion plans of banks now had to be accompanied by a statement depicting the distribution of operational bank branches in underbanked districts, as well as semi-urban and rural centres (RBI, 2007).

[8]While the rule for classifying districts as underbanked was followed for the vast majority of districts, the RBI amended this rule for a total of 9 districts in 2006. Thus, 6 districts were classified as underbanked, even though their branch density exceeded the national average, while 3 districts were not classified as underbanked, even though their branch density fell below the national average. For addition details, see RBI's master circular on branch authorisation, issued on August 3, 2005 (available at https://rbi.org.in/Scripts/BS_CircularIndexDisplay.aspx?Id=2408).

average branch density – $\overline{BranchPC}$ – as an arbitrary threshold in the spirit of a RD design to causally identify the impact of the branch expansion policy.

With $\overline{BranchPC}$ serving as the discontinuity threshold for a district's underbanked status, the running variable of interest – $Runvar_d$ – is defined as:

$$Runvar_d = BranchPC_d - \overline{BranchPC} \quad (1)$$

Districts are underbanked if $Runvar_d < 0$ or the district's bank branch density in 2005 was less than the national average. Figure 2 shows the distribution of $Runvar_d$, with a significant mass of districts around the threshold 0. For instance, 304 districts (211 underbanked and 93 non-underbanked) fell within a bandwidth of 20 around the discontinuity threshold. Reducing the bandwidth to 15 (10) results in 231 (156) districts lying within the neighbourhood of 0. The presence of a large set of districts around the discontinuity threshold provides both statistical power to detect treatment effects, and also limits concerns regarding the external validity of our findings. Appendix Figure A1 shows that underbanked districts were geographically dispersed across the country.

Finally, while the policy applied to both state-owned and private banks, we expect private banks to be disproportionately affected owing to their limited operations in districts classified as underbanked. At the beginning of 2005, 45% of Indian districts lacked a private bank branch, while 60% of "underbanked" districts had no operations by private banks. On the contrary, government-owned banks had branches across all districts. This made the BAP binding on private banks, which were also mandated by the policy to have at least 25% of their branches in semi-urban or rural centres. Aggregate trends in Appendix Figure A2 point to higher private bank branch openings after adoption of the BAP: the median private bank branch density increased from 0.66 branches to 2.85 branches (per million population) between 2005 and 2010. The fraction of districts without a private bank branch also fell below 20% by 2010.

3 Data and Empirical Strategy

This section describes the primary datasets used in the paper and the empirical strategy to causally identify the impact of the bank branch expansion on manufacturing investment.

3.1 Data

We use data from three major sources: the Annual Survey of Industries, the Economic Census, and the Basic Statistical Returns.

3.1.1 Manufacturing Establishment Data

We use data from the Annual Survey of Industries (ASI) to identify the impact of bank branch expansion under the BAP on manufacturing investment. The ASI is a nationally representative survey undertaken every year by the National Sample Survey Organisation (NSSO), covering registered manufacturing enterprises in India. The unit of observation is the manufacturing establishment (and not the firm). The ASI has two components: a census component whereby establishments employing over 100 workers are covered every year, and a survey component, whereby the ASI uses a stratified random sample each year to survey establishments employing less than 100 workers.[9] The ASI by design excludes enterprises not registered under either the Factories Act, 1948 or the Companies Act, 1956, making it a dataset pertaining exclusively to registered enterprises.[10]

The ASI provides rich data on fixed capital, plant and machinery, raw materials, output, workers employed and wages paid. Additional information on loans and interest payments are also provided, although there is no information on the source of credit. The ASI includes establishment identifiers for all years but provides district identifiers only between 1998 and 2007. District identifiers allow us to determine whether an enterprise was located in an underbanked ("treated") district. We use the procedure outlined in Martin et al. (2017) to construct our primary sample: a decade long unbalanced establishment-level panel between 2001 and 2011, covering over 10,000 unique manufacturing establishments. As the BAP was initiated in 2005, this provides us with 4 years of data prior to the intervention, and 6 years post-intervention. We use data till 2011 for the main analyses since the RBI introduced a new branching policy that year, encouraging banks to open branches in relatively small urban centres.[11]

Our primary outcome of interest is capital expenditures, defined as the difference between the closing and the opening value of an establishment's plant and machinery in a year, scaled by the average value of the establishment's plant and machinery during the year. Specifically, for establishment i in year t, we define capital expenditures as:

[9] Such establishments are typically surveyed once every 3 years.
[10] These two legal statutes governs the operations of registered enterprises in India.
[11] See Chakraborty et al. (2021) for details on the new policy.

$$Capex_{it} = \frac{Plant_{i,t} - Plant_{i,t-1}}{0.5 \times Plant_{i,t-1} + 0.5 \times Plant_{i,t}} \quad (2)$$

where $Plant_{i,t}$ is establishment i's value of plant and machinery in period t, net of depreciation. The principal advantage of formulating capital spending in this manner is that the variable is bounded between -2 and 2, reducing sensitivity to outliers (Berton et al., 2018). We opt to focus on plant and machinery as our primary measure of manufacturing capital as it captures establishments' productive assets. However, we confirm robustness to using aggregate fixed assets (land and buildings, in addition to plant and machinery) as our measure of capital. In addition to capital investment, we also consider other outcomes such as credit growth, output, and workers employed. All growth variables are defined as per equation (2). All nominal (INR) values are deflated to 2011 values using a wholesale price index deflator for manufacturing commodities and top-coded at the 1% level to limit the influence of outliers.

Appendix Table B1 presents summary statistics from the ASI for our primary sample: establishments situated in districts located within a narrow window around the discontinuity threshold. Similar to most firm-level data, Appendix Table B1 documents a large right tail for a number of variables of interest. The average establishment has plant and machinery (fixed assets), net of depreciation, worth INR 29 (40) million, but the median establishment machinery (fixed assets) is INR 1.2 (3.2) million. Similarly, while the mean establishment size in terms of hired workers is 90, the median establishment size is 20. Two-thirds of the establishments satisfied the administrative criterion of "micro" enterprises, while another quarter qualified as "small".[12] Over 80% of the establishments satisfied the definition of small-scale industries, making them eligible for subsidized bank credit.[13] The median establishment age is 14 years and 14 percent of establishments were owned by publicly listed corporations.

Average annual capital expenditures, net of depreciation, during this period equaled -.002, – signifying a net reduction in plant and machinery values in an year. The 75th percentile value is 0.03. This implies that the median establishment did not engage in any capital spending during the year to offset the depreciation in the value of plant and machinery. Capital expenditures are inherently lumpy, and we define a binary variable

[12]We use administrative definitions for classifying establishments as micro, small, medium and large enterprises. In 2005, establishments with plant and machinery worth less than INR 2.5 million were classified as micro-enterprises; between INR 2.5 and 5 million as small enterprises; between INR 5 and 10 as medium enterprises; and exceeding INR 10 million as large enterprises. We use pre-treatment maximum values of establishment plant and machinery to classify enterprises into these 4 categories.

[13]Small-scale enterprises are those whose investment in plant and machinery do not exceed INR 10 million.

$AnyCapex_{it}$ to equal one if the closing value of plant and machinery exceeded the opening value, or $Capex_{i,t} > Capex_{i,t-1}$. Attesting to the inherently lumpy nature of capital investments, we see that only a fourth (third) of the establishments engaged in any positive investment in plant and machinery (fixed assets), net of depreciation, in a given year.

While the ASI does not record the source of credit, it collects data on outstanding loans for establishments. Based on closing and opening values of outstanding establishment loans, we find annual loan growth to be 4 percent or INR 1.2 million.[14] The median establishment however saw no loan growth. Along the extensive margin, 38% of establishments had closing values of outstanding loans in excess of opening values, reflecting a net increase in outstanding credit. Despite being registered establishments, almost a fourth of the establishments had no outstanding credit during the year.[15] Entry into credit markets in a year is also limited – 2.5% of establishments reported having no outstanding credit at the beginning of the accounting period, but a positive loan balance at the end of the accouting year (classified as new loans). Using interest payments undertaken by establishments during the year, the cost of credit for the median establishment is 14%, while the average cost of credit is 24%.[16]

The ASI separately reports three categories of workers: hired workers, contractual workers, and supervisors. The former involve workers engaged in manufacturing tasks and directly hired by the establishment. These workers are eligible for social security benefits, mandatory bonus, and protected from dismissal in large firms by stringent labor laws. Contract workers are manufacturing workers hired on contractual terms by the establishment, and ineligible for the benefits and job security available to hired workers. Finally, supervisors are employees not directly involved in manufacturing tasks but responsible for overall management and supervision. We use the number of supervisors (as a share of total employees) as a measure of the establishment's managerial capital, and the number of hired workers on the establishment's payroll to measure establishment size. The median (average) establishment had 20 (89) hired workers, 0 (27) contract workers, and 2 (10) supervisors. The ratio of supervisors to total employees, used as a measure of managerial capital, was 0.11, or about 1 supervisor for every 10 employees.

[14] Average annual outstanding loans in this period equaled INR 27.9 million.

[15] We classify an establishment to have no outstanding credit if it reports no outstanding loans for both the opening and closing values in a year.

[16] We use the ASI data on annual interest expenses and scale it by opening value of outstanding loans to impute the rate of interest.

3.1.2 Economic Census

The Economic Census (EC) is an enumeration of all business establishments operating in India. It is typically conducted once every 8 years and covers all sectors of the economy. The two most recent ECs pertain to the years 2013 and 2005 respectively.[17] While there is very limited information on establishment performance and characteristics, the EC does provide granular geographic identifiers for establishments, in addition to establishment size. The EC geographic identifiers can be mapped to Census tracts using the SHRUGS database constructed by Asher et al. (2021). As select rural Census tracts were also mapped to pincodes in the 2011 decennial Census, the EC permits the mapping of establishments to pincodes. We use this information to map rural EC establishments to the nearest private bank branch to test whether the BAP affected manufacturing establishments' distance to the nearest private bank branch.

3.1.3 Basic Statistical Returns

We use data from the Basic Statistical Returns (BSR), hosted by the RBI, to assess the impact of the BAP on bank branches, deposits, loans, interest rates, non-performing assets, and staffing decisions. The publicly available BSR data provides annual aggregates of district-level deposits and loans for commercial banks. The data is disaggregated by bank ownership and sectoral allocation of credit, allowing us to compare branch openings and credit disbursement across underbanked and non-underbanked districts, and also by bank group. To gauge new branch openings, we use publicly available information on commercial bank branch opening dates between 2001 and 2011. We also aggregate proprietary branch-level data on interest rates, employment, and non-performing loans to the district level. Results using the BSR data are detailed in Appendix A.

3.2 Empirical Strategy

The use of an arbitrary threshold – national average bank branch density – to classify districts as "underbanked" lends itself to causal identification using a RD design. If districts are unable to manipulate their treatment status (evident from Figure 1), and pre-treatment district covariates vary smoothly on either side of the threshold, we can causally identify the impact of the branch expansion policy on manufacturing investment, using estab-

[17] The data for the EC conducted in 2020 has not yet been released.

lishments in non-underbanked districts as a valid counterfactual for "treated" establishments in underbanked districts. We also exploit the inherent time-variation in the onset of the BAP to combine the RD design with a DiD design (effectively, a "differences-in-discontinuity" design postulated in Grembi et al. (2016)). The DiD framework allows us to exploit the panel structure of the establishment-level data and condition on establishment fixed effects. This allows us to flexibly partial out time-invariant unobserved establishment characteristics such as accounting standards, or local networks, which can affect establishments' access to credit. While it is plausible that such unobservable establishment characteristics would be "balanced" across underbanked and non-underbanked districts within a narrow window of the threshold, using a differences-in-discontinuity framework enables us to explicitly account for them using establishment fixed effects.

Akin to a DiD design, our empirical strategy compares establishment outcomes before and after the policy intervention across underbanked and non-underbanked districts. In the spirit of a RD design, we also restrict the comparison of establishment outcomes to establishments operating in districts within a narrow window of the discontinuity threshold. Our primary estimating equation takes the form:

$$Y_{ijdt} = \alpha_i + \delta_{jt} + \beta Underbanked_d \times Post_t + f(Runvar_d) + \gamma \mathbf{X}_{ijdt} + \epsilon_{ijdt} \tag{3}$$

where Y is the outcome of interest for establishment *i*, operating in industry *j*, located in district *d*, and year *t*. α_i denotes establishment fixed effects, while δ_{jt} denotes industry-year fixed effects, corresponding to the 2-digit industry (*j*) in which the establishment operates. This restricts the comparison of establishment outcomes to the same broad industry category and year. **X** includes establishment and district time-varying covariates. We control for a quadratic in firm age, and flexibly account for establishment size by including five bins of establishment size, interacted with year dummies.[18] Pre-treatment district demographic and economic covariates from 2004, interacted with a post-treatment indicator, are also included.[19]

$Underbanked_d$ is a dummy equaling 1 if establishment *i* is located in a district classified

[18]Establishment size is measured using the average number of workers hired between 2001 and 2005. Five establishment size bins are used: less than 10 workers, 10-25 workers, 25-50 workers, 50-100 workers and more than 100 workers.

[19]The covariates considered are population density; labor force participation and unemployment rate; fraction of self-employed, salaried and causal workers; fraction of workers employed in farm, manufacturing, trade, construction and services sectors; fraction of adults with secondary or higher education; fraction of rural population; gender ratio; fraction of Muslim population; logged per capita household consumption. The covariates are sourced from National sample Survey (2004-05) employment-unemployment household survey.

as "underbanked". Local variation in establishments' exposure to the treatment arises through the variation in districts' underbanked status at the discontinuity threshold. *Post* is a dummy equaling 1 for years after 2005, when the BAP comes into effect. Similar to the RD design, we include a linear polynomial in the running variable, $f(Runvar)$, interacted with the post-treatment and underbanked indicators (Cingano et al., 2016; Grembi et al., 2016).[20] This ensures that the treatment effect (β) is estimated at the discontinuity threshold, where establishments in underbanked ("treatment") and non-underbanked ("control") districts are most comparable. Standard errors are clustered by district, the level at which our treatment varies. Regressions are weighted with establishment-specific weights provided by the ASI.[21]

Conditional on the smoothness of pre-treatment covariates around the discontinuity threshold, establishments in non-underbanked districts serve as a valid counterfactual for establishments in underbanked districts. To ensure the pre-treatment comparability of treatment and control units, we restrict our primary sample to establishments in districts located within a neighbourhood of 15 (bank branches per million persons) around the discontinuity threshold. The bandwidth is selected using the data-driven optimal bandwidth method of Calonico et al. (2020).[22] We also show robustness to a range of alternate bandwidths between 10 and 20 bank branches per capita.

A causal interpretation of β is subject to the standard assumption in a DiD specification: namely enterprise outcomes across underbanked and non-underbanked districts should have evolved comparably in the absence of the policy intervention. While the counterfactual is fundamentally untestable, we use an event-study framework to test whether outcomes of interest exhibited parallel trends across underbanked and non-underbanked districts prior to the BAP. Specifically, we estimate: Equation (**??**) estimates the treatment effect corresponding to each year in the sample, with 2005 ($k = -1$) – the year in which the BAP was announced – serving as the benchmark year. If establishment outcomes were comparable across underbanked and non-underbanked districts prior to the BAP, then $\beta_k = 0 \; \forall \; k \in \{-5, ..., -2\}$.

[20]Namely, we include $Runvar_d \times Post_t$ and $Runvar_d \times Underbanked_d \times Post_t$ in all our specifications. Establishment fixed effects results in the omission of $Runvar_d$ and its interaction with the underbanked indicator.

[21]The weights equal the inverse of the sampling probability. For establishments surveyed every year, the assigned weight is 1.

[22]In the absence of a prescribed method for computing the optimal bandwidth in specifications combining a RD design with a DiD ("differences-in-discontinuity specification), we use the optimal MSERD bandwidth for the year 2011, using the method of Calonico et al. (2020). For the sake of comparison, the optimal bandwidth used by Young (2017) to study the same policy intervention is 13, while Khanna and Mukherjee (2021) uses an optimal bandwidth of 20.

3.3 Pre-Treatment Covariate Balance

Before discussing our main findings, we empirically confirm that underbanked and non-underbanked districts were "balanced" along pre-treatment observable characteristics. This would substantiate the validity of the RD design, and attest to the comparability of treatment and control units. Appendix Figures B1 and B2 undertake covariate balance checks using pre-treatment district covariates from the nationally representative household employment-unemployment survey, conducted in 2004-05 by the NSSO (also referred to as the National Sample Survey (NSS) 2004-05 in this paper). These include demographic factors such as population, urbanization and education, as well as employment characteristics and household consumption. Visually, there is no evidence of any discontinuity across the 18 covariates within a narrow window of the discontinuity threshold, and the discontinuity estimates are also not statistically significant. Appendix Figures B3 and B4 replicate this exercise for pre-treatment establishment-level manufacturing outcomes.[23] Similar to the district covariates, we find no evidence of any significant differences in manufacturing outcomes across underbanked and non-underbanked districts in the pre-treatment period. Appendix Tables B2-B5 confirm these results using linear regressions.[24] Collectively, Appendix Figures B1-B4 and Tables B2-B5 affirm that within a narrow window of the discontinuity threshold and prior to the policy intervention, a) underbanked districts were observationally equivalent to non-underbanked districts; and b) manufacturing outcomes were statistically indistinguishable across underbanked and non-underbanked districts.

4 Results

We now present our key findings. First, we document the impact of the BAP on financial intermediation in underbanked regions. Next, we identify the impact of the policy intervention on manufacturing investment. Subsequently, we explore the distributional aspects of the policy and explore mechanisms explaining the observed findings.

[23]For manufacturing establishments, we collapse the pre-treatment data by computing within-establishment averages between 2000 and 2004.

[24]We regress the observable characteristic of interest on the underbanked indicator, conditional on a linear polynomial in the running variable and state fixed effects. Establishment-level regressions also includes 2-digit industry fixed effects.

4.1 Financial Deepening in Underbanked Districts

We use district-level data from the BSR and the differences-in-discontinuity design discussed in Section 3.2 to identify the impact of the BAP on bank branches and financial intermediation in previously under-served regions. As our unit of observation is the district, we estimate specification (3) using district and year fixed effects. District fixed effects allow us to flexibily control for time-invariant observable and unobservable district characteristics affecting banks' choice of location. All specifications restrict the sample to a bandwidth of 15 (bank branches per million persons) around the discontinuity threshold, with the standard errors being clustered by district.

Column (1) of Table 1 shows that relative to observationally equivalent non-underbanked districts, the average underbanked district had 5 additional private bank branches following the BAP. The coefficient is statistically significant at the 1% level and unchanged upon adding district-level covariates, interacted with the post-treatment indicator (column (2)). The coefficient is sizeable in magnitude and reflects a 25 percent increase in the number of private bank branches, when compared to the pre-BAP mean in non-underbanked areas. Columns (3)-(6) of Table 1 show that the BAP had little impact on bank branches for state-owned banks and regional rural banks. The latter is particularly reassuring as regional rural banks were exempted from the policy and the coefficients serve as a placebo check. Columns (7) and (8) show that underbanked districts had between 7 and 10 additional bank branches in the post-treatment period, driven primarily by the increase in private banks.[25] Compared to the "social banking" era, a conservative back of the envelope calculation indicates that the BAP's impact was approximately a fourth of the impact of the state-driven push to expand banking infrastructure in rural unbanked locations between 1970 and 1990.[26] The expansion in bank branches in underbanked districts following the BAP is consistent with the findings of Young (2017), Khanna and Mukherjee (2021) and Cramer (2022)

A causal interpretation of the point estimates in Table 1 is subject to the assumption that banking outcomes across underbanked and non-underbanked districts would have been comparable in the absence of the policy intervention. While the counterfactual cannot

[25]Relative to the 149 bank branches in the average non-underbanked district between 2001 and 2005, this quantifies as an increase of 5.88 percent.

[26]Burgess and Pande (2005) notes that around 30,000 bank branches were opened over a 20 year period between 1969 and 1990. This equates to approximately 84 new branches opened per district over a 2 decade period. The number needs to be interpreted with caution as a number of these districts were subsequently divided into smaller districts during the 1990s, so the aggregate average effects of the social banking program most likely reflect an upper bound.

be observed, the event-study plot in Figure 3 confirms the absence of any significant differential pre-treatment trend in private bank branch expansion in underbanked areas. Consistent with the fact that banks had to establish proposed branches within a year of approval of their expansion plan, there is a sharp increase in private bank branches in underbanked districts following the policy intervention in 2005, which persists till the end of our sample period in 2011.

The BAP clearly stated that the central bank would be assessing compliance with the policy both in terms of bank branches opened, and the volume of financial intermediation undertaken by banks in underbanked areas. To this effect, Panel A of Appendix Table A1 uses the differences-in-discontinuity specification to identify the impact of the BAP on credit disbursement by private banks. The outcome variable in the odd-numbered columns is the amount of outstanding credit; in even-numbered columns, the number of loan accounts. To account for large values of credit disbursement and a number of 0s in the pre-2005 period owing to limited operations by private banks, we use an inverse hyperbolic sine transformation of the outcome variable.

Columns (1) and (2) point to a positive treatment effect on private bank credit disbursement, although the point estimates are noisy (p-values .159 and .113). Columns (3) and (4) identify a positive treatment effect on private bank credit for manufacturing activities, with the intensive margin coefficient being significant at the 5% margin, and the extensive margin coefficient at the 10% margin. We use the proprietary BSR data in columns (5) and (6) to identify a positive and significant treatment effect for priority sector lending undertaken by private banks. This shows that private banks responded to the BAP by expanding their physical presence in underbanked areas through bank branches, and expanding credit to the priority sector.[27] Finally, columns (7) and (8) consider loans issued to manufacturing establishments within the priority sector. We identify a positive treatment effect along the extensive margin, significant at the 10% level (p-value .064), while the intensive margin coefficient is positive, albeit statistically non-significant (p-value .103).

Recent research recommends interpreting coefficient magnitudes from an inverse hyperbolic sine transformed outcome variable with caution when the raw outcome variable has a substantial mass at 0. In this regard, Appendix Table A2 re-estimates the impact of the BAP on credit disbursement by private banks with the credit outcomes in levels. The findings are qualitatively similar, but the coefficients are imprecisely estimated. Nonetheless, we identify an aggregate INR 1.2 billion additional outstanding loans from private banks in underbanked districts (p-value .09), a fifth of which is accounted for by manufactur-

[27] These are loans issued to rural village industries, as well as small and micro-enterprises.

ing credit (p-value .203). Relative to the pre-BAP mean in non-underbanked districts, the treatment effect amounts to a 12 percent increase in private bank manufacturing credit. Approximately 70 percent of this was accounted for by private bank credit to small and micro-enterprises in underbanked areas [column (7)].

Appendix Figure A3 shows event-study plots estimated using specification (**??**), which identifies the average annual treatment effect of the BAP on private bank credit. The coefficients are benchmarked to the year 2005 – the year in which the BAP was announced, and the vertical lines denote 95% confidence intervals. Consistent with our identifying assumption, there is no evidence of differential pre-treatment trends in aggregate private bank credit, or private bank manufacturing credit across underbanked and non-underbanked districts. There is an increase in both total private bank credit, and private bank manufacturing credit following the policy intervention, which becomes statistically significant at the 5% level in the final three years of our sample period (2009-11). Similar to the muted impact of the BAP on the expansion of state-owned banks in underbanked areas, Panel B of Appendix Table A1 shows that financial intermediation by state-owned banks too remained unaffected in response to the BAP. This limit concerns that the increased credit disbursement by private banks identified in Panel A of Appendix Table A1 emanated from a secular expansion of bank lending in these districts in response to increased demand.

Our empirical strategy compares private bank credit before and after the policy intervention, across underbanekd and non-underbanked districts. This makes it challenging to ascertain whether the expansion in private bank credit to under-served regions came at the cost of credit expansions in non-underbanked areas. Aggregate trends in the top row of Appendix Figure A4 shows an increase in private bank credit across both underbanked and non-underbanked areas, with a sharper increase in manufacturing credit in underbanked districts following the BAP in 2005. This partially limits concerns that the expansion in private bank credit to underbanked regions was driven by a substantial reallocation of credit from non-underbanked areas. The RD plot in the top-left panel of Appendix Figure A5 further supports this contention: while the private bank credit-deposit ratio is significantly larger in underbanked districts, the average credit-deposit ratio is 0.51.[28] As the average credit-deposit ratio in underbanked districts was less than 1 in 2010, it implies that private banks could have supplied the additional credit in treated areas through local deposits, and would not be required to reallocate credit from other regions through internal capital markets.

[28]The mean credit-deposit ratio for private banks in non-underbanked districts in 2010 was 0.29. The RD coefficient is 0.22 (p-value .004)

4.2 Bank Branch Expansion and Manufacturing Investment

We exploit the time variation in the onset of the BAP and the panel structure of the ASI to combine a DiD approach with a sharp RD design to causally identify the impact of financial deepening on manufacturing investment. Our primary outcome of interest is investment in plant and machinery, defined in equation (2). Unless otherwise specified, all estimates are undertaken using the sample of establishments located in districts within a bandwidth of 15 around the discontinuity threshold.

Column (1) of Table 2 shows a parsimonious specification including only establishment and year fixed effects, and identifies a near 5 percentage point increase in manufacturing investment, with the coefficient being statistically significant at the 5% level. Column (2) replaces the year fixed effects with 2-digit industry-year fixed effects, with little impact on the point estimate. The industry-year fixed effects absorb industry-specific time-varying demand and productivity shocks common to all establishments operating within the broad industry category and year. This implies comparing manufacturing investment for enterprises operating in the same broad industry and year. Column (3) adds establishment-specific covariates, including a quadratic in establishment age, and also dummies for establishment size, interacted with year dummies.[29] Column (4) further includes district covariates, interacted with a post-treatment indicator. This increases the coefficient magnitude slightly and improves its precision to the 1% level. Column (5) replaces the quadratic in establishment age with age fixed effects, while column (6) replaces the 2-digit industry-year fixed effects with 3-digit industry-year fixed effects, limiting our comparison to an even smaller set of establishments in each year. The coefficient corresponding to $Underbanked_d \times Post_t$ is stable in terms of both magnitude and statistical significance in all the above specifications.

Our preferred specification is column (4) of Table 2 which includes establishment, 2-digit industry-year, and establishment size-year fixed effects, along with establishment and district covariates. The point estimates suggests that the average manufacturing establishment in an underbanked district witnessed a 6 percentage point increase in capital expenditures in the post-treatment period, relative to manufacturing establishments in observationally equivalent non-underbanked districts. Compared to the pre-treatment mean in control regions, this implies an additional INR 1.8 million investment in plant

[29]We include 5 dummies for establishment size, based on the pre-treatment average number of workers hired. We also consider the following establishment specific covariates: a dummy for whether the establishment is located in a rural area; a dummy for whether the establishment uses computers for accounting; and dummies for ownership type.

and machinery for establishments in underbanked districts in the post-treatment period.[30] This reflects a substantial increase when considering that the median value of establishment plant and machinery in the pre-treatment period was INR 0.9 million in non-underbanked districts.

Appendix Table B6 considers alternate functional forms and outcomes of interest. Column (1) shows that our results are robust to measuring capital expenditures as the logged difference between closing and opening values of net plant and machinery. Section 3.1.1 discusses the lumpiness of capital spending and notes that less than a fourth of the establishments undertook any positive capital spending in a given year. To this effect, the outcome of interest in column (2) is a binary equaling 1 if the closing value of establishment net plant and machinery exceeded the opening value in a year, and 0 otherwise. We find that the treatment increased establishments' likelihood of engaging in any positive capital spending by 5 percentage points – a 20% increase. With 57 manufacturing establishments operating in the pre-treatment period in the average non-underbanked district, this implies that approximately 3 additional establishments engaged in positive spending on plant and machinery in the post-treatment period in underbanked districts. Columns (3)-(5) show that the results are comparable if we use net fixed assets to measure capital spending. In line with higher capital spending, columns (6)-(7) also identify a positive impact on the use of raw materials.

We next verify whether the observed increase in capital spending is accompanied by increased credit access for manufacturing establishments. Appendix Table B7 reports the impact of the BAP on manufacturing establishments' access to credit. An important caveat is that while the ASI reports outstanding loans, it does not separate establishment credit across bank and non-bank sources, or by bank group. Consequently, our outcomes compare aggregate credit growth for establishments across underbanked and non-underbanked districts. We would expect this to increase if credit from private banks supplemented other credit sources and enabled establishments to undertake higher capital spending.

The outcome variable in column (1) of Appendix Table B7 is credit growth, defined as in equation (2), using the opening and closing values of outstanding loans, while column (2) measures credit growth as the logged difference in closing and opening values of outstanding loans in a year. We identify a positive and statistically significant increase for

[30] In the pre-treatment period, the mean establishment value of plant and machinery (average of net opening and closing values) equaled INR 29.8 million for establishments in non-underbanked districts, located within a bandwidth of 15 around the discontinuity threshold. A 6 percentage point increase thus amounts to INR 1.82 million higher spending on plant and machinery.

both outcomes. Column (1) shows a 13 percentage point increase in credit growth for manufacturing establishments in underbanked districts – equivalent to INR 3.7 million – about twice the increase in capital spending. With 140 manufacturing establishments operating in the average underbanked district, the coefficient points to a INR 513 million increase in aggregate manufacturing credit. As Section 4.1 identified an increase in manufacturing credit from private banks equivalent to INR 278 million, it suggests that the BAP induced increase in private bank manufacturing credit accounted for approximately 54 percent of the aggregate increase in manufacturing credit in under-served areas.

Column (2) of Appendix Table B7 shows that our results are similar if we measure credit growth as a logged differences. Columns (3)-(5) show that the increase in credit is primarily along the intensive margin, with little impact of the treatment on entry of manufacturing establishments into credit markets. Lack of information on loan applications made by manufacturing establishments preclude us from assessing whether this was due to a lack of demand, or a denial by lenders to extend credit to such establishments. Column (6) finds no impact of the policy intervention on the aggregate cost of credit facing manufacturing establishments.

Figure 4 presents event study plots, identifying the average annual treatment effect for three key outcomes of interest: capital expenditures, the likelihood of any positive capital spending, and credit growth. To ensure consistency with our preferred baseline specification, we include establishment, 2-digit industry-year and size-year fixed effects, as well as establishment and district-level controls. Standard errors are clustered by district and the sample is restricted to establishments located in districts within a bandwidth of 15 around the discontinuity threshold. Figure 4 lends support to our identifying assumption and finds all three outcomes of interest to exhibit comparable trends across underbanked and non-underbanked districts prior to 2005. In the aftermath of the BAP, there is a visible jump in the point estimates. For capital expenditures (top-left panel), the coefficient estimates are positive and statistically significant at the 5% level for the years 2007, 2008 and 2009, while they are statistically significant at 10% level for the remaining post-treatment years. The increase in credit growth (top-right panel) begins in 2006 and is the largest for the years 2008-11, coinciding with the expansion in manufacturing credit from private banks in underbanked districts (Appendix Figure A3). The presence of parallel trends and the sharp uptick in capital investments and credit growth following the onset of the BAP allows us to assign a causal interpretation to the baseline findings in Table 2.

4.2.1 Robustness Checks

We subject our baseline results to a number of robustness checks. Table 3 shows that our results are unchanged if we do not weight the specifications with establishment-specific weights (column 1); or two-way cluster standard errors by district and industry (column 2). Column (3) excludes the 9 districts for which the RBI did not precisely follow the assignment rule laid out in equation (**??**). The results are unaffected by this sample restriction.

All specifications till now restrict the primary sample till 2011 due to a new branching policy unveiled that year (discussed in Section 3.1.1). Column (4) relaxes this restriction and uses data till 2014, and finds the long-term impact on manufacturing investment to be comparable to the medium term effects. Column (5) undertakes a placebo test by restricting the sample to 2005 – the year of introduction of the BAP – and defines the post-treatment period as starting from 2002.[31] The point-estimate obtained using this pseudo-treatment is attenuated towards 0, assuaging concerns that the identified treatment effect can be attributed to an overall positive trend in manufacturing investment, coinciding with the timing of the policy intervention.

Our preferred specification uses a bandwidth of 15 (bank branches per million persons) around the discontinuity threshold. We choose this using the data-based optimal bandwidth selection methods recommended by Calonico et al. (2020).[32] Figure 5 shows that our findings are invariant to alternate bandwidth choices. Specifically, we re-estimate our baseline specification for bandwidths between 10 and 20 and plot the coefficients in Figure 5. For all 20 bandwidths, the coefficient estimate is between .03 and .06, and only 3 out of 20 coefficients are statistically insignificant at the 10% level (the largest p-value being .13).[33] This affirms that our results are not contingent on any specific bandwidth. The most conservative bandwidth of 10 identifies a 5 percentage point increase in manufacturing investment in underbanked districts, with the coefficient being significant at the 5% level.

Appendix Figure B5 plots the coefficients obtained from using a cross-sectional RD spec-

[31]This provides us with 4 years of pre-treatment data, and 3 years of post-treatment data.

[32]The optimal MSERD bandwidth using a cross-sectional RD specification for the year 2012 is 14.359. The outcome of interest is capital expenditures in plant and machinery, and the RD specification controls for 3-digit industry fixed effects, state fixed effects, establishment and district covariates. Standard errors are clustered by district.

[33]The first coefficient in Figure 5 is estimated using a bandwidth of 10. Subsequent specifications are re-estimated after incrementally increasing the bandwidth by 0.5. The last specification uses a bandwidth of 20.

ification for each year between 2001 and 2011, using optimal MSERD data driven bandwidths suggested by Calonico et al. (2020). The largest MSERD bandwidth in Appendix Figure B5 is 14.359; the smallest, 6.590.[34] Even for the smallest MSERD bandwidth (in 2008), we identify a 3 percentage point increase in capital spending for manufacturing establishments in underbanked districts (p-value 0.16). Consistent with the event-study plot in Figure 4, the increase in capital spending for manufacturing establishments in underbanked districts is evident only after 2005, once the BAP comes into effect. 5 out of 6 post-treatment coefficients are positive and comparable in magnitude to the coefficients obtained in Figure 4, with 4 being statistically significant at the 5% margin or better.

Figure 6 shows that our findings are not driven by any single state or industry. We establish this by re-estimating our baseline specification after dropping one state/industry at a time. As seen from both panels of Figure 6, the coefficients are not sensitive to the exclusion of any single state or industry – all the coefficient estimates remain positive, centred around 0.06, and statistically significant at the 10% level or better. This reassures us that the positive treatment effect on manufacturing investment was not driven by some confounding state or industry-specific place-based policy, the timing of which also coincided with the policy intervention of interest.

4.3 Distributional Impacts of Bank Branch Expansion

Our baseline results show that districts witnessing an expansion in financial infrastructure see higher capital investment and credit growth for registered manufacturing establishments. We now explore distributional implications to assess whether financial deepening also resulted in an alleviation of credit constraints. We begin by exploring treatment heterogeneity by establishment size. To avoid any contamination of establishment size by the policy intervention, we compute establishment size using the average number of workers hired between 2001 and 2004 and classify establishments as "large" or "small" based on the median establishment size in the pre-treatment period (16 workers).[35] Using

[34] All cross-sectional specifications include a linear polynomial in the running variable, state and 3-digit industry fixed effects, and establishment and district covariates. All specifications are weighted using a triangular kernel and establishment-specific weights. Standard errors are clustered by district.

[35] Specifically, we use the pre-treatment median establishment size for establishments located in non-underbanked districts and within a bandwidth of 15 around the discontinuity threshold.

this classification, we estimate the triple-difference specification:

$$Y_{ijdt} = \alpha_i + \delta_{jt} + \pi_1 Underbanked_d \times Post_t + \pi_2 Underbanked_d \times Large_i \times Post_t + f(Runvar_d) + \gamma \mathbf{X}_{ijdt} + \epsilon_{ijdt} \quad (4)$$

The double-difference coefficient π_1, in Equation (4), compares capital investments across underbanked and non-underbanked districts for smaller establishments. The triple difference coefficient (π_2) identifies the differential effect on capital spending within underbanked districts for larger establishments.

Columns (1) and (5) of Table 8 estimate treatment heterogeneity in capital expenditures and credit growth across "large" establishments, respectively. We find positive and significant coefficients on the double-difference term (π_1). The triple interaction coefficient (π_2) is negative, albeit statistically significant only for credit growth. We cannot reject the null of $\pi_1 + \pi_2 = 0$ for credit growth, pointing to comparable growth in outstanding loans for large manufacturing establishments across underbanked and non-underbanked districts.

Columns (2) and (6) use thresholds of 10, 25, 50 and 100 workers to identify treatment heterogeneity within each establishment bin. The double difference coefficient shows that establishments employing less than 10 workers in underbanked districts witnessed 8 (27) percentage point higher capital spending (credit growth) post-treatment, relative to those in non-underbanked districts. The triple interaction coefficients are all negative, albeit imprecisely estimated.[36] For capital expenditures, the sum of the double and triple interaction terms differ significantly from 0 for establishments employing between 10 and 25 workers (p-value: .068). Manufacturing credit growth in underbanked districts however is driven primarily by the smallest establishments employing less than 10 workers, followed by those employing 10-25 workers.

Columns (1)-(2) and (5)-(6) of Table 8 show that manufacturing investment and credit growth in underbanked districts was driven by relatively smaller establishments for whom credit constraints are also most likely to bind. While firm size is widely used as an indicator of being credit-constrained, employment decisions are endogenous: firms can choose to remain small either because it is optimal, or in response to market distortions. This is particularly relevant in the Indian context as size thresholds have been extensively employed to determine firms' eligibility for subsidized credit (Banerjee and Duflo, 2014). To

[36]The triple interaction coefficients estimating heterogeneity for establishments employing between 10 and 25, and 25 and 50 workers are attenuated towards 0 for capital expenditures, signifying no evidence of a differential effect.

this effect, we consider heterogeneity across both establishment size and age (Criscuolo et al., 2019; Hadlock and Pierce, 2010). The intuition is that younger firms require time to scale up and are initially small due to operational and logistical constraints (including limited credit availability).

Columns (3) and (7) of Table 8 split our sample into 4 mutually exclusive groups: small and young (omitted category); small and old; large and young; and large and old. We use the pre-treatment median establishment size to classify establishments as small or large. Establishments are classified as young if their operations started after 1992.[37] Consistent with Criscuolo et al. (2019), columns (3) and (7) show that the positive treatment effects are driven by establishments which are both small *and* young, with the triple interaction coefficients for both columns being negative.[38] Overall, capital expenditures and credit growth for small and young establishments in underbanked districts were 13 and 31 percentage points higher in the post-treatment period.

Farre-Mensa and Ljungqvist (2016) argue that firm size and age capture life-cycle effects of firms, and are inaccurate measures of financial constraints. Instead, they recommend firms' listing status as a more appropriate signal for being financially constrained.[39] We use information on the establishment's organization to create the binary variable *Listed* if the establishment is owned by a corporation which is publicly listed. Consistent with credit constrained establishments increasing their capital investments in areas witnessing financial deepening, columns (4) and (8) show that the positive treatment effects on capital investment and credit growth are driven entirely by establishments which are *not* publicly listed.

We check the robustness of these findings to other definitions of firm size. Columns (1)-(3) of Appendix Table B8 show comparable findings if pre-treatment establishment fixed assets are used to determine establishment size. In particular, columns (2) and (3) use administrative definitions based on the value of establishment plant and machinery to show that the treatment effects are concentrated amongst small establishments, and establishments qualifying as small-scale industries. This echoes Banerjee and Duflo (2014), who document that a relaxation in the administrative cutoff for small-scale industries resulted in the alleviation of credit constraints. The point estimates in column (3) show that

[37]We use this year as the cutoff as a major overhaul of the Indian economy was undertaken in 1991, encouraging private competition. Using this cutoff implies that establishments classified as young were at most 13 years old at the time of the policy intervention.

[38]When comparing the sum of the double and triple interaction coefficients, we identify a positive impact for capital spending (credit growth) for large and young (small and old) establishments.

[39]In Farre-Mensa and Ljungqvist (2016), being publicly unlisted is a necessary condition for being financially constrained, but not a sufficient condition.

such establishments in underbanked districts increased their capital investments in plant and machinery by an additional INR 0.13 million after the BAP. Comparable results are obtained in Appendix Table B9, columns (1)-(3) when the outcome of interest is credit growth.

Column (4) of Appendix Table B8 considers heterogeneity by establishments' tangible assets. Tangible assets refer to the value of land and building owned by the establishment, often used as collateral to secure credit. If locational proximity to borrowers aids information collection, we would expect a lesser role for collateral in loan covenants (Fisman et al., 2017). Column (4) offers partial support for this hypothesis: while the triple interaction coefficient identifying treatment heterogeneity across establishments with high (above-median) ex-ante collateral values is not significantly different from 0, the uninteracted coefficient corresponding to establishments with low collateral is positive and statistically significant at the 5% level. This signifies that the ability to offer collateral in the form of immovable physical assets was not a necessary condition for undertaking capital investment in underbanked districts. Credit growth in underbanked districts was driven by establishments with relatively low collateral values (Appendix Table B9, column 4). Finally, column (5) of Appendix Tables B8 and B9 show the positive treatment effects on manufacturing investment and credit growth to be driven by individual proprietorships and family owned enterprises, which again are likely to be smaller and face binding credit constraints.[40]

Collectively, we see that manufacturing investment and credit growth in underbanked districts was concentrated amongst smaller establishments – particularly, small and young establishments, and establishments not publicly listed. An extensive literature studying the finance-growth nexus shows that these establishments are also most likely to face binding credit-constraints. Consequently, our results support the contention that the expansion in local financial infrastructure aided the alleviation of credit-constraints and allowed these enterprises to invest in productive capital. Higher credit growth for small and young establishments, and establishments not publicly listed points to credit allocation towards firms with limited "hard" information, and for whom the collection of "soft" information would be critical for effective financial intermediation (Aggarwal and Hauswald, 2010; Liberti and Petersen, 2019). We discuss this channel in greater detail while exploring mechanisms explaining the increase in capital spending and credit

[40]The ASI precludes the linking of establishments to parent firms but does provide broad ownership categories. We use individual proprietorships and family-owned enterprises as our benchmark category and explore treatment heterogeneity across establishments classified as partnerships, private limited companies, government-owned/aided enterprises and public limited (listed) companies.

growth for smaller establishments.

4.4 Quality of Credit Intermediation

Our baseline results document a higher disbursement of manufacturing credit from private banks in underbanked districts following the BAP. This is accompanied by increased capital spending and credit growth amongst registered manufacturing establishments, driven by establishments most likely to face binding credit constraints. Relative to government banks, private banks are considered to have superior loan officer incentives, corporate governance, and adhere to profit-maximizing motives. They are also less susceptible to political capture. This leads us to examine how the expansion in private bank credit in underbanked districts affected the overall quality of financial intermediation. While enhanced career incentives for loan officers and the lack political interference can improve borrower quality through better screening and monitoring, an increased appetite for risk can also lead to higher loan delinquency.

Following Jayaratne and Strahan (1996), we use loan delinquency as a proxy for lending quality and compare non-performing loans (as a fraction of total loans) across underbanked and non-underbanked districts. As loan delinquency is often a function of time, we compare the share of non-performing loans in March 2016 – a decade after the policy intervention – using RBI's proprietary data on branch-level non-performing assets.[41] Using a cross-sectional RD specification, the top-right panel of Appendix Figure A5 finds no difference in the share of non-performing loans across underbanked and non-underbanked districts at the discontinuity threshold.[42] While this does not point to an improvement in the quality of financial intermediation by private banks, it also rules out that the expansion in private bank credit was accompanied by a deterioration in lending quality.

Next, we use ASI's establishment-level data to identify whether the increase in manufacturing investment was concentrated amongst productive establishments. A key role of banks is to shape societal allocation of resources by selecting firms which receive financing (Bai et al., 2018). While we cannot distinguish across sources of credit, an increase in capital expenditures by enterprises with relatively high ex-ante creditworthiness would

[41] We aggregate the branch-specific share of non-performing loans to the district by computing loan-weighted averages of branch NPAs.

[42] We opt for a cross-sectional sharp RD specification as a number of districts had no private bank operations in the prior to 2005, which leads to the share of non-performing loans being undefined in such instances.

be consistent with the contention that private banks were more effective at allocating credit to borrowers of higher quality. It would also be consistent with the overall findings of Table 2, as we would expect financial deepening to affect manufacturing investment only if firms had projects with net positive returns.

We use four pre-treatment measures of firm quality: namely marginal product of capital, output per worker, pre-treatment interest rates, and interest coverage ratio. The first two measures proxy for establishment quality, while the last two capture credit risk and delinquency. Appendix Table B10 broadly supports the explanation that the increase in manufacturing investment in underbanked districts was driven by creditworthy establishments. We identify a positive coefficient on the triple interaction term corresponding to establishments with higher marginal product of capital (column 1), and negative coefficients for establishments facing higher ex-ante interest rates (column 3), and establishments whose interest payments exceeded annual sales at least once between 2001 and 2005 (columns 4). The double-difference coefficients in the latter two instances are positive and significant at the 1% level, signifying that the increase in manufacturing investment in underbanked districts was driven by establishments with low cost of borrowing, and no record of delinquency in the medium-term prior to the treatment. These findings are also consistent with the absence of any aggregate increase in non-performing loans for private banks in underbanked districts. There is however no evidence of treatment heterogeneity across establishments with relatively high pre-treatment output per worker (column 2): contrary to expectations, the triple interaction coefficient is negative, albeit statistically non-significant, and the double interaction coefficient suggests that capital spending was also undertaken by establishments with relatively low pre-treatment output per worker.[43]

4.5 Output and Profitability

Our empirical findings identify a positive impact of bank branch expansion on manufacturing investment and credit growth. This is driven by establishments most likely to face binding credit constraints. There is also suggestive evidence that credit was allocated towards establishments of better quality. We now identify the impact of the BAP on manufacturing output, profitability and employment. If capital investments are productivity enhancing, we would expect the BAP to also increase manufacturing output and profits. Capital spending can also increase profitability without affecting output by lowering the

[43]This could also reflect relatively lower capital investment by bigger firms in underbanked districts since bigger firms are also more likely to have higher labor productivity or output per worker.

cost of production. Effect on employment at firm level can be ambiguous depending on whether capital and labour are substitutes, or complementary.

Appendix Table B11 shows the impact of the treatment on the reported value of output and establishment profits. Profits are defined as output, net of depreciation and input costs, where the latter takes into account interest expenses and labour costs. As profits can be negative, we use an inverse hyperbolic sine transformation to measure establishment profits. Since the measurement of inputs and labour costs can be noisy, we also define a binary indicator equaling 1 if the establishment reported positive profits during the year. Columns (1)-(3) suggest that output and profitability was unaffected by the policy intervention. Contrary to our expectations, the point estimates are negative, albeit imprecisely estimated. It is however possible that output and profits respond to capital investments with a lag, and the restriction of the sample to 2011 hinders our ability to detect an effect. In columns (4)-(6), we extend the sample to 2014 and identify long-run treatment effects on output and profits. While the coefficients remain insignificant, the point estimates for profits are now positive. We explore this further in columns (7)-(9) using the full sample, but splitting the post-treatment period into two: between 2006 and 2010, and from 2011 to 2014. For all three outcomes of interest, the treatment effect for the 2011-2014 period is positive, and statistically significant at the 5% level for the binary measure of profits. Specifically, column (9) shows that establishments in underbanked districts are 5 percentage points more likely to report positive profits during this period – a 6 percent increase in profitability.[44] Column (7) points to a near 8 percent increase in establishment output between 2011 and 2014, although the confidence intervals are too wide to reject a null effect.

Overall, Appendix Table B11 shows that capital investments positively affect enterprise performance, especially in terms of profitability, but the returns accrue only over the long-term. Appendix C examines in greater detail the short and medium-term impact of capital investment on manufacturing output. Drawing from the literature on managerial capital, we find positive treatment effects on output over the medium term, but only for establishments with relatively high managerial capital. This also holds among the more credit constrained enterprises. The findings affirm to the critical role played by managerial capital in firm operations. We find no significant effect of financial deepening on firms' labor demand: if anything, the evidence points to a reduction in the number of contractual workers hired by establishments. When considering that labour costs did not increase following the BAP, this indicates that enterprises were possibly substituting labor with

[44]In the pre-treatment period, 83 percent of establishments reported positive profits in control districts. The treatment effect is computed as $.048/.83 = .06$

capital. The results are discussed in detail in Appendix Section C.1 and C.2.

5 Mechanisms

This section explores plausible mechanisms explaining the increase in manufacturing investment and credit growth in smaller establishments in response to financial deepening. We consider five channels: physical proximity of lenders to borrowers; increased staffing by lenders to collect borrower information; creditors' specialization in lending to small manufacturers; increased lender competition lowering the cost of credit; and aggregate demand.

5.1 Distance to Manufacturing Establishments

Screening and monitoring costs incurred by lenders are typically increasing in distance (Aggarwal and Hauswald, 2010; Chen et al., 2022; Chernenko et al., 2022; Gilje et al., 2016; Hollander and Verriest, 2016). Physical proximity to borrowers aids in the collection of "soft" information which is not easily substitutable (Nguyen, 2019), and critical with regard to small unlisted borrowers for whom there is limited "hard" information (Aggarwal and Hauswald, 2010; Liberti and Petersen, 2019). A reduction in the cost of information acquisition can improve both the volume, and the quality of information collected by lenders (Chen et al., 2022). Thus, physical proximity can reduce the costs of financial intermediation, contributing to higher credit disbursement towards informationally opaque borrowers. Consequently, we test whether the BAP reduced the physical distance between lenders and manufacturing establishments in underbanked districts.[45]

The empirical challenge here is that while bank branch addresses are publicly listed, the ASI only provides location identifiers at the level of district. This precludes us from testing whether the policy reduced establishments' distance to private banks in underbanked districts. We overcome this by using data from the Economic Census (EC), which allows us to map a subset of rural business establishments to pincodes. We combine this with the mapping of bank branches to pincodes, and compute the minimum Euclidean distance of establishments to private and government-owned bank branches.[46]

[45]Despite the banking expansion, if new bank branches are concentrated around existing branches, it would not result in a reduction in the distance between bank branches and establishments.

[46]We use the dataset compiled by Agarwal et al. (2021) to obtain pincodes of private bank branches in India. Agarwal et al. (2021) scrape the physical addresses of over 120,000 bank branches operating in

We use the 2013 EC and a cross-sectional RD specification to test whether manufacturing establishments' average distance to private banks was significantly lower in underbanked districts. We begin by aggregating the establishment-level data to the district and compare the median establishment-private bank distance across underbanked and non-underbanked districts. As the ASI covers only registered manufacturing establishments, and the EC provides no information on establishments' registration status, we exclude manufacturing establishments employing less than 5 workers.[47] The left panel of Figure 7 restricts the sample to districts within a window of 15 (bank branches per million persons) around the discontinuity threshold and identifies a significant reduction in the median distance to private banks in underbanked districts. Specifically, the median distance to private banks for manufacturing establishments declined by 6 kilometres (km) in underbanked districts (relative to a mean distance of 13 km in non-underbanked districts). The right panel of Figure 7 shows no corresponding impact for establishments' distance to government-owned banks.

We next use the establishment-level data from the EC to verify the aggregate findings of Figure 7. This allow us to control for establishment characteristics such as industry of operation, and also explore heterogeneity by establishment size. We estimate local linear regressions of the form:

$$DistPvtBank_{ijds} = \alpha_s + \delta_j + \beta Underbanked_d + f(Runvar_d) + \gamma X_{id} + \epsilon_{ijds} \tag{5}$$

The unit of observation in Equation (5) is establishment i, located in district d of state s, and operating in industry j. The outcome of interest is the minimum distance to a private bank branch. We include state and industry fixed effects (α_s and δ_j) along with controls for establishment ownership and additional pre-treatment district covariates. $f(Runvar_d)$ denotes a linear polynomial in the running variable, and its interaction with the *Underbanked* indicator. The coefficient of interest is β, comparing the minimum distance to private banks across manufacturing establishments in underbanked and non-underbanked districts at the discontinuity threshold. The sample is restricted to establishments employing at least one worker and located in districts within a neighbourhood of 15 around the discontinuity threshold. Standard errors are clustered by district.[48]

India in 2016. We use the pincodes of bank branches which began operation prior to 2012, and obtain the latitude and longitude of the centroid of each pincode, which is then used to compute the Euclidean distance between two pincodes.

[47] In the pre-treatment period, establishments employing less than 5 workers comprised the bottom decile of the ASI's distribution of workers.

[48] This effectively excludes establishments where the owner is also the sole worker, or self-employed manufacturing enterprises. As such establishments are also most likely to be unregistered, this sample

Column (1) of Table 5 shows a 4.5 km reduction in the minimum distance to a private bank in underbanked districts – a 30 percent decline.[49] Columns (2)-(4) show heterogeneity by establishment size. We find the decline in distance to the nearest private bank branch to be strongest for establishments employing between 5 and 25 workers.[50] There is however no significant reduction in the distance to private banks for larger establishments hiring in excess of 25 workers (column 4), or establishments' distance to government-owned banks (columns 5-8).

Finally, using Equation (4) we show that the increase in capital spending and credit growth is driven by establishments located in districts where establishments are in close proximity to private banks. We compute the within-district median establishment-private bank distance from the EC and define the binary variable *High Dist. Pvt. Bank* to equal one if the within-district median distance exceeds the median value across all districts.[51] Columns (1) and (4) of Table 6 confirm that the increase in capital spending and credit growth is driven by establishments operating in underbanked districts where the median establishment-private bank distance is relatively small. The triple interaction coefficient corresponding to *High Dist. Pvt. Bank* is negative and large for both capital expenditures and credit growth, and statistically significant for capital expenditures.

If information acquisition is a key friction for financial intermediation and physical proximity to borrowers aids the acquisition of information by lenders, then this channel is likely to be most beneficial for smaller informationally opaque firms. Consistent with this hypothesis, columns (2) and (5) of Table 6 show higher manufacturing investment and credit growth for relatively smaller establishments (less than 25 workers) in underbanked districts with a relatively low establishment-private bank distance. In contrast, columns (3) and (6) identify no impact of the policy intervention on capital spending and credit growth for larger establishments (hiring more than 50 workers), irrespective of whether they were located in districts where the distance to private banks was smaller.

While underbanked and non-underbanked districts are observationally equivalent within a narrow window around the discontinuity threshold, the entry of private banks in underbanked districts is likely to suffer from selection, as banks were free to enter any underbanked district. Additionally, since we can only map rural establishments to bank

restriction makes the EC more comparable to the ASI.

[49]The mean establishment-private bank distance in non-underbanked districts is 13.42 km.

[50]Conditional on employing at least one worker, establishments employing at least 25 workers constitute the top decile of EC establishments. The set of establishments employing between 5 and 25 workers account for 40 percent of EC establishments.

[51]We restrict the sample to establishments employing at least 5 workers to achieve comparability with the ASI.

branches, the findings of Table 6 should be interpreted with caution.[52] Nonetheless, the mapping of manufacturing establishments to private bank branches in the EC shows that the BAP significantly reduced establishments' distance to private bank branches in underbanked districts. Suggestive evidence also shows the increase in manufacturing investment and credit growth to be driven by establishments in treated districts where the distance to private bank branches was relatively small. This indicates that the physical proximity of private banks to borrowers aided credit growth and capital investment for these establishments. While we cannot distinguish whether this was due to the improved screening of borrowers ex-ante, or better monitoring ex-post, the results broadly underline the role of local credit markets in facilitating financial intermediation (Aggarwal and Hauswald, 2010; Nguyen, 2019).

5.2 Increased Staffing by Banks

We next examine whether the BAP affected the staffing decision of banks. If banks were entering underbanked districts to comply with the BAP and engaged in financial intermediation, we would also expect to observe increased hiring of loan officers, responsible for evaluating and processing loan applications. Aggarwal and Hauswald (2010) in particular note the vital role played by loan officers in collecting "soft information" and subjective intelligence with regard to informationally opaque small borrowers. To this effect, we use proprietary data from the BSR to identify the impact of the policy intervention on the staffing decisions of banks across underbanked and non-underbanked districts.

The proprietary BSR data separates bank employees into two categories – officers and clerks. The former comprises of skilled bank employees who are typically hired through competitive examinations and are responsible for overall branch management, and responsible for lending decisions. To this effect, we aggregate the number of officers and employees to the level of district in each year and use our primary empirical specification to identify the impact of the BAP on staffing decisions of private and state-owned banks.

Column (1) of Table 7 shows that private banks in underbanked areas had 35 additional areas in the aftermath of the BAP. This is a near 36 percent increase, relative to the control group mean prior to 2005, and the coefficient is statistically significant at the 1% level. Column (3) shows that in aggregate, private banks had 52 additional employees in under-

[52]The *High Dist. Pvt. Bank* would remain unaffected as long as the distance to private banks from urban establishments was not significantly larger in districts where the distance to private banks for rural establishments was relatively small.

banked districts – a 24 percent increase. The coefficients in columns (1) and (3) indicates that 2 out of every 3 additional employees engaged by private banks in underbanked districts were officers. While we cannot specifically observe loan officers or the actual processing of loan applications by banks, the increased staffing of branches by officers indicate that private banks were investing in the necessary human capital to undertake lending to small borrowers in underbanked districts. Importantly, columns (2) and (4) show no impact of the BAP on staffing levels of state-owned banks, either in terms of officers, or overall employees. The coefficients are a third in magnitude and statistically non-significant. This again underlines that private banks primarily responded to the BAP, through an increase in bank branching, lending activities, and increased staffing of bank branches. If officers engaged in the acquisition of "soft information" and monitoring of relatively smaller manufacturing establishments, it would be consistent with the increase in credit growth and capital spending for these establishments identified in Section 4.3.

5.3 Comparative Advantage of Lenders

We next explore whether specific attributes of entrant banks can explain the increase in manufacturing investment. We examine three attributes in this regard: first, if the district witnessed entry by a small private bank; second, if the district witnessed entry by a private bank specializing in lending to small borrowers; and third, if the district witnessed entry by a private bank specializing in lending to small *manufacturing* borrowers. Each of the above hypothesis essentially tests for comparative advantage of entrant banks in small lender operations, and is motivated from studies showing small banks to possess informational advantages in lending to small borrowers (Bai et al., 2018; Berger et al., 2005; Liberti and Petersen, 2019). Resultantly, we examine whether the increase in capital spending by small and micro-enterprises can be explained by the entry of small private banks in underbanked districts.

