The Optimal Monetary Policy Response to Belief Distortions: Model-Free Evidence*

Jonathan J. Adams† Symeon Taipliadis‡
University of Florida University of Florida

March 3, 2025

Working Paper
Link to Most Current Version

Abstract

Some inflation forecast errors are predictable. Economic theory predicts that these belief distortions affect the business cycle. How should monetary policy respond? We investigate this question with a model-free approach using high-frequency monetary policy shocks and a structural VAR method to identify the effects of shocks to belief distortions. Belief distortion shocks are contractionary: if households become overly pessimistic about inflation, then unemployment and deflation follow. Intuitively, the optimal policy response is to ease. This is most effective with short-term rates; we find that a 1 p.p. increase in the belief distortion is optimally offset by a 0.85 p.p. surprise interest rate decrease. Monetary policy targeting longer-term rates is less effective but also useful.

JEL-Codes: E52, E30, D84, E70
Keywords: Monetary Policy, Belief Distortions, Optimal Policy, Policy Counterfactuals, Shocks

*Replication code available at GitHub.
†email: adamsjonathanj@gmail.com, website: jonathanjadams.com
‡email: s.taipliadis@ufl.edu

1 Introduction

People are poor forecasters. When their forecasts are inconsistent with the full information rational expectation, this is referred to as a belief distortion. In the aggregate, belief distortions are large and vary over time.[1] Theoretical and applied work conclude that shocks driving belief distortions can affect the business cycle. How should monetary policy respond?

In this paper, we evaluate how belief distortions affect the macroeconomy, and calculate the optimal monetary policy response from aggregate time series, without assuming any particular theoretical model. To do so, we apply the McKay and Wolf (2023) method, which allows for the calculation of counterfactual policies from macroeconomic time series in a way that is not subject to the Lucas critique. The method requires two ingredients: estimated impulse response functions (IRFs) to monetary policy shocks and to belief distortions. For monetary policy, we borrow from Swanson (2023), which uses high frequency data around Fed events to identify shocks corresponding to three policy instruments: the target rate, forward guidance, and large scale asset purchases. For belief distortions, we identify shocks in two ways. First, we apply the semi-structural approach developed in Adams and Barrett (2024), which decomposes VAR innovations to find the component that causes empirical forecasts to deviate from the rational expectation. Second, we construct a reduced-form shock, which is simply the statistical innovation to an estimated belief distortion.

Specifically, we study belief distortions over inflation. Household inflation expectations are well known to violate full information rational expectations (FIRE), and the difference varies over the business cycle.[2] The theoretical literature suggests that belief distortion dynamics can be driven by structural shocks and have large effects on the economy.[3] Recent empirical evidence (Ascari et al., 2023; Adams and Barrett, 2024) finds that these shocks have large, robust effects: inflation belief distortions are *contractionary*. When households' forecasts rise too high, prices and real activity fall. These findings constitute an empirical puzzle, at odds with standard theory, which is one reason model-free evidence is so desirable for addressing this policy question.

We learn that monetary policy should ease after a pessimistic belief distortion shock. A shock that causes households to make inflation forecasts that are higher than the FIRE forecast depresses real activity, so intuitively the optimal monetary policy response is stimulative. There are multiple tools that central banks have at their disposal to respond; we

[1]Some examples of recent evidence include Bianchi et al. (2022), Bianchi et al. (2024), and Farmer et al. (2024).

[2]See D'Acunto et al. (2023) or D'Acunto et al. (2024) for recent literature surveys.

[3]Some examples focusing on belief distortions over different variables include Ascari et al. (2023) (inflation), Candia et al. (2023) (exchange rates), Bhandari et al. (2024) (unemployment), and Maenhout et al. (2025) (GDP growth)

find that traditional interest rate policy is most effective. Specifically, we estimate that interest rates should fall roughly one-to-one in response to a structural shock that raises the inflation belief distortion. The optimal response is also negative – but half as large – for the reduced-form shock. And while the target rate is the most effective single tool, incorporating stimulative forward guidance and asset purchases can do even better.

Our results contribute to a mostly theoretical literature studying optimal monetary policy without FIRE. Adams (2024) proves that the belief distortion is a sufficient statistic for the the optimal policy response to deviations from FIRE; in a New Keynesian model, the policy rule for interest rates is increasing in inflation and income belief distortions. Focusing on belief distortions to guide policy is valuable for central banks, because it provides guidance without needing to specify the precise mechanism by which FIRE fails. Many such mechanisms abound, and give conflicting policy prescriptions.[4] A few recent papers explicitly study monetary policy with exogenous shocks to expectations, which most closely matches our structural approach; examples include Ascari et al. (2023) and Neri (2023), whose models predict that the real economy contracts after shocks to inflation expectations over the short and long run, respectively.

2 Model-Free Method

Our empirical strategy consists of four steps: measuring belief distortions, estimating the impulse response functions (IRFs) to belief distortion shocks, estimating the IRFs to monetary policy shocks, and calculating the optimal policy counterfactual. In this section, we outline each step in detail.

2.1 Identification of Shocks to Belief Distortions

In order to study the optimal monetary policy response to belief distortions, we need to estimate the IRFs to belief distortion shocks. To do so, we consider two methods: a *structural shock*, and a *reduced-form shock*.

The *structural shock* is an identification strategy motivated by theory. We follow the approach of Adams and Barrett (2024); the identifying assumption is that structural belief shocks are the only shocks to cause forecasts to depart from rational expectations. For example, if individuals respond to productivity shocks with rational expectations, but also independently exhibit stochastic belief distortions, this method correctly identifies the latter shock from the former.

[4]This literature is enormous. Some recent examples studying optimal monetary policy when FIRE fails due to behavioral constraints include Hommes et al. (2019) (heuristics), Gabaix (2020) (cognitive discounting), and Iovino and Sergeyev (2023) (level-k thinking). Examples with information frictions include Angeletos and La'O (2019), Benhima and Blengini (2020), and Angeletos et al. (2020).

