Empirical Macro (With Fiscal Application)

Ethan Ilzetzki

March 17, 2025

Macroeconomic Experiments

Macroeconomics poses a difficult challenge for the experimental ideal.

- Ethics/politics makes controlled experiments impossible. Even if possible
  - Compliance
  - Selection
- Natural experiments hard to come by

Sources of Variation

- Time Series
  - SVAR / Local Projections: timing/structural restrictions
    - Sims (1979) for monetary policy.
    - Blanchard and Perotti (2002)
  - Narrative Instruments
    - Barro / Ramey and Shapiro military spending shocks.
    - Romer and Romer (2010)
    - IMF (2011)
  - Hybrid (Proxy VAR)
    - Mertens and Ravn (2011, 2013)
    - Gertler and Karadi (2015)
- Cross Sectional
  - Wilson (2012)
  - Chodorow Reich et al (2012)
- Panel
  - Nakamura and Steinsson (2014)

The Experimental Ideal

- We are trying to assess the impact of a specific *treatment*.
  - A shock to the supply of money, a cut in taxes, a transition to democracy
- Some countries/regions/individuals are treated in certain times $T = 1$ and others are the *control* group $T = 0$.
- Let $y^i_{t,T}$ be the *potential* outcome of unit i at time t if his treatment status is T.
- Then the effect of treatment is $y^i_{t,1} - y^i_{t,0}$.
- Problem: we only observe $y^i_{t,1} or\ y^i_{t,0}$.

The Experimental Ideal (II)

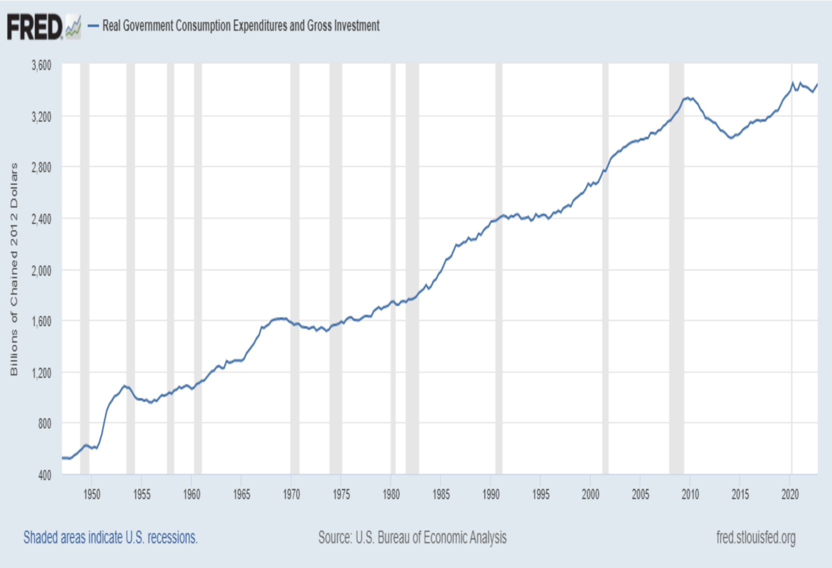
- We observe some observations with $T = 1$ and others with $T = 0$.

$$\begin{aligned}
& E\left[y^i_{t,1} | T = 1\right] - E\left[y^i_{s,0} | T = 0\right] \\
= \; & \underbrace{E\left[y^i_{t,1} - y^i_{t,0} | T = 1\right]}\text{Effect of treatment on the treated} \\
& + \underbrace{E\left[y^i_{t,0} | T = 1\right] - E\left[y^i_{s,0} | T = 0\right]}\text{Selection Bias}
\end{aligned}$$

- The great benefit of a randomized controled trial is that treatment status is independent of potential outcomes, so that $E\left[y^i_{t,0} | T = 1\right] - E\left[y^i_{s,0} | T = 0\right] = 0$.
- IV, diff-in-diff, regression discontinuity design, ect. are all attempts at creating quasi-random assingment.
- The main challenge is arguing that selection bias is zero.

Time Series

G and GDP in the Time Series



SVAR

- Write our system of endogenous variables as as Vector Autoregression

$$AY_t = \sum_{k=1}^{j} C_k Y_{t-k} + Bu_t,$$

where $Y_t = (y_t, g_t ...)'$.