We use proprietary bank lending data from the BSR between 2000 and 2005 to classify banks into the above three categories. "Small" banks are defined as those whose aggregate loan portfolio is less than the median loan portfolio during this period. Banks specializing in small loans are defined as those where the average loan size is less than the median loan size across all banks.[53] Banks specializing in lending to small manufacturing enterprises are those which have a relatively high (above median) share of small borrowers within their manufacturing portfolio.

[53]The average loan size is computed as total outstanding loans, divided by total loan accounts.

Column (1) of Table **??** offers weak evidence in support of the first hypothesis. While the triple interaction term corresponding to entry by small private banks is positive, it is statistically insignificant. The sum of the coefficients is positive and significant at 5% level (p-value .013), pointing to a 6 percentage point increase in manufacturing investment in underbanked districts witnessing entry by a small private bank. We find no evidence to support the second hypothesis (column 2), since the triple interaction term for entry by private banks specializing in small loans is negative, large, and statistically significant

Finally, the triple interaction coefficient in column (3) is positive and statistically significant, indicating that the increase in capital spending is undertaken by establishments located in districts which witnessed entry by banks with a relatively high share of small manufacturing loans. This is consistent with the hypothesis that manufacturing investment increased in districts witnessing entry by private banks specializing in lending to small manufacturing establishments. Moreover, the double-difference coefficient in column (3) is attenuated towards 0 and not statistically significant. This affirms no difference in manufacturing investment across underbanked and non-underbanked districts, for the subset of districts not witnessing entry by a private bank specializing in lending to small manufacturing borrowers.

Given the flexibility accorded to banks to locate in underbanked districts, a number of district-specific factors could be correlated with a bank's decision to begin operations in a district in response to the BAP. While all our specifications control for district observables, we cannot rule out unobservable factors determining the entry of select private banks. Nonetheless, columns (1) and (3) of Table **??** offer suggestive evidence consistent with the explanation that the specialization of banks' lending operations matters for capital investment by manufacturing establishments.[54]

5.4 Cost of Credit

An expansion in bank branches can increase competition amongst financial institutions, which in turn can facilitate financial intermediation, both through incumbents' incentive to preserve their market share, and a reduction in the cost of credit (Carlson et al., 2022). Consequently, lender competition can increase capital investments by lowering the marginal cost of capital. If borrowing costs were higher for smaller informationally

[54] As we are comparing lending rates in equilibrium, it is possible that lenders had initially lowered rates but increased competition for credit amongst borrowers had put upward pressure on bank interest rates, resulting in a null result.

opaque firms, it is possible that these firms gained disproportionately from the reduction in lending rates owing to heightened lender competition after the policy. We examine this channel using the proprietary BSR data and compare average lending rates across underbanked and non-underbanked districts in the post-treatment period (2011), using a cross-sectional sharp RD specification.[55] The choice of this empirical strategy is driven by the absence of private banks in almost half of the districts prior to 2005, resulting in the interest rate being undefined in such instances.

The top-left panel of Appendix Figure A5 provides no evidence of a reduction in lending rates of private banks in underbanked vs non-underbanked districts. This holds for state-owned and private banks (bottom-left panel). The discontinuity estimates are small and statistically insignificant. The absence of a change in district-level lending rates is corroborated in the establishment-level data as well. The ASI collects information on annual interest payments made by establishments (irrespective of credit source), which can be used to impute establishment-level interest rates. Consistent with the district-level results, column (6) of Appendix Table B7 and B12 offer no evidence of a decline in overall lending rates or those faced by small establishments in response to the BAP. This makes it unlikely that the increase in manufacturing investment and credit growth in underbanked districts emanted in response to lower borrowing costs from heightened lender competition.

5.5 Aggregate Demand

The final channel examined is aggregate demand. This is particularly relevant given the findings in Young (2017), who showed that the BAP increased farm productivity and economic activity (measured using night-lights). Thus, financial deepening can boost regional economic activity, resulting in higher local demand through general equilibrium effects, which in turn can boost manufacturing investment. If smaller establishments cater to local demand, this can explain the increased capital spending by smaller manufacturing enterprises in underbanked districts.

To assess whether our results are driven by higher local demand, we explore treatment heterogeneity across tradable and non-tradable industries. If the increase in manufacturing investment is solely an upshot of increased local demand, we would expect the

[55] The administrative BSR data provides information on the weighted average lending rate charged by each bank branch. The average lending rate in the district is computed as the loan volume weighted mean across all branches.

treatment effects to be driven by establishments operating in non-tradable industries. We follow Mian and Sufi (2014) and use the geographic dispersion of industries to classify them as tradable and non-tradable.[56] The intuition is that industries with greater geographic dispersion are likely to be non-tradable.

Column (5) of Appendix Table B10 fails to identify treatment heterogeneity across establishments in industries with relatively low geographic dispersion (tradables). The point estimate is positive but the confidence intervals are sufficiently wide to rule out a significant differential effect. The double-interaction term is positive and significant at 10% level (p-value .057), implying that capital spending in underbanked districts increased by 5 percentage points for establishments operating in non-tradable industries. Thus, while we cannot rule out that part of the increase in capital spending was in response to higher local demand, the positive and statistically significant treatment effect for establishments operating in industries with low geographic dispersion indicates that the aggregate demand channel cannot be the sole explanation for our main findings.

6 Aggregate Effects of Bank Branch Expansion

In this section, we estimate the aggregate effects of financial deepening due to the BAP. This allows us to identify the impact of bank branch expansions on firm entry, closures, and aggregate employment. Section 3.1.1 mentions that our primary sample only includes establishments which were observed at least once before, and after the policy intervention. This limits our ability to identify whether the bank branch expansion affected the entry or exit of establishments. For instance, if bank entry also facilitated the entry of new establishments, resulting in higher capital spending, the current results would be an under-estimate of the true impact of financial deepening on manufacturing investment.

We examine this by aggregating the establishment-level data to the district-industry (3-digit) level and use the following specification:

$$Y_{jdt} = \alpha_d + \delta_{jt} + \beta Underbanked_d \times Post_t + f(Runvar_d) + \gamma \mathbf{X}_{dt} + \epsilon_{jdt} \tag{6}$$

[56]We use data from the Economic Census of 2005 for this exercise. The Economic Census provides total number of workers hired by every business establishment, irrespective of their registration status, allowing us to obtain aggregate estimates of employment at the industry-district level. We first compute the share of manufacturing workers for a specific industry within each district (as share of total manufacturing employment for that industry). We next sum the square of these shares within industries to construct industry-specific measures of geographic dispersion.

The unit of observation in Equation (6) is the district-industry (3 digit-level) pair (dj), observed in year t. We continue to use our difference-in-discontinuity specification and include district fixed effects (α_d) to absorb district-specific time-invariant factors affecting the outcomes of interest. We include industry-year fixed effects (3-digit industry), limiting our comparison to outcomes within similar industry groups in the same year, with the identifying variation arising from changes in districts' underbanked status. The sample is restricted to districts within a bandwidth of 15 from the discontinuity threshold; standard errors are clustered by district.

We find an increase at district-level in the three main outcomes of our interest – capital expenditures, fraction of establishments undertaking any positive capital spending, and credit growth (columns (1)-(3) of Table 9). Capital expenditures are again limited to plant and machinery. The aggregate district-level findings are consistent with the establishment-level results: there is a 5.4 percentage point increase in capital investments, and a 3.3 percentage point increase in the fraction of establishments engaging in capital spending. This corresponds to an aggregate INR 18 million increase in capital investments for establishments in underbanked districts (in the average 3-digit industry) in the post-treatment period. Credit growth too increases by 10 percentage points, signifying an increase in outstanding loans in underbanked districts, equivalent to INR 33 million.

Columns (4)-(6) consider total output, employment and revenue productivity. While the point estimates are positive, the confidence intervals are too wide to rule out a null effect. For employment, the coefficient suggests a 12 percent increase in the hiring of workers, although the coefficient is only significant at 15% level (p-value: .141). Relative to the pre-treatment mean in non-underbanked districts, this amounts to an additional 120 workers employed in the average industry in underbanked districts. Finally, we identify a positive and significant effect on the total number of establishments currently operating (column 7). There is however no impact on the fraction of establishments closed (column 8). In the absence of establishment closures, an increase in total establishments points to higher aggregate entry in underbanked districts in response to the policy intervention. Compared to the dependent variable mean in non-underbanked districts, the coefficient signifies 2 additional establishments operating in each 3-digit industry in underbanked districts in the post-treatment period. Overall, Table 9 shows that the BAP resulted in higher aggregate manufacturing investment and credit growth in underbanked districts. This is accompanied by higher entry of manufacturing establishments. We also find suggestive evidence supporting higher aggregate manufacturing employment in response to an expansion in financial infrastructure and manufacturing credit in these districts.

7 Conclusion

Using firm level panel data from registered manufacturing sector enterprises in India and exploiting a bank branch expansion policy in 2005 that led to an increase in private bank branches and credit in underbanked districts, we find an increase in capital investment and credit growth by manufacturing firms located in the targeted regions. The increase is driven by small and young firms, which are more likely to be credit constrained. At the same time, we do not find a reduction in credit quality. We test various mechanisms that can explain our findings and find evidence in support of two main channels.

First, we find a reduction in physical distance between small borrowers and lenders in underbanked areas after the policy. This is accompanied by greater staffing of officers who are responsible for lending decisions in these bank branches. Given that small borrowers are also more likely to be informationally opaque, our results suggest that increased proximity to banks enabled gathering of soft information on these borrowers and better screening by private banks. Second, we find that the increase is driven by banks which specialize in lending to small manufacturing units. Taken together, these results show that a reduction in physical proximity to banking institutions can reduce frictions in the credit market, specifically for the credit constraint firms, by lowering information and monitoring costs. Thus, increased access to banking has distributional consequences with smaller firms, having higher returns to capital, benefiting more from increased access to bank branches.

References

Sumit Agarwal, Meghana Ayyagari, Yuxi Cheng, and Pulak Ghosh. Road to stock market participation. *Available at SSRN 3897168*, 2021.

Sumit Aggarwal and Robert Hauswald. Distance and private information in lending. *Journal of Financial Economics*, 23(7):2757–2788, 2010.

Sam Asher, Tobias Lunt, Ryu Matsuura, and Paul Novosad. Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in india using the shrug open data platform. *The World Bank Economic Review*, 35(4):845–871, 2021.

Meghana Ayyagari, Asli Demirguc-Kunt, and Vojislav Maximovic. Who creates jobs in developing countries? *Small Business Economics*, 43:75–99, 2014.

John (Jianqui) Bai, Daniel Carvalho, and M. Phillips, Gordon. The impact of bank credit on labour reallocation and aggregate industry productivity. *Journal of Finance*, 73(6): 2787–2836, 2018.

V. Banerjee, Abhijit and Esther Duflo. Giving credit where it is due. *Journal of Economic Perspectives*, 24(3):61–80, 2010.

V. Banerjee, Abhijit and Esther Duflo. Do firms want to borrow more? testing credit constraints using a directed lending program. *The Review of Economic Studies*, 81(2): 572–607, 2014.

Thorsten Beck, Asli Demirguc-Kunt, and Ross Levine. Smes, growth, and poverty: Cross-country evidence. *Journal of Economic Growth*, 10(3):199–229, 2005a. ISSN 13814338, 15737020. URL http://www.jstor.org/stable/40216076.

Thorsten Beck, Asli Demirguc-Kunt, and Vojislav Maximovic. Financial and legal constraints to growth:does firm size matter? *Journal of Finance*, 60(1):137–177, 2005b.

Thorsten Beck, ASLI Demirguc-Kunt, Luc Laeven, and Ross Levine. Finance, firm size, and growth. *Journal of money, credit and banking*, 40(7):1379–1405, 2008.

Allan N. Berger, Nathan H. Miller, and Rajan Raghuram G. Petersen, Mitchell A., and C. Stein, Jeremy. Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial Economics*, 76(2):237–269, 2005.

Allen N Berger and Gregory F Udell. Relationship lending and lines of credit in small firm finance. *Journal of business*, pages 351–381, 1995.

Fabio Berton, S. Presbitero, Andrea, and Matteo Richardi. Banks, firms and jobs. *Review of Financial Studies*, 31(6):2113–2156, 2018.

and Mahajan Aprajit and McKenzie David and Roberts John Bloom, Nicholas. Why firms in developing countries have low productivity? *American Economic Review: Papers and Proceedings 2010*, 100(2):619–623, 2010.

Nicholas Bloom, Aprajit Mahajan, David Mckenzie, and John Roberts. Does management matter? evidence from india. *Quarterly Journal of Economics*, 128(1):1 – 51, 2013.

Nicholas Bloom, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. What drives differences in management practices? *American Economic Review*, 109(5):1648 – 1683, 2019.

M. Bruhn, Dean Karlan, and Antoinnete Schoar. What capital is missing in developing countries? *American Economic Review: Papers Proceedings*, 100:629 – 633, 2010.

M. Bruhn, Dean Karlan, and Antoinnete Schoar. The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in mexico. *Journal of Political Economy*, 126(2):635 – 687, 2018.

Miriam Bruhn and Inessa Love. The real impact of improved access to finance: Evidence from mexico. *The Journal of Finance*, 69(3):1347–1376, 2014.

Robin Burgess and Rohini Pande. Do rural banks matter? Evidence from the Indian social banking experiment. *American Economic Review*, 95(3):780–795, 2005.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210, 2020.

Mark Carlson, Sergio Correia, and Stephen Luck. The effects of banking competition on growth and financial stability: Evidence from the national banking era. *Journal of Political Economy*, 130(2):462–520, 2022.

Indraneel Chakraborty, Apoorva Javadekar, and Rodney Ramcharan. The real effects of bank branching: Evidence from india. *University of Miami Business School Research Paper No. 3805031*, 2021.

Dequi Chen, Yujing Ma, Xiumin Martin, and Roni Michaely. On the fast track: Information acquisition costs and information production. *Journal of Financial Economics*, 143 (2):794–823, 2022.

Yehning Chen, Rachel J Huang, John Tsai, and Larry Y Tzeng. Soft information and small business lending. *Journal of Financial Services Research*, 47(1):115–133, 2015.

Sergey Chernenko, Isil Erel, and Robert Prilmeier. Why do firms borrow directly from non-banks? *Review of Financial Studies*, 35(11):4902–4947, 2022.

Federico Cingano, Marco Leonardi, Julian Messina, and Giovanni Pica. Employment protection legislation, capital investment and access to credit: evidence from italy. *The Economic Journal*, 126(595):1798–1822, 2016.

Ximena Clark, David Dollar, and Alejandro Micco. Port efficiency, maritime transport costs, and bilateral trade. *Journal of development economics*, 75(2):417–450, 2004.

Fe Cramer, Kim. Bank presence and health. *Unpublished Working Paper*, 2022.

Chiara Criscuolo, Ralf Martin, G. Overman, Henry, and John Van Reneen. Some causal effects of an industrial policy. *American Economic Review*, 109(1):48–85, 2019.

Suresh De Mel, David McKenzie, and Christopher Woodruff. Returns to capital in microenterprises: Evidence from a field experiment. *The quarterly journal of Economics*, 123 (4):1329–1372, 2008.

Marcel Fafchamps and Matthias Schündeln. Local financial development and firm performance: Evidence from morocco. *Journal of Development Economics*, 103:15–28, 2013.

Joan Farre-Mensa and Alexander Ljungqvist. Do measures of financial constraints measure financial constraints? *The Review of Financial Studies*, 29(2):271–308, 2016.

Alessandra Fenizia. Managers and productivity in the public sector. *Econometrica*, 90(3): 1063 – 1084, 2022.

Raymond Fisman, Daniel Paravisini, and Vikrant Vig. Cultural proximity and loan outcomes. *American Economic Review*, 107(2):457–492, 2017.

Julia Fonseca and Adrian Matray. The real effects of banking the poor: Evidence from brazil. *NBER Working Paper*, 30057, 2022.

Arturo José Galindo and Alejandro Micco. Creditor protection and credit response to shocks. *The World Bank Economic Review*, 21(3):413–438, 2007.

Erik P Gilje, Elena Loutskina, and Philip E Strahan. Exporting liquidity: Branch banking and financial integration. *The Journal of Finance*, 71(3):1159–1184, 2016.

Xavier Giroud and M. Mueller, Holger. Capital and labour reallocation within firms. *Journal of Finance*, 70(4):1767–1804, 2015.

Greer K. Gosnell, John A. List, and Robert D. Metcalfe. The impact of management practices on employee productivity: A field experiment with airline captains. *Journal of Political Economy*, 128(4):1195 – 1233, 2020.

Jeremy Greenwood, Juan M Sanchez, and Cheng Wang. Financing development: The role of information costs. *American Economic Review*, 100(4):1875–1891, 2010.

Veronica Grembi, Tommaso Nannicini, and Ugo Troiano. Do fiscal rules matter? *American Economic Journal: Applied Economics*, pages 1–30, 2016.

Nandini Gupta and Raghav Dehejia. Financial development and micro-entrepreneurship. *Journal of Financial and Quantitative Analysis*, 57(5):1834–1861, 2021.

Charles J. Hadlock and Joshua R. Pierce. New evidence on measuring financial constraints: Moving beyond the kz index. *Review of Financial Studies*, 23(5):1909–1940, 2010.

Stephan Hollander and Arnt Verriest. Bridging the gap: the design of bank loan contracts and distance. *Journal of Financial Economics*, 119(2):399–419, 2016.

Jith Jayaratne and E. Strahan, Philiip. The finance-growth nexus: Evidence from branch deregulation. *The Quarterly Journal of Economics*, 111(3):639–670, 1996.

Yan Ji, Songyuan Teng, and Robert M Townsend. *Branch expansion versus digital banking: The dynamics of growth and inequality in a spatial equilibrium model*. National Bureau of Economic Research, 2021.

Gaurav Khanna and Priya Mukherjee. Political punishment and financial safety nets: Evidence from india's demonetization. *Unpublished Working Paper*, 2021.

Anjini Kochar. The distributive consequences of social banking: A microempirical analysis of the indian experience. *Economic Development and Cultural Change*, 59(2):251–280, 2011.

Mauricio Larrain and Sebastian Stumpner. Capital account liberalization and aggregate productivity: The role of firm capital allocation. *The Journal of Finance*, 72(4):1825–1858, 2017.

Jose Maria Liberti and A. Petersen, Mitchell. Information: Hard and soft. *Review of Corporate Finance Studies*, 8(1):1–41, 2019.

Kanika Mahajan and Shekhar Tomar. Bridging the miles: Spatial factors in job application and selection. Technical report, Working Paper, 2023.

A. Martin, Leslie, Shanti Nataraj, and E. Harrison, Ann. In with big, out with small: Removing small-scale reservations in india. *American Economic Review*, 107(2):354–386, 2017.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

Atif Mian and Amir Sufi. What explains the 2007-2009 drop in employment? *Econometrica*, 82(6):2197–2223, 2014.

Kaivan Munshi and Mark Rosenzweig. Why is mobility in india so low? social insurance, inequality, and growth. Technical report, National Bureau of Economic Research, 2009.

Q. Nguyen, Hoai-Luu. Are credit markets still local? evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1):1–32, 2019.

A. Petersen, Mitchell and G. Rajan, Raghuram. Does distance still matter? the information revolution in small business lending. *Journal of Finance*, 57(6):2533–2570, 2002.

Mitchell A Petersen and Raghuram G Rajan. The benefits of lending relationships: Evidence from small business data. *The journal of finance*, 49(1):3–37, 1994.

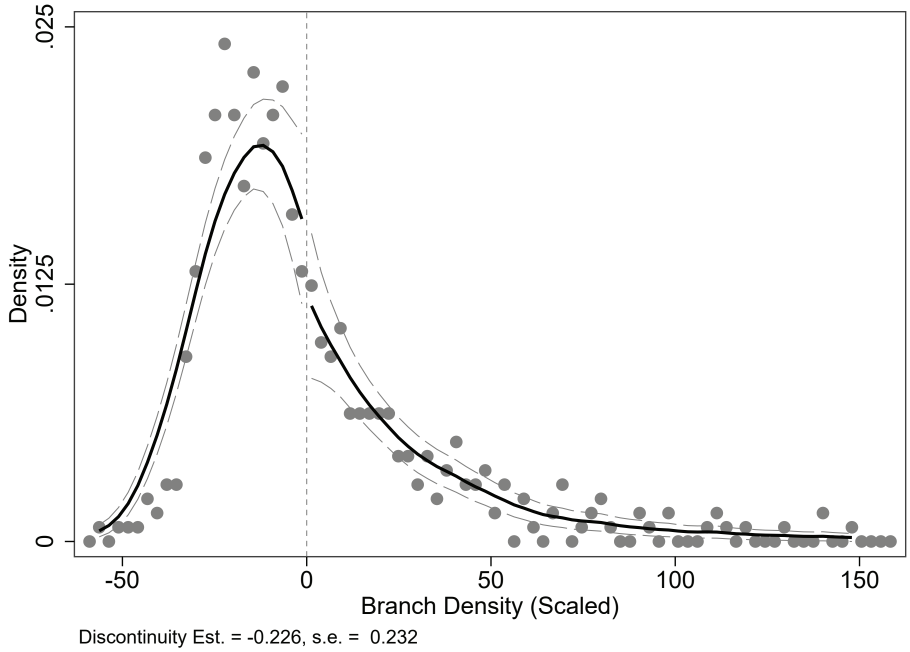
Ni Qin and Dongmin Kong. Access to credit and entrepreneurship: Evidence from china. *Economic Development and Cultural Change*, 71(1):295–331, 2022.

RBI. *Master Circular on Branch Authorisation*. Reserve Bank of India, Mumbai, India, 2005.

RBI. *Master Circular on Branch Authorisation*. Reserve Bank of India, Mumbai, India, 2007.

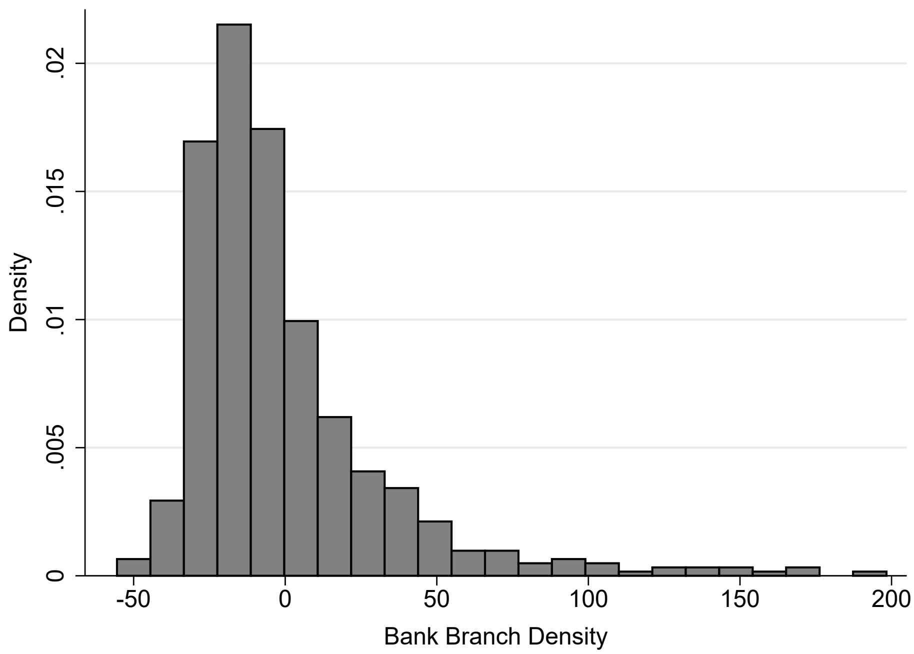
Nathaniel Young. Banking and growth: Evidence from a regression discontinuity analysis. *EBRD Working Paper 2017*, 2017.

Figure 1: Selection of Districts Into Underbanked Status: McCrary Test



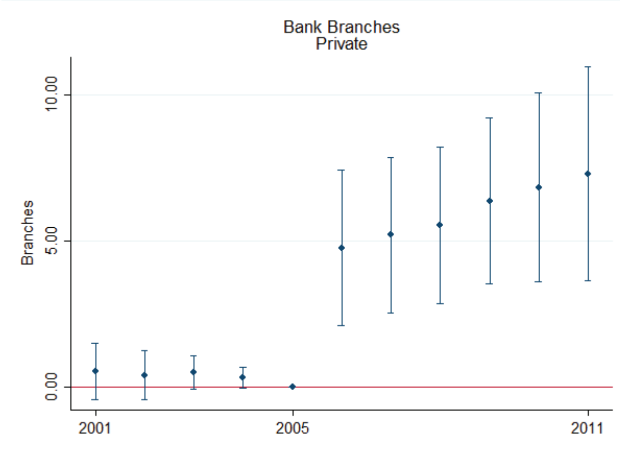
Notes: The above figure tests for bunching of the running variable around the threshold of 0 using the McCrary test (McCrary, 2008). The solid line shows the local polynomial estimate, while the dashed lines show the 95% confidence intervals.

Figure 2: Distribution of Running Variable



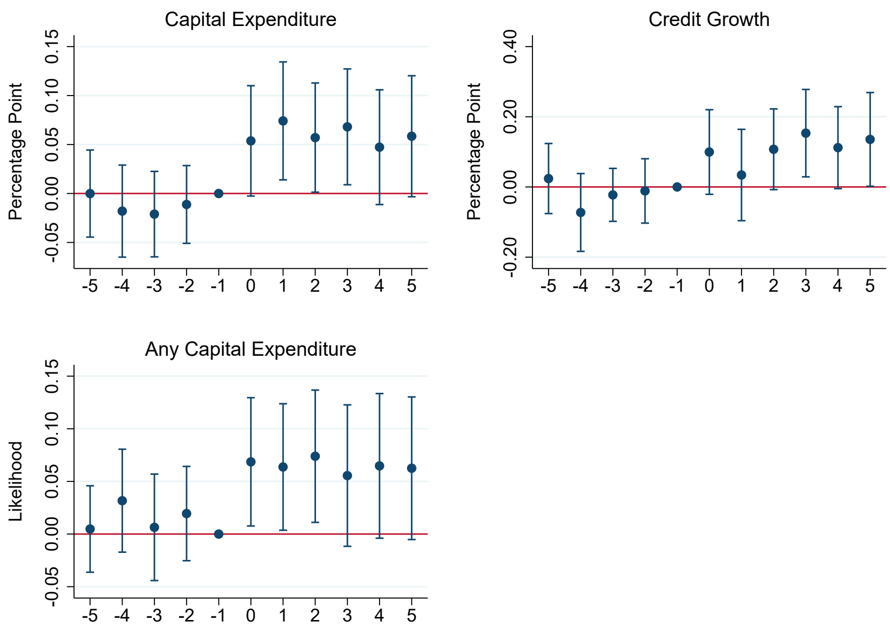
Notes: This figure shows the distribution of the running variable where the running variable is defined at the district level as $Runvar_d = BranchPC_d - \overline{BranchPC}$. *BranchPC* refers to the bank branch density in district *d* in 2005, while $\overline{BranchPC}$ is the national average bank branch density in 2005. Districts are classified as "underbanked" if $Runvar_d < 0$ – located to the left of the threshold 0.

Figure 3: Private Bank Branches in Underbanked Districts: Event Study Plot



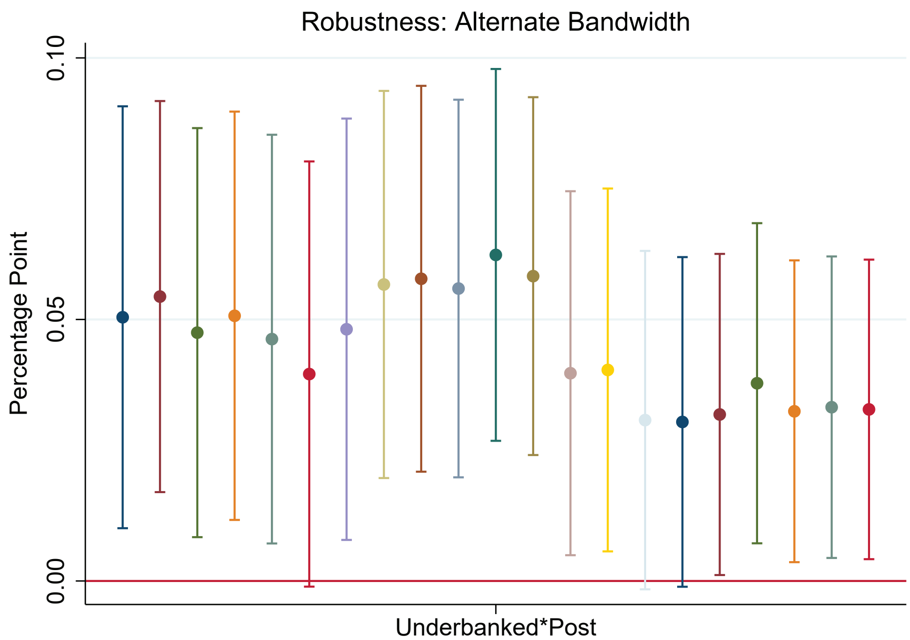
*Notes:*The above figures shows the event-study plots comparing private bank branches across underbanked and non-underbanked districts, before and after the BAP. The unit of observation is the district. The outcome of interest is the number of private bank branches in the district. The year 2005, in which the BAP was introduced, is omitted and forms the reference period. The vertical lines correspond to 95% confidence intervals. All specifications include district and year fixed effects, in addition to district covariates. Standard errors are clustered by district.

Figure 4: Capital Expenditures and Credit Growth in Underbanked Districts: Event-Study Plots



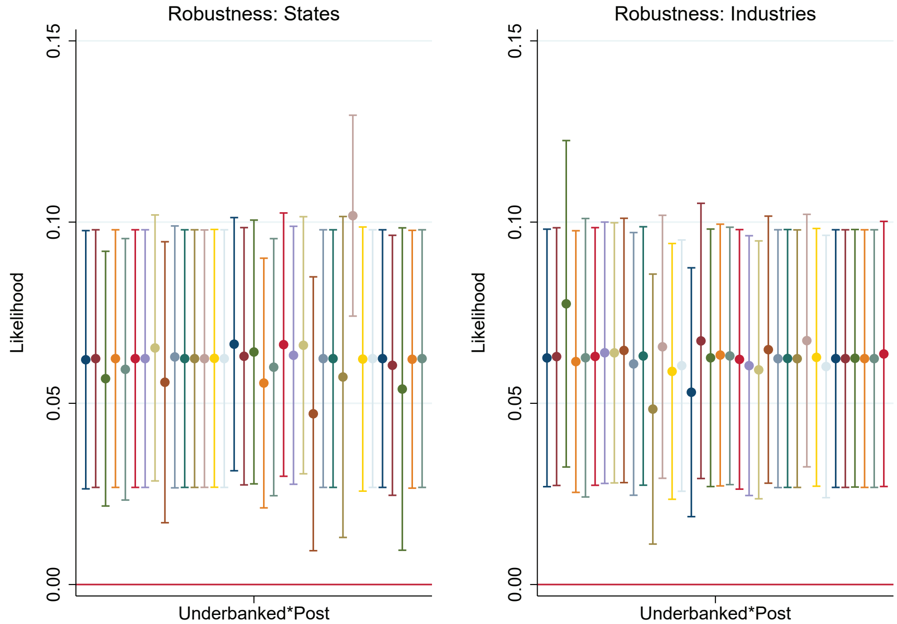
Notes: The above figure presents event-study plots estimating the effect of the BAP on capital expenditures (top left panel), any capital expenditures (bottom left panel), and credit growth (top right panel) in underbanked districts, using a difference-in-discontinuity design. The unit of observation is the manufacturing establishment. Capital expenditures refer to investment in plant and machinery. The solid line represents the average annual treatment effects, and the capped bars denote the 95% confidence intervals. The treatment effects are benchmarked to the year 2005 (omitted time period=-1) – the year in which the treatment is initiated. All specifications include establishment and industry-year fixed effects, a linear polynomial in the running variable interacted with the district underbanked indicator and a post-treatment indicator, establishment and district covariates. All specifications are weighted using establishment-specific weights. The sample in each instance is restricted to districts within a bandwidth of 15 around the discontinuity threshold. Standard errors are clustered by district.

Figure 5: Manufacturing Investment in Underbanked Districts: Robustness to Alternate Bandwidths



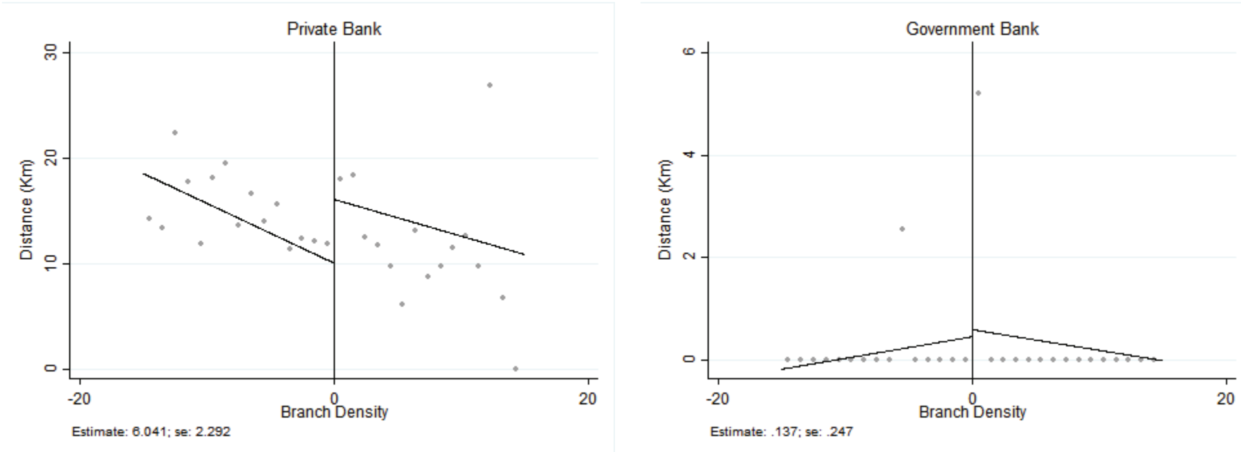
Notes: The above figure shows the robustness of the baseline results to alternate bandwidths. The unit of observation is the manufacturing establishment and the outcome of interest is capital expenditures (investment in plant and machinery). All specifications include establishment and industry-year fixed effects, a linear polynomial in the running variable interacted with the district underbanked indicator and a post-treatment indicator, establishment covariates, and district covariates. All specifications are weighted using establishment-specific weights. The sample in the first specification estimated is restricted to districts within a bandwidth of 10 around the discontinuity threshold; subsequent specifications sequentially increases the bandwidth by 0.5, till the final specification, which equals a bandwidth of 20 around the discontinuity threshold. The vertical lines denote the 95% confidence intervals. Standard errors are clustered by district.

Figure 6: Manufacturing Investment in Underbanked Districts: Robustness to Dropping Individual States and Industries



Notes: The above figures shows the robustness of the baseline results to the dropping of individual states and industries. The unit of observation is the manufacturing establishment and the outcome of interest is capital expenditures (investment in plant and machinery). All specifications include establishment and industry-year fixed effects, a linear polynomial in the running variable, establishment, and district covariates. All specifications are weighted using establishment-specific weights. Specifications are estimated by dropping one state (two-digit industry) at a time. The sample is restricted to districts within a bandwidth of 15 around the discontinuity threshold. The vertical lines denote the 95% confidence intervals. Standard errors are clustered by district.

Figure 7: Minimum Distance of Manufacturing Establishments to Bank Branches



Notes: The above figures compare the minimum distance of business establishments to bank branches across underbanked and non-underbanked districts. The unit of observation is the district. Minimum establishment distances to bank branches are aggregated to the district. Establishment locations are based on the Economic Census of 2013. The sample is restricted to rural manufacturing establishments, hiring at least 5 workers. The left panel shows the minimum distance to private banks; the right panel shows the minimum distance to government-owned banks. The vertical line denotes the national average threshold. Districts to the left of the cutoff are classified as "underbanked". RD coefficients and standard errors, computed using the methodology of Calonico et al. (2020) are shown at the bottom of each figure. RD coefficients are estimated using a linear polynomial and weighted using a triangular kernel. All specifications include state-region fixed effects and the sample is restricted to districts within a bandwidth of 15 around the threshold. Standard errors are clustered by state-region.

Table 1: Bank Branches in Underbanked Districts

	Private Banks		Government Banks		Regional Rural Banks		All Banks	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	4.795***	5.653***	1.608	3.697	.361	.555	6.764	9.906*
	(1.659)	(1.442)	(4.264)	(3.778)	(.739)	(.690)	(5.836)	(5.155)
Observations	2505	2505	2505	2505	2505	2505	2505	2505
R^2	.89	.91	.97	.98	.99	.99	.96	.97
Control Mean	18.81	18.81	111.49	111.49	18.47	18.47	148.77	148.77
Covariates	N	Y	N	Y	N	Y	N	Y

Notes: This table compares average bank branches across underbanked and non-underbanked districts. The unit of observation is the district. Columns (1)-(2) estimate the treatment effect for private banks; columns (3)-(4), government banks; columns (5)-(6), regional rural banks; columns (7)-(8), all banks. All specifications include district and year fixed effects, in addition to a linear polynomial in the running variable. Even-numbered columns also include district covariates. The sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 2: Manufacturing Investment in Underbanked Districts

	Capital Expenditures					
	(1)	(2)	(3)	(4)	(5)	(6)
Underbanked × Post	.046**	.043**	.047**	.062***	.063***	.063***
	(.019)	(.020)	(.020)	(.022)	(.022)	(.022)
Observations	71542	71542	71542	71542	71536	71522
R^2	.37	.38	.38	.38	.38	.39
Control Mean	-.03	-.03	-.03	-.03	-.03	-.03
Firm FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Size-Year FE	N	N	Y	Y	Y	Y
Industry-Year FE	N	Y	Y	Y	Y	Y
Firm Controls	N	N	Y	Y	Y	Y
District Controls	N	N	N	Y	Y	Y
Age FE	N	N	N	N	Y	N

Notes: This table estimates the treatment effect of the BAP on manufacturing investment. The unit of observation is the manufacturing establishment. The outcome of interest is capital expenditures, defined as in (2). Capital expenditures is restricted to capital expenditures in plant and machinery. All specifications include establishment fixed effects, and a linear polynomial in the running variable. Column (1) includes establishment and year fixed effects. Columns (2)-(5) replace year fixed effects with 2-digit industry-year fixed effects, while column (6) considers 3-digit industry year fixed effects. Size-year fixed effects are also included in columns (2)-(6). Size refers to pre-treatment establishment size, measured using the average number of workers hired by the establishment in the pre-treatment period. Column (3) includes establishment specific covariates while column (4) adds in district covariates. Column (5) replaces the quadratic in establishment age with age fixed effects. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 3: Manufacturing Investment in Underbanked Districts: Robustness and Placebo Checks

	Capital Expenditures				
	Unweighted	District-industry Cluster	Exclude Districts	Long Term Effects	Placebo
	(1)	(2)	(3)	(4)	(5)
Underbanked × Post	.039**	.063***	.046*	.062***	
	(.016)	(.021)	(.027)	(.021)	
Underbanked × Post 2001					-.003
					(.030)
Observations	71542	71484	68648	85633	38813
R^2	.26	.38	.38	.35	.47
Control Mean	-.02	-.03	-.04	-.02	-.02

Notes: This table shows robustness of the treatment effect of BAP on manufacturing investment to alternate specifications and placebo tests. The unit of observation is the manufacturing establishment. The outcome of interest is capital expenditures, defined as in (2). Capital expenditures is restricted to those in plant and machinery. All specifications include establishment, 2-digit industry-year fixed effects, a linear polynomial in the running variable, establishment and district covariates. Column (1) excludes establishment weights; column (2) clusters the standard errors by district-industry; column (3) excludes the 9 districts for which the underbanked rule was violated; column (4) extends the sample till the year 2014; column (5) restricts the sample to the years between 1998 and 2005 and considers the period after 2001 to comprise of the post-treatment period. All specifications except column (1) include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district, with the exception of column (2), where they are twoway custered by district and industry. Significant levels: *10%, **5%, and ***1%

Table 4: Manufacturing Investment and Credit Growth in Underbanked Districts: Heterogeneity by Establishment Size, Age and Listing Status

	Capital Expenditures				Credit Growth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	.071**	.081**	.126***	.079***	.254***	.274***	.313***	.144***
	(.032)	(.041)	(.044)	(.023)	(.066)	(.082)	(.085)	(.047)
Underbanked × $Size$ > Median × Post	-.019				-.215***			
	(.039)				(.079)			
Underbanked × $10 > Size \leq 25$ × Post		-.003				-.164		
		(.056)				(.112)		
Underbanked × $25 > Size \leq 50$ × Post		-.005				-.266*		
		(.064)				(.153)		
Underbanked × $50 > Size \leq 100$ × Post		-.061				-.115		
		(.061)				(.135)		
Underbanked × $Size > 100$ × Post		-.061				-.202**		
		(.051)				(.079)		
Underbanked × Large, Young × Post			-.019				-.279**	
			(.055)				(.117)	
Underbanked × Large, Old × Post			-.107**				-.269***	
			(.051)				(.089)	
Underbanked × Small, Old × Post			-.101**				-.103	
			(.050)				(.121)	
Underbanked × Listed × Post				-.137***				-.124
				(.037)				(.098)
Observations	71542	71542	71542	71542	53666	53666	53666	53666
R^2	.38	.38	.38	.38	.34	.34	.34	.34
Control Mean	-.03	-.03	-.03	-.03	.04	.04	.04	.04

Notes: This table estimates the heterogeneity in the treatment impact of BAP on manufacturing investment across establishments which are most likely to be credit constrained. The unit of observation is the manufacturing establishment. The outcome of interest in columns (1)-(4) is manufacturing investment; in columns (5)-(8), credit growth. Growth variables are defined as in equation (2); manufacturing investment is restricted to investments in plant and machinery. All specifications include establishment and industry-year fixed effects, a linear polynomial in the running variable, and establishment and district covariates. *Median* refers to the median establishment size; $Size$ refers to establishment size in the pre-treatment period, defined as the number of workers hired. *Large* and *Small* refer to establishments with above and below median sizes. *Young* refers to establishments which started operations after 1992. *Listed* refers to establishments which are publicly listed. All specifications are weighted using establishment-specific weights. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 5: Minimum Distance of Manufacturing Establishments to Bank Branches

	Private Banks				Government Banks			
	Workers ≥ 1	Workers ≥ 5	5< Workers ≤ 25	Workers > 25	Workers ≥ 1	Workers ≥ 5	5< Workers ≤ 25	Workers > 25
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked	-4.557***	-4.530**	-5.066***	-1.249	1.115***	1.008***	1.226***	-.697
	(1.386)	(1.766)	(1.753)	(1.636)	(.341)	(.350)	(.352)	(.532)
Observations	377694	71467	60691	10768	377694	71467	60691	10768
R^2	.36	.40	.41	.30	.07	.09	.09	.12
Control Mean	13.42	12.54	13.03	9.76	.93	.72	.73	.66

Notes: This table estimates the impact of the BAP on the minimum distance between a manufacturing establishment and a bank branch, using a cross-sectional regression discontinuity design. The unit of observation is the manufacturing establishment in the Economic Census of 2013. The outcome of interest is the minimum distance between the establishment and a private (government) bank branch in columns (1)-(4) [(5)-(8)]. Columns (1) and (5) restrict the sample to establishments that hire at least one worker; columns (2) and (6) exclude establishments hiring less than 5 workers; columns (3) and (7) restrict the sample to establishments hiring between 5 and 25 workers; columns (4) and (8) exclude establishments hiring less than 25 workers. All specifications include establishment ownership controls, state and industry fixed effects, in addition to district-specific pre-treatment covariates. The sample is restricted to ruralestablishments located in districts within a neighbourhood of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 6: Manufacturing Investment and Credit Growth in Underbanked Districts: Heterogeneity by Distance to Private Bank Branch

	Capital Expenditures			Credit Growth		
	All	Small Est.	Large Est.	All	Small Est.	Large Est.
	(1)	(2)	(3)	(4)	(5)	(6)
Underbanked × Post	.076***	.112***	-.003	.138***	.154**	.072
	(.019)	(.026)	(.035)	(.053)	(.067)	(.058)
Underbanked × High Dist. Pvt. Bank × Post	-.126**	-.138*	-.086	-.079	-.093	.072
	(.064)	(.080)	(.084)	(.105)	(.132)	(.153)
Observations	62556	27904	34608	46271	20441	25805
R^2	.38	.42	.29	.34	.41	.25
Control Mean	-.03	-.04	.00	.04	.05	.02

Notes: This table estimates heterogeneity in the impact of BAP on manufacturing investment across distance to private banks. The unit of observation is the manufacturing establishment. The outcome of interest in columns (1)-(3) is manufacturing investment; in columns (4)-(6), credit growth. Growth variables are defined as in equation (2); manufacturing investment is restricted to investments in plant and machinery. All specifications include establishment and industry-year fixed effects, a linear polynomial in the running variable, and establishment and district covariates. *High Dist. Pvt. Bank* refers to districts with high (above median) distance of manufacturing establishments to the nearest private bank branch. Columns (1) and (4) include all establishments; columns (2) and (5) restrict the sample to small establishments; columns (3) and (6) restrict the sample to large establishments. *Small* establishments refer to those hiring under 25 workers; *Large* establishments are those hiring in excess of 50 workers. All specifications are weighted using establishment-specific weights. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 7: Officers and Employees in Underbanked Districts

	Officers		Employees	
	Private Banks	State-Owned Banks	Private Banks	State-Owned Banks
	(1)	(2)	(3)	(4)
Underbanked	35.571***	9.184	52.357***	18.417
	(13.467)	(14.297)	(16.465)	(28.601)
Observations	2505	2505	2505	2505
R^2	.81	.98	.89	.99
Control Mean	97.887	470.792	217.290	1764.064

Notes: This table estimates the impact of the BAP on staffing levels of private and state-owned banks in underbanked districts. The unit of observation is the district. All specifications include district and year fixed effects, along with a linear polynomial in the running variable, and district covariates. The sample is restricted to districts within a neighbourhood of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 8: Manufacturing Investment and Credit Growth in Underbanked Districts: Heterogeneity by Establishment Size, Age and Listing Status

	Capital Expenditures				Credit Growth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	.071**	.081**	.126***	.079***	.254***	.274***	.313***	.144***
	(.032)	(.041)	(.044)	(.023)	(.066)	(.082)	(.085)	(.047)
Underbanked × $Size >$ Median × Post	-.019				-.215***			
	(.039)				(.079)			
Underbanked × $10 > Size \leq 25$ × Post		-.003				-.164		
		(.056)				(.112)		
Underbanked × $25 > Size \leq 50$ × Post		-.005				-.266*		
		(.064)				(.153)		
Underbanked × $50 > Size \leq 100$ × Post		-.061				-.115		
		(.061)				(.135)		
Underbanked × $Size > 100$ × Post		-.061				-.202**		
		(.051)				(.079)		
Underbanked × Large, Young × Post			-.019				-.279**	
			(.055)				(.117)	
Underbanked × Large, Old × Post			-.107**				-.269***	
			(.051)				(.089)	
Underbanked × Small, Old × Post			-.101**				-.103	
			(.050)				(.121)	
Underbanked × Listed × Post				-.137***				-.124
				(.037)				(.098)
Observations	71542	71542	71542	71542	53666	53666	53666	53666
R^2	.38	.38	.38	.38	.34	.34	.34	.34
Control Mean	-.03	-.03	-.03	-.03	.04	.04	.04	.04

Notes: This table estimates the heterogeneity in the treatment impact of BAP on manufacturing investment across establishments which are most likely to be credit constrained. The unit of observation is the manufacturing establishment. The outcome of interest in columns (1)-(4) is manufacturing investment; in columns (5)-(8), credit growth. Growth variables are defined as in equation (2); manufacturing investment is restricted to investments in plant and machinery. All specifications include establishment and industry-year fixed effects, a linear polynomial in the running variable, and establishment and district covariates. *Median* refers to the median establishment size; $Size$ refers to establishment size in the pre-treatment period, defined as the number of workers hired. *Large* and *Small* refer to establishments with above and below median sizes. *Young* refers to establishments which started operations after 1992. *Listed* refers to establishments which are publicly listed. All specifications are weighted using establishment-specific weights. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table 9: Aggregate Effects of Bank Branch Expansion

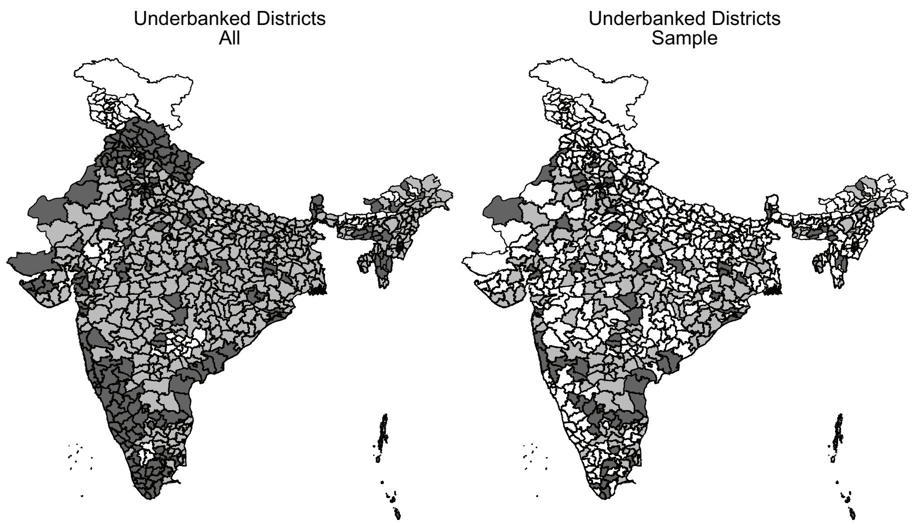
	Capital Expenditure (1)	Any Capital Expenditure (2)	Credit Growth (3)	Output (Log) (4)	Workers (Log) (5)	Revenue TFP (6)	Total Establishments (Log) (7)	Fraction Closed (8)
Underbanked×Post	.054**	.033	.103**	.018	.121	.032	.140**	-.032
	(.024)	(.023)	(.045)	(.087)	(.082)	(.040)	(.057)	(.026)
Observations	17962	18152	15908	18015	17898	16128	19388	19388
R^2	.07	.12	.06	.36	.35	.10	.33	.19
Control Mean	.03	.28	.07	1586.71	.97	-.03	14.59	.07

Notes: This table estimates the impact of BAP on aggregate district-level outcomes. The unit of observation is the district-industry (3-digit). The outcomes of interest in columns (1) and (3) are constructed using equation (2). The outcomes of interest in columns (2) and (8) are fractions. The remaining outcomes are log transformed. Capital expenditures refer to expenditures in plant and machinery. All specifications include district and 3-digit industry-year fixed effects, in addition to district-specific controls. Standard errors are in parentheses, clustered by district.
Significant levels: *10%, **5%, and ***1%

A Branch Authorisation Policy and Banking Outcomes

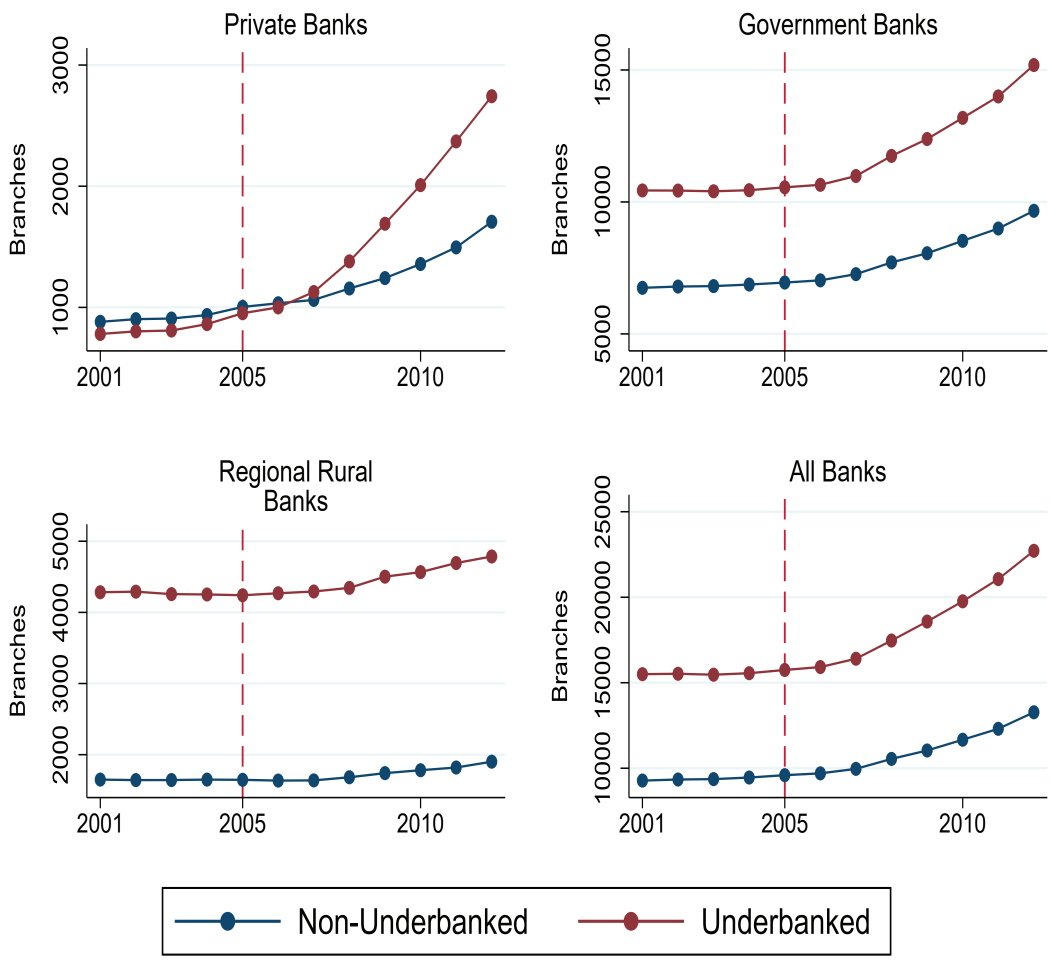
A.1 Figures

Figure A1: Geographical Distribution of Underbanked Districts



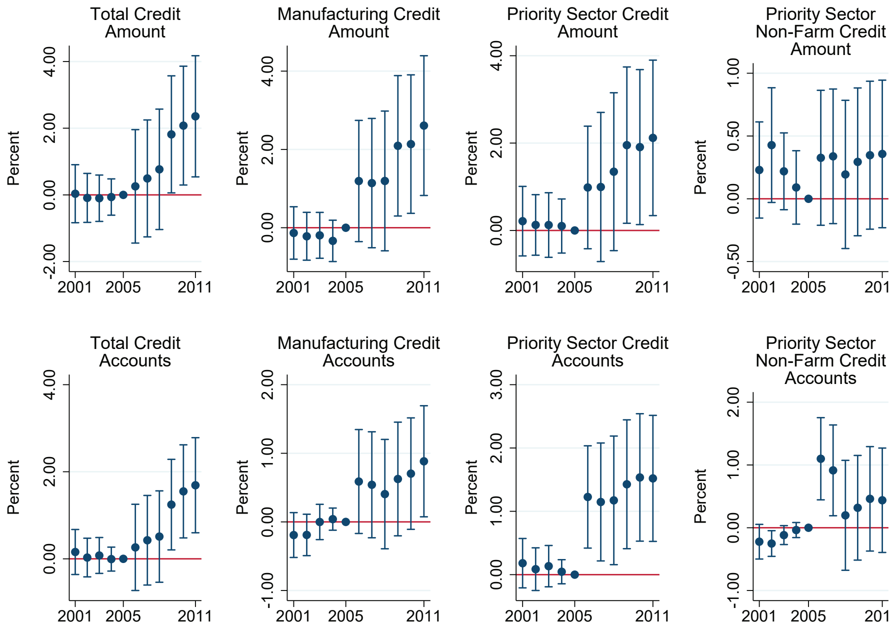
Notes: This figure shows the geographical distribution of underbanked districts. The left-panel shows all districts; the right panel shows districts which form our primary sample, which are located within a bandwidth of 15 bank branches per capita from the discontinuity threshold. The darker shades depict underbanked districts; lighter shades depict non-underbanked districts. Districts in white are excluded from the sample.

Figure A2: Bank Branches in Underbanked Districts: Unconditional Trends



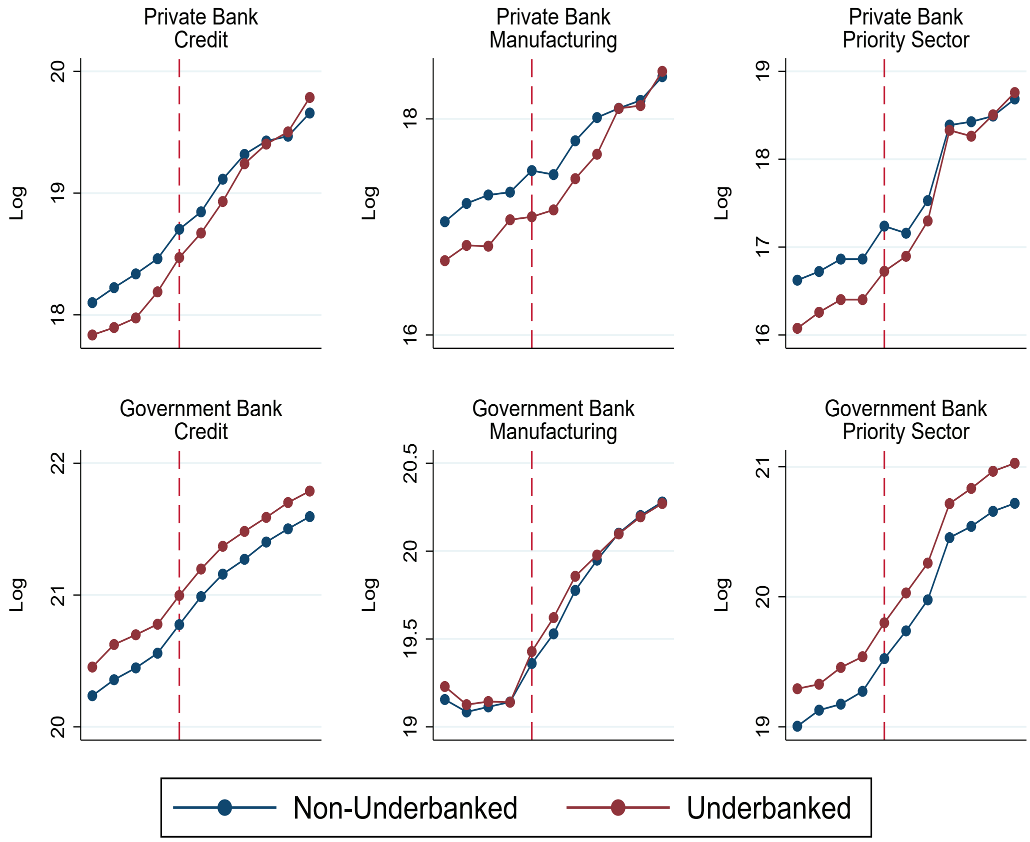
*Notes:*The above figure presents unconditional trends in the number of bank branches across underbanked and non-underbanked districts between 2001 and 2011. The sample is restricted to districts within a bandwidth of 15 around the discontinuity threshold. The red vertical line corresponds to the year 2005. The top-left panel has private bank branches; the top right panel, government bank branches; the bottom-left panel, regional rural banks; the bottom-right panel, total bank branches.