To identify the effects of structural shocks, we estimate an n-dimensional VAR including the surveyed forecast $f^{y,h}$:

$$\begin{pmatrix} f_t^{y,h} \\ x_t \end{pmatrix} = \sum_{j=1}^{J} B_j^s \begin{pmatrix} f_{t-j}^{y,h} \\ x_{t-j} \end{pmatrix} + w_t^s \tag{1}$$

where w_t^s are reduced-form innovations related to structural shocks ε_t^s by

$$w_t^s = A^s \varepsilon_t^s$$

We encode the structural belief shock in the first entry of the shock vector ε_t^s, so this method identifies $A_{1\cdot}^s$, the first column of the impact matrix A^s.[5]

The alternative *reduced-form shock* is the statistical innovation in the belief distortion. In other words, it is the residual in a regression of belief distortion $d_t^{y,h}$ on its lags and other controls; it appears in the first row dimension in the estimated innovation w_t. This shock is reduced-form: it may be driven by any number of structural shocks at time t that cause measured expectations to depart from the rational expectation.

In this second case, we estimate another n-dimensional VAR, replacing the surveyed forecast from equation (1) with the belief distortion $d_t^{y,h}$. In the next section, we describe how the belief distortion is measured. With its inclusion, the VAR is

$$\begin{pmatrix} d_t^{y,h} \\ x_t \end{pmatrix} = \sum_{j=1}^{J} B_j^r \begin{pmatrix} d_t^{y,h} \\ x_t \end{pmatrix} + w_t^r \tag{2}$$

again, w_t^r are reduced-form innovations related to structural shocks ε_t^r by

$$w_t^r = A^r \varepsilon_t^r$$

for some matrix A^r.

The reduced-form shock is also not an *uncorrelated* shock: it covaries with the other residuals, to which monetary policy may already be responding. But the belief distortion shock is linearly independent from other dimensions of the VAR. If monetary policymakers ignore belief distortions, then they ignore a business cycle driver that demands a policy response. Studying the reduced-form shock captures the additional policy response required by this typically overlooked dimension.

[5]This method contrasts with a traditional approach that orders expectation shocks first in a Cholesky decomposition. For this causal ordering to identify a belief distortion shock, the necessary assumption is that the only shock that can affect forecasts contemporaneously is an exogenous expectation shock. Estimation with this causal ordering method typically finds that inflation expectation shocks are expansionary (Leduc et al., 2007; Clark and Davig, 2011).

2.2 Measuring Belief Distortions

The *belief distortion* is the difference between average expectations in the economy, and the appropriate full information rational expectation. Define the belief distortion $d_t^{y,h}$ over quantity y at horizon h by

$$d_t^{y,h} \equiv f_t^{y,h} - re_t^{y,h} \tag{3}$$

where $f_t^{y,h}$ denotes the time t average forecast of y_{t+h}, and $re_t^{y,h}$ denotes the corresponding rational expectation. For some information set Ω_t, the rational expectation is the conditional expectation of y_{t+h} given Ω_t:

$$re_t^{y,h} = \mathbb{E}_t[y_{t+h}|\Omega_t]$$

The average forecast $f_t^{y,h}$ is taken from survey data, but the rational expectation $re_t^{y,h}$ must be estimated. We estimate $re_t^{y,h}$ by selecting an information set Ω_t, and projecting y_{t+h} on the variables in Ω_t. Specifically, we estimate a regression of the following form:

$$y_{t+h} = \sum_{j=0}^{J} \left(\alpha_j f_{t-j}^{y,h} + \beta_j x_{t-j} \right) + \upsilon_{t+h} \tag{4}$$

where x_t is a vector of macroeconomic variables, and υ_{t+h} is the forecast error. Including $f_t^{y,h}$ and its lags guarantees that our rational expectation includes all information contained in the surveyed forecasts. We take the predicted value from regression (4) as the rational expectation $re_t^{y,h}$.

2.3 Calculating Optimal Policy

To calculate optimal monetary policy, we follow the method pioneered in McKay and Wolf (2023). To do so, we require 3 ingredients: IRFs to belief distortions, IRFs to monetary policy shocks, and a welfare criterion.

The IRFs to belief distortion shocks are given by the $n \times 1$ impulse response functions $\phi_s(k)$ to shock s. The IRFs to n_m monetary policy shocks are given by the $n \times n_m$ impulse response function $\phi_m(k)$. We keep these functions abstract for the moment, but Section 3 will describe how we estimate them.

The welfare criterion is a function of the IRFs to the various shocks. In our baseline approach, we use a welfare criterion that depends on the horizon-H conditional variance of unemployment u_t and inflation π_t, with weighting parameter λ:

$$\mathcal{W}_s = \lambda V_s^u(H) + (1-\lambda) V_s^\pi(H) \tag{5}$$

where $V_s^u(H)$ and $V_s^\pi(H)$ denote the variance of unemployment and inflation, respectively, that is due to all shocks at horizons no more than H. The variance of quantity x is

$$V_s^x(h) = \sum_{k=0}^{H} Var(x_{t+k}|s_t) = \sum_{k=0}^{H} \left(\phi_w^x(k)\right)^2 Var(w_t)$$

where $\phi_s^x(k)$ is the IRF of x to the shock s at horizon k. Therefore, it is possible to express the welfare criterion as a function of the IRFs to the various shocks. For example, the welfare loss due to a shock s over H horizons is

$$\mathcal{W}_s = \sum_{k=0}^{H} \left(\lambda \left(e_u \phi_w(k)\right)^2 + (1-\lambda)\left(e_\pi \phi_w(k)\right)^2\right) \tag{6}$$

where e_u and e_π are the basis vectors selecting unemployment and inflation from the vector $\phi_w(k)$.

