Nonlinearities in Monetary Policy Transmission

Paul Bousquet*

July 14, 2024
[Click here for latest version]

Abstract

Most empirical and quantitative estimations of impulse response functions (IRFs) rule out the presence of nonlinearities by construction. Namely, it's possible that big shocks are not a scaled version of small shocks (size effect) or that responses following a positive shock do not mirror those from a negative shock (sign effect). Using a Local Projection Instrumental Variable (LP-IV) framework, I estimate impulse responses of standard macro variables to monetary policy shocks and look for evidence of nonlinearities. Specifically, I use monetary policy surprise series as an instrument and decompose the measured shock into regimes based on how the Fed Funds Rate changed in a given month to allow for nonlinearity. Focusing on output and inflation, I find extensive evidence of size effects and asymmetries for large changes in policy that work in opposite directions. I show that a common alternative to linear DSGE models, a New Keynesian model with downward rigidities in price and wage adjustments, cannot be modified to remotely produce similar IRFs. These results suggest better understanding is needed about how policy and economic fundamentals interact both in the data and our models.

*Department of Economics, University of Virginia, ptb8zf@virginia.edu.

Introduction

Much of the evidence on the effects of monetary policy has been produced from model-based and empirical approaches that impose a linear relationship between shocks and real variables. Namely, while standard impulse response functions (IRFs) can be non-linear functions of time, it's typically the case that if the impulse response to a shock s_t is $\{y_{t+h}\}$, the impulse response to αs_t is simply $\{\alpha y_{t+h}\}$. This rules out non-linearity from a shock's *size* (big and small don't have proportional effects) and *sign* (positive and negative don't have symmetric effects). To produce any applicable policy recommendations, it must be determined whether this simplification is consequential.

Formally, we can put nonlinearities into two categories: size and sign effects. While little attention has been paid to the potential for size effects, Economists dating back to the Great Depression have explored the idea that expansionary and contractionary monetary policy have asymmetric properties. One popular narrative is that during a downturn there is little central banks can do to create an appetite for lending and spur broader economic activity, commonly paired with the analogy that "pushing on a string" is futile. Grigoli and Sandri (2023) summarize the body of empirical evidence on this thesis as being generally mixed but slightly supportive, albeit with an abundance of statistical insignificance. Still, there is currently no knowledge on the intersection of size and sign. If any sort of asymmetric economic hindrances exist, knowing how the magnitude and direction of policy interacts with each other in transmission should more directly reveal the exact nature of these constraints.

I first estimate impulse response functions using LP-IV, the Stock and Watson (2018) Instrumental Variable extension of Jordà (2005) Local Projections, and allow for the possibility for size and sign effects by introducing 4 interest rate change regimes (combinations of big vs. small, positive vs. negative) and instrumenting each with established monetary surprise shock series. For all variables, I find size effects exist with both interest rate cut and hikes and sign effects for big changes. As shown in Table 1, size effects at a horizon of 12 months are extensive. Returning to the "string" narrative, the sign effect results imply that for policy changes of modest size, which constitute the vast majority, there is not enough data to observe meaningful asymmetry, and the (statistically insignificant) estimates that are produced point in the other direction (greater magnitudes for small cuts). For large changes, output-related variables (consumption and industrial production) do respond more to hikes but the opposite is true for inflation. These results are insensitive to a myriad of perturbations in the specification.

Table 1: Difference in YoY growth between large shock and scaled small shock

	Cut Size Effect	**Hike Size Effect**
Consumption	-1.8%	-2.3%
Industrial Production	-6%	-3.5%
CPI	-1.4%	0.8%

All estimates significant at 95% level

I then show that a workhorse New Keynesian model cannot match these findings. To allow for consequential nonlinearities, the model includes asymmetric frictions in price and wage setting and is estimated to second order via a Metropolis-Hastings routine and particle filter from Fernandez-Villaverde and Rubio-Ramirez (2007), producing a posterior distribution implying downward-rigid prices and wages found in micro-evidence and past quantitative work (e.g., Kim and Ruge-Murcia, 2009; Aruoba et al., 2017). A comparable LP-IV specification combined with model-simulated data is used to produce Bayesian IRFs and also search for parameter combinations that minimize deviation from the results from US data. These illustrations show the model can produce non-linear effects of monetary policy, but they appear on impact and quickly vanish. The implication is that there is not enough internal propagation in this model to generate the sustained nonlinearities observed empirically. Namely, a Taylor Rule, even with a lagged interest rate, does not create a sufficient amount of persistence on its own.

The primary contribution of this paper is presenting new facts about the historical effects of monetary policy and our default model's ability to account for them. First, we observe that non-linear behavior is most potent in the medium run. Earlier effort has been made to modify New Keynesian models to yield "hump shaped" impulse responses, but whether it's possible to do the same for non-linear differences based on the type of shock is unclear. We also see that for small shocks, there is no "pushing on a string" dynamic between policy and real variables. This possibly suggests policymakers do not face asymmetric constraints to stimulate the economy until the goal for expansion becomes sufficiently large. For big changes in policy, such asymmetric constraints are observed for output: big hikes amplify the contractionary effect and big cuts have a proportionally smaller effect. The opposite relationship exists for inflation. Some may posit that (downward) nominal rigidities could rationalize all the above, but a New Keynesian model with such frictions fails on all fronts. The implications of these results immediately point to new, necessary directions for future research.

Related Literature: This work connects to and clarifies a vast empirical literature assessing the effects of monetary policy using time series data paired with a high-frequency (e.g., Bauer and Swanson, 2023) or orthogonalization (e.g., Romer and Romer, 2004) shock identification strategy rather than specifying a general equilibrium model. Ramey (2016) provides a survey of various method-shock series combinations and documents prevalent "puzzles", implications of impulse responses that conflict with standard intuition. In general, this literature is noisy – one can find a well-cited paper suggesting that macro variable x responds in y direction for any combination of x and y. I sidestep the puzzle rabbit hole altogether by focusing on *relative* effects, which haven't received much theoretical or applied attention. This paper is also the first in this setting, to my knowledge, to leverage the wealth of research that has come out about Local Projections the past few years. Namely[1], the differences between LPs and vector autoregressive estimation (VAR) have been made clear. Plagborg-Møller and Wolf (2021) shows the methods yield

[1]Other important considerations include bias in estimates (Gonçalves et al., 2024; Herbst and Johannsen, 2024), bias in standard errors (Plagborg-Møller and Montiel Olea, 2021; Herbst and Johannsen, 2024), and general identification (Stock and Watson, 2018; Jordà, 2023)

the same population estimates of impulse responses and in finite samples the trade-ff boils down to flexibility (LP) vs. efficiency (VAR). Even more recent work by Montiel Olea et al. (2024) reveals the cost of the efficiency gains from VARs can be prohibitive: they are comfortably insensitive to misspecification if and only if the coefficient of interest has similar variance to its LP analogue. I use a penalized approach (Barnichon and Brownlees, 2019) to retain the appealing properties[2] of LPs with some added efficiency. These advantages are also extended to a LP-IV framework, which has been underutilized. The various monetary shock series that have been developed still have unresolved interpretability issues (e.g., Brennan et al., 2024), but what is consistent across all approaches is that they are developed in order to capture relevant changes in monetary policy while maintaining exogenous variation. Thus, it's more appropriate to treat these shock series as instruments. Moreover, by instrumenting actual changes to the interest rate target, we can argue the coefficients produced in our estimation are representative of the effects of monetary policy in general, rather than just policy surprises.

Broadly speaking, the set of popular methods has remained the same over the past decade. In addition to providing insight on the usability of past research by accounting for recently revealed shortcomings, this paper is one of few to focus on nonlinearities. Even fewer have taken this focus to LPs, with much of the group (e.g., Tenreyro and Thwaites, 2016; Ascari and Haber, 2022) incorporating a methodology for state-dependent IRFs that induces overwhelming bias (Gonçalves et al., 2024) and seemingly none accommodating both size and sign effects. Angrist et al. (2016) gauges possible asymmetries using a unique, inverse probability weighting identification scheme, but their results are not statistically significant and the crucial overlap assumption[3] plainly does not hold. Additionally, some remain committed to a preference of VARs and could perhaps point to evidence from Stock and Watson (2018) suggesting inevitability is a relatively innocuous assumption for this application. However, the monetary shocks identified by Debortoli et al. (2023), which is the most general VAR-based approach I'm aware of (e.g., relaxes ordering assumption, permits general nonlinear functions of variables as controls), show that being explicitly tied to a Taylor Rule that does not account for the zero lower bound (ZLB) is an insurmountable hurdle to inference, especially given the limited sample size. Barnichona and Matthes (2018) estimate non-linear impulse responses through a vector moving average representation with some structural restrictions and present substantive evidence of sign effects with inflation moving counter to what should persist in a "pushing on a string" world, in line with my findings. Another contribution in this realm is using standard deviations of effect sizes to provide a novel depiction of nonlinearities, making them more readily comparable to estimations of DSGE models by removing any meaningless distortion from scaling differences between time series and model-simulated data.

Finally, this research also adds to the literature on model-based accommodations of nonlinearities. One well-documented cause of model nonlinearities is the ZLB. In particular, the standard New Keynesian model

[2] Also superiority in dealing with non-stationarity (Plagborg-Møller and Montiel Olea, 2021; Plagborg-Møller et al., 2024) and relaxing VAR's structural assumptions (namely, recursiveness and inevitability). Ramey (2016) and Mavroeidis (2021) find recursiveness is not palatable.