Figure A3: Private Bank Credit in Underbanked Districts: Event-Study Plots



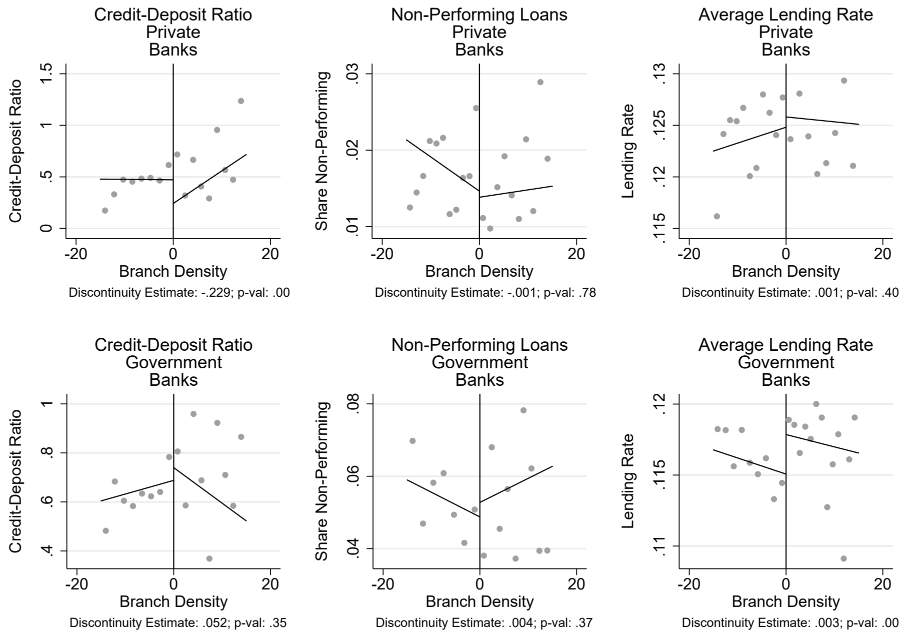
Notes: The above figures shows event-study plots comparing financial intermediation by private banks across underbanked and non-underbanked districts, before and after the BAP. The unit of observation is the district. The outcome of interest in the top-panel is outstanding private bank loan amounts for the categories mentioned; in the bottom panel, loan accounts. The outcome of interest is inverse hyperbolic sine transformed. The year 2005, in which the BAP was introduced, is omitted and forms the reference period. The vertical lines correspond to 95% confidence intervals. All specifications include district and year fixed effects, in addition to district covariates. The sample is restricted to a bandwidth of 15 around the discontinuity threshold. Standard errors are clustered by district.

Figure A4: Credit Disbursement in Underbanked Districts: Unconditional Trends



Notes: The above figures present unconditional trends in logged credit disbursement across underbanked and non-underbanked districts for select credit categories. The vertical line corresponds to the year 2005 in which the BAP was introduced. The sample is restricted to districts within a bandwidth of 15 around the discontinuity threshold. The top panel corresponds to private banks; the bottom panel considers government banks. The left-panel considers aggregate credit disbursement by banks, inclusive of loans to farm, manufacturing, trade and service sectors, along with personal loans to households.

Figure A5: Banking Outcomes in Underbanked Districts: RD Plots



Notes: This figure shows RD plots comparing credit-deposit ratios, lending rates and non-performing loans across underbanked and non-underbanked districts. The vertical line denotes the national average threshold. Districts to the left of the cutoff are classified as "underbanked". RD coefficients and standard errors are computed using the methodology of Calonico et al. (2020). RD coefficients are estimated using a linear polynomial and weighted using a triangular kernel. All specifications include state-region fixed effects and the sample is restricted to districts within a bandwidth of 15 around the threshold. Credit-deposit ratios is the amount of private bank credit in the district, scaled by total private bank deposits. Loan volume weighted average lending rates are for the year 2011; non-performing loans for 2016. Non-performing loans are expressed as a share of total outstanding loans. The top row corresponds to private banks; the bottom row, government banks. Standard errors are clustered by state-region.

A.2 Tables

Table A1: Credit Disbursement in Underbanked Districts: By Sectors and Bank Groups

Panel A: Private banks

	Total		Manufacturing		Priority Sector		Micro and Small Credit	
	Amt.	Act.	Amt.	Act.	Amt.	Act.	Amt.	Act.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	1.339	.896	1.904**	.693*	1.437*	1.249**	1.468	.697*
	(.948)	(.563)	(.870)	(.390)	(.851)	(.486)	(.897)	(.375)
Observations	2505	2505	2505	2505	2505	2505	2505	2505
R^2	.86	.89	.85	.89	.86	.90	.81	.85
Control Mean	6339.910	.035	2306.507	.001	539.630	.002	219.591	.000

Panel B: Government banks

	Total		Manufacturing		Priority Sector		Micro and Small Credit	
	Amt.	Act.	Amt.	Act.	Amt.	Act.	Amt.	Act.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	-.100	-.017	.020	-.187*	-.040	.041	.116	-.015
	(.066)	(.064)	(.153)	(.113)	(.093)	(.088)	(.290)	(.153)
Observations	2505	2505	2505	2505	2505	2505	2505	2505
R^2	.99	.99	.96	.92	.98	.98	.90	.95
Control Mean	23899.050	.094	9201.978	.006	4341.187	.032	1550.907	.001

Notes: This table shows credit disbursement across underbanked and non-underbanked districts in the post-treatment period. The unit of observation is the district. Panel A considers private bank credit; panel B credit from government banks. Outcome variables are transformed using an inverse hyperbolic sine transformation. Odd-numbered columns show bank credit along the intensive margin (outstanding loans); even-numbered columns, bank credit along the extensive margin (total loan accounts). The outcome of interest in columns (1) and (2) is total credit; in columns (3) and (4), manufacturing credit; in columns (5) and (6), loans to the priority sector; and columns (7) and (8), loans to micro and small enterprises. All specifications include district and year fixed effects, in addition to a linear polynomial in the running variable and district covariates. The sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Control variable means for outstanding loan amounts are in millions. Standard errors in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

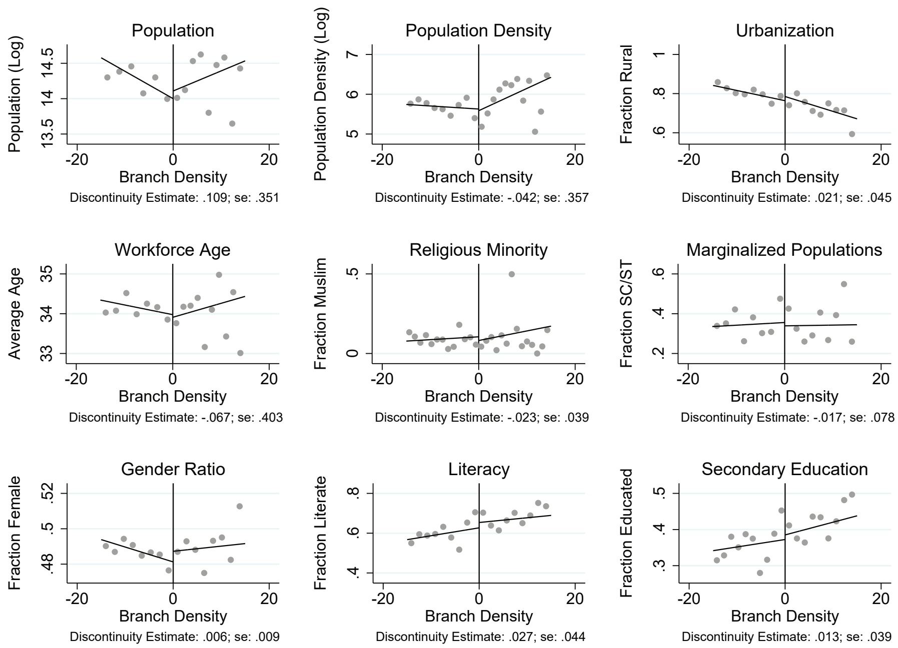
Table A2: Private Bank Credit in Underbanked Districts: Levels Specification, by Sectors and Bank Groups

	Total		Manufacturing		Priority Sector		Micro and Small Credit	
	Amt.	Act.	Amt.	Act.	Amt.	Act.	Amt.	Act.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	1257.284*	5515.649	278.350	263.905*	422.701	1110.599	194.753	144.439
	(737.419)	(4081.622)	(218.080)	(159.700)	222.665)	(2060.735)	(153.619)	(182.034)
Observations	2505	2505	2505	2505	2505	2505	2505	2505
R^2	.80	.75	.77	.52	.64	.78	.58	.48
Control Mean	6339.910	.035	2306.507	.001	539.630	.002	219.591	.000

Notes: This table credit disbursement across underbanked and non-underbanked districts in the post-treatment period. The unit of observation is the district. Outcome variables are in levels (millions of rupees). Odd-numbered columns show private bank credit along the intensive margin (outstanding loans); even-numbered columns, private bank credit along the extensive margin (total loan accounts). The outcome of interest in columns (1) and (2) is total credit; in columns (3) and (4), manufacturing credit; in columns (5) and (6), loans to the priority sector; and columns (7) and (8), loans to micro and small enterprises. All specifications include district and year fixed effects, in addition to a linear polynomial in the running variable and district covariates. The sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

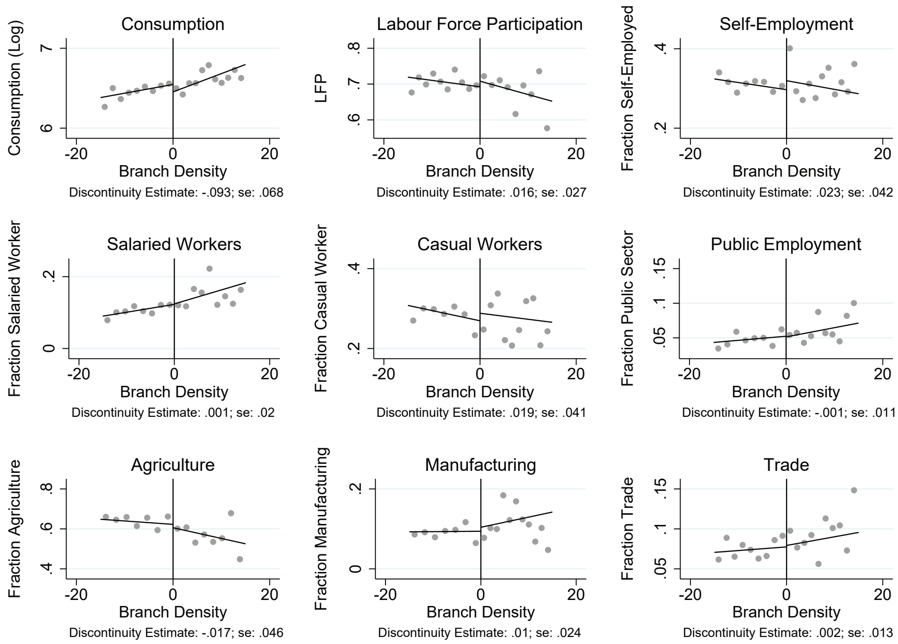
B Appendix: Additional Figures and Tables

Figure B1: Pre-Treatment Covariate Balance: District Demographics



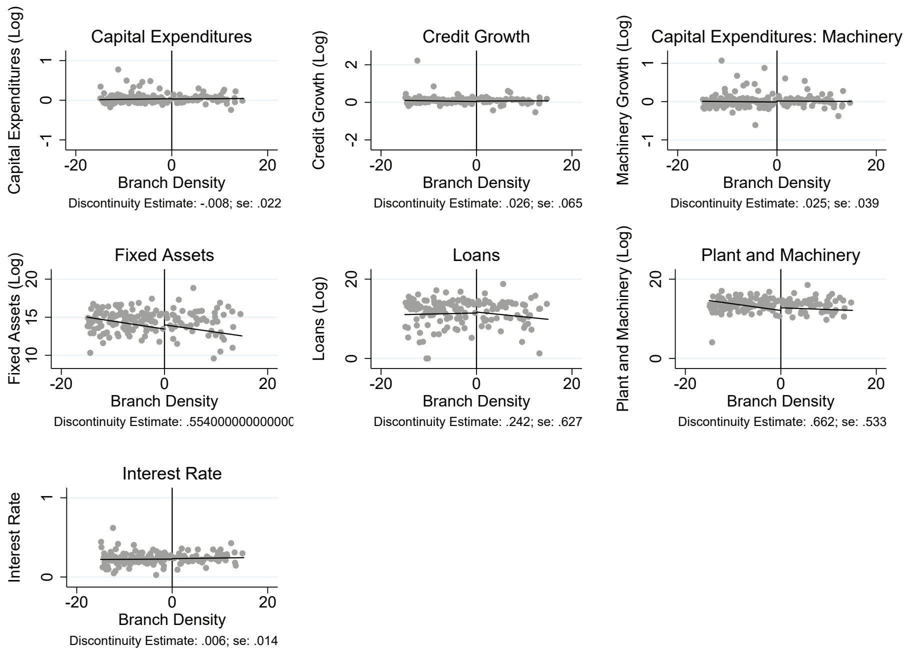
Notes: The above figure shows the pre-treatment covariate balance across district-specific demographic covariates. The vertical line denotes the national average threshold. Districts to the left of the cutoff are classified as "underbanked". The sample is restricted to a bandwidth of 15 around the threshold. The discontinuity estimates and standard errors corresponding to each figure are computed as outlined in Calonico et al. (2020).

Figure B2: Pre-Treatment Covariate Balance: District Demographics



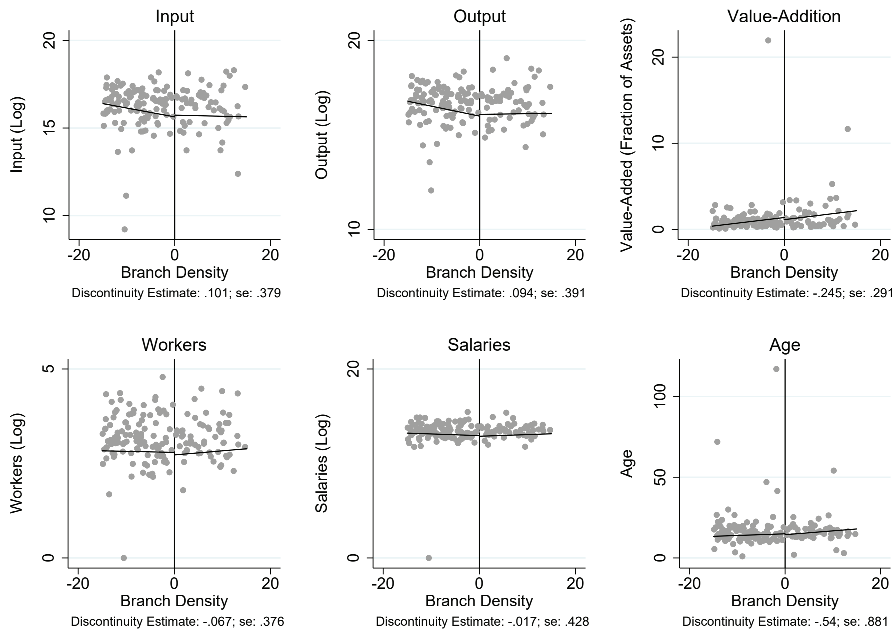
Notes: The above figure shows the pre-treatment covariate balance across district-specific demographic covariates. The vertical line denotes the national average threshold. Districts to the left of the cutoff are classified as “underbanked”. Consumption refers to households monthly per capita consumption. The sample is restricted to a bandwidth of 15 around the threshold. The discontinuity estimates and standard errors corresponding to each figure are computed as outlined in Calonico et al. (2020).

Figure B3: Pre-Treatment Covariate Balance: Manufacturing Enterprise Financial Characteristics



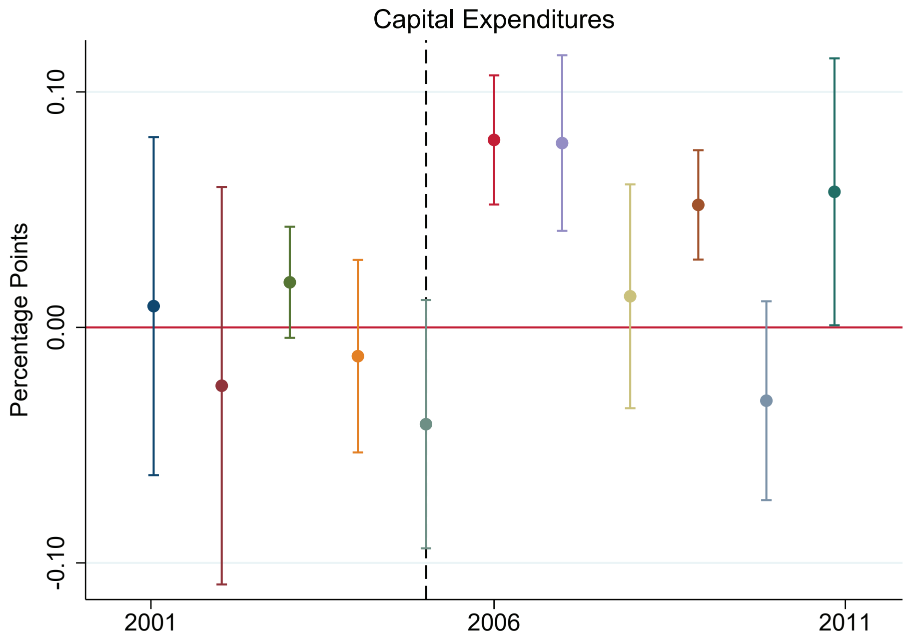
Notes: The above figure shows the pre-treatment balance across manufacturing enterprise characteristics. The vertical line denotes the national average threshold. Districts to the left of the cutoff are classified as "underbanked". The sample is restricted to a bandwidth of 15 around the threshold. The discontinuity estimates and standard errors corresponding to each figure are computed as outlined in Calonico et al. (2020).

Figure B4: Pre-Treatment Covariate Balance: Manufacturing Enterprise Labor and Production Characteristics



Notes: The above figure shows the pre-treatment balance across manufacturing enterprise characteristics. The vertical line denotes the national average threshold. Districts to the left of the cutoff are classified as "underbanked". Consumption refers to households monthly per capita consumption. The sample is restricted to a bandwidth of 15 around the threshold. The discontinuity estimates and standard errors corresponding to each figure are computed as outlined in Calonico et al. (2020).

Figure B5: Capital Investment in Underbanked Districts: Annual Cross-Sectional RD Estimates



Notes: The above figure shows the treatment effect on capital investment using a cross-sectional RD specification for each year. The first coefficient estimate corresponds to the year 2001; the last coefficient estimate, 2011. The unit of observation is the manufacturing establishment. The outcome of interest is capital expenditures, defined as in specification (2). Capital expenditures is restricted to plant and machinery (net of depreciation). All specifications include state and 3-digit industry fixed effects, in addition to establishment and district covariates. A linear polynomial in the running variable is also included. The estimates are weighted using a triangular kernel and establishment-specific multiplier weights. MSERD bandwidths, computed using the procedure outlined in Calonico et al. (2020), are used to compute the optimal bandwidth for each year. The broken vertical line denotes the year 2005, the year of introduction of the BAP. Standard errors are clustered by district. Vertical lines denote 95% confidence intervals.

Table B1: Summary Statistics: Manufacturing Establishments

	N	Mean	SD	P25	P50	P75
Capital Expenditures – Machinery	69071	-0.002	0.398	-0.162	-0.105	0.030
Any Capital Expenditure – Machinery	70285	0.269	0.443	0.000	0.000	1.000
Capital Expenditures	70285	0.023	0.328	-0.129	-0.066	0.060
Any Capital Expenditure	70285	0.318	0.466	0.000	0.000	1.000
Loan Growth	54066	0.042	0.753	-0.207	0.000	0.266
Any Loan Growth	69891	0.383	0.486	0.000	0.000	1.000
New Loan	70285	0.025	0.157	0.000	0.000	0.000
No Loan	69891	0.226	0.418	0.000	0.000	0.000
Interest Rate	54046	0.244	0.280	0.071	0.144	0.282
Plant and Machinery (INR)	70285	28.679	102.425	0.195	1.204	8.082
Fixed Assets (INR)	70285	40.335	124.906	0.799	3.241	15.985
Raw Materials (INR)	57602	15.970	45.519	0.618	2.459	8.999
Land and Buildings (INR)	70285	11.709	34.696	0.203	1.120	5.353
Assets (INR)	69890	127.688	383.564	4.284	14.663	60.823
Loans (INR)	69891	27.442	92.475	0.096	2.076	10.479
Hired Workers	70285	89.372	481.087	8.000	20.000	63.000
Contract Workers	70240	26.975	354.670	0.000	0.000	5.000
Supervisors	70240	10.029	81.256	1.000	2.000	6.000
Salaries – Hired Workers (INR)	70285	4.693	11.952	0.298	0.829	2.921
Salaries – Contract Workers (INR)	70240	0.940	3.069	0.000	0.000	0.174
Salaries – Supervisor (INR)	70240	2.442	7.487	0.057	0.235	1.119
Output (INR)	70285	182.752	480.885	5.877	23.301	109.496
Value-Addition (INR)	70285	35.866	106.813	1.216	3.850	16.743
Value-Addition (Share of Assets)	70285	0.820	2.612	0.186	0.322	0.570
Rural	70285	0.438	0.496	0.000	0.000	1.000
Computer Use	70285	0.592	0.492	0.000	1.000	1.000
Age	70285	17.841	14.709	8.000	14.000	23.000
Young Establishment	70285	0.449	0.497	0.000	0.000	1.000
Micro-Enterprise	70050	0.649	0.477	0.000	1.000	1.000
Small Enterprise	70050	0.272	0.445	0.000	0.000	1.000
Medium Enterprise	70050	0.029	0.168	0.000	0.000	0.000
Large Enterprise	70050	0.050	0.218	0.000	0.000	0.000
Small-Scale Industries	70050	0.808	0.394	1.000	1.000	1.000

Notes: This table shows the summary statistics for registered manufacturing establishments. The sample is restricted to establishments situated in districts located within a bandwidth of 15 around the discontinuity threshold. Rupee values are in constant 2005 millions of rupees. Growth variables are defined as in equation (2). Micro, small, small-scale, medium and large enterprises are defined according to administrative definitions.

Table B2: Covariate Balance Across Pre-Treatment District Demographic Covariates

	Population (Log)	Population Denisty (Log)	Fraction Rural	Workforce Age	Fraction Muslim	Fraction Marginalized	Fraction Female	Fraction Literate	Fraction Educated
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Underbanked	.339	.089	-.024	.389	-.046*	-.007	.002	-.041	-.006
	(.242)	(.171)	(.035)	(.428)	(.025)	(.045)	(.008)	(.027)	(.031)
Observations	228	228	228	228	228	228	228	228	228
R^2	.83	.86	.66	.58	.91	.82	.52	.84	.78
Control Mean	2.106	543.41	.74	34.11	.11	.35	.49	.67	.41

Notes: This table shows the pre-treatment covariate balance across district-level demographic covariates. The unit of observation is the district. *Underbanked* is a dummy equaling 1 if the district's per capita bank branch density in 2005 was less than the national average bank branch density. The sample is restricted to districts within a bandwidth of 15 on either side of the national average threshold. Workforce age is the average age of workers in the district; marginalized castes refer to the fraction of *Dalits* (Scheduled Castes) and *Adivasis* (Scheduled Tribes) in the district; educated refers to the fraction of adults with secondary or higher education. All specifications include a linear polynomial in the running variable and its interaction with a district's underbanked status. All specifications include state fixed effects and standard errors are clustered by state-region. Significant levels: *10%, **5%, and ***1%.

Table B3: Covariate Balance Across Pre-Treatment District Economic Covariates

	LFP	Fraction Self-Employed	Fraction Salaried Workers	Fraction Casual Workers	Fraction Farm Activities	Fraction Manufacturing Activities	Fraction Trade Activities	Fraction Public Employment	Per Capita Consumption (Log)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Underbanked	.018	.007	-.026	.004	.006	.012	.004	-.007	.013
	(.027)	(.021)	(.020)	(.033)	(.042)	(.022)	(.013)	(.009)	(.063)
Observations	228	228	228	228	228	228	228	228	228
R^2	.62	.82	.60	.74	.66	.55	.55	.67	.77
Control Mean	.69	.31	.15	.28	.57	.11	.09	.06	752.72

Notes: This table shows the pre-treatment covariate balance across district-level economic covariates. The unit of observation is the district. *Underbanked* is a dummy equaling 1 if the district's per capita bank branch density in 2005 was less than the national average bank branch density. The sample is restricted to districts within a bandwidth of 15 on either side of the national average threshold. All specifications include a linear polynomial in the running variable and its interaction with a district's underbanked status. Per capita consumption is the district's average household monthly per capita consumption. All specifications include state fixed effects and standard errors are clustered by state-region. Significant levels: *10%, **5%, and ***1%.

Table B4: Pre-Treatment Balance of Manufacturing Characteristics Across Underbanked Districts

	Capital Expenditures (Log) (1)	Any Capital Expenditures (2)	Plant Machinery Investment (Log) (3)	Credit Growth (Log) (4)	Interest Rate (5)
Underbanked	.004	-.039	-.025	-.014	-.001
	(.022)	(.027)	(.026)	(.029)	(.021)
Observations	21110	22079	20620	16044	17077
R^2	.01	.05	.01	.00	.02

Notes: This table shows balance across pre-treatment manufacturing characteristics using the ASI data. The unit of observation is the manufacturing establishment. All specifications include state and 2-digit industry year fixed effects, along with a linear polynomial in the running variable. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table B5: Pre-Treatment Balance of Manufacturing Characteristics Across Underbanked Districts

	Fixed Assets (Log) (1)	Plant Machinery (Log) (2)	Loans (Log) (3)	Input (Log) (4)	Workers (Log) (5)	Salaries (Log) (6)	Output (Log) (7)	Value-Addition (8)	Age (9)
Underbanked	-.595	-.398	-.818	-.242	.051	-.019	-.149	.444	.022
	(.447)	(.510)	(.712)	(.253)	(.131)	(.174)	(.208)	(.444)	(1.004)
Observations	22079	21462	21962	22079	22079	22079	22079	22079	21321
R^2	.15	.15	.16	.16	.08	.09	.14	.06	.08

Notes: This table shows balance across pre-treatment manufacturing characteristics using the ASI data. The unit of observation is the manufacturing establishment. All specifications include state and 2-digit industry year fixed effects, along with a linear polynomial in the running variable. Value-addition is defined as establishment value-addition, scaled by establishment assets. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and × ***1%

Table B6: Manufacturing Investment in Underbanked Districts: Alternate Outcome Variables

	Capital Expenditures: Plant and Machinery		Capital Expenditures			Raw Materials	
	Log Difference	Pr(Any Capex = 1)	Capital Expenditure	Log Difference	Pr(Any Capex = 1)	Expenditure	Log Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Underbanked × Post	.075***	.051**	.032*	.036**	.043*	.113***	.149***
	(.026)	(.024)	(.017)	(.017)	(.025)	(.040)	(.055)
Observations	71542	71542	71542	71542	71542	57673	57673
R^2	.37	.44	.38	.38	.45	.35	.35
Control Mean	-.003	.233	.003	.013	.268	.070	3.438

Notes: This table shows the robustness of the baseline specification to alternate functional forms and outcome variables. The unit of observation is the manufacturing establishment. All specifications include establishment, industry year fixed effects, and establishment size-year fixed effects, in addition to a linear polynomial in the running variable, establishment and district covariates. Columns (1) and (2) restricts capital expenditures to capital expenditures in plant and machinery; capital expenditures in columns (3)-(5) is net fixed assets; expenditures in columns (6)-(7) refers to raw materials. Columns (1), (4) and (7) measure capital expenditures as the logged difference in closing and opening values; the outcome in columns (2) and (5) is a dummy equaling 1 if the establishment undertook any positive capital spending during the year. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table B7: Credit Growth for Manufacturing Establishments in Underbanked Districts

	Credit Growth	Credit Growth (Log)	Any Credit Growth	New Loan	No Loan	Interest Rate
	(1)	(2)	(3)	(4)	(5)	(6)
Underbanked × Post	.128***	.600***	-.013	.011	.009	.013
	(.043)	(.139)	(.027)	(.008)	(.018)	(.018)
Observations	53666	71138	71138	71542	71138	53645
R^2	.34	.33	.44	.35	.71	.58
Control Mean	.043	.188	.381	.025	.224	.241

Notes: This table estimates the treatment effect on manufacturing credit growth. The unit of observation is the manufacturing establishment. All specifications include establishment, industry-year fixed effects, a linear polynomial in the running variable, district and establishment covariates. The outcome of interest in column (1) is credit growth, defined as in equation (2); in column (2), logged difference in closing and opening values of outstanding loans; column (3), a dummy equaling 1 if the closing value of loans exceeded the opening value; column (4), a dummy equaling 1 if the establishment had no outstanding loans through the year; in column (5), a dummy equaling 1 if the establishment had no outstanding credit at the beginning of the accounting period, but positive outstanding loans at the year-end; in column (6), the imputed interest rate. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table B8: Manufacturing Investment in Underbanked Districts: Heterogeneity by Establishment and Industry Characteristics

	Capital Expenditures				
	(1)	(2)	(3)	(4)	(5)
Underbanked × Post	.084**	.068**	.076***	.087**	.065*
	(.035)	(.030)	(.026)	(.038)	(.039)
Underbanked × High Capital × Post	-.035				
	(.040)				
Underbanked × $Small$ × Post		-.014			
		(.041)			
Underbanked × $Medium$ × Post		-.149**			
		(.058)			
Underbanked × $Large$ × Post		-.088			
		(.054)			
Underbanked × Non-SSI × Post			-.101**		
			(.047)		
Underbanked × High Collateral × Post				-.042	
				(.050)	
Underbanked × Partnership × Post					-.005
					(.047)
Underbanked × Private Ent. × Post					.036
					(.047)
Underbanked × Govt. × Post					.020
					(.103)
Underbanked × Listed × Post					-.094**
					(.047)
Observations	71542	71280	71280	71417	71542
R^2	.38	.38	.38	.38	.38
Control Mean	-.03	-.03	-.03	-.03	-.03

Notes: This table estimates the treatment heterogeneity on manufacturing investment across establishment fixed assets, tangibility, ownership and listing status. The unit of observation is the manufacturing establishment. The outcome of interest is capital expenditures. Capital expenditures refers to capital spending on plant and machinery, defined as in equation (2). All specifications include establishment and industry year fixed effects, along with a linear polynomial in the running variable, as well as establishment and district covariates. *High Capital* refers to establishments whose fixed assets exceed the median pre-treatment fixed assets. Administrative definitions are used to classify establishments as *Small*, *Medium*, *Large* and *SSI* (Small Scale Industries), based on their pre-treatment establishment fixed capital. *High Collateral* refers to establishments whose value of land and buildings exceed the pre-treatment median. *Listed* is a dummy for establishments owned by publicly listed corporations. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and × ***1%

Table B9: Credit Growth in Underbanked Districts: Heterogeneity by Establishment and Industry Characteristics

	Credit Growth				
	(1)	(2)	(3)	(4)	(5)
Underbanked × Post	.193**	.164***	.171***	.170**	.208**
	(.084)	(.058)	(.046)	(.067)	(.083)
Underbanked × High Capital × Post	-.089				
	(.101)				
Underbanked × *Small* × Post		-.045			
		(.077)			
Underbanked × *Medium* × Post		-.223			
		(.164)			
Underbanked × *Large* × Post		-.131*			
		(.076)			
Underbanked × Non-SSI × Post			-.178***		
			(.067)		
Underbanked × High Collateral × Post				-.061	
				(.079)	
Underbanked × Partnership × Post					-.064
					(.112)
Underbanked × Private Ent. × Post					-.087
					(.116)
Underbanked × Govt. × Post					-.196
					(.160)
Underbanked × Listed × Post					-.224**
					(.107)
Observations	53666	53507	53507	53597	53666
R^2	.34	.34	.34	.34	.34
Control Mean	.04	.04	.04	.04	.04

Notes: This table estimatestreatment heterogeneity in credit growth across establishment fixed assets, tangibility, ownership, and listing status. The unit of observation is the manufacturing establishment. The outcome of interest is credit growth. All specifications include establishment and industry year fixed effects, along with a linear polynomial in the running variable, as well as establishment and district covariates. *High Capital* refers to establishments whose fixed assets exceed the median pre-treatment fixed assets. Administrative definitions are used to classify establishments as *Small*, *Medium*, *Large* and *SSI* (Small Scale Industries), based on their pre-treatment establishment fixed capital. *High Collateral* refers to establishments whose value of land and buildings exceed the pre-treatment median. *Listed* is a dummy for establishments owned by publicly listed corporations. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table B10: Manufacturing Investment in Underbanked Districts: Heterogeneity by Pre-Treatment Establishment Characteristics

	Capital Expenditures				
	(1)	(2)	(3)	(4)	(5)
Underbanked × Post	.027	.085**	.116***	.064***	.050*
	(.031)	(.035)	(.031)	(.021)	(.028)
Underbanked × High MRPK × Post	.062*				
	(.035)				
Underbanked × High Output × Post		-.050			
		(.047)			
Underbanked × High Interest × Post			-.058		
			(.043)		
Underbanked × ICR < 1 × Post				-.057	
				(.092)	
Underbanked × Tradable × Post					.034
					(.037)
Observations	69918	71445	61198	71542	68826
R^2	.38	.38	.38	.38	.38
Control Mean	-.03	-.03	-.03	-.03	-.03

Notes: This table estimates the treatment heterogeneity on manufacturing investment across borrower and industry characteristics. The unit of observation is the manufacturing establishment. The outcome of interest is capital expenditures, defined as in equation (2). All specifications include establishment and industry year fixed effects, along with a linear polynomial in the running variable, as well as establishment and district covariates. *High MRPK* refers to establishments with relatively high (above median) marginal product of capital; *High Output* refers to establishments with relatively high (above median) output per worker; *High Interest* refers to establishments with relatively high (above median) interest rates; $ICR < 1$ refers to establishments whose annual interest payments exceeded annual sales at least once in the pre-treatment period; *Tradable* refers to establishments operating in industries with relatively low geographic dispersion. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table B11: Bank Branch Expansion, Manufacturing Output and Establishment Profits

	Medium-Term				Long-Term				
	Output (Log)	Profits (IHS)	Profit > 0	Output (Log)	Profits (IHS)	Profit > 0	Output (Log)	Profits (IHS)	Profit > 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Underbanked × Post	-.092	-.421	-.002	-.073	.127	.020			
	(.117)	(.655)	(.021)	(.106)	(.584)	(.019)			
Underbanked × $1(2006 \leq Year \leq 2010)$							-.131	-.168	.008
							(.113)	(.618)	(.020)
Underbanked × $1(Year > 2010)$							.076	.893	.048**
							(.117)	(.727)	(.024)
Observations	71542	69436	71542	90097	87347	90097	90097	87347	90097
R^2	.84	.51	.50	.83	.48	.47	.83	.48	.47
Control Mean	146.70	11.45	.84	146.70	11.45	.84	146.70	11.45	.84

Notes: This table estimates the impact of BAP on manufacturing wages. The unit of observation is the manufacturing establishment. All specifications include establishment, industry-year fixed effects, a linear polynomial in the running variable, district and establishment covariates. All outcome variables are logged. Outcomes in columns (1)-(4) are total wage payments; in columns (5)-(8), average daily wages. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table B12: Cost of Credit for Manufacturing Establishments in Underbanked Districts: Heterogeneity by Establishment Size and Age

	Interest Rate			
	(1)	(2)	(3)	(4)
Underbanked × Post	.039	.063**	.032	.007
	(.025)	(.031)	(.030)	(.021)
Underbanked × $Size >$ Median × Post	-.046			
	(.029)			
Underbanked × $10 > Size \leq 25$ × Post		-.062*		
		(.036)		
Underbanked × $25 > Size \leq 50$ × Post		-.088*		
		(.046)		
Underbanked × $50 > Size \leq 100$ × Post		-.107**		
		(.047)		
Underbanked × $Size > 100$ × Post		-.042		
		(.044)		
Underbanked × Large, Young × Post			.014	
			(.040)	
Underbanked × Large, Old × Post			-.071*	
			(.036)	
Underbanked × Small, Old × Post			.013	
			(.029)	
Underbanked × Listed × Post				.055
				(.044)
Observations	53645	53645	53645	53645
R^2	.58	.58	.58	.58
Control Mean	.24	.24	.24	.24

Notes: This table estimates the treatment heterogeneity on imputed interest rates for manufacturing enterprises. The unit of observation is the manufacturing establishment. The outcome of interest is the imputed interest rate. All specifications include establishment and industry year fixed effects, along with a linear polynomial in the running variable, as well as establishment and district covariates. $Size$ refers to the pre-treatment average number of employees employed by the establishment. *Large* and *Small* refer to establishments with above and below median sizes. *Young* refers to establishments which started operation after 1992. Listed establishments are those which are publicly listed. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

C Appendix: Bank Branch Expansion, Manufacturing Employment, and Output – Role of Managerial Capital

This section examines the impact of the branch expansion policy on manufacturing employment and output. Section 4.5 showed muted effects of financial deepening on establishment profits over the medium-run, but positive effects over the long-run. In light of the large literature documenting the positive role of managerial capital on firm performance, we examine whether financial deepening affected manufacturing output over the medium-run, conditional on the availability of managerial capital.

C.1 Workers and Wages

Columns (1)-(4) of Appendix Table C1 show no impact of financial deepening on firms' labor demand. The point estimates are negative and not precisely estimated. Appendix Table C2 shows that the lack of an impact on employment cannot be attributed to higher labor costs in underbanked districts. If anything, the evidence points to a decline in labor costs, with a significant reduction in contract worker wages. A possible explanation would be that manufacturing establishments were substituting labor with additional capital. As contractual employees are not protected by labor regulations, they are likely to form the first margin of labor adjustment.

C.2 Manufacturing Output

Appendix Table B11 identified a null effect of BAP on manufacturing output in the medium-run. Bloom (2010) points that lack of managerial talent (along with financing constraints) is a key factor depressing firm productivity in developing economies, while Bloom et al. (2019) use firm surveys to show that managerial capital can explain up to 20% of the variation in cross-country firm productivity. Fenizia (2022) combines administrative data and quasi-exogenous rotation of managers in Italy to show that improvements in the allocation of managerial capital can raise output by over 6 percent. Experimental evidence from Bloom et al. (2013), Bruhn et al. (2018) and Gosnell et al. (2020) also document a positive impact of managerial capital on firm performance and entrepreneurship. This leads us to examine whether the absence of managerial capital precluded establishments from fully realizing the gains from higher capital spending in response to improved access to credit.

The role of managerial capital becomes salient in our context as the branch expansion policy was targeted towards under-developed regions. While measuring managerial capital is challenging (Bruhn et al., 2010), a loose proxy using educational attainment shows that only a third of the working-age population in underbanked districts had completed secondary education in 2005. Consequently, if the supply of skilled labor remained fixed over the medium-run and there is limited mobility of labor across regions (Mahajan and Tomar, 2023; Munshi and Rosenzweig, 2009), the absence of skilled labor can preclude enterprises from fully realizing the gains from increased capital investment.

We test this hypothesis using ASI's disaggregation of employees into workers and supervisors, with the latter engaged in non-manufacturing tasks such as overall management and supervision of plant operations. We define an establishment's managerial capital as the ratio of supervisors to total employees ($ShSuper$). Establishments are classified as having "high" managerial capital if their pre-treatment ratio of supervisors ($ShSuper$) exceeds the pre-treatment median value of $ShSuper$ across all establishments.[1] Appendix Table C3 shows a positive coefficient on the triple interaction term in column (2), albeit it is imprecisely estimated (p-value: .145). A further disaggregation of establishments' managerial capital to the top two quartiles in column (3) yields positive coefficients for both triple interaction terms, with the one corresponding to establishments with the high-

[1]Between 2001 and 2005, $ShSuper$ for the median establishment in non-underbanked districts was 0.079. The median establishment hired 2 supervisors, (relative to 22 employees).

est managerial capital being statistically distinguishable from 0 at the 10% level (p-value .058). The sum of the coefficients suggest that manufacturing output increased by 14 percent in the post-treatment period in underbanked districts for establishments with the highest managerial capital.[2]

Since Table 8 shows that the increase in capital investment and credit is driven by relatively small establishments, column (4) of Appendix Table C3 shows the estimates for establishments below the median pre-treatment establishment size (20 workers). Column (4) shows that the results in column (3) are driven by relatively smaller establishments; within this sub-sample, output increases by almost 30 percent for establishments with the highest level of managerial capital. There is no corresponding increase in output for larger establishments in columns (5), irrespective of managerial capital. These findings line up with the results in Table 8, where capital investment and credit growth for the largest establishments were unaffected by the BAP.

Collectively, Appendix Table C3 underlines the critical role of managerial capital: while financial deepening can indeed alleviate credit constraints and lead to higher manufacturing investment, complementary inputs in managerial capital are necessary for capital investment to boost manufacturing output.

[2] We confirm in Appendix Table C4 that the increase in manufacturing investment and credit growth was also higher for establishments with relatively high managerial capital; while the triple interaction coefficients are imprecisely estimated, the sum of the coefficients for both capital investment and credit growth are significantly different from 0 for establishments with the highest level of managerial capital (p-values of .038 and .013 respectively).

Table C1: Manufacturing Employment in Underbanked Districts

	Total Employees (Log)) (1)	Total Workers (Log) (2)	Contract Workers (Log) (3)	Supervisors (Log)) (4)	Supervisors Share (5)
Underbanked × Post	-.035	-.033	-.143	-.064*	-.002
	(.056)	(.064)	(.109)	(.037)	(.005)
Observations	71542	71542	71496	71496	71229
R^2	.92	.90	.77	.88	.66
Control Mean	105.92	87.55	29.47	8.38	.11

Notes: This table estimates the impact of BAP on manufacturing employment. The unit of observation is the manufacturing establishment. All specifications include establishment, industry-year fixed effects, a linear polynomial in the running variable, district and establishment covariates. The outcome variables in columns (1)-(4) is logged. The outcome of interest in column (1) is total employees; in column (2), total number of workers; column (3), total contract workers; column (4), the number of supervisors; in column (5), supervisors as a share of total employees. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table C2: Labor Compensation in Underbanked Districts

	Total Wages (Log)				Daily Wages (Log)			
	All Employees	Total Workers	Contract Workers	Supervisors	All Employees	Total Workers	Contract Workers	Supervisors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Underbanked × Post	-.091	-.107	-.548	-.187	-.048*	-.036	-.127**	-.052
	(.096)	(.120)	(.446)	(.286)	(.026)	(.029)	(.050)	(.043)
Observations	71542	71542	71496	71496	71313	70999	24278	63427
R^2	.87	.78	.73	.78	.86	.82	.66	.77
Control Mean	6.95	3.92	.72	1.79	194.22	152.13	136.13	496.30

Notes: This table estimates the impact of BAP on manufacturing wages. The unit of observation is the manufacturing establishment. All specifications include establishment, industry-year fixed effects, a linear polynomial in the running variable, district and establishment covariates. All outcome variables are logged. Outcomes in columns (1)-(4) are total wage payments; in columns (5)-(8), average daily wages. All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table C3: Manufacturing Output in Underbanked Districts

	Output (Log)				
	(1)	(2)	(3)	(4)	(5)
Underbanked × Post	-.024	-.152	-.153	-.424	-.091
	(.105)	(.137)	(.138)	(.333)	(.140)
Underbanked × $High\ Super$ × Post		.211			
		(.145)			
Underbanked × $Sh.\ Super^{Q3}$ × Post			.130	.354	.097
			(.168)	(.375)	(.178)
Underbanked × $Sh.\ Super^{Q4}$ × Post			.302*	.696*	.105
			(.158)	(.372)	(.179)
Observations	63624	63556	63556	13822	49687
R^2	.85	.85	.85	.77	.85
Control Mean	181.29	181.29	181.29	181.29	181.29

Notes: This table estimates the impact of BAP on manufacturing output. The unit of observation is the manufacturing establishment. All specifications include establishment, 2-digit industry-year fixed effects, a linear polynomial in the running variable, district and establishment covariates. The outcome of interest is logged output. $High\ Super$ is a dummy equaling 1 if the establishment has a relatively high pre-treatment share of supervisors. $Sh.\ Super^{Q3}$ is a dummy equaling if the establishment's pre-treatment share of supervisors falls in the third quartile; $Sh.\ Super^{Q4}$ is a dummy equaling if the establishment's pre-treatment share of supervisors falls in the top quartile; All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and ***1%

Table C4: Capital Investment and Credit Growth in Underbanked Districts: Heterogeneity by Managerial Capital

	Capital Expenditures			Credit Growth		
	(1)	(2)	(3)	(4)	(5)	(6)
Underbanked × Post	.060**	.027	.026	.051	.055	.053
	(.023)	(.029)	(.029)	(.044)	(.076)	(.076)
Underbanked × $High\ Super$ × Post		.052			.004	
		(.045)			(.090)	
Underbanked × $Sh.\ Super^{Q3}$ × Post			.033			-.051
			(.049)			(.103)
Underbanked × $Sh.\ Super^{Q4}$ × Post			.076			.094
			(.054)			(.093)
Observations	63624	63556	63556	48752	48705	48705
R^2	.38	.38	.38	.34	.34	.34
Control Mean	-.02	-.02	-.00	.04	.04	.00

Notes: This table estimates the treatment heterogeneity in manufacturing investment and credit growth by establishments' managerial capital. The unit of observation is the manufacturing establishment. All specifications include establishment, industry-year fixed effects, a linear polynomial in the running variable, district and establishment covariates. The outcome of interest in columns (1)-(3) is capital expenditures, defined as in (2); in columns (4)-(6), credit growth. Capital expenditures is restricted to capital expenditures in plant and machinery. $High\ Super$ is a dummy equaling 1 if the establishment has above median pre-treatment share of supervisors. $Sh.\ Super^{Q3}$ is a dummy equaling if the establishment's pre-treatment share of supervisors falls in the third quartile; $Sh.\ Super^{Q4}$ is a dummy equaling if the establishment's pre-treatment share of supervisors falls in the top quartile; All specifications include establishment-specific weights and the sample is restricted to districts located within a bandwidth of 15 around the discontinuity threshold. Standard errors are in parentheses, clustered by district. Significant levels: *10%, **5%, and × ***1%

SOVEREIGN DEFAULT AND TAX-SMOOTHING IN THE SHADOW OF CORRUPTION AND INSTITUTIONAL WEAKNESS

Marina Azzimonti* **Nirvana Mitra†**

November 2023

Emerging countries exhibit volatile fiscal policies and frequent sovereign debt crises, that significantly diminish the well-being of their citizens. International advisors typically suggest developed-world solutions as a remedy. We argue that the root of the problem lies in the institutional environment, which does not incentivize responsible policymaking, particularly tax-smoothing practices. Focusing on democratic representation and control of corruption, our dynamic political-economy bargaining model shows that nations with weaker institutions experience frequent default episodes and greater economic volatility. Our results are in line with stylized facts from a panel of 58 countries between 1990 and 2022. Through counterfactual experiments, we find that while emerging economy policymakers might favor moderate reforms to improve democratic representation, achieving the institutional depth seen in developed countries is politically unfeasible, despite its clear advantages for citizens.

JEL Codes: D72, E43, F34, E62, F41

Keywords: Sovereign Debt Crises, Tax Smoothing, Representation and Accountability, Corruption, Sovereign Default, Fiscal Pro-Cyclicality, Emerging Markets

*FRB of Richmond Email: marina.azzimonti@gmail.com

†CAFRAL & Shiv Nadar IoE Email: nirvana.mitra@gmail.com

Non-Technical Summary

This paper contributes to the effort to explain the characteristics of emerging economy business cycles and distinguish them from those in advanced economies. It introduces a unified framework for both types of countries and focuses on the existence and deepening of democratic institutions to explain the differences. The paper first documents that countries with strong representation and accountability, and the ability to control corruption are also favored by foreign investors because of lower chances of default on their sovereign debt, as evident from their lower credit ratings and interest rate spreads. Most of the advanced countries (by World Bank classification) are characterized by strong institutional quality. Furthermore, emerging market economies react to unfavorable access to credit by borrowing more when credit costs are lower. This further magnifies the cycle and increases volatility in their economic fundamentals. They borrow into a boom and have limited access to credit during busts. This is reflected in their countercyclical trade balance and fiscal balance. On the other hand, advanced economies can borrow during busts, enabling them to smooth taxes and consumption over time. Advanced economies seldom default. This allows them to borrow in recessions and depend less on tax revenues, even when their spending requirements are higher.

We explain these differences using a dynamic model of fiscal policy and sovereign default, and exogenously variable institutions. We find that even if default is a possibility, countries with very strong institutions do not default. They borrow less and endogenously prefer to adjust their borrowing and taxes rather than default. Countries with weaker institutions borrow more, tax more, and prefer to default in bad times as borrowing becomes more expensive and increasing tax revenues become unfeasible. We term this as a political default in this paper.

This behavior is incentivized by the ability of politicians to divert public resources for private benefit. Our model economy is divided into a discrete and finite number of regions. Each region is uniformly populated and represented by one politician in the central legislature. These politicians bargain over policy. Stronger institutions or better representation are modeled as the requirement of more votes to pass a policy mix that consists, among other things, of a region-specific kickback or favor that is exclusive to the region. This captures the tendency of many economies with weak institutions to use favors for political gain (buying votes). In good times, with fewer votes required to pass a policy, each politician is completely aware of the importance of their vote, and that they can negotiate for more favors by wielding their importance to sustain the government. This results in endogenously higher favors at the cost of lower levels of services provided to the average citizen.

Other than representation, we model control of corruption as the ability to convert tax revenues into private favors. Technically, control of corruption is the price of favors every period. Even if the above mechanism encourages resources to be spent on favors in countries with weaker institutions, strict control of corruption can strongly discourage such squandering. Singapore, for example, is a country with lower political representation but almost draconian anti-corruption measures. Most of the advanced economies in our sample have both strong political representation and control of corruption, and the emerging economies are weak along both dimensions. Improvement along both fronts tends to reduce spreads and volatility. Foreign investors only care about default and both policies tend to reduce chances of default. However, the executive's preferences for policy are almost always at the cost of the welfare of the average citizen. It is only with a greater depth of institutions that the preferences of the executive and the average citizen align completely. This finding characterizes the political economy hindrances to the development of insti-

tutions. International institutions focus on capacity development as if it is the knowledge about the cost, benefits, and consequences of good policy that is lacking without emphasizing the political incentives for their implementation.

The model is calibrated to Argentina for the period between 1990-2018. Argentina was characterized by weak institutions, and lower control of corruption during this sample period and had been categorized as a serial defaulter. We use this calibrated model to conduct two numerical exercises. First, we study the impact of institutions on the response to a recession shock. We find that our model economy with stronger institutions behaves similarly to advanced economies in the data. Responding to a negative productivity shock, an economy with weak institutions/low control of corruption raises taxes and lowers debt (if not default) as debt becomes expensive. Labor supply dampens, aggravating the effect of the shock on GDP. The other type, on the other hand, lowers taxes to mitigate the effect of the shock. It funds the excess spending during a recession using cheaper credit. Second, we use the anecdotal examples of transitioning to democracy for both Argentina and Chile from their dictatorships in the late twentieth century. While starting from a higher level of debt we find that Chilean institutional depth contributed to a less volatile economy with lower chances of default. Argentine institutions allowed the use of government-sponsored welfare schemes as political instruments for the distribution of favors resulting in rampant misuse and default at regular intervals. Not only in the transition, these economies also differ in the long run. Their ergodic distribution properties also suggest that countries with weaker institutions tend to borrow more, tax more, and use the resources as political favors resulting in a significantly higher incidence of politically induced default episodes.

This paper uses a real model of fiscal policy to study the contribution of institutions and the political incentives at play in shaping them. We abstract from inflation, trade policies, and a variety of fees that the executive can use to its advantage in building its network of favors-for-favors. Our paper is also silent on the ways institutions can be improved over time in an economy. In our model, institutions can be improved endogenously in a recession where favors are not possible anyway. In reality, we find that inclusive political institutions improve over time, even for the country at the lower end of the spectrum, albeit slowly. This, we conjecture, is a culmination of many such downturns that add cumulatively to institutional development. To expedite this process, international organizations should encourage institution-building both implicitly and explicitly as attached conditionalities, and increase the transparency of fiscal and monetary policies so that reversal of policies that are responsible for sound institutions is not possible, even if they are politically feasible.

1 Introduction

Fiscal policy responses to shocks are widely different across countries. Developed economies smooth taxes, borrowing in recessions, and paying back during booms[1]. This is illustrated in Table 1 (in blue), where a positive correlation is observed between the primary balance and output.[2] This behavior, which dampens economic cycles, is consistent with optimal debt determination (see Barro (1979)).

Table 1: Business Cycle Moments and Institutions

Moments	Developed	Emerging
Economic and Fiscal Variables		
$\sigma(y)$	5.0	7.3
$\sigma(c)/\sigma(y)$	1.1	1.2
$\sigma(g)/\sigma(y)$	0.9	1.3
Fitch Ratings	3.8	2.9
$\rho(NX, y)$	0.0	-0.1
$\rho(PB, y)$	0.3	0.0
Representation and Accountability (R&A)		
Average	1.2	0.1
Volatility (st dev)	0.1	0.2
Control of Corruption (CC)		
Average	1.4	-0.2
Number Countries	30	28

Notes: Yearly sample, 1990-2022, where y =GDP, c =private consumption, and g = public consumption. All real and log-linearly de-trended. PB is the primary balance (Rev-Exp+rb), and NX are net exports (X-I), both as $\%$ of GDP. Correlations represent regression coefficients. R&A and CC in the range [-2.5,2.5]. See Appendix A.3 for sources and further computation details.

In emerging economies, borrowing is acyclical on average, and pro-cyclical for certain countries (see Figure 1). The latter means that borrowing rises during expansions, facilitated by easy access to international capital markets, and decreases during recessions, often as a result of a sovereign default.[3] Evidence of emerging countries' reduced capacity to mitigate shocks can be seen in the more pronounced fluctuation in private and public consumption relative to output, a higher volatility of output itself, as shown in Table

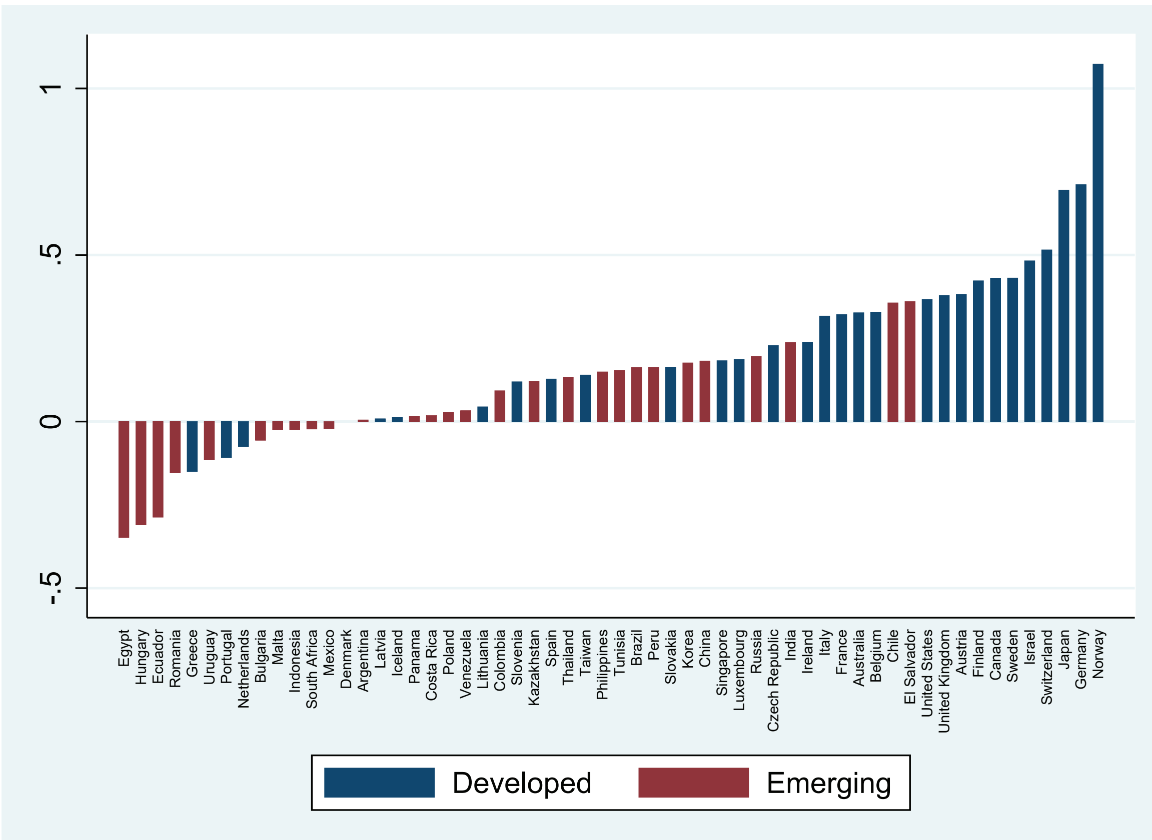
[1] See the studies by Sargent and Velde (1995) or Hall and Sargent (2020)).

[2] The negative relationship between net exports and GDP is consistent with this as well, indicating that countries tend to import goods (e.g. borrow) in recessions.