Counterfactual policies can be studied by constructing alternative impulse response functions to minimize the welfare loss (6). The central insight from McKay and Wolf (2023) is that this can be done in a way that satisfies the Lucas critique by manipulating only the covariance between the shock s and monetary policy shocks. Specifically, this is done by constructing a counterfactual rule for monetary policy shock m:

$$m_t = \psi s_t$$

where ψ is an $n_m \times 1$ vector. Therefore we can construct the alternative impulse response function $\phi_\psi(k)$ by adding the IRF to the monetary policy shock to the IRF to the belief distortion shock:

$$\phi_\psi(k) = \phi_s(k) + \phi_m(k)\psi \tag{7}$$

With this approach, the welfare loss to shock w can be written as a function of the policy vector ψ by

$$\mathcal{W}_w(\psi) = \sum_{k=0}^{H} \left(\lambda e_u \left(\phi_w(k) + \phi_m(k)\psi\right)\right)^2 + \left((1-\lambda)e_\pi \left(\phi_w(k) + \phi_m(k)\psi\right)\right)^2$$

and the optimal policy vector ψ minimizes $\mathcal{W}_w(\psi)$.[6]

The crucial assumption to estimate policy counterfactuals that do not violate the Lucas critique is that the shocks are unanticipated. This is almost certainly true for the high frequency monetary policy shocks. And this is also true if the assumptions used to identify

[6]Appendix A.2 explains the calculation in detail. We follow McKay and Wolf (2023) and truncate the IRFs at $H = 60$ months, but consider alternatives in Section 5.

the structural shock hold. But in the case of the reduced-form shock, it is important that the VAR is close to fundamental. For this to be a reasonable assumption, we will need to use a VAR that is not especially small, and includes data on yields and forecasts.

3 Data

Our baseline VAR model is as standard as possible. As in Gertler and Karadi (2015), Bauer and Swanson (2023), and others, we include the following monthly series: the log consumer price index (CPI), the log of industrial production (IP), the Gilchrist and Zakrajšek (2012) excess bond premium (EBP), and the 2 year treasury yield (TREAS). We also include two additional variables necessary for our procedure: the unemployment rate (UNEMP), and the log of the 1-year-ahead implied CPI forecast or belief distortion, depending on whether we are estimating model (1) or (2). We apply the Akaike information criterion (AIC), which selects 7 monthly lags for our baseline model. We consider alternative lag lengths in Section 5.

When measuring belief distortions, we must construct a CPI forecast from surveyed forecasts of inflation, and also estimate the rational expectation. To calculate implied CPI forecasts we use the monthly median household forecast from the Michigan Survey of Consumers. The survey reports an inflation forecast in percentage points, so we construct a 1-year-ahead CPI forecast by:

$$f_t^{CPI,12} = (1 + f_t^{\pi,12}) \times CPI_t$$

The Michigan Survey data begin in 1978 which restricts our dataset for the main VARs to January 1978 – May 2024.[7]

When estimating the rational expectation, we estimate a regression specified by equation (4). On the right-hand side, we include contemporaneous and up to four lags of the variables from our baseline VAR (CPI, industrial production, unemployment, 2-year Treasury yield, excess bond premium, and the 1-year-ahead implied CPI forecast) as well as the PPI commodity index, 10-year Treasury yield, and the Wu and Xia (2016) shadow fed funds rate. The adjusted R-squared is 0.9989 implying the estimated rational expectation is an accurate predictor.

We source the monetary policy shocks from Swanson (2023), which uses high-frequency data around Federal Reserve events to derive three distinct shocks: the target rate (FFR), forward guidance (FG), and large-scale asset purchases (LSAP). Roughly speaking, these

[7]While the Survey has included a question on expected inflation since 1948, the question had a qualitative format where respondents needed to report if they expect inflation to go up, down, or stay the same. The survey revised this question in 1966 in bins format.

shocks are identified from their effects on short, medium, and long-term rates, respectively. Notice that while monetary policy shocks dataset does not start until 1988, this does not restrict the data used to estimate the VAR model. The effects of monetary policy shocks are estimated from their impact on the reduced-form residuals of the model.[8]

4 Results

This section describes the results of our optimal policy analysis. We first present the impulse response functions to belief distortion shocks and to monetary policy shocks. Then using these estimates, we calculate optimal policy responses.

4.1 Impulse Response Functions

Figure 1 reports estimated IRFs for the two variables that are relevant to the optimal policy analysis – CPI and unemployment – as well as the belief distortion itself. Both the reduced-form and structural shock are scaled to cause 1-year-ahead CPI forecasts to depart from their rational expectations by 1 percentage point.[9] For example, if the average inflation forecast is 3%, and the rational expectation is 2%, this would be a one percentage point belief distortion.