[3] Probability of every possible treatment (change in i_t) should be bounded away from 0 and 1 at all points in state space

has several counterintuitive implications about optimal policy when the economy is sitting at the bound (e.g., Eggertsson, 2011; Christiano et al., 2011; Wieland, 2019). Bonciani and Oh (2023) find that many of these "paradoxes" are resolved by allowing the central bank to conduct a wider array of open market operations, which provides a path for future progression of this research for both the empirical and quantitative exercises. For now, these results most directly connect to the smaller subset that have explored the role in asymmetric frictions in price and wage setting. Past work with firm-level data, dating as far back as Keynes (1936) and Tobin (1972) and continuously reconfirmed since, that there are relatively more frictions of revising prices and wages downward. Kim and Ruge-Murcia (2009) is the canonical reference for introducing more costly downward price and wage rigidities into a New Keynesian model, but they assume a non-stochastic monetary policy rule. Aruoba et al. (2017) solve a more general version of the model, but do not focus on impulse responses. Lee (2023), the closest companion paper, uses an asymmetric investment constraint and simpler downward wage adjustment asymmetry and finds both sign and size asymmetries. Oh (2020) finds that under uncertainty shocks (unexpected increases in the second moments of exogenous processes) Calvo pricing yields more non-linear behavior resulting from these shocks compared to symmetric menu costs. This also provides direction for future work to employ a more general menu cost structure (Reiter and Wende, 2024) that can also accommodate features of Calvo pricing and match stylized facts of more complicated state-dependent pricing models.

Empirical Methodology and Results

Framework and Data

Impulse responses generated by local projections (LP), pioneered by Jordà (2005), are a collection of coeficients $\{\hat{\alpha}_h\}$, where each $\hat{\alpha}_h$ comes from a regression of y_{t+h} on time t control variables X_t and shocks s_t

$$y_{t+h} = \alpha_h s_t + \beta_h X_t + \epsilon_{t,h}$$

Because these regressions are independent of one another, this is a non-linear impulse response function (IRF). However, this is only non-linear with respect to time, not the shock itself. We want to relax this imposed linearity and allow responses to vary by the type of shock that occurred at $t = 0$. Linearity in this context means big shocks and small shocks have nearly proportional effects and positive and negative shocks (of the same size) produce symmetric effects. The usual LP framework can be modified to accommodate potential nonlinearities by decomposing our shock into several shock series based on what "regime" of monetary policy we are in. We specify 4 regimes by creating a single threshold within our two dimensions of size and sign:

1.BH (Big Hike) 2.BC (Big Cut) 3.SH (Small Hike) 4.SC (Small Cut)

Formally, regimes can be represented as indicator variables and the vanilla LP decomposition becomes

$$y_{t+h} = \underbrace{\alpha_h^R R_t s_t}_{(\alpha_h^{BH} r_{1,t} + \ldots + \alpha_h^{SC} r_{R,t}) s_t} + \beta_h X_t + \epsilon_{t,h}$$

where α_h^R is a 1×4 matrix of coefficients and R_t is 4×1 matrix of the corresponding regime indicator variables, which can be generalized to more regimes as needed. One interpretation of this setup is that there are 4 shocks[4], one of which will be "activated" in each period.

The first step towards implementation is finding a suitable shock series s_t to represent exogenous variation in interest rates. We are in essence seeking to recover the "treatment effect" of monetary policy and therefore must overcome the classic identification concerns presented by simultaneity (outcome variables influencing the probability of treatment) and anticipation (the model will be misspecified if we're capturing responses to something other than our s_t). There have been a litany attempts to construct monetary surprise measures that reflect unanticipated change relative to what people expect, with many using a high-frequency window around monetary policy announcements to plausibly argue the data is solely capturing responses to central bank action. Because these measures can be normalized in different ways (Acosta, 2023) or have interpretation that is sensitive to units or specification (Brennan et al., 2024), it's less appropriate to include them in estimation procedures directly as regressesors. Rather, we can leverage the fact that they are constructed to satisfy the usual IV assumptions of relevance and exogeneity (a sufficient condition for the exclusion restriction). Another important advantage is by instrumenting actual changes to the interest rate target, the estimation is more representative of the general effects of monetary policy, rather than just policy surprises. Thus, we proceed with the LP-IV approach put forward by Jordà et al. (2015) and Stock and Watson (2018).

We also need to explicitly define our regimes, namely take a stance on what constitutes a "big" change. We use a cutoff of a 10% jump from the previous fed funds target, which balances sample sizes between regimes nicely. The cost is no changes close enough to the zero lower bound (ZLB) are classified as small. However, percent change is more in line with the treatment of shocks in structural setups. There is also an instrument validity problem with the monetary shock series when using magnitudes to define regimes, partially due to a scarcity of hikes greater than 25 basis points. Conversely, under our chosen specification I find that the Romer and Romer (2004) series is a strong instrument for all regimes except "small cut", for which I instead use Bauer and Swanson (2023).

As previously mentioned, LP faces efficiency concerns[5] and there can also be large variability in the coefficients from one horizon to the next. To address both concerns, the main visualizations are generated by 2SLS combined with the penalized LP approach of Barnichon and Brownlees (2019), who approximate the shock coefficients using

[4] In a pure LP setup, there would also be a fifth regime for when there is no change in the LP target. However, because we are using LP-IV and thus instrumenting changes in the target, there is no such regime in LP-IV and thus this is omitted from the initial setup for clarity

[5] An advantage of LP-IV is since we're able to use a different shock series for each regime, we can simply select the strongest instruments for efficiency gains. This is a similar dynamic to components of VAR models that are only estimated over subsamples (Gertler and Karadi, 2015)

B-spline basis functions and minimize a ridge loss function with a particular penalty matrix that shrinks the IRFs towards a polynomial of a given order as the regularization parameter λ grows. Following their guidance on the sensitivity of inference in light of penalization and cross-validation induced complications, I fix λ at a mild level and use standard errors from an even more "under-smoothed" estimation (with $.1^{*}\lambda$). Most applications of local projections estimate standard errors using a Newey-West adjustment, which theoretically accounts for heteroskedasticity and autocorrelation (HAC), but Herbst and Johannsen (2024) finds this procedure yields biased estimates. Rather than introducing unneeded complexity by using a more sophisticated HAC-robust method, Plagborg-Møller and Montiel Olea (2021) shows that using the usual Huber-White heteroskedasticity-robust standard errors paired with including a sufficient number of lagged control variables is sufficient.

Following Ramey (2016), we consider outcome variables at a monthly frequency of the Consumer Price Index (CPI), industrial production, 1 year treasury yields, excess bond premium (Favara et al., 2016), and add real consumption expenditures. Control variables also include lagged interest rates, monetary policy uncertainty (Husted et al., 2020), an indicator for the ZLB, and a healthy number lags (12) of both outcome and controls following our discussion of standard errors. Data is sourced from FRED unless noted otherwise and the maximum sample periods are retained, which is trimmed down to 1988-2019 because that's the sample period of the Bauer and Swanson (2023) series. We focus on CPI and the joint picture of output painted by industrial production and consumption in order to take the findings directly to models. Outcome variables are cumulative log differences, yielding an approximate percent change interpretation: $\widehat{\alpha}_h$ is represents the percent change in levels h periods after a shock at t. At $h = 12$, this takes a nice form of year over year growth.

Results

Now that means of estimating the LP-IV coefficients has been established, the focus turns to illustrating possible nonlinearities, which we refer to as size and size effects. The most straightforward way to think of these effects is as functions of parameters. For the simple case of plotting in levels, the objects of interest are clear

$$\text{Size Effect}_h : \hat{\alpha}_h^B - \hat{\alpha}_h^S \qquad \text{Sign Effect}_h : \hat{\alpha}_h^P + \hat{\alpha}_h^N$$

In other words, a size effect exists if we can conclude the difference in the big and small (regime) coefficients are distinct from 0 and a sign effect exists if positive and negative coefficients have different magnitudes. This is complicated slightly by wanting to allow for both types of non-linearities simultaneously: we want to see if a size effect exists for both cuts and hikes and a sign effect exists for both big and small changes, in other words 4 graphs per outcome variable. Grouping is thus by type of nonlinearity, rather than outcome. For inference, we use 95% confidence bands and thankfully get standard errors for free from the variance-covariance matrix.

Figure 1 shows[6] that significant size size effects exist for all variables and both positive and negative monetary policy changes at some horizon. Big cuts yield proportionally smaller values of all of our variables of interest in the medium to long-run. The same is true for big hikes relative to small hikes with the exception for inflation, where it appears that the conventional relationship between output and inflation is weakened for larger changes. For sign effects, Figure 2 shows that while not much can be said for small changes, there is asymmetry for large changes.

Figure 1

In terms of drawing conclusions, the extent of these nonlinearities is arguably more open to interpretation than desirable because there could be differing opinions about how much attention should be paid to a given %-variable-horizon combination. To remove a bit of this subjectivity, the Appendix shows these same graphs but in terms of standard deviations, rather than percent change in levels since $t = 0$. The original figures show size effects exists for all variables under both cuts and hikes, and consulting the plots in the appendix reveals this amounts to a more than 2 standard deviation away realization of the small change coefficients in all cases. That level of magnitude leaves no ambiguity that the difference in effects is significant.