[3] This phenomenon, especially prevalent in Latin America during the 80s and 90s, has been noted by studies including Kaminsky et al. (2004), Talvi and Végh (2005), Ilzetzki and Vegh (2008), and more recently Kaas et al. (2020) and Bianchi et al. (2023). However, our sample period from 1990 to 2022 reveals a significant decrease in the number of countries showing procyclicality.

1 and the higher frequency of sovereign debt crises (see Mitchener and Trebesch (2023)). Even within emerging economies, some countries are 'serial defaulters' (like Argentina), while others (like Chile) have been less prone to default.

Figure 1: Correlation between Primary Balance and GDP - benchmark sample



**Notes:* See Appendix A.3 for sources and computation details.

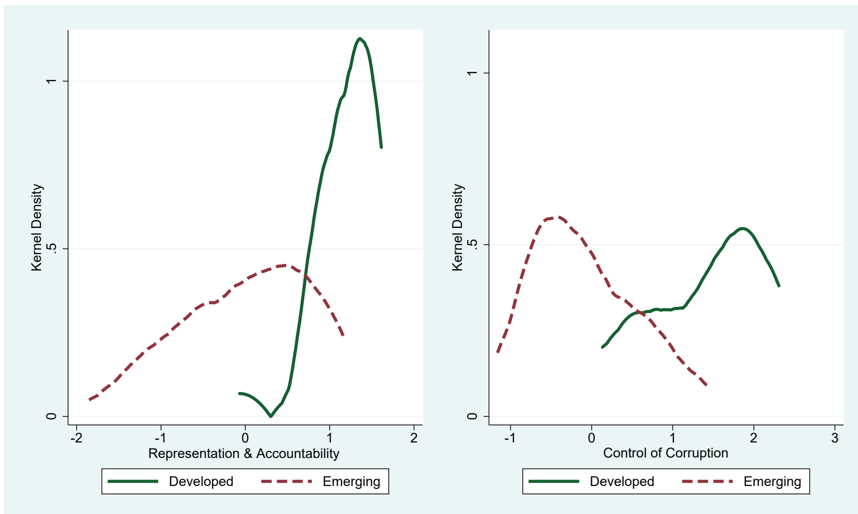
Observing such diverging outcomes, international organizations typically prescribe policies to serial defaulters that proved successful in developed countries, such as reducing deficits and the size of the public sector. For example, IMF loans include conditions emphasizing fiscal austerity and budgetary reforms, including structural reforms at times [4]. However, policymakers quickly deviate from such prescriptions or do not implement them at all, as they become politically unfeasible. Unfortunately, it is not enough that policymakers understand what good policies are; they need to find it optimal to implement them. In this paper, we argue that the institutional framework in which policies are decided is as important as, if not more than, the policy prescriptions themselves. We concentrate on two key institutional characteristics, namely the degree of democratic *representation and accountability* and the *control of corruption*. We propose that these elements play a significant role in guiding fiscal policy and debt management strategies. Once we summarize how these two factors are measured in the data, we proceed to describe their role in our quantitative model.

Our empirical proxies are obtained from the World Bank's Worldwide Governance Indicators. The degree of representation and accountability (R&A) captures public perceptions about (i) the extent to which citi-

[4]See https://www.imf.org/en/About/Factsheets/Sheets/2023/IMF-Conditionality

zens can participate in selecting their government through free elections, freedom of expression, freedom of association, and a free media (also known as 'vertical accountability') and (ii) de-facto checks and balances (or 'horizontal accountability')[5]. Positive scores indicate stronger democratic institutions (see additional details in Appendix A). The second indicator we use is 'control of corruption' (CC), capturing the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as the capture of the state by elites and private interests. Smaller values indicate that excessive patronage, nepotism, and 'favors-for-favors' are widespread. Table 1 presents summary statistics for these series.

Figure 2: Distribution of R&A (left) and CC (right) across countries



**Notes:* World Bank's WGI, 58 countries, 1990-2022 (country list in Online Appendix).

Figure 2 shows the distributions of R&A (left) and CC (right) across 58 countries between 1990 and 2022. On average, developed economies (solid green line) are more representative and exhibit less corruption than emerging ones (dashed maroon line).[6] However, there is some overlap between the CC series, indicating that some developed economies have levels of corruption closer to those in emerging countries, and vice-versa.[7]

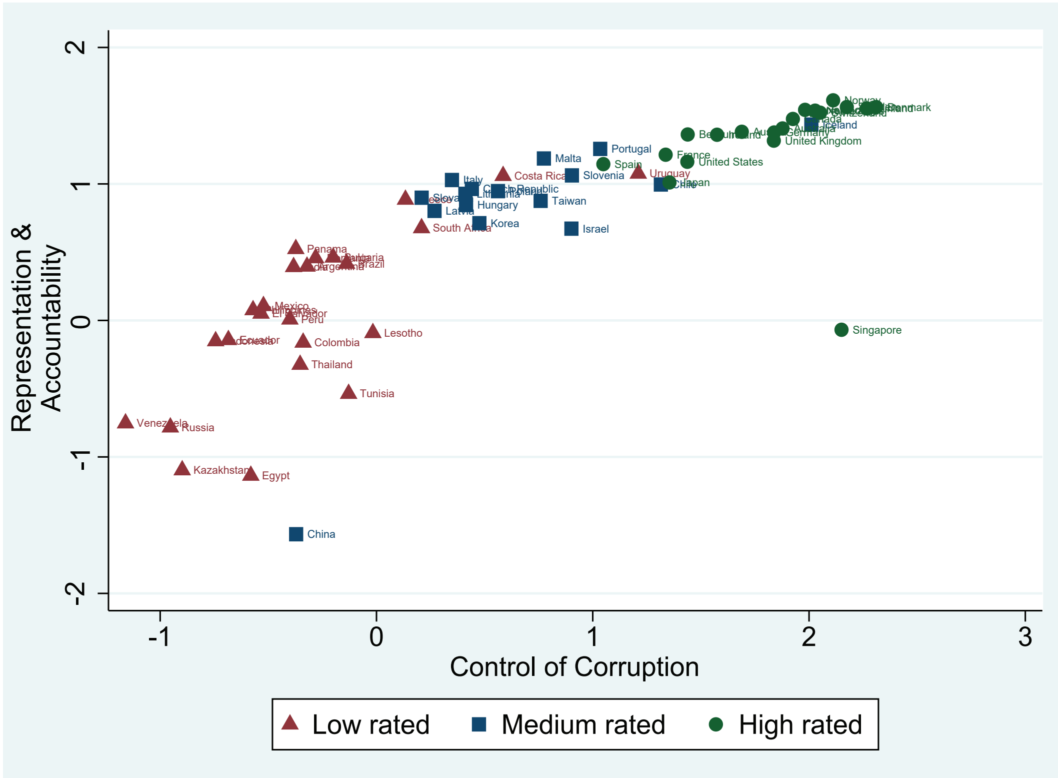
Figure 3 depicts the average values of R&A plotted against CC for each country, revealing a positive correlation. In other words, governments with stronger democratic institutions typically exhibit less corruption (with Singapore being an exception). Different colors are used to emphasize the country's default risk, which is calculated using the long-term average of Fitch ratings (see Appendix A.2 for details). Coun-

[5]The concepts of vertical and horizontal accountability were first introduced by O'Donell (1998). While 'representation and accountability' is a *subjective* or de-facto measure combining both vertical and horizontal accountability, there exist *de-jure* measures that independently capture each dimension. Because they are based on enacted legislation and power structure (see Lührmann et al. (2020)), they change less frequently and more abruptly than R&A. Our benchmark measure captures partial changes, as perceived by the population. On average, these measures are highly correlated with those constructed by the WGI.

[6]The sample only includes countries for which we also had economic and fiscal variable observations, as well as Fitch sovereign debt ratings, as described later. The unrestricted distributions of CC and R&A are shown in the Online Appendix, Figure 1. Clearly, our sample misses countries with the weakest institutions.

[7]The only developed country with negative R&A scores is Singapore. See Tables 1 and 2 in the Online Appendix.

Figure 3: Default risk, representation and accountability and corruption



**Sources*: World Bank's "Worldwide Governance Indicators (WGI)" and Fitch Ratings, 58 countries, 1990-2022. See Appendix A.3 for details.

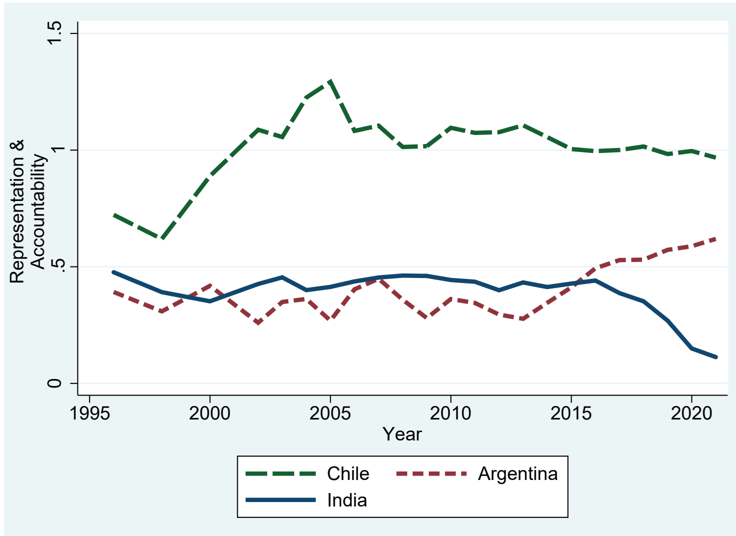
tries are categorized into three groups: those with high ratings (marked by green circles), medium ratings (marked by blue squares), and low ratings (marked by red triangles). Lower ratings indicate a higher probability of default. Countries with higher representation that keep corruption under control generally have good sovereign credit ratings, which is predominantly seen in developed economies. It is worth noting that our sample does not include countries with very low R&A and CC scores, due to either their lack of participation in international financial markets or insufficient economic data.[8]

Table 1 presents the standard deviation of representation and accountability for the two groups of countries, underscoring the volatile nature of R&A in emerging markets. There is also notable disparity within this group, as illustrated by Figure 4. The government of Chile scores significantly higher than Argentina's and India's, reaching a R&A mean that is much closer to that of developed countries in recent years. Despite having similar paths until 2015, Argentina and India's R&A scores diverged soon after, highlighting the dynamic nature of institutional development.

Taken together, the evidence above indicates that: (i) fiscal policy and macroeconomic outcomes are more volatile in emerging countries (there is less tax smoothing), (ii) default risk is lower in countries with strong representation and accountability and low levels of corruption, and (iii) institutional strength evolves over time. While developed economies look similar both in terms of their institutional setting and risk of de-

[8]In the figure, we are only including countries for which we have institutional and economic data, as well as Fitch ratings' observations. The scatter plot including all countries with existing R&A and CC scores for this period can be found in Figure 2 of the Online Appendix.

Figure 4: Representation & Accountability over time (Argentina, Chile, and India)



fault, emerging economies are more heterogeneous. In this paper, we build a theory consistent with these observations, where institutional strength and the pervasive effects of corruption take center stage.

Our study builds on the dynamic political economy bargaining model developed in Azzimonti and Mitra (2023), to explore how democratic representation and corruption impact policymakers' behavior towards tax smoothing and preventing sovereign debt crises.[9] The population is composed of symmetric groups representing diverse regions, industries, castes, and ethnic or religious groups, each with a leader that has political influence. Group leaders bargain periodically over public goods spending, financing through distortionary taxes and/or external borrowing, and over whether to repay or default. One group leader is randomly selected to make a policy proposal, but needs the support of a 'minimum winning coalition' (mwc)—formed by a subset of these groups— to implement policy. In order to secure support, the proposer can promise 'political favors.' These can take the form of targeted public goods, permits for resource exploitation, bribes, subsidies, nepotism, pork spending, favorable exchange rates, or policies affecting non-economic dimensions such as ethnic or religious in-group favoritism, same-sex marriage and abortion laws, or transgender rights. A key characteristic is that these favors are costly to provide and can be targeted to benefit a specific group. The 'control of corruption' is an institutional characteristic determining how easy it is to transform public revenues into favors.

The size of the minimum winning coalition (mwc) in the model plays a crucial role in determining the number of groups whose preferences are considered when making policy decisions, capturing the notion of 'representation and accountability.' A small mwc allows the proposer to pass policies with the support of only a few groups, ignoring the effects of their actions on other groups. Because political favors are financed with government resources, a common pool problem arises (where all groups contribute, but only a few enjoy the proceeds of taxation). As a result, the use of favors can be widespread, generating deadweight losses[10]. When the mwc equals the total population, the model collapses to that of a benevolent

[9] In Azzimonti and Mitra (2023), we augmented the bargaining model proposed by Battaglini and Coate (2008) by integrating sovereign default, akin to Arellano (2008) and Cuadra et al. (2010).

[10] Choi et al. (2021) provides micro evidence of how excess ARRA stimulus grants extended to politically connected firms in the U.S. lead to fewer jobs.

planner without commitment. This captures a society where the preferences of all groups are considered, and favors are never used. International risk-neutral lenders provide loans to the government, and in turn, price sovereign bonds based on the government's default risk. This pricing mechanism ensures that the lenders discipline government borrowing, offering lower prices when the likelihood of default is high. To accurately assess the risk, lenders must consider, in addition to the state of the economy and the current stock of debt, the current and future institutional environment as well as the level of corruption[11]. Because of this, regime changes have important implications for debt sustainability.

The latter is illustrated with a quantitative counterfactual experiment in which we evaluate the macroeconomic consequences of a permanent regime change that strengthens institutions in two countries (e.g. increases the size of the mwc), but asymmetrically, using Argentina and Chile as case studies. Both countries were under the rule of military juntas in the 1970s, but transitioned to democracies in 1983 and 1989, respectively. While Argentina and Chile share similar geographic, cultural, and historical characteristics, Chile was more successful in ensuring well-functioning checks and balances– or horizontal accountability— after democracy was reinstated. In our model, this is captured by a higher mwc size. We show that our model can help rationalize the different trajectories in Argentina's and Chile's economies in subsequent years. In particular, we find that Argentina suffers recurrent sovereign debt crises and significant volatility in private and public consumption, whereas Chile does not, consistently with the data (see Mitchener and Trebesch (2023)). A key insight is that while both countries can default at any time, institutional strength makes this option less desirable for Chilean policymakers than for Argentinean ones.

More generally, our analysis reveals that policymakers tend to smooth taxes when R&A is high, as in developed countries, but less so when R&A is low, as in emerging countries, *even when strategic default is permitted in both cases and policymakers are equally patient*. Specifically, we find that policymakers in countries with strong institutions behave responsibly, reducing debt in good times and avoiding the default region. In contrast, when the government is less representative (but still democratic), government officials borrow excessively in good times to give out favors and find it beneficial to default when recessions hit, as commonly observed in emerging markets. In other words, our model generates 'political defaults,' namely that a politician chooses to default when a benevolent planner would not. As a result, low R&A countries exhibit worse tax smoothing (and sometimes even procyclical debt), recurrent sovereign debt crises, and frequent default events[12].

Addressing corruption can mitigate some impacts of inadequate checks and balances. Yet, our analysis reveals that reforms enhancing representation and accountability are significantly more effective in stabilizing economic shocks than similar efforts to curb corruption. While anti-corruption measures could theoretically improve resource allocation by lowering default risks and economic volatility, our findings suggest that politicians in environments with extremely weak democratic structures initially prioritize strengthening these structures. This preference stems from the fact that, although both strategies might lead to comparable outcomes in terms of taxation, public expenditure, and debt repayment in a political equilibrium, anti-corruption efforts limit the use of political favors by those in power. However, extensive reforms in representation and accountability to match those in developed countries are politically unattainable. Politi-

[11] In a model with nominal debt—denominated in domestic currency—policy risk, emanating from less developed and volatile domestic institutions, would lead to excess returns due to currency risk and a positive violation of UIP, as documented in Kalemli-Özcan and Varela (2021).

[12] While autocratic regimes are not the main focus of interest in our applications, we show how the model can be expanded to analyze them. An important difference between democratic regimes with low R&A and autocratic regimes is that the latter exhibit less political turnover. We develop this further in Appendix C.2.

cians are more inclined to implement incremental reforms, just large enough to receive favorable treatment from lenders, but allowing them to maintain their capacity to distribute political favors.

The rest of the paper is organized as follows. Section 2 summarizes how our work fits in the literature. In Section 3, we describe the economic environment and define the politico-economic equilibrium. Section 4 describes the calibration strategy. We compare the transition to democracy in Argentina and Chile by simulating a once-and-for-all institutional reform in Section 5 and describe tax smoothing in emerging and developed economies through the lens of our model in Section 6. In Section 7, we analyze a 'crackdown on corruption' and show how it can be different from a reform that increases representation. Section 8 concludes.

2 Related Literature

The literature on fiscal policy cyclicality has bifurcated into two distinct branches: (i) the *traditional public finance* literature and (ii) the more quantitative *sovereign default* literature. Our paper aims to connect them by providing a unifying theory of fiscal policy cyclicality and sovereign default decisions.

The traditional public finance literature has predominantly focused on studying developed economies, assuming the existence of a 'commitment technology' for debt repayment. The seminal work of Barro (1979)[13], followed by the studies of Aiyagari et al. (2002) and Barseghyan et al. (2013), have highlighted tax smoothing and counter-cyclical debt as key implications of their models (these are summarized in Yared (2019)). While consistent with the cyclical properties of fiscal policy in developed countries like the US, their predictions are inconsistent with the empirical behavior observed in emerging economies (as documented by Gavin and Perotti (1997), Talvi and Végh (2005), and Ilzetzki and Vegh (2008)). An exception is Ilzetzki (2011), who generates procyclical fiscal policies in a model subject to political distortions, but assuming commitment to repay debt (and abstracting from default risk, inconsistent with Figure 3).

At the other extreme, the sovereign default literature—with the pioneering work of Arellano (2008)— has centered on explaining the behavior of emerging economies, with particular emphasis on Latin America. An important innovation in these models is allowing the government to strategically default. Within this environment, Cuadra et al. (2010) shows that debt increases in booms rather than recessions and taxes are pro-cyclical (see more recent work by Kaas et al. (2020) and Martinez et al. (2022)). The latter is a direct consequence of counter-cyclical sovereign default risk: When default risk rises the government needs to raise tax rates to finance public goods[14] (see Vegh and Vuletin (2015) for empirical evidence on this channel). At the end of the day, then, what makes emerging economies different from developed ones in these papers boils down to one important assumption: whether countries have a commitment technology to repay debt obligations or not[15]. The European sovereign debt crisis, however, demonstrated that even developed nations, such as Greece, may occasionally opt for strategic default. In this paper, we provide a unifying

[13]Also see Lucas and Stokey (1983) and recent work by Hall and Sargent (2020) for a discussion of its empirical validity.

[14]Espino et al. (2020) shows, in a model with nominal debt, how governments can use a combination of both inflation and distortionary taxes to finance debt. The interplay between default risk and inflation is also shown in Mihalache (2020) in an environment with long-term debt.

[15]In addition, it is typically assumed that policymakers in emerging markets are 'more impatient' than those in developed markets, incorporated through differences in their discount factors.

theory that can encompass the behavior of countries across different stages of development[16].

We build on the dynamic bargaining political economy model developed in Azzimonti and Mitra (2023) to study how different degrees of representation and accountability along with poor control of corruption affect fiscal policy decisions. We relate to the dynamic bargaining literature in political economy, which typically abstracts from default (see Baron and Ferejohn (1989), Battaglini and Coate (2008), Azzimonti et al. (2016), or Drazen and Ilzetzki (2023)).[17] A recent paper by Cusato Novelli (2020) incorporates default in a bargaining model, allowing the author to simultaneously match the observed default frequencies and debt-to-output ratios, something that traditional sovereign default models typically fail to achieve [18]. Our model can also match these moments. The main difference between our papers is that we consider a production economy (rather than an endowment one) where taxes are distortionary[19]. This generates a tax-smoothing motive—as in the traditional public finance literature— a key ingredient to our analysis of the effects of institutional strength on fiscal policy cyclicality and the impact of politics on business cycle fluctuations. Another important difference between our models is that we allow for democratic representation to fluctuate over time as seen in the data. This feature of our model can explain why countries with low democratic representation display more volatile fundamentals than their developed counterparts. Both democratic representation and corruption, as well as their volatility, *amplify* the business cycle and increase the likelihood of sovereign debt crises.

It is important to note that while we incorporate dynamic bargaining, ours is not a model of debt restructuring, where the government renegotiates repayment with foreign creditors (see Yue (2010)). Our study focuses instead on internal political bargaining, where the key actors are domestic politicians, who often must agree on tax increases to avoid a default. This internal bargaining process is influenced by the institutional context, specifically the levels of democratic representation and corruption within the government.

We offer three novel quantitative contributions relative to previous work. First, we explain the tax-smoothing behavior of both developed and emerging countries assuming the same discount factor, which has not been done in previous quantitative studies (e.g. they need to assume a significantly smaller value for the discount factor of emerging economies' government to generate default in equilibrium). There is a body of work analyzing the role of political distortions on default decisions within the sovereign default literature (see Hatchondo and Martinez (2010) and Chang (2007)[20]). Hatchondo et al. (2009) study how differences in re-election probabilities that affect the effective discount factor of a government, can impact default incentives. Chatterjee and Eyigungor (2019) and Cotoc et al. (2022) study default incentives under endogenous re-election probabilities, but again assume relatively low discount factors. In our bargaining model, the discount factor equals the inverse of the risk-free rate as in standard macroeconomic models, allowing us

[16]While there is heterogeneity across countries on institutional development, we abstract from inequality within the country. See DErasmo and Mendoza (2016) or DErasmo and Mendoza (2021) for studies where default has redistributive effects.

[17]Andreasen et al. (2019) considers default but in a voting model and abstracting from productive inefficiencies associated with the tax system.

[18]See a discussion of the difficulties in obtaining this in default models with a benevolent government in the literature reviews by Tomz and Wright (2007) and Aguiar and Amador (2014).

[19]In Mendoza and Yue (2012a), a default amplifies economic fluctuations through a production channel. A default reduces the ability of producers to obtain the working capital needed to import intermediate goods. An inefficiency arises after a default since domestic intermediate goods are poor substitutes for imported ones. Taxes are lump-sum in their environment, so a tax-smoothing motive is not considered.

[20]Relatedly, Herrera et al. (2020) models the relationship between popularity (e.g., political booms) and financial crises.

to replicate the tax-smoothing behavior of developed economies when either institutions are strong (e.g. the mwc equals the whole population) or corruption levels are low, or both. By weakening institutions (e.g. lowering the size of the mwc), we can approximate the cyclicality behavior of emerging economies without the need to change the degree of impatience of policymakers. In the appendix, we illustrate how democratic representation and political turnover are not equivalent by analyzing an autocracy. That is, a situation where the size of the mwc is low but the autocrat remains in power forever. We also show that the dynamic behavior of our model is not equivalent to one where the discount factor is low. Second, we investigate how institutional reform can improve tax smoothing in an emerging country using Chile and Argentina as case studies. Third, we conduct a policy experiment to highlight how an improvement in the technology to control corruption differs from an institutional reform strengthening democratic representation –an exercise that has implications for international policy-making.

3 The model

This is an infinite-horizon model where time is discrete. The economy has a domestic sector, with competitive firms and workers, a government that decides policy, and international investors that buy debt from the government.

3.1 Economic Environment

The domestic economy is populated by a continuum of infinitely lived agents uniformly distributed across n groups, representing diverse regions, industries, castes, and ethnic or religious groups, each with a population normalized to 1. Agents in group i derive utility from private consumption c, labor l, a public good g, and political favors $\tilde{f}_i$. These can be thought of as pork spending (bridges to nowhere), nepotism, bribes, or targeted public goods. Generically, we think of these as 'corrupt public practices.' That is activities that benefit a subset of the population at the expense of all groups. The instantaneous utility of an agent in group i satisfies

$$U(c, l, g, f_i) = u(c, l) + \pi v(g) + \tilde{f}_i, \tag{1}$$

where π denotes the weight of pure public goods relative to the consumption-leisure aggregate. Agents discount the future at a rate β. In line with small open economy models of sovereign default, we consider imperfect capital markets that prevent individuals from using assets to smooth aggregate fluctuations. This is akin to assuming that agents are hand-to-mouth—as the model abstracts from idiosyncratic shocks—implying that $c = wl$, with w denoting the economy-wide wage rate.

Firms produce a single non-storable consumption good, y, using the linear technology $y = h(z, d)l$. The economy is subject to aggregate TFP shocks, where $z \in Z$ follows a first-order Markov process $\mu(z'|z)$. We denote the long-run value of TFP by $\bar{z}$. When $z > \bar{z}$, the economy is in *good times*, and when $z < \bar{z}$ the economy is in a *recession*. The function $h(z, d)$, increasing in z, takes different forms in states of repayment ($d = 0$) and default ($d = 1$).

Firms are competitive and maximize profits, subject to a distortionary policy τ,

$$\max\,(1 - \tau)y - wl.$$

Policy τ can be interpreted as labor income taxes, revenue taxes, or other costs proportional to the firm's output which are mandatory to operate the production technology and generate revenues for the government (permits, fees, etc.). The behavior of international lenders is modeled following Arellano (2008). Specifically, we assume that an infinite number of identical risk-neutral international lenders are present. These lenders can borrow and lend at a risk-free rate r from the international capital market and can also buy government debt at a price q. Because this is a perfectly competitive market, lenders earn no profits in equilibrium.

3.2 Government Sector

The government collects revenue $Rev(\tau) = \tau y$ and uses its proceeds to fund public goods and political favors. In addition, it can issue one-period non-contingent real bonds that can be bought and sold in international markets. The net resources raised by the government from capital markets are denoted by $qb' - b$, where a positive value of b' is the face value of the current debt of the government and q is the price of new bonds. The government can choose to strategically default on debt b, a decision we denote with $d = 1$. If the government reneges on its current debt obligations, it is excluded from international credit markets for a stochastic number of periods determined by probability θ, with credit standing downgraded to $\Omega = 0$. Since $b' = 0$ in such case, it must finance expenditures solely with $Rev(\tau)$ until access to markets is re-gained. The government budget constraint is

$$Rev(\tau) - g + (1-d)\Big[qb' - b\Big] \geq \sum_i f_i, \tag{2}$$

with $f_i \geq 0$ and $\sum_i f_i$ corresponds to the total amount of resources devoted to political favors. We assume that engaging in corrupt practices is costly: For each dollar taken from the budget, only a fraction $1/\phi$ can be consumed. Hence, when a policymaker spends f_i dollars in favors, agents only receive $\tilde{f}_i = \frac{f_i}{\phi}$. The institutional variable ϕ is our model counterpart for 'control of corruption.' Higher values correspond to better technologies used to crack down on corruption in the public sector.

When the government borrows from abroad, additional resources (measured in consumption goods) are imported. When the government pays down debt, resources are exported. The aggregate resource constraint is

$$y = c + g + \sum_i f_i + nx,$$

where $nx = x - i$ denotes net exports (e.g. exports minus imports). Since there is a single production good in the world, the exchange rate is equal to 1 at all times. In our model, trade deficits have a one-to-one mapping with government deficits.

Bargaining process

Each group in society has a leader with political influence, who has a 'seat' in the bargaining table where policy is decided. Policy proposals need the support of enough leaders, $m \leq n$, to be implemented. This process gives each one of them some veto power and opens the possibility for corrupt public practices. We refer to m as the size of the 'minimum winning coalition' or mwc. We allow for m to change stochastically over time, capturing the evolution of representation and accountability illustrated by Figure 4. More specifically, we assume that m follows a first-order Markov process with transition probability $p(m'|m)$. We define permanent increases or decreases in its long-run average $\bar{m}$ as "regime changes," reflecting fundamental shifts in the political landscape or institutional structures. On the other hand, temporary shocks to m are interpreted as fluctuations in the balance of power that may arise due to electoral outcomes (where one group controls all branches of government) or the fluctuating relative importance of certain groups at different periods of time.

We use the bargaining protocol from Azzimonti and Mitra (2023), which extends the dynamic legislative bargaining model (with exogenous status quo) of Battaglini and Coate (2008) to incorporate strategic default. Group leaders meet at the beginning of a period and one of them is chosen (at random) to make a policy proposal. Since individuals are identical in all regions, the identity of the leader of each group is irrelevant. A proposal is given by

$$\Phi(\Omega) = \begin{cases} \{\tau, g, b', d, f_1, f_2, ..., f_n\} & if\ \Omega = 1 \\ \{\tau, g, f_1, f_2, ..., f_n\} & if\ \Omega = 0 \end{cases} \tag{3}$$

When the country is excluded from credit markets, $\Omega = 0$, it is 'in default', indexed by $d = 1$ and unable to borrow $b' = 0$. Otherwise, when $\Omega = 1$, d and b' are free choices.

If the proposal succeeds in obtaining the support of m leaders, the policy is implemented. If the proposal fails to obtain enough support, leaders move to the next proposal round in which a new proposer is chosen at random. If no agreement can be reached in $T \geq 2$ proposal rounds, an outsider is appointed to choose a reference (symmetric) policy.

3.3 Politico-Economic Equilibrium

The aggregate state variables at the outset of any period are the stock of debt b, the TFP shock z, the size m of the mwc, and the credit standing $\Omega \in \{0, 1\}$. Let $\mathbf{s} = \{z, m, \Omega\}$, and the full state-space be denoted by $\Pi = \{\mathbf{s}, b\}$, We start by describing a competitive equilibrium given government policy $\Phi(\Omega)$. In equilibrium, variables depend on $\Phi(\Omega)$ and Π. Throughout the paper, we write Φ (omitting Ω) to simplify notation.

Firms maximize profits, implying that $w(\mathbf{s}, \Phi) = (1 - \tau)h(z, d)$. To characterize the agents' problem, we make further assumptions about the utility functional forms.

Suppose that $u(c, l)$ is of the GHH family (see Greenwood et al. (1988)),

$$u(c, l) = \frac{1}{1-\sigma}\left(c - \frac{l^{1+\gamma}}{1+\gamma}\right)^{1-\sigma} \quad \text{and} \quad v(g) = \frac{g^{1-\sigma}}{1-\sigma},$$

where $\sigma > 0$ captures the degree of risk aversion and $\gamma > 0$ represents the Frisch elasticity of labor supply.

Because agents are hand-to-mouth, it is easy to show that the individual labor supply $l(\mathbf{s}, \Phi)$ and optimal consumption $c(\mathbf{s}, \Phi)$ are independent of debt decisions,

$$l(\mathbf{s}, \Phi) = [h(z, d)(1 - \tau)]^{\frac{1}{\gamma}} \quad \text{and} \quad c(\mathbf{s}, \Phi) = (1 - \tau)h(z, d)l(\mathbf{s}, \Phi).$$

As a result, aggregate output is also independent of debt, $y(\mathbf{s}, \Phi) = nh(z, d)l(\mathbf{s}, \Phi)$.

International lenders make zero profits when $\Omega = 1$ (that is, when the country can actually borrow). Their break-even bond prices satisfy

$$q(\mathbf{s},\Phi) = \int_{(z',m')\in\Psi(z',m')} \left[\frac{1-d(z',m',b')}{1+r}\right]\partial z'\partial m'|(z,m), \tag{4}$$

where $\Psi(z',m') = \{(z',m') : d(z',m',b') = 0\}$ is the repayment set for the government and $d(z',m',b')$ is tomorrow's equilibrium default choice (which only depends on tomorrow's exogenous states and today's borrowing in the political equilibrium, as will become clearer in the next section). Note that q depends on s, since z and m are Markov processes.

It is useful to define the budget balance excluding favors

$$B(\Pi,\Phi) = \begin{cases} Rev(\mathbf{s},\Phi) - g + (1-d)\Big[q(\mathbf{s},\Phi)b' - b\Big] & if\ \Omega = 1 \\ Rev(\mathbf{s},\Phi) - g & if\ \Omega = 0 \end{cases} \tag{5}$$

with $Rev(\mathbf{s},\Phi) = \tau y(\mathbf{s},\Phi)$. The budget constraint of the government collapses to $B(\Pi,\Phi) \geq \sum_i f_i$.

We focus on a symmetric Markov-perfect equilibrium, implying that any proposer in round $k \in \{1,2,...,T\}$ selects identical policies. The full decision-making problem of the proposer, which is quite heavy in terms of notation, is relegated to Appendix B.1. Since the other $n-1$ legislators are ex-ante identical, the $m-1$ coalition members needed to pass the legislation are randomly selected from them. To secure their consent, the proposer offers favors $f_i = f$ to members of the selected mwc and $f_i = 0$ to members outside of it. This, again, is a result of symmetry. The proposer keeps the remaining $f_p = B(\Pi,\Phi) - (m-1)f$, which may be higher than what is given to other mwc members, $f_p \geq f$.

Following Azzimonti and Mitra (2023), we show in Appendix B.2 that the proposal is accepted immediately, in $k = 1$, and that the proposer's problem is equivalent to one where the welfare of the average member of the mwc is maximized[21]

$$\begin{aligned} &\max_{\Phi} U\Big(c(\mathbf{s},\Phi), l(\mathbf{s},\Phi), g\Big) + \frac{B(\Pi,\Phi)}{\phi m} + \beta \mathbb{E}_{\mathbf{s}'} J(\Pi') \\ &\text{s.t.} \quad B(\Pi,\Phi) \geq 0, \end{aligned} \tag{6}$$

with $J(\Pi')$ denote the expected value of continuation utility for the proposer given that the next period's Markov-perfect equilibrium policy is $\Phi'(\Pi')$,

$$J(\Pi') = U\Big(c(\mathbf{s}',\Phi'(\Pi')), l(\mathbf{s}',\Phi'(\Pi')), g'\Big) + \frac{B(\Pi',\Phi'(\Pi'))}{\phi n} + \beta\,\mathbb{E}_{\mathbf{s}''} J(\Pi''). \tag{7}$$

Note that the left-hand side of the objective function in eq. (6) is not equal to $J(\Pi)$, from eq. (7). This happens because the proposer can use the budget to keep f_p for certain in the current period. In the future, the proposer may not even be in the mwc, implying that he or she would receive less in expectation (see Appendix B.2 for a derivation of the objective function above). This force will create incentives for the proposer to over-spend and over-borrow relative to a benevolent planner, a feature already highlighted in Battaglini and Coate (2008). In this paper, we emphasize how this characteristic also affects incentives to default, and how these incentives become stronger in the shadow of corruption (low ϕ) and institutional weakness (low $\bar{m}$).

The size of the mwc, given by m, is a key parameter of our model: it captures how many votes the proposer

[21] A key assumption for the proof to go through is the linearity in f. Most dynamic bargaining papers, including Battaglini and Coate (2008) and Barseghyan et al. (2013), need to make this assumption to make the problem tractable.

needs in choosing a policy. As m grows closer to n, the policymaker must internalize the preferences of all members of society. When $m = n$, there is full representation and accountability, and institutions are strong. On the contrary, when $m = 1$ there are no independent institutions and the proposer chooses an allocation that devotes significant resources towards her group. Due to the common pool problem, where all members of society contribute towards government revenues, but only one group can appropriate a large share of it, this situation corresponds to one with low $R\&A$ in the data. Similarly, a government can have varied degrees of control over corruption. A combination of low $R\&A$ and low CC implies maximum incentives to divert resources towards favors, as evident from eq. (6). As we will show later, institutional reform through a change in $\bar{m}$ or ϕ will have different implications for the continuation value, and hence different equilibrium outcomes.

3.4 Characterization and intuition

Suppose that $\gamma, \sigma \geq 1$ and $h(z,d) = z$ (this last assumption matters for calibration, but not at this point). It is easy to show that $Rev(\tau, z) = \tau n z^{1+\frac{1}{\gamma}}(1-\tau)^{\frac{1}{\gamma}}$. Recall that the budget devoted to corrupt public practices is $B = Rev(\tau, z) - g + (1-d)\Big[q(z,m,b')b' - b\Big] \geq 0$. There are two possibilities: either the proposer promises political favors in exchange for support, $B > 0$, or not $B = 0$ (so the constraint binds).

Case 1: When $B > 0$,

$$g^*(m) = (\pi \phi m)^{\frac{1}{\sigma}}.$$

Public good provision is independent of debt and default decisions and increases with the size of the mwc. This implies that societies with weak institutions, $m << n$, will tend to under-provide public goods. Taxes are chosen to equate the marginal cost of taxation in terms of consumption (net of the disutility of labor) to the marginal increase in the provision of political favors:

$$\left(\frac{\gamma}{1+\gamma}[(1-\tau)z]^{\frac{1+\gamma}{\gamma}}\right)^{-\sigma} = \frac{n}{m}\left[1 - \frac{\tau}{\gamma(1-\tau)}\right] \quad => \tau^*(m,z).$$

The resulting tax rate is also independent of debt decisions. When $m < n$, the marginal gain in political favors associated with an increase in τ is larger than that of a planner. As a result, societies with weak institutions will tend to over-tax and over-spend on political favors relative to the planner's preferred allocations. The expressions above illustrate that whether the country is in default or not does not affect g or τ in this case. Additional borrowing is used only to increase the budget devoted to corrupt public practices. This also implies that the incentives to repay debt are unaffected by public good provision and tax distortions.[22]

Case 2: When $B = 0$, that is no longer true. Here, borrowing and default decisions affect the provision of g:

$$g = Rev(\tau, z) + (1-d)\Big[q(z,m,b')b' - b\Big],$$

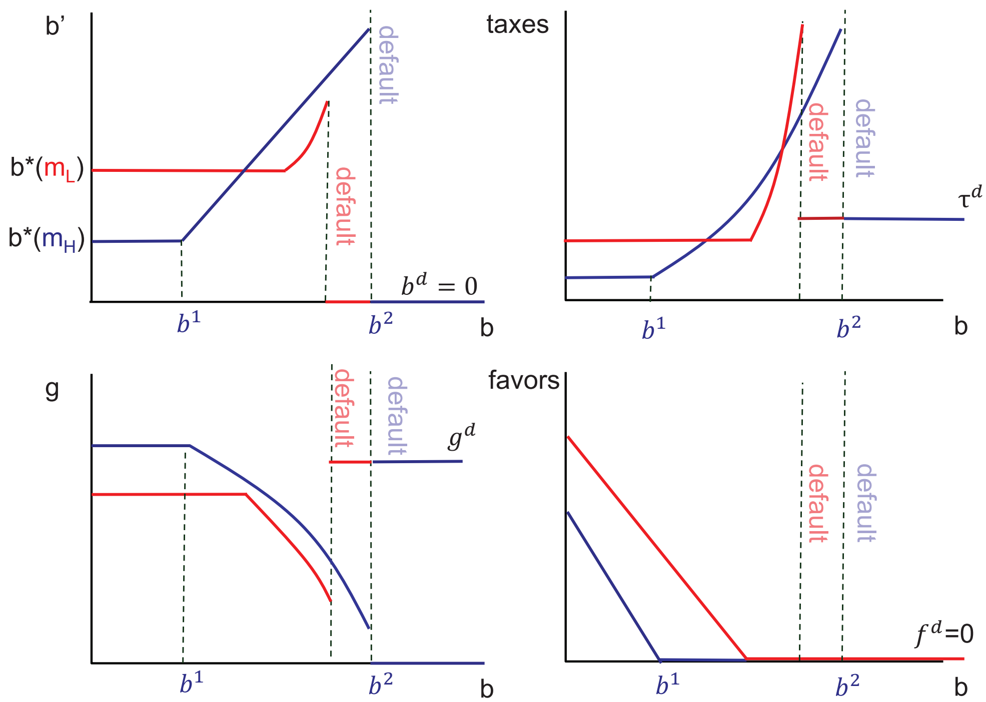
[22]These results are a direct consequence of the linearity in f_i. Relaxing this assumption would make policy functions smoother, and the results less stark. However, the bargaining model's level of complexity grows significantly with non-linearity.

and taxes (through g),

$$\left(\frac{\gamma}{1+\gamma}[(1-\tau)z]^{\frac{1+\gamma}{\gamma}}\right)^{-\sigma} = \pi g^{-\sigma} n \left[1 - \frac{\tau}{\gamma(1-\tau)}\right].$$

The LHS of the equation above is the same as before, but the RHS is now the marginal gain in additional public good provision financed by an additional tax dollar. The only effect of m in this case comes through the endogenous price of debt $q(z, m, b')$. When the country is in default ($d = 1$), however, g and τ are independent of institutional strength.

Figure 5: Policy functions for borrowing, taxes, government spending, and political favors



To better understand how the strength of institutions affects fiscal policy in the political equilibrium, we display two sets of policy functions, for $m_L < m_H$, in Figure 5. The figure is constructed assuming that governments have access to financial markets, $\Omega = 1$. We cannot derive all the theoretical properties of the model (as the decisions to borrow and default are quite involved analytically), so these functions are stylized representations of those obtained through numerical simulations.

First, let us focus on the functions in blue, corresponding to a government with a high value of m_H. When b is low, the proposer chooses $B > 0$, constant taxes $\tau^*(m_H, z)$ and public good provision $g^*(m_H)$, as derived in Case 1. From any $b < b^1$, debt jumps immediately to $b^*(m_H)$, reflecting the proposer's eagerness to spend fiscal capacity on political favors, even if that means financing them via borrowing. At b^1, the constraint on B starts to bind and $B = 0$ thereafter. As the stock of debt rises and, if more debt is to be used to extend political favors without changing g or taxes, the price of bonds declines steeply (not shown). This makes it more costly to finance expenditures via public debt. As a result, the government is forced to increase taxes along with higher borrowing, and lower public good provision in order to pay back debt. At b^2, rolling over the level of debt becomes too costly, so the country defaults (setting $d = 1$). At this point, $b = b' = 0$, freeing fiscal capacity to reduce taxes to τ^d and increase public spending to g^d. Note that the proposer does not waste resources on political favors when in default. This is because policymakers agree on policy when in autarky in this example[23].

Now compare the blue policy functions (for m_H) with the red ones (for m_L). When debt is low, policymakers in countries with low mwc requirements spend more on political favors and less on g, as derived above (see Case 1). In addition, $b^*(m)$ is larger, implying that countries where institutions are weaker, borrow much more when it is cheap to do so. This implies that debt accumulates faster. The price schedule for debt is also steeper as the same debt implies higher chances for default for m_L. Repayment implies cutting back on favors that are valued more. This can be seen by the fact that the default point moves to the left (e.g. $b^2(m_L) < b^2(m_H)$). In other words, countries with weaker institutions are unable to sustain as much debt as those with higher m. This is consistent with the data in Figure 3, where we found that countries with larger R&A scores had lower default risk. In Appendix 5, we also show that they also have higher levels of debt/GDP in our sample. Finally, note that these differences arise only when corrupt practices give private gains to the proposer and other members of the mwc. As ϕ grows, the gains from providing $B > 0$ dissipate and policies converge to those of a benevolent planner. Hence, weak institutions only distort policies and allocations in the shadow of corruption in this environment.

4 Quantitative Model

This section summarizes the calibration strategy and model fit. In addition, we illustrate how governments with different mwc sizes respond to a recession. Counterfactual experiments are presented in subsequent sections.

4.1 Calibration

We calibrate the model to Argentina during the period 1993-2022. A list of the economic and fiscal variables used, together with data sources, can be found in Appendix A.3. We start by describing the exogenously determined parameters and then move to the calibrated ones.

A period in the model is one quarter[24]. The risk aversion parameter in the utility function is set to $\sigma = 2$, following the literature. The inverse of the Frisch elasticity of labor supply is $\gamma = 2$. The risk-free interest rate, $r = 0.55\%$, equals the value of the real 3-month U.S. T-bill interest rate for the period under consideration. The discount factor, β, is set to match the inverse of the gross risk-free rate in the model $\beta = \frac{1}{1+r} = 0.9945$. It is worth noting that the discount factor takes the same value as in standard macroeconomic models. We do not need to assume an extremely low value for β—as most sovereign default models do[25]—, because the effective degree of impatience in our model depends on m/n. This will be explained in more detail below. We fix the corruption parameter $\phi = 1$ in the benchmark model, but later relax this to study the effects of corruption.

[23]This also implies that the values of g and τ obtained in our model right after default are identical to the ones that would arise in a standard default model without politics (e.g. one where choices are made by a benevolent planner). This is developed further in Appendix C.1, where we contrast our bargaining model to the standard case, but assuming a lower discount factor as in Arellano (2008).

[24]We convert annual data to quarterly series by repeating the annual value for each quarter when quarterly data is unavailable.

[25]For example, Arellano (2008) sets β to 0.8.

The exogenous productivity shock is assumed to follow an $AR(1)$ process of the form:

$$z_{t+1} = (1 - \zeta_z)\bar{z} + \zeta_z z_t + \epsilon^z_{t+1} \tag{8}$$

where $\mathbb{E}\epsilon^z_{t+1} = 0$ and $\mathbb{E}(\epsilon^z_{t+1})^2 = \sigma^2_z$. The income process parameters ζ_z and σ_z are chosen by fitting the above AR(1) process to real GDP per employed person, per year (HP-filtered using a parameter of 100). The average value of the process $\bar{z}$ is normalized to 1. The fitted AR(1) process is discretized to 21 possible realizations of the productivity shock using Tauchen and Hussey (1991). Table 2 reports the resulting values of ζ_z and σ_z used in the simulations.

In terms of the political process, we assume that there are $n = 20$ groups in the population. This is a normalization trading off computational accuracy and time[26]. The size of the mwc, represented with m, is also assumed to follow an AR(1) process of the form:

$$m_{t+1} = (1 - \zeta_m)\bar{m} + \zeta_m m_t + \epsilon^m_{t+1}. \tag{9}$$

Similar to the productivity shock, we assume $\mathbb{E}\epsilon^m_{t+1} = 0$ and $\mathbb{E}(\epsilon^m_{t+1})^2 = \sigma^2_m$. The values for ζ_m and σ_m are chosen by estimating an AR(1) process on the normalized series of R&A for Argentina for the period under consideration, displayed in Figure 4. Inspection of this figure suggests that R&A follows a mean reverting process for most of the samples. We first normalize the data to take values between 1 and 20, and discretize the fitted AR(1) process to take integer values in the same range. This maps the data to the model, where a higher value of m would imply more representation and accountability in the fiscal policy decision process.

Table 2: Calibration Targets

Parameter	Value	Target	Description
σ	2		CRRA
γ	2		Frisch Elasticity
β	0.9945	$\frac{1}{1+r}$	FOC
r	0.0055		90 day U.S. Treasury
θ	0.0385	6.5 Years of Exclusion	
$\bar{z}$	1		Normalized
n	20		Normalized
ϕ	1		Normalized
ζ_z	0.925	Persistence Real GDP	AR(1)
σ_z	0.017	Volatility of Real GDP	
ζ_m	0.954	Persistence of R&A	AR(1)
σ_m	0.234	Volatility of R&A	
α_0	-0.364	$\mathbb{E}$(Spreads) = 7.05%	Jointly Calibrated
α_1	0.403	$\frac{\text{Debt}}{\text{GDP}} = 50\%$	
π	1.2	$\frac{\text{Spend}}{\text{Y}} = 0.14$	
$\bar{m}$	4.77		

[26] It is without loss of generality since what matters for computational results is the ratio m/n.

Following Chatterjee and Eyigungor (2012), the probability of re-entry into the market following default, θ, is set to 0.0385. This is approximately equal to $\frac{1}{\theta} \cong 26$ quarters (6.5 years) of exclusion from financial markets after a default event. This value is consistent with the estimates by Gelos et al. (2011) and Richmond and Dias (2009). We also assume that default entails a loss in productivity (see Rose (2005), Bocola (2016)). While we do not formally model these, we take a reduced-form approach following Chatterjee and Eyigungor (2012). In particular, we assume that labor productivity takes the following form[27]:

$$h(z,d) = \begin{cases} z & \text{if } d = 0 \\ z - \max\{0, \alpha_0 z + \alpha_1 z^2\},\ \alpha_1 \geq 0 & \text{if } d = 1 \end{cases} \tag{10}$$

We calibrate α_0, α_1, $\bar{m}$ and π jointly by minimizing a quadratic loss function to match: average spreads, average of external debt to GDP, and the average value of government spending to GDP in the data. We use 150 equally spaced grid points for the borrowing level ranging from 0 to 100 percent of average GDP. The model is simulated for one million model periods. The values of the initial and final five thousand periods are discarded before using the series to compute the relevant moments. The bottom of Table 2 reports the parameters resulting from the calibration procedure.

4.2 Model Fit

We compute long-run moments from the simulation, and report their values in the second column of Table 3 along with their data counterparts. The numerical algorithm is described in the Section 2 of the Online Appendix. The moments reported below the horizontal line are the ones targeted in the calibration, and the ones above the line were not matched by design.

The model performs fairly well in matching the data. An important feature of our model is that we can match the *pro-cyclicality* of fiscal policies earlier documented by Talvi and Végh (2005). These are reflected by the strong positive correlation between government spending and GDP, as well as the negative correlation between net exports and GDP. A $\rho\left(\frac{NX}{y}, y\right) < 0$ means that in recessions (e,g, a decline in y), $NX = x - i$ increases. Because in our model next exports are just external borrowing, this implies that the country borrows in booms, not in recessions (which is the opposite of what optimal tax-smoothing behavior would prescribe). The relative volatility of consumption to output is greater than 1 (as in most emerging economies) and close to that observed in Argentina. Tax rates in the model are almost acyclical, as in the data, for the period under consideration[28].

There are two dimensions where our model does not perform as well. First, the volatility of output is a bit lower in the model than it is in the data. Second, the correlation of spreads to GDP is smaller and of the opposite sign to that in the data. The latter might be a consequence of using short-term debt, which results in a steep bond price schedule. In a recession, debt reduction dominates the effects of increasing spreads to avoid a default. Another reason could be that the series for m and z are uncorrelated in the model, but could be correlated in the data. See Mendoza and Yue (2012b) for a theory that can deliver counter-cyclical spreads.

[27]The curvature of the default cost function is disciplined by the parameters α_0 and α_1. If $\alpha_0 > 0$, and $\alpha_1 = 0$, then the cost of default is proportional. If $\alpha_1 > 0$, and $\alpha_0 = 0$, then the cost rises more than proportionally to the rise in productivity. If $\alpha_0 < 0$, and $\alpha_1 > 0$, then for $z < -\frac{\alpha_0}{\alpha_1}$, the default cost is 0, but the cost rises more than proportionally for higher realizations of z.

[28]This calculation also includes periods of default.

Table 3: Model Fit

Moment	Data: Argentina	Benchmark	
$\frac{\sigma(c)}{\sigma(y)}$	1.26	1.29	
ρ(y,c)	0.97	0.81	Pro-cyclical
ρ(y,g)	0.79	0.70	Pro-cyclical
$\rho(\frac{NX}{y}, y)$	-0.57	-0.37	Pro-cyclical
ρ(y,tax)	0.03	-0.004	
$\rho(r - r^{*}, y)$	-0.30	0.08	
$\sigma(y)$	5.0%	3.4%	
σ(Spreads)	2.9%	4.7%	
$\mu\left(\frac{Debt}{y}\right)$	50%	50%	Matched by Construction
$\mu\left(\frac{g}{y}\right)$	14%	13%	Matched by Construction
$\mu(r - r^{*})$	7.05%	7.11%	Matched by Construction

4.3 Tax smoothing in the shadow of corruption

The canonical models determining *optimal* fiscal policy state that the government should increase debt rather than increase taxes in response to a temporary shock (that lowers revenues or increases spending needs). This results from the assumption that raising tax revenue generates distortions in the economy, whereas selling bonds does not (see Barro (1979), Lucas and Stokey (1983)). The "tax smoothing" hypothesis implies that government debt should be counter-cyclical (borrow in recessions) and taxes should exhibit low volatility. Taxes should eventually rise to repay the debt, but should not track the shock process.

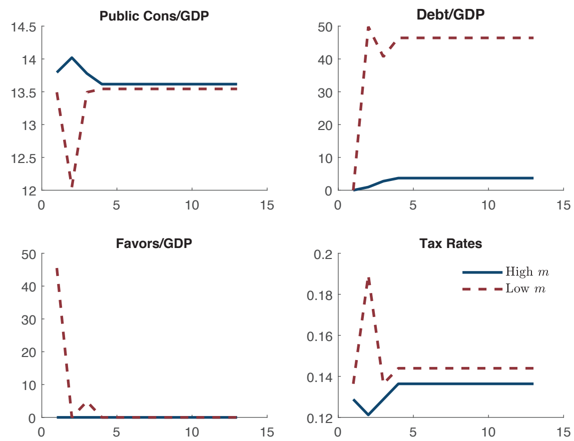
In this model, the correlation between the fiscal balance (revenues minus expenditures) and output is -0.39 instead: the government borrows in booms rather than in recessions, a pro-cyclical policy response. To understand intuitively why this happens, we simulate the response to a negative TFP shock in period 2, assuming that aggregate productivity recovers immediately in period 3. In this experiment, we keep the size of the mwc constant at its long-run average for the calibrated economy, $m_t = \bar{m} = 4.77$. We analyze the response for two different initial levels of debt: a low one $b_0 < b^1$ and a high one $b_0 > b^1$.

Low Initial Debt

Suppose that the economy starts with no initial debt, $b_0 = 0$. This corresponds to the case where $b_0 < b^1$ in Section 3.4. The evolution of the endogenous variables for our benchmark economy is displayed in Figure 6, with a dashed-red line. In period 0, before the recession hits, borrowing (defined as qb') jumps immediately and corrupt public practices skyrocket (favors/GDP go up). The economy enters the recession (period 2) with high debt, so spreads rises to more than 6 percent. Unable to borrow cheaply, policymakers sharply decrease the provision of public goods and hike taxes. This reduces the labor supply, since $l = [h(z,d)(1-\tau)]^{\frac{1}{\gamma}}$ in equilibrium, exacerbating the effects of the negative shock: GDP not only falls because z goes down but also because higher taxes reduce the labor supply l. Rather than stabilizing the economy, the government *amplifies* the effects of a recession.

It is useful to contrast this to the response of a government with stronger institutions. To that end, we

Figure 6: Response to a recession, low b_0



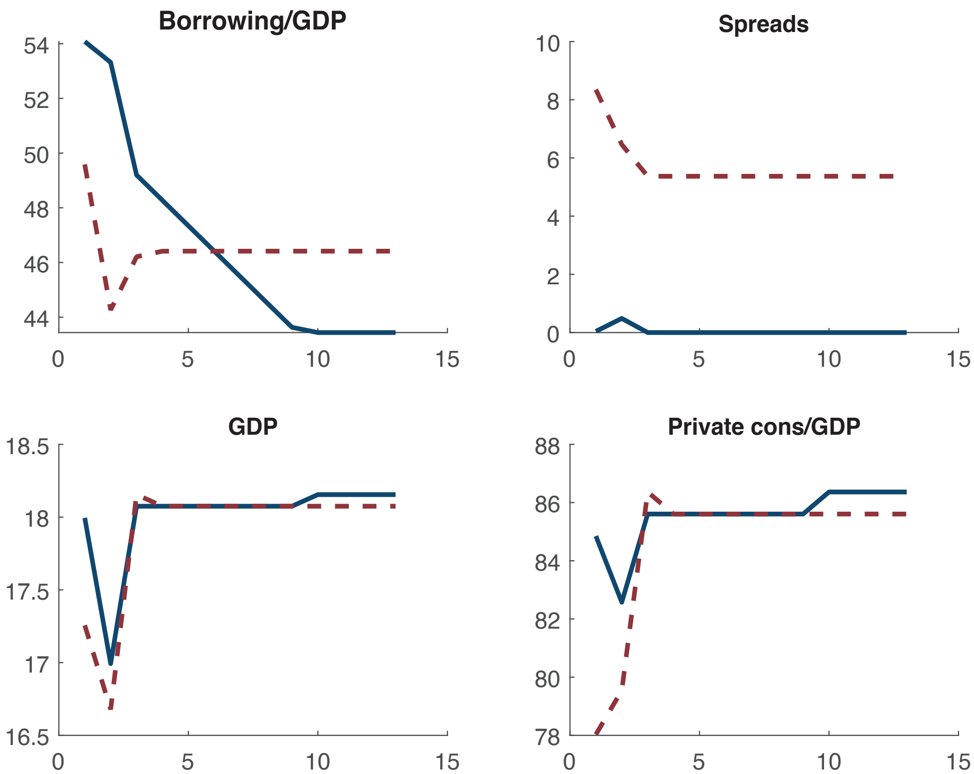
re-compute the model assuming $\bar{m}_H = 20$ and repeat the experiment (time series displayed in blue). As described in Section 3.4, a government with higher m behaves more responsibly. First, fewer resources are devoted to corrupt public practices (favors are zero in this case). Second, borrowing doesn't increase as much initially. When the recession hits, the government can raise debt levels at a lower cost than in the benchmark case (i.e., spreads do not jump), so it can afford to lower τ, as tax-smoothing would prescribe. The decline in GDP is smaller, indicating the government's role in stabilizing the economy.

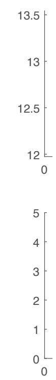
High Initial Debt

We conduct the same exercise assuming initial debt to be about 70 percent of GDP, corresponding to $b_0 \in (b^1, b^2)$ in Section 3.4. Figure 7 shows that both economies start reducing debt in the first period, but the benchmark economy (in red) does this at a much faster rate because of the higher spreads it faces. The benchmark economy imposes higher taxes and has lower GDP and private and public consumption. When the recession hits in period 2, taxes increase considerably more than in the $\bar{m}_H$ economy, even though its legacy debt in the current period is lower. This happens because markets price a higher default probability in the economy with weak institutions, despite its lower debt levels.

Once the benchmark economy recovers, taxes drop as the cost of borrowing becomes lower, and debt stabilizes at a constant level of around 45 percent of GDP. Instead of further reducing the stock of debt, the government chooses to spend the additional resources on political favors, which increase sharply after the recession ends. In the case of high debt, there is imperfect tax smoothing in both economies. This shows up in the decline in both private and public consumption to GDP ratios and increases in taxes during the recession. However, the negative effect of the shock is stronger in the economy with weak institutions. For initial debt levels above b^2, the benchmark economy responds to a negative TFP shock by outright defaulting. The impulse responses are omitted, but available upon request.