- A, B, C: $N \times N$ matrices.
- u_t: $N \times 1$ vector of $N(0,1)$ orthogonal "structural" shocks.
- A gives contemporaneous relationships between variables in Y_t

SVAR (2)

Estimating VAR using OLS

$$Y_t = \sum_{k=1}^{K} A^{-1} C_k Y_{t-k} + A^{-1} B u_t.$$

we obtain estimates of $A^{-1}C_k$ and of

$$\varepsilon_t = A^{-1} B u_t.$$

$A^{-1}C_k$ tells us how past changes in Y_t affect current Y_t.

Can iterate forward any shock from u_t to see its effects on Y_{t+h}

But so far can't identify A

SVAR (3)

Let Σ be the variance-covariance matrix of ε_t:

$$\begin{aligned}\Sigma &= E\left[\varepsilon_t \varepsilon_t'\right] = E\left[A^{-1}Bu_t\left(A^{-1}Bu_t\right)'\right] \\ &= E\left[A^{-1}Bu_t u_t' B'\left(A^{-1}\right)'\right] \\ &= A^{-1}BB'\left(A^{-1}\right)'\end{aligned}$$

We can estimate Σ using the variance covariance matrix of $\hat{\varepsilon}_t$.

As Σ is symmetric, we have $\frac{N(N+1)}{2}$ equations.

We can use them to estimate (at most) $\frac{N(N+1)}{2}$ unknowns in A and B.

There is a total of N^2 unknowns in A and B.

We need $N^2 - \frac{N(N+1)}{2} = \frac{N(N-1)}{2}$ restrictions (identifying assumptions) on the elements of A and B.

SVAR (4)

3x3 example

Have $\frac{N(N+1)}{2} = 6$ equations to estimate $N^2 = 9$ parameters.

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} \\ a_{21} & 1 & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}.$$

Use theory to restrict 3 additional parameters.

Cholesky ordering assumes A is lower diagonal, so that

$$A = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix}$$

Using SVAR to identify fiscal shocks

E.g. let $Y_t = (y_t, g_t, T_t)'$

Use institutional data (tax code) to estimate the automatic responses of g_t and τ_t to y_t

- $t_t = a_{ty} y_t$. Get a_{ty} from micro data and information about the tax code.
- $g_t = a_{gy} y_t$ (typically $a_{gy} = 0$).

Then

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} \\ a_{gy} & 1 & a_{23} \\ a_{ty} & a_{32} & 1 \end{bmatrix}.$$

Still need an additional assumption. Make some assumption about how g_t and τ_t react to each other (specifically which moves first).

Blanchard and Perotti (2002); Perotti (2004)

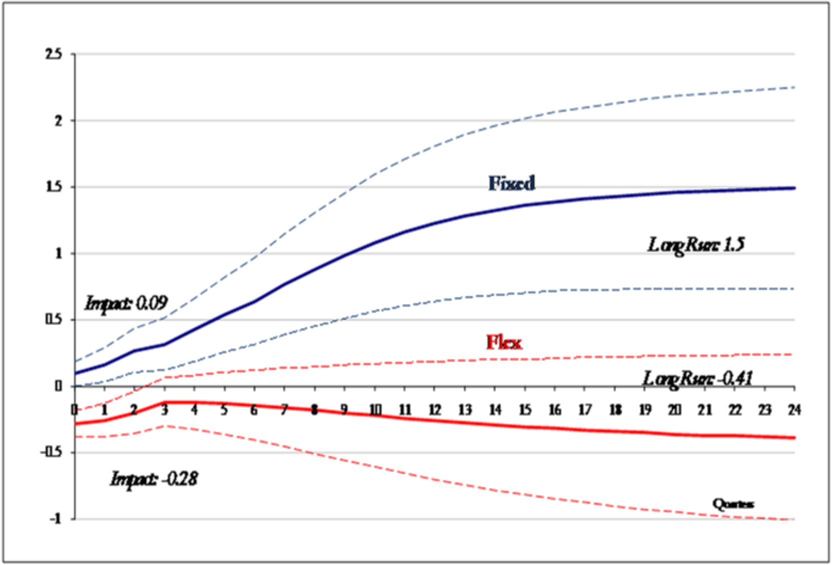
Shocks to government expenditure cause

- an increase in output (with multiplier close to 1)
- an increase in consumption
- an increase in wages and hours worked.

Shocks to tax rates cause

- a decrease in output with multiplier ˜0.75.
  - More recent estimates much closer to zero.

Ilzetzki, Mendoza and Vegh (2013)



Critiques of SVAR

Timing assumptions arbitrary.

Fundamental shocks unknown to econometrician, but may predicted by agents.

The VAR assumption is restrictive; bias due to model mis-specification.

Local Projections

Jordá (2005): Local Projections: estimate impulse response directly

- Run a separate regression for each impulse response horizon h:

$$Y_{t+h} = \sum_{k=1}^{j} \phi^h Y_{t-k} + \nu_t.$$

The n$^{\text{th}}$ row of ϕ^h gives the response of all variables in Y to a change in its n$^{\text{th}}$ element (variable), h periods ahead.

LP is flexible

- If we have an external instrument for a variable in Y, can instrument directly.
- Can use the "identified" residual from a structural VAR as an instrument here too.
- Can interact Y with a variable to estimate a dif in dif.

Local Projections vs. VARs

Local projections often give “unsmooth” impulse responses with large(er) standard errors.

VARs have (larger) bais because of model mis-specification.

Plagborg-Møller and Wolf (2021): VAR(K) and LP give same impulse responses as $K \to \infty$

22

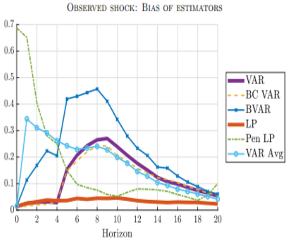


Figure 4: Median (across DGPs) of absolute bias of the different estimation procedures, relative to $\sqrt{\frac{1}{21}\sum_{h=0}^{20}\theta_h^2}$.

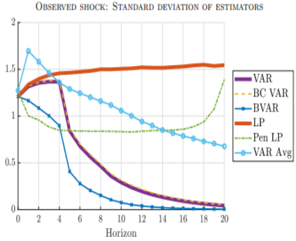
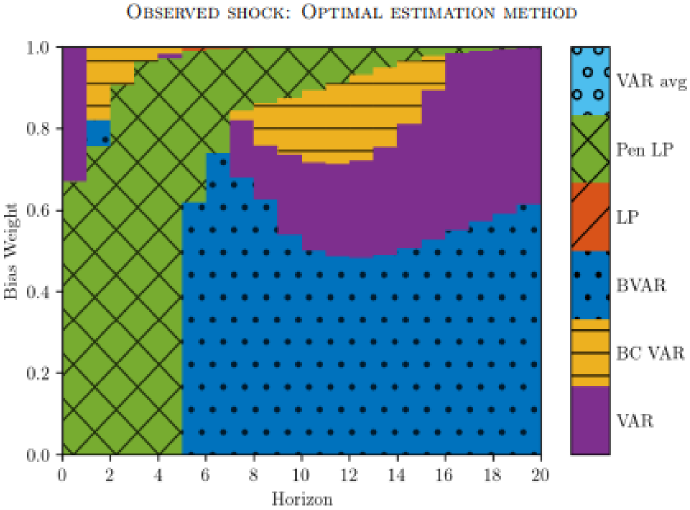


Figure 5: Median (across DGPs) of standard deviation of the different estimation procedures, relative to $\sqrt{\frac{1}{21}\sum_{h=0}^{20}\theta_h^2}$.

Bias vs. Precision in LP vs. VAR



Source: Li, Plagborg-Moller, Wolf (2022)

Local Projections vs. VARs

Local projections often give "unsmooth" impulse responses with large(er) standard errors.

VARs have (larger) bais because of model mis-specification.

Plagborg-Møller and Wolf (2021): VAR(K) and LP give same impulse responses as $K \to \infty$

22

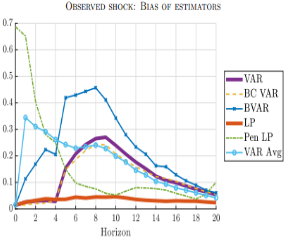


Figure 4: Median (across DGPs) of absolute bias of the different estimation procedures, relative to $\sqrt{\frac{1}{21}\sum_{h=0}^{20}\theta_h^2}$.

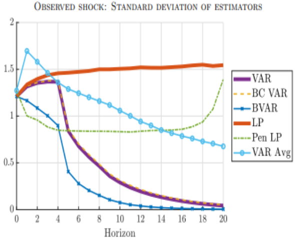
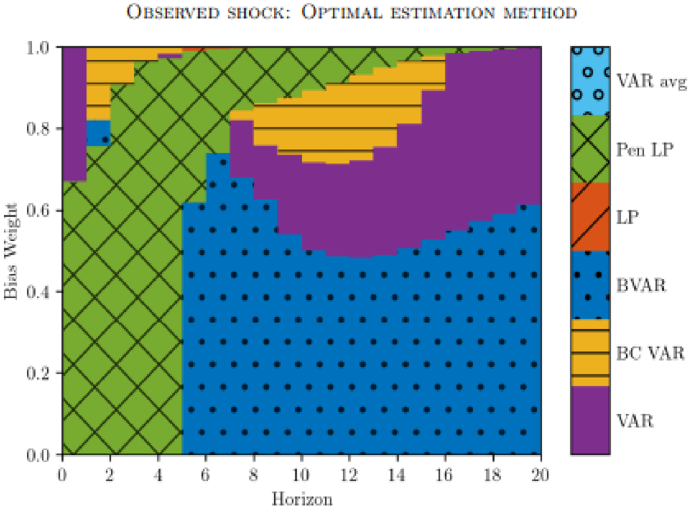


Figure 5: Median (across DGPs) of standard deviation of the different estimation procedures, relative to $\sqrt{\frac{1}{21}\sum_{h=0}^{20}\theta_h^2}$.

Bias Variance Tradeoff in LP and VARs



Source: Li, Plagborg-Moller, Wolf (2022)

Coverage Ratios of LP and VARs

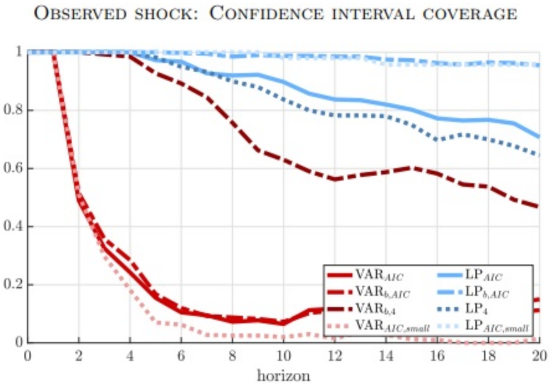


Figure 6.2: Fraction of DGPs (both stationary and non-stationary) for which the confidence interval coverage probability is above 80 per cent, by estimation procedure; the b subscript in the figure legend indicates bootstrap confidence intervals. The subscript "AIC" indicates lag length selection via the AIC, "4" indicates four lags, and "small" indicates a small system, containing only shock and outcome of interest.

Source: Montiel Olea, Plagborg-Moller, Qian, Wolf (2025)

Fiscal Natural Experiments

Biggest shocks to US public spending: Military

Barro: Use military expenditure as instrument for total government expenditure.

- Better yet: use news of military buildup (Ramey and Shapiro).
- Identify large scale military buildups (major US wars + Reagan expansion)
- Use reporting in Business Week magazine to identify timing of expectations of military buildups.

Argument: US wars fought on foreign soil, no direct impact on US GDP.

Ramey (2009)

Shocks to government expenditure cause

- an increase in output (with multiplier close to 1)
- a *decrease* in consumption
- an *increase* in hours worked, but a *decrease* in wages

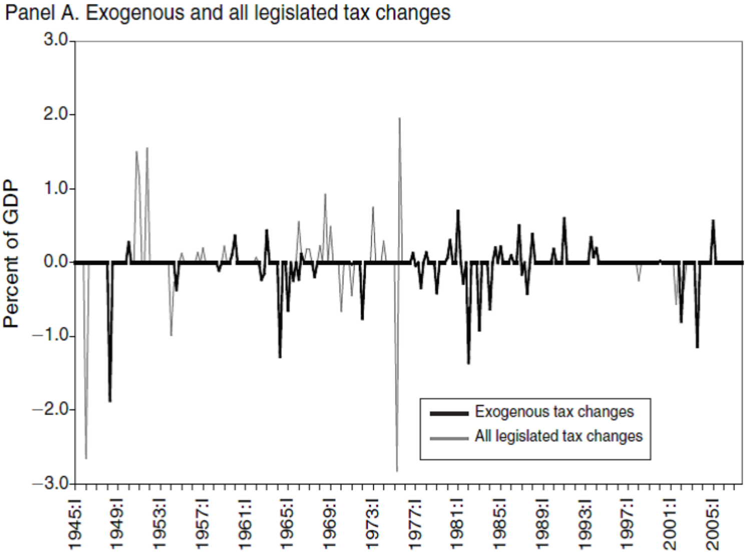
Critiques of Military Natural Experiments

Exclusion restriction?

External validity?

Two observations?

Narrative Approach: Romer and Romer



Classifying Tax Changes

Countercyclical tax change

Government expenditure related tax changes

Deficit-reducing tax changes

Tax changes for long-run efficiency

Romer and Romer (2010)

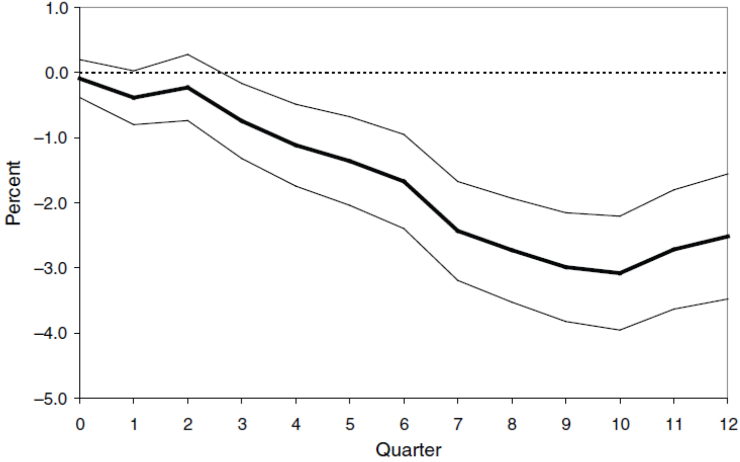


FIGURE 4. ESTIMATED IMPACT OF AN EXOGENOUS TAX INCREASE OF 1 PERCENT OF GDP ON GDP
(*Single equation, no controls*)

Reconciling SVAR and Narrative Evidence

Mertens and Ravn (2011) and Stock and Watson (2018) propose an SVAR identification that does not rely on "arbitrary" assumptions: Proxy VAR

Use time series of narrative evidence (e.g. Romer and Romer, 2010) as in instrument for *innovations* in fiscal policy.

- Regress the Romer and Romer (2012) series on ε_t^T.

Weaker assumption than Romer and Romer

- Their series needs to be uncorrelated to unforecastable shocks to y_t (u_t^y) not to y_t.
- Their series needs to be correlated with u_t^T, but not perfectly so.

Other application: Gertler and Karadi (2015) use high frequency movements in monetary policy as proxies in a monetary VAR

- See Miranda-Agrippino nad Ricco (2021) for critique and extension.

Problem with Proxy VARs

Instrumenting one component of u_t with another gives causal identification of a component of A matrix

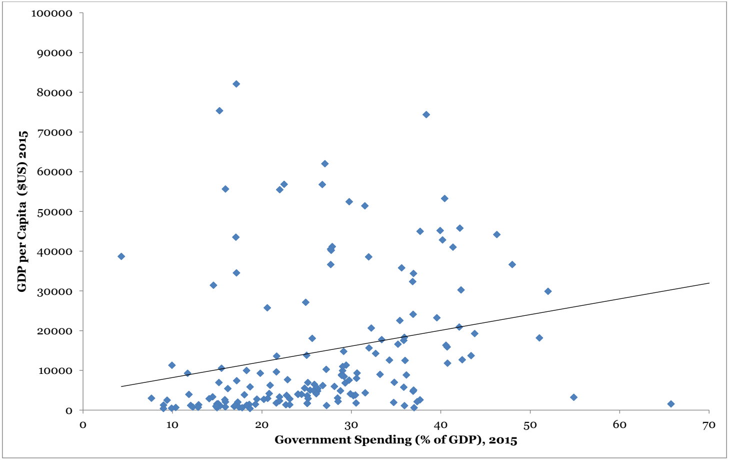
But C_k matrices estimated with OLS regression.