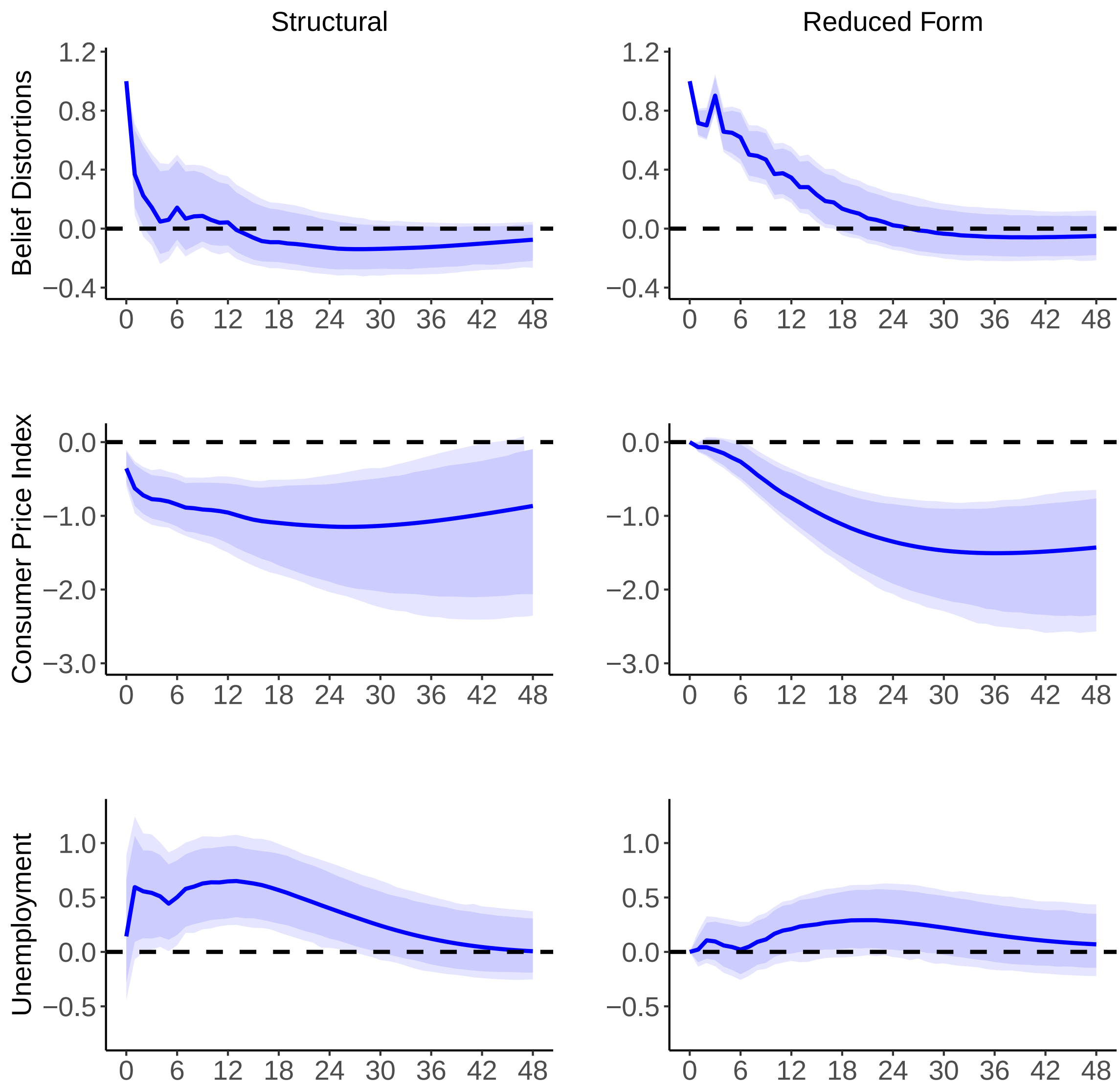
Both belief distortion shocks are contractionary. Unemployment rises and prices fall. Unreported in the figure, unemployment and the bond premium rise, industrial production declines, and Treasury yields fall, likely reflecting that the Fed responds to the contractionary shock by lowering interest rates.[10] These patterns are consistent with those estimated by Adams and Barrett (2024).

The two shocks also differ in noticeable ways. After a structural shock, the belief distortion disappears within half a year, while the reduced-form shock has a long, persistent effect on the belief distortion, which takes two years to decay. The reduced-form shock also has delayed effects on the economy: prices take several months to fall and unemployment takes at least half a year to rise. In contrast, the structural shock causes a contraction almost immediately. Finally, the structural shock has a much larger effect on the real economy, causing a 0.6 p.p. peak increase in unemployment, twice as large as that of the reduced-form shock. Given that these shocks have different quantitative effects, they will require different policy responses.

[8] Furthermore, we follow Swanson's baseline approach and truncate the series after February 2020 to exclude the COVID-19 pandemic.

[9] To be more precise, the shock increases the implied belief distortion on log CPI by 0.01, which we refer to as 1 percentage point.

[10] Appendix B reports these additional impulse response functions.



Notes: The figure shows the impulse responses to a positive belief distortion shock that causes 1-year-ahead CPI reported forecasts to depart from their rational expectations by 1 percentage point. Bootstrapped confidence intervals on 1,000 bootstrap replications are reported at the 95% and 90% level.

Figure 1: Impulse Responses to Belief Distortion Shocks

To conduct the optimal policy calculation, we also need to estimate the IRFs for the monetary policy shocks.[11] Figure 2 plots the impulse responses to the target rate, forward guidance, and large-scale asset purchase shocks. Each shock is normalized such that the 2-year Treasury yield increases by 1 percentage point on impact. The first row includes

[11] We do so with SVAR-IV (Stock and Watson, 2012; Mertens and Ravn, 2013; Gertler and Karadi, 2015). Appendix A.1 gives details.

the yield to illustrate how differently identified shocks have different effects on interest rate policy over the medium run. Otherwise, the figure plots the IRFs to inflation and unemployment, which are the only functions needed for the optimal policy calculation.

The policy shocks have distinct effects, reflecting the different dimensions of monetary policy that they capture. The Target Rate shock resembles a textbook interest rate shock: it raises yields in the short run, reducing inflation and real activity. The Forward Guidance shock also raises yields, but over a long horizon. In the short run it is inflationary and expansionary, but in the medium run the effects reverse. The LSAP shock increases yields in the first year, but then reduces them in the following years; the effect is strongly contractionary.

4.2 Optimal Monetary Policy Response

We apply the methodology outlined in Section 2 to the data described in Section 3 in several ways. We calculate how optimal monetary policy responds to belief distortion shocks using both reduced-form and structural methodologies. We first calculate the optimal response for each type of monetary policy shock (Target, FG, LSAP) separately. Then, we consider the optimal policy response using pairs of these policies, and finally using all three together. Table 1 presents these results, displaying the appropriate entries in the estimated optimal policy vector ψ.

The optimal policy rules are largely intuitive, and in most cases are qualitatively consistent between the reduced-form and structural methodologies. Therefore we discuss the structural shocks first, and return to the differences with reduced-form shocks below.

The optimal target rate response to a structural belief distortion shock is as expected: the shock is contractionary, so the optimal response is to loosen monetary policy. The coefficient -0.845 implies that if the shock increases the inflation belief distortion by 1 percentage point, then interest rates should be reduced enough to lower the two year yield by 0.845 percentage points. This policy is nearly one-to-one, because a target rate shock increases unemployment by roughly the same amount as a belief distortion shock. So offsetting the belief distortion with monetary policy mostly negates the distortion to the real economy.