[6] IRF plots are often scaled to represent responses to a 25 basis point change in i_t. We scale so our size regimes correspond to 50 and 25bps, respectively. Intuitively, these plots then show the difference between getting hit with a 50 or 25bp change while sitting at $i_t \in [2.5, 5]$. The findings do not materially differ but this is in principle a necessary step because our instruments have differing abilities as an Δi_t proxy

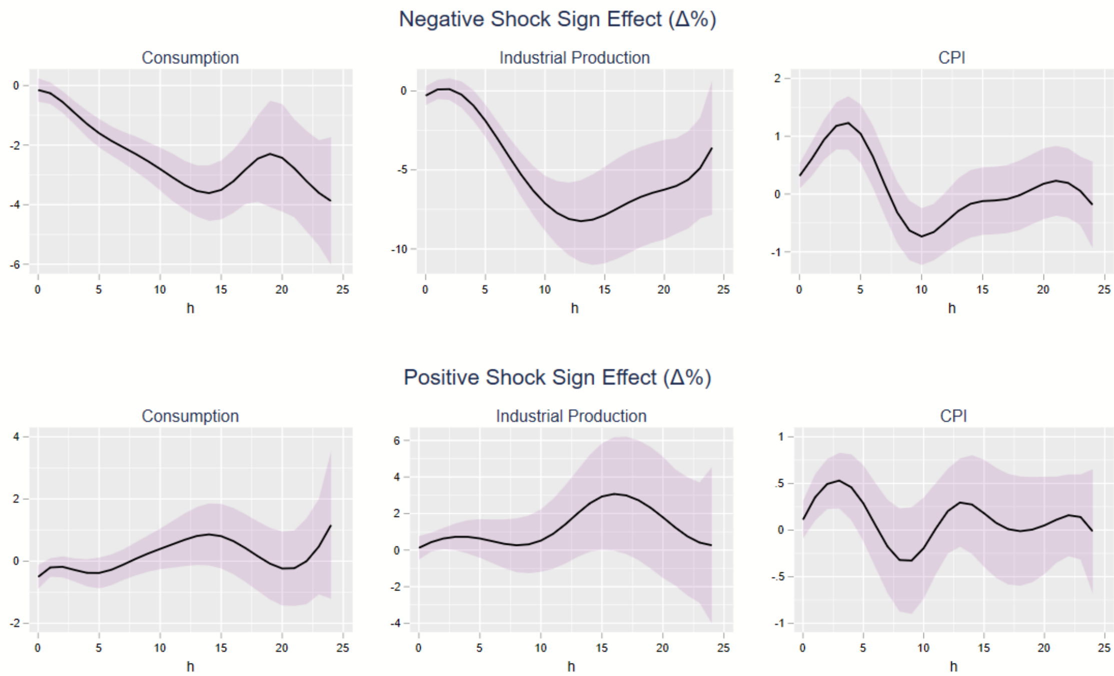


Figure 2

Sensitivity

Before gauging how models holds up to these data-based findings, it's important to have a sense of what, if anything, can make these results weaken when pushed. Changing the lag order, choice of monetary surprise measure, adding and removing controls, estimating in differences vs. levels[7], different measures for inflation, bias-correcting point estimates (Herbst and Johannsen, 2024) and using LP instead of LP-IV in general do not yield IRFs with different interpretations, even under various combinations of these factors. The one dimension that has revealed some delicacy is sample size[8], in line with Ramey (2016). This is unsurprising given the relatively few months there is data ($n/k \approx 2$) and the even fewer periods with consequential central bank intervention (e.g., only the Romer and Romer (2004) measure includes the hikes of the early 80's). Future work will extend these procedures to VAR and Bayesian methods. We can also consider alternative measures of interest rates to address the zero lower bound. One option would be using a non-linear filtering procedure (e.g., Farmer, 2021) to construct a shadow interest rate, or a measure of how interest rates "would have moved" if the ZLB didn't bind.

A more involved critique of model-free IRF estimation is an inability to account for state-dependence. For example, many past efforts try to allow for responses in boom and bust cycles to be asymmetric. With respect to

[7] Estimating in differences and summing for the cumulative effect in levels produces very similar IRFs to estimating in levels directly, despite that some of the outcome variables are highly non-stationary (McCracken and Ng, 2016) . This is in line with Plagborg-Møller and Montiel Olea (2021), who show LP is, in general, remarkably robust to the presence of unit roots and non-stationary variables.

[8] For instance, increasing the size threshold to 12.5% for our regimes yields much different sample distributions. The results generally hold under this definition but are less pronounced for inflation

interest rate shocks of a given size, another worry could be that beliefs about the future path of policy may not be updated in the same for different histories of action. The econometric concern is that these local projection coefficients amount to weighted averages and these weights could be biased if the joint distribution of the shock and state space has disparate behavior from a product of their marginals. In a regression context, this essentially amounts to the difference between including a variable as a control and additionally interacting it with the shock. Jordà (2023), following the revelation of Gonçalves et al. (2024) that the previous default methodology can severely distort IRFs, provides a framework incorporating interaction terms to estimate state-dependent effects. While this is not feasible for LP-IV because of the curse of dimensionality, applying this approach in the LP analogue does not affect the conclusions. Gonçalves et al. (2024) themselves suggest non-parametric estimation, which has an over-parameterization problem with or without instrumenting (i.e., in either case, control variables must be shed). The non-parametric estimation performed thusfar focusing on boom and bust dependence has yielded similar size effect results. Future work will extend to IV by using a parametric first stage and also include a focus on the ZLB, since lifting off from the bound is classified as a large change under our definitions but may be dissimilar to changes in unconstrained states. It's also worth noting that Caravello and Martínez-Bruera (2024) show that under fairly mild assumptions, LP estimands do recover the aforementioned unbiased weighted average under general state-dependence. In light of earlier discussion, this implies if state-dependence is of great concern the bias-reducing properties of LP should be even more appealing relative to a VAR.

A separate issue that interacts with the LP-IV specification is the validity of the instruments themselves. Jarocinski and Karadi (2020) and Acosta (2023) present evidence that when using monetary policy surprise measures, it may be important to not assume that Fed announcements only affect people's beliefs about interest rates, i.e., they argue it's important to account for a "Fed information effect". Koo et al. (2024) are the first to document formally how this affects inference when using impulse responses generated by LP-IV. The monetary policy surprise literature has yet to reach a consensus on several key issues (Acosta, 2023; Brennan et al., 2024), crucially including how justifiable different identification strategies are across the array of shock series. Future work will seek to address these concerns, including a procedure that bootstraps the difference between a purely linear specification and ours. Jacobson et al. (2024) presents important work on the importance of temporal aggregation bias – the idea that IRFs can be skewed based on differences in measurement between a high frequency shock and a low frequency variable. Adding a control for the day of the month when a Fed announcement occurred did not affect the results, but because this is a relatively new critique, it's still unclear the best way to guard against its presence. One possibility could be adding high frequency asset prices as an outcome variable to have a better sense of the extent of information and related fast-moving transmission effects.

Quantitative Benchmarking

Model and Estimation[9]

To compare to the IRFs estimated directly on US data, a basic point of reference would be using a model that features meaningful nonlinearities to generate data and then run the same regressions. I estimate the model of Kim and Ruge-Murcia (2009), who add on asymmetric price and wage setting frictions to a standard New Keynesian model. Specifically, a firm seeking to change its price at a rate different than steady-state inflation face a linear-exponential (Linex) adjustment cost that takes the form of

$$\Phi_t^p(\pi_t) = \frac{\phi_p}{\psi_p^2}\left(e^{-\psi_p(\pi_t-\pi^*)} + \psi_p(\pi_t - \pi^*) - 1\right)$$

For $\psi_p > 0$, it's more costly to decrease prices than raise them (downward-rigid), for $\psi_p < 0$ prices are upward-rigid, and the function limits to symmetric adjustment costs as $\psi_p \to 0$. Nominal wage adjustment costs take on the same structure. Past estimation of this model have found evidence of downward rigidity in prices and wages, consistent with empirical evidence. I follow the modifications of Aruoba et al. (2017) and allow for stochastic shocks to price markups and the interest rate setting rule.

Using the same sample period of US data as the empirical IRFs, I performed a second-order, Bayesian estimation of the model via a standard random walk Metropolis-Hastings algorithm and a particle filter outlined in Fernandez-Villaverde and Rubio-Ramirez (2007). I use the distribution of parameters generated by this exercise to simulate data and run the same Local Projection Instrumental Variable (LP-IV) procedure to create various IRFs, resulting in two main illustrations of how this model generates nonlinearity. The first are relatively standard "Bayesian IRFs": data is simulated for each set of parameters, LP-IV is used to estimate the IRF for each, and then I plot the median outcome variable response at each horizon with bands given by the 10th and 90th percentile of responses, called a credible set. For the second illustration, I take the posterior mode of all parameters and then vary both asymmetry parameters (one at a time, in both directions, and then both at once in the same direction) while keeping everything else fixed, then simulate data and plot IRFs for each combination. This is meant to shed some light on what the relevant comparative statics for the parameters we most care about might resemble. I also attempt to directly match the empirical results by using Metropolis-Hastings to search over the parameter space and find a vector that generates IRFs which minimize loss with respect to the original estimation.

The results show that this while the model can generate nonlinearities, in general the observed asymmetric effects for both size and sign occur on impact and then quickly dissipate, unlike what we saw with US data.

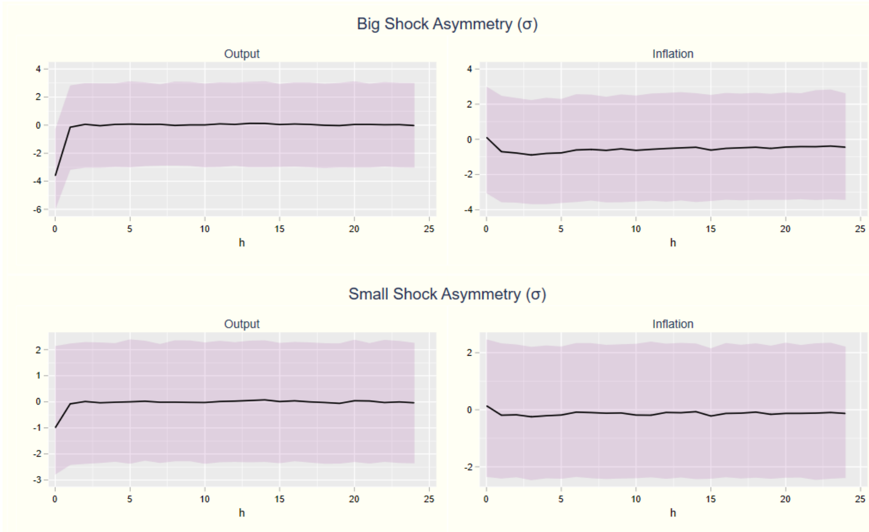
[9]The full list of model equations and more detail on the estimation can be found in the Appendix.