Solution: **internal instrument** rather than proxy

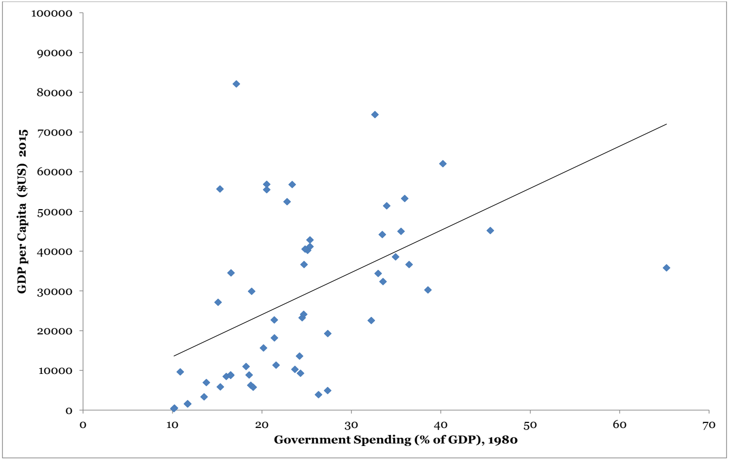
- Include instrument as an element of Y_t
- Order instrument first in the Cholesky decomposition.

Cross Section

G and GDP in the Cross Section



G (lagged) and GDP in the Cross Section



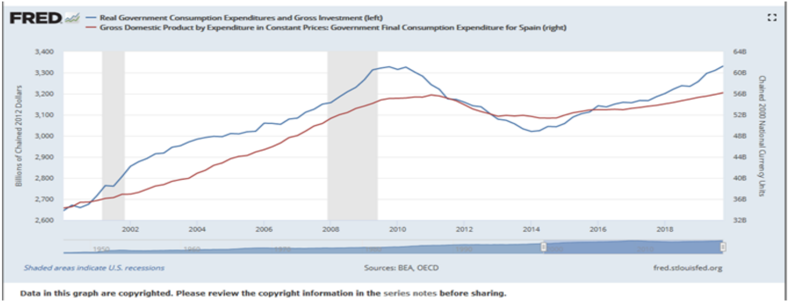
Fixed Effects

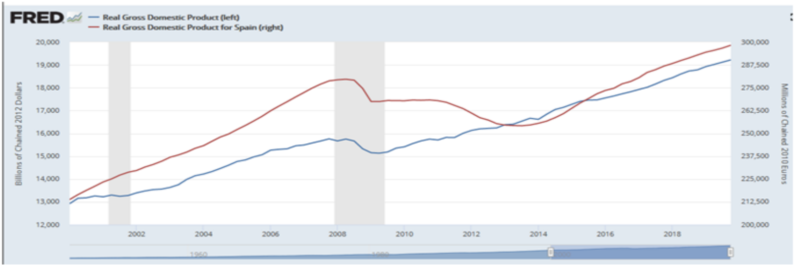
- Panel data allows the use of time and cross sectional (country/region) fixed effects:

$$Y_{it} = \alpha_i + \alpha_t + G_{it} + \varepsilon_{it}$$

  - Time fixed effects ($alpha_t$) capture all common variation over time.
  - Cross-sectional fixed effects ($alpha_i$) capture all time-invariant country-specific variation.
  - Leaves time*country variation.
  - Remaining concerns: *differential* sources of variation in G between countries across time that are correlated with other drivers of GDP.

G and GDP in Differences across Time in Two Countries





Shift Share (Bartik) Instruments

Interact the time-invariant expected intensity of treatment z_i...

with the total treatment over time z_t...

to create an instrument $S_{it} = z_i z_t$

Key example: Bartik interacts the initial share of employment in an industry in a State z_i with the national employment growth in that State z_t to see how industry decline affects employment.

Panel Estimates of Fiscal Multipliers

Nakamura and Steinsson (2012)

- Shift Share instrument
- Interact the size of military industry across states (z_i) with total military spending across time (z_t).
- Find multipliers > 1

Kraay (2010)

- World Bank loan disbursement schedule determined in advance
- Certain types of World Bank loans allowed for cyclical reasons, other for long-term reasons.

Ilzetzki (2024)

- (Leave one out) national purchases of aircraft of a given type in US during WWII as instrument for aircraft purchases from each plant.
- Argument: Demand for aircraft from specific plant endogenous, but for aircraft types driven by military strategy.

Shift-share critiques

Goldsmith-Pinkham, Sorkin, Swift (2020): Think of shift-share as an average "randomly assigned" *shares* weighted by the *shocks*

Borusyak, Jaravel Hull (2020): Think of shift-share as an average of "randomly assigned" *shocks* weighted by the *shares*

Adão, Kolesár, Morales (2019): Correlation in residuals across cross-sectional observations with similar "shares" leads to mis-specified standard errors.