Figure 7: Response to a recession, high b_0



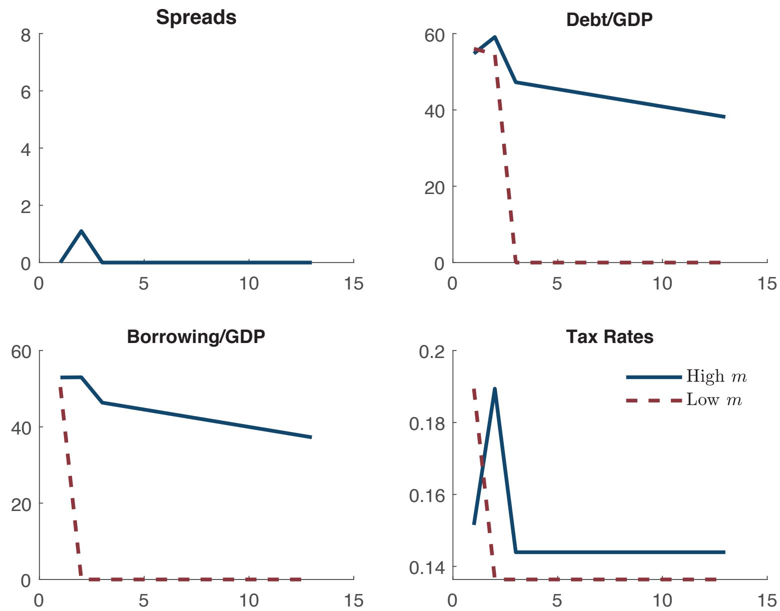


Political Defaults

We now consider a large recession, where TFP goes down significantly in $t = 2$ (for only one period). As in the previous case, countries start the simulation with a high level of debt (of around 70 percent). Figure 8 shows that the country with strong institutions m_H (solid blue) faces a large recession with an austerity measure (e.g. a tax hike), and chooses not to default. Spreads increase marginally.

The country with weak institutions m_L (dashed maroon), instead, prefers to default in the second period rather than adopt an austerity measure. This illustrates the heterogeneous response of countries with different political environments, shedding light on the behavior of European nations following the global financial crisis. Greece (a country with a R&A of 0.89, close to that of an emerging economy) defaulted in March 2012, whereas Germany (a country with a R&A of 1.38, m_H in our example) did not. Our finding is related to Arellano and Bai (2016), which shows that a country subject to fiscal restrictions, in particular the inability to increase taxes, will be forced to default when facing a large recession. They refer to those as 'fiscal defaults.' In their paper, taxes are exogenously fixed through a budget rule. In our model, taxes are flexible but the politician prefers not to increase them when institutions are weak, giving rise to a 'political default.' A planner, in this case, would choose an austerity measure instead.

Figure 8: Response to a "Great Recession," high b_0

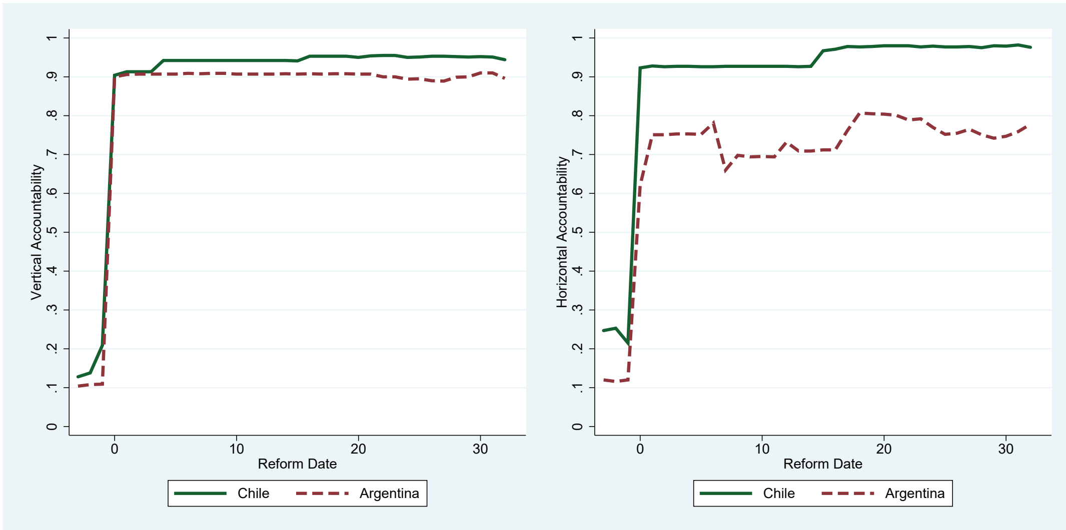


5 Transition to Democracy: Chile vs Argentina

Chile and Argentina were under the rule of military juntas in the late 1970s. Democratic governance was reinstated in both nations in the 1980s, with Argentina holding its inaugural democratic election in 1983 and Chile following suit in 1989. Free speech, free association, free press, and free elections were fully restored in both countries soon after. Yet, intriguingly, the trajectories of fiscal and economic indicators diverged significantly in the post-transition period. Argentina experienced a series of sovereign debt crises and high volatility in output, whereas Chile experienced economic stability and low sovereign debt spreads. Economists attribute this divergence to economic reforms, particularly market liberalization. However, both nations embarked on similar economic liberalization journeys. Political scientists, on the other hand, emphasize the nuances in institutional reforms following the transition to a free democracy as the main differentiating factor. In this section, we evaluate whether heterogeneity in institutional reform can help explain the different trajectories of these two countries. More specifically, we evaluate the macroeconomic consequences of a permanent regime change that strengthens institutions in both countries, but asymmetrically. In particular, we assume that Chile's $\bar{m}$ is higher than Argentina's.

This choice is informed by reading accounts of how the institutional reform was implemented in these countries (see O'Donell (1998) and Wigell (2017)), as well as data on the evolution of sub-components of representation and accountability. In his seminal work, O'Donell (1998) distinguishes between 'vertical' and 'horizontal' accountability. Vertical accountability (VA) refers to the citizens' capacity to hold the government accountable, evidenced by their freedom to establish political parties and engage in open elections. He argues that both countries were successful in ensuring VA. Horizontal accountability (HA), instead, refers to an effective system of separation of powers and checks and balances that impose constraints on how policy can be used to further politicians' private and political goals. According to the author, Chile was

Figure 9: Vertical and Horizontal Accountability



Notes: VA and HA obtained from the V-Dem dataset.

more successful than Argentina in securing strong HA, because key veto players could block the incumbent from governing by decree through legislative and judiciary independence. Figure 9 provides evidence of this claim, by displaying the evolution of of HA and VA from the V-Dem (Varieties of Democracy) dataset[29]. Period 0 indicates the year in which the first democratic election was held. While VA is similar after free elections took place, Argentina's HA is significantly lower than Chile's following the transition.

Simulating the transition

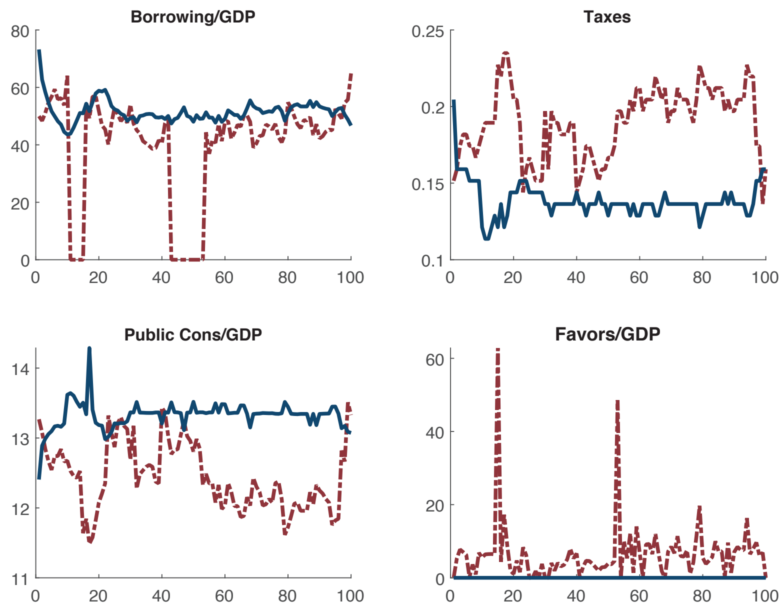
In terms of our model, stronger HA is captured by a larger number of groups with veto power, $\bar{m}$. We computed our model for two different values of $\bar{m}$: Argentina's corresponds to the benchmark of $\bar{m}_A = 4.77$, whereas Chile's is higher at $\bar{m}_C = 8.8$ (with $n = 20$ in both cases). We keep the volatility of m to that of the benchmark[30]. We assume that both countries move from a dictatorship (with $\bar{m}_i = 1$) to a democracy in period 0, and simulate the two economies for 100 quarters assuming that they face the same sequence of TFP shocks. We discipline the initial debt-to-output ratios of the two countries using IMF data for the years in which the institutional reform happened. Argentina's debt/GDP in the initial period is 55.8 percent (e.g. the observed value in 1983) and in Chile is 77.1 percent (e.g. the observed value in 1989).

Figure 10 plots the evolution of fiscal and economic variables. Chile (displayed with a solid blue), has a higher stock of debt when the reform happens, but starts reducing debt/GDP soon after. Its borrowing, $q'B'$, reaches a smaller value than Argentina's (red dashed line) in less than 20 quarters. The decrease in debt obligations is achieved not by increasing taxes in the first period—which are actually smaller than Argentina's— or by initially reducing public spending, but instead by choosing not to waste tax dollars on

[29]Note that our measure of R&A, voice and accountability, includes both dimensions. Another difference is that R&A is based on public perceptions (a de facto measure), whereas VA and HA are constructed from practices that are legally recognized (a de jure measure).

[30]This is also lower for Chile. Lowering the volatility for m will only strengthen our results.

Figure 10: Evolution of policies and allocations to institutional reform



political favors. Faced with a higher number of veto players and a large stock of debt, the Chilean government behaves more *responsibly*: the distortions associated with engaging in corrupt public practices are not worth the benefit to the proposer, as they need to be distributed among a larger mwc. Note that in this experiment we did not change ϕ: the reform did not make it more difficult to engage in corruption, it just made it less appealing. The elimination of f_i is an endogenous choice. Knowing that future policymakers will also not engage in corruption, further increases the incentives to reduce indebtedness today (because of better prices on debt), implying that the proposer is also more *patient* when $\bar{m}$ is higher. As a result of this, the Chilean government can better smooth taxes and consumption when recessions happen (and consistent with Figure 7). Spreads are low for Chile throughout and experiences no default episodes.

Argentina's post-transition experience is completely different. Rather than reducing borrowing, the government quickly engages in corrupt practices: we see positive values of f_i through the simulation. This is consistent with Wigell (2017), who documents that in the beginning of the 1990s, President Menem (Argentina) started manipulating targeted benefits as a lynchpin for mobilizing political support[31]. Intuitively, a lower $\bar{m}$ worsens the common pool problem associated with government expenditures. Fewer members of the mwc benefit from "priority" projects of the government at the cost of taxes and debt repayment obligations on the entire population. Favors are financed with higher taxes, and the government under-spends on public good provision. Markets attribute a higher probability of default to the Argentinean government, so spreads are much higher. When shocks hit, the economy experiences sovereign debt crises (e.g. jumps in spreads) and even two default episodes soon after. In other words, Argentina becomes a country where 'to default is the default.' The experiment points to the reason for significantly higher volatility in policy and

[31] In Chile, by contrast, targeted programs remained under firm technocratic control. We can interpret f_i as the use of targeted programs for political support, as we do not have heterogeneity among agents (e.g. there are no other redistributive programs).

allocations, as well as lower private and public consumption than in Chile.

6 Tax-smoothing: Emerging vs. Developed countries

The experiment in the previous section suggests that the ability to smooth policy and avert sovereign debt crises is strongly linked to institutional strength in our model. Is this insight more general or was it simply an artifact of arbitrary initial conditions or TFP shock realizations? To study this, we simulate the model for a large number of periods and eliminate the first 5000 quarters before computing statistics. The resulting long-run business cycle moments for the two countries are shown in Table 4. Argentina's is displayed in the first column and Chile's in the second. These aim to capture generic emerging economies with varying mwc sizes. In addition, we include a simulation for a hypothetical developed economy with strong institutions by setting $\bar{m} = n$ (this also corresponds to the solution of a benevolent planner without commitment to repay).[32]

Table 4: Business Cycle Moments in the Long Run

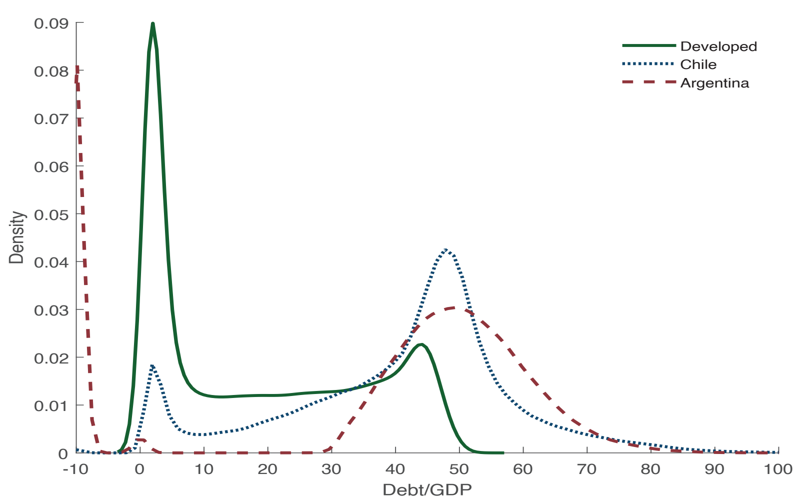
Moment	Argentina $\bar{m} = 4.77$	Chile $\bar{m} = 8.8$	Developed $\bar{m} = 20$
Institutions			
R&A (mean)	4.77	8.80	20.0
Instability (St Dev)	0.78	0.78	0
Favors/GDP (mean)	4.52%	0.10%	0
Fiscal Policy			
Spreads	7.08%	0.06%	0.001%
$\sigma(tax)$	0.04	0.01	0.01
$\sigma(g)/\sigma(y)$	1.49	1.25	0.83
$\mu(b/Y)$	50.4%	41.1%	18.3%
$\rho(TB, y)$	-0.37	-0.20	0.55
$\rho(FB, y)$	-0.31	-0.21	0.65
$\mu(tax)$	17.69	13.77	13.63
Macro Outcomes			
Sp. Crisis	9.85%	4.12%	2.82%
$\sigma(c)/\sigma(y)$	1.29	1.14	0.84

The model predicts monotonicity in fiscal policies and macroeconomic outcomes. Spreads are highest in Argentina, where institutional quality is the lowest. Government spending relative to GDP and tax rates are more volatile in Argentina than in Chile, illustrating differences in their ability to smooth taxes. As expected,

[32] To make things comparable, we are only changing $\bar{m}$ across the three simulations. The TFP shock process and realizations, as well as other parameters in the model are constant. In other words, we are not re-calibrating these economies in each column to match spreads. Because Argentina has the worst outcomes of the three, using it as a benchmark is a lower bound on the ability of economies with stronger institutions to smooth taxes.

Chile exhibits higher volatility of taxes and government spending compared to the developed economy. Our results explain the heterogeneity observed in the data across the spectrum of emerging economies with varying levels of institutional strength. Some of them are serial defaulters, or prone to a spread crisis[33] like Argentina, while others are significantly less prone to such crises like Chile. In our simulations, Argentina has a 9.85 percent chance of falling into a spread crisis, while Chile is much closer to a developed country with a 4 percent possibility of the same. Favors as a percentage of GDP are higher on average in Argentina, implying a significantly higher targeted use of public resources for political gains. This number is negligible in the other two cases with higher $\bar{m}$. The developed economy exhibits a strong positive correlation between trade balance and GDP, but this statistic is negative for both Chile and Argentina. This implies that the developed country can borrow in bad times without having to resort to abrupt tax increases to finance spending. This is also reflected in the sign and magnitude of the fiscal balance (FB) in all the columns, and in the magnitude of the ratio of the volatility of consumption to that of GDP. Highly volatile taxes spill over into private consumption expenditures, making them more volatile than output. That is, policy amplifies the cycle in countries with weak institutions like Argentina.

Figure 11: Ergodic debt distribution.



Why are spreads so much higher in Argentina? Figure 11 displays the long-run distribution of debt to GDP in the three economies. The average value of debt to GDP is highest for Argentina (red dashed line), at around 50 percent. The other two economies have significantly lower values of debt to GDP. Moreover, the developed economy has a significant mass at zero, indicating a desire to save in good times and use debt only in recessions[34]. Argentina defaults very frequently, as seen at the large mass in -10 (the arbitrary coded value for debt under default, for this plot's readability). Chile and the developed economy have no mass at that point, indicating that they rarely default. After gaining market access (post-default, indicated by the small mass at 0), Argentina jumps to a higher value of debt-to-GDP (of around 30 percent). This is

[33]Following Trebesch and Zabel (2017), we code a spread crisis when we observe: (i) an increase in spreads compared to the previous period above the 99th percentile of the spread distribution or (ii) an annualized spread of 10 percentage points or higher.

[34]See Bianchi and Sosa-Padilla (2023) for an alternative environment where accumulating reserves could help smooth business cycle fluctuations.

not the case for Chile, whose average debt-to-GDP ratio is lower (blue dotted line). A mass at 0 implies that in some parts of the state space Chile would even like to save. The developed economy with strong institutions has a distribution of debt/GDP similar to Chile's, the only difference is that the mass at zero is larger[35].

Figure 12: Default regions.

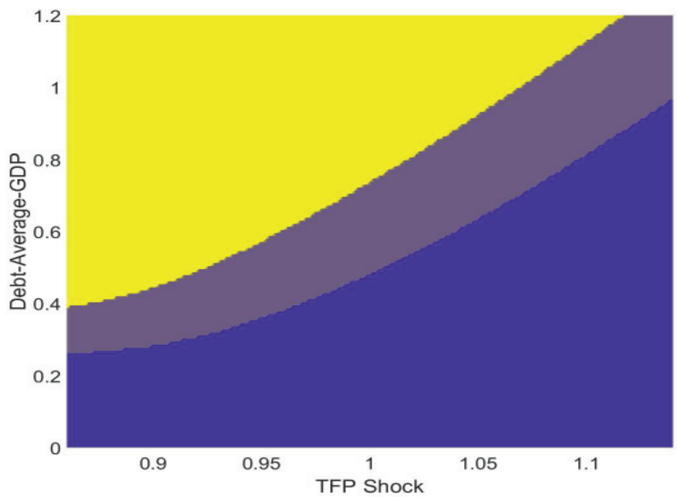


Figure 12 shows repayment and default regions in the state space for the three economies. The darker-shaded regions represent combinations of TFP shocks and legacy debt-to-GDP ratios where the government would optimally repay. The darkest region in the bottom right corner represents the repayment region for Argentina, while the combination of the two darker regions represents the repayment set for the other two countries. If the country finds itself in the light yellow region, it defaults. In addition to having higher debt on average (shown in the previous figure), Argentina defaults at a larger region of the state space, making it more likely to end up in the default region. The default region for the other two economies is the same. However, since Chile still borrows more than the developed economy, it is more likely to default.

This result is driven by two forces that change when $\bar{m}$ increases. On the one hand, stronger institutions reduce the gains of engaging in corrupt public practices (as political favors must be shared with a larger number of groups in society). On the other hand, it also increases the likelihood that the proposer is in the mwc in the future (e.g. receives political favors). This makes the proposer endogenously more patient. In Appendix C.2, we split these two forces by comparing our benchmark to an autocracy where the proposer is in power forever. That is a situation in which $m = 1$, but the autocrat knows it will be in the mwc with probability 1 (which we interpret as the case of China). We invite an interested reader to visit this extension.

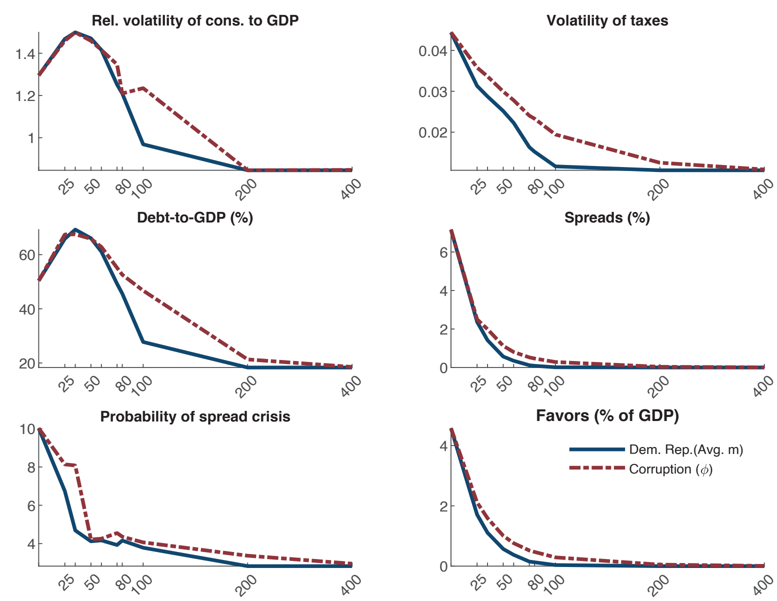
[35]Following the tradition in the sovereign default literature, we computed the model restricting debt to be positive (e.g. ruling out savings). This is not binding in the calibrated benchmark economy (Argentina), but it can become binding in some states of the word for economies with higher values of $\bar{m}$ (such as the two examples we presented above). However, even if we allowed for such behavior, it is unlikely that borrowing will stop completely as in Aiyagari et al. (2002). With imperfect political institutions (lower representation and corruption), the asset floor is endogenously determined, which would make borrowing optimal at most points of the state space.

7 Implementing a 'crackdown on corruption'

In the previous counterfactual experiments, we kept the control of corruption fixed at the normalized value of $\phi = 1$. In this section, we want to study whether a 'crackdown of corruption' (e.g. an increase in ϕ) is as effective as a reform that strengthens representation and accountability (e.g. an increase in $\bar{m}$). An increase in ϕ reduces the benefits from engaging in corrupt public practices, since for each dollar taken from the budget, the proposer and each of the coalition members only get $1/\phi$ of their respective shares. An increase in $\bar{m}$ reduces the per-member amount received through corrupt public policies, but also affects the probability of being part of the mwc in the future. We evaluate which of the two reforms is more palatable to the proposer at different stages of institutional development, and compare this to the welfare gains of an average citizen in the population.

We simulate the economy where alternative values of ϕ or $\bar{m}$ are increased by {10, 25, 35, 50, 60, 75, 80, 100, 200, 400} percentage points. When ϕ is increased $\bar{m}$ is kept constant, and vice-versa. The long-run response of endogenous variables and their volatilities are shown in Figure 13. The dashed maroon line depicts crackdowns on corruption (e.g. increases in ϕ) whereas the blue solid line depicts strengthening the representation and accountability (e.g. increases in $\bar{m}$).

Figure 13: Effect of Increasing m and ϕ by the Same Magnitude on the Endogenous Variables



Notes: Numbers on the x axis indicate percentage increase in $\bar{m}$ from 4.77 or ϕ from 1. When ϕ is increased $\bar{m}$ is kept constant and vice-versa.

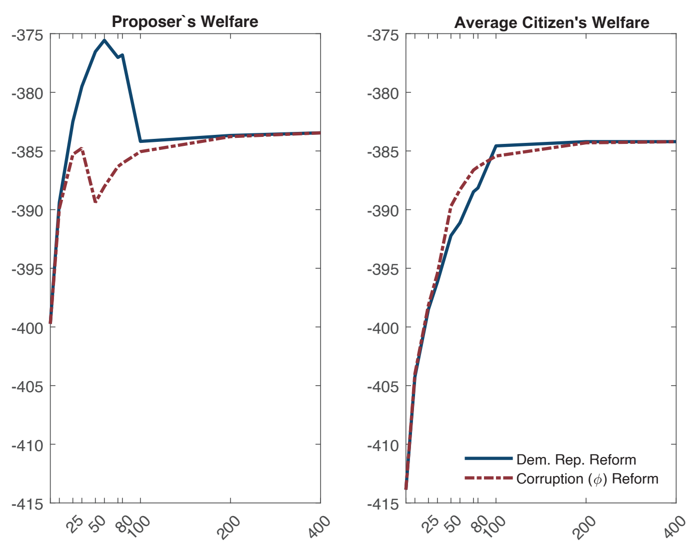
When the percentage increase in $\bar{m}$ is relatively small, debt issuance goes up because the fiscal space increases as a result of a reduction in spreads. A further increase, however, reduces debt and also the relative volatility of consumption to GDP. The volatility of taxes is always decreasing, regardless of whether ϕ or $\bar{m}$

are increased, indicating that either reform improves tax smoothing. Corrupt public practices, spreads and the probability of spread crises also decrease monotonically for all degrees of such reform[36]. An increase in ϕ leads to qualitatively similar implications, but are much smaller in magnitude compared to that of $\bar{m}$ when it comes to reducing the volatility of consumption and taxes. In general, the performance of the economy is more responsive to increases in $\bar{m}$ than to increases in ϕ. Intuitively, volatility in m is equivalent to policy uncertainty as in Azzimonti (2018). In a case with lower $\bar{m}$, each member in the coalition carries more weight, and hence fluctuations in the coalition size lead to larger swings in fiscal policy variables.

7.1 Reform when institutions are weak

The analysis above relates to long-run outcomes under alternative institutional environments. Apart from affecting these, the $\bar{m}$ reform also impacts the intertemporal price of f_i, and hence the desirability of the proposer to implement such reform. Figure 14 shows the welfare of the proposer (left panel) and that of an average citizen (right panel) assuming that either a crackdown on corruption or an $\bar{m}$ reform is implemented in the current period. The value at the origin represents welfare if no reform is undertaken. While the average citizen is always better off when there is less corruption or stronger representation (e.g. welfare is increasing in ϕ as well as $\bar{m}$ reforms), the proposer's welfare is non-monotonic. This happens because the group with greater political power benefits most from institutional weakness in the short run.

Figure 14: Effect of Increasing m or ϕ on Welfare



Notes: Numbers on the x axis indicate percentage increase in $\bar{m}$ from 4.77 (solid-blue) or ϕ from 1 (dashed-maroon). When ϕ is increased $\bar{m}$ is kept constant and vice-versa.

To understand the intuition, it is useful to look at eq.(6): a higher m reduces the weight on favors in the current period. But it also increases the chances of being chosen as a member of the mwc in the following

[36]For spreads, debt-to-GDP and favors average values of a one-million-period long simulation are reported.

period, thereby increasing the chances of receiving political favors in the future[37]. Therefore, the relative price of favors in the current period with respect to the future increases which makes the proposer optimally reduce favors in the current period and back-load it for the future. This price effect is relevant only because favors are positive at a lower level of development– when both the parameters are low. Increasing ϕ however, increases the price of favors both in the current and the future periods keeping the relative price unchanged. At lower values of $\bar{m}$, the effect of a reduction in volatility dominates the relative price effect, and hence the proposer is better off if $\bar{m}$ goes up. Increasing $\bar{m}$ by more than 60 percent leads to a drop in the proposer's welfare. Although there is higher fiscal space, borrowing starts to decline as the utility from favors for each coalition member is lower. Here the relative price effect dominates the volatility effect[38].

Both the rise and the corresponding fall, due to lower favors, in the proposer's welfare are modest when a corruption reform is undertaken. Since favors are directly targeted, such a reform never encourages borrowing as much as in the $\bar{m}$ reform to pay for favors. Consequently this results in lower taxes to finance favors in the ϕ-reform regime. An increase in $\bar{m}$ of 400 percent or more has the same effect on welfare as an increase in ϕ of the same magnitude. At that point, moments look similar to a developed country. They are characterized by relative volatility of consumption to GDP of less than 1, less volatile taxes, lower spreads, and no spending in political favors, as shown in Figure 13.

From the viewpoint of the average citizen, a ϕ reform is associated with higher welfare than a $\bar{m}$ reform in the initial bursts of reform (up to 70 percent). Lower taxes and lower debt allow for higher consumption and welfare. However, given a choice, the proposer would rather implement an $\bar{m}$ reform to attain favorable treatment from the lenders, move to the peak of the proposer's welfare (when $\bar{m}$ is increased by about 60 percent), and stay put. The average citizen would rather have a 60 percent crackdown on corruption, than a 60 percent increase in representation. This shows that political economy frictions may come in the way of an emerging/developing economy transitioning into a developed country. It is important to emphasize that foreign lenders price debt by taking into account only the risk of repayment, without any concern about the use or abuse of the lent resources.

7.2 Reform when institutions are stronger

In this paper, we have reported long-term averages of $R\&A$ and CC for all the countries in our sample. However, over time many of them came out of dictatorships and there had been significant gains in terms of institutional development.

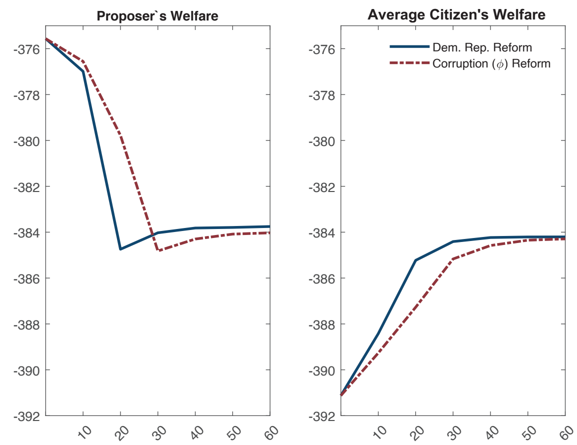
Argentina, for example, shows a current $R\&A$ measure that is 60 percent higher than its long-term average. Given higher accountability, these countries need to make promises of reform to the electorate and carry them out to stay in power and get re-elected. The new president-elect of Argentina, Javier Milei, promises to put an end to corruption during his tenure, along with many other lofty reform promises. In Figure 15 we plot the welfare implications for such promises and evaluate their credibility. More specifically, we evaluate if a corruption reform is a promise that can be fulfilled.

We start from $\bar{m}$ that is already 60 percent higher and also corresponds to the peak for the proposer in Figure 14. Evidently, the proposer would not want to make any reforms at this point, as they would result in lower

[37]The unconditional probability of being a member of the mwc in the following period is given by $\frac{m-1}{n}$.

[38]After an 80 percent reform, welfare starts to marginally improve again because taxes drop, and consumption and GDP rise. At this point, f_i drops to zero.

Figure 15: Effect of Increasing m and ϕ from peak $\bar{m}$



Notes: Numbers on the x axis indicate percentage increase in $\bar{m}$ from 7.63 or ϕ from 1. When ϕ is increased $\bar{m}$ is kept constant and vice-versa.

welfare. Increasing ϕ, however, has an associated welfare loss that is smaller for the proposer. The average citizen, however, would be better off improving democratic institutions. Intuitively, with already favorable treatment from the lenders, a ϕ-reform does not have the price effect that back-loads favors. This results in higher borrowing and higher taxes than a similar $\bar{m}$ reform that results in lower aggregate welfare.

8 Conclusion

We introduce a fiscal policy model that explains how countries with different levels of institutional development respond to macroeconomic shocks, considering the option of strategic default. The model focuses on two key institutional factors: representation and accountability (R&A) and control of corruption (CC). Empirically, we find that countries with strong R&A and CC have higher sovereign bond ratings, less tax volatility, and more counter-cyclical fiscal policies. In contrast, countries with weaker R&A and CC face higher economic volatility and more frequent sovereign debt crises.

Our theory highlights the significant impact of political favoritism in shaping these economic outcomes. When R&A,—represented by $\bar{m}$ (the size of the mwc) in our model—is low, governments tend to allocate excessive resources for political favors, exacerbating the common pool problem. A low CC environment facilitates the misuse of tax revenues for corrupt public practices. Lenders, recognizing these risks, adjust their lending terms accordingly, often leading to higher borrowing costs and fiscal instability in countries with weak institutions. We can explain economic patterns in countries with varying institutional strengths without assuming a lower subjective discount factor, as often done in sovereign default literature. We also generate 'prudent,' counter-cyclical fiscal responses in stronger institutional settings without assuming a debt repayment commitment. Using the quantitative model, we can explain the divergent economic paths of Argentina and Chile following the end of the dictatorship rules in the 1980s. A critical observation from our model is that while both could default at any time, Chile's stronger institutions make default less appealing for its policymakers.

We compare the effects of permanent reforms in representation and accountability (R&A) against sustained anti-corruption efforts. We find that in countries with very low institutional development, small improvements in R&A can lead to increased political favoritism, at the cost of lower welfare to the average citizen. This happens as international lenders, encouraged by perceived improvements, offer better loan terms, but the funds are often used for political gains. Such countries typically cease R&A improvements once they reach a level that benefits the ruling proposer, with no incentive for further reform. We find that in emerging countries, the incentives of the average citizen and those of the politicians never align perfectly. From the perspective of public policy, it is important to keep these considerations in mind. A one-size-fits-all policy may not be the best solution for economies at every stage of development.

References

Aguiar, M. and M. Amador (2014). Sovereign debt. In *Handbook of International Economics Vol 4*, pp. 647–87. North-Holland.

Aiyagari, S. R., A. Marcet, T. J. Sargent, and J. Seppälä (2002). Optimal taxation without state-contingent debt. *Journal of Political Economy 110*(6), 1220–1254.

Andreasen, E., G. Sandleris, and A. Van der Ghote (2019). The political economy of sovereign defaults. *Journal of Monetary Economics 104*, 23–36.

Arellano, C. (2008). Default risk and income fluctuations in emerging economies. *American Economic Review 98*(3), 690–712.

Arellano, C. and Y. Bai (2016). Fiscal austerity during debt crises. *Economic Theory 64*.

Azzimonti, M. (2018). Partisan conflict and private investment. *Journal of Monetary Economics 93*, 114–131.

Azzimonti, M., M. Battaglini, and S. Coate (2016). The costs and benefits of balanced budget rules: Lessons from a political economy model of fiscal policy. *Journal of Public Economics 136*, 45–61.

Azzimonti, M. and N. Mitra (2023). Political constraints and sovereign default. *Journal of International Money and Finance*.

Baron, D. P. and J. A. Ferejohn (1989). Bargaining in legislatures. *American political science review 83*(4), 1181–1206.

Barro, R. (1979). On the determination of public debt. *Journal of Political Economy 82*(6), 1095–1117.

Barseghyan, L., M. Battaglini, and S. Coate (2013). Fiscal policy over the real business cycle: A positive theory. *Journal of Economic Theory 148*(6), 2223–2265.

Battaglini, M. and S. Coate (2008). A dynamic theory of public spending, taxation, and debt. *American Economic Review 98*(1), 201–36.

Bianchi, J., P. Ottonello, and I. Presno (2023). Fiscal stimulus under sovereign risk. *Journal of Political Economy 0*(ja), null.

Bianchi, J. and C. Sosa-Padilla (2023, 7). Reserve Accumulation, Macroeconomic Stabilization, and Sovereign Risk. *The Review of Economic Studies*, rdad075.

Bocola, L. (2016). The pass-through of sovereign risk. *Journal of Political Economy 124*(4), 879–926.

Chang, R. (2007). Financial crises and political crises. *Journal of Monetary Economics 54*(8), 2409–2420.

Chatterjee, S. and B. Eyigungor (2012). Maturity, indebtedness, and default risk. *American Economic Review 102*(6), 2674–99.

Chatterjee, S. and B. Eyigungor (2019). Endogenous political turnover and fluctuations in sovereign default risk. *Journal of International Economics 117*, 37–50.

Choi, J., V. Penciakova, and F. Saffie (2021). Political connections, allocation of stimulus spending, and the jobs multiplier.

Cotoc, J., A. Johry, and C. Sosa-Padilla (2022). Sovereign spreads and the political leaning of nations. Working paper.

Cuadra, G., J. M. Sanchez, and H. Sapriza (2010). Fiscal policy and default risk in emerging markets. *Review of Economic Dynamics 13*(2), 452–469.

Cusato Novelli, A. (2020). Sovereign default, political instability and political fragmentation. *Review of International Economics*.

Drazen, A. and E. Ilzetzki (2023). Kosher pork. *Journal of Public Economics 227*, 105001.

DErasmo, P. and E. G. Mendoza (2016). Distributional incentives in an equilibrium model of domestic sovereign default. *Journal European Economic Association 14*(1), 7–44.

DErasmo, P. and E. G. Mendoza (2021). History remembered: Optimal sovereign default on domestic and external debt. *Journal of Monetary Economics 117*, 969–989.

Espino, E., J. Kozlowski, F. M. Martin, and J. M. Sanchez (2020). Domestic policies and sovereign default. *FRB St. Louis Working Paper* (2020-17).

Gavin, M. and R. Perotti (1997). Fiscal policy in latin america. *NBER macroeconomics annual 12*, 11–61.

Gelos, R. G., R. Sahay, and G. Sandleris (2011). Sovereign borrowing by developing countries: What determines market access? *Journal of international Economics 83*(2), 243–254.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988). Investment, capacity utilization, and the real business cycle. *The American Economic Review*, 402–417.

Hall, G. J. and T. J. Sargent (2020, May). Debt and taxes in eight u.s. wars and two insurrections. Working Paper 27115, National Bureau of Economic Research.

Hatchondo, J. C. and L. Martinez (2010). The politics of sovereign defaults. *FRB Richmond Economic Quarterly 96*(3), 291–317.

Hatchondo, J. C., L. Martinez, and H. Sapriza (2009). Heterogeneous borrowers in quantitative models of sovereign default. *International Economic Review 50*(4), 1129–1151.

Henisz, W. J. (2000). The institutional environment for economic growth. *Economics & Politics 12*(1), 1–31.

Herrera, H., G. Ordonez, and C. Trebesch (2020). Political booms, financial crises. *Journal of Political Economy 128*(2).

Ilzetzki, E. (2011). Rent-seeking distortions and fiscal procyclicality. *Journal of Development Economics 96*(1), 30–46.

Ilzetzki, E. and C. A. Vegh (2008, July). Procyclical fiscal policy in developing countries: Truth or fiction? Working Paper 14191, National Bureau of Economic Research.

Kaas, L., J. Mellert, and A. Scholl (2020). Sovereign and private default risks over the business cycle. *Journal of International Economics 123*, 103293.

Kalemli-Özcan, . and L. Varela (2021). Five facts about the uip premium. Technical report, National Bureau of Economic Research.

Kaminsky, G. L., C. M. Reinhart, and C. A. Végh (2004). When it rains, it pours: Procyclical capital flows and macroeconomic policies. *NBER Macroeconomics Annual 19*, 11–53.

Lucas, R. E. and N. L. Stokey (1983). Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics 12*(1), 55–93.

Lührmann, A., K. L. Marquardt, and V. Mechkova (2020). Constraining governments: New indices of vertical, horizontal, and diagonal accountability. *American Political Science Review 114*(3), 811820.

Martinez, L., F. Roch, F. Roldan, and J. Zettelmeyer (2022). Sovereign debt. *IMF Working Paper 22/122*.

Mendoza, E. G. and V. Z. Yue (2012a, 04). A General Equilibrium Model of Sovereign Default and Business Cycles*. *The Quarterly Journal of Economics 127*(2), 889–946.

Mendoza, E. G. and V. Z. Yue (2012b). A general equilibrium model of sovereign default and business cycles. *The Quarterly Journal of Economics 127*(2), 889–946.

Mihalache, G. (2020). Sovereign default resolution through maturity extension. *Journal of International Economics*, 103326.

Mitchener, K. and C. Trebesch (2023). Sovereign debt in the 21st century. *Journal of Economic Literature*.

O'Donell, G. (1998). Horizontal accountability in new democracies. *Journal of Democracy 9*(3).

Richmond, C. and D. A. Dias (2009). Duration of capital market exclusion: An empirical investigation. *Available at SSRN 1027844*.

Rose, A. K. (2005). One reason countries pay their debts: renegotiation and international trade. *Journal of development economics 77*(1), 189–206.

Sargent, T. J. and F. R. Velde (1995). Macroeconomic features of the french revolution. *Journal of Political Economy 103*(3), 474–518.

Talvi, E. and C. A. Végh (2005). Tax base variability and procyclical fiscal policy in developing countries. *Journal of Development Economics 78*(1), 156–190.

Tauchen, G. and R. Hussey (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica: Journal of the Econometric Society*, 371–396.

Tomz, M. and M. Wright (2007). Do countries default in bad times? *Journal of the European Economic Association 5*.

Trebesch, C. and M. Zabel (2017). The output costs of hard and soft sovereign default. *European Economic Review 92*, 416–432.

Vegh, C. A. and G. Vuletin (2015). How is tax policy conducted over the business cycle? *American Economic Journal: Economic Policy 7*, 327–370.

Wigell, M. (2017). Political effects of welfare pluralism: Comparative evidence from argentina and chile. *World Development 95*, 27–42.

Yared, P. (2019). Rising government debt: Causes and solutions for a decades-old trend. *Journal of Economic Perspectives 33*, 115–40.

Yue, V. Z. (2010). Sovereign default and debt renegotiation. *Journal of International Economics 80*(2), 176–187.

A Data Appendix

We collect data on (1) economic and policy variables, (2) institutional quality, and (3) sovereign credit ratings. Economic and policy data series are annual, ranging from 1990 to 2022, institutional data starts in 1996 and Fitch ratings in 1995. The three sets of variables, their sources, and detrending methods are described below. There are 58 countries for which we have data on all three variables which will be referred to as "benchmark" data, a subset of our full database.

A.1 Institutional Quality

We use two indicators for institutional quality: the strength of Representation and Accountability (R&A) and the control of corruption (CC). These are part of the World Bank's "World Governance Indicators" (obtained from the WGI Website).

- $R\&A$ is the degree of **Representation and Accountability**, proxied by the WGI variable "Voice and accountability," which captures perceptions of the extent to which a country's citizens can participate in selecting their government, as well as freedom of expression, freedom of association, and a free media. The variable ranges between [-2.5,2.5], with lower values indicating weaker R&A.
- CC is **Control of Corruption** (CC), obtained from the WGI variable with the same name, and captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. The variable ranges between [-2.5,2.5], with lower values indicating poorer CC.

Alternative measures: The WGI provides alternative measures of institutional quality. These are 'Government Effectiveness (GE),' 'Regulatory Quality (RQ),' and ' Rule of Law (RL)'. Their long run values are highly correlated with CC: $Corr(CC, RL) = 0.96$, $Corr(CC, GE) = 0.97$ and $Corr(CC, RQ) = 0.93$. Because of this, we simply use CC in our analysis.

In terms of de-facto measures, there is an alternative measure called 'political constraints (PolCon)' developed by Henisz (2000), that we have used in previous work (see Azzimonti and Mitra (2023)). The series in only available until 2015, and this is why we use R&A instead. But the two are highly correlated: $Corr(CC, PolCon) = 0.85$.

A.2 Fitch Ratings

Sovereign ratings are obtained from 'Fitch Ratings' (see website and Sovereign Rating Criteria). They are available since 1995. The Sovereign Issuer Default Ratings (IDs) are evaluations that anticipate the ability and willingness of a government to fulfill its debt obligations to both private-sector creditors and public debt securities, regardless of ownership, in a timely and complete manner. Fitch uses a combination of quantitative analysis and qualitative judgments to evaluate sovereign credit risk (e.g., their proprietary Sovereign Rating Model and the Qualitative Overlay). In addition, Fitch considers the impact of the sovereign's policies and actions on the country's overall economic performance, which can have significant effects on the sovereign's creditworthiness and can, in turn, be influenced by it. Using a series of factors, Fitch summarizes the rating at each point in time on the scale AAA to CCC+. We transform these ratings to a 0-4 scale. We discard all the countries for which we have less than 22 observations and then compute the average for each country. The third column of Online Appendix Tables 1 and 2 summarizes these values for developed and emerging economies.

We want to emphasize that structural factors, including 'political risk', are incorporated into Fitch's ratings. The company develops its measure of political risk incorporating Governance Quality (Government effectiveness, Rule of law, Control of corruption, Voice and accountability, and Business environment) and Political Stability & capacity (Political stability and capacity, Legitimacy of regime, Conflict/war risk, Debt payment record, Risk to economic policy). They also include past default decisions and other economic indicators when rating a country's default risk.

Construction of Figure 3: For each year, we we give a value of 4 to the countries with AA or above ratings, 3.5 to A, 3 to BBB and BB, 2 to B, 1.5 to CCC, 1 to CC, 0.5 to C, and 0 to DDD or below. We then average these out for all years in which we have observations and eliminate countries with less than 22 observations.

A.3 Fiscal Variables and Economic Aggregates

Variables are obtained from FRED, Eurostat, the World Bank, the IMF, OECD, Fitch Solutions, Statistical Offices and Financial Ministries.

A.3.1 Variable Definitions

- c: is "Private final consumption, USD (2010 prices)." Computed as the sum of all household spending on goods and services within the economy. It also includes spending by non-profit institutions serving households (NPISHs). Adjusted to 2010 prices. Built from NIPA.
- GDP Deflator: Gross Domestic Product: Implicit Price Deflator, Index 2012=100, Annual, Seasonally Adjusted. Downloaded from FRED, the variable name is GDPDEF.
- FB: is "Budget balance, % of GDP". It is constructed as fiscal revenue less fiscal expenditure as a percentage of GDP, $FB = \frac{Rev-Exp}{y}$, showing the net position or fiscal balance.
- g: corresponds to "Government final consumption, USD (2010 prices)". The sum of all government spending on goods and services within the economy. Also includes spending on government community services. Adjusted to 2010 prices. Built from NIPA.
- PB: is "Primary balance, % of GDP." Primary balance refers to the budget balance excluding interest payments on government debt as a percentage of GDP, $PB = \frac{Rev-Exp+rb}{y}$.
- tax: "Total revenue, % of GDP. " Fiscal revenue is government revenue or income through taxation, non-tax revenue (including asset sales), financial resourcing, and capital receipts. This variable is obtained from the OECD website.
- $Totb/y$: "Total government debt, % of GDP." Total government debt refers to the sum of both internal debt and foreign external debt issued by the government. Series obtained from the IMF.

- TB: is 'Net exports of goods services, % of GDP." It is the value of goods and services exported or sold abroad minus those that are imported or bought abroad. Net exports = exports - imports and $TB = \frac{X-I}{y}$. Variable expressed as a percentage of GDP. Built from NIPA.
- y: is "Real GDP, USD (2010 prices)."

Average values of private and public consumption to GDP, total debt to GDP, and tax revenues to GDP are displayed in Table 5

Table 5: Auxiliary Moments by Economic Group

Moment	Developed	Emerging
Means		
c/Y (%)	53.6	61.5
g/Y (%)	19.7	14.2
b/Y (%)	65.9	60.0
tax (%)	38.7	24.8
Cyclicality		
$\rho(g, y)$	0.37	0.56
$\rho(c, y)$	0.86	0.94
$\rho(PB/y, y)$	0.28	0.04

A.3.2 Variable Manipulations

Series are available from 1990 to 2022, although this is an unbalanced panel. We discard all the countries for which we have less than 22 observations in all variables except taxes (the series tends to have few observations for this variable). The list of countries in our sample satisfying the criteria can be found in Tables 1 and 2 of the Online Appendix.

We log-linearize: real GDP y, real private consumption c, and real public consumption g. That is, for each country, we regress $\ln y$ on time t using $\ln y_{it} = \alpha + \beta t + \epsilon_{it}$. Using the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$, we compute the residual $\hat{\epsilon}_{it} = \ln y_{it} - \hat{\alpha} - \hat{\beta}t$ and this is our de-tended measure for GDP. The same method is used to de-trend private and public consumption. We use the resulting series to compute: (i) the standard deviations: $\sigma(g)$, $\sigma(c)$, $\sigma(y)$ and (ii) the cyclicalities $\rho(g, Y)$, $\rho(c, y)$. The fiscal balance FB, the primary balance PB, the debt-to-output ratio b/y, and the trade balance TB are neither logged nor de-trended, as they are expressed as ratios (of output). When calculating their cyclicalities, only the cyclical component of real GDP is used for consistency. The cyclicality $\rho(g, y)$ corresponds to the estimated coefficient $\hat{\beta}$ of the OLS regression: $g_{it} = \alpha + \beta y_{it} + \epsilon_{it}$ (where g and y are log-linearized). Cyclicalities for all other variables are computed analogously.

B Theoretical Details

The policy proposal $\Phi(\Omega)$ depends on whether the government has access to capital markets $\Omega = 1$ or it is excluded from them, $\Omega = 0$. More specifically, $\Phi(1) = \{\tau, g, b', d, f_1, f_2, ..., f_n\}$ and $\Phi(0) = \{\tau, g, f_1, f_2, ..., f_n\}$, with $b' = 0$ and $d = 1$ when $\Omega = 0$. As explained in the main text, the proposer delivers the same amount of political favors f to each member of the mwc and the remaining f_p to him or herself. Policy Φ yields welfare $V^P(\Pi, \Phi)$ to the proposer, $V^I(\Pi, \Phi)$ to members of the mwc and, $V^O(\Pi, \Phi)$ to those outside of it.

B.1 The proposer problem

In round k, the proposer chooses Φ^k to maximize

$$\begin{aligned} &\max_{\Phi^k} V^{Pk}(\Pi, \Phi^k) \equiv U\Big(c(\mathbf{s}, \Phi^k), l(\mathbf{s}, \Phi^k), g\Big) + \frac{f_p}{\phi} + \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi', \Phi'|\Phi^k) \\ \text{s.t.}\quad & V^{Ik}(\Pi, \Phi^k) \geq J^{k+1}(\Pi, \Phi^{k+1}) \\ & f_p = B(\Pi, \Phi^k) - (m-1)f \geq 0 \\ & \tau, g, f \geq 0. \end{aligned} \tag{B.1}$$

The first constraint is an incentive compatibility constraint (IC) stating that the value of accepting the proposal for members of the mwc, $V^{Ik}(\Pi, \Phi^k)$, must be at least as large as the value of rejecting it, $J^{k+1}(\Pi, \Phi^{k+1})$. The second condition ensures the feasibility of government policies, requiring that the budget balance excluding favors is sufficient to cover the cost of providing political favors to the coalition members. Favors must be non-negative. The expected future value in the following period is $J(\Pi', \Phi'|\Phi^k)$, conditional on the current proposal being accepted.

We define the objects in the IC constraint next. Since members of the mwc obtain f in political favors,

$$V^{Ik}(\Pi, \Phi^k) = U\Big(c(\mathbf{s}, \Phi^k), l(\mathbf{s}, \Phi^k), g\Big) + \frac{f}{\phi} + \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi', \Phi'|\Phi^k)$$

when a proposal is accepted. If the proposal is rejected, we move to round $k+1$ where a new proposer is elected with probability $\frac{1}{n}$. If the member of the current minimum winning coalition mwc_k becomes a proposer in $k+1$, he/she will receive $V^{Pk+1}(\Pi, \Phi^{k+1})$. Note that the states do not change, as all proposal rounds happen within a period. With probability $\frac{m-1}{n}$, a member of mwc_k belongs to mwc_{k+1}, in which case they get $V^{Ik+1}(\Pi, \Phi^{k+1})$. With probability $\frac{n-m}{n}$ a member of mwc_k is not in mwc_{k+1}, which delivers $V^{Ok+1}(\Pi, \Phi^{k+1})$. Recall that anyone outside of the mwc obtains

$$V^{Ok}(\Pi, \Phi^k) = U\Big(c(\mathbf{s}, \Phi^k), l(\mathbf{s}, \Phi^k), g\Big) + \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi', \Phi'|\Phi^k),$$

as in that case $f_i = 0$. The continuation value of rejecting a proposal to any member of the current mwc, $J^{k+1}(\Pi, \Phi^{k+1})$, is computed by taking expectations over these three possibilities.

$$J^{k+1}(\Pi, \Phi^{k+1}) = \frac{1}{n} V^{P,k+1}(\Pi, \Phi^{k+1}) + \frac{m-1}{n} V^{I,k+1}(\Pi, \Phi^{k+1}) + \frac{m-n}{n} V^{O,k+1}(\Pi, \Phi^{k+1}). \tag{B.2}$$

Re-arranging the expression above,

$$J^{k+1}(\Pi, \Phi^{k+1}) = U\Big(c(\mathbf{s}, \Phi^{k+1}), l(\mathbf{s}, \Phi^{k+1}), g^{k+1}\Big) + \frac{B(\Pi, \Phi^{k+1})}{\phi n} + \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi' | \Phi^{k+1}). \tag{B.3}$$

where Φ^{k+1} is the optimal policy of the next round's proposer (given any continuation value).

B.2 Simplified problem

The proposer will choose Φ^k such that the IC constraint holds with equality, $V^{Ik}(\Pi, \Phi^k) = J^{k+1}(\Pi, \Phi^{k+1})$. This implies political favors f satisfy

$$\frac{f}{\phi} = J^{k+1}(\Pi, \Phi^{k+1}) - U\Big(c(\mathbf{s}, \Phi^k), l(\mathbf{s}, \Phi^k), g\Big) - \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi', \Phi' | \Phi^k).$$

Replacing this in the definition of f_p (see B.1), we obtain

$$f_p = B(\Pi, \Phi^k) - (m-1)\phi\left[J^{k+1}(\Pi, \Phi^{k+1}) - U\Big(c(\Pi, \Phi^k), l(\Pi, \Phi^k), g\Big) - \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi', \Phi' | \Phi^k)\right]$$

substituting f_p in $V^{Pk}(\Pi, \Phi^k)$, we can write the objective function as

$$\max_{\Phi^k} U\Big(c(\mathbf{s}, \Phi^k), l(\mathbf{s}, \Phi^k), g\Big) + \frac{B(\Pi, \Phi^k)}{\phi m} + \beta\, \mathbb{E}_{\mathbf{s}'}\, J(\Pi', \Phi' | \Phi^k).$$

Note that we dropped $J^{k+1}(\Pi, \Phi^{k+1})$ from the expression. This is possible because it is unaffected by Φ^k, so it can be considered a constant for the optimization problem above.

On the equilibrium path, proposal rounds $2, ..., T$ does not occur. This reflects the assumption that any coalition member accepts a policy mix that ensures that he/she is weakly better off than waiting for the next round. We can thus write the proposer's problem omitting k. Moreover, since $c(.)$, $l(.)$, and $B(.)$ are independent of f, the only policy variables to be determined in the problem above are—with a slight abuse of notation—, $\Phi(1) = \{\tau, g, b', d\}$ when the government has access to credit markets and $\Phi(0) = \{\tau, g\}$ otherwise. The proposer problem becomes

$$\begin{aligned} &\max_{\Phi} U\Big(c(\mathbf{s}, \Phi), l(\mathbf{s}, \Phi), g\Big) + \frac{B(\Pi, \Phi)}{\phi m} + \beta \mathbb{E}_{\mathbf{s}'} J(\Pi') \\ \text{s.t.} \quad & B(\Pi, \Phi) \geq 0, \end{aligned} \tag{B.4}$$

with

$$J(\Pi') = U\Big(c(\mathbf{s}', \Phi'(\Pi')), l(\mathbf{s}', \Phi'(\Pi')), g'\Big) + \frac{B(\Pi', \Phi'(\Pi'))}{\phi n} + \beta\, \mathbb{E}_{\mathbf{s}''}\, J(\Pi''). \tag{B.5}$$

where $\Phi'(\Pi')$ denotes next period's political *Markov Perfect equilibrium* outcomes. Up to this point, we have not mentioned anything about the continuation values. The derivation of $J(\Pi')$ is similar to that of $J^{k+1}(\Pi, \Phi^{k+1})$. In the next period, the proposer retains proposal power with probability $\frac{1}{n}$, obtaining welfare $V^P(\Pi', \Phi'(\Pi'))$. With probability $\frac{m-1}{n}$, it will be a member of next period's mwc, receiving $V^I(\Pi', \Phi'(\Pi'))$. With probability $\frac{n-m}{n}$ the current proposer is not in tomorrow's mwc, which delivers $V^O(\Pi', \Phi'(\Pi'))$. The proposer calculates $J(\Pi')$ as the expected value of $V^P(\Pi', \Phi'(\Pi'))$, $V^I(\Pi', \Phi'(\Pi'))$, and $V^O(\Pi', \Phi'(\Pi'))$.

C Extensions

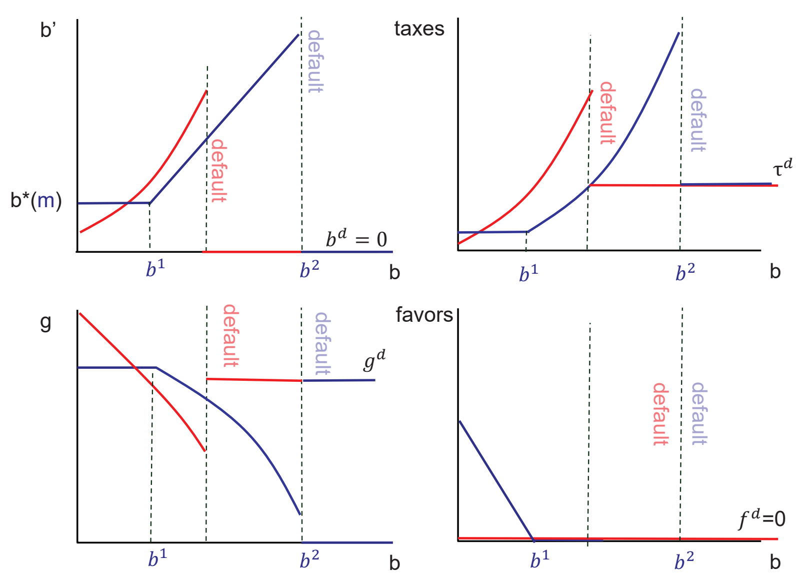
C.1 Low β case

The sovereign default literature typically assumes that the government is significantly more impatient than foreign lenders, imposing $\beta(1+r) < 1$. Without this assumption, the government would not borrow in the long run and default would never happen in equilibrium. Since r is a fixed risk-free rate, this restrictive assumption often requires the value of β to be unusually low. In our model, institutional weakness and corruption make the government *endogenously* impatient. As a result, we do not need restrictions on β, which can take a standard value in macro (including $\beta(1+r) = 1$). In this section, we compare our baseline political-economy model against the standard sovereign default model with an impatient benevolent planner, which we will refer to as the 'low β' case, evaluating their similarities and differences.

The competitive equilibrium given policy is the same in both environments because the domestic private sector and international capital markets are identical. The government budget constraint is also the same, as we consider identical policy instruments in both cases. The only difference lies in the maximization problem of the policymaker. In our benchmark model, the proposer's problem is given by eq. (6). The maximization problem of the planner in the 'low β' case is, instead,

$$\begin{aligned} J(\Pi) = \quad & \max_{\Phi} U\Big(c(\mathbf{s}, \Phi), l(\mathbf{s}, \Phi), g\Big) + \frac{B(\Pi, \Phi)}{\phi n} + \beta_L \mathbb{E}_{\mathbf{s}'} \, J(\Pi') \\ \text{s.t.} \quad & B(\Pi, \Phi) \geq 0, \end{aligned} \tag{C.1}$$

Figure 16: Policy functions, benchmark (blue) vs 'low β' (red)



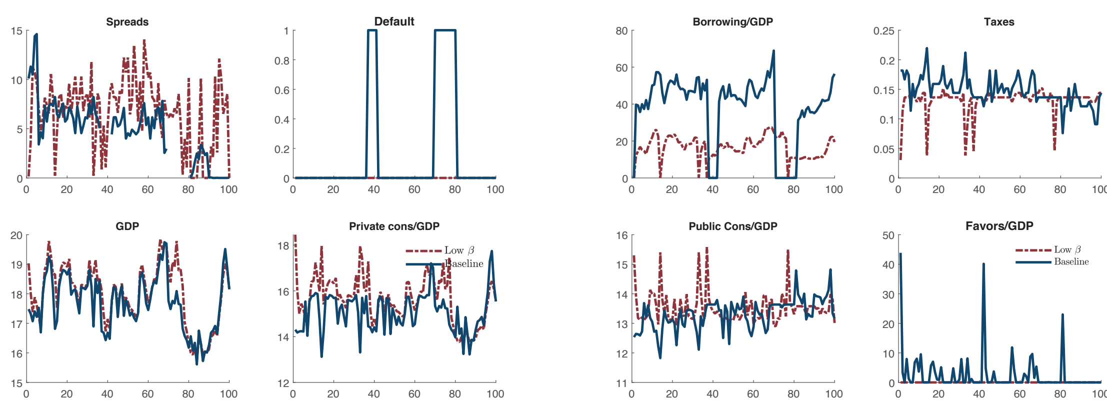
There are two main differences: First, if the planner chooses to give out political favors $B(\Pi, \Phi) > 0$, these need to be distributed among all groups (so there is an 'n' instead of an 'm' in the denominator). Second, the discount factor of the policymaker is β_L, lower than that of the lenders and domestic economic agents,

β. Even though we have included the corruption parameter, this takes the value of $\phi = 1$ in the benchmark calibration in both cases.

Because in the 'low β' case $m = n$ (by construction), $f_i = 0$ for all states of the world: a benevolent planner does not find it optimal to distort the economy to engage in corrupt public practices, regardless of how impatient it may be. As a result, $B = 0$ and we are in Case 2 of Section 3.4. Additionally, when the country is in default, $b = b' = 0$, implying that the choices of τ and g are independent of β_L. The first-order conditions are also independent of m, implying that, in default, taxes and public good provision are the same regardless of whether we are in the benchmark economy or the 'low β' case. Figure 16 illustrates this point. In that figure, the benchmark case is depicted in blue whereas the 'low β' case is shown in red.

Things are different when the government has access to international capital markets. While it is still the case that there is no revenue devoted to political favors $B = 0$, the government is more impatient. This will yield a steeper debt accumulation profile and default at a value of debt smaller than b_2. In equilibrium, this will result in lower levels of sustainable debt as there is no additional incentive to defer default and enjoy favors; that is the reason behind the inability to match spreads and debt/GDP ratio in those models. In our benchmark economy, for low levels of debt, there is a jump in b' but then a flatter borrowing profile for intermediate levels of debt. This allows the government to borrow more before a default is triggered. As a result, our model delivers higher debt/GDP ratios for similar spreads, making it easier to match the data.

Figure 17: Endogenous variables compared to a low-β economy



To understand this further, we simulated both economies for a finite number of periods. For the 'low-β' model (red dashed line) we set the value of β to 0.82 to match the average value of spreads in the benchmark. The initial debt for both economies is set to 0. We plot the time series of both economies for the same realizations of the TFP shocks in Figure 17. The one for the baseline model is the blue solid line.

There is a stark difference between the two economies. The baseline model borrowing immediately jumps to 60% of debt-GDP in period 1. The other one also rises, but at a much slower rate, and never reaches the level achieved by the benchmark economy. Since spreads are the same on average, higher debt in the benchmark economy requires higher taxes to repay the debt, as shown in Figure **??**. Higher taxes also imply lower GDP and lower private and public consumption. They are also lower in the benchmark as a share of GDP because of positive favors paid out. In the 'low-β' economy favors are zero by construction. Although default probabilities are the same in the two models by construction, our simulation generates two default events for the baseline economy. These are defaults due to political shocks that are absent in a 'low-β' economy.

Long-run moments associated with the low β case are displayed in the second column of Table 6 (the first column shows the results for the benchmark economy). The 'low-β' model sustains a lower level of debt than the baseline economy, even though the average spreads are the same. This implies that at higher levels of debt, the bond prices for the 'low-β' economy are lower, indicating a steeper bond price schedule. The subjective discount factor, therefore, has a sharper impact on spreads compared to m. This explains why in models of short-term debt and no possibility of political favors, it is difficult to match debt and spreads simultaneously to the data. This difference in the slope of the bond price schedule is because, in models with favors, curvature on the utility function makes favors easier to cut in a recession than increase taxes. If there are no favors, taxes need to be raised or there has to be a default to smooth consumption.

Table 6: Long-run moments: Model Extensions

Moment	Benchmark	Low β	Autocracy
Institutions			
R&A (mean)	4.77	20	1
Instability (St Dev)	0.78	0	0
Favors/GDP (mean)	4.58%	0	35.58%
Impatience β	0.99	0.74	0.99
Fiscal Policy			
Spreads	7.11%	7.00%	0.004%
$\mu(tax)$	17.69%	13.67%	43.2%
$\mu(g/Y)$	13.1%	13.5%	7.2%
$\sigma(tax)$	0.044	0.023	0.018
$\sigma(g)/\sigma(y)$	1.48	1.67	0
$\mu(b/Y)$	50.1%	21.7%	65.2%
$\rho(TB, y)$	-0.37	-0.65	-0.35
$\rho(FB, y)$	-0.31	-0.58	-0.34
Macro Outcomes			
Sp. Crisis	9.85%	40.1%	3.59%
$\sigma(y)$	3.34%	4.04%	2.70%
$\sigma(c)/\sigma(y)$	1.29	1.48	0.46

C.2 Autocracies: the case of China

In our model, the number of groups with veto power m (e.g. in the mwc) affects not only how tempting it is to engage in corrupt public practices but also the probability of being in the mwc in the future. As m declines, the proposer becomes naturally more impatient, as this reduces the likelihood of getting f_i next period. While we interpret the case of $m = 1$ as an autocracy, there are examples in which a group remains in power for long periods (e.g. China). In such a case, the proposer does not need to engage other groups to implement policies and also knows that it will likely be in power for a long period. To study such an environment, which is not covered in the main text, we simulated our economy assuming $m = 1$, but also that the one group in power will be in power forever. That is, there is full persistence. We understand this is unrealistic, but it could give us insights into how intermediate but highly persistent cases would work. This case also explains how the political frictions in this paper are different from turnovers.

The third column on Table 6 displays the results. The first thing to note is that a significant part of the budget is devoted to political favors, which reach a whooping 35% of GDP (versus 5% in the benchmark case). Because the autocrat is endogenously more patient (e.g. there is no risk of losing power), default risk is practically zero. This allows the government to better smooth taxes and public good provision, and sustain a higher level of debt to GDP. The volatility of output and consumption is also smaller. The autocrat can implement tax-smoothing policies but at the cost of frequently engaging in corrupt public practices.

CAREER INCENTIVES AND EMPLOYEE PRODUCTIVITY IN STATE-OWNED ENTERPRISES: EVIDENCE FROM INDIA

Tanisha Agarwal* **Samarth Gupta†** **Kaushalendra Kishore‡**

December, 2023

State-owned enterprises (SOEs) are large and important organizations in many economies but suffer from low labour productivity IMF (2020). Can SOEs improve their labour by enhancing career concerns for their employees? We show that exogenous change in opportunities to influence career progression significantly improves the performance of employees of state-owned banks in India. In particular, we find that when banking employees get more exposure to senior management, who can influence their promotion decisions, they increase credit expansion on both intensive and extensive margins. Further, this expansion happens through increased productivity, and not costly factors such as liberal screening, lower interest rates or higher resource allocation. Our results show that reforms in performance review processes, which allow workers to signal effort to supervisors in state-owned firms, may yield substantial productivity gains.

JEL Codes: J33, M51, M54, D24 **Keywords:** Career concerns, State-owned enterprises, Firm productivity

*Graduate Student, The University of Texas at Austin

†Assistant Professor, Amrut Mody School of Management, Ahmedabad University. Email- samarth.gupta@ahduni.edu.in

‡Research Director, CAFRAL. Email- Kaushalendra.kishore@cafral.org.in. The views expressed are those of the authors and do not necessarily reflect those of CAFRAL

Non-Technical Summary

In economies worldwide, State-Owned Enterprises (SOEs) play a pivotal role, contributing significantly to various sectors. However, these entities often grapple with inefficiencies, especially concerning low labor productivity. Unlike their private counterparts, SOEs frequently lack robust incentive structures, relying primarily on promotions or career motivations to stimulate employee performance. This paper studies the efficacy of career concerns in public sector banks in India uses a natural experiment to establish a causal relationship between career concerns and employee productivity.

The findings reveal that when bank employees get more exposure to senior management, who can influence their promotion decisions, they increase credit expansion on both intensive and extensive margins. Further, this expansion happens through increased productivity, and not costly factors such as liberal screening, lower interest rates or higher resource allocation. Our results show that reforms in performance review processes, which allow workers to signal effort to supervisors in state-owned firms, may yield substantial productivity gains.

Why do State-Owned Enterprises (SOEs) face challenges in labor productivity compared to private firms? The study recognizes that while private firms use monetary incentives to enhance worker effort, SOEs primarily rely on promotions and career motives. However, the evaluation of employees for promotions suffers from leniency bias, with a high percentage receiving excellent ratings. Additionally, the absence of layoffs for poor performance creates a floor on the downside risk for employees. In light of these factors, the study aims to explore whether career concerns can effectively improve labor productivity within SOEs.

To study the incentive structures within SOEs the research uses data from Indian banks and the Lead Bank Scheme. Initiated in 1969, the Lead Bank Scheme designates one state-owned bank as the lead bank for each district to promote credit access to priority sectors. To monitor the lead banks of a state, the Reserve Bank of India appoints another state-owned bank at the state level. These banks are called the convener banks, and their task is to review the performance of the lead bankers in every district in that state. Thus, there are some districts where the same bank acts as both the lead bank and the convener bank. We refer to these lead banks as ‘aligned’; otherwise they are referred to as ‘non-aligned’ .

Why is the alignment of lead and convener banks crucial, and how does it impact the study’s hypothesis? In quarterly review meetings headed by the CEO of the convener bank, aligned districts enable the CEO to observe the output of her own employees, fostering interactions with senior management crucial for promotions and career advancements. The study hypothesizes that aligned lead bankers, with an opportunity to showcase effort and productivity to their CEO, would have a higher incentive for career enhancement.

Does communication between lead bankers and the CEO contribute to increased effort among employees? The research finds that alignment substantially increases credit inclusion in rural areas, with a 35% surge at the intensive margin and a 30% rise at the extensive margin. The heightened credit expansion is primarily concentrated in the agricultural sector, aligning with the priority sector lending's sectoral focus. The analysis extends to various drivers of credit markets, revealing that productivity metrics, including accounts per loan officers, witness improvement after districts become aligned. These results reinforce the hypothesis that enhanced communication between the lead banker and the CEO triggers increased effort, compelling loan officers to extend more credit.

The study's implications for the Indian banking industry suggest that enhancing management practices to provide employees with more opportunities to demonstrate effort, particularly in promotion-related activities, could boost productivity.

In conclusion, the research highlights the effectiveness of career concerns as a motivator for SOE employees and demonstrates the positive impact of organizational alignment on labor productivity. The findings provide valuable insights for policymakers seeking to enhance the performance of state-owned entities, with implications for improved social welfare and economic development.

1 Introduction

State-owned enterprises (SOEs) are often large and important organizations in many economies but suffer from various inefficiencies such as low labour productivity (IMF, 2020). While private-sector firms use monetary incentives or promotions to improve worker's effort, SOEs rarely offer bonuses based on subjective performance evaluations. This leaves SOEs only promotions or career motives to incentivize their employees.

Can career concerns be a strong motive to improve SOE employees? SOEs usually have a single port of entry and minimal lateral hiring, which make career concerns for existing employees a strong motivator. However, contract termination is rare in SOEs, which puts a floor on the downside risk for the employees. Further, even non-performing employees may achieve career growth, although with some delay in promotion. These factors would reduce the power of incentives generated by career concerns within SOEs. Thus, *prima facie*, the effect of career concerns on SOE employees is ambiguous. If career concerns are indeed effective in SOEs, then SOEs can improve their labour productivity by improving their review processes, which would further have enormous welfare implications.

In this paper, we study the role of career motives in a novel setting of Indian public sector banks. Studying within firm incentive structures in SOEs is a challenging task for several reasons. First, finding exogenous variation in incentives within firms is hard; hence most studies rely on conducting experiments within firms, which may be difficult to implement for the researchers. Secondly, SOEs have fewer incentives and are also more reluctant to share within-firm data compared to private firms, which are constantly looking to improve their productivity because of market competition. We overcome these challenges by using within-firm granular data from Indian banks and exploiting a natural experiment that induces exogenous variation in employees' ability to demonstrate higher effort and productivity to supervisors.

We analyze the Lead Bank Scheme of India to answer our questions. This scheme was started in 1969 to improve credit access to priority sectors such as agriculture, small businesses and marginalized sections of society.[1] Under this scheme, each district is assigned one state-owned bank (also called public sector bank) to promote Lead Bank Scheme outreach to these priority sectors in that district. This bank is called the lead bank for that district, and its activities are managed by the mid-management level employee, whom we call the lead banker (he). To monitor the lead banks of a state, the Reserve Bank of India appoints another state-owned bank at the state level. These banks are called the

[1]We discuss priority sector lending in greater detail in section 2.

convener banks, and their task is to review the performance of the lead bankers in every district in that state. Thus, there are some districts where the same bank acts as both the lead bank and the convener bank. We refer to these lead banks as 'aligned'; otherwise they are referred to as 'non-aligned'.

The review of Lead Banks occurs through quarterly meetings which are headed by the CEO (she) of the convener bank. In aligned districts, the CEO of the bank can observe the output of her own employees, while in non-aligned districts, the lead banker's output is reviewed by the CEO of a different bank. Institutional arrangements in Indian sector banks make such interactions with own CEO and senior management crucial for promotions and career advancements. Specifically, each employee in the bank receives confidential performance appraisal reports from his supervisors. At the time of promotion evaluations, a committee takes the decision on whether or not to promote the employee based on these reports and an interview with the employee. However, Khandelwal (2010) finds that nearly 80 - 90% of the appraisees get 'excellent' ratings. This leniency bias (Prendergast, 1999) renders these reports useless for separating performers from non-performers. Our discussions with former bank officials suggest that recommendations by senior officers and the discretion of supervisors in the form of informal oral recommendations become critical in evaluation by the committee. Further, since promotions at the middle management level are not automatic, and only a fraction of employees get promoted at a time, these recommendations are extremely valuable for early promotion. Apart from promotions, the employees in the banks also look to enhance their careers by getting appointed in coveted locations and departments for which they again rely on informal recommendations from senior managers in their bank.

We hypothesize that an aligned lead banker will have a higher incentive to demonstrate his effort and productivity to the CEO of his own bank. But a non-aligned lead banker has no such opportunity. Thus, the alignment of a district determines the employee's motive to enhance his career through better performance (Holmstrom, 1999). We exploit exogenous change in *alignment* to study how variation in career concerns affects the performance of the employees. Alignment of a lead bank changes mainly due to the splitting of old states in which case the newly formed states get new convener banks. Importantly, Lead Banks do not change in our sample. Thus, all variation in alignment change comes from administrative or political factors at the state, which are unlikely to be influenced by or correlated with the district-level performance of banks. Our granular dataset allows us to use a specification with a rich set of fixed effects such as district-time fixed effects to control for district-level time-varying characteristics such as local demand (this is analogous to using firm-time fixed effects in Khwaja and Mian (2008)), bank-time fixed effects to control for bank-level time-varying characteristics, and bank-district fixed effect to control for, among other factors, the appointment of a particular bank as lead bank in that district.

We expect the impact of alignment to be the highest in rural areas, and in particular, agriculture, since the focus of priority sector lending is toward the agricultural sector.[2] Further, since agriculture lending comprises only 7% of urban lending, we expect no impact of alignment on urban areas. Consistent with this reasoning, the results show an average increase of 35% in the amount of loans and an increase of 30% in the number of loans by aligned lead banks relative to other banks in rural areas. Sector-wise analysis also confirms that most of the increase is driven by agricultural credit, and not industry, personal and trade. There is no corresponding increase for the urban markets, as expected.

Our conjecture is that after alignment change, the lead banker would exert higher effort, which may occur by pushing his subordinate loans officers to reach out to more borrowers and disburse more credit. If loan officers are indeed working harder, then their average productivity should increase. We find that in rural areas, for aligned banks, credit amount per loan officer increases by 32% and number of accounts per loan officer increases by 28%. We then test whether other factors that could drive credit, such as lower lending rates, liberal screening of borrowers or number of loan officers also change for aligned lead banks. We find no change in Weighted Average Lending Rate (WALR), ratio of non-performing assets (NPAs) and total number of district-level loan officers. Thus, the increase in credit supply occurs mostly through an improved effort by banking officers of the district.

To further test the effort channel, we study the impact of competition by private banks on lending by aligned banks. Poaching new customers or expanding to new ones may be difficult if the competitor is a private bank, which are more productive relative to state-owned banks. Thus, marginal returns of exerting effort may be lower if the market share of the private banks is higher. We test this by including the interaction of alignment indicator with the market share of private banks in our bank-district level regressions. Increase in rural lending by the aligned banks and productivity of employees is lower in the districts with higher market share of private banks. This result provides further evidence that the increase in lending by the aligned banks is coming from the effort channel.

We then study if the increase in credit by the aligned lead banks translates to an increase

[2] 45% of all priority sector loans are earmarked toward the agricultural sector.

in total credit at the district level. We find an increase in total supply of credit in the aligned districts both at the intensive margin and extensive margin. This suggest that aligned lead banks are not just poaching old customers of existing banks but also giving credit to new customers. Sector-wise split suggests an increase in credit in all sectors. No significant change occurs for urban areas.

Finally, we conduct additional tests to check the robustness of our results. First, we rule out pre-trends in credit disbursal by aligned lead banks. As an alternative identification strategy, we compare treated lead banks against non-treated lead banks from other districts to find expected increase in credit disbursal by treated lead banks. We then look at the impact of alignment on non-lead banks and find no increase in credit as expected. As a placebo test, we study how deposits change because of alignment of lead banks, and find no evidence of increase in deposits. This is expected since deposit generation is beyond the scope of the Lead Bank scheme.

Our paper makes several contributions to the literature. First, we contribute to the empirical literature which studies the impact of career concerns on productivity of employees within organizations. The existing literature has analyzed the impact of career concerns for bureaucrats (Bertrand et al., 2020; Karachiwalla and Park, 2017) and employees in private firms Manthei et al. (2023). Bertrand et al. (2020) show that bureaucratic leaders may be incentivized by glittering prizes which come in the form of last career promotion making them the top bureaucrats in the country. Karachiwalla and Park (2017) show that higher wage increases for promotion are associated with better performance of teachers in government schools in China. Ours is the first study to show that career concerns can be effective incentives for SOE employees.

We also contribute to the literature on the productivity of SOEs. Most of the current literature has focused on the impact of privatization on the productivity of SOEs (Barberis et al., 1996; Estrin and Pelletier, 2018; Gupta, 2005; Hsieh and Song, 2015). However, privatization is an arduous reform and faces several political constraints. Instead, management reforms may serve as an alternative to improve SOE performance without a change in ownership, because the latter may also change the socially oriented objectives that the SOEs pursue rather than just profit maximization. Among these reforms, Kala (2023) shows the importance of giving managers of SOEs in India more autonomy in creating higher value. Our paper provides empirical evidence that better performance review processes, motives for career advancements and better levers at influencing own prospects may indeed be very effective in improving employee productivity. While our study is focused on banks, the issue of low employee productivity and poor incentives for employees plague most SOEs. Thus, our paper is also important in the specific context of India where SOEs form a large part of the economy and employ many workers.[3] State-owned banks alone employed about eight hundred thousand employees in 2019 as per the economic survey of India. Improving employee productivity in SOEs can thus have an enormous impact on social welfare.

The subject of credit inclusion for poor and marginalised sections in India has received considerable attention. Lead Bank Scheme itself originated during the social banking period with the aim of extending credit to under-represented sections. We add to the exiting literature on credit inclusion (Burgess and Pande, 2005; Burgess et al., 2005; Cole, 2009) by studying the Lead Bank Scheme and show that improving incentives of bank employees can improve credit delivery to priority sectors.

The rest of the paper is as follows. Section 2 discusses the Lead Bank scheme in India and describes the institutional setting. Section 3 describes the data. Section 4 discusses our identification strategy and section 6 discusses the results. In section 6, we conduct several robustness tests and section 7 concludes.

2 Institutional details and lead bank scheme

Indian banking sector comprises of a mix of private banks and state-owned or government-owned banks banks (also called public sector banks).[4] These banks usually have a national presence and are quite large in terms of their asset size. Both private banks and state-owned banks are listed on stock markets, but the state-owned banks are mostly owned by the government. On the other hand government usually do not have any direct ownership in the private banks.

2.1 Lead Bank Scheme

The Lead Bank Scheme was introduced in 1969 to address geographic disparity in credit availability in India. The Reserve Bank of India (RBI) was concerned that commercial banks did not have adequate presence in rural areas and lacked rural orientation (Gadgil, 1969). To address this concern, RBI adopted a service area approach where one state-

[3] As per Kala (2023), SOEs employed over 1.14 million people in the year 2009.

[4] There are also some other types of banks such as cooperative banks and regional rural banks, but these banks form a very small fraction total bank credit.

owned bank in each district was assigned the role to promote credit supply to priority sectors in that district. This commercial bank is known as the Lead Bank of the district. The lead banks assign dedicated personnel at the district level whom we refer to the Lead Bank manager or Lead Banker to oversee the activities of the lead bank scheme in their districts. These lead bank managers are middle-management level employees with about 15-20 years of experience.

In order to monitor the activities of lead banks, RBI appointed another state owned bank for each state 1977. This bank is known as the Convener Bank of the state. The convener bank of a state monitors the lead banks of all the districts in that state by conducting quarterly meetings with the lead bank managers. These meetings are headed by the CEO of the convener bank. In the absence of the CEO, these meetings are headed by another senior manager of the bank. The convener bank of a state and all the lead banks of districts in the state collectively form the State-Level Bankers' Committee (SLBC). In the year 2016, the final year of our sample, there were 26 banks which were assigned the responsibility of lead banks for all the districts in India and 18 banks which were assigned the responsibility of convener banks for the different states in India.

This institutional set up implies that their are some districts for which the lead and convener banks are same which we call as aligned districts; else the districts are referred to as non-aligned.

2.2 Activities of lead banks

The main role of the lead banks is to promote lending to priority sectors in their respective districts (RBI, 2021). As per RBI's regulations, 40% of each bank's Adjusted Net Bank Credit is reserved for priority sectors which include sectors such as agriculture, micro, small and medium enterprises, housing and education (Banerjee et al., 2004). Of the total priority sector loans, agricultural sector receives 45%, while 18% is reserved for micro enterprises (RBI, 2022). In addition, 30% of the priority sector loans should be given to weaker sections of the society.

For the execution of the lead bank scheme, each lead bank in consultation with other financial institutions in that district prepares an annual credit plan. The job of the lead bankers is to implementation of these credit plans. All banks operating in a district have to meet respective credit allocation targets. However, the lead banker only has his own sub-ordinates and resources at his disposal to achieve these targets. The loan officers are entry-level employees who follow the directions of the lead banker.

The lead banker also coordinates with other banks and local bureaucrats and politicians to promote socially beneficial activities such as digital payments and financial literacy. For example, the lead banks in the state of Odisha organized credit *melas* in 6 districts in October 2019 (Odisha, 2019), where bank employees conducted programs on financial literacy. Lead bankers also interact closely with high ranking government officials through various fora and inform government agents about the institutional and infrastructural bottlenecks faced by banks in credit supply and financial inclusion.

2.3 Appointment of lead banks and convener banks

RBI has adopted the following criteria in choosing a bank as a Lead Bank of a given district (RBI, 1972).

- Number of branches of the bank—The bank which has higher number of branches in the district receives priority in being appointed as Lead Bank of the district.
- Resources of the bank in the district—For resources, assets and liabilities are taken into account while selecting a Lead Bank for a district.
- Contiguity of districts with the same Lead Bank—RBI tries to take into account the proximity of districts while assigning leads banks as it may help the banks in their operation.

In a nutshell, typically the banks with the largest market share in the district is appointed the lead bank.

For selecting a Convener Bank, RBI considers the regional orientation of the bank. For example, when the state of Telangana was formed out of Andhra Pradesh in 2014, its convenership was allotted to State Bank of Hyderabad which had strong presence in Telangana, whereas Andhra Bank was retained as the Convener Bank of Andhra Pradesh. Map of districts and state tagged by their Lead and Convener Bank, respectively, can be found here.

Thus, Lead and Convener Bank appointment is not driven by district-level demand factors but mostly by supply-side capabilities of the bank.

2.4 Impact of alignment on career concerns of lead bankers

The officers in the state-owned banks are selected through competitive exams and these jobs are highly coveted as they pay well and have high career stability.[5] The banks usually have only one port of entry, observe low exit rate and exhibit rare or non-existent lateral entries.

The lead banker in a district is a mid-management level employee of the lead bank. These lead bankers have experience of 15-20 years and are engaged specifically for Lead Bank Scheme activities. Thus these bankers have climbed the organization ladder for more than 15 years and are very unlikely to exit the organization. Since bonuses are limited, the incentives for these employees come from promotions and appointments in coveted locations and departments within the bank.

The appraisal process in these banks is as follows. The supervisor of each employee writes an annual confidential performance appraisal report for the employee. In some cases, the supervisor may also write an appraisal report for a particular task assigned to the employee. When the employee is up for promotion, a committee takes the decision on whether or not to promote the employee based on these reports and an interview with the employee. The promotions at the middle and senior management level are not automatic and only a fraction of employee get promoted in the first shot. At the middle management level, only about one third of the employees may get promoted in their first attempt. The remaining employees have to wait for one or more years for their promotion.

A key challenge in the appraisal process is that the appraisal reports and ratings suffer from a high leniency bias (Prendergast, 1999). As per the Khandelwal Committee Report, 80 - 90% of the appraisees get an 'excellent' rating (Khandelwal, 2010) and this bias renders these reports ineffective for screening performers from non-performers. In such instances, informal oral recommendations and the discretion of senior officers, at the request of the employee, are utilized to decide on promotion decisions. Because only a fraction of employees get promoted at a time, these recommendations are extremely valuable for early promotion. Apart from promotions, the employees in the banks also look to get appointed in coveted locations and departments for which they again rely on informal recommendations from senior managers in their bank.

As discussed above, the leads bankers at the district level are monitored by the CEO

[5] Our sample is from 1999 to 2016. Hence the lead bankers in the period of our sample would have been as hired as entry-level officers in an era when the Indian economy was still quite socialist in nature and jobs in state-owned enterprises considered amongst the best jobs in India.

(or occasionally a very senior manager) of the convener bank of the state through SLBC meetings which are held quarterly. Thus the aligned lead bankers get to demonstrate performance and productivity to the CEO of his own bank whereas the non-aligned lead bankers interact with the CEO of a different bank. These review meetings are an opportunity for the aligned lead bankers to impress the CEO of his own bank with his performance. A good impression on the CEO and any recommendation by her can go a long way in enhancing the career of the lead bankers. Thus, being aligned gives the lead bankers an opportunity to enhance his career through his performance.

Based on the above arguments, our key hypothesis is that aligned lead bankers will perform better than non-aligned lead bankers. This hypothesis may not hold true because of several reasons. First, the power of incentives may not be high enough as the CEO will not fire the non-performers limiting the down side risk for such employees. Secondly, while the non-performers may not get promoted immediately, they generally will also get promoted albeit with some delay. Finally, the aligned lead bankers instead of becoming more productive, may engage in unproductive influence activities such as "buttering up" the CEO (De Janvry et al., 2023; Milgrom and Roberts, 1988; Milgrom, 1988) to garner his favours. This final countervailing force may in fact make the aligned lead bankers less productive.

3 Data and Summary Statistics

We use Basic Statistical Returns (BSR) Data of the Reserve Bank of India. This dataset provides branch-level credit and deposit statistics of scheduled commercial banks and regional rural banks in India from 1999-2016.[6] Since our treatment is at the bank-district-year level, we consider bank-district-year as the unit of observation for our analysis. We observe the number of accounts and the total amount of loans or credit outstanding in a given year by each branch across sectors (agricultural, industry, transport, professional services, trade, etc.) and population centers (rural, semi-urban, urban and metropolitan).[7] We can also observe metrics of asset quality such as weighted average lending rates (WALR) and ratio of non-performing assets (NPA). We will club semi-urban, urban and metropolitan branches together and call them urban branches. We conduct our analysis

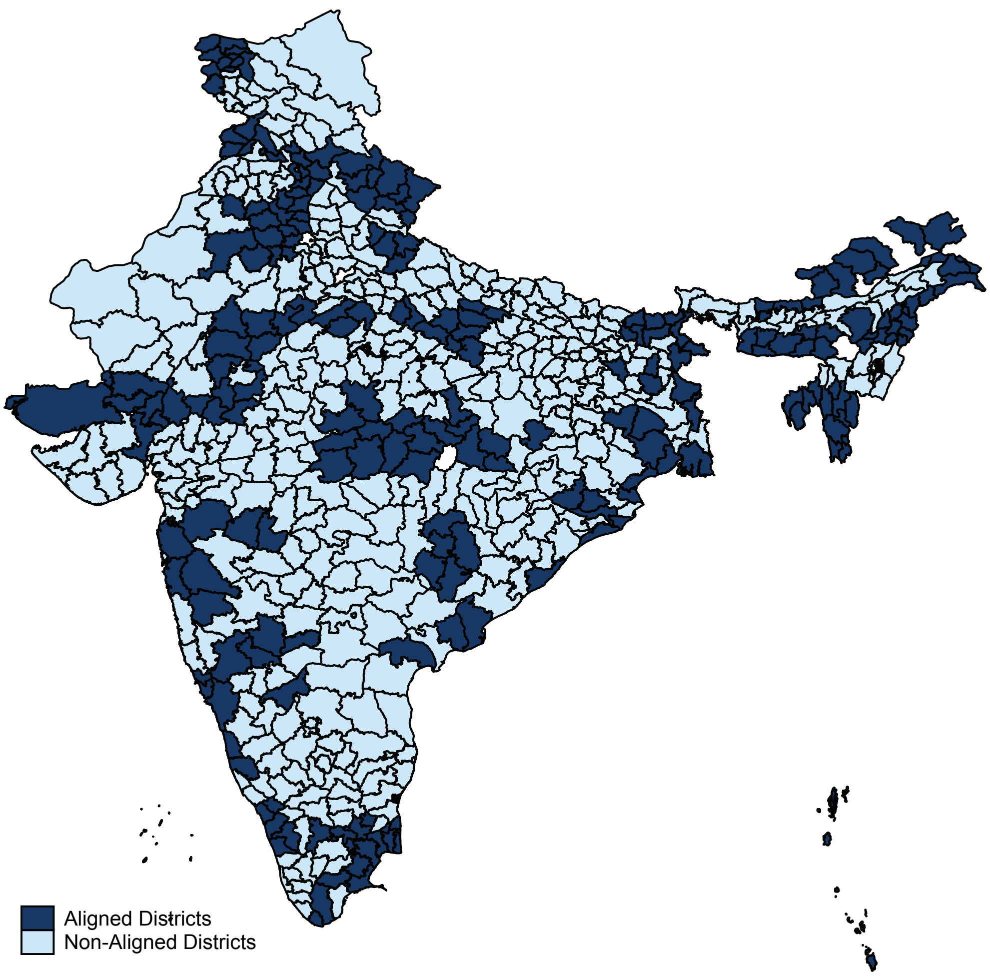
[6]In India the financial year starts on 1st April and ends on 31st March. So we have annual data from March 31, 1999, to March 31, 2016.

[7]According to RBI, rural areas are defined as centres with a population of less than 10,000. Similarly, semi-urban areas are those with population between 10,000 and 100,000, urban areas are those with population between 100,000 to 1 million and metropolitan areas are those with population of more than 1 million.

separately for rural and urban branches for reasons which will be discussed below.

We develop the panel data on alignment by collecting information on Leads and Conveners from websites of various SLBCs. In order to track changes in Leads and Conveners across years, we collect the notifications for Lead Bank Scheme available from RBI's website. Around 44% of districts are aligned. Figure 1 shows the map of districts in India tagged by their status of alignment in the year 2011.

Figure 1: Districts by Alignment



We first look at the distribution of lending by four major sectors- agriculture, industry, personal and trade - in rural and urban branches (Table 1). In rural areas, agricultural lending comprises of 53% (72%) of total amount (accounts) of credit. Contrarily, in urban areas, agricultural lending comprises only of 8% (29%) of total amount (accounts) credit. Industry loans form a larger share in urban areas (45%) than in rural areas (7%).

Table 2 shows the average loan size in urban and rural areas. As expected, average loan

Table 1: Sectoral Share- Mean (in %)

	(1) Agri	(2) Industry	(3) Personal	(4) Trade
Panel A: Credit				
Rural	53.28	7.05	16.84	15.72
Urban	7.94	44.58	16.34	9.42
Panel B: Accounts				
Rural	72.37	1.96	12.61	6.13
Urban	29.39	3.24	51.59	5.54

This table calculates each sector′s credit share as a % of total credit in Panel A for both Rural and Urban subsamples. Panel B reports the share in terms of number of accounts.

size is much larger in urban areas than rural areas. More importantly, even for agricultural sector, urban loans are more than twice the average size of agricultural rural loans.

Table 2: Average Loan Size

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Rural	70.4	57.6	129.2	89.7	123.4
Urban	416.4	128.7	4719.2	139.3	636.1

Average loan size for all loans is calculated by regions - Rural & Urban in column 1, and then it is further segregated by the various sectors- Agriculture, Industry, Personal and Trade.Note that this number is in INR thousands.

The above statistics suggest that lead bankers will focus mostly on rural areas than urban areas for the following reasons. The main objective of a lead bank is to give out more priority sector loans, which refer to lending to agriculture, micro, small and medium enterprises, housing, education, export, etc. However, agriculture receives 45% of priority sector loans, while 20% is reserved for small industrial units. In addition, within these sectors, attention is paid to financially excluded and weaker sections of society, which are primarily located in rural areas. Since agricultural credit comprises 53% loans in rural areas and only 8% loans in urban areas, it is natural that the lead banker will focus on rural areas. Hence we expect that as lead banks become aligned, they will give more credit in rural areas than in urban areas.

Table 3 provides the share of credit that the lead banks provide in rural and urban areas. We see that in rural (urban) area leads banks have a markets share of 35.6% (27.1%). This is consistent with the selection of Lead Banks reflecting supply-side orientation of the bank and indicates the capacity of Lead Banks to influence credit delivery in the district.

Table 3: Lead Banks Share- Mean (in %)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Credit					
Rural	35.6	35.3	35.5	35.5	35.1
Urban	27.1	27.9	26.8	26.7	25.3
Panel B: Accounts					
Rural	32.4	33.1	30.5	32.5	30.9
Urban	26.4	28.4	25.5	26.0	24.5

Lead Banks' average lending share in terms of Credit Amount (Panel A) and Credit Account (Panel B) is reported by regions and sectors.

Finally, table 4 and 5 provides the summary statistics of the important variables separately for aligned lead banks, non-aligned lead banks and non-lead banks. On overage the leads banks are much larger than non-lead banks as they lend more and also has higher number of branches. The WALR is quite similar for the lead and non-lead banks. But the non-lead banks have lower NPA ratio. This is because private banks are non-lead banks and they have lower NPA.

Table 4: Summary Statistics for Rural Area

	(1) Aligned Lead	(2) Non-Aligned Lead	(3) Non-Lead
Credit	1047.8	1172.5	413.8
(INR million)	(1792.1)	(1702.3)	(1596.0)
Accounts	13236.6	16725.1	5953.0
	(16764.8)	(17127.3)	(13033.4)
Loan Officers	27.83	31.78	12.34
	(25.40)	(25.78)	(19.68)
WALR	11.82	11.78	11.82
	(1.688)	(1.678)	(2.694)
NPA ratio	0.0763	0.0811	0.0709
	(0.115)	(0.114)	(0.132)
Branches	13.31	14.79	6.231
	(11.24)	(10.77)	(10.01)

The table reports summary statistics for different subsamples - 1. Lead Banks in Aligned Districts , 2. Lead Banks in Non-Aligned districts , and 3. Non-Lead Banks. Note that the variable Credit is reported in INR millions.

Table 5: Summary Statistics for Urban Area

	(1) Aligned Lead	(2) Non-Aligned Lead	(3) Non-Lead
Credit	2466.7	7444.0	2384.3
(INR million)	(5744.5)	(52795.8)	(21984.9)
Accounts	12550.9	15686.5	5669.0
	(19825.3)	(21564.1)	(103702.6)
Loan Officers	39.70	62.28	24.34
	(48.03)	(195.4)	(139.4)
WALR	12.14	12.30	12.40
	(1.838)	(1.690)	(2.503)
NPA ratio	0.0773	0.0792	0.0647
	(0.115)	(0.112)	(0.126)
Branches	9.247	12.25	4.198
	(12.10)	(17.89)	(8.415)

The table reports summary statistics for different subsamples - 1. Lead Banks in Aligned Districts , 2. Lead Banks in Non-Aligned districts , and 3. Non-Lead Banks. Note that the variable Credit is reported in INR millions.

4 Empirical Strategy

Our data is rich enough to allow several empirical specifications. We choose the following which, we believe, allows for the most rigour and interpretability to estimate the impact of alignment on a bank within a district.

$$y_{bdt} = \beta.1\{\text{AlignedLead}\}_{bdt} + \phi_{bt} + \phi_{dt} + \phi_{bd} + \epsilon_{bdt} \quad (1)$$

where, y_{bdt} is our dependent variable for bank b in district d in year t. $1\{AlignedLead\}_{bdt}$ is an indicator which takes value 1 if the lead bank b in district d is aligned in the year t; i.e. the Convener bank is the same as Lead in that district for that year.

Credit market related outcomes of a bank in a given district can be influenced by large number of factors, such as demand and supply for credit, temporal changes within a bank or variation in capacity across banks in the market. We include bank-year, ϕ_{bt}, district-year, ϕ_{dt}, and bank-district, ϕ_{bd}, fixed effects. Inclusion of bank-year dummies addresses time-varying changes for a bank such as organizational changes, lending strategies or bank capital. District-year effects can account for time-varying local demand and supply factors, which can influence credit markets. District-year effects also absorb endogeneity of a district being aligned. Assignment of a bank as the Lead of a district depends on the pre-existing supply-side capacity of the bank in that region. To account for time-invariant resource differentials across banks within the local market, we include bank-district fixed effects. Since Lead banks did not change in the time period of our observations, bank-district effects also account for endogeneity in Lead bank selection.

4.1 Identification

In equation 1, the coefficient on $1\{\text{AlignedLead}\}_{bdt}$, β, measures the impact of alignment on credit outcomes of a bank in a district. Identification of this impact requires alignment indicator and the unobserved error term, ϵ_{bdt}, to be uncorrelated; i.e. $E(1\{\text{AlignedLead}\}_{bdt}).\epsilon_{bdt} = 0$ However, the estimate of β would be biased if the change in alignment of a lead bank was influenced by unobservable bank-district-time factors. Thus, it is important to study the reasons for change in alignment.

In the time period of our study, change in alignment occurs due to the following reasons:

1. Formation of a New State—When a new state is formed, it may be assigned a differ-

ent Convener compared to the mother state. In such cases, the alignment status of districts in new state may change. For the period of this study, four new states were formed—Chattisgarh, Jharkhand, Uttarakhand and Telangana. Each of these states received a new Convener.

2. Change in Convenership of a State—RBI also changed convenership for Manipur in 2004 from Union Bank of India to State Bank of India.

In total there are 58 districts which change alignment. Out of these 58 districts, there were 44 districts which changed to aligned from non-aligned and 14 districts which changed to non-aligned from aligned.

For the estimate of β in equation 1 to be biased, residual bank-district-year components should influence state-level changes. It seems highly unlikely that a banking unit within a district leads to a change at the state-level. Further, most Lead Banks once appointed do not change and have remained the same since inception of the Lead Bank Scheme.[8] Inclusion of district-year and bank-districts effect account for these factors. Thus, what changes for the Lead Banks is only the opportunity to communicate with own CEO. The indicator for alignment of Lead Bank in a district is, thus, independent of residual component ϵ_{bdt}. Nevertheless, we also conduct robustness checks to lend weight to our estimation.

5 Results

5.1 Comparing Lead Banks against other banks within districts

As discussed above, priority sector lending is oriented toward rural markets. Thus, we present results separately for rural and urban areas.

Lead-Bank Performance in Rural Markets Panel A of table 6 shows the impact on credit disbursal in rural branches of Lead Banks under equation 1. Column (1) shows that the total credit disbursed by Lead Banks increases by 0.299 log points after alignment. We dis-aggregate the impact across four main sectors which constitute 91% of total credit market in rural branches.

[8]Some Lead Banks changed due to merger of banks. However, those cases were not responsible for alignment change since the Convener bank of that district was neither of the two merging banks; i.e. alignment remained 0 before and after merger.

Table 6: Bank-District Impact (Only Rural)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Log (1+Credit)					
$AlignedLead_{bdt}$	0.299***	0.290**	0.337*	0.136	-0.120
	(0.09)	(0.12)	(0.18)	(0.13)	(0.22)
Observations	83101	83101	83101	83101	83101
R-Squared	0.932	0.885	0.787	0.889	0.808
Panel B: Log(1+NOAC)					
$AlignedLead_{bdt}$	0.263***	0.235**	0.298***	0.175*	-0.060
	(0.08)	(0.10)	(0.11)	(0.10)	(0.12)
Observations	83101	83101	83101	83101	83101
R-Squared	0.949	0.926	0.870	0.910	0.890

Sample restricted to rural areas. Sample is trimmed at 1 and 99 percentile to remove the effect of outliers. Each specification controls for bank-district, bank-year and district-year fixed effects. Standard errors are clustered at the bank-district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

For the agricultural sector (column 2), the coefficient on Aligned Lead Bank is 0.290 with a standard error of 0.12—rural agricultural credit by a Lead Bank increases by 33% after it becomes aligned. Lending for industry also increases by 0.337 log points though the effect is only significant at 10% level. No effect appears for Personal and Trade loans.[9]

Panel B of table 6 shows the impact of alignment on the opening of new accounts in rural areas by Lead Banks. The total number of new accounts increases by 0.263 log points after a Lead Bank becomes aligned. This effect translates into nearly 30%. We also find a significantly positive increase in new accounts for agriculture, industry and personal sectors.

[9]Our choice of clustering standard errors at district level follows from Abadie et al. (2017). Results remain statistically similar if we cluster standard errors within bank-district strata.

Lead Bank Performance in Urban Markets Though PSL covers the needs of urban areas, PSL-oriented sectors such as agriculture occupy a smaller proportion in urban areas. For eg; agriculture credit comprises only 7% share in overall disbursal in urban branches. Thus, we expect a small or no impact on urban credit markets. Table 7 shows the results of Lead Bank performance in urban markets. No significant increase occurs for either total credit disbursal or new accounts of Lead Banks in urban areas after alignment.

Table 7: Bank-District Impact (Only Urban)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Log (1+Credit)					
$AlignedLead_{bdt}$	0.075	-0.005	-0.025	-0.133	0.052
	(0.08)	(0.16)	(0.17)	(0.09)	(0.13)
Observations	179906	179906	179906	179906	179906
R-Squared	0.921	0.834	0.865	0.904	0.853
Panel B: Log(1+NOAC)					
$AlignedLead_{bdt}$	0.037	0.051	0.142	-0.106	0.030
	(0.08)	(0.13)	(0.10)	(0.07)	(0.08)
Observations	179906	179906	179906	179906	179906
R-Squared	0.931	0.887	0.868	0.919	0.885

Sample restricted to urban areas. Sample is trimmed at 1 and 99 percentile to remove the effect of outliers. Each specification controls for bank-district, bank-year and district-year fixed effects. Standard errors are clustered at the district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

5.2 Mechanisms

Next we examine the channel that led to increase in credit in rural areas.

5.2.1 Employee Productivity

If our hypothesis is true, then lead banker would work harder and push the loans officers to reach out to more borrowers and disburse more credit. We should then see an increase in productivity of the loans officers in the aligned lead banks. We use two metrics of productivity—ratio of total lending and loan officers, and ratio of total credit accounts and loan officers. Column (1) shows the results for log of ratio of credit and loan officers. The coefficient in column (1) is 0.285 with a standard error of 0.08. Column (2) reports that the corresponding impact on log ratio of number of accounts and loan officers is 24.9%, which is significant at 1% level. These coefficients are very close to those in credit regressions (see table 6), suggesting that almost the entire increase in credit may be driven by increase in loan officer productivity.

5.2.2 Lending Rates

Lower lending rates may also increase credit uptake as opposed to higher effort on the part of loan officers. Column (3) reports the impact on Weighted Average Lending Rate (WALR). The coefficient is -0.051 which is very small and is statistically insignificant.

5.2.3 Effort substitution

While loan officers may increase effort in prospecting for new loans, they might simultaneously also reduce effort in screening or monitoring the loans to increase credit uptake. This will affect the asset quality and the level of non-performing assets. To detect this, we use ratio of Non-Performing Assets (NPAs) as a dependent variable in Column (4). NPAs remain nearly unchanged after the bank becomes aligned, suggesting that loan officers are not reducing their effort in screening or monitoring the loans.

5.2.4 Loan Officers

Finally, in column (5), we report the impact on total number of loan officers appointed in Lead Banks after alignment. The coefficient is 0.014 and the effect is insignificant. This result rules out allotment of more resources to banks after alignment change.

5.2.5 Competition from Private Banks

Another method to test our hypothesis of higher effort would be to check whether the impact of alignment decreases when facing more efficient competitors. This is because any residual demand in the market may already have been fulfilled by the more efficient banks. So the marginal benefit of effort would be lower in such markets resulting in lower effort in equilibrium, which would further reduce any credit increase because of alignment. Given that private sector banks in India are more productive (Sensarma, 2006), in districts with higher share of rural credit provided by private banks, the effect of alignment should be lower. We use the following specification for this hypothesis:

$$\begin{aligned} y_{bdt} \;=\; & \beta.1\{\text{Aligned Lead Bank}\}_{bdt} + \gamma.1\{\text{Aligned Lead Bank}\}_{bdt} \times \%\text{Pvt.Rural Credit} \\ & + \phi_{bt} + \phi_{dt} + \phi_{bd} + \epsilon_{bdt} \end{aligned} \tag{2}$$

Table 8 shows the results for log of credit, log of accounts, log of credit per loan officers and log of account per loan officers in columns (1), (2), (3) and (4) respectively. Across all specifications, the coefficient on Aligned Lead Bank indicator remains significant and positive. However, the interaction of Aligned Lead Bank indicator and share of private sector rural credit is significantly negative.

A one percent increase in share of private banks in rural credit attenuates the increase in total credit, total credit per loan officers, total accounts and total accounts per loan officers by 1.9%, 1.4%, 1.6% and 1.1%, respectively. Another implication of this result is the presence of slack or unmet demand in credit markets in India, which is lower when competition from private bank increases.

Table 8: Alignment Effect in Markets with Efficient Competitor

	Log(Cr)	Log(N)	Log(Cr/LO)	Log(N/LO)
$AlignedLead_{bdt}$	0.311***	0.271***	0.310***	0.271***
	(0.09)	(0.08)	(0.08)	(0.06)
$AlignedLead_{bdt}$ ×	-0.019***	-0.014***	-0.016***	-0.011***
% Rural Lending by Pvt.	(0.01)	(0.00)	(0.00)	(0.00)
Observations	83101	83101	82997	82997

Sample restricted to rural areas. Sample is trimmed at 1 and 99 percentile to remove outliers. Each specification controls for bank-district, bank-year and district-year fixed effects. Following Abadie et al. (2017), standard errors are clustered at the bank-district level. ***/**/* denote significance at the 1/5/10 percent level, respectively. Log(Cr), Log(N), Log(N/LO) represents Log(Credit), Log(1+NoACs) and Log(1+NoACs/LO) repsectively.

5.3 District-level Impact

Does alignment of Lead Banks also impact aggregate market outcomes at the district-level? Lead Banks constitute the biggest banking firms in their respective districts, with branch share of nearly 20-25%. Thus, an increase in Lead Bank productivity may have broader impact as well. To test for aggregate impact in the district, we use the following specification:

$$y_{dt} = \beta.1\{\text{AlignedDist}\}_{dt} + \phi_d + \phi_t + \phi_{st} + \epsilon_{dt} \tag{3}$$

$1\{\text{AlignedDist}\}_{dt}$ is the indicator which takes value 1 if district d is aligned at time t, and 0 otherwise. We include district, year and state-year fixed effects. The coefficient on $1\{\text{AlignedDist}\}_{dt}$ indicates the impact on y_{dt} when the district becomes aligned. We are unable to include district-year effects since it is perfectly correlated with $1\{\text{AlignedDist}\}_{dt}$. While this may not be a concern if the alignment change of a district is uncorrelated with time-varying demand for credit, we acknowledge the absence of controls for demand for banking services may produce biased results.[10]

[10] Such a control was possible in equation 1 through the inclusion of district-year effects.

District-level Rural Impact Panel A of table 9 provides the impact on district-level credit disbursal after the district becomes aligned. Total credit disbursal in the district improves by 0.320 log points with a statistically significant impact at 1% level across all sectors. Panel B shows similar results for accounts as well. The total number of accounts increased by 0.207 log points at the district-level with the corresponding impact for agriculture, industry, personal and trade sectors at 0.184, 0.307, 0.287 and 0.117 log points, respectively.

Table 9: District-Level Impact (Only Rural)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Log (1+Credit)					
$AlignedDist_{dt}$	0.320**	0.208**	0.487**	0.391	0.267**
	(0.14)	(0.10)	(0.19)	(0.24)	(0.12)
Observations	10590	10590	10590	10590	10590
R-Squared	0.970	0.969	0.860	0.950	0.897
Panel B: Log(1+NOAC)					
	Total	Agri	Industry	Personal	Trade
$AlignedDist_{dt}$	0.207***	0.184***	0.307***	0.287*	0.117**
	(0.07)	(0.06)	(0.10)	(0.17)	(0.06)
Observations	10590	10590	10590	10590	10590
R-Squared	0.970	0.967	0.869	0.919	0.933

Sample restricted to rural areas. Sample is trimmed at 1 and 99 percentile to remove outliers. Each specification controls district, year and state-year fixed effects. We include coefficients on district-year fixed effects from equation 1 as proxy for time-varying demand for financial services in the district. Following Abadie et al. (2017), standard errors are clustered at the district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

As noted above, these effects may be biased without controlling for time-varying demand at the district level, which may explain the high magnitude of the coefficients.

District-level Urban Impact We also test for the corresponding equilibrium impact in urban markets. Table 10 provides the results. Similar to what we observed for bank-district effects, we find no district-level impact on higher credit disbursal or number of new accounts as shown in Panels A and B, respectively.

Table 10: District-Level Impact (Only Urban)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Log (1+Credit)					
$AlignedDist_{dt}$	0.007	0.010	0.047	-0.124	-0.160
	(0.03)	(0.08)	(0.07)	(0.12)	(0.17)
Observations	10241	10241	10241	10241	10241
R-Squared	0.987	0.971	0.966	0.984	0.968
Panel B: Log(1+NOAC)					
	Total	Agri	Industry	Personal	Trade
$AlignedDist_{dt}$	-0.009	0.078	0.043	-0.123	-0.124
	(0.03)	(0.08)	(0.06)	(0.13)	(0.18)
Observations	10241	10241	10241	10241	10241
R-Squared	0.980	0.967	0.925	0.975	0.947

Sample restricted to urban areas. Sample is trimmed at 1 and 99 percentile to remove outliers. Each specification controls district, year and state-year fixed effects. We include coefficients on district-year fixed effects from equation 1 as proxy for time-varying demand for financial services in the district. Following Abadie et al. (2017), standard errors are clustered at the district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

6 Robustness Checks

Our identification assumption is that the unobserved bank-district-year effects do not influence the alignment of districts or its change. Since the change in alignment occurs due to state-level factors such as the formation of new states, which are plausibly exogenous to local-level factors, we believe that the estimation strategy has credibility. Thus, local-level factors are unlikely to influence alignment. To further validate our results, we provide some robustness checks in this section.

6.1 Pre-Trends in Credit Variables

Pre-existing trends in credit disbursal or account generation may confound our results. Such a trend may occur for various reasons. Lead Bank personnel could have anticipated the change in alignment, or there could be coinciding unobservable factors unrelated to the alignment change which introduce a pre-trend in credit markets. We test this concern by using the following specification:

$$\begin{aligned} y_{bdt} \; = \; & \beta_{-1}\text{Before}^{-1} * 1\{\text{Lead}\}_{bdt} + \beta_0 \text{Before}^0 * 1\{\text{Lead}\}_{bdt} + \\ & \beta_{+1}\text{After}^{+}1 * 1\{\text{Lead}\}_{bdt} + \phi_{bt} + \phi_{dt} + \phi_{bd} + \epsilon_{bdt}. \end{aligned} \tag{4}$$

We replace the $AlignedLead_{bdt}$ dummy in equation 1 with three dummy variables - $Before^0$, $Before^{-1}$ and $After^{+1}$, which, take value 1 for the year of alignment change, exactly one year before the year of alignment change and for all years after the year of alignment change respectively, and 0 otherwise. We interact these variables with the indicator for the lead bank. β_{-1}, β_0 and β_1, therefore, estimate the impact one year before, the year of and the years after alignment change, relative to the years before alignment change. If $\beta_{-1} > 0$, the our results are biased upwards.

Table 11 presents the results for equation 4 with log of credit amount disbursed as y_{bdt}. β_{-1} remains statistically insignificant across all columns, indicating that credit disbursal for any sector by Lead Banks immediately before alignment change was not different from the years prior to that. Importantly, we also find β_0 to be statistically indistinguishable from 0, which indicates that the impact of alignment does not occur instantly. There may be two possible reasons for this. First, some changes in alignment may have occurred in the middle of the year, and so there is not enough time to see the impact in the same year. Secondly, it may take time for the lead banker to push his loan officers to disburse more credit.

Table 11: Pre-Trend Analysis for Amount (Only Rural)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
$Before^{-1}$	0.024	0.133	-0.052	-0.081	0.086
	(0.08)	(0.12)	(0.18)	(0.10)	(0.17)
$Before^{0}$	0.115	0.103	0.093	-0.064	0.153
	(0.09)	(0.13)	(0.24)	(0.11)	(0.24)
$After^{+}1$	0.291***	0.268*	0.196	0.174	-0.347
	(0.10)	(0.14)	(0.21)	(0.14)	(0.27)
Observations	83394	83394	83394	83394	83394
R-Squared	0.939	0.890	0.786	0.892	0.810

Sample restricted to rural areas. Sample is trimmed at 1 and 99 percentile to remove outliers. Each specification controls for bank-district, bank-year and district-year fixed effects. Following Abadie et al. (2017), standard errors are clustered at the bank-district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

In table 12, we use log of accounts as the dependent variable for equation 4. We find β_{-1} insignificant for the number of total, agricultural and industry accounts. For personal and trade sectors, there exist some pre-trends but the effect is negative.

Table 12: Pre-Trend Analysis for Accounts (Only Rural)

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
$Before^{-1}$	-0.003	0.104	-0.021	-0.178**	-0.161*
	(0.07)	(0.09)	(0.12)	(0.08)	(0.09)
$Before^{0}$	0.107	0.115	0.130	-0.107	-0.077
	(0.07)	(0.10)	(0.14)	(0.09)	(0.14)
$After^{+1}$	0.273***	0.263**	0.237**	0.186*	-0.232*
	(0.08)	(0.11)	(0.11)	(0.11)	(0.13)
Observations	83394	83394	83394	83394	83394
R-Squared	0.953	0.929	0.871	0.911	0.890

Sample restricted to rural areas. Sample is trimmed at 1 and 99 percentile to remove outliers. Each specification controls for bank-district, bank-year and district-year fixed effects. Following Abadie et al. (2017), standard errors are clustered at the bank-district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

6.2 Comparison of Lead Banks across districts

Specification in equation 1 provides the impact on Lead Banks with respect to all other banks within districts which received the treatment of alignment change. As an alternative strategy, we compare treated Lead Banks against Lead Banks in other districts. The Lead Banks in other districts did not change alignment, i.e. they were either aligned or non-aligned throughout our sample. Thus, all the other Lead Banks serve as a control group for the treated Lead Banks. Panel A and B of table 13 provides the results on credit and amount for this quasi-experiment. We restrict our sample to rural areas and control for district, state-year and bank-year effects.[11]

Table 13: Comparison of Lead Banks across districts

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Log (1+Credit)					
$AlignedDist_{dt}$	0.124**	0.142*	-0.065	0.034	0.076
	(0.06)	(0.08)	(0.16)	(0.07)	(0.12)
Observations	10207	10207	10207	10207	10207
R-Squared	0.958	0.951	0.793	0.892	0.823
Panel B: Log(1+NOAC)					
$AlignedDist_{dt}$	0.113***	0.118**	0.096	0.102*	0.077
	(0.04)	(0.06)	(0.09)	(0.06)	(0.09)
Observations	10207	10207	10207	10207	10207
R-Squared	0.955	0.940	0.844	0.880	0.871

Sample restricted to Lead Banks in rural areas. Each specification controls for district, year, state-year and bank-year fixed effects. Standard errors are clustered at the bank-district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

Consistent with our hypothesis, we find that the Lead Banks increase total credit disbursal by 0.124 log points and agricultural credit by 0.142 log points. Trade sector loans have statistically strong results as well. For total number of accounts and agricultural accounts, we see an increase of 0.113 and 0.118 log points, respectively.

[11] This specification has similar concerns as in equation 4; i.e. we are unable to control district-year effects since they will be perfectly collinear with the treatment variable. This will not be a concern as long as unobserved temporal demand is uncorrelated with change in alignment status.

6.3 Comparison of Non-Lead Banks across districts

Non-lead banks do not attend quarterly meetings, and thus do not have the opportunity to communicate even if own CEO is attending. Thus, we expect these banks to not experience any effect due to alignment change. To test that, we restrict our sample to non-lead bank district observations, and compare non-lead banks of districts where alignment changes against non-lead banks of other districts. We regress credit and accounts on the alignment indicator for the district. Table 14 shows the results for amount and accounts for non-lead banks.

Table 14: Comparison of Non-Lead Banks across districts

	(1) Total	(2) Agri	(3) Industry	(4) Personal	(5) Trade
Panel A: Log (1+Credit)					
$AlignedDist_{dt}$	-0.016	-0.025	0.135	0.033	0.235
	(0.07)	(0.11)	(0.15)	(0.09)	(0.15)
Observations	73221	73221	73221	73221	73221
R-Squared	0.909	0.856	0.737	0.863	0.766
Panel B: Log(1+NOAC)					
	Total	Agri	Industry	Personal	Trade
$AlignedDist_{dt}$	-0.064	-0.018	-0.035	-0.028	0.034
	(0.06)	(0.08)	(0.09)	(0.08)	(0.08)
Observations	73221	73221	73221	73221	73221
R-Squared	0.929	0.903	0.833	0.886	0.860

Sample restricted to non-lead banks in rural areas. Each specification controls for bank-district, bank-year and state-year fixed effects. Standard errors are clustered at the district level. ***/**/* denote significance at the 1/5/10 percent level, respectively.

Each regression controls for district, state-year and bank-year fixed effects. As expected, we see null effects in every regression. These results also rule out any negative spillover on the competitors in the market.

6.4 Impact on Deposits

Credit delivery is the main goal under the Lead Bank Scheme. As a placebo check, we look at whether the deposits get affected too. Table 15 shows that the deposits remain unaffected due to the change in alignment.

Table 15: Deposits - Bank District (Only rural)

	(1) Log(Deposits Amount)	(2) Log(Deposit Accounts)
$AlignedLead_{bdt}$	-0.063	-0.008
	(0.06)	(0.05)
Observations	83019	83019
R-Squared	0.962	0.964

Standard errors in parentheses.Each specification controls for bank-year,district-year and bank-district fixed effects. $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

This is unsurprising since Lead Bank personnel are unlikely to exert effort in mobilizing fund intake.

7 Conclusion

SOEs across the world are plagued with low productivity. While several scholars argue for privatization as a panacea to this issue, political compulsions and costly legal process in ownership transfer make this solution difficult to implement. In contrast, the evidence on administrative and management reforms in SOEs is scant due to a lack of adequate research design or data.

In this paper, we show that public sector employees may significantly increase effort when given the opportunity to strongly signal their productivity to supervisors, who can influence promotion and transfer decisions. We analyze the Lead Bank Scheme of the RBI, which aims at expanding district-level credit oriented toward some priority sectors. Under this scheme, RBI assigns plan implementation responsibilities to employees of one bank in each district, known as Lead Bank. Lead Bank personnel of a state convey the problems and challenges faced in their tasks in quarterly meetings. RBI assigns the responsibility of overseeing and conducting these meetings to the CEO of another bank, known as the Convener of the state. In some districts, where Lead and Convener banks are same, therefore, Lead Bank personnel are reviewed by own CEO. We define the districts where Lead and Convener belong to the same bank, as aligned districts. Discretion and informal oral recommendations by senior management plays a crucial role in promotion and transfer decisions. Aligned lead bankers, therefore, have a higher incentive to signal effort and productivity.

To empirically test this theory, we exploit exogenous changes in the alignment status of a Lead Bank. We use the BSR data from RBI to observe credit disbursal from banks. At the bank-district level i.e. the level of intervention of the Lead Bank Scheme, credit inclusion in rural areas increase by 0.299 and 0.263 log points at the intensive and extensive margin, respectively. Further, consistent with the sectoral focus of the priority sector lending, most of the credit expansion occurs in the agricultural sector. We test for various drivers of credit markets. While productivity metrics of credit and accounts per loan officers improve after change in alignment, other drivers which may affect supply of credit, such as lending rates, asset quality, and the total number of loan officers, remain unchanged. These tests validate our hypothesis that when the lead banker is able to communicate with the CEO, he exerts more effort and pushes the loans officers to give out more credit.

Our results also provide interesting implications regarding the banking industry in India. Good management practices can improve the performance of firms in developing countries (Bloom et al., 2010). The banking industry in India has also received attention from this debate (Khandelwal, 2010). Broadly, our results suggest that if banks can provide more levers for employees to demonstrate effort in a manner that affects their promotions, productivity is likely to increase. However, the net benefit would require balancing these productivity gains against the cost of instituting such a review mechanism, which is beyond the scope of the paper. Understanding this trade-off holistically may provide valuable insights into organization design of banks in India.

References

Abadie, A., Athey, S., Imbens, G., Wooldridge, J., 2017. When should you adjust standard errors for clustering? URL: http://arxiv.org/abs/1710.02926, doi:10.3386/w24003.

Banerjee, A.V., Cole, S., Duflo, E., 2004. Banking reform in india. India Policy Forum 1.

Barberis, N., Boycko, M., Shleifer, A., Tsukanova, N., 1996. How does privatization work? evidence from the russian shops. Journal of Political Economy .

Bertrand, M., Burgess, R., Chawla, A., Xu, G., 2020. The glittering prizes: Career incentives and bureaucrat performance. The Review of Economic Studies .

Bloom, N., Sadun, R., Reenen, J.V., 2010. Recent advances in the empirics of organizational economics. Annual Review of Economics 2, 105–137. URL: http://www.annualreviews.org/doi/10.1146/annurev.economics.050708.143328, doi:10.1146/annurev.economics.050708.143328.

Burgess, B.R., Pande, R., 2005. Do rural banks matter ? evidence from the indian social banking experiment. The American Economic Review 95, 780–795.

Burgess, R., Wong, G., Pande, R., 2005. Banking for the poor : Evidence from india. Journal of the European Economic Association 3, 268–278.

Cole, S., 2009. Financial development, bank ownership, and growth: Or, does quantity imply quality? Review of Economics and Statistics 91, 33–51. doi:10.1162/rest.91.1.33.

De Janvry, A., He, G., Sadoulet, E., Wang, S., Zhang, Q., 2023. Subjective performance evaluation, influence activities, and bureaucratic work behavior: Evidence from china. American Economic Review .

Estrin, S., Pelletier, A., 2018. Privatization in developing countries: what are the lessons of recent experience? The World Bank Research Observer .

Gadgil, D., 1969. Organizational Framework for Implementation of Social Objectives. Reserve Bank of India.

Gupta, N., 2005. Partial privatization and firm performance. The Journal of Finance .

Holmstrom, B., 1999. Managerial incentive problems: A dynamic perspective. The Review of Economic Studies 66, 169–182. doi:10.1111/1467-937X.00083.

Hsieh, C.T., Song, Z.M., 2015. Grasp the large, let go of the small: The transformation of the state sector in China. Technical Report. National Bureau of Economic Research.

IMF, 2020. State owned enterprises: The other government.

Kala, N., 2023. The impacts of managerial autonomy on firm outcomes. Technical Report. National Bureau of Economic Research.

Karachiwalla, N., Park, A., 2017. Promotion incentives in the public sector: Evidence from chinese schools. Journal of Public Economics .

Khandelwal, A., 2010. Report of the committee on hr issues of public sector banks.

Khwaja, A.I., Mian, A., 2008. Tracing the impact of bank liquidity shocks: Evidence from an emerging market. American Economic Review .

Manthei, K., Sliwka, D., Vogelsang, T., 2023. Talking about performance or paying for it? a field experiment on performance reviews and incentives. Management Science .

Milgrom, P., Roberts, J., 1988. An economic approach to influence activities in organizations. American Journal of sociology .

Milgrom, P.R., 1988. Employment contracts, influence activities, and efficient organization design. Journal of political economy .

Odisha, S., 2019. Proceedings of the 156th state level bankers' committee for the quarter ended june 2019.

Prendergast, C., 1999. The provision of incentives in firms. Journal of economic literature .

RBI, 1972. Report of the banking commission.

RBI, 2021. Master circular - lead bank scheme.

RBI, 2022. Master directions – priority sector lending.

Sensarma, R., 2006. Are foreign banks always the best? comparison of state-owned, private and foreign banks in india. Economic Modelling 23, 717–735.

CREDIT CONSTRAINTS, BANK INCENTIVES, AND FIRM EXPORT: EVIDENCE FROM INDIA

Yogeshwar Bharat * **Raoul Minetti†**

December 22, 2023

The aim of this paper is to investigate the impact of firms' financial fragility and banks incentives on firms' decision to exit the export market. We draw information from the Prowess database on a large sample of Indian businesses between 2005 and 2020 and we obtain bank data from the Reserve Bank of India. Estimation results indicate that more indebted firms are associated with a high probability of exiting the export market. However, when we focus only on bank borrowing, we find that firms with high levels of bank debt (over total assets) are characterized by a lower probability of abandoning the export sector. By interacting our measures of financial fragility with a state-owned bank dummy, we also show that highly indebted firms borrowing from state-owned banks are associated with an even lower probability of exiting the export market. Finally, when we employ the change in the priority sector lending regulation to test the causality of our results and avoid endogeneity concerns, we provide evidence that firms borrowing from banks that were missing their priority sector targets are characterized by a significantly lower probability of abandoning foreign markets.The study did not find any significant effect of policy change on firms trying to enter the export market. Using an indirect definition of productivity showed that the policy change did not affect the productivity of the leveraged firms.

JEL Classification: F1, G20, G21, O53

Keywords: International Trade, Banking, Soft Budget Constraints, India

*CAFRAL, Research Department, Reserve Bank of India Main Building, Mumbai, yogeshwarbharat@gmail.com

†Department of Economics, Michigan State University, minetti@msu.edu

Non-technical summary

There has been a growing attention to studying the factors that affect firms' entry into and exit from the export sector. The dynamics related to the entry and exit decisions of firms are different and have variation across economies. The role that the banking industry plays in influencing export entry and exit decisions is the topic of this paper. We look into how the influence of credit market factors on a firm's likelihood of survival in the export sector, especially the incentives of lending banks? We contribute to the empirical literature on the trade-finance linkages by establishing a causal relationship between banks' lending incentives and the exit/survival decision of exporting firms. Interestingly, our study shows that a financially fragile firm is more likely to stay in the export market if they can borrow from any type of bank. Firms have two means of financing their activities: internally from retained earnings and externally by borrowing from banks and other financial institutions. The ease with which firms can access financial markets can have an important effect on their decisions. A firm may experience considerable difficulties in making decisions about exports on both the extensive and intensive margins if it is unable to secure sufficient funding.

Our study is interesting in the sense that we look at an emerging economy and explore how heterogeneity in bank ownership has an impact on firm export dynamics. For our analysis, we exploit a rich panel data of Indian firms. We find similar pattern for firms in the extensive and intensive margin. We start our analysis by checking the basic hypothesis that firms that are highly leveraged have a higher probability of leaving the export market. We find that controlling for various firm characteristics, financially fragile firms have a higher probability of leaving the export market compared to less fragile firms. We get the same result when we control for liquidity.

Our analysis shows some interesting results when we define firms financial fragility in form of their bank leverage, we find that when we control for various firm level controls, the firms that are more financially fragile has a lower probability of leaving the export market compared to a less fragile firm. The result was more pronounced for firms that were able to borrow from at least one public sector bank. We find similar result with respect to firms that were able to borrow from more than one banks. But no such results were observed for firms that were borrowing from only private sector banks. Such result clearly highlights the heterogeneity in banks behaviour and how relationship of firms with banks of different ownership has an impact on their decision to stay/exit the

export market. This process of providing credit to financially inefficient exporting firms and keeping them afloat in the export sector creates a soft budget constraint and crowding out efficient firms from exporting.

We also employ advance causal inference tools and a policy shock to establish a causal relation between firm financial fragility, bank relation and firms decision to stay/exit the export market. There can be some type of sorting between banks and firms in terms of their relationship, which can contaminate and bias our result. To treat the contamination, we are using a policy change in priority sector lending (like directed lending program) in 2015 that mandated schedule commercial banks to book export credit as a part of priority sector lending. My study hypothesizes that after the policy change, there is an exogenous shock to credit supply (increase in credit supply), and state-owned banks to fulfill their priority sector lending target provide credit to already leveraged firms in the export sector. State-owned banks see these highly leveraged firms as low-hanging fruits to meet their priority sector lending target.

To find the impact of priority sector lending policy change on the probability of financially fragile firms leaving the export market, we use an advance causal inference method to study the effect of the policy change on firms' decision to stay or leave the export sector. We found that that on average firms borrowing from public sector bank has a 1.3% lower probability of leaving the export sector and it decreases by a further 2.7% if the firm is fragile. Additionally, we find no significant effect of the policy change on firms entering export. Similarly, we didn't find any effect of policy change on firms borrowing from privately owned banks. The policy didn't show any long-term impact on the productivity of the firms. It establishes the claim that public sector banks are using already exporting leveraged firms as a low-hanging fruit to fulfill their lending targets.

Our paper highlights the issue of soft budget constraints among exporting firms in an emerging economy where banking sector is heterogeneous in nature and most of the firms rely on banks for credit. One such heterogeneity was captured in form of bank ownership in India, and we see how this led to crowding out of efficient firms from the export sector. Some policy suggestions that can be suggested in case of India is to re-look at the ownership of Indian banking sector and have a proactive auditing system within the banking system to stop supporting such financially inefficient firms.

1 Introduction

The factors that affect firms' entry into, and exit from, the export sector have received growing attention in recent years. The dynamics of entry and exit decisions are different. For instance, when entering the export sector, a firm has to overcome relevant start-up costs,[1] whereas in the case of exit, the firm has already overcome that barrier but perhaps has found that it could not sustain the running costs to cater to both the domestic and the export market.[2] The main question explored in this paper is the role played by the banking sector in affecting entry and exit decisions from export. Specifically, we ask: How is the likelihood of a firm's survival in the export sector influenced by credit market factors, especially the incentives of lending banks? We address this question by studying the experience of an emerging economy, India, where credit market frictions are allegedly very severe.

Generally, firms have two means of financing their activities: internally from retained earnings and externally by borrowing from banks and other financial institutions. The ease with which firms can access financial markets can have an important effect on their decisions. If a firm cannot gain access to adequate financing, it can be significantly hampered in its export decisions both on the extensive and intensive margins. Indeed, credit constraints have been shown to be an important determinant of international trade flows.[3].

To carry out our analysis, we exploit rich panel data at the firm level from India. Since the liberalization of its economy in the early 1990s, India has seen a large increase in export participation among its firms (imports were already high before the liberalization process). Further, with the advent of the new millennium and the dot-com boom, Indian exports have taken off, led by the service sector. We contribute to the empirical literature on the trade-finance linkages by establishing a causal relationship between banks' lending incentives and the exit/survival decision of exporting firms for a panel of around 8,000 Indian firms, which constitute about 40% of total exports and contribute about 45% to the value added in manufacturing.[4] We use information about the leverage ratio and liquidity ratio of the firms as an indicator of their financial fragility. We also use infor-

[1]Melitz, 2003
[2]Ferri et al., 2019
[3]Manova 2008, 2013; Minetti and Zhu 2011
[4]Kale, 2017

mation on bank borrowing by these firms, including the type of their lending banks. The Indian banking system is largely dominated by state-owned banks in its loan and deposit markets. A broad literature has demonstrated that state-owned banks have different incentives in their lending decisions and standards relative to privately owned banks.[5]

The main hypothesis of this paper is that, through state-owned bank, the government created a conducive setting to a problem of soft budget constraints for exporters. In India, the government makes conscious efforts to support various industries and micro, small and medium enterprises (MSMEs) by mandating state-owned banks to offer easy loans to them. The Reserve Bank of India mandates "domestic scheduled" banks to provide easy loans to these enterprises under the Priority Sector Lending program.[6] The goal of granting easy loans to such firms by state owned banks may induce a soft budget constraint issue, whereby a highly leveraged firm is able to remain in the export market even though it wouldn't in a non-distorted credit market (i.e., borrowing under a situation with no mandates).

The results are robust to various estimation approaches. In particular to assuage concerns about the endogeneity of export we exploit as a quasi-experiment a policy change occurred in April 2015. This policy change was related to lending by domestic scheduled banks under the priority sector lending program of the Reserve Bank of India. The idea of priority sector lending was to facilitate a holistic development of the economy through the support of the banking sector. Till 2015, export credit was considered as part of the priority sector only for foreign banks with 20 or less bank branches. But, since micro, small or medium enterprise and agriculture sector comprise a sizable part of India's export sector, it was thought that it would be appropriate to extend export credit as a part of priority sector lending also to domestic scheduled commercial banks to facilitate the flow of funds to exporting firms. Thus, in April 2015 the Reserve Bank of India mandated that all domestic scheduled commercial banks would also be allowed to register export credit as priority sector lending. This policy change allegedly created a distortion in credit access for exporting firms. In particular, it induced an environment conducive to soft budget constraints for these firms (banks were looking for easy lending options), leading to even inefficient firms being able to remain in the export market.

[5] Minetti et.al. , 2021

[6] Reserve Bank of India Notification, 2012

After controlling for liquidity and other relevant firm attributes and accounting for the endogeneity of firms borrowing from state-owned banks, we find that the probability of exiting the export market is 1.6% lower for firms borrowing from a bank. That probability shrinks by a further 2.7% for a highly leveraged firm that borrows from state-owned banks that are missing their priority sector lending target.

We conduct several robustness checks and extensions of the baseline results. First, we establish that firms trying to enter the export market do not benefit from the policy change. The results for entering firms are indeed insignificant, showing that the policy change does not affect entrants. This conclusion carries through when we consider non-exporting firms and non-exporting small and medium enterprises and agricultural firms (which we use as a placebo test). We also look at the effect of policy change on firms borrowing from private banks who are missing their target and the result shows no significant effect on exit rate of leveraged exporting firms. Second, we look at the effect of the policy change on the productivity of firms, by using an indirect definition of productivity. We do not find any long-term effect of the policy change on firms' productivity. This suggests that the policy change in priority sector lending helping leveraged firms remain in the export market by a creating soft budget constraint, rather than enhancing their long-term productivity. We further divided the data into manufacturing and service sector industries, and the result showed that the policy change had a significant effect on leveraged firms from both sectors. Hence the impact was uniform across sectors.

The remainder of this paper proceeds as follows. Section 2 provides a review of literature. Section 3 covers the institutional details of the Indian economy, while section 4 describes the data. Section 5 specifies the empirical design, and section 6 presents the results. Section 7 contains robustness check for the empirical model. Section 8 concludes and discusses avenues for further research.

2 Literature

This paper is related to two main strands of the current literature. First, it adds to the research on firm export dynamics (Section 2.1). Second, it contributes to the literature on finance and trade, and more specifically to those studies investigating the role played by financial constraints and banking development in shaping firms' export activities (Section

2.2).

2.1 Firm Export Dynamics

There is a growing body of literature on export dynamics. This literature looks at the extensive margin of exports and studies what factors determine a firm's decision to enter, exit, or continue in an export market. Eaton et al. (2007) showed a significant turnaround in export markets, finding that nearly half of exporters were new entrants. Further, these new exporters are generally very small compared to average exporters, and most of them exit the market in the following year. This paper did not look at why firms are failing to continue exporting and are leaving the export market or why they are continuing in it. At the same time, Fernandes et al. (2015) analyzed micro-level data across several countries to study exporter behavior and how it varied across countries of different sizes and at various stages of development. If high-productivity firms are highly constrained and fail to invest in case of a developing economy, exporters should be relatively small in such economies (Bento and Restuccia, 2014; Hsieh and Klenow, 2009; Hsieh and Olken, 2014). This paper looks at a similar issue where most exporters in India are small and medium enterprises, which is attributed to credit constraints to productive firms. Another critical aspect of export dynamics also depends on the institutional strength of the countries to which the firms are exporting. Araujo et al. (2015) covered this issue in their paper and showed that firms are more comfortable doing business in countries with strong institutions. Chaney (2014) developed a model for international network formation, where firms gather information about future partners from current ones. Besedes et al. (2014) explored how credit constraints in the origin country affect import growth at the product level in the European Union and the United States. These papers looked at institutions and levels of development of home and foreign countries for export dynamics. We plan to look at individual firms' characteristics and analyze the credit constraint side of a firm and what role this plays in the firm deciding to exit or continue in an export market.

2.2 Financial imperfections and firms' export activities

Several papers have examined the nexus between firms' financial conditions and internationalization. Manova (2013) and Chaney (2016) develop theoretical models demonstrating that when liquidity constraints plague financial markets, whether a firm is credit constrained or not may influence its decision to export and the volume of foreign sales. Greenaway et al. (2007), for a panel of UK manufacturing firms over the period 1993-2003, empirically explore the link between firms' financial health and their export market participation decisions and find that exporters exhibit better financial health than non-exporters. By using survey data on Italian manufacturing firms, Minetti and Zhu (2011) estimate the impact of credit rationing on firms' export and find that the probability of exporting is 39% lower for rationed firms and that rationing reduces foreign sales by more than 38%. Similar results are provided by Manova et al. (2015), who show that credit constraints restrict international trade and affect the pattern of multinational activity. More recently, some studies have focused on the role of bank finance and bank-firm relationships for firms' export activities. Paravisini et al. (2015) estimate the elasticity of exports to credit using bank-firm level data from Peru and show that credit shocks affect the intensive margin of export. Minetti et al. (2019) investigate the effect of financial constraints on firms' participation in domestic and international supply chains. Ferri et al. (2019), using a sample of European manufacturing firms, investigate the nexus between bank-firm relationships and firms' export activities and find that the contraction of firms' export was milder when banks had access to up-to-date soft information on firms' export prospects.

A second strand of related studies has examined the impact of financial development on international trade. Beck (2002), using a large panel of countries over the period 1966-1995, suggests that more developed financial systems promote export in industries with increasing returns to scale. Becker et al. (2013) confirm these results by providing evidence that a developed financial system facilitates exports, especially in industries where fixed costs are high. Recently, some papers have also looked at the role played by the banking system. For instance, Claessens and van Horen (2021) investigate whether banking development and the presence of foreign banks affect firms' export activities. By analyzing a large set of countries over the period 1995-2007, they find that foreign banks increase export activities and that the entry of a foreign bank boosts export to its country of origin. Similarly, Minetti et al. (2021), using data from a large panel of countries over the period 1997-2014, examine the impact of countries' banking structure and regulations on export dynamics. Their results suggest that bank-oriented financial systems boost the number of exporters, although banks in lower income countries tend to reduce the dynamism of the export sector by slowing down exporters' entry and exit. This finding is in line with the finance literature highlighting that banks' tendency to protect the position of incumbent exporters is particularly strong for some types of financial intermediaries, such as domestic and public-owned banks in developing countries (Allen and Gale, 2001;

Main, 2006). Especially when operating in lax regulation environment, they may be more sensitive to the pressure of domestic governments and related parties to protect the export activity of incumbent businesses.

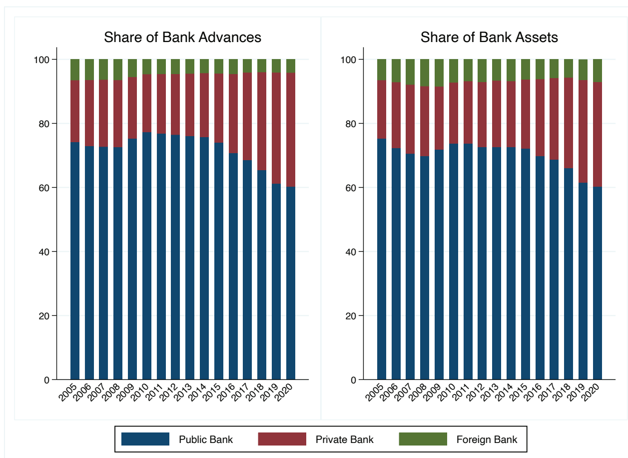
This paper tries to merge these branches of the literature in an emerging economy where the financial sector isn't developed and banks are a significant source of credit for firms. The study becomes more interesting for India, where the majority share of loans disbursed is by state-owned banks. The existing literature primarily examines regions that are historically known to have an efficient banking system. Hence, credit constraint is only an issue from the demand side and not from the supply side. Most of the export dynamics literature emphasized firms trying to enter the export market, and not much study related to exiting firms. We specifically look at firms in the export market and factors that can affect their exit. We are trying to combine firms' decisions to leave/stay in the export market with the type of banks they borrow. The analysis uses panel data from India. The reason for using data from an emerging economy was to look at the effect of credit constraints on a firm's decision to exit or continue in an export market. Here, the question becomes even more interesting as the issue of credit constraint is facilitated not only from the firm side but also from the bank side due to the type of banks present in the Indian banking system - discussed this part under the Indian banking system. We will do a causal analysis of the impact of borrowing from a state-owned bank on a highly leveraged firm to stay in the export market. We will further see whether such an effect is limited to firms already in the export market or those trying to enter. Another contribution is the novel data we created for the study; we have made data that contains firm and bank-related financial and non-financial information. Along with this, the data also have the mapping of firms borrowing from various banks. This mapping makes the data unique as it helps find the type of banks a firm is borrowing. The idea is to create a generic model that could be replicated to any emerging economy or even a set of emerging economies.

3 Institutional Background

3.1 Banking Sector in India

In recent years, the Indian banking sector has witnessed the emergence of many private banks and several large foreign banks. Since the liberalization of the banking sector and economy at large in 1991, the Reserve Bank of India has periodically distributed banking licenses to new banks and allowed foreign banks to operate retail banking in India. However, by and large, the banking sector is still dominated by the state-owned banks (these are corporate banks in which the government is the majority shareholder, also known as public sector banks). For example, around two thirds of total deposits are with the state-owned banks, which disburse approximately 70% of total loans and advances.[7] Also, state-owned banks hold the majority of assets; this trend not only seen for total assets but also loans and advances. This share is declining over time, albeit it still remains very high. Figure 1 shows the share of total assets as well as that of loans and advances over time.[8]

Figure 1: Share of Bank Assets



Source: Reserve Bank of India

Clearly, the share of assets and the share of bank advances have declined over time for state-owned banks. However, the combined share is around 70%, which shows a great dependence of firms and households on state-owned banks for access to credit. Hence, it is important to look at the effect of state-owned banks' lending to firms and households, respectively, compared to private banks (foreign banks can be neglected given their low penetration and share). State-owned banks were heavily regulated by the Reserve Bank of India, which has led to a constraint in borrowing for different firms. It was addressed to some degree with a change in the lending policy by the Reserve Bank of India after 1997 in line with the Nayak Committee recommendations. These recommendations provided more flexibility to the state-owned banks to disburse loans than before. Although even after these recommendations, the Reserve Bank of India still retained an instrumental

[7] Reserve Bank of India 2018
[8] Reserve Bank of India, 2018

role in determining bank lending policy (public sector) to individual borrowers. Banerjee, Cole, and Duflo (2004) suggested that state-owned banks in India were reluctant to engage in fresh lending decisions. Inertia plays a crucial role in explaining the behavior of the loan officers in state-owned banks. It is worth noting that the government owns public banks, so the loan officer is treated as a public servant. There is no incentive structure in place to reward loan officers bringing in more loans, while, on the other hand, there is a chance of them being charged with corruption or some form of impropriety.[9] Banerjee, Cole, and Duflo (2004) further validated this in their study.

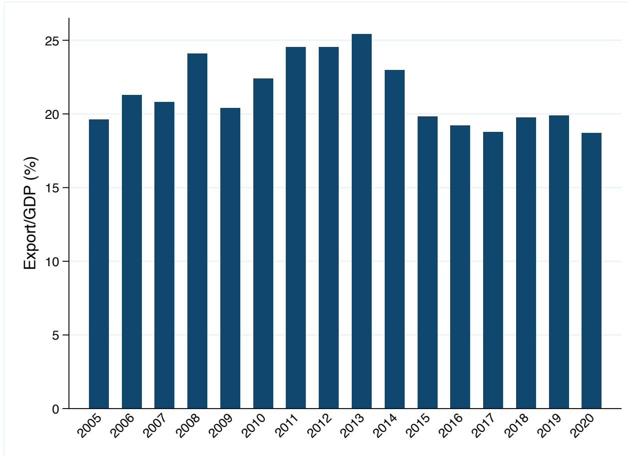
3.2 Indian Exporting Market

The Indian economy has gained considerable momentum over the last two decades, achieving and sustaining a high annual GDP growth rate. This high growth rate can, in part, be attributed to the growing contribution of the export sector to the economy. Historically, post-independence, India has focused on self-reliance, which has meant a minimal dependence on international trade as a source of income. Due to this, a large number of people still live in abject poverty. To address this, the central government has sought to improve the well-being of people by adopting a strategy of 'import-substituting' industrialization. The government developed a complex, extensive, and often costly system of price controls and quantitative restrictions to implement this.[10]

Since the start of the millennium, the volume of export in India has been increasing, and this can be seen through exports share of GDP, which has been growing over time. Figure 2 shows exports as a share of GDP for India since 2005. From the figure, the share of exports has been growing (albeit a drop in 2009 due to the financial recession of 2007), and on average, exports account for around a fifth of the total GDP of India. This makes it essential to study the export market and factors affecting the exporting decision of various firms in India. Another characteristic of the Indian exporting market is the share of micro, small and medium enterprises (MSME here onward) involved in exporting activities; this has also been seen for other economies, be it emerging (Eaton et al., 2008) or developed (Minetti and Zhu, 2011). In the case of India, the share of MSME exports to total exports has been increasing over time, particularly since the global financial crisis of 2007.

[9] Kapoor et al., 2012
[10] Mukherjee et al., 2012

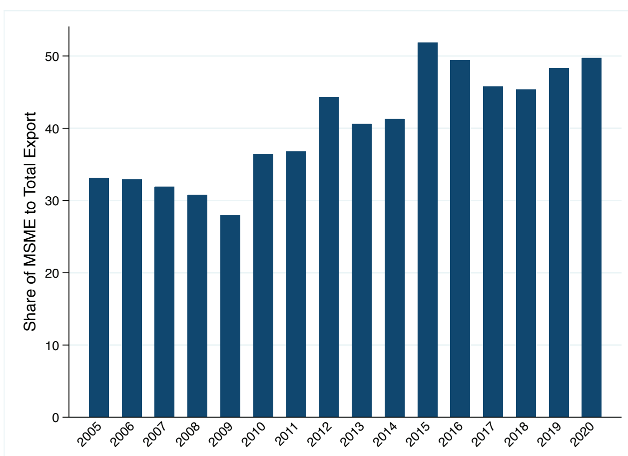
Figure 2: Share of Export to GDP



Source: Directorate General of Commercial Intelligence and Statistics (DGCIS)

Since 2005, the share was primarily flat, declined in 2008 and 2009, after that, it increased and reached about 50% of total export. Figure 3 validates this fact, showing that from 2012 onward, the increase has been much more significant. Hence, it has become essential to look at the banking sector more carefully given the inability of these firms to easily borrow from the bond or equity markets, leaving banks as their best bet for accessing credit.

Figure 3: Share of MSME Export to Total Export



Source: Directorate General of Commercial Intelligence and Statistics (DGCIS)

3.3 Priority Sector Lending

Priority sector lending is an important role given by the Reserve Bank of India to the banks for providing a specified portion of the bank lending to a few specific sectors, like agriculture and allied activities, micro and small enterprises, poor people for housing, students for education, and to other low-income groups and weaker sections. This is essentially meant to support the all-round development of the economy as opposed to focusing only on the financial sector. At a meeting of the National Credit Council held in July 1968, it was emphasized that commercial banks should increase their involvement in the financing of priority sectors, namely agriculture and small-scale industries etc.

Presently, Priority sector lending consists of the following categories:

1. Agriculture
2. Micro, small and medium enterprises (MSMEs)
3. Export credit
4. Education
5. Housing
6. Social Infrastructure
7. Renewable Energy
8. Others

Additional details related to priority sector lending are in Appendix A.

4 Data

The data for this paper are from the Prowess database from the Center for Monitoring the Indian Economy (CMIE), a private think-tank that provides firm-level data on all companies that are traded on India's major stock exchanges (Bombay Stock Exchange and National Stock Exchange) and several other public sector undertakings, it also has data of firms that are not publicly traded. The Prowess database comprises of rich panel data and is updated on a regular basis. The Prowess database has been used in several studies, including Bertrand et al. (2002), Khanna and Palepu (1999), Fisman and Khanna (2004), Topalova (2007), Kapoor et al. (2012), and Goldberg et al. (2010).

The Prowess database contains information primarily from the annual financial statements and balance sheets of listed and non-listed companies. One benefit of this database is that it contains information on manufacturing as well as on service sector firms. Since India's service sector contributes a major share to GDP, it is really important to see how the service sector industries fare in the export market. The coverage of the Prowess database is quite extensive, whereby all the firms put together account for 75% of corporate taxes and 95% of the excise duty collected by the Indian government. [11] For all these firms, Prowess contains detailed information (compiled from audited annual accounts, stock exchanges, company announcements, etc.) on 1500 items, including quantitative information on firms production, sales, export earnings, profitability, liabilities, assets, capital, cash flow, expenditure on capital goods, raw materials, power and fuel, labor, ownership, age, etc. It also contains detailed data on financial variables, like the amount of borrowing, bank borrowing, other financial institutional borrowing, and secured and unsecured debt. The database also categorizes firms by industry according to the 5-digit NIC classification. The list of firms spans the entire industrial composition of the Indian economy.[12] In this paper, we use firm-level data from 2005 to 2020 to cover a large span of time period, including the great financial recession of 2007 in the the data. Along with using the firm-level data, this database also provides extensive information on the banking side too, like, age, government ownership, tier 1 capital, non-performing assets, deposits, borrowing, liabilities, cash flow, investments, etc. Along with the information related to the firm and bank-specific variables, we were able to map the list of banks that are lending to firms. Since the ownership of the banks is known to us, we can easily see the effect on firms of borrowing from different types of banks.

Firm-level information is presented in Table 1 through Table 3. Table 1 provides the total number of firms along with the breakdown by sector. In total, there are $8,128$ unique firms in the sample that were active over a period of 16 years and the sample is approximately equally distributed between manufacturing and the service sector. This is helpful for the study as India is essentially a service-led economy, so we can see the effect of the service sector on the exporting part.

[11] Kapoor et.al.(2012)
[12] Prowess CMIE website

Table 1
Total Number of firms and their distribution sector wise

	Number	Percent	Share of Sales
Manufacturing	4,254	51.1	83.6
Services	4,070	48.9	16.4
Total	8,324	100	100

Table 2
Firm ownership status

	Number	Percent	Share of Sales
Private	8,030	96.5	69.4
Public	294	3.5	30.6
Total	8,324	100	100

Similarly, Table 2 shows the breakdown of firms as per their ownership status, where it can be seen that most of them ($\approx 96\%$) are privately owned.

Figure 4: Number of Firm, Entry and Exit rates respectively

Figure 4 illustrates the trend of the total number of exporting firms in the sample, along with their exit and entry rates. The trend clearly shows that the number of exporters remained the same throughout the time period. For the exit and entry rate too, we can see that the rate did not vary a lot in the first half, but some movement around and after 2013 can be seen, when they either peaked or troughed, another abrupt jump can be seen in 2017 due demonetization. Similarly, Table 3 provides descriptive statistics for some of the key variables that we have used in our study as firm-level controls, like total sales, total income, export earnings, profit after tax, total liabilities, total debt, cash flow, age, etc.

The paper looks at the effect of banks and their lending to various firms and how bor-

Table 3
Summary Statistics of firms in the database (taking initial, middle and final time period)

Variable[24]	Year	Mean	SD	Min	Max
Total Income	2005	2,013	21,048	0	1,153,380
	2012	5,857	56,818	0	3,351,871
	2020	13,525	101,370	0	4,615,428
Total Sales	2005	1,968	20,883	0	1,152,064
	2012	5,664	55,614	0	3,312,205
	2020	11,878	98,767	0	4,638,168
Export Earning	2005	394	2,564	0	100,621
	2012	1,634	16,487	0	882,226
	2020	2,168	29,357	0	1,478,650
Total Debt	2005	4,549	56,812	-16,689	3,680,789
	2012	13,789	153,903	-20,409	10,242,089
	2020	40,687	463,221	-58,856	20,109,070
Total Liability	2005	4,865	58,391	0	3,625,682
	2012	14,732	156,984	-43	12,783,548
	2020	42,694	482,469	-84	26,740,300
Total Capital	2005	302	2,458	0	77,280
	2012	467	2,945	0	112,785
	2020	1,075	8,404	0	521,562
Total Cash flow	2005	30	1,693	-31,639	90,327
	2012	223	5,838	-78,446	371,455
	2020	451	10,852	-97,819	634,258
Age	-	31	18	0	162

[24] All variables with monetary value are measured in Million Indian National Rupee

rowing from different types of banks (state-owned vs. privately-owned) affects firms' decision to stay in or to exit the export sector. Figure 5 clearly shows that if a firm is borrowing from at least one state-owned bank, then their exit rate is lower than those firms that are not borrowing from state-owned banks. The figure shows that the exit trend for both types of firms follows the same trend, and the exit rate peaks around 2011 for both types of firms (consistent with the overall trend). Since 2013, there has been a constant decline in the exit rate, but the decline is relatively steeper for firms borrowing from at least one public bank than those borrowing from no state-owned banks.

Figure 5: Exit rates for firms borrowing from at least public bank

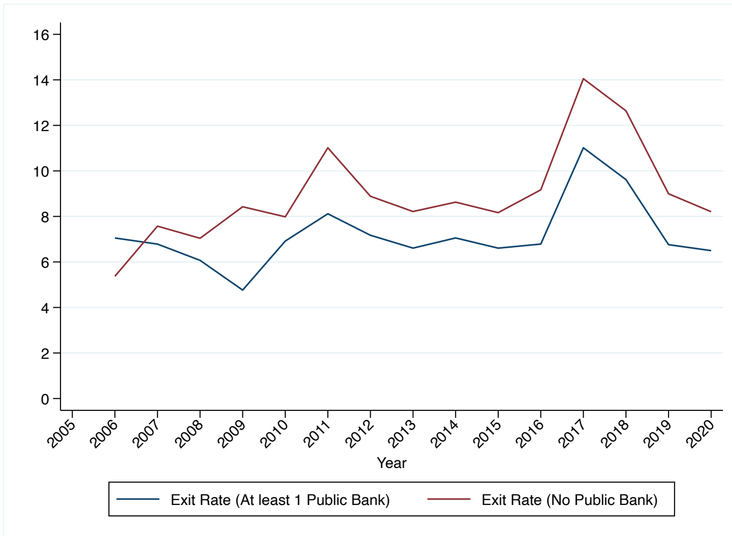
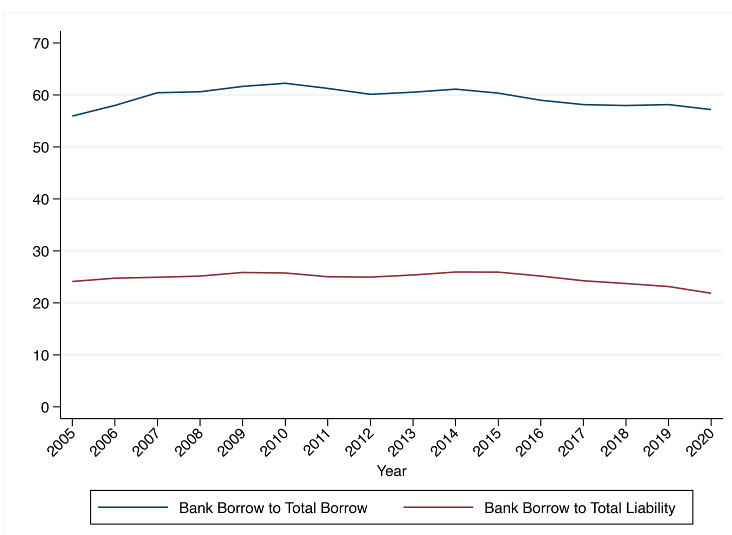


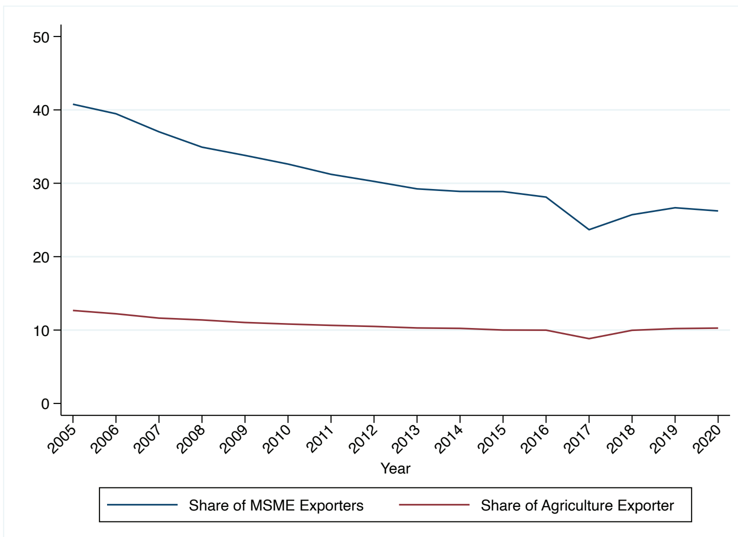
Figure 6 shows the share of bank borrowing to total borrowing and the total liabilities of firms. This shows that, on average, bank borrowing was around 60% of total borrowing, which makes its very important to look at the effect of bank borrowing on these firms as a channel for accessing credit. For India, being an emerging market economy where the majority of exporting firms are micro, small, or medium enterprises (as seen from Figure 3, also seen from the data in figure 7), access to credit through banks is an important channel, with Figure 6 showing its importance.

Figure 6: Bank borrowing to Total Borrowing for firms



It is also important to note that the financial market is not well developed in India and hence access to credit through banks is an important channel, and here, state-owned banks play the most important role (Figures 1 and 5, respectively). Figure 7 shows the share of micro, small, or medium firms and agricultural firms in the data and, on average, they comprise more than half of the total firms available in the data-set. Although, the share of micro, small, or medium firms have oscillated over time, the share of agricultural firms has remained constant over time.

Figure 7: Share of MSME and Agriculture exporter



An important aspect of the study is to look at the banking side and the supply of credit to various firms. The bank data are presented in the Prowess database, but more detailed data are looped in from the Reserve Bank of India's asset and liability database of various scheduled commercial banks from the fiscal year 2019-20. This database is used to get information regarding extensive asset and liability information for various 'scheduled commercial' banks, which can finally be used to calculate priority sector lending for various banks and to check which banks are missing their priority sector target and which do not. This will help us to check how those banks who miss their priority sector target react to the policy change be it public or private sector banks (foreign banks are not of great importance). We have merged this data with prowess database to create an extensive database for this study.

Along with the variables that were available in the data-set, we have also created some of other variables to use in the analysis. The dependent variable used in the study is a dummy variable defined as conditional on exporting in period $t-1$, whether the firm is continuing to export or is exiting the export market (1 for exiting and 0 for surviving). The exit probability is used to understand the effect of credit constraint on a firm in exiting the exporting market. Further, we use the information of the type of banks a firm is borrowing and how it helps to create a soft budget constraints for firm who are borrowing from state-owned banks compared to not borrowing from them. We also define leverage ratio (debt to capital ratio) an indicator for financial fragility in a broader and narrower sense (bank borrowing to total asset) and liquidity ratio (cash flow to asset ratio) in the general form, as used in most of the literature.

5 Empirical Strategy

It is reasonable to expect that credit constraints, size of firm and ownership would affect both the extensive and intensive margins of trade. Additionally, it will also affect the firms decision to enter, exit or continue in the market as well. Also, the size and ownership of the firm are important factor whether a firm will be able to borrow from market and banks. In this study we are using borrowing from banks (primarily state-owned) as an indicator for credit constraint. The firms who are not able to borrow from these banks are more credit constraint and will need to borrow from market (Banerjee and Duflo (2014)) which will affect their exporting and production decision. In the preliminary analysis, we try to capture the correlation/pattern between firms' leverage, their decision to stay/exit the export market and how borrowing from banks are affecting those decisions. These preliminary analysis are not a causal inference, later we will use a policy change to establish causality.

Preliminary specification for analyzing the effects of credit constraints on the extensive margin of export for firm f in industry i at time t is:

$$E_{fit} = \beta leverage_{ft-1} + \gamma X_{ft-1} + \varphi_i + \psi_t + \epsilon_{fit} \tag{1}$$

Where E_{fit} is the probability of a firm to exit the export market in t conditional on exporting in $t-1$. Whereas $leverage_{ft-1}$ is the standard leverage ratio that acts as a proxy for financial constraint. X_{ft-1} is a vector of firm-specific characteristics like size, profitability, age, ownership, etc. φ_f and ψ_t are firm and time dummy.

In the above specification, size is defined as the log of sales, and age is just the age of the firm. The expected sign for these variables should be negative as large and old firms tend to exit less if in the market. The variable of interest is 'leverage' as it captures financial fragility; an expected co-efficient should be positive, demonstrating highly leveraged firms tend to exit the market. Further, the effect of leverage may not be very obvious here, as there might be firms that are highly liquid in nature and can survive even when highly leveraged. So, we controlled for liquidity in the equation to see the effect of leverage in the presence of liquidity. The specification will be:

$$E_{fit} = \beta leverage_{ft-1} + \delta liquidity_{ft-1} + \gamma X_{ft-1} + \varphi_i + \psi_t + \epsilon_{fit} \tag{2}$$

The co-efficient for liquidity is expected to be negative as more liquid firms will tend to stay in the export market in the short run. The argument is that the coefficient for leverage is still positive, showing the credit constraint nature of a particular firm and how it is affecting the export decision. Leverage ratios are a very common measure of financial constraints in the finance literature. Buch et.al. (2014) used tangible assets as a proxy for the fixed cost of exporting, and we can argue that tangible asset ratio is synonymous to leverage ratio. Higher tangible asset ratio means more stressed firm leading to a higher probability of exit. Buch et.al. (2014) also used cash flows (similar to liquidity) as a source of internal fund which can be used to relax financial constraint.

It is important to note that firms can get loans from the bank to relax their credit constraint and hence it becomes really important to have a look at the pattern of borrowing for these firms. As, it was discussed in the earlier section that in India more than 70% of the total loan is disbursed by the state-owned banks and remaining share is majorly serviced by handful of large private sector banks. Now, it is very important to see the effect on firms that are borrowing from a bank or have a relation with a bank on their decision to stay/exit the export market. Can it be the case that firms that are highly leveraged in terms of bank borrowing (narrower definition of leverage) are staying in the export market compared to less leveraged one. If this is the case then it becomes imperative to look at the pattern of borrowing for the firms, like the type of bank they are borrowing from, number of banks they are borrowing from etc. In order to address the question of whether or not firms borrowing from a particular type of banks are less likely to leave than the one that are not able to borrow, I include the bank borrowing dummy (or number bank borrowed from) along with the interaction between the leverage ratio and bank borrowing dummy, yielding the following specification:

$$E_{fit} = \beta leverage_{ft-1} + \delta liquidity_{ft-1} + \gamma X_{ft-1} + \tau bank_{ft} + \rho(leverage_{ft-1} * bank_{ft}) + \varphi_i + \psi_t + \epsilon_{fit} \quad (3)$$

Here, $bank_{ft}$ is a dummy for the type of bank the firm f is borrowing in period t and $leverage_{ft-1} * bank_{ft}$ is the respective interaction between the two. The coefficient for interaction term should be negative, implying that those firms which are borrowing from a particular type of banks are less likely to exit the export market compared to the one which are not borrowing. We have used different types of bank borrowing dummies to test the hypothesis like firms borrowing from at least one public bank, firm borrowing from only public bank, firm borrowing from more than one bank, and firms borrowing from only private banks. Also, for leverage ratio, we have used the standard definition of debt to capital ratio, asset to capital ratio. Alongside these indicators we have added bank borrowing indicators such as bank borrowing to asset ratio (which is used as a narrower definition to leverage ratio). This result will validate the claim that firms that are borrowing from banks are less likely to leave the export market despite being highly leveraged.

5.1 Policy Change

As discussed earlier, priority sector lending is an important role given by the Reserve Bank of India to various scheduled commercial banks for an all round development of the economy as opposed to focusing on financial sector only. Providing export credits to exporting firms was one of the several categories where bank can provide loan under priority sector lending. Till 2015, export credit was part of priority sector for foreign banks with 20 or less bank branches, but in April 2015, the Reserve Bank of India mandated that domestic banks were also allowed to register export credit as priority sector lending. The idea of priority sector lending was to facilitate a holistic development of the economy and India being an emerging economy, this can be achieved through banking sector. Export credit was initially assigned to foreign banks only for priority sector, the reason being their low penetration to small towns and villages where most of the fund for priority sector can be allocated. But, as we saw earlier that micro, small or medium enterprises, and the agriculture sector comprise a sizable part of India's export part; hence it made sense to add export credit to domestic banks as part of priority sector lending to facilitate more funds to exporting firms.

This policy shift creates an exogenous shock in the credit market and for firms seeking credit will increase substantially in form of export credits.[13] It is important to see how the banks that are missing their priority sector targets are reacting to loaning to leveraged firms after the policy shift. It can be argued that these firms that are already in the exporting market and have past relation with the banks can be seen as low hanging fruits for the banks to fulfill their target. Since, this is an exogenous shock, it can be used to treat the endogeneity that was there in bank's lending pattern to different types of firm across various industries. It is important to see how the banks that are missing their priority sector targets are reacting to loaning to leveraged firms already in the export market after the policy shift. The analysis will be a difference in differences analysis, with the policy shock of 2015 used as a time variance for exogenous shock and the treatment groups are those exporting firms that are borrowing from state-owned banks that are missing their priority sector target.
The specification with adding priority sector policy change will be:

$$\begin{aligned}E_{fit} = \beta leverage_{ft-1} + \delta liquidity_{ft-1} + \tau PSL_{ft} + \rho(leverage_{ft-1} * PSL_{ft}) + \\ \nu(leverage_{ft-1} * Post_t) + \eta(leverage_{ft-1} * Treat_f) + \gamma X_{ft-1} + \varphi_i + \psi_t + \epsilon_{fit} \quad (4)\end{aligned}$$

Here, E_{fit} , $liquidity_{ft-1}$, $leverage_{ft-1}$, and X_{ft-1} have the same definition as earlier equations, where as PSL_{ft} is defined as priority sector dummy. PSL_{ft} is the dummy variable that takes value 1 if any of the state-owned bank from which a firm is borrowing misses its priority sector lending target after 2015 and 0 otherwise. This creates those firms borrowing from defaulting state-owned banks as treatment firms against those who are not borrowing from such banks. $Treat_f$ takes a value of 1 if a firm belongs to the treatment group and 0 otherwise. Similarly, $Post_t$ takes a value of 1 if time period is post-2015 and 0 otherwise. The coefficients of interest are τ and ρ respectively that provide us with the average treatment effect on the treated (τ) and a differential impact of the policy change on the treated group with higher bank leverage (ρ). The sign for τ and ρ are expected to be negative, implying that the highly leveraged firms that are borrowing from state-owned banks that are missing on their priority sector targets have a lower probability of leaving the export market than the firms that are not able to borrow from these banks.

[13]Reserve Bank of India

6 Results

6.1 Results without bank information

Table 4 presents the baseline results for the effect of the leverage ratio on firms decision to exit or continue in the export market. Here, we are looking at different definitions of leverage ratio (asset to capital, debt to capital, or bank borrowing to asset ratios), while in the result we are controlling for firm-specific variables [14], year, region and firm fixed effects. The results for firm controls are quite intuitive and they make sense with respect to the exit decision of firms. Similarly, when we look at the leverage ratio coefficients, in the first 2 columns, they are positive, showing that 75th percentile leveraged firm has a 2.1 percentage point more probability of leaving the export market compared to a 25th percentile firm; whereas in the last column, firms that are more leveraged in terms of bank borrowing have a lesser chance of exiting compared to those that cannot. A 75th percentile leveraged firm has a 1.7 percentage point less probability of leaving the export market compared to a 25th percentile firm.

Table 4
Regression output for Exiting firm and leverage ratio

Variable	(1) Exit Exporter	(2) Exit Exporter	(3) Exit Exporter
Total Asset to Capital	0.0089**		
	(0.004)		
Debt to Capital		0.0089**	
		(0.004)	
Bank Borrowing to Total Asset			-0.0071***
			(0.002)
Firm Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes
Observations	31,550	34,614	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Clearly, there is a difference in the number of observations when we look at the result for the leverage ratio related to debt and capital and that from bank borrowing, which is due to the fact that there are some firms that are not reporting their borrowing information, it

[14]Note: Firm controls consists of age, log of sales, profit to sales ratio, ownership status

is assumed in the study that these firms have no source of bank borrowing. It is important to point out that in all results, we are using firm-level controls along with the region, year, and firm level dummy to capture the respective fixed effects.

In Table A1, we include different indicators of liquidity ratio (cash flow in the financial activity to asset ratio and total cash flow to asset ratio) along with different measures of the leverage ratio. It is to be kept in mind that in all these results, we are using firm-level controls and the region, year, and firm specific dummies for fixed effects. Adding the liquidity ratio to the equation shows that the result with respect to the leverage ratio becomes more stronger (when we take it as the ratio of debt to capital and otherwise), showing that there is a higher probability of exit for highly leveraged firms. A 75th percentile leveraged firm has a 2.5 percentage point higher probability of exiting the export market vis à vis 25th percentile file. However, when we look at the leverage ratio in terms of the bank borrowing to asset (narrower definition), then the results are still significant and negative but smaller than the earlier result due to the effect of liquidity. It is important to note that the liquidity ratio is not significant in columns 3 and 6, but economically it makes sense. Overall, we can say that adding a liquidity ratio to the model further consolidates the fact that the more leveraged (financially constrained) firms are more likely to exit the export market compared to lesser leveraged firms.

6.2 Include type of bank to the model

In the previous subsection, we saw that more-leveraged firms are more likely to exit the export market compared to others. However, when we look at leverage in terms of bank borrowing, it showed that leveraged firms have a lower probability to leave export market. Now, we further examine what happens to the firms' exit decision when we add information about type of bank a firm is borrowing into the model and interact it with the leverage ratio (both debt as well as borrowing one). Borrowing from the type of bank can be seen as a measure of credit constraint (Banerjee and Duflo (2014)), hence it can affect a firm's decision to continue or exit the export market.

6.2.1 Firms borrowing from at least one state-owned bank

In India where more than 70% of total loans are given out by state-owned banks, it is interesting to see the effect of borrowing from a state-owned bank compared to not borrowing from a state-owned bank at all. This can show what happens to a leveraged firm when they borrow from at least one state-owned bank. In Table 5, we included a dummy variable for borrowing from at least one state-owned bank and also interacting with other measures of leverage ratio (be it from a debt or borrowing side), the interaction term shows the effect of borrowing from at least one state-owned bank on a leveraged firm's decision to exit or continue in the export market.

Table 5
Regression output: Leverage Ratio and Public Bank dummy and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Total Asset to Capital	0.0089**			0.0090**		
	(0.004)			(0.004)		
Bank Dummy*Leverage ratio	0.0118			0.0115		
	(0.016)			(0.016)		
Debt to Capital		0.0089**			0.0090**	
		(0.004)			(0.004)	
Bank Dummy*Leverage ratio		0.0118			0.0115	
		(0.016)			(0.016)	
Bank Borrowing to Total Asset			-0.0054**			-0.0052**
			(0.002)			(0.002)
Bank Dummy*Leverage ratio			-0.0210*			-0.0206*
			(0.011)			(0.012)
Cash flow in financial activity to Asset	-0.0010**	-0.0010**	-0.0011			
	(0.001)	(0.001)	(0.001)			
Total cash flow to Asset				-0.0018***	-0.0018***	-0.0005
				(0.000)	(0.000)	(0.001)
Atleast 1 Bank Dummy	-0.023***	-0.023***	-0.029***	-0.023***	-0.023***	-0.029***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,230	34,301	24,709	31,550	34,614	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

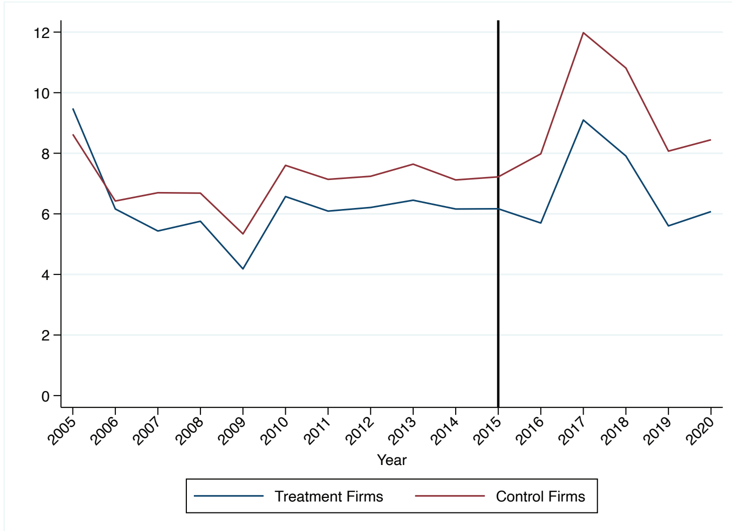
Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

The results clearly show that firms borrowing from at least one state-owned bank have a lower probability of exiting the export market compared to those not borrowing from state-owned banks. The probability is lower by 2.3 percentage points when leverage ratio is calculated using debt to capital and around 2.9 percentage point when leverage is calculated using bank borrowings. It is important to turn our attention to the leverage ratio and their respective differential effect with respect to type of bank a firm is borrowing from. Interestingly, none of the interaction terms with any of the leverage terms were significant, other than the one with respect to bank borrowing. This means that when borrowing from at least one state owned bank, a highly leveraged firm has an additional 3.5 percentage point lower probability of leaving the export market compared to a less leveraged firm. Whereas, in the case of the leverage ratio using capital for its definition, the interaction terms were not significant, but the leverage ratio were significant, implying that leveraged firms that are borrowing or not borrowing from a state-owned banks do not differ significantly in terms of the probability of them exiting the export market. Other such cases are discussed in appendix B.

6.3 Priority Sector Lending policy change

The lending pattern of state-owned banks to firms is not exogenous in nature and there are some sort of endogeneity to the type to firms they are lending, be it related to the type of industries, relation with banks, objective of state-owned banks. In short, we can say that there is some kind of sorting happening between banks and firms. In this subsection, we will look at the policy change of 2015 in the priority sector lending, which is used for the difference-in-differences analysis to establish causality and to treat endogeneity in the previous results as it is clear that banks might be interested in providing loans to particular type of firms. Our coefficient of interest will be the interaction term between the priority sector lending dummy and leverage ratio in terms of bank borrowing. To further motivate the result, figure 8 shows the effect of credit shock on treated and control firms.

Figure 8: Effect of Priority sector lending shock on different firms



Here treated firms are those that are borrowing from at least one state-owned bank that is missing their priority lending target. In contrast, control firms are those firms that are borrowing from state-owned banks that are not missing their lending target. The plot also establishes the pre-intervention parallel trend between control and treatment group, hence fulfilling the difference in differences assumption. Clearly, the plot shows that treated firms have a lower probability of leaving the export market than control firms since the policy change. We further motivate ourselves for the parallel trend using the event study diagram shown in figure 9. Since, we do not have selective treatment timing in our study, hence the event study does validate the pre-intervention parallel trend assumption. It can also be argued that the treatment and control groups will evolve uniformly over time if there are no intervention. The only difference they have in accessing credit is the state-owned banks ability to fulfill their priority sector target. This will not have any significant difference for exporting firms as these firms cannot get benefit from banks missing their target, hence it is rational to assume pre-intervention parallel trend between control and treatment group and it is validated in figure 8 (parallel trend) and 9 (event study). We further tested assumptions for time varying covariates in appendix D.

Figure 9: Estimates of policy change effect on exit rate using event-study design

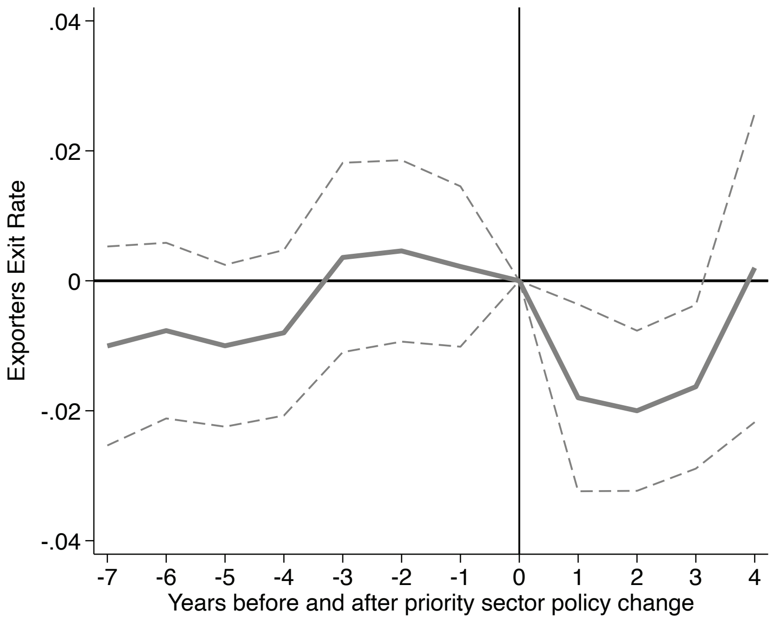


Table 6 includes the dummy variable for firms borrowing from state-owned banks missing their priority sector target and its interaction with the leverage ratio (both definitions). Clearly, the priority sector dummy was negative in all cases and significant in the case where leverage was defined as bank borrowing to total asset, implying firms borrowing from banks missing their priority sector target have a lower probability of exiting the export market because they now have access to easy credit compared to other firms. The results show that the firms' probability to exit the market is approximately 1.3 percentage points lower (this result is valid for the cases where leverage ratio is used as bank borrowing to total assets) than for those firms not able to borrow from banks that are missing their targets. Looking at the result for the leverage ratio and their respective interaction with the dummy variables, the interaction with the leverage ratio in terms of bank borrowing is significant. This means that if a firm is leveraged in terms of bank borrowing and that bank is missing their priority sector target, then it will further reduce the firm's probability of exiting the export market by approximately 2.6 percentage points (in addition to 1.3 percentage points when not borrowing). Whereas, in the case of leverage ratio using debt to capital for its definition, the interaction term was not significant but the leverage ratio was significant, implying no difference to the leveraged firms that are borrowing from banks missing their target or not missing target. This result shows that banks are actually looking for low-hanging fruits to fulfill their priority sector target and for doing so, they are actually providing credit to firms that are more leveraged (in terms of bank borrowing), hence effectively creating a soft budget constraint.

Table 6
Regression output: Leverage ratio and priority sector lending and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Debt to Capital	0.00370***	0.00367***				
	(0.001)	(0.001)				
PSL Dummy*Leverage ratio	0.00482	0.00516				
	(0.018)	(0.018)				
Total Asset to Capital			0.01491***	0.01478***		
			(0.003)	(0.003)		
PSL Dummy*Leverage ratio			-0.00423	-0.00315		
			(0.022)	(0.022)		
Bank Borrowing to Total asset					-0.00671***	-0.00719***
					(0.002)	(0.003)
PSL Dummy*Leverage ratio					-0.56514***	-0.56471***
					(0.20)	(0.20)
PSL Dummy	-0.00442	-0.00389	-0.00594	-0.00542	-0.01321*	-0.01351*
	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Cash Flow in Fin. Act. to Asset	-0.00132**		-0.00098**		-0.00114	
	(0.001)		(0.000)		(0.001)	
Total Cash Flow to Asset		-0.00221***		-0.00169***		-0.00080
		(0.000)		(0.001)		(0.001)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26,250	25,898	23,884	24,107	18,821	18,938

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status
Note: All standard errors are clustered at industry level
Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The banks that are missing their target have to buy priority sector certificates or invest in rural infrastructure development fund, which have a rate of return equivalent to the baseline bank rate (which is generally lower than the existing lending rate under priority sectors). This actually creates a situation where banks may not be willing to provide loans to highly leveraged firms when they are close to their priority sector target, i.e., missing their target by only a small percent; while on the contrary, their willingness increases when they are far from hitting their target. Eventually they will not care about their target achievement if they are far from the target, where an outside option of buying a priority sector certificate or invest in rural infrastructure development fund is considered better than investing in some risky project. Tables A7 tried to test this hypothesis that banks closer or far from their target may not lend to these leveraged firms, while the banks in the middle will do it. Here, the priority sector lending missing value was put under a 5% bracket to see how banks react to their lending decision once they are far away from their target percentage. The hypothesis held well and the interaction term for bank borrowing and the priority sector dummy were significant and negative for 5% - 10% and 10% - 15%, respectively, and non-significant for 0 - 5% and above 15%, respectively. This is an interesting result as it shows that banks that are not far away from achieving their target but not very close either are the ones creating a soft budget constraint for those highly leveraged firms in the form of providing easy credit for them.

6.4 Placebo Test

This part looks at the placebo test for the policy change. We investigate the effect of the priority sector lending policy change on non-exporting firms and medium and small enterprises, and agricultural firms that are not exporting. As the policy change primarily affected exporting firms and in particular medium and small enterprises, and agricultural firms were also affected. It is important to see that the non-exporting firms in general and under these specific groups were not affected in any sense by the change in policy. It might have been the case that banks missing their targets might have provided extra loans to non-exporting firms rather than providing loans to newly added firms as a form of export credit. Hence, non-exporting medium and small enterprises, and agricultural firms and non exporting firms in general present to us a good case of fake treatment group for our placebo test.

Tables 7 shows the effect of policy change on non exporting firms. We further see the effect of this policy change on non-exporting MSMEs and agricultural firms in table A8 and A9 respectively. These results show whether the change in policy helped them to shift from being a non-exporter to an exporter, or whether they remained the same or were even adversely affected. We start the placebo test with overall firms that are not exporting and borrowing from state-owned bank that are missing their priority sector target as a fake treatment group. Table 7 shows the effect of policy change on non-exporting firms, and the results shows no effect of policy change on firms decision to move from non-export to export sector. The priority sector dummy and its interaction with leverage ratio (in terms of bank borrowing) are not significant. Although the interaction of leverage ratio (in term of debt to capital) is negative and significant, implying highly leveraged firms in overall debt that borrow from state-owned banks that are missing their priority sector target have lower probability to be non-exporting. However, our study is based on firms leverage in terms of bank borrowing and banks lending pattern to such firms, it can be safely said that such a policy change didn't have any effect on leveraged firms decision to move from non exporter to exporter.

Table 7
Regression Result: Leverage ratio, priority sector lending, their interaction for Non Exporters

	(1)	(2)	(3)	(4)
Dependent Variable: Never Exporter				
Debt to Capital	0.00205**	0.00206**		
	(0.001)	(0.001)		
PSL Dummy*Leverage ratio	-0.03253**	-0.03106**		
	(0.016)	(0.015)		
Bank Borrowing to Total Asset			-0.03582***	-0.03291***
			(0.010)	(0.011)
PSL Dummy*Leverage ratio			-0.41261	-0.26957
			(0.355)	(0.352)
PSL Dummy	-0.00384	-0.00361	-0.00920	-0.01204
	(0.010)	(0.010)	(0.012)	(0.012)
Cash flow in financial activity to Asset	-0.00858**		-0.00253***	
	(0.004)		(0.001)	
Total cash flow to Asset		-0.02592***		0.00096
		(0.008)		(0.004)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	52,970	55,594	38,292	38,562

Note: Note: Firm controls consists of public bank dummy, log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status
Note: All standard errors are clustered at industry level
Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

Tables A8 and A9 discuss the effect of policy change on specific types of non exporter which can be seen as fake treatment groups.Table A8 shows the effect of the policy change on non-exporting medium and small enterprises, and it is clear from the results that the policy change had no effect on the decision of these non-exporting firms. Both the priority sector dummy and its interaction with the leverage ratio (in terms of the debt to capital or bank borrowing to the total assets) coefficients are insignificant; hence we could not infer anything from these coefficients. Similarly, table A9 shows the effect of policy change on non exporting agriculture firms, it is clear from the results that the policy change had no effect on the decision of those non-exporting firms. The coefficients of the priority sector dummy and its interaction with the bank borrowing to asset ratio were insignificant, and hence any economic or statistical inference does not make sense. The result establishes that the policy change did not have a significant effect on the medium and small enterprises, and agricultural sector non-exporters. The placebo test reveals a zero impact on the fake treatment groups; hence, it supports our parallel trend assumption for the model too.

7 Robustness check

7.1 Effect of policy change on firms entering export sector

The study was primarily based on firms that are already in the export market and at any time t they are deciding whether to stay in or exit the export market. The results show that firms that are highly leveraged and that borrow from state-owned banks have a lower probability of exiting the export market compared to other firms. But, this does not establish that firms that are not in the export markets do not experience the same effect. In this part, we wanted to test the hypothesis that firms that want to enter the export market are not able to benefit from the change in policy related to priority sector lending, and it's only those already in the export market that are benefited. Table 8 shows the regression results, where the dependent variable was changed from exiting firms to the probability of a firm entering the export market to see the effect of policy change on their decision to enter the market.

Table 8
Regression Result: Leverage ratio, priority lending, their interaction

	(1)	(2)	(3)	(4)
Dependent Variable: Entry Dummy				
Debt to Capital	0.00032	0.00030		
	(0.001)	(0.001)		
PSL Dummy*Leverage ratio	-0.02725*	-0.03030**		
	(0.016)	(0.016)		
Bank Borrowing to Total Asset			0.05905	-0.00571
			(0.044)	(0.004)
PSL Dummy*Leverage ratio			0.13650	0.12500
			(0.130)	(0.100)
PSL Dummy	-0.00055	-0.00207	-0.00140	-0.00170
	(0.006)	(0.006)	(0.007)	(0.007)
Cash flow in financial activity to Asset	0.00077		0.00256	
	(0.001)		(0.002)	
Total cash flow to Asset		0.00550		-0.0182
		(0.010)		(0.040)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	19,256	20,452	13,875	14,078

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The results show that there was no significant effect on firms' entry decisions due to the change in policy. The priority sector dummy and its interaction with the bank borrowing to asset ratio were not significant, hence showing that the policy change had no effect on entering firms (although the interaction of the priority sector dummy and debt to capital was significant and negative, which further established the fact that the policy change had no effect on entering firms). These results clearly show that banks (primarily state-owned banks) are going for low hanging fruits in the form of highly leveraged firms that are already in the export market to fulfill their priority sector lending target.

7.2 Effect of policy change on private banks missing priority lending target

Another robustness check that we do is to see the effect of the policy change on the lending behavior of private banks. The hypothesis is that policy change does not affect the exit rate of exporting firms borrowing from private banks who are missing their priority sector lending (and not state-owned banks). If this hypothesis holds, then we can conclude that the policy change affected highly leveraged exporting firms' exit rates that are borrowing from state-owned banks and missing their priority sector target. To test the hypothesis, we run the same regression with a change in priority sector lending dummy, i.e., the dummy is defined as those firms that are borrowing from at least one private bank who are missing their priority sector target after the policy shock are assigned value of one otherwise zero. Table 9 shows the regression results for different definitions of leverage. Our leverage ratio of interest is bank borrowing to the total asset (leverage in terms of bank borrowing) and its interaction with priority sector dummy. The result shows that private banks missing their priority sector lending target have no significant effect on firms' exit rate when we control firms' leverage in terms of bank borrowing. The interaction term is not significant, hence showing that highly leveraged firms borrowing from private banks who are missing their priority sector target do not have a lower exit rate than firms that are less leveraged. Although the leverage ratio in bank borrowing is negative and significant, it shows that leveraged firms have a lower probability of leaving the export market. The robustness results are more promising when we look at the result of leverage in terms of debt to capital ratio, here both leverage and its interaction are positive and significant. Hence, firms borrowing from private banks that miss their priority sector lending target have a higher probability of exiting. Therefore, we can say that the policy change does not positively affect firms' decision to stay in the export market if they borrow from private banks that are missing their priority sector lending target. This result shows that private banks are not lending to highly leveraged firms that are already present in the export market to fulfill their priority sector lending target. It further strengthens the result that state-owned banks are lending to leveraged exporting firms to meet their priority sector lending target.

Table 9
Regression Result: Leverage ratio, priority lending private only, their interaction and public bank dummy

	(1)	(2)	(3)	(4)
Dependent Variable: Exit Dummy				
Debt to Capital	0.00390***	0.00393***		
	(0.001)	(0.001)		
PSL Private Bank*Leverage ratio	0.07610***	0.07980***		
	(0.020)	(0.020)		
Bank Borrowing to Total Asset			-0.00703***	-0.00762***
			(0.003)	(0.003)
PSL Private Bank*Leverage ratio			-0.22172	-0.22629
			(0.300)	(0.300)
PSL Private Bank	-0.01980***	-0.01992***	-0.01042	-0.01196
	(0.007)	(0.007)	(0.010)	(0.010)
Public Bank Dummy	-0.01484***	-0.01462***	-0.02846***	-0.02860***
	(0.005)	(0.005)	(0.005)	(0.005)
Cash flow in Financial Activity to Asset	-0.00138		-0.00143	
	(0.001)		(0.001)	
Total Cash Flow to Asset		-0.00202***		-0.00070
		(0.001)		(0.001)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	18,751	19,160	13,798	13,872

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

We further divided our sample to manufacturing and service sector. Using the sub sample, we check the effect of policy change in priority sector lending on leveraged firms borrowing from state-owned banks that are missing their priority sector target. We document the result of the sub-sample in appendix C.

7.3 Productivity indicator, bank borrowing and effect of policy change

This part looks at the effect of the leverage ratio on various indicators of productivity. We further look at what happens when we look at those firms that are borrowing from state-owned banks. we used a few indicators for productivity, namely the log of the sales to capital ratio and the sales to asset ratio. The rationale of using sales to capital as an indicator of productivity is that it gives an indication of the level of production for which sales are used as a proxy for the level of capital or investment made by a firm. In this case, we define productivity in the form of the sales that a firm makes for the level of investment it makes. Another indicator that we looked as an indicator of productivity is the sales to asset ratio; this can be seen to be similar to the sales to capital ratio but here I looked at productivity in the form of the total sales made by a firm relative to the total assets of the firm. These two indicators were used to see how the leverage and borrowing pattern of firms affect their productivity.

Table 10 shows the effect of the leverage and borrowing pattern on the different indicators of productivity.The table shows that firms that are borrowing from state-owned banks were less productive compared to those not borrowing from state-owned banks (this was true for both the productivity variables that we used in our analysis). This made sense, as state-owned banks lend to those firms that are generally highly leveraged (as we have seen earlier) and their objective of lending to such firms is not profit maximization but welfare maximization; hence they lend to sectors that are low in profitability and also productivity. When we looked at the leverage ratio as debt to capital, we found that more-leveraged firms were less productive, but when we looked at the leverage ratio as bank borrowing, more-leveraged firms were more productive, indicating that when firms are leveraged and borrowing from a bank, then they are more productive compared to other firms (it is important to note that these numbers were not significant statistically but we were trying to look at economic significance). The coefficient for the priority sector lending dummy variable was insignificant, although the signs were the opposite for the different productivity variables. Similarly, the coefficient of interaction of the priority sector lending dummy and leverage ratio was not significant but it had a negative effect on productivity when we looked at bank borrowing as an indicator of the leverage ratio. Although the effect of the bank borrowing leverage ratio had a positive effect on productivity, firms that were benefiting from the priority sector lending were less productive than those firms that were not using the benefits of priority sector lending. It is important to note that most of the coefficients were statistically insignificant, but economically these coefficients make sense when we look at the relation between productivity, leverage, and the bank borrowing status.

Table 10

Regression Result: Several indicators of productivity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variables	Sales Capital (log)		Sales Capital (log)		Sales to Asset		Sales to Asset	
Debt to Capital	-0.02094	-0.02031			-0.00549***	-0.00392***		
	(0.027)	(0.027)			(0.002)	(0.001)		
PSL Dummy*Leverage ratio	0.30834	0.30681			-0.01990	-0.03124		
	(0.245)	(0.245)			(0.035)	(0.033)		
Bank Borrowing to Total Asset)			-0.00048	0.00011			0.88825	0.67053
			(0.010)	(0.010)			(0.818)	(0.701)
PSL Dummy*Leverage ratio			-0.33167	-0.29594			-0.90279	-0.77600
			(0.956)	(0.958)			(0.815)	(0.748)
PSL Dummy	-0.01736	-0.02892	-0.03775	-0.03727	0.04456	0.02348	0.04345	0.03154
	(0.042)	(0.042)	(0.054)	(0.054)	(0.046)	(0.035)	(0.064)	(0.059)
Public Bank Dummy	-0.13426***	-0.14609***	-0.03344	-0.03555	-0.15255***	-0.08221**	-0.07652*	-0.07159*
	(0.051)	(0.052)	(0.056)	(0.056)	(0.044)	(0.032)	(0.043)	(0.042)
Cash flow in financial activity to Asset	0.00765		0.00425		0.25189		0.13499	
	(0.008)		(0.005)		(0.158)		(0.116)	
Total cash flow to Asset		0.00613		0.00332		0.56238***		0.31154**
		(0.017)		(0.004)		(0.103)		(0.137)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32,955	33,334	23,704	23,720	32,960	33,339	23,708	23,724

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

8 Conclusion

This paper, we contribute to the empirical literature on trade-finance linkages. In particular, we studied the causal impact of credit constraints on exporting firms with regard to their decision to exit or continue in the export market and how borrowing from state-owned banks created a soft budget constraint to less efficient firms. The main findings in this paper are that more-leveraged firms (in terms of debt to capital) have a disadvantage over less leveraged firms, and on an average, they have a 2.1 percentage point higher probability of exiting the export market. These results support the conclusion that those leveraged firms that are able to borrow, be it from banks or any other institution, have a lower probability of exiting the market (on an average of 1.5 percentage points) compared to a less leveraged firm. Given the banking system of India, we used the firms' borrowing information from various types of banks (state-owned and private) as a dummy for credit availability. The results for the bank borrowing information showed that leveraged firms borrowing from at least one public sector (state-owned) bank have an additional 3.5 percentage lower probability of exiting the export market compared to less leveraged firm, hence showing that these firms have better access to credit than the other firms. This result also holds for leveraged firms that borrow from more than one bank (additional 4% lower probability of exiting the export market compared to less leveraged firms. However, we did not find any significant result for firms that were borrowing from only state-owned banks.

The causal relation was established using a difference-in-differences approach and the exogenous shock used was the change in priority sector lending policy for "domestic scheduled" commercial banks in April 2015. The result showed that, on average, a firm borrowing from at least one state-owned band that is missing its priority sector target has a 1.3 percentage lower probability of exiting the market and, in addition, a 2.6% lower probability of exiting the export market for highly leveraged firms compared to less leveraged firm. Along with this, when a public bank dummy was added, it was found that firms borrowing from these banks have a lower probability of exiting the export market in the range of 2.3 percentage points to 2.9 percentage points, depending on the variable used for the leverage. The policy shock related to priority sector lending turned out an important shock as it helped to purge out the endogeneity in the lending pattern of banks to various firms.

The robustness checks for the various indicators established the validity of the results. We did various robustness checks in the form of checking how firms that are entering the export market react to the policy change. The robustness check showed that this shock did not have any significant effect on these firms; hence establishing that the policy mostly affected the firms already in the export market and their exit/stay in decision. A similar robustness check was done for non-exporting firms, non-exporting medium and small enterprise, and agricultural firms. We have used these firms as placebo test for our results. The results for these kind firms were not significant; hence it helped us to conclude that only exporting MSMEs and agricultural firms were the ones affected by the change in policy (and this should be the case as the policy adds export credit to these sectors). We looked at some of the productivity indicators and analyzed the effect of the leverage and borrowing patterns on productivity. The results showed that firms borrowing from state-owned banks are less productive compared to firms not borrowing from them. The result also showed that firms that are borrowing from banks and have high leverage are more productive, although the results were not significant. It was further shown that firms with high leverage and that were borrowing from banks but using the benefit of priority sector lending were less productive than firms not using the benefit of priority sector lending, but it is to be noted that the results here were not significant either. We finally looked at the effect of policy change across sectors and the results were similar across sectors. The policy change showed a significant negative effect on leveraged firms to exit the export market (lower probability to exit) for manufacturing and service sector. The result indicates that the policy change had a uniform effect across sector.

The study clearly shows that the presence of public sector (state-owned) banks in the Indian economy is creating a distortion in the credit market from the supply side. This distortion is helping highly leveraged firms to stay in the export market, which would have not been the case in an ideal situation. The public sector (state-owned) bank lending pattern is creating a soft budget constraint for these leveraged firms. This hypothesis was further confirmed using an exogenous policy shock in the form of priority sector lending, along with policy analysis, while various robustness checks also established the proposition of a soft budget constraint for these leveraged firms.

There are a few things that we can look further in a separate studies; for instance, it is important to look at the spatial distribution of those firms that are exporting along with the distribution of bank branches for public and non state-owned banks. This will, however, help us to see the pattern of borrowing for these firms and allow us to deduce any spatial relation between exporting firms and their respective credit access. The bank branch location along with the firms' location will be helpful for us in understanding the relation between firms borrowing and their credit access. Another important aspect that can be looked at is the quality of the bank (as we are looking at the supply side of credit). We have financial information about banks (like NPA, tier 1 capital, total loan, borrowing, government ownership etc.), which will help us to create a variable/index for the quality of banks that a particular firm is borrowing from and how that is creating a credit constraint problem for the firms.

References

[1] Albornoz, Facundo; Fanelli, Sebastián and Hallak, Juan Carlos "Survival in export markets" *Journal of International Economics* 102 (2016) 262–281

[2] Allen, F. and Gale, D. "Comparing financial systems" *MIT press* (2001)

[3] Alfaro, Laura; Kalemli–Ozcan, Sebnem and Sayek, Selin "FDI, Productivity and Financial Development" *The World Economy*, (2009), 111–135

[4] Araujo, Luis; Mion, Giordano and Ornelas, Emanuel "Institutions and export dynamics" *Journal of International Economics* 98 (2016) 2 –20

[5] Bannerjee, A; Cole, S. A, and Duflo, E"Banking Reform in India, Vol. 1" *Brookings Institution Press and National Council of Applied Economic Research*, 277–323

[6] Bannerjee, A and Duflo, E "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program" *Review of Economic Studies*, October April 2014, 81 (2), 572–607

[7] Beck. T "Financial development and international trade: Is there a link?" *Journal of International Economics*, 2002, 57 (1), 107–131

[8] Becker, B.,;Chen, J. and Greenberg, D. "Financial development, fixed costs, and international trade" *The Review of Corporate Finance Studies*, 2013, 2 (1), 1–28

[9] Bento, Pedro and Diego, Restuccia "Misallocation, Establishment Size, and Productivity" *NBER Working Paper No. 22809*, 2016

[10] Besedes, T.; Prusa, T. "Ins, outs, and the duration of trade" *Canadian Journal of Economics*, 2006, 39(1) 266 –295

[11] Buch, Claudia M.; Kesternich, Iris; Lipponer, Alexander and Monika Schnitzer "Financial Constraints and Foreign Direct Investment: Firm-Level Evidence" *Review of World Economics*, 2014, 150 393 –420

[12] Chaney, T. "The network structure of international trade " *American Economic Review*, 2014, 104(11) 3600 –3634

[13] Chaney, T. "The network structure of international trade " *Journal of Economic Dynamics and Control*, 2016 72 141 –154

[14] Claessens, S. and Van Horen, N. "Foreign banks and trade " *Journal of Financial Intermediation*, 2021 45 100856

[15] Das, Sanghamitra; Roberts, Mark J. and Tybout, James R. "Market Entry Costs, Producer Heterogeneity and Export Dynamics" *Econometrica*, 2007, 75(3) 837 –873

[16] Das, S. and Dutta, A. "A Study on NPA of Public Sector Banks in India" *IOSR Journal of Business and Management*, 2014, 16(11) 75 –83

[17] Eaton, Jonathan; Samuel Kortum and Francis Kramarz "Dissecting Trade: Firms, Industries, and Export Destinations" *American Economic Review Papers and Proceedings*, 2004, 94 150 –154

[18] Eaton, Jonathan; Samuel Kortum and Francis Kramarz "An Anatomy of International Trade: Evidence from French Firms", 2007 Working Paper, New York University, Department of Economics.

[19] Eaton, Jonathan; Eslava, Marcela; Kugler, Maurice and Tybout, James "Exports Dynamics in Colombia: Firm-Level Evidence" *NBER Working Paper No. 13531*, 2007

[20] Fernandes, Ana M.; Freund, Caroline and Pierola, Martha Denisse "Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database" *Journal of Development Economics*, 2016 729 –744

[21] Feenstra, Robert C.; Li, Zhiyuan and Yu, Miaojie "Exports and Credit Constraints Under Incomplete Information: Theory and Evidence from China" *Review of Economics and Statistics*, 2014 96(4), 121 –137

[22] Ferri, G.; Minetti, R. and Murro, P. "Credit Relationships in the great trade collapse. Micro evidence from Europe" *Journal of Financial Intermediation*, 2019 40, 100809

[23] Greenaway, D.; Guariglia, A. and Kneller, R."Financial factors and exporting decisions" *Journal of International Economics*, 2007 73(2), 377 –395

[24] Hsieh, Chang-Tai and Klenow, Peter J."Misallocation and manufacturing TFP in China and India" *Quarterly Journal of Economics*, 2009 124(4), 1403 –1448

[25] Hsieh, Chang-Tai and Olken, Benjamin A."The Missing "Missing Middle"" *Journal of Economic Perspective*, 2014 28(3), 89 –108

[26] Ibrahim, M. Syed and Thangavelu, Rangasamy"A Study on the Composition of Non-Performing Assets (NPAs) of Scheduled Commercial Banks in India" *Journal of Finance and Bank Management*, 2014 2(1), 31 –48

[27] Kapoor, Mudit; Ranjan, Priya and Raychaudhuri, Jibonayan"The Impact of Credit Constraints on Exporting Firms: Empirical Evidence from India" *World Economy*, 2017 40, 2854 –2874

[28] Kale, Deeksha"Could Expansions in Directed Lending Programs Hurt Small Businesses? Evidence from a Policy Change in India" *SSRN Working Paper No. 2809713*, 2017

[29] Khanna, T. and Fisman, R. "Facilitating development: The role of business groups" *World Development*, 2004 4(1), 609 –628

[30] Khanna, Tarun and Palepu, Krishna "Emerging Market Business Groups, Foreign Investors, and Corporate Governance " *NBER Working Paper No. 6955*, 1999

[31] Manova, Kalina"Credit constraints, equity market liberalizations and international trade " *Journal of International Economics*, 2008 76(1), 33 –47

[32] Manova, Kalina"Credit Constraints, Heterogeneous Firms, and International Trade " *Review of Economic Studies*, 2013 80(2), 711 –744

[33] Melitz, Marc J"The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity " *Econometrica*, 2003 71(6), 1695 –1725

[34] Minetti, R.; Mulabdic, A., Ruta; M. and Zhu, S. C."Financial structures, banking regulations, and export dynamics" *Journal of Banking & Finance*, 2021 124, 106056

[35] Minetti, R.; Murro, P.; Rotondi, Z., and Zhu, S. C."Financial constraints, firms' supply chains, and internationalization" *Journal of the European Economic Association*, 2019 17(2), 327 –375

[36] Minetti, Raoul, and Susan C. Zhu"Credit Constraints and Firm Export: Microeconomic Evidence from Italy" *Journal of International Economics*, 2011 83, 109 –125

[37] Mukherjee, Shameek, and Mukherjee, Shahana"Overview of India's Export Performance: Trends and Drivers" *IIM Bangalore Research Paper*, 2012 363

[38] Paravisini, D.; Rappoport, V.; Schnabl, P. and Wolfenzon, D."Dissecting the effect of credit supply on trade: Evidence from matched credit-export data" *The Review of Economic Studies*, 2015 82(1), 333 –359

[39] Sant'Anna,P. and Zhao,J."Doubly Robust Difference-in-Differences Estimators" *Journal of Econometrics*, 2020 219(1),101 –122

[40] Topalova, P."Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts" *NBER Working Paper No. 11614*, 2007

[41] Tendulkar, Suresh. D "Indian export and economic growth performance in asian perspective" *Indian Council for Research on International Economic Relations*, 2000

Appendix

A Institutional Detail

A1 Indian Banking Sector

The major issue that has engulfed the Indian banking system over last decade is the ever-growing number of non-performing assets (NPAs here onward). There are many factors that may contribute to the rise in NPAs, but market failure, willful defaults, poor follow-up and supervision, non-cooperation from banks, poor legal framework, lack of entrepreneurial skills, etc. are a few to point out. Dutta (2014) studied the growth of NPAs in public and private sector banks over time and analyzed the sector-wise non-performing assets of these banks. Similarly, Ibrahim et al. (2014) analyzed the concept of NPAs, and the components of loan assets in the state-owned, private sector, and other foreign banks. The reason that NPAs are important for the study of firms' financial constraint is that the state-owned banks are the ones with a very high proportion of NPAs and on the other hand they are the ones responsible for the majority of lending in the Indian banking system. Since, they are themselves constrained in their ability to provide loans, this will have a direct effect on the borrowing of firms in general and even more so exporting firms.

Due to the rise in NPAs among state-owned banks primarily over the past decade, the Reserve Bank of India and the Indian government are planning to do a major restructuring of the banking system. One plan is to merge several small state-owned banks into one large state-owned bank. A similar exercise was performed for the State Bank of India, which was completed in 2017. The hope is to create a more efficient system while addressing the over-staffing issues and making it easier to track all the NPAs and stressed loans as these will then be under one umbrella bank. Since, most of the stressed banks are primarily state-owned banks, privatization of these banks is also a feasible option as currently these banks are not working with the objective of profit maximization but rather they are working with an objective of welfare maximization and supporting sick industries and state-owned enterprises that are not at all profit making. Above all, indeed perhaps most important of all, is the need to look at the governance system of these state-owned banks and here, some reformation of the governance structures of these banks is critical. The Nayak Committee (1997) report on bank governance is a good starting point. Further, there should be a strategy to incentivize and penalize individuals and organizations for any worthwhile initiative as well as for wrong doing.

A2 Indian Export Market

Post 1991, the gradual liberalization of the Indian economy was characterized by policy reforms that created a favorable environment for India's export market to flourish and evolve into an engine of social and economic growth. The process gathered further momentum with India signing the Marrakesh Treaty, which brought into existence the World Trade Organization (WTO) on January 1,1995.[15] In spite of these major changes in India's formal stance toward international integration, globalization, as in the past, is still looked upon with suspicion and apprehension. Hence, it was seen as an inevitable imposition rather than accepted as able to make a healthy contribution to the development process of the country. However, the last two decades have seen India transformed from a closed economy pre-liberalization to an important player in the global international trade.

A3 Priority Sector Lending

The description of the priority sectors was later formalized in 1972 on the basis of the report submitted by the Informal Study Group on Statistics relating to advances in the priority sectors constituted by the Reserve Bank in May 1971. At a meeting of the Union Finance Minister with the Chief Executive Officers of the state-owned banks held in March 1980, it was agreed that banks should aim to raise the proportion of their advances made to priority sectors to 40% by March 1985. Subsequently, on the basis of the recommendations of the Working Group on the Modalities of Implementation of Priority Sector Lending and the Twenty Point Economic Programme by Banks (Chairman: Dr. K. S. Krishnaswamy), all commercial banks were advised of the need to achieve a target of priority sector lending determined at 40% of aggregate bank advances by 1985. Sub-targets were also specified for lending to agriculture and the weaker sections within the priority sectors. Since then, there have been several changes in the scope of priority sector lending and the targets and sub-targets applicable to various bank groups.

On the basis of the recommendations made in September 2005 by the Internal Working

[15] Tendulkar, 2000)

Group (Chairman: Shri C. S. Murthy), set up in the Reserve Bank to examine, review, and recommend changes, if any, in the existing policy on priority sector lending, including the segments constituting the priority sector, sectors, targets, and sub-targets, etc. and the comments/suggestions received thereon from banks, financial institutions, public, and the Indian Banks' Association (IBA), it was decided to include only those sectors as part of the priority sectors that impact large sections of the population, the weaker sections, and sectors that are employment-intensive, such as agriculture, and micro and small enterprises.[16]

The outline for priority sector lending as indicated by the Reserve Bank of India for domestic and foreign banks are the following:

a. **Domestic Bank:** 40% of Adjusted Net Bank Credit (ANBC)[17] or Credit Equivalent Amount of Off-Balance Sheet Exposure, whichever is higher. Foreign banks with 20 branches and above have to achieve the Total Priority Sector Target within a maximum period of five years starting from April 1, 2013 and ending on March 31, 2018 as per the action plans submitted by them and approved by Reserve Bank of India

b. **Foreign banks with less than 20 branches:** 40% of Adjusted Net Bank Credit or Credit Equivalent Amount of Off-Balance Sheet Exposure, whichever is higher; to be achieved in a phased manner by 2020

B Regression Results for type of banks

B0.1 Firms borrowing from only state-owned bank

We further refined the bank borrowing dummy to firms borrowing only from state-owned banks, in order to show what happens to a leveraged firm when they borrow from only state-owned banks and not other banks. In Table A3, we included a dummy variable for borrowing from only state-owned banks and also its interaction with other measures of the leverage ratio (both capital and borrowing definitions). The interaction term shows the effect of borrowing from only state-owned banks on a leveraged firm on its decision to exit or continue in the export market.

[16] Reserve Bank of India
[17] ANBC = Net Bank credit in India + Investment in Non-SLR category

Table A3
Regression output: Leverage Ratio and only Public Bank dummy and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Total Asset to Capital	0.0110**			0.0112**		
	(0.005)			(0.005)		
Bank Dummy*Leverage ratio	-0.062				-0.0682	
	(0.045)			(0.045)		
Debt to Capital		0.0110**			0.0112**	
		(0.005)			(0.005)	
Bank Dummy*Leverage ratio		-0.062			-0.0782	
		(0.045)			(0.045)	
Bank Borrowing to Total Asset			-0.046**			-0.052**
			(0.002)			(0.002)
Bank Dummy*Leverage ratio			-0.082			-0.079
			(0.140)			(0.140)
Cash Flow in Financial Activity to Asset	-0.001**	-0.001**	-0.0008			
	(0.000)	(0.000)	(0.001)			
Total Cash Flow to Asset				-0.0018***	-0.002***	-0.0002
				(0.000)	(0.000)	(0.001)
Only Public Bank	-0.0082**	-0.0082**	-0.003	-0.0079**	-0.0078**	-0.003
	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,230	31,550	34,301	34,614	24,709	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The results showed that firms borrowing from only state-owned banks had significant leverage ratios related to capital (columns 1, 2, 4, and 5) and a lower probability of exiting the export market compared to other firms (probability is around 0.82% and 0.79% respectively) . These results were not significant for the leverage ratio when defined in terms of bank borrowing, but the sign indicates it still makes economic sense. Looking at the results for the leverage ratio and their respective interaction with the dummy variables, none of the interaction terms with any of the leverage term was significant, but here also the sign makes economic sense (leveraged firms borrowing from only state-owned banks have a lower probability of exiting the export market). Whereas, the leverage ratio for both sets of definition were significant, implying that leveraged firms (in terms of capital) have a higher probability of exiting by approximately 2.6 percentage points and leveraged firms (in terms of bank borrowing) have a lower probability of exiting the export market by approximately 1.2 percentage points.

B0.2 Firms borrowing from more than bank

We look in this subsection at the effect of firms borrowing from more than one bank in two ways. Firstly, we see the effect of the dummy variable for firms borrowing from more than one bank and simultaneously its interaction with the leverage ratio (defined in both senses). Secondly, we use the number of banks a firm is borrowing from as a variable and look at the results from that.

Table A4 includes a dummy variable for firm borrowing from more than one bank and its interaction with the leverage ratio (both definitions). Clearly, the bank dummy is strongly significant and negative, implying firms borrowing from more than one bank have a lower probability of exiting the market because they can sustain their credit constraint by borrowing from multiple banks.

Table A4
Regression output: Leverage Ratio and more bank dummy and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Total Asset to Capital	0.0086**			0.0088**		
	(0.004)			(0.004)		
Bank Dummy*Leverage ratio	0.0034			0.0031		
	(0.020)			(0.020)		
Debt to Capital		0.0086**		0.0088**		
		(0.004)		(0.004)		
Bank Dummy*Leverage ratio		0.0034		0.0031		
		(0.020)		(0.020)		
Bank Borrowing to Total Asset			-0.0041**			-0.0047**
			(0.002)			(0.002)
Bank Dummy*Leverage ratio			-0.018*			-0.0176*
			(0.010)			(0.010)
Cash Flow in Financial Activity to Asset	-0.001**	-0.001**	-0.001			
	(0.000)	(0.000)	(0.001)			
Total Cash Flow to Asset				-0.0015***	-0.0015***	-0.001
				(0.000)	(0.000)	(0.001)
More Bank Dummy	-0.023***	-0.023***	-0.029***	-0.023***	-0.023***	-0.029***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,230	31,550	34,301	34,614	24,709	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The result shows that the firms' probability of exiting the market is approximately 2.3 percentage points lower than those firms borrowing from one bank or no bank (this is the result when we use leverage in terms of capital and it is approximately 2.9 percentage points when we use leverage in terms of borrowing). Looking at the result for the leverage ratios and their respective interaction with the dummy variables, none of the interaction terms with any of the leverage terms was significant, other than the one with bank borrowing. This means that if a firm is leveraged with respect to bank borrowing and borrowing from more than one bank then, it will further reduce the firms probability of exiting the market by approximately 4.3 percentage points. Whereas, in the case of the leverage ratio using capital for its definition, the interaction term was not significant but the leverage ratio was significant, implying that there is no difference between leveraged firms that are borrowing from one bank or no bank and more than one bank, whereby the baseline case shows a higher probability of exiting the export market by approximately 2.1 percentage points.

Table A5 shows the results when we used the number of banks a firm borrows from as a variable instead of a dummy, hence there was no interaction to investigate. Clearly, when we look at the number of bank variables, it showed a strongly significant result, and the coefficient was negative in nature. This implies that if a firm's borrowing from banks increases by a unit, then the probability of the firm exiting from the export market declines by 1.6 percentage points (in all cases of leverage ratio). Similarly, the results for the leverage ratio were also significant and consistent with the other results, showing economic significance too. The leverage ratio (in terms of capital) implies that a change in the leverage ratio (defined in terms of capital) will increase the probability of exiting the export market by approximately 2.4 percentage points and in case of leverage ratio (defined in terms of bank borrowing), it will decrease the probability by approximately 1.47 percentage points. It is to be kept in kind that for all the results listed above, we are also using different definitions of liquidity as control and firm-level controls along with the region, time, and firm fixed effects.

Table A5

Regression output: Leverage ratio and borrowing from more than one bank

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Total Asset to Capital	0.0102**			0.00101**		
	(0.005)			(0.005)		
Debt to Capital		0.0102**			0.0101**	
		(0.005)			(0.005)	
Bank Borrowing to Total Asset			-0.0058**			-0.0061**
			(0.003)			(0.003)
Cash Flow in Financial Activity to Asset	-0.0010**	-0.0010**	-0.0010			
	(0.000)	(0.000)	(0.001)			
Total Cash flow to Asset				-0.0016***	-0.0016***	-0.0003
				(0.000)	(0.000)	(0.001)
Number of Bank	-0.005***	-0.005***	-0.004***	-0.005***	-0.005***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,230	34,301	24,709	31,550	34,614	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

B0.3 Firms borrowing from private banks only

In this part, we look at the effect on leveraged firm when they borrow from only private banks and not other banks. In Table A6, we included a dummy variable for borrowing from only private banks and also its interaction with other measures of the leverage ratio (both capital and bank borrowing definitions). The interaction term shows the effect of borrowing from only private banks on a leveraged firm on its decision to exit or stay in the export market. The results showed that firms borrowing from only private banks had no significant effect on firms decision to stay or leave the export market, although the sign is negative.

Table A6
Regression output: Leverage Ratio and Private Bank dummy and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Total Asset to Capital	0.0159***	0.0159***				
	(0.003)	(0.003)				
Bank Dummy*Leverage ratio	-0.0401***	-0.0399***				
	(0.010)	(0.010)				
Debt to Capital			0.0159***	0.0159***		
			(0.003)	(0.003)		
Bank Dummy*Leverage ratio			-0.0401***	-0.0399***		
			(0.010)	(0.010)		
Bank Borrowing to Total Asset					-0.0081**	-0.0079**
					(0.004)	(0.004)
Bank Dummy*Leverage ratio					-0.1968	-0.1918
					(0.250)	(0.250)
Cash Flow in Financial Activity to Asset	-0.0001		-0.0001		-0.002	
	(0.000)		(0.000)		(0.002)	
Total Cash Flow to Asset		-0.0010		-0.0010		-0.0010
		(0.001)		(0.001)		(0.001)
Only Private Bank Dummy	-0.0078	-0.0088	-0.0078	-0.0089	-0.0120	-0.0110
	(0.007)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,230	31,550	34,301	34,614	24,709	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Looking at the results for the leverage ratio and their respective interactions with the dummy variables, the interaction terms with respect to leverage in terms of bank borrowing was not significant, although the sign was negative. The result that is interesting is the interaction of private bank dummy with overall leverage ratio. The result is significant and negative, implying that a firm that is highly leveraged in general sense has a lower probability of leaving the export market if it is borrowing from private banks only. The result is counter intuitive in nature but a possible reason can be that these firms are not highly leveraged in terms of bank borrowing but they have diversified their debt to other instruments hence making it less prone to risk as compared to firms that are leveraged only in terms of bank borrowing. Private bank see those firms as a less risky investment option compared to firms that are leveraged in terms of bank borrowing. The leverage ratio for both sets of definition were significant, implying that highly leveraged firms (in terms of capital) have a higher probability of exiting by approximately 3.8 percentage points and highly leveraged firms (in terms of bank borrowing) have a lower probability of exiting the export market by approximately 2 percentage points.

C Effect of policy change among sectors

This subsection looks at the effect of policy change in priority sector lending on firms borrowing and their decision to stay or leave the export market by different sectors (manufacturing and service sectors). We will look at manufacturing and service sectors in this part. The good thing about the Prowess database is its distribution of firms across sectors, table 1 showed that the firms are equally divided among service and manufacturing sectors. Hence, it is important to look at the effect of policy change across sectors and see whether policy change is affecting firms across both sectors or one sector in particular. Following sub subsections talk about the effect of policy change on leveraged firms across different sectors on their likelihood to stay or leave the export market.

C0.1 Effect of policy change in manufacturing sector

This part looks at the effect of policy change on leveraged firms to stay or leave the export market in the manufacturing sector. Table A10 shows the regression results, which capture the impact of firms' borrowing from a state-owned bank that is missing their priority sector target on probability to exit the export market. The priority sector dummy is not significant when we control leverage in terms of bank borrowing, although the sign is negative. Although, the sign is negative and significant when we look at the dummy's interaction with the leverage ratio (in terms of bank borrowing). It establishes the claim that leveraged manufacturing firms that are borrowing from state-owned banks missing their priority sector target have a lower probability of leaving the export market. Hence, the policy change significantly affects highly leveraged manufacturing firms in terms of bank borrowing to stay in the export market. The result is not significant when we look at the interaction of priority sector dummy with leverage in terms of debt to capital ratio, showing that policy change does not affect firms' exit rate from the export market irrespective of their overall leverage. Results from table A10 establish the claim that manufacturing firms that are leveraged in terms of bank borrowing are positively affected when it comes to staying in the export market if they are borrowing from at least one state owned bank that is missing its priority sector target.

Table A10
Regression Result: Effect of policy change on manufacturing firms

	(1)	(2)	(3)	(4)
Dependent Variable: Exit Dummy				
Debt to Capital	0.01122***	0.01132***		
	(0.003)	(0.003)		
PSL Dummy*Leverage ratio	0.00224	0.00258		
	(0.018)	(0.018)		
Bank Borrowing to Total Asset			-0.00450***	-0.00405***
			(0.001)	(0.001)
PSL Dummy*Leverage ratio			-0.59869***	-0.59913***
			(0.210)	(0.210)
PSL Dummy	-0.01180*	-0.01161*	-0.01026	-0.01082
	(0.006)	(0.006)	(0.010)	(0.010)
Cash flow in Financial Activity to Asset	-0.00074*		-0.00140	
	(0.000)		(0.001)	
Total Cash Flow to Asset		-0.00170***		-0.0019**
		(0.001)		(0.001)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	13,481	13,894	9,434	9,742

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C0.2 Effect of policy change in service sector

As we have seen the effect of policy change on leveraged exporting manufacturing firms, this part of the section looks at the effect of policy change on leveraged firms to stay or leave the export market in the service sector. Table A11 shows the regression results, which capture the impact of firms' borrowing from a state-owned bank that is missing their priority sector target on probability to exit the export market. The priority sector dummy is significant when we control leverage in terms of bank borrowing, and the sign is negative. This means that firms have a lower probability of leaving the export market if they borrow from a state-owned bank that is missing its priority sector target. Along with this, the sign is negative and significant when we look at the dummy's interaction with the leverage ratio (in terms of bank borrowing). It establishes the claim that leveraged service sector firms that are borrowing from state-owned banks missing their priority sector target have a lower probability of leaving the export market. Hence, the policy change significantly affects highly leveraged firms in service sectors in terms of bank borrowing to stay in the export market. The result is also significant when we look at the interaction of priority sector dummy with leverage in terms of debt to capital ratio, showing that policy change does affect firms' exit rate from the export market conditional of their total leverage. Results from table A11 establish the claim that service sector firms that are leveraged in terms of bank borrowing (and also in case of overall leverage) are positively affected when it comes to staying in the export market if they are borrowing from at least one state owned bank that is missing its priority sector target.

Table A11
Regression Result: Effect of policy change on service sector firms

	(1)	(2)	(3)	(4)
Dependent Variable: Exit Dummy				
Debt to Capital	0.00309***	0.00305***		
	(0.000)	(0.000)		
PSL Dummy*Leverage ratio	-0.03462*	-0.03546*		
	(0.018)	(0.018)		
Bank Borrowing to Total Asset			-0.07261	-0.07810
			(0.140)	(0.140)
PSL Dummy*Leverage ratio			-0.60022*	-0.61116*
			(0.340)	(0.340)
PSL Dummy	0.01240	0.01170	-0.02592*	-0.02454*
	(0.010)	(0.010)	(0.015)	(0.014)
Cash flow in Financial Activity to Asset	-0.00382***		-0.02065	
	(0.001)		(0.020)	
Total Cash Flow to Asset		-0.00310**		-0.05460
		(0.001)		(0.050)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	10,584	10,844	8,672	8,824

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

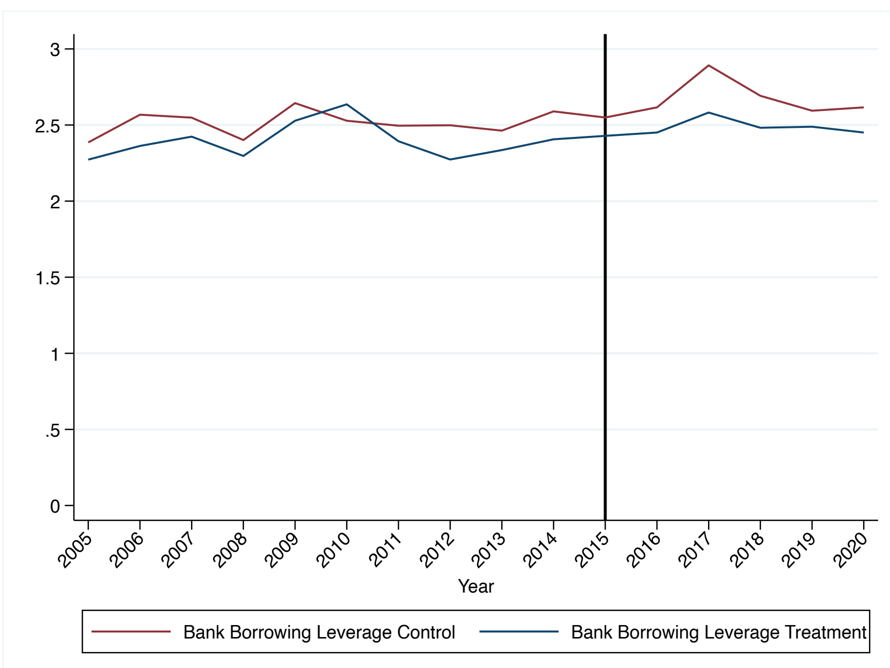
Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

D Extra assumptions for time varying covariates

We have tested our essential assumption for difference in differences regression in figure 8 and 9 respectively in form of pre-intervention trend plot and event study regression results. Both of them establishes the parallel trend assumption for the outcome variable between treatment and control groups. However only parallel trend of outcome variable is not sufficient for a consistent estimate of two way fixed effect estimate (as we are trying to do in our study) when we have time varying covariates. Even in case of non staggered intervention, a time varying covariate can lead to a biased estimate of average treatment effect in our study. Our study has a time varying covariate in terms of bank borrowing to total asset (indicator of financial fragility) and we are looking at its effect interacted with priority sector dummy on firms probability to stay/exit the export market. Clearly, we don't want our leverage covariate to be a 'bad control' resulting in a biased estimate of our result. Sant'Anna and Zhao (2020) suggest additional assumption to be tested in order to make the difference in differences estimate unbiased. The extra assumption for time varying covariates are homogeneous treatment effect in X i.e. not a bad time varying control and no X-specific trends in both groups i.e. assume parallel trends in terms of X. We will try to test these assumptions and argue why these assumption hold in our study.

We first test parallel trend assumption for X in both groups. Our treatment group are firms that are borrowing from state-owned banks that are missing their priority sector lending target and control as firms that borrows from state-owned bank who are not missing their target. Our covariate of interest in the study is an indicator of leverage in terms of bank borrowing to total asset, hence we test parallel trend assumption for this variable and rule out any X-specific trend between treatment and control. Figure D1 shows that bank borrowing to total asset is following a somewhat parallel trend over time and the trend has not been affected by the policy change.

Figure D1: Parallel Trend Assumption for Bank Borrowing to Total Asset



Clearly, the parallel trend assumption of the covariate of our interest is established. Now we turn our attention to the homogeneous treatment effect in X assumption, this means that treatment doesn't affect influence of X. It is hard to test this assumption given our study design, we try to argue this assumption and use figure D1 to establish the claim. We can see from figure D1 that the trend of covariate of our interest doesn't change substantially post intervention implying the treatment had no effect on X. Now, in order to rule out 'bad control' argument, it is not possible for the treatment to affect the influence of X as the treatment is on the credit supply side and the one getting affected are the banks and not the firms, hence policy change is not affecting the effect of leverage on firms decision to exit/stay in the export sector. Its effect is similar to what we had before treatment and only channel through which the change is seen is through policy change. The influence of leverage on exit rate is independent of the policy change and policy change is further amplifying the survival of leveraged firms. Hence, we can rule out the 'bad control' argument and both assumption for time varying covariates hold.

E Tables for Appendix

Table A1
Regression output: with leverage ratio, liquidity ratio

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Total Asset to Capital	0.0107**			0.0104**		
	(0.005)			(0.005)		
Debt to Capital		0.0107**			0.0104**	
		(0.005)			(0.005)	
Bank Borrowing to Total Asset			-0.0049**			-0.0062**
			(0.002)			(0.003)
Cash flow in financial activity to Asset	-0.0010**	-0.0010**	-0.0010			
	(0.000)	(0.000)	(0.001)			
Total cash flow to Asset				-0.0014***	-0.0014***	-0.0004
				(0.000)	(0.000)	(0.001)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	31,230	34,301	24,709	31,550	34,614	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2
Regression output: Leverage Ratio and Public Bank dummy and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Exit Dummy						
Debt to Capital	0.0129***	0.0129***				
	(0.004)	(0.004)				
Bank Dummy*Leverage ratio	0.0134	0.0132				
	(0.020)	(0.020)				
Total Asset to Capital			0.0129***	0.0129***		
			(0.004)	(0.004)		
Bank Dummy*Leverage ratio			0.0134	0.0132		
			(0.020)	(0.020)		
Bank Borrowing to Total Asset					-0.0066*	-0.0065*
					(0.003)	(0.003)
Bank Dummy*Leverage ratio					-0.0203**	-0.0198**
					(0.01)	(0.01)
Cash Flow in Financial Activity to Asset	-0.0002		-0.0002		-0.0018	
	(0.002)		(0.002)		(0.002)	
Total Cash Flow to Asset		-0.0014		-0.0014		-0.0012
		(0.002)		(0.002)		(0.001)
At least 1 Bank Dummy	-0.0292***	-0.0296***	-0.0292***	-0.0296***	-0.0304***	-0.0305***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34,301	34,614	31,230	31,550	24,709	24,732

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7

Regression output: Leverage ratio, priority sector lending cutoffs and their interaction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: Exit Dummy								
	0-5%	0-5%	5-10%	5-10%	10-15%	10-15%	Above 15	Above 15
Debt to Capital	0.00383***		0.00403***		0.00412***		0.00394***	
	(0.001)		(0.001)		(0.001)		(0.001)	
PSL Dummy*Leverage ratio	0.0746***		-0.00623		0.00591		-0.04100	
	(0.025)		(0.005)		(0.025)		(0.035)	
Bank Borrowing to Total Asset		-0.00692**		-0.00683**		-0.00621**		-0.00677**
		(0.003)		(0.003)		(0.003)		(0.003)
PSL Dummy*Leverage ratio		-0.4431		-0.72426***		-1.04649**		-0.55362
		(0.324)		(0.248)		(0.401)		(0.403)
PSL Dummy	-0.01523***	-0.00998	0.00096	0.00662	0.00735	0.01007	-0.00540	0.00473
	(0.005)	(0.010)	(0.006)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)
Total cash flow to Asset	-0.00264***	-0.00043	-0.00266***	-0.00047	-0.00266***	-0.00047	-0.00265***	-0.00048
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry*Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,898	18,938	25,898	18,938	25,898	18,938	25,898	18,938

Note: Firm controls consists of log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** p <0.01, ** p <0.05, * p <0.1

Table A8
Regression Result: Leverage ratio, priority sector lending, their interaction for Micro, small and medium enterprises

	(1)	(2)	(3)	(4)
Dependent Variable: Never Exporter				
Debt to Capital	-0.00312	-0.00281		
	(0.009)	(0.009)		
PSL Dummy*Leverage ratio	-0.02072	-0.01701		
	(0.015)	(0.015)		
Bank Borrowing to Total Asset			0.50722***	0.49186***
			(0.170)	(0.170)
PSL Dummy*Leverage ratio			0.06622	0.19880
			(0.300)	(0.280)
PSL Dummy	-0.01717	-0.01646	-0.01838	-0.02084
	(0.013)	(0.013)	(0.016)	(0.016)
Cash flow in financial activity to Asset	-0.09522***		-0.05970***	
	(0.016)		(0.016)	
Total cash flow to Asset		-0.19438**		-0.16034***
		(0.080)		(0.055)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	15,938	16,550	12,142	12,328

Note: Firm controls consists of public bank dummy, log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A9
Regression Result: Leverage ratio, priority sector lending, their interaction for agriculture firms

	(1)	(2)	(3)	(4)
Dependent Variable: Never Exporter				
Debt to Capital	0.00170	-0.00291		
	(0.018)	(0.018)		
PSL Dummy*Leverage ratio	-0.05410***	-0.04971**		
	(0.020)	(0.020)		
Bank Borrowing to Total Asset			-0.58258**	-0.56780**
			(0.280)	(0.280)
PSL Dummy*Leverage ratio			0.66426	0.61423
			(0.600)	(0.600)
PSL Dummy	0.01912	0.00555	-0.03920	-0.03924
	(0.016)	(0.016)	(0.025)	(0.025)
Cash flow in financial activity to Asset	0.00134**		-0.04816*	
	(0.001)		(0.025)	
Total cash flow to Asset		-0.23420***		-0.15981
		(0.050)		(0.110)
Firm Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Industry*Year FE	Yes	Yes	Yes	Yes
Observations	7,078	7,185	5,925	5,988

Note: Firm controls consists of public bank dummy, log of age, log of sales, profit to sales ratio, borrowing to asset, and ownership status

Note: All standard errors are clustered at industry level

Note: Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

CAFRAL
CENTRE FOR ADVANCED FINANCIAL RESEARCH AND LEARNING
Promoted by Reserve Bank of India
C-8, 8th Floor, Reserve Bank of India, Bandra-Kurla Complex, Bandra (East), Mumbai 400051