Bayesian IRFs

For the set of draws that came out of our Metropolis- Hastings routine, I simulated data of 400 observations for each group of parameters to align with the US data sample size. I then created IRFs using LP-IV estimation, where the instrument in this case is $z_t = i_t - \mathbb{E}_{t-1}[i_t]$ to match the construction of monetary surprise measures in the literature. Analogous control variables are included (lagged interest rates, zero lower bound, unemployment, output and interest rate variance) and plot in terms of standard deviations to abstract away from any trivial differences in levels between model-simulated and US data. One caveat is I do not use the aforementioned penalized 2SLS approach because it significantly reduces the computation time. However, Bayesian LPs can be thought of as another type of "smooth" LP, so this procedure is still capture the spirit of previous the previous illustrations.

For size effects, we largely see extremely precise estimates around 0, suggesting in this model a big interest rate change change does in fact amount to a scaled small change. The exception is the immediate impact on output: interest rate cuts are proportionally less stimulative and hikes are proportionally less contractionary on impact. This effect dissipates quickly, likely because when a big monetary policy shock takes the economy away from target (r^*), the central bank takes a large corrective action in the next period (consecutive big shocks are unlikely). For

sign effects, the standard errors are much larger. So what we can say is that the evidence is suggestive against the presence of sign effects, but we cannot be as confident as before. Some of the estimation performed for the illustrations in the next section reveals a likely reason for the wide credible sets. The estimates for sign effects for a given set of parameters tends to fluctuate around 0. This leads to lots of variance when pooling the results from thousands of sets of simulated data. In light of this context, we can be a bit more confident that this model is not generating meaningful nonlinearity.



Sketch of Comparative Statics

We also have an illustration of how size and sign effects vary as we make price and wage-setting frictions more and less asymmetric (both individually and jointly). This results in 12 graphs which were informative but not particularly eye-popping, so the illustrations are left in the Appendix, along with a table of concise descriptions and links to each. Each plot contains a solid black line at the posterior mode and then dotted, colorful lines to represent the IRFs for different values of the asymmetry parameters. For each graph, the way in which the parameters was changed was consistent: the dotted lines exclusively represent values that get increasingly far away from the mode in the same direction (e.g., the first figure shows how the IRFs change as price-setting frictions are made progressively more asymmetric).

For size effects, it seems like the partial derivative evaluated at the mode (for both asymmetry parameters) is relatively monotonic: if any one of the dotted lines was far from the black line, it typically implied that the other lines were in the space in between. However, the deviation from the mode line was is hardly ever consequential. Almost all graphs contained at least one deviation at $h = 0$, but otherwise tracked the mode line well. For sign effects, as alluded to in the previous subsection, the mode lines here are a lot more volatile and tend to bounce around 0. This also meant that the behavior of the dotted lines (as we varied the asymmetry parameters) was more pronounced, but ultimately these are seemingly random, one-time deviations. There was still general monotonicity (dotted lines all moving away from mode line in same direction) but it was definitely less consistent.

Therefore, we draw the same conclusions as from the Bayesian IRFs: we can tweak the asymmetry parameters to get whatever kind of non-linear behavior we want on impact, but we cannot produce sustained sign and size effects. The most serious long-term nonlinearity is for big shocks and inflation, with an sign effect that fluctuates around 0 but is still persistently negative. Even if we take that volatile point estimate seriously, this is the opposite direction of asymmetry we found in US data.

Fitting Empirical Results Directly

One thing lost in the credible sets of the Bayesian IRFs is keeping track of individual parameter estimates. For example, a given parameter vector from our Metropolis-Hastings could be in the 99th percentile for h=0 and the 1st percentile for h=1. We can instead try to find $\boldsymbol{\theta}^*$ that minimizes loss with respect to the empirical IRF coefficients. This is unfortunately not as simple as minimizing a traditional objective function. We must simulate multiple sets of data, which means thousands of regressions (51*the number of sets) at each step of the minimization algorithm. Given the length of the parameter vector, this would be overly time consuming and almost certainly get stuck at a local minimum. Instead, we do a loss-minimization procedure via a quasi Metropolis-Hastings algorithm. For each iteration, we take a draw as we would for a traditional Metropolis-Hastings and compute the loss. The difference is the loss is completely deterministic (the same sets of shocks are used to simulate data) so we accept or reject purely based on having a lower or higher objective function value.

The point of this exercise is to strip out the features of the data that are not of great importance and just try to generate these nonlinearities. We strip out even more by just focusing on the medium run (a year out), rather than trying to match the entire IRF. The barometer is not matching these point estimates exactly[10] but simply to see if the effects can be non-trivially pulled in the proper directions. It does not appear this is possible in a meaningful sense. This procedure was able to produce a parameter combination that cut the loss in half relative to the posterior mode, but this combination is extremely fragile and its success was at the expense of extreme volatility. This combination involved an IES of over 1000 ($1/\tau < 10^{-3}$) and the loss would change dramatically for minor perturbations to the parameters. Namely, this peculiarly happened when adjusting the steady state levels of output and inflation growth in either direction, suggesting the routine found an arbitrary combination of parameters which coincidentally produced these nonlinearities. Looking at the entire IRF confirms this intuition: the values at a year out were only matched thanks to wild swings from one horizon to the next. It's thus unsurprising that introducing penalization for the results 11 and 13 months out for smoothness results in even poorer performance.

Because this procedure will only span the parameter space asymptotically, these results will never be complete in some sense. This algorithm will continue to be ran with different starting points

[10] In any event, industrial production and consumption have to be combined to make a comparison

Conclusion

We found significant nonlinearities in the transmission of monetary policy to real variables that a workhorse New Keynesian model was unable to match. Specifically, big changes in interest rates depress the effects on inflation but for output the results match a "pushing on a string" story: big cuts have proportionally less stimulative effects but big hikes have proportionally more contractionary effects. but depresses the effects on inflation. For small shocks, the presence of asymmetry cannot be identified from the data. Adding downward rigidities in price and wage setting to a representative agent New Keynesian model comes nowhere close to replicating these findings. There, the nonlinearities appear on impact and quickly dissipate, whereas we observe empirically the most significant effects occur roughly a year after a given interest rate change. These results are novel along several fronts, including approach, visualization, and incorporation of recent literature about impulse response methods.

The clear next step is to determine what can successfully account for what we found. Some immediate modifications to the model to generate more internal propagation could include habit formation and price indexation with respect to lagged prices, rather than steady state. On a deeper level, any direction needs to carefully consider the equilibrium effects of policy adjustment. For example, Lee (2023) estimates a similar model with comparable findings and suggests that perhaps the Fed should refrain from making large changes. But this prescription does not account for anticipation effects, and Stein and Sunderam (2018) presents compelling theoretical evidence that gradualism is not possible without forward guidance. This leads to a natural potential avenue of directly measuring how forward guidance affects transmission. There has been some work done on this front, but given the differences in nonlinearities between output in inflation, more attention needs to be paid to the response of firms, which have well documented inattentiveness to policy change (Coibion et al., 2018). Another direction to account for the unique traits of inflation responses could be recognizing the non-uniform coordination of policy because of the global marketplace, which our models say would lead to disparate incentives for domestic and foreign firms. In any event, an overarching goal for the future will be quantifying welfare implications, as the optimal policy under a model that is able to match these observed trends may be very different.

References

Acosta, Miguel. 2023. “The Perceived Causes of Monetary Policy Surprises.” February, Working paper.

Angrist, Joshua D., Oscar J. Jordà, and Guido M. Kuersteiner. 2016. “Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited.”

Aruoba, S. Boragan, and Thomas Drechsel. 2024. “Identifying Monetary Policy Shocks: A Natural Language Approach.” Working paper.

Aruoba, S. Borağan, Luigi Bocola, and Frank Schorfheide. 2017. “Assessing DSGE model nonlinearities.” *Journal of Economic Dynamics and Control*, 83 34–54.

Ascari, Guido, and Timo Haber. 2022. “Non-Linearities, State-Dependent Prices and the Transmission Mechanism of Monetary Policy.” *The Economic Journal*, 132(641): 441–478.

Barnichon, Regis, and Christian Brownlees. 2019. “Impulse Response Estimation by Smooth Local Projections.” *The Review of Economics and Statistics*, 101(3): 522–530.

Barnichon, Régis, and Geert Mesters. 2024. “Innovation Powered Narrative Inference.” July, Working paper.

Barnichona, Regis, and Christian Matthes. 2018. “Functional Approximation of Impulse Responses.” *Journal of Monetary Economics*, 99 41–55.

Bauer, Michael D., and Eric T. Swanson. 2023. “An Alternative Explanation for the "Fed Information Effect".” *American Economic Review*, 113(3): 664–700.

Bonciani, Dario, and Joonseok Oh. 2023. “Revisiting the New Keynesian Policy Paradoxes under QE.” *European Economic Review*, 154.

Brennan, Connor M., Margaret M. Jacobson, Christian Matthes, and Todd B. Walker. 2024. “Monetary Policy Shocks: Data or Methods?” finance and economics discussion series, Federal Reserve Board, Washington, D.C.

Caravello, Tomás E., and Pedro Martínez-Bruera. 2024. “Disentangling Sign and Size Non-linearities.” June.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. “When Is the Government Spending Multiplier Large?” *Journal of Political Economy*, 119(1): .