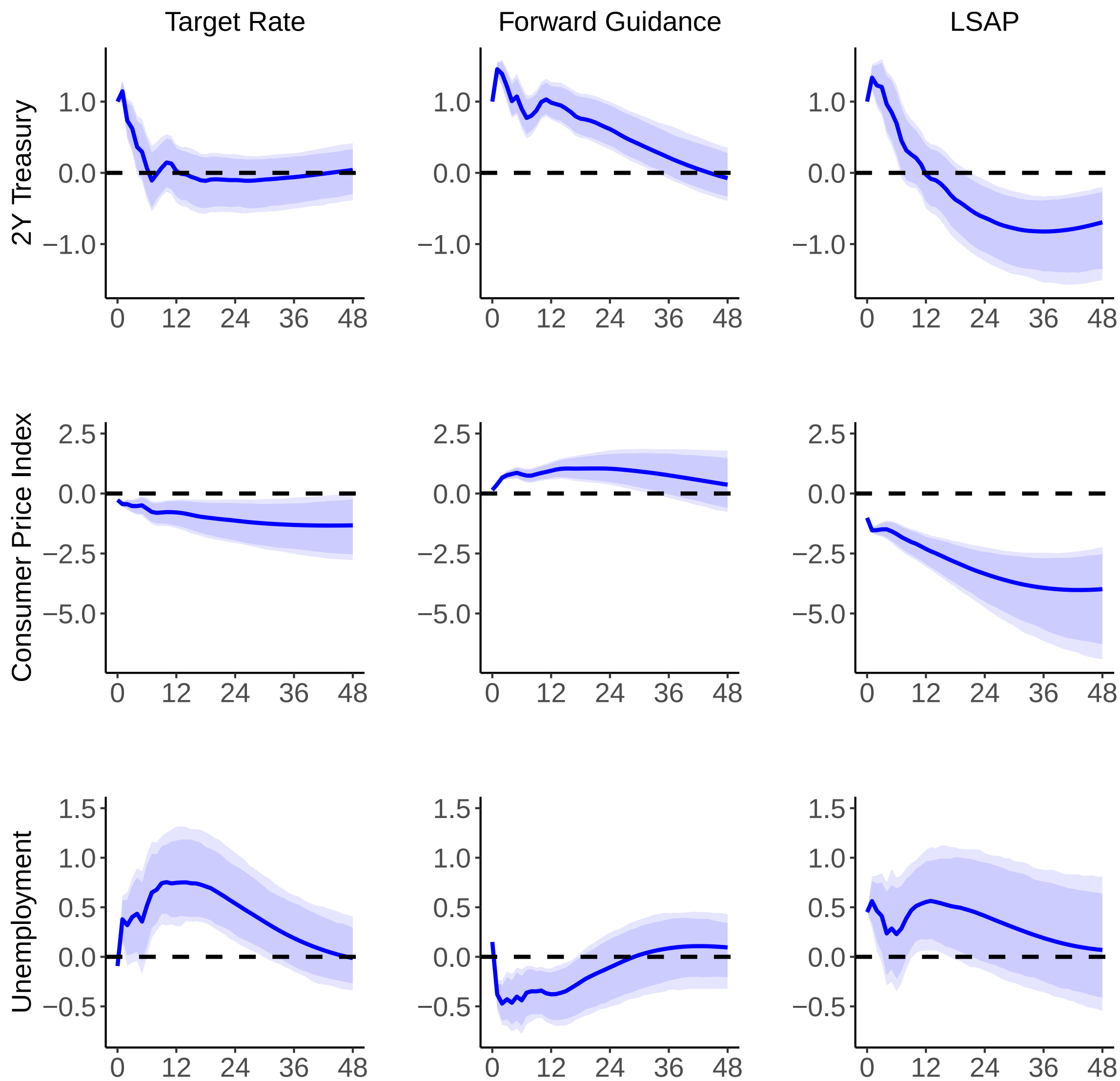
This offsetting policy rule can be seen clearly in Figure 3, which plots the IRFs to the belief distortion shocks under three scenarios: the baseline result, a counterfactual under the optimal policy that uses the target rate alone, and a counterfactual under the optimal policy using all three monetary tools. In the lower right plot, the belief distortion shock would ordinarily increase unemployment by more than 5% for many months, but when the target rate policy is used, then unemployment increases by half as much and returns rapidly.

The deflationary impact is also moderated, reducing the magnitude of the impact by

	Structural Methodology				Reduced-form Methodology			
	Target	FG	LSAP	R^2	Target	FG	LSAP	R^2
Independent Tools	-0.845			0.948	-0.435			0.806
	(0.301)				(0.344)			
		1.417		0.731		0.24		0.082
		(0.463)				(0.404)		
			-0.949	0.847			-0.415	0.322
			(0.354)				(0.338)	
Pairwise Tools	-0.678	0.411		0.973	-0.518	-0.26		0.872
	(0.25)	(0.34)			(0.478)	(0.479)		
		0.71	-0.65	0.946		0.202	-0.401	0.38
		(0.429)	(0.256)			(0.348)	(0.309)	
	-0.649		-0.262	0.962	-0.418		-0.042	0.808
	(0.283)		(0.233)		(0.315)		(0.247)	
All Tools	-0.459	0.429	-0.283	0.988	-0.544	-0.28	0.05	0.875
	(0.208)	(0.27)	(0.179)		(0.443)	(0.442)	(0.212)	

Notes: Columns 1-4 (5-8) report estimates from the structural (reduced-form) methodology. Target stands for Target Rate tool, FG for Forward Guidance, and LSAP for Large Scale Asset Purchases. R^2 statistics reported in columns 4 and 8 measure the share of the welfare loss that is eliminated by adopting the optimal policy. Bootstrapped standard errors from 1,000 iterations are reported in parentheses. Regressions in Panel A use single monetary tools, while in Panel B and Panel C use tools in pairs and a combination of all three tools, respectively.

Table 1: Optimal Monetary Policy Response

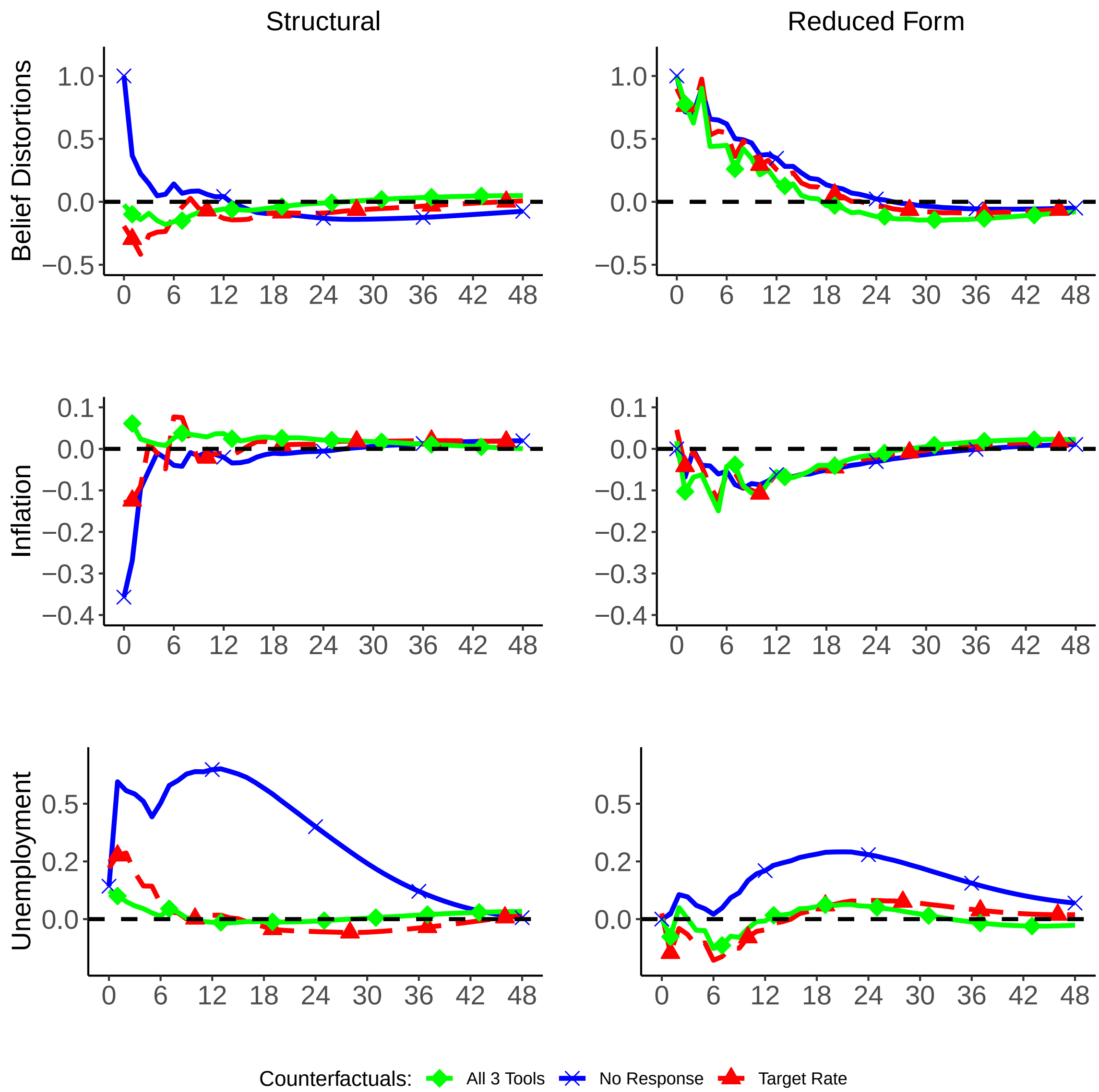


Notes: The figure shows the impulse responses to a monetary policy shock that raises the 2Y Treasury Yields by 100 basis points on impact. Bootstrapped confidence intervals on 1,000 bootstrap replications are reported at the 95% and 90% level. Each column reports impulse responses of different shock type (target rate, forward guidance, and large-scale asset purchases).

Figure 2: Impulse Responses to Monetary Policy Shocks

more than half. This can be seen in the second row of Figure 3, which plots the inflation IRFs under the three scenarios. We plot inflation instead of the CPI here because inflation is the object that enters the welfare criterion (5). The inflation IRF is simply the first difference of the CPI IRF plotted in Figure 1.

The target rate is close to the perfect policy instrument to address the structural belief



Notes: The figure shows the effects of belief distortion shocks under three scenarios: no monetary policy reaction, optimal reaction with target rate policy, and optimal reaction when all three monetary tools are available (target rate, forward guidance and large scale asset purchases). The first column reports estimated responses from the reduced-form methodology and the second column reports estimated responses from the structural methodology.

Figure 3: Counterfactual Monetary Policy Response to Belief Distortions

distortion shock; Figure 3 shows that it eliminates most of the resulting variation in unemployment and inflation. We summarize this effectiveness with the R^2 statistic in Table 1. That the $R^2 = 0.948$ indicates that the optimal target rate policy eliminates $\sim 95\%$ of the welfare loss due to the shock. The FG and LSAP policies are less successful, but still effective with R^2 statistics of 0.731 and 0.847 respectively.[12] The reason that the target rate is the most effective policy is because its CPI and unemployment IRFs (Figure 2) most closely match the shape of the belief distortion IRFs (Figure 1). For example, forward guidance is less effective at moderating the unemployment response because the effect of FG on unemployment reverses sign after 2 years, and LSAP is less effective because its impact on unemployment decays almost immediately.

A policy response coordinating across multiple instruments is always more effective than using any of the instruments individually. Table 1 also reports these results, showing the optimal rule for pairwise combinations in the second block of rows, and the rule for all three in the final block of rows. In all cases, the coefficients on the contractionary instruments (target rate and LSAP) are negative, while the coefficient on the expansionary forward guidance instrument is positive. Figure 3 shows that the use of all three instruments is even more effective, almost entirely eliminating the response of inflation and unemployment to structural belief distortion shock.

The first row of Figure 3 provides another useful lesson: this negation of the effects of the belief distortion shock is achieved by directly negating the belief distortion entirely. Under the optimal policy counterfactual, the structural shock has almost no effect on the inflation belief distortion. This is not a necessary result of the method. Rather, it occurs because expansionary monetary policy reduces belief distortions. This is mainly because forecasts are inelastic: a sudden interest rate reduction increases future inflation, but expectations do not move one-for-one.

How general are the conclusions from the structural shock? The identifying assumption is that the shock is the single dimension of the data along which forecasts deviate from FIRE. In a model with entirely exogenous belief distortions it exactly identifies the shock, but it could be that belief distortions are endogenously determined from behavioral expectations or other frictions. Adams and Barrett (2024) show that even under many of these alternatives, the structural shock is still contractionary. And even if belief distortions are entirely endogenous, we can learn from the reduced-form shocks.

The optimal policy response to the reduced-form shock is similar, but with quantitative differences. After a reduced-form shock, the optimal policy is also to tighten the target rate, but by less (-0.435) because the effect on unemployment is roughly half that of the structural shock (Figure 1). Across the board, policy is less effective at moderating the reduced-form shock: the R^2 statistics in Table 1 are smaller, the coefficients are less statistically significant, and the counterfactual impulse response functions in Figure 3 are further from zero. This is because the shapes of the IRFs to the reduced-form shock are

[12] This result is thematically similar to evidence from D'Acunto et al. (2022), who find that forward guidance is relatively ineffective at managing household inflation expectations.

not as closely spanned by the monetary policy IRFs as in the case of the structural shock. Specifically, the long delay from the reduced-form shock's impact to the peak unemployment response is difficult to replicate with any combination of the unemployment IRFs in Figure 2. The FG and LSAP instruments are particularly bad at matching this shape; thus in Table 1 their coefficients are small when all tools are included, and when the target rate is excluded the R^2 statistics are especially small. Still, the main conclusion remains: expansionary interest rate policy is an effective response to the belief distortion shock

5 Robustness Checks

This section presents a variety of robustness checks. In each case, we calculate the optimal policy response of the target rate shock to a structural or reduced-form belief distortion shock.

Our first check is to consider alternative policy objectives. In the baseline, monetary policy has a dual mandate: minimize both unemployment and inflation volatility. When the welfare criterion places all the weight on unemployment ($\lambda = 1$) the results are nearly unchanged from the baseline, but when all the weight is on inflation ($\lambda = 0$) the optimal response to a structural shock is even more aggressive, but the response to a reduced-form shock is near zero. This is because the reduced-form belief distortion shock creates a very persistent deflation, which is nearly orthogonal to the inflationary effect of the target rate shock.

In the baseline we used the AIC to select a 7-lag structure for the VARs. We also consider 1 quarter and 1 year worth of lags. In general, shorter lag lengths yield more aggressive monetary policy responses. We also adopt an alternative estimation of the rational expectation component of the belief distortion with additional lags; the results are mostly unchanged from the baseline.

Many other choices related to timing had limited effects on our conclusions. Changing the maximum IRF horizon from the 60 month baseline to 24 or 120 months had little effect on the optimal policy. This was also true when we ended the sample on December 2019 before COVID-19.

Finally, we considered an entirely different type of monetary policy shock. The Swanson (2023) shocks are identified from high-frequency data around monetary policy events. We also employed a narrative approach: the Aruoba and Drechsel (2024) shocks are identified residuals to the decisions predicted by natural language processing of Federal Reserve staff documents. The narrative shocks have smaller effects on the real economy, so the optimal policy coefficients are larger in magnitude.

	Structural		Reduced-form	
	Target	R^2	Target	R^2
Baseline Model	-0.845	0.948	-0.435	0.806
	(0.301)		(0.344)	
Inflation Targeting ($\lambda = 0$)	-1.006	0.718	0.075	0.004
	(0.263)		(0.284)	
Employment Targeting ($\lambda = 1$)	-0.843	0.955	-0.439	0.857
	(0.339)		(0.354)	
VAR with 3 lags	-1.046	0.892	-0.524	0.894
	(0.456)		(0.341)	
VAR with 12 lags	-0.599	0.861	-0.186	0.275
	(0.262)		(0.229)	
Belief Distortion estimation	-	-	-0.378	0.773
with 12 lags	-		(0.237)	
Excl. COVID-19 Era	-0.828	0.736	-0.677	0.793
	(0.239)		(0.475)	
24-Month Truncation of	-0.874	0.954	-0.378	0.764
Welfare Objective	(0.366)		(0.291)	
120-Month Truncation of	-0.822	0.913	-0.397	0.596
Welfare Objective	(0.303)		(0.377)	
Small VAR	1.861	0.745	0.504	0.317
	(0.926)		(0.408)	
Aruoba-Drechsel Monetary	-2.092	0.705	-0.912	0.845
Policy Shock	(1.059)		(0.942)	

Notes: Columns 1 and 3 report estimates of the optimal policy response for the target rate shock to a belief distortion estimated with the structural or reduced-form approach respectively. The R^2 statistics reported in columns 2 and 4 measure the share of the welfare loss that is eliminated by adopting the optimal policy. Bootstrapped standard errors from 1,000 iterations are reported in parentheses.

Table 2: Robustness Checks

6 Conclusion

This paper demonstrates that belief distortions exert contractionary effects on the macroeconomy, increasing unemployment and reducing inflation. Using the McKay and Wolf (2023) method to evaluate model-free counterfactuals, we estimate the optimal monetary policy response to these distortions. Broadly: central banks should ease when inflation expectations become overly pessimistic. Short-term interest rates are the most effective tool.

Our findings emphasize the value of belief distortions as a sufficient statistic for policy, circumventing the need for specific models or assumptions about why FIRE fails. This approach is straightforward to implement and addresses well-known macroeconomic inefficiencies. We studied household inflation belief distortions, but this method will be valuable to apply to expectations by other agents or regarding other variables.

References

Adams, Jonathan J, "Optimal Policy Without Rational Expectations: A Sufficient Statistic Solution," *University of Florida mimeo*, 2024.

Adams, Jonathan J. and Philip Barrett, "Shocks to inflation expectations," *Review of Economic Dynamics*, October 2024, *54*, 101234.

Angeletos, George-Marios and Jennifer La'O, "Optimal Monetary Policy with Informational Frictions," *Journal of Political Economy*, June 2019, p. 704758.

_ , Luigi Iovino, and Jennifer La'O, "Learning over the business cycle: Policy implications," *Journal of Economic Theory*, November 2020, *190*, 105115.

Aruoba, S. Borağan and Thomas Drechsel, "Identifying Monetary Policy Shocks: A Natural Language Approach," May 2024.

Ascari, Guido, Stefano Fasani, Jakob Grazzini, and Lorenza Rossi, "Endogenous uncertainty and the macroeconomic impact of shocks to inflation expectations," *Journal of Monetary Economics*, April 2023.

Bauer, Michael D. and Eric T. Swanson, "A Reassessment of Monetary Policy Surprises and High-Frequency Identification," *NBER Macroeconomics Annual*, May 2023, *37*, 87–155. Publisher: The University of Chicago Press.

Benhima, Kenza and Isabella Blengini, "Optimal Monetary Policy when Information is Market-Generated," *The Economic Journal*, May 2020, *130* (628), 956–975.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, "Survey Data and Subjective Beliefs in Business Cycle Models," *The Review of Economic Studies*, May 2024, p. rdae054.

Bianchi, Francesco, Sydney C. Ludvigson, and Sai Ma, "Belief Distortions and Macroeconomic Fluctuations," *American Economic Review*, July 2022, *112* (7), 2269–2315.

_ , _ , and _ , "What Hundreds of Economic News Events Say About Belief Overreaction in the Stock Market," April 2024.

Candia, Bernardo, Olivier Coibion, and Yuriy Gorodnichenko, "The macroeconomic expectations of firms," in Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, January 2023, pp. 321–353.

Clark, Todd E. and Troy Davig, "Decomposing the declining volatility of long-term inflation expectations," *Journal of Economic Dynamics and Control*, July 2011, *35* (7), 981–999.

D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber, "What do the data tell us about inflation expectations?," in Rüdiger Bachmann, Giorgio Topa, and Wilbert van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, January 2023, pp. 133–161.

D'Acunto, Francesco, Daniel Hoang, and Michael Weber, "Managing Households' Expectations with Unconventional Policies," *The Review of Financial Studies*, April 2022, *35* (4), 1597–1642.

_ , Evangelos Charalambakis, Dimitris Georgarakos, Geoff Kenny, Justus Meyer, and Michael Weber, "Household Inflation Expectations: An Overview of Recent Insights for Monetary Policy," May 2024.

Farmer, Leland E., Emi Nakamura, and Jón Steinsson, "Learning about the Long Run," *Journal of Political Economy*, October 2024, *132* (10), 3334–3377. Publisher: The University of Chicago Press.

Gabaix, Xavier, "A Behavioral New Keynesian Model," *American Economic Review*, August 2020, *110* (8), 2271–2327.

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

Gilchrist, Simon and Egon Zakrajšek, "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, June 2012, *102* (4), 1692–1720.

Hommes, Cars, Domenico Massaro, and Matthias Weber, "Monetary policy under behavioral expectations: Theory and experiment," *European Economic Review*, September 2019, *118*, 193–212.

Iovino, Luigi and Dmitriy Sergeyev, "Central Bank Balance Sheet Policies Without Rational Expectations," *The Review of Economic Studies*, November 2023, *90* (6), 3119–3152.

Leduc, Sylvain, Keith Sill, and Tom Stark, "Self-fulfilling expectations and the inflation of the 1970s: Evidence from the Livingston Survey," *Journal of Monetary Economics*, March 2007, *54* (2), 433–459.

Maenhout, Pascal J., Andrea Vedolin, and Hao Xing, "Robustness and dynamic sentiment," *Journal of Financial Economics*, January 2025, *163*, 103953.

McKay, Alisdair and Christian K Wolf, "What can time-series regressions tell us about policy counterfactuals?," *Econometrica*, 2023, *91* (5), 1695–1725.

Mertens, Karel and Morten O Ravn, "The dynamic effects of personal and corporate income tax changes in the United States," *American economic review*, 2013, *103* (4), 1212–1247.

Neri, Stefano, "Long-term inflation expectations and monetary policy in the euro area before the pandemic," *European Economic Review*, May 2023, *154*, 104426.

Runkle, David E, "Vector autoregressions and reality," *Journal of Business & Economic Statistics*, 1987, *5* (4), 437–442.

Stock, James H. and Mark W. Watson, "Disentangling the Channels of the 2007-2009 Recession," May 2012.

Stock, James H and Mark W Watson, "Identification and estimation of dynamic causal effects in macroeconomics using external instruments," *The Economic Journal*, 2018, *128* (610), 917–948.

Swanson, Eric T., "The Macroeconomic Effects of the Federal Reserve's Conventional and Unconventional Monetary Policies," August 2023.

Wu, Jing Cynthia and Fan Dora Xia, "Measuring the macroeconomic impact of monetary policy at the zero lower bound," *Journal of Money, Credit and Banking*, 2016, *48* (2-3), 253–291.

Online Appendix

A Implementation Details

This appendix describes our methods in detail.

A.1 Estimating IRFs

Throughout the appendix, we denote with y_t the vector of the VAR model. In our baseline results, this takes the form $y_t = [FCST_t, CPI_t, IP_t, UNEMP_t, EBP_t, TREAS_t]'$ when we use the structural methodology and $y_t = [BD_t, CPI_t, IP_t, UNEMP_t, EBP_t, TREAS_t]'$ when we use the reduced-form methodology. These variables are the log of implied 1-year-ahead CPI forecasts (FCST), the log of consumer price index (CPI), the log of industrial production index (IP), the unemployment rate (UNEMP), excess bond premium (EBP), 2Y treasury yields (TREