Ten Years after the Financial Crisis:

What Have We Learned from

the Renaissance in Fiscal Research?

by

Valerie A. Ramey

University of California, San Diego and NBER

CAFRAL Annual Conference, December 10-11, 2018

Progress on All 3 Methodological Fronts

1. Theory

Incorporation of sticky prices & wages, hand-to-mouth consumers, lower bounds on policy interest rates, currency unions, variety of financing methods, effects of anticipated fiscal policy, debt sustainability.

2. Empirical methods

Identification via natural experiments, narrative methods, Bartik-style instruments; proxy SVAR/external instrument methods, local projection methods for estimating dynamic responses, standardization of methods for computing multipliers, incorporation of state dependence.

3. Data

Newly constructed historical and cross-sectional data sets within countries, narrative instruments for panels of countries, exploitation of the rich new data created by the variety of policymakers' fiscal responses to the crisis.

Outline

1. Summary of leading empirical approaches.

   A. Strengths and weaknesses.
   B. Some pitfalls in calculating multipliers.

2. Summary of estimates at this point in time.

3. Were multipliers higher in the wake of the financial crisis?

4. Conclusions.

1. Summary of Leading Empirical Approaches

A. Strengths and Weaknesses

Type	Advantages	Disadvantages
1. Aggregate country-level time series evidence	Estimates are directly informative about national-level multipliers.	- Identification of exogenous policy shocks is often challenging. - Estimates are based on historical happenstance. - Difficult to net out other fiscal changes. - It is difficult to construct counterfactuals.
2. DSGE models, estimated or calibrated	- Estimates are directly informative about national-level multipliers. - Estimates can be used to form counterfactuals.	- Identification is based on strong assumptions about the model structure and shocks processes.

A. Strengths and Weaknesses (continued)

Type	Advantages	Disadvantages
3. Subnational geographic cross-sectional or panel evidence	- Identification is often much easier and stronger – uses applied micro identification. - Many different possible data sets.	- Estimates only *relative effects*, so not directly informative about national multipliers (national factors are differenced out). - Requires additional identification assumptions to translate subnational to national multipliers, typically DSGE model for identification.
4. Individual industry, firm or household estimates of key parameters (such as MPC).	- Identification is often much easier and stronger – uses applied micro identification. - Many different possible data sets.	- Only estimates some key micro parameters, so not directly informative about national multipliers. - A DSGE model is required to translate the micro parameter estimates to national predictions about multipliers.

B. Some Pitfalls in Calculating Multipliers

- A recent lesson learned from the literature is that an important source of the wide range of multiplier estimates is due to differences in the method for calculating multipliers.

- I will highlight two commonly used methods that often lead to upward bias in multipliers.

Illustration for government purchases multipliers

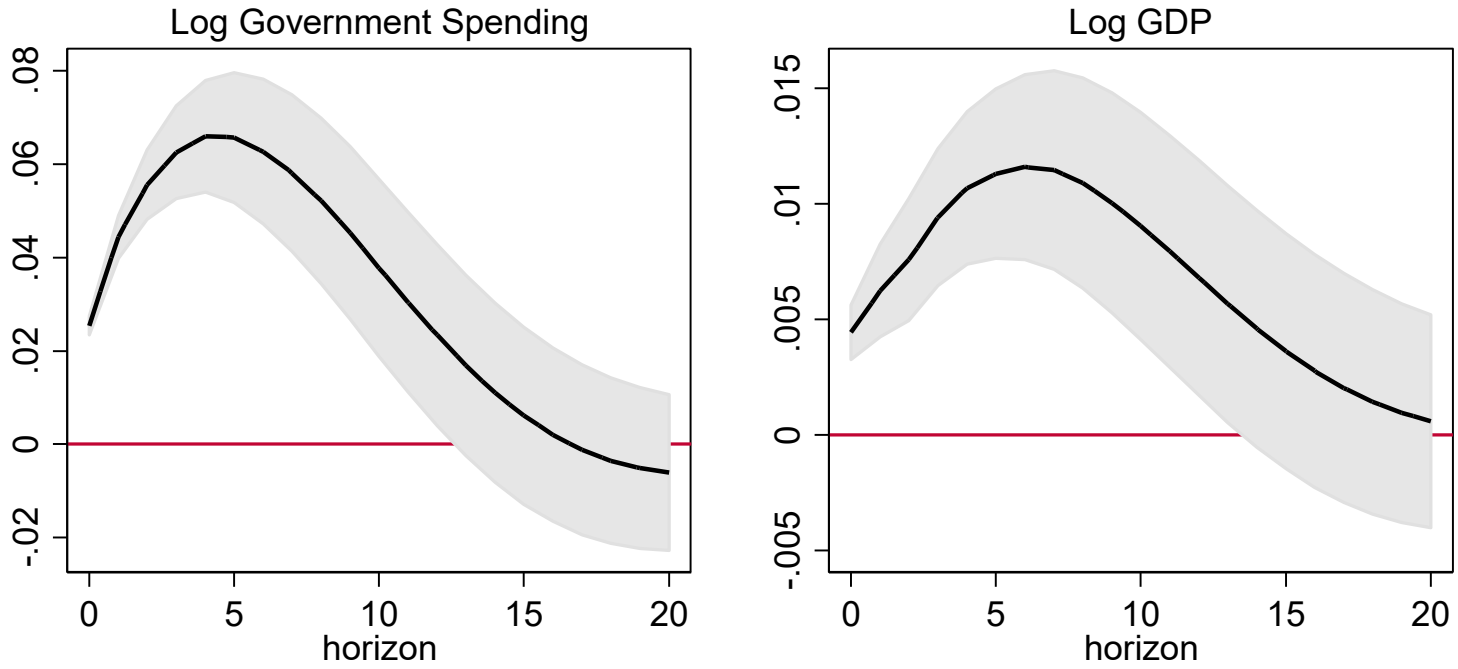
- Structural VAR (SVAR) using Blanchard-Perotti identification, which orders government spending first.

- Quarterly data from 1939:1 – 2015:4.

- 5 variables:
  - log real total government spending per capita
  - log real GDP per capita
  - log real federal tax receipts per capita
  - 3-month Treasury bill interest rate
  - inflation rate.

- 4 lags.

Estimated responses to government spending shock.



95-percent confidence bands

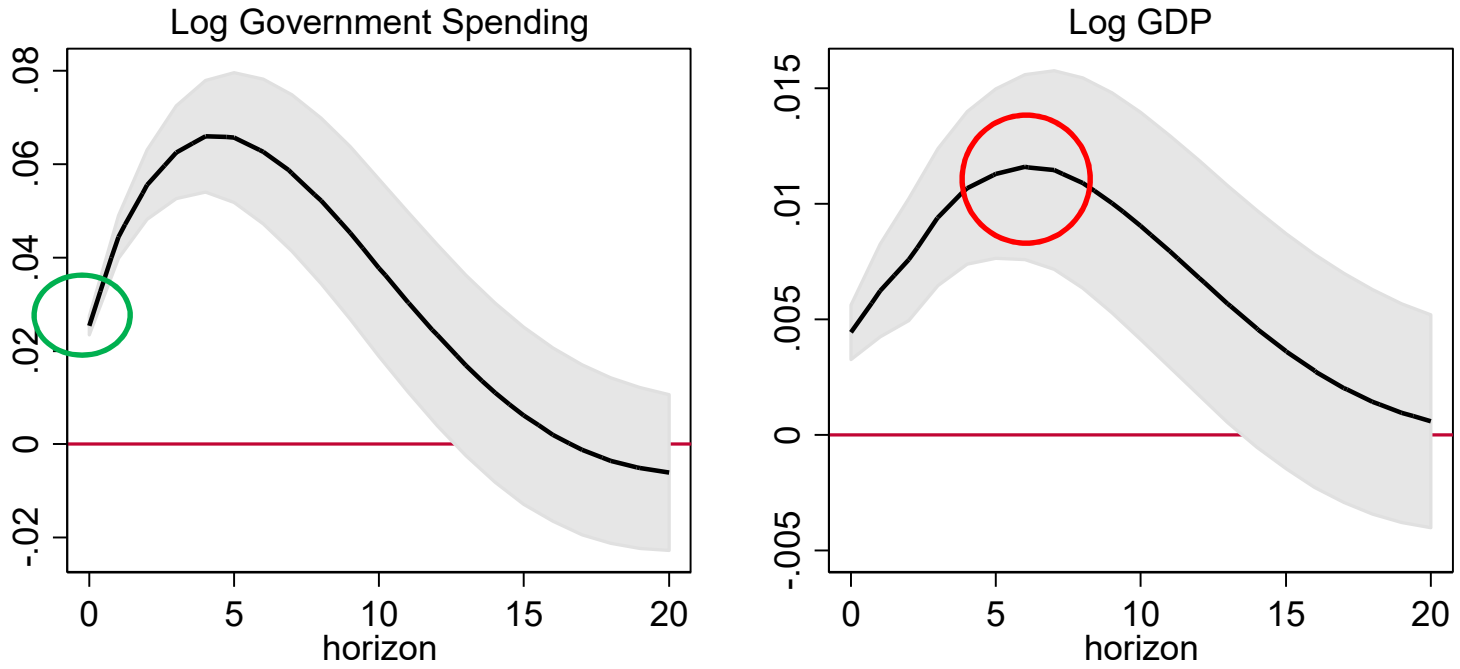
Calculating Multipliers

- How do we use the dynamic responses of the log variables to multipliers to answer the question:

  How much does GDP rise when government spending rises by $1?

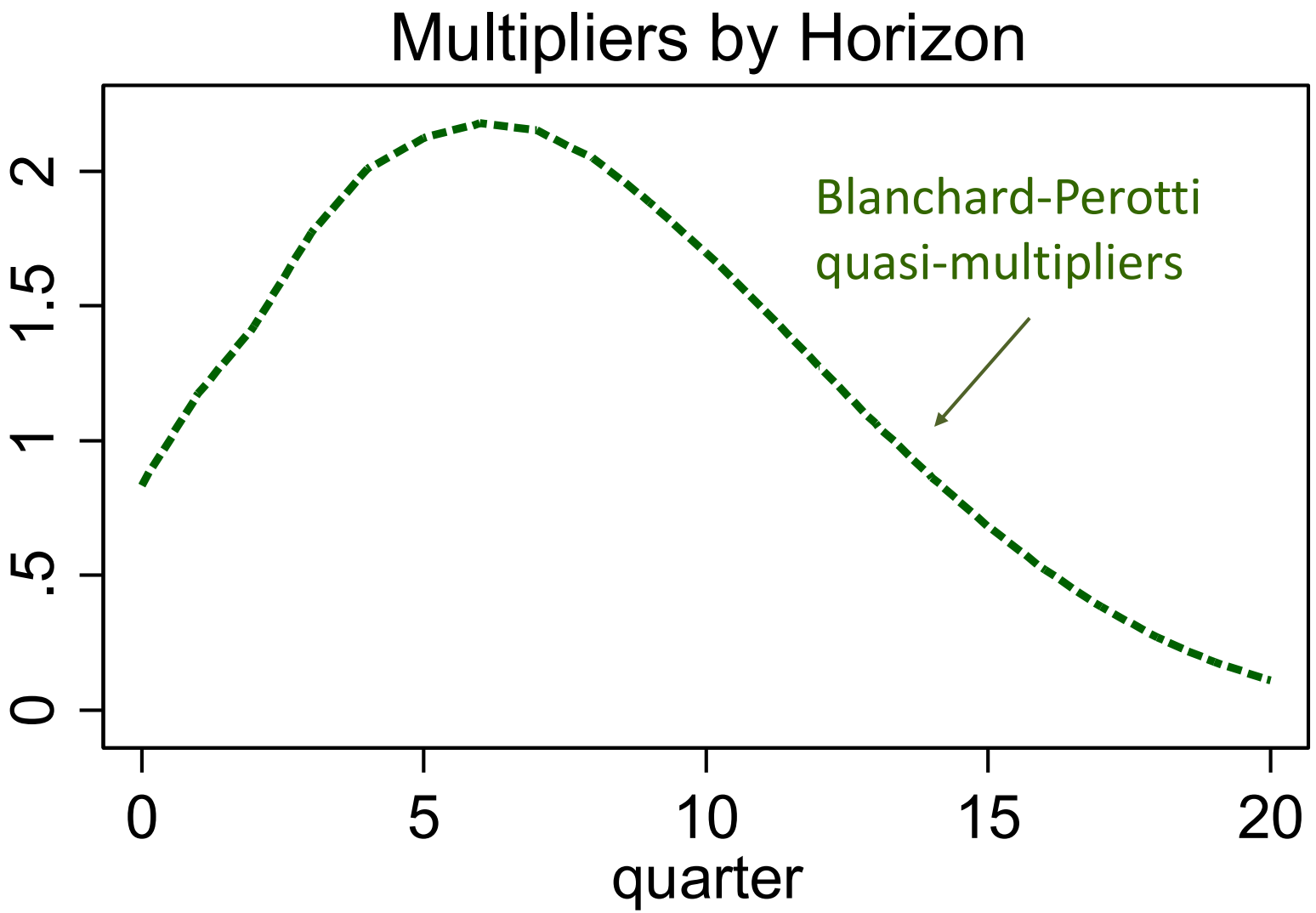
- I will show how seemingly small changes in the method can lead to large changes in the multiplier.

1. Blanchard-Perotti (2002) Quasi-Multipliers



a. Compute the ratio of the log GDP response at horizon h to the impact response (i.e. horizon 0) log government spending.

b. Convert elasticities (since logs) to $ multipliers by multiplying the ratio in (a) by the sample average GDP/Gov (4.8 in this sample).

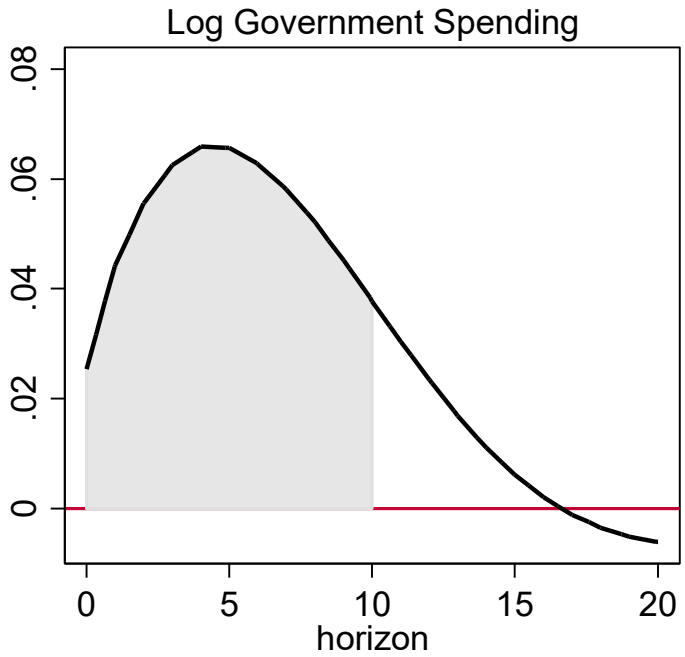


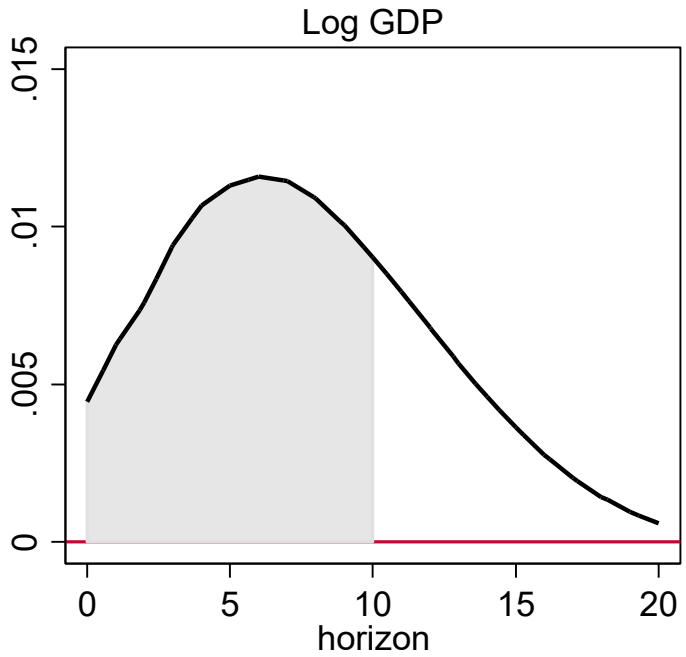
Blanchard and Perotti's quasi-multipliers are above 2 at the peak.

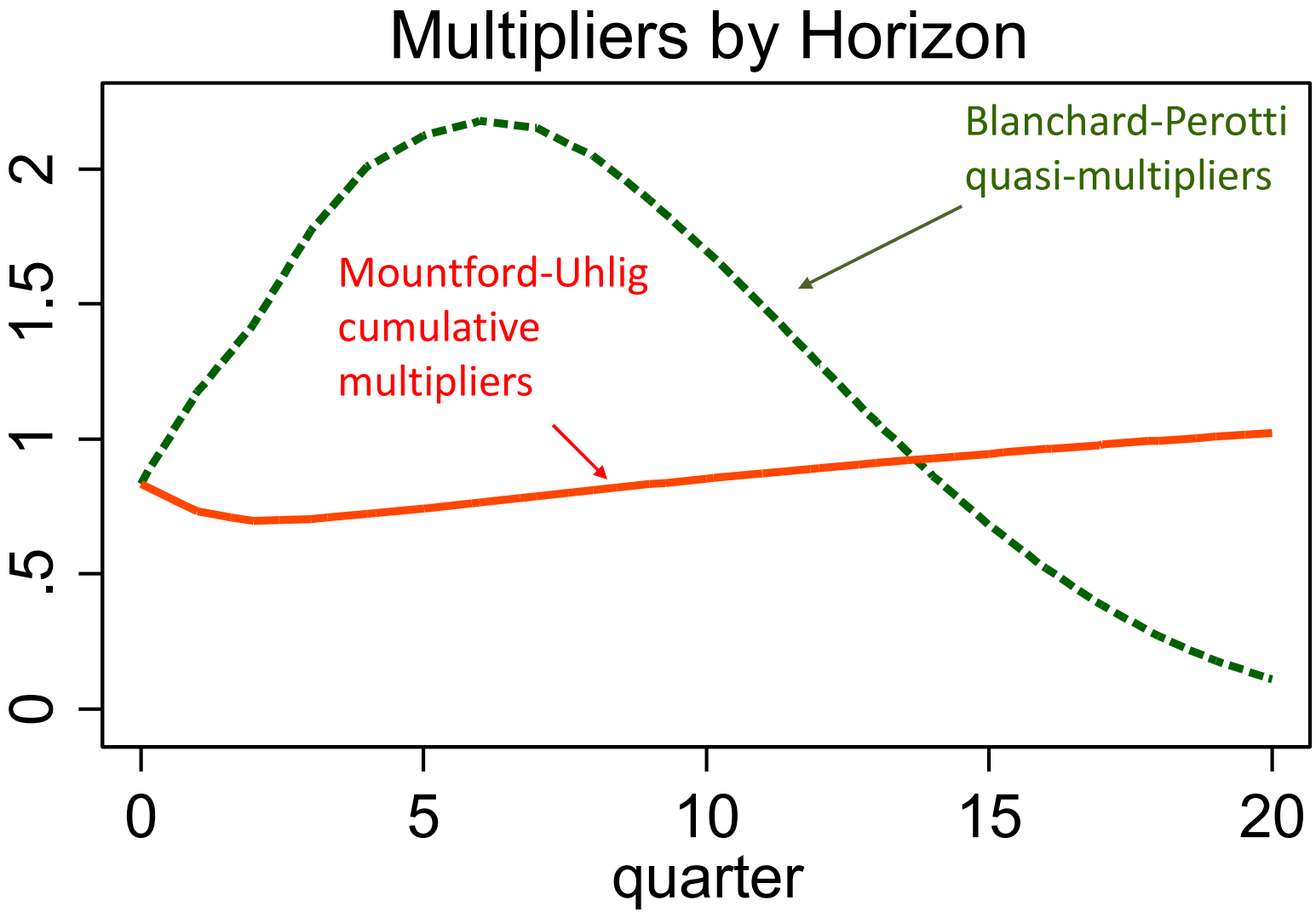
Mountford-Uhlig Method

a. Compute the ratio of the present value of the cumulative responses.

b. Convert elasticities to $ multipliers, as with previous method.






Multipliers by Horizon
Blanchard-Perotti
quasi-multipliers
Mountford-Uhlig
cumulative
multipliers
0
.5
1
1.5
2
0
5
10
15
20
quarter

Elasticities versus Multipliers

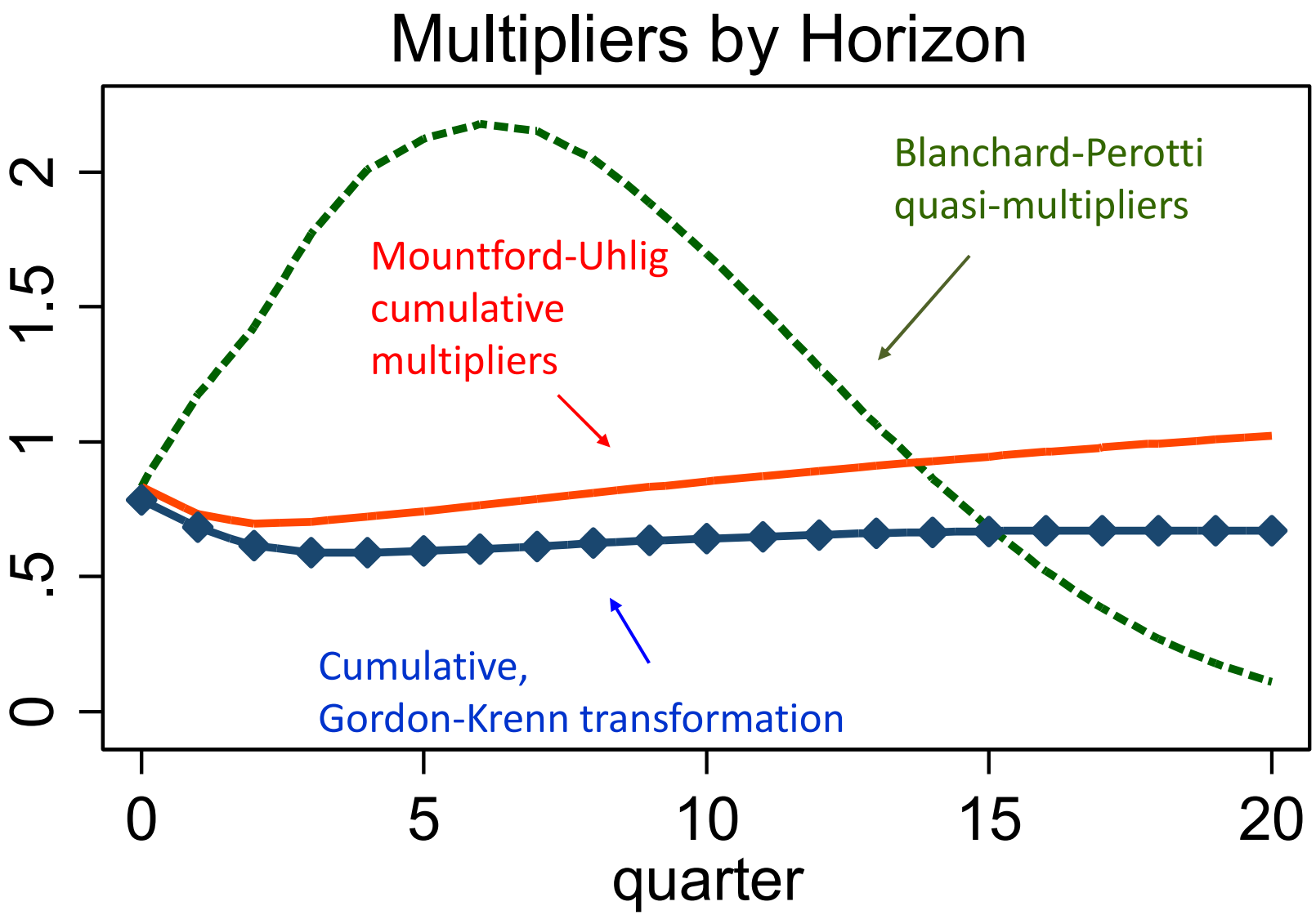
- The standard log specification yields estimates of elasticities, not multipliers.

- The standard method has been to convert elasticities to multipliers by multiplying by the sample average of Y/G.

$$\frac{dY}{dG} = \frac{dln(Y)}{dln(G)} \cdot \frac{\overline{Y}}{G}$$

- Alternative method: Transform the variables to same units before estimation. Hall, Barro-Redlick divided changes in Y and G by lagged Y.

Gordon-Krenn: $\frac{Y_t}{Y_t^P}$, $\frac{G_t}{Y_t^P}$


Multipliers by Horizon
Blanchard-Perotti
quasi-multipliers
Mountford-Uhlig
cumulative
multipliers
Cumulative,
Gordon-Krenn transformation
0
.5
1
1.5
2
0
5
10
15
20
quarter

Additional issues

- Sims and Wolff (2018) show that the conversion factor also tends to bias multipliers differentially, making them seem much higher during recessions.

- Most tax multipliers reported are based on the legislative forecast of the budget impact, not taking into account dynamic feedback. If tax cuts raise GDP, then the decline in actual revenue will be muted.

Bottom Line

- Many of the big differences in reported multipliers are not due to the estimation method, sample, etc.

- Rather, they are due to the reporting of quasi-multipliers that don't take into account the dynamic path of government spending or to the use of ad hoc conversion factors to deal with estimates based on logs.

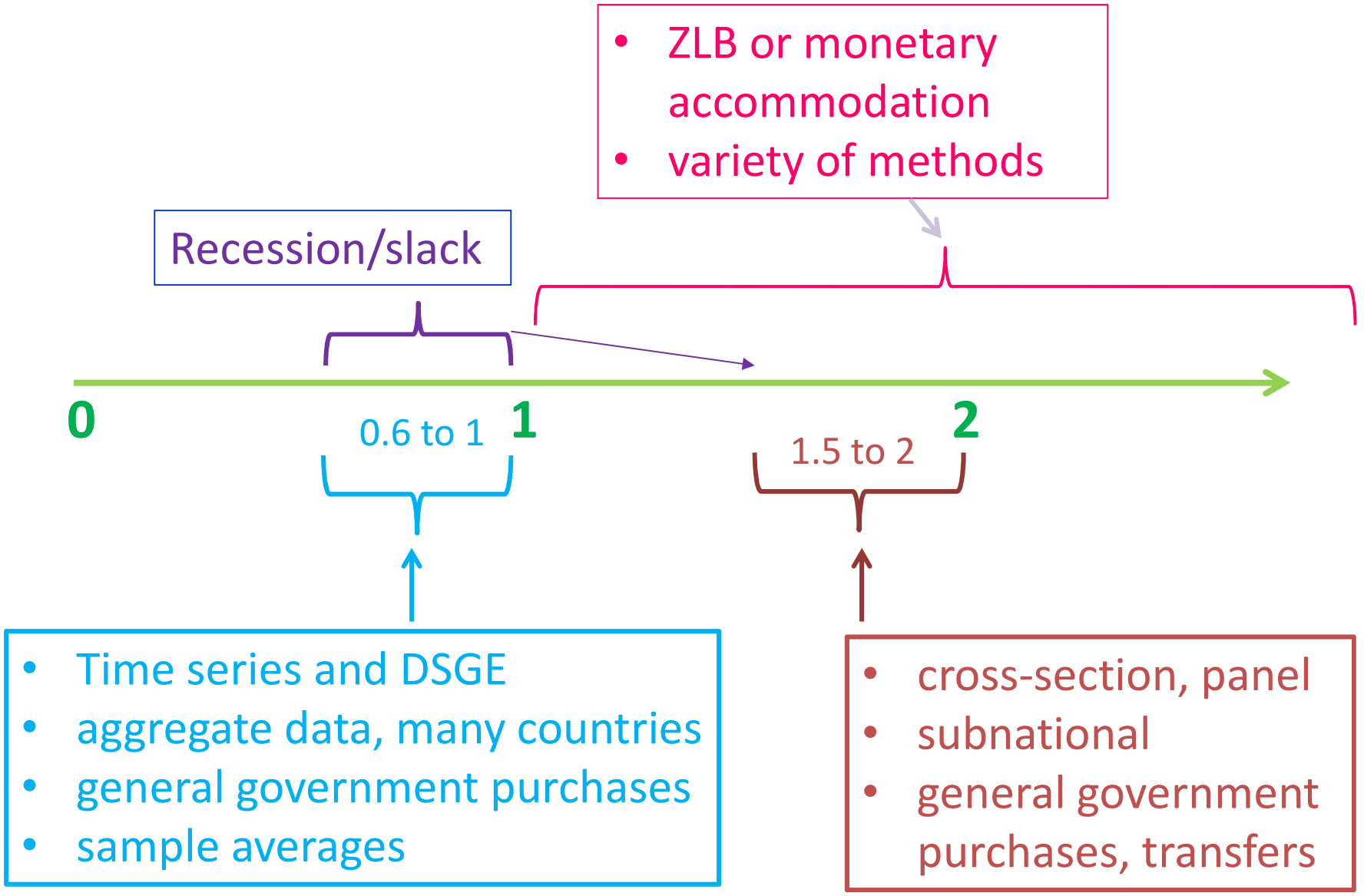
- I will highlight some other pitfalls when I discuss the cross-state estimates below.

2. Current Range of Leading Estimates of Fiscal Multipliers

Scope of summary:

- Multipliers within the first two to five years.
- Estimates based on a variety of methodologies: time series models, narratives, and estimated New Keynesian DSGE models.
- Excludes estimates that do not use the best practices.
- Ranges shown are for the majority of the estimates, but don't include some notable outliers produced with good methodology.

Multipliers on Government Purchases



Summary of Estimates Relevant for Emerging Economies

- By exchange rate regime (Iltzetzki, Mendoza, Vegh (2013))

  - Fixed exchange rates: 0.1 on impact to 1.4 in long-run
  - Flexible exchange rates: 0.1 to -0.7.

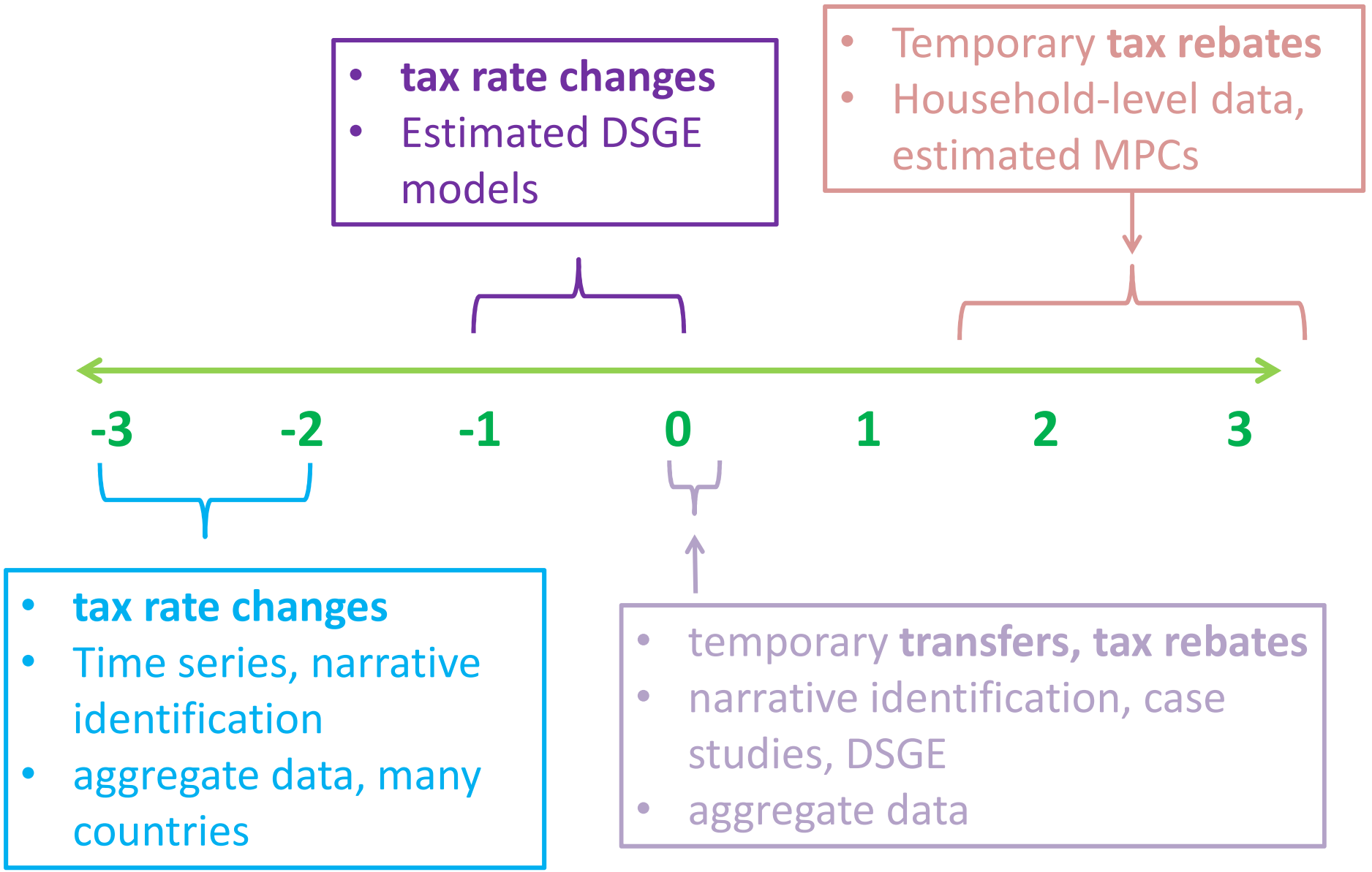
- By debt-to-GDP ratio

  - Multipliers are negative when debt-to-GDP ratio is above 60 percent.

- Infrastructure

  -Small multipliers on impact, 1.6 in longer run. (Iltzetzki et al. (2013), Leff Yaff (2018), Boehm (2018))

Multipliers on Taxes and Transfers



Summary of Aggregate Evidence

1. Most multipliers on government consumption spending estimated with linear models on national data are below 1. These apply to both the BP and defense spending estimates and DSGE estimates.

2. The tax change multipliers are estimated to be between -2.5 and -3 across numerous countries, but less than 1 in DSGE models.

3. The fiscal consolidation multipliers depend on whether the consolidation was mostly through spending or taxes, with the effects greater in magnitude for tax based consolidations.

4. Public investment multipliers could be higher: Iltzeki, Mendoza, Vegh (2013) find 1.5 to 1.6.

3. Were multipliers higher in the wake of the financial crisis?

Evidence in Favor of Higher Multipliers

- Recall that theory and some empirical evidence suggests that multipliers may be greater than one during periods of monetary accommodation, such as at the ZLB.

- The latest cross-state estimates of the impact of the ARRA suggest big employment and output multipliers.

  Chodorow-Reich (forthcoming) standardizes and synthesizes the ARRA evidence and estimates multipliers from 1.7 to 2 on output or 2 job-years created per $100K.

Chodorow-Reich's Synthesis of ARRA Multipliers

- Chodorow-Reich (forthcoming) standardizes and synthesizes the ARRA evidence and estimates multipliers from 1.7 to 2 on output or 2 job-years created per $100K.

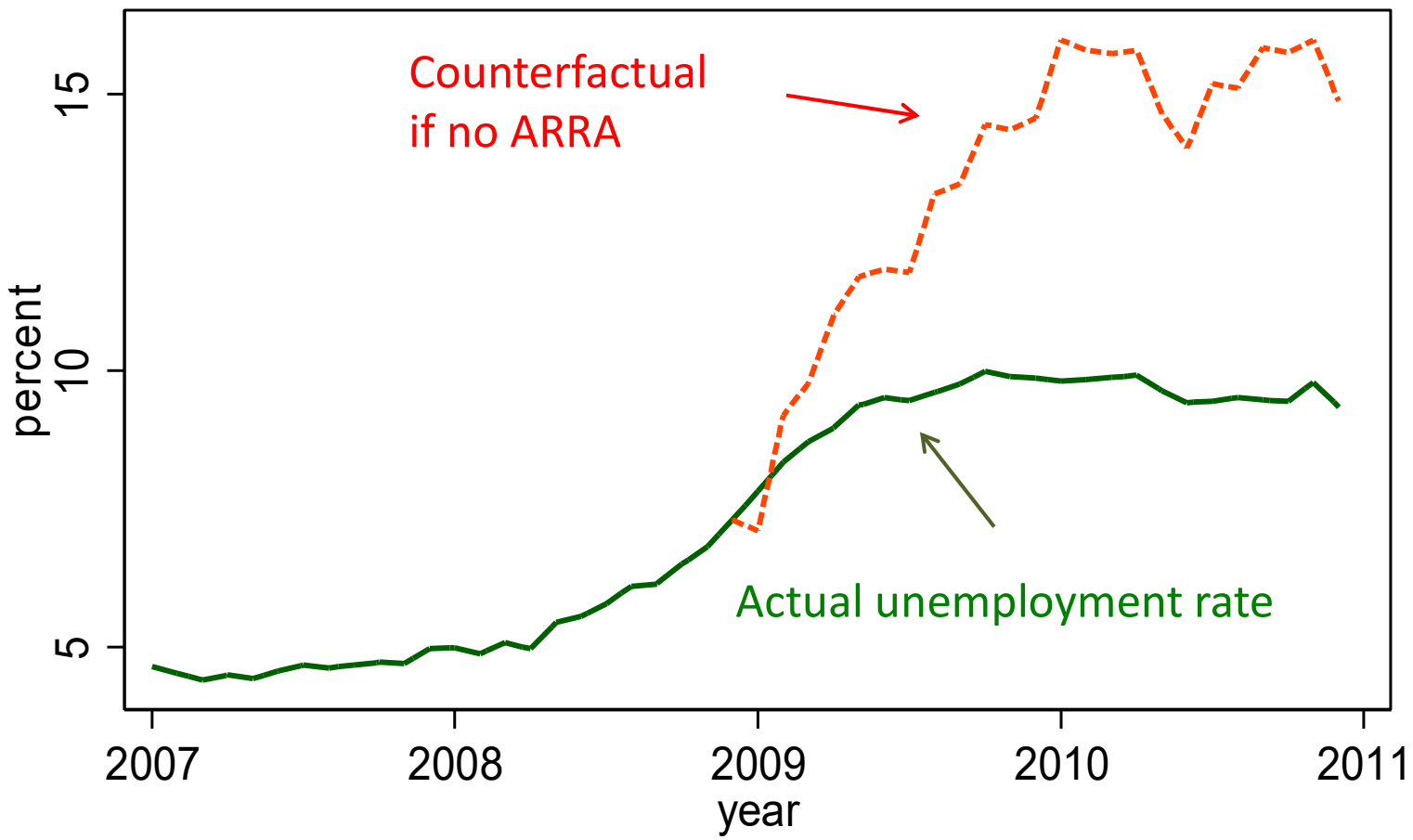
- Subnational multipliers are not the same as aggregate multipliers for a variety of reasons.

- However, Chodorow-Reich (forthcoming) uses theoretical arguments from Farhi-Werning to argue that at the ZLB, the cross-state multipliers for externally financed spending are *lower bounds* on the aggregate multiplier.

Plausibility: A Counterfactual

- Chodorow-Reich claims that his estimates are a lower bound on the national estimates.
- To check the plausibility, I use Chodorow-Reich's impulse response estimates to calculate the following counterfactual:
- What would the unemployment rate have been had there been no ARRA? This is a partial equilibrium analysis, which ignores general equilibrium effects.
- I use the stimulus shock together with his IRF estimates to calculate the implied additional employment, and then use it to find the implied unemployment rate.

Aggregate Implications of the Cross-State Multiplier Estimates



The counterfactual estimates are based on Chodorow-Reich (forthcoming) estimates of the effects of the ARRA on employment by month through December 2010.

Re-examining Chodorow-Reich's Estimates

Table 1—: ARRA Example

	Dependent variable:				
	Job years per $100K spent				GSP
	(1)	(2)	(3)	(4)	(5)
Endogenous variable:					
Total ARRA spending	2.29	2.22	1.82	2.01	1.53
	(0.71)	(1.22)	(0.69)	(0.59)	(1.19)
Instruments	*FMAP*	*DOT*	*DM*	ALL	ALL
Estimator	2sls	2sls	2sls	2sls	2sls
First stage coefficient	0.36	1.66	6.76	.	.
First stage F statistic	35.9	9.8	52.0	46.1	129.3
First stage R^2	0.40	0.23	0.55	0.73	0.87
Hansen J statistic p-value				0.76	0.34
Observations	50	50	50	50	50

Notes: The table reports cross-state regressions of the effect of ARRA spending on employment (columns 1-4) or gross state product (column 5) during 2009 and 2010. ARRA spending is instrumented using pre-recession Medicaid spending (*FMAP*), Department of Transportation formula (*DOT*), and other pre-recession forumlae (*DM*) as described in the text. All specifications also control for the employment change from December 2007 to December 2008 normalized by the December 2008 population 16+, gross state product (GSP) growth from the fourth quarter of 2007 to the fourth quarter of 2008, and the December 2008 ratio of employment to the population 16+. In columns (1)-(4) Total ARRA spending and the instruments are normalized by the December 2008 population 16+. In column (5), Total ARRA spending and the instruments are normalized by 2008Q4 GSP. Eicker-White standard errors in parentheses. Following AEA guidelines, symbolic indicators of significance are omitted.

95% confidence band: (0.85, 3.17)

Relevant instruments

Passes over-identification test

Two Distinct Questions

- The ARRA studies answer the following question:

  "How much extra employment was induced in the average state by each $1 of ARRA spending by the federal government?"

- The question relevant for national multipliers is:

  "How much extra aggregate employment was generated by each $1 of government spending induced by ARRA spending by the federal government?"

Why the Cross-State ARRA multipliers don't Aggregate

- They are not nationally representative.

  - The ARRA studies use per capita variables by state but don't weight their regressions by state population, i.e., they give North Dakota the same weight as California.

  - If treatment effects are heterogeneous across states, then the unweighted estimates won't be nationally representative.

- They don't account for all government spending.

  Much of the ARRA consisted of federal transfers to states. Several studies have found that induced state spending was more than one-for-one (e.g. Leduc and Wilson (2017)).

What happens if we correct the estimates?

- Using Chodorow-Reich's replication files, I re-estimate his model but weight each state by population and use total state and local induced spending.

- The estimates are for job-years per $100K but that is approximately equal to the output multiplier.

	Chodorow-Reich estimate	Weighted by population	All govt spending, weighted by pop
Multiplier	2.01	1.15	0.89
Robust s.e.	(0.59)	(0.72)	(0.45)

Bottom line: now the ARRA multiplier estimates look like the average historical aggregate estimates.

My conclusions about cross-state and household estimates

There is no "applied micro free lunch" for macroeconomists.

- The cross-state multipliers and household MPCs are easier to identify and estimate precisely because they are micro parameters, not macro parameters!

- There is a wide chasm between those estimates and the ones we need for macroeconomics.

- Theory can be helpful, but the assumptions of the theory are just an additional set of identifying assumptions.

Conclusions

- There really has been a renaissance in fiscal research.
- We know much more now than we did ten years ago.
- Garden-variety government purchases multipliers are probably between 0.6 and 1, though there are a few credible estimates above 1.
- Tax rate change multipliers are probably between -2 and -3.
- Multipliers on infrastructure and during periods of monetary accommodation are probably above one, possibly substantially above one, but more research should be done to assess the robustness of these results.