Generated Instruments with Validation Data

Xinran Liu & Dingli Wang

Example

March 16, 2022 FOMC announcement

Policy decision

+25 bp

Increase in the federal funds target range.

Market observation

↓ 5-year BE

The five-year breakeven inflation rate fell sharply that day.

Does tighter monetary policy *cause* inflation compensation to fall?

Why do we need an instrument?

Around an FOMC announcement, asset prices can move for several reasons.

Policy news

Markets learn that the policy path is tighter or looser than expected.

Other news

Markets may also learn about:

- the economic outlook,
- inflation risks,
- risk premia.

Observed yield movements alone do not tell us which force is driving the response.

$$\Delta y = \text{response to policy news} + \text{response to other news}.$$

The conventional conclusion: real yields rise more

Nominal yields combine real yields and inflation compensation:

$$\Delta y^N = \Delta y^R + \Delta BE.$$

Across meetings, for a common policy-repricing variable D, write

$$\begin{aligned} \Delta y^N &= \beta_N D + \varepsilon_N, \\ \Delta y^R &= \beta_R D + \varepsilon_R, \\ \Delta BE &= \beta_{BE} D + \varepsilon_{BE}. \end{aligned} \qquad \Longrightarrow \qquad \boxed{\beta_{BE} = \beta_N - \beta_R}.$$

Real financing conditions

$\beta_R > 0$: real yields rise.

Inflation compensation

$\beta_R > \beta_N$: inflation compensation falls.

Conventional IV: real yields rise more than nominal yields

	Change in basis points for a 10-bp OIS repricing		
Communication window	Nominal yield	Real yield	Inflation comp.
Statement	11.4	**15.5**	−4.2
Press conference	10.4	**14.2**	−3.8
Full communication	12.4	**16.4**	−4.0

Real yields rise: $\beta_R > 0$ **Inflation compensation falls:** $\beta_{BE} < 0.$

The estimates use 92 FOMC meetings and a 10-bp OIS normalization. They are not the March 2022 event response and require exact exclusion.

Individual estimates are separately rounded to one decimal place.

Can we interpret this pattern as a policy effect?

The structural equation for each yield outcome is

$$Y^k = \beta_k D + \varepsilon_k, \qquad k \in \{N, R, BE\}.$$

Our generated high-frequency instrument Z identifies the policy response under **exact exclusion**:

$$\boxed{\mathrm{Cov}(Z, \varepsilon_k) = 0.}$$

Research question: Do the conclusions $\beta_R > 0$ and $\beta_R > \beta_N$ survive when exact exclusion is relaxed?

Our contribution: validation-calibrated sensitivity analysis

1. **Learn the scale of error.**
   External validation data quantify the error in the generated instrument.

2. **Allow what validation cannot determine.**
   A sensitivity parameter limits the unknown relationship between instrument error and structural disturbances.

3. **Draw robust conclusions.**
   We construct identified sets and use weak-IV-robust inference, rather than relying on the conventional IV ratio.

The empirical objects and the three event windows

Policy repricing and instrument

D: change in the 1-year OIS rate.

$X = (MP1, MP2, ED2, ED3, ED4)'$.

$Z = \widehat{w}_A' X$: one rate factor from the five surprises.

Yield outcomes

Y^N: change in the 5-year nominal yield.

Y^R: change in the 5-year TIPS real yield.

$Y^{BE} = Y^N - Y^R$.

We use the **same 92 FOMC meetings** in three intraday windows:

$$\underbrace{\text{Statement}}_{\text{release}} \qquad \underbrace{\text{Press conference}}_{\text{chair's remarks}} \qquad \underbrace{\text{Full communication}}_{\text{both}}.$$

PCA summarizes the five surprises but does not guarantee exclusion. Our analysis allows the resulting factor to remain imperfect.

These event windows are distinct from the three historical samples on the next slide.

Three historical samples, three different jobs

1994–2003
Construction

Observe five surprises X.

Estimate PCA weights $\widehat{w}_A$.

Fix $Z_t = \widehat{w}_A' X_t$.

2004–2010
Validation

Observe Z and benchmark S.

Estimate the scale τ_V of instrument error.

2011–Jun 2026
Target

Observe Z, D, Y.

Transport the error scale.

Infer policy responses β.



Why these dates? 1994 begins the announcement sample; 2004 separates construction from validation; 2011 marks the start of regular FOMC press conferences. 2026 is the data cutoff.

What's S? S denotes the Jarociński–Karadi (JK) monetary-policy instrument, which serves as the external benchmark in the validation sample.

Validation identifies error scale, not its target relationship

In the validation sample, compare the generated factor with a policy benchmark:

$$\boxed{Z = \lambda S + u.}$$

What validation tells us

The scale of the unexplained part of Z:

$$\boxed{\tau_V = \mathrm{Var}_V(u).}$$

What validation does not tell us

Its association with the target disturbance:

$$\boxed{\mathrm{Corr}_T(u, \varepsilon_k).}$$

We observe S in the validation sample, *not* as a competing IV in the target sample.

How we allow the generated instrument to be imperfect

Carry the validation error scale into the target period using one of two assumptions:

$$\underbrace{\tau_T = \tau_V}_{\text{constant error variance}} \qquad \text{or} \qquad \underbrace{\frac{\tau_T}{\mathrm{Var}_T(Z)} = \frac{\tau_V}{\mathrm{Var}_V(Z)}}_{\text{constant relative error share}}.$$

Then allow the instrument error to covary with other yield news:

$$\boxed{\left|\mathrm{Corr}_T(u, \varepsilon_k)\right| \leq \bar{\rho}.}$$

Interpretation of $\bar{\rho}$

$\bar{\rho} = 0.50$ means that the *absolute correlation* may be as large as 0.50. It does **not** mean that 50% of the IV is invalid.

From one IV estimate to an identified set

Under exact exclusion, the moment condition yields a ratio:

$$\mathrm{Cov}_T(Z, Y - \beta D) = 0 \quad \Longrightarrow \quad \beta^{IV} = \frac{\mathrm{Cov}_T(Z, Y)}{\mathrm{Cov}_T(Z, D)}.$$

With imperfect exclusion and a valid benchmark policy component, write $\varepsilon = Y - \beta D$ and allow

$$\left|\, \mathrm{Cov}_T(Z, Y - \beta D) \right| \;\le\; \bar{\rho} \sqrt{\tau_T \, \mathrm{Var}_T(Y - \beta D)}.$$

The identified set therefore contains all β satisfying

$$\boxed{\mathcal{B}(\bar{\rho}; \tau_T) = \left\{ \beta : \; \left|\, \mathrm{Cov}_T(Z, Y - \beta D) \right| \le \bar{\rho} \sqrt{\tau_T \, \mathrm{Var}_T(Y - \beta D)} \right\}.}$$

For each candidate β: can the allowed instrument error explain the remaining instrument–residual covariance?

Inference: Testing candidate coefficients

Instead of estimating an IV ratio, we test each candidate β directly.

For $e(\beta) = Y - \beta D$, define

$$\boxed{\widehat{g}_j(\beta) = j\widehat{\mathrm{Cov}}_T(Z, e(\beta)) - \bar{\rho}\sqrt{\widehat{\tau}_T}\,\widehat{sd}_T(e(\beta)), \quad j = \pm 1.}$$

Uncertainty from two samples

$$\widehat{\sigma}_j^2(\beta) = \underbrace{\frac{\widehat{V}_{T,j}(\beta)}{n_T}}_{\text{Target sample}} + \underbrace{\frac{\widehat{V}_{V,j}(\beta)}{n_V}}_{\text{Validation sample}}.$$

Inference accounts for uncertainty in both market-response moments and the estimated IV error scale.

Theorem: Weak-IV-robust inference

We invert the two moment-inequality tests to construct the confidence set:

$$\mathcal{C}^B_{1-\alpha}(\bar{\rho}) = \left\{ \beta : \max_{j=\pm 1} \frac{\widehat{g}_j(\beta)}{\widehat{\sigma}_j(\beta)} \leq z_{1-\alpha/2} \right\}.$$

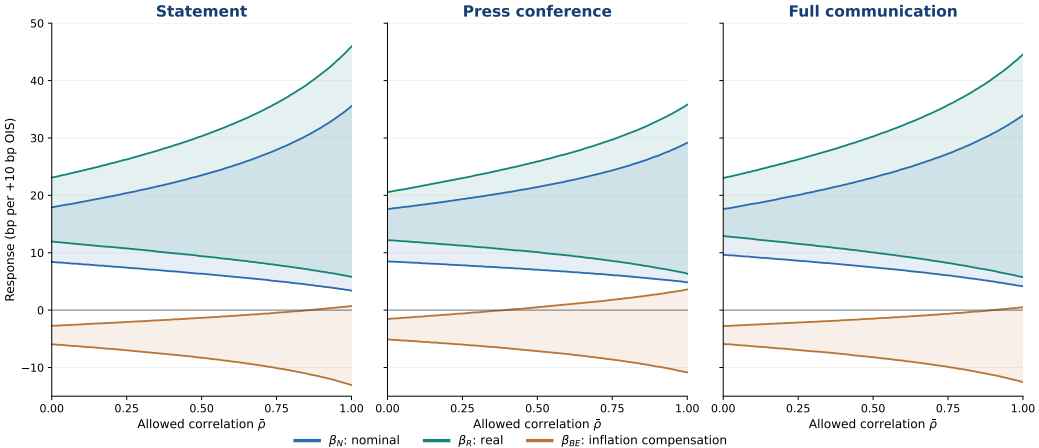
Theorem: Uniform coverage

Under the transport, sensitivity, and sampling regularity assumptions,

$$\liminf_{n_T, n_V \to \infty} \inf_{P \in \mathcal{P}_M} \Pr_P \left(\beta_0 \in \mathcal{C}^B_{1-\alpha}(\bar{\rho}) \right) \geq 1 - \alpha.$$

The coverage guarantee holds even when the first stage is weak or exactly zero.

Continuous sensitivity: all three yield responses



Shaded regions: 95% marginal confidence sets · Constant relative error share · HAC(2)

Which conclusions survive when exclusion is relaxed?

Does the stronger ordering $\beta_R > \beta_N$ remain supported?

Allowed $\bar{\rho}$	Statement	Press conference	Full communication
0.25	**Supported**	**Supported**	**Supported**
0.50	**Supported**	**Not supported**	**Supported**
0.75	**Supported**	**Not supported**	**Supported**

The decline in inflation compensation is more sensitive to IV imperfection, especially in the press-conference window.

How much exclusion violation can the conclusion tolerate?

At what value of $\bar{\rho}$ does the 95% confidence set first include zero for β_{BE}?

Communication window	Threshold $\bar{\rho}^*$
Statement	0.86
Press conference	0.39
Full communication	0.90

The negative inflation-compensation response loses statistical support first in the press-conference window.

The real-yield and nominal-yield responses are more robust than their ordering.

Conclusion: the real-yield rise is more robust

1. **Conventional IV:** real yields rise more than nominal yields, implying lower inflation compensation.
2. **Validation data:** estimate the scale of instrument error, while $\bar{\rho}$ controls its unknown target-sample correlation with the disturbance.
3. **With imperfect exclusion:** positive nominal and real yield responses are comparatively robust; a decline in inflation compensation is more fragile.

External validation tells us which substantive monetary-policy conclusions survive when a generated IV need not be perfectly valid.

Backup: Why the first principal component?

We follow the conventional one-factor monetary-policy instrument used in the empirical literature:

$$X = (MP1, MP2, ED2, ED3, ED4)', \qquad Z_t = \widehat{w}_A' X_t.$$

- PCA summarizes correlated rate surprises; **it does not guarantee exclusion**.
- Our question is whether conclusions from this commonly used factor survive when it is imperfect.
- Useful additional checks: individual surprises, current-policy versus forward-guidance factors, and alternative factor weights.

We do not claim that combining potentially contaminated surprises creates an exogenous instrument.

Backup: The three intraday FOMC windows

The same 92 target-sample meetings are evaluated in each of three windows:

Window	Measurement interval
Statement	10 minutes before to 20 minutes after statement release.
Press conference	10 minutes before to 60 minutes after the press conference begins.
Full communication	10 minutes before statement release through 60 minutes after the press conference begins.

This is separate from the historical construction, validation, and target samples.