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2018. “How Do Firms Form Their Expectations? New Survey Evidence.” *American Economic Review*, 108(9): 2671–2713.

Debortoli, Davide, Mario Forni, Luca Gambetti, and Luca Sala. 2023. "Asymmetric Monetary Policy Tradeoffs." Working paper.

Eggertsson, Gauti B. 2011. "What Fiscal Policy Is Effective at Zero Interest Rates?" *NBER Macroeconomics Annual*.

Farmer, Leland E. 2021. "The discretization filter: A simple way to estimate nonlinear state space models." *Quantitative Economics*, 12(1): 1–45.

Favara, Giovanni, Simon Gilchrist, Kurt F. Lewis, and Egon Zakrajšek. 2016. "Updating the Recession Risk and the Excess Bond Premium."

Fernandez-Villaverde, Jesus, and Juan F. Rubio-Ramirez. 2007. "Estimating Macroeconomic Models: A Likelihood Approach." *Review of Economic Studies*, 74(4): 1059–1087.

Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics*, 7(1): 44–76.

Gonçalves, Sílvia, Ana María Herrera, Lutz Kilian, and Elena Pesavento. 2024. "State-dependent local projections."

Grigoli, Francesco, and Damiano Sandri. 2023. "Monetary Policy and Credit Card Spending." IMF Working Paper, Jan.

Herbst, Edward P., and Benjamin K. Johannsen. 2024. "Bias in local projections." *Journal of Econometrics*, 240(1): .

Husted, Lucas, John Rogers, and Bo Sun. 2020. "Monetary Policy Uncertainty." *Journal of Monetary Economics*, 115 20–36.

Jacobson, Margaret M., Christian Matthes, and Todd B. Walker. 2024. "Temporal Aggregation Bias and Monetary Policy Transmission." March, Working Paper.

Jarocinski, Marek, and Peter Karadi. 2020. "Deconstructing Monetary Policy Surprises—The Role of Information Shocks." *American Economic Journal: Macroeconomics*, 12(2): 1–43.

Jordà, Óscar, Moritz Schularick, and Alan Taylor. 2015. "Leveraged bubbles."

Jordà, Òscar. 2023. "Local Projections for Applied Economics."

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95(1): 161–182.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. Cambridge: Macmillan Cambridge University Press.

Kim, Jinill, and Francisco J. Ruge-Murcia. 2009. "How much inflation is necessary to grease the wheels?" *Journal of Monetary Economics*, 56 365–377.

Kolesár, Michal. 2024. "Many Instruments and Judges." Lecture notes.

Koo, Bonsoo, Seojeong Lee, and Myung Hwan Seo. 2024. "What Impulse Response Do Instrumental Variables Identify?" April, Working paper.

Lee, Dongjae. 2023. "Essays on Macroeconomic Asymmetry and Monetary Policy." Ph.D. dissertation, Georgetown University, Washington, DC.

Mavroeidis, Sophocles. 2021. "Identification at the Zero Lower Bound." *Econometrica*.

McCracken, Michael W., and Serena Ng. 2016. "FRED-MD: A Monthly Database for Macroeconomic Research." *Journal of Business & Economic Statistics*, 34(4): .

Montiel Olea, José Luis, Mikkel Plagborg-Møller, Eric Qian, and Christian K. Wolf. 2024. "Double Robustness of Local Projections and Some Unpleasant VARithmetic."

Oh, Joonseok. 2020. "The Propagation of Uncertainty Shocks: Rotemberg versus Calvo." *International Economic Review*, 61(3): 1097–1113.

Plagborg-Møller, Mikkel, Dake Li, and Christian K. Wolf. 2024. "Local Projections vs. VARs: Lessons From Thousands of DGPs."

Plagborg-Møller, Mikkel, and José Luis Montiel Olea. 2021. "Local Projection Inference is Simpler and More Robust Than You Think." *Econometrica*, 89(4): 1789–1823.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89(2): 955–980.

Ramey, Valerie. 2016. "Macroeconomic Shocks and Their Propagation." In *Handbook of Macroeconomics*. Chap. 2.

Reiter, Michael, and Adrian Wende. 2024. "Generalized Rotemberg Price-Setting." June, Working Paper.

Romer, Christina D., and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review*, 94(4): 1055–1084.

Stein, Jeremy C., and Adi Sunderam. 2018. “The Fed, the Bond Market, and Gradualism in Monetary Policy.” *The Journal of Finance*, 73(3): .

Stock, James H., and Mark W. Watson. 2018. “Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments.” *Volume 128, Issue 610*, 128(610): 917–948.

Tenreyro, Silvana, and Gregory Thwaites. 2016. “Pushing on a String: US Monetary Policy Is Less Powerful in Recessions.” *American Economic Journal: Macroeconomics*, 8(4): 43–74.

Tobin, James. 1972. “Inflation and unemployment.” *American Economic Review*, 62 1–18.

Wieland, Johannes. 2019. “Are Negative Supply Shocks Expansionary at the Zero Lower Bound?” *Journal of Political Economy*, 127(3): 973–1007.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*.: MIT Press.

Appendix

Empirical IRFs using Standard Deviations

In a linear regression, coefficients are the estimated effect of a marginal (size), positive (sign) change. If we normalize our previous definitions by the standard deviation of the coefficient corresponding to this linear "default", we have an alternative formulation of size and sign effects in terms of standard deviations instead of percent change in levels at a given horizon. For example, if $\frac{\hat{\alpha}_{BC}-\hat{\alpha}_{SC}}{\sigma_{SC}} = 3$, the interpretation is that the big cut coefficient amounts to a 3 standard deviations away realization of the small cut coefficient. Additional intuition can be gleaned by noticing that if we instead normalized by the standard deviation of the entire (original) definition, we would simply have a t-statistic. This approach has the advantage of the y-axis having a uniform representation across all outcomes of interest and arguably removes some of the subjectivity implicit in deciding what % constitutes a meaningful effect for a given variable-horizon combination. Put differently, this representation sends a similar signal to the results of a hypothesis test (is there enough evidence from data to infer these parameters are drawn from distinct distributions), but unlike a t-statistic the units lend themselves more to economic meaning (moment of the distribution for our baseline coefficient, rather than a general normal distribution).

Sign-effects in terms of standard deviations are shown in figure 3 and sign effects in figure 4. As previously advanced, these plots are more indicative of the extent of nonlinearity relative to the first set of graphs and we can more strictly and formally adhere to the rules of thumb applied to interpreting the first set of graph. We can infer there is a size effect if the estimate is bounded away from 0, infer there is no size effect if we have a precise estimate near 0, and cannot say anything if we have point estimates far from 0 but wide confidence intervals. It should come as no surprise that the graphs are very similar[11].

[11] What non-trivial differences in shape exist are because the normalization is with respect to the standard deviation from an unpenalized LP to guard against inducing upward bias (in magnitude)

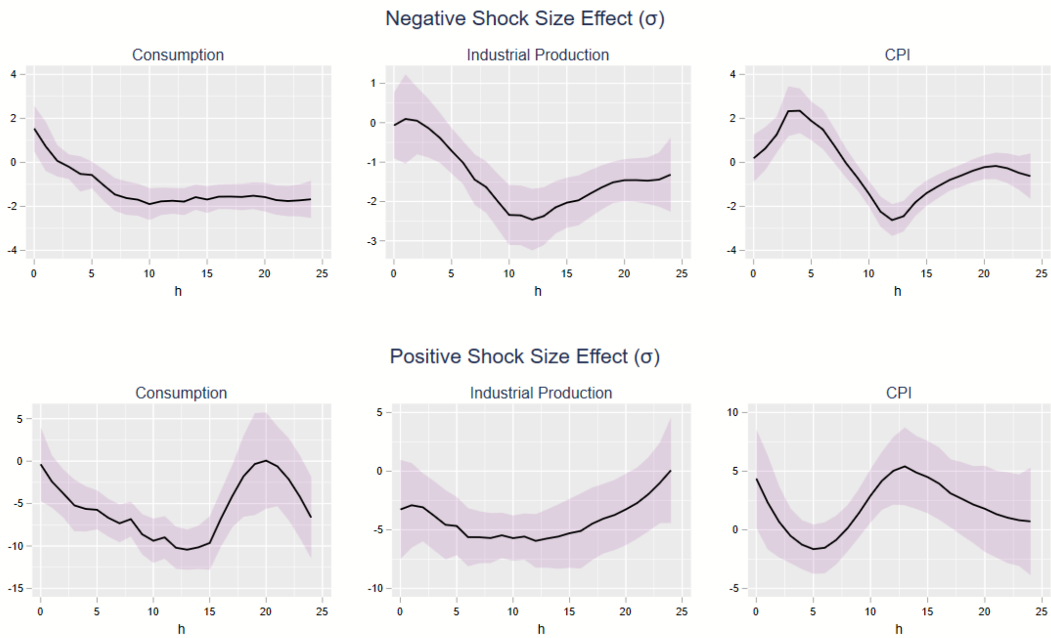


Figure 3

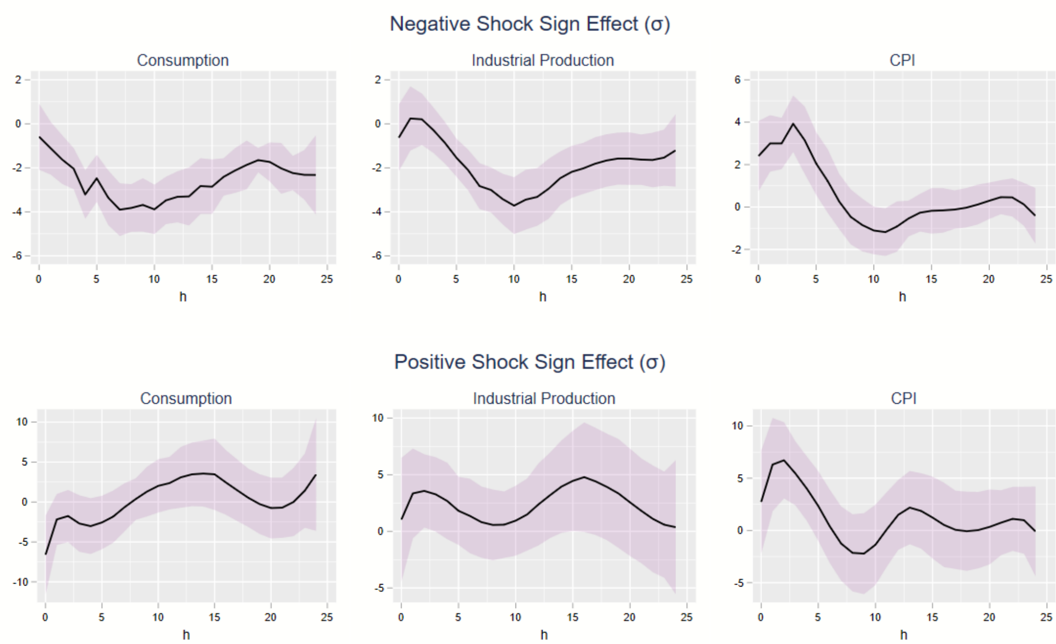


Figure 4

Full Model Equations

Description	Equation	#
Consumption Euler Equation	$1 = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\tau}\frac{R_t}{\Pi_{t+1}\tilde{A}_{t+1}}$	(1)
Definition for Real Wages	$\Delta_t^w = \frac{W_t}{W_{t-1}}\cdot\tilde{A}_t$	(2)
Resource Constraint	$\frac{G_t-1}{G_t}\cdot Y_t + C_t = Y_t(1-\Phi_t^p) + W_t Y_t\cdot\Phi_t^w$	(3)
Wage Equation, Household's problem	$\frac{\chi_h}{\lambda_w}\cdot W_t^{-\tau}C_t^{\tau}Y_t^{\frac{1}{\nu}} + (1-\Phi_t^w)\left(1-\frac{1}{\lambda_w}\right) =$	(4)
	$\Delta_t^{w_{nom}}\cdot\Phi_t'^w - \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\tau}\frac{\Pi_{t+1}R_t}{\tilde{A}_{t+1}}W_{t+1}^2 Y_{t+1}\cdot\Phi_{t+1}'^w$	
Price Equation, Intermediate Firms problem	$(1-\Phi_t^p) + \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\tau}\Phi_{t+1}'^p\frac{\Pi_{t+1}}{\tilde{A}_{t+1}}\cdot Y_{t+1}\tilde{A}_{t+1} = \frac{\mu_t}{\Lambda_t} + \Phi_t'^p\Pi_t$	(5)
Hours Equation	$W_t = (1-\Phi_t^p) + \mu_t$	(6)
Adjustment Costs, Nominal Wages	$\Phi_t^w = \frac{\phi_w}{\psi_w^2}\left(e^{-\psi_w(\Delta_t^{w_{nom}}-\gamma\cdot\pi^*)} + \psi_w(\Delta_t^{w_{nom}}-\gamma\pi^*)-1\right)$	(7)
Adjustment Costs, Prices	$\Phi_t^p = \frac{\phi_p}{\psi_p^2}\left(e^{-\psi_p(\Pi_t-\pi^*)} + \psi_p(\Pi_t-\pi^*)-1\right)$	(8)
Derivative, Adjustment Costs Nominal Wages	$\Phi_t'^w = \frac{\phi_w}{\psi_w}\left(1-e^{-\psi_p\left(\Delta_t^{w_{nom}}-\gamma\pi^*\right)}\right)$	(9)
Derivative, Adjustment Costs to Prices	$\Phi_t'^p = \frac{\phi_p}{\psi_p}\left(1-e^{-\psi_p\left(\Pi_t-\pi^*\right)}\right)$	(10)
Taylor Rule	$R_t = \exp(r_t); \quad r_t = \rho_r r_{t-1} + (1-\rho_r)r_t^* + \sigma_r\varepsilon_r$	(11)
TFP Growth	$\tilde{A}_t = \exp(a_t); \quad a_t = (1-\rho_a)\log\gamma + \rho_a a_{t-1} + \sigma_a\varepsilon_a$	(12)
Government Spending Shocks	$G_t = \exp(g_t); \quad g_t = (1-\rho_g)\log g^* + \rho_g g_{t-1} + \sigma_a\varepsilon_g$	(13)
Price Markup Shock	$\Lambda_t = \exp(\lambda_t); \quad \lambda_t = (1-\rho_p)\log\lambda_{p_{ss}} + \rho_p\lambda_{t-1} + \sigma_p\varepsilon_p$	(14)
Output change	$\Delta_t^y = Y_t/Y_{t-1}$	(15)
Nominal wage change	$\Delta_t^{w_{nom}} = \Delta_t^w\Pi_t$	(16)
Interest rate target	$r_t^* = \log\left(\frac{\gamma}{\beta}\cdot\pi^*\right) + \psi_1\left(\pi_t - \log\pi^*\right) + \psi_2\left(\Delta_t^y + a_t - \log\gamma\right)$	(17)

Technical Details of Estimation

- The priors are largely from Aruoba et al. (2017). Because of the difference in sample period, I scaled down the priors for annualized output growth (μ_y) and inflation (μ_π), as well as β. In fact, it's actually not possible for this model to generate a steady state that matches the data. Namely, steady state interest rates are $\mu_y/4 + \mu_\pi + 400(\beta^{-1} = 1)$. If μ_y and μ_π are picked to match inflation data, you must pick $\beta > 1$ to match interest rate data.

- Related^, the authors start the M-H algorithm at the mode of the linear model and then manually append starting values for the asymmetry parameters and fix the diagonals of the inverse hessian at 4. Asymptotically, there's nothing wrong with this, but I'm going to play around with this to see how sensitive the algorithm is to initial values and priors. From what I've done so far, it seems like there's sensitivity on both fronts. I'm also digging further into the particle filtering procedure.

- A running "diary" of some things I've discovered working with this model/codes can be found here
- For consistency in the comparative static illustrations, it was necessary to make sure this mode line was the same across plots, but that meant the same series of shocks would need to be used for all sets of simulated data, which could paint a misrepresentative picture for a short sample size. So I plotted the median estimate for 100 samples for each parameter group (for simulation i, seed was set to $i = \{1, \ldots, 100\}$).
  - It'd be nice to do this for the Bayesian IRFs, but that would take a month to run. For these, I even had to be parsimonious with what I did – because each loop of the local projections file performs 25*number of outcome variables calls to regress, I randomly selected 10,000 draws of the post burn-in M-H output.

Additional Sensitivity Concerns

A prerequisite to ensure the same shock series can instrument multiple regimes is including the regimes in instrument construction. This ensures that the same shock series can instrument multiple regimes. In other words, this extension results in 4 instruments which each contain mostly 0 values. Recent work by Barnichon and Mesters (2024) shows this is potentially problematic because of the short sample size. Specifically, it makes our approach more sensitive to the possibility that our instruments are correlated with other structural shocks. This is not a problem if there is no "central bank information effect", in other words if people don't act as if the central bank has superior information about the underlying economy. The debate on this issue is ongoing (see, e.g., Acosta, 2023), but the takeaway is whatever validity exists for this criticism is compounded by our chosen approach. Other than using an aforementioned non-parametric approach, another option could be committing to using different instruments for each regime. This alternative is plagued by the fact that (i) even if these shock series are less correlated than they probably should be (Brennan et al., 2024), the remaining collinearity could still be an issue and (ii) more practically, we don't have a unique strong instrument for all 4 regimes. With these drawbacks in mind, this approach still shows the core results hold, albeit weakened in essentially all cases. Another alternative is using one monetary policy surprise series to instrument changes in the fed funds rate in general and then attaching the regime indicators afterwards (see, e.g., Wooldridge, 2002). The drawback of this approach is it somewhat violates the spirit of treating the regimes as fundamentally different and relaxing linearity when possible. Here, the results are largely hold across various surprise series, with the exception that there is more evidence that big cuts have more expansionary effects on output on impact. Barnichon and Mesters (2024) themselves suggest an approach which would require the probability of being in a given regime is strictly greater than 0 at all points in the space-space, which is untenable.

Comparative Static Figures

Size Effects

	Description	Anything Interesting? (all at $h = 0$)	Link
1	$\psi_p \uparrow$	(slightly) amplifies negative size effect for hikes on π at $h = 0$	Figure 5
2	$\psi_p \downarrow$	(slightly) amplifies all $h = 0$ size effects except for hike on π	Figure 6
3	$\psi_w \uparrow$	(slightly) increases the positive size effect for cuts on Y at $h = 0$	Figure 7
4	$\psi_w \downarrow$	No.	Figure 8
5	$\psi_p \uparrow$, $\psi_w \uparrow$	amplifies size effect (-) of cuts on Y, depresses size effect (+) of hikes on Y	Figure 9
6	$\psi_p \downarrow$,$\psi_w \downarrow$	(slightly) amplifies negative size effect for hikes on π at $h = 0$	Figure 10

Sign Effects

	Description	Anything Interesting? (all at $h = 0$)	Link
1	$\psi_p \uparrow$	depressed all $h = 0$ sign effects except for small changes on π	Figure 11
2	$\psi_p \downarrow$	low values reversed the direction of the sign effect for big changes on π.	Figure 12
3	$\psi_w \uparrow$	Depresses small change on Y size effect and amplifies everything else	Figure 13
4	$\psi_w \downarrow$	(slightly) amplified sign effect of big changes on π and small changes on Y	Figure 14
5	$\psi_p \uparrow$, $\psi_w \uparrow$	depressed sign effect of small changes on Y and amplified everything else	Figure 15
6	$\psi_p \downarrow$,$\psi_w \downarrow$	reversed the direction of sign effect for big changes on π	Figure 16

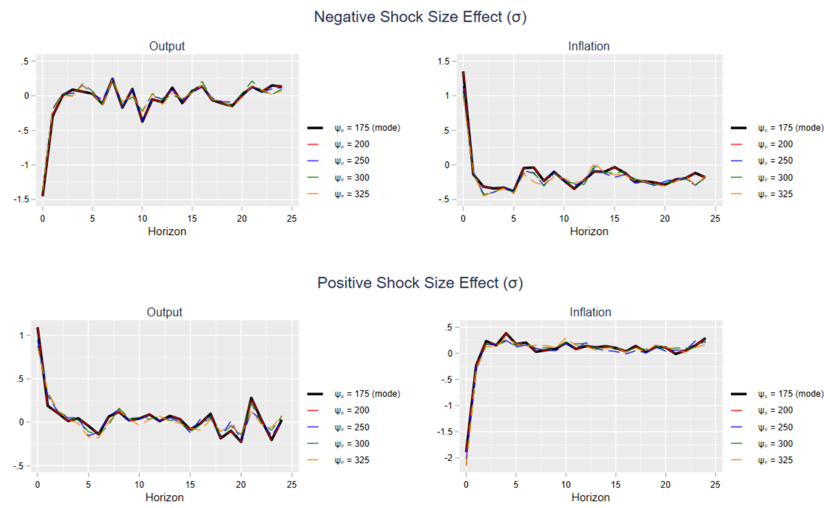


Figure 5: (click to go back to tables)

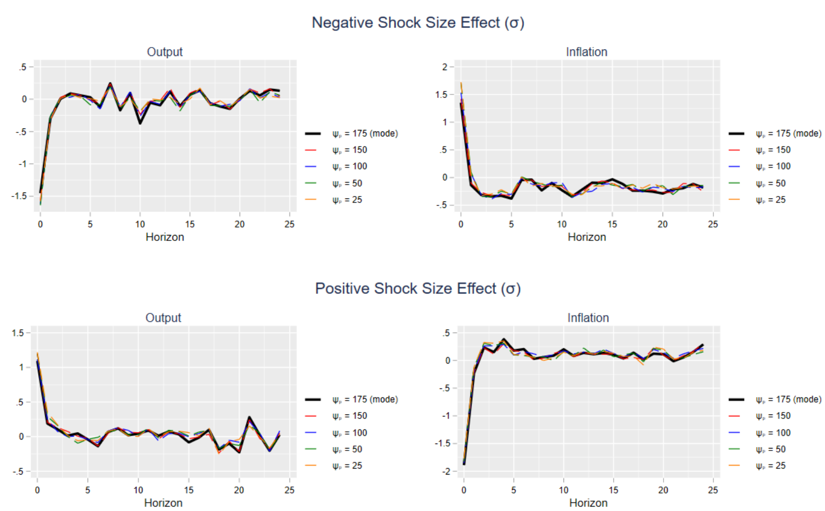


Figure 6: (click to go back to tables)

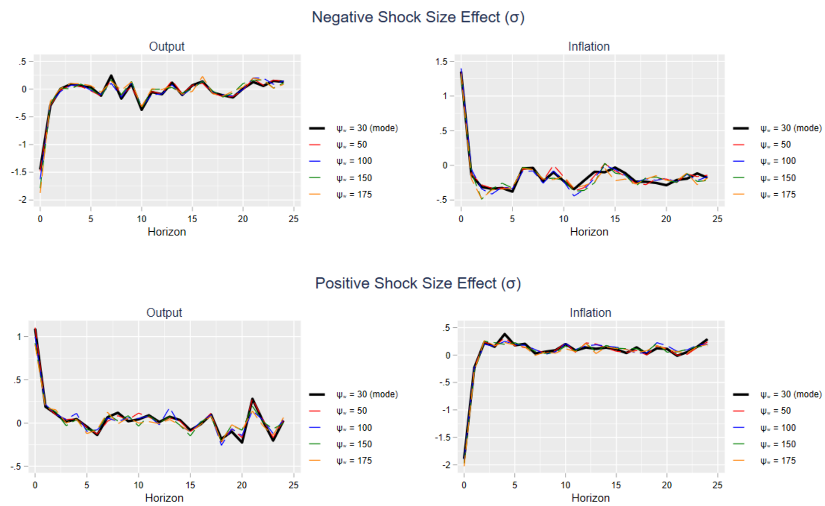


Figure 7: (click to go back to tables)

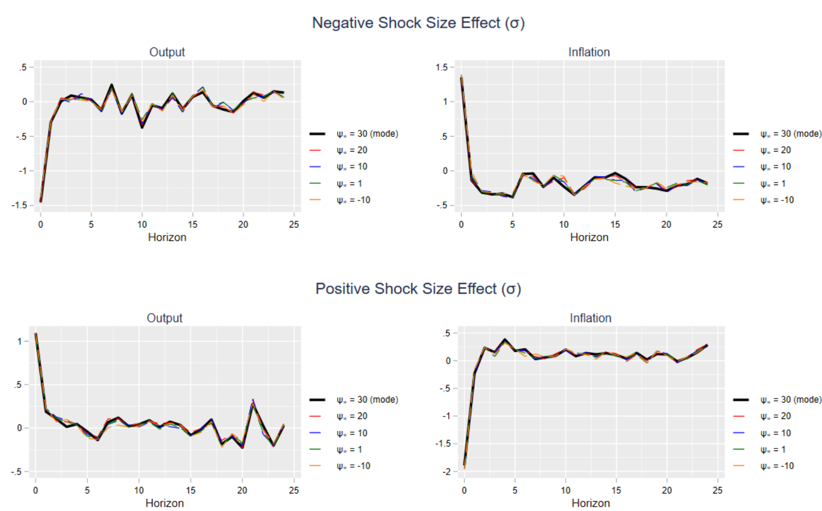


Figure 8: (click to go back to tables)

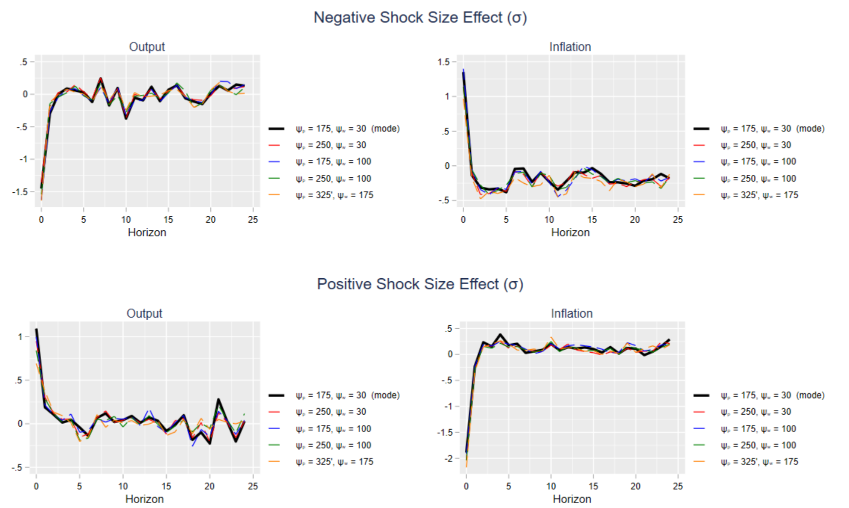


Figure 9: (click to go back to tables)

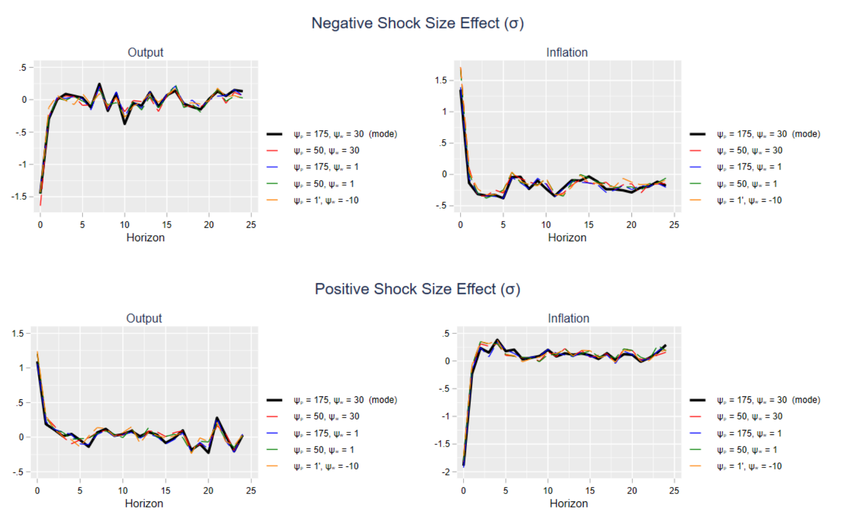


Figure 10: (click to go back to tables)

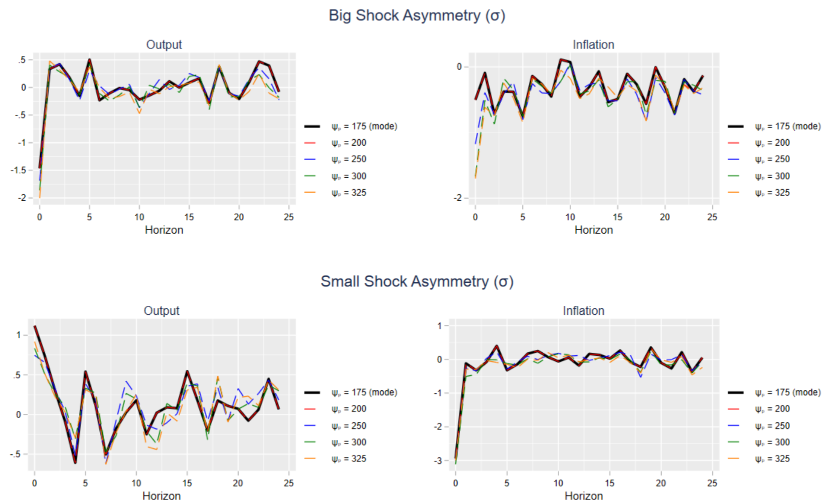


Figure 11: (click to go back to tables)

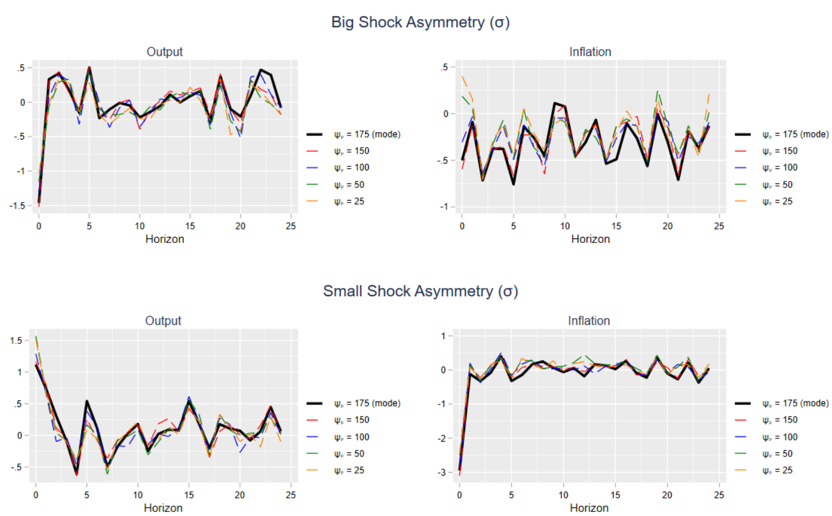


Figure 12: (click to go back to tables)

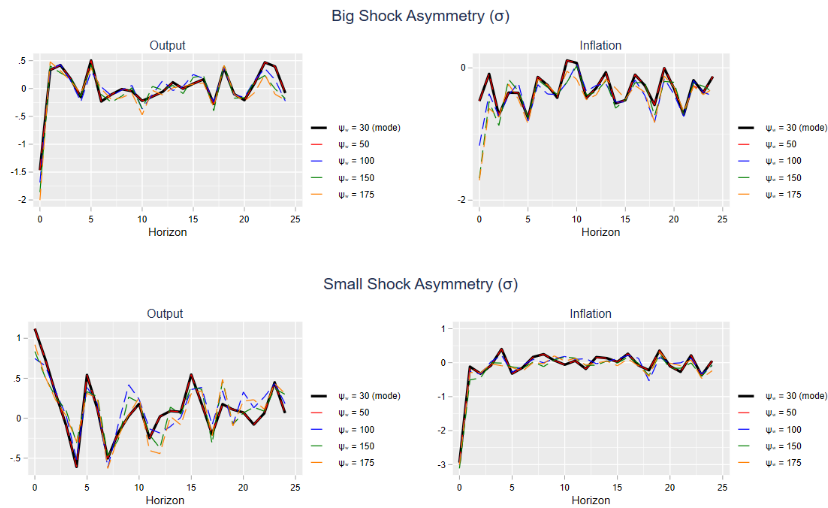


Figure 13: (click to go back to tables)

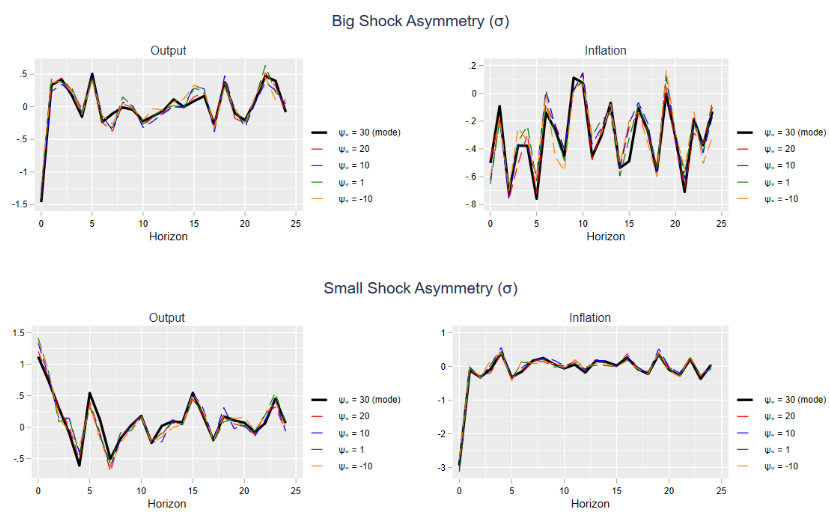


Figure 14: (click to go back to tables)

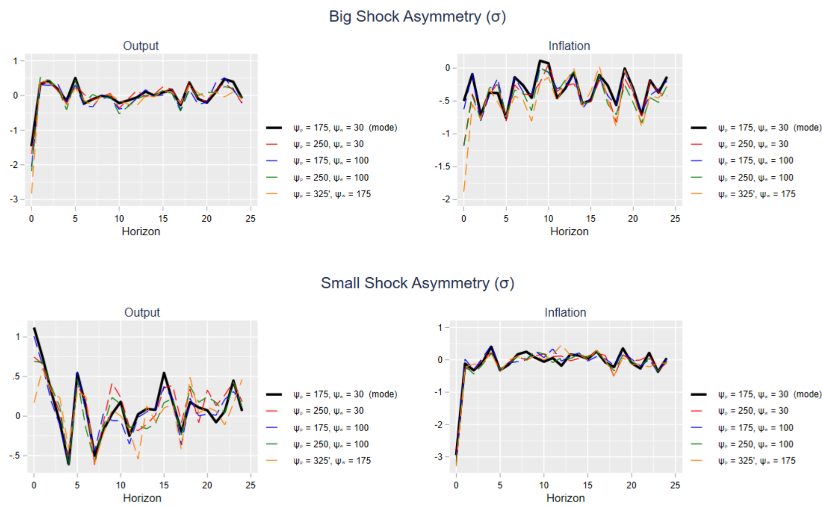


Figure 15: (click to go back to tables)

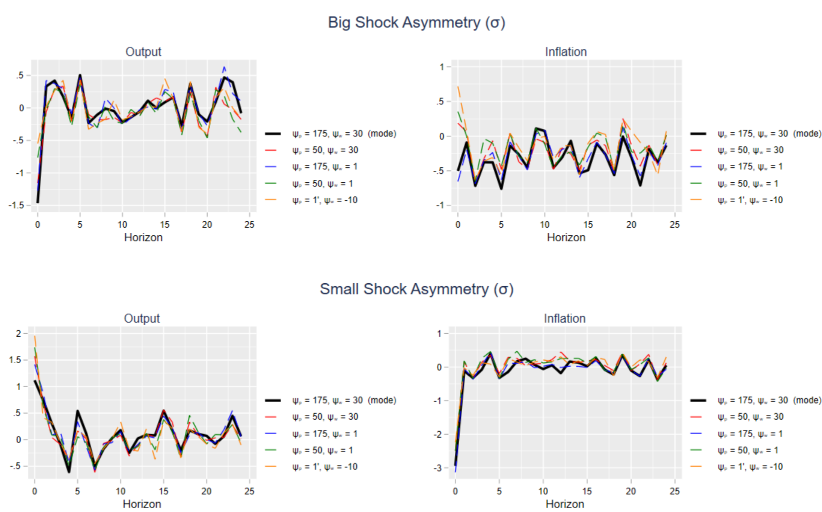


Figure 16: (click to go back to tables)

Rougher Running Notes about Sensitivity

- Issues with Romer and Romer
- ZLB
- Negative weights and IV
- T.E. Heterogeneity
- Zeros
- General Multiple IV

  IV itself can be fragile and extending to multiple IV can only compound these issues. However, in our setting, there's some reason to have faith in the approach. First, a general result about the consistency of 2SLS with multiple instruments applies in the case where instrument strength grows with sample size and we have faith in our shock series exogeneity properties (Kolesár, 2024). On the first point, any weakness in instruments in application is completely ascribable to the short sample size: if we had an infinitely long sample, the shock series we use would amount to a continuous distribution. Anecdotally, we can also point to the fact that the regimes that generally seem more "instrumentable" across various surprise measures have been for small changes (i.e., larger sample sizes). For the second, in the high frequency case this is clearly not an issue. More formally, both the conditional and unconditional expectation of the instruments is 0. Even in states of high uncertainty, where we might consider any outcome a "surprise", the forecast errors[12] and thus the instruments are mean-zero. This may be more of an issue in the Romer and Romer (2004) series, where there is strong evidence to suggest the staff Forecasts of the Fed are more modal and thus this "no forecast errors in expectation" property is less likely to hold Aruoba and Drechsel (2024). However, there is also a separate issue: perhaps we don't expect markets to price as efficiently in times that are perceived as more stable (rational inattention) or that Fed forecasts have a sort of emotional presentation bias based on the state of the economy. In other words, our instruments could suffer from some sort of systematic skew that thretends their validity.

Actually referenced in body of paper and could be expanded on

- Monetary surprise Measure
- Sample Size
- State-Dependence

[12] At least under the efficient markets hypothesis

- Fed Information Effect
- Temporal Aggregation bias