Semiparametric Estimation of Monetary Policy Effects: String Theory Revisited

Joshua D. Angrist[⋆] Òscar Jordà[†] Guido Kuersteiner[§]

[⋆]MIT and NBER
[†]Federal Reserve Bank of San Francisco; University of California, Davis
[§]University of Maryland

Disclaimer:
The views expressed herein are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.

September 18, 2014

Intellectual journey

Rob Engle: ACD to monetary policy interventions

Jim Hamilton: ACH + OP + impulse responses as a difference in forecasts

Clive Granger: Direct forecasting

...
Local projections: RA

...
Local projections: RA + IV

...
Local projections: IPW + RA

Taking local projections a step further

What is the causal effect of monetary policy on economic outcomes?

- Re-randomization through inverse propensity score weighting
- Parametric model of the p-score, but conditional response model of the outcome unspecified (fully flexible)

Motivation

Governor Eccles: [...] one cannot push on a string. We are in the depths of a depression and ... beyond creating an easy money situation through reduction of discount rates and through the creation of excess reserves, there is very little if anything that the reserve organization [Federal Reserve Board] can do toward bringing about recovery. I believe that in a condition of great business activity that is developing to a point of credit inflation, monetary action can very effectively curb undue expansion.

- Testimony before the House Committee of Banking and Currency. March 18, 1935

Two important questions

Is the Fed’s accelerator as effective its brakes?

Answer: NO

Did the Great Recession *substantially* modify the Fed’s ability to support the economy before the era of unconventional monetary policy began?

Answer: NO

Setup

$\chi_t = (D_t, y_t, x_t)$ where D_t is the policy variable, y_t is the outcome and x_t are additional inputs to the policy function

Policy function: $D_t = D(z_t, \psi, \epsilon_t)$

z_t is a subset of χ_t that includes lags χ_{t-j}

Connection to VARs

Stylized structural VAR:

$$\begin{pmatrix} y_t \\ D_t \end{pmatrix} = \begin{pmatrix} a_{yy} & a_{yd} \\ a_{dy} & a_{dd} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ D_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ c_{dy} & 1 \end{pmatrix} \begin{pmatrix} u_t \\ \epsilon_t \end{pmatrix}$$

Policy coefficients: $\psi = (a_{dy}, a_{dd}, c_{dy})'$

Variables in the policy equation: $z_t = (y_{t-1}, D_{t-1})'$

Impulse response:

$$R(y^{\psi}_{t,1}, d_j - d_0) = a_{yd}(d_j - d_0)$$

Potential outcomes

Potential outcomes $y^{\psi}_{t,l}(d_j)$ are y_{t+l} if $D_t = D(.) = d_j$ for all $\psi \in \Psi$ and all possible realizations d_j of $\mathcal{D}$

The causal effect of a policy intervention is defined as:

$$y^{\psi}_{t,l}(d_j) - y^{\psi}_{t,l}(d_0)$$

An unobservable random variable. Can calculate interesting moments, e.g.,

Average Treatment Effect (ATE): $\Lambda^l_j = E(y^{\psi}_{t,l}(d_j) - y^{\psi}_{t,l}(d_0))$

Randomized policy experiments

Suppose $D_t \in \{0, 1\}$

Average effect of policy intervention, ATE:

$$\hat{\Lambda}_{GM} = \frac{1}{n_1} \sum_t D_t (y_{t+l} - y_t) - \frac{1}{n_0} \sum_t (1 - D_t)(y_{t+l} - y_t)$$

In regression form:

$$y_{t+l} - y_t = \alpha_0 D_t + \alpha_1 (1 - D_t) + v_{t+l}$$

$$\hat{\Lambda}_{GM} = \hat{\alpha}_1 - \hat{\alpha}_0$$

Observational data

ATE:

$$
\begin{aligned}
E\left[(y_{t,l}(1) - y_t) - (y_{t,l}(0) - y_t)\right] = \\
E[E\left[y_{t,l}(1) - y_t | D_t = 1; \chi_t\right] - \\
E\left[y_{t,l}(0) - y_t | D_t = 0; \chi_t\right] = \\
\Lambda^l \quad \text{for all } l \geq 0
\end{aligned}
$$

For example, use regression control (local projections):

$$
y_{t+l} - y_t = D_t \alpha_1^l + \alpha_0^l (1 - D_t) + D_t X_t \beta_1^l + (1 - D_t) X_t \beta_0^l + v_{t+l}
$$

$$
\hat{\Lambda}^l = (\hat{\alpha}_1^l - \hat{\alpha}_0^l) + \underbrace{(\bar{X}_t^1 \hat{\beta}_1^l - \bar{X}_t^0 \hat{\beta}_0^l)}_{=0?}
$$

Identification assumptions

Selection on observables:

$$y_{t,l}^{\psi}(d_j) \perp D_t | z_t, \psi \quad \forall l \geq 0 \text{ and } \forall d_j \text{ and } \forall \psi \in \Psi$$

Overlap: The propensity score

$$P(D_t = d_j | z_t) = p_t^{d_j}(z_t, \psi)$$

is such that

$$0 < p_t^{d_j} < 1 \quad \forall z_t$$

Inverse propensity score weighting

Selection on observables means

$$E(y_{t,l}\mathbf{1}\{D_t = d_j\}|z_t) = E(y_{t,l}(d_j)|z_t)p_t^{d_j}(z_t, \psi)$$

hence

$$E(y_{t,l}(d_j)|z_t) = \frac{E(y_{t,l}\mathbf{1}\{D_t = d_j\}|z_t)}{p_t^{d_j}(z_t, \psi)}$$

Therefore

$$\Lambda_j^l = \frac{E(y_{t,l}\mathbf{1}\{D_t = d_j\}|z_t)}{p_t^{d_j}(z_t, \psi)} - \frac{E(y_{t,l}\mathbf{1}\{D_t = d_0\}|z_t)}{p_t^{d_0}(z_t, \psi)}$$

Group means + IPW

Suppose treatment $D_t \in \{0, 1\}$. The analogous estimator to Hirano, Imbens and Ridder (2003) is:

1. estimate the propensity score (e.g. probit): $\hat{p}_t \equiv p(D_t = 1 | z_t, \hat{\psi})$
2. define weights $w_t = \frac{D_t}{\hat{p}_t} + \frac{(1-D_t)}{(1-\hat{p}_t)}$
3. estimate ATE using IPW based on the group means estimator and the weights

$$\hat{\Lambda}^l_{IPW} = \frac{\sum_t w_t D_t (y_{t+l} - y_t)}{\sum_t w_t D_t} - \frac{\sum_t w_t (1 - D_t)(y_{t+l} - y_t)}{\sum_t w_t (1 - D_t)}$$

Remarks

p-score: more weight on treated/untreated obs not predicted to be treated/untreated

Overlap critical. Two polar cases:

- perfect sorting $\rightarrow$ IPW returns the original regression
- random sorting $\rightarrow$ original regression as good as randomly assigned

Covariates enter through p-score.

No need to specify the conditional mean model for the outcome

Estimation

Outcome vector $\equiv Y_t$

$$h_{t,j}(\psi) = E\left[Y_t\left(\frac{\mathbf{1}\{D_t = d_j\}}{p_t^{d_j}(z_t,\psi)} - \frac{\mathbf{1}\{D_t = d_0\}}{p_t^{d_0}(z_t,\psi)}\right)\right]$$

Let $h_t(\psi) = (h'_{t,1}, ..., h'_{t,J})'$ and $\hat{h}_t = h_t(\hat{\psi})$

$$\arg\min_{\Lambda}\left(\frac{1}{T}\sum_{t=1}^{T}\hat{h}_t - \Lambda\right)'\Omega^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\hat{h}_t - \Lambda\right)$$

In cross-section: Hahn (1998), Hirano, Imbens, Ridder (2003), Cattaneo (2010)

Causal effect of monetary policy

New features of the analysis with IPW

Model probability of intervention rather than identifying the "shock"

Use mixed-frequencies: financial data (daily), macro controls (monthly)

Responses tailored to the type of intervention (+0.25 v. -0.25 relative to no change benchmark)

No joint model for the outcomes and the interventions $\rightarrow$ flexibility

Causal effect of monetary policy

Propensity score specification

Policy intervention: fed funds target adjustments grouped into $\{-0.5, -0.25, 0, 0.25, 0.5\}$ percent changes

Determinants/predictors of policy:

- fed funds futures contracts + Hamilton (2008) + adjustments (since 1989)
- inflation (PCEPI), unemployment rate (two lags)
- target level, last target change, LTC in an FOMC meeting, FOMC meeting

Two samples: pre-crisis 1989–2005; full 1989–2008

Propensity score specification passes overlap and specification tests

Term structure responses

Pre-crisis sample: 1989-2005

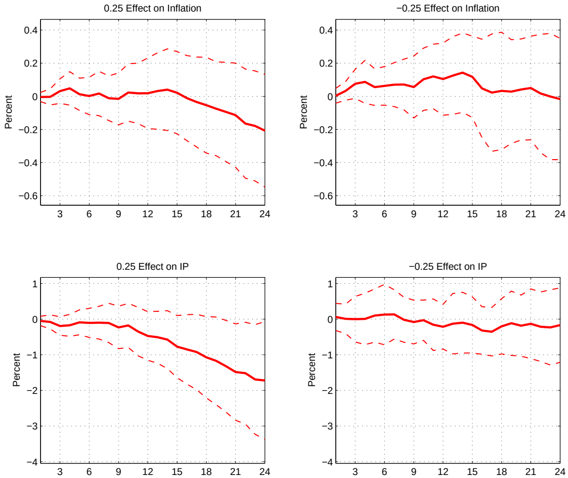
	+0.25				-0.25			
Months	6	12	18	24	6	12	18	24
FFR	0.4**	0.8**	0.7*	0.4	-0.2**	-0.3	-0.4	-0.3
3-m	0.4**	0.7**	0.5	0.3	-0.1	-0.2	-0.3	-0.2
2-y	0.4**	0.5**	0.2	0.2	0.1	-0.1	-0.2	0.0
5-y	0.4**	0.4**	0.1	0.2	0.2	-0.0	-0.1	0.1
10-y	0.3**	0.3**	0.1	0.3*	0.2	0.0	-0.1	0.1

Qualitatively similar results with full sample

Macro outcomes

pre-crisis sample: 1989-2005

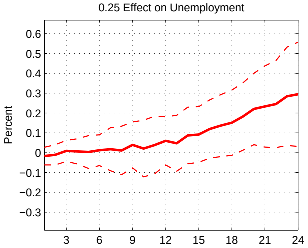
Target change: + 0.25 Target change: -0.25



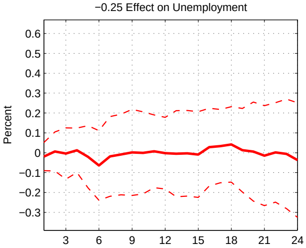
Macro outcomes (cont.)

pre-crisis sample: 1989-2005

Target change: + 0.25



Target change: -0.25



Takeaways

Raising the fed funds target:

- propagates through the term structure more normally
- reduces economic activity after 2 years with visible effects on IP and UR
- has a more muted effect on prices but eventually in the right direction

Lowering the fed funds rate:

- does not propagate through the term structure as well
- has no effect on economic activity or inflation

Results with the full sample are very similar

Conclusions

Establishing causality is always difficult. IPW is arguably the best we can do with observable information.

New methods are straightforward and flexible

Effects of monetary policy on the economy:

- asymmetric in normal times (raise $\neq$ decrease)
- similarly ineffective during the crisis
- crisis sample analysis (not shown) confirms asymmetry

With instruments combine methods (e.g., Jordà, Schularick and Taylor 2014 "Betting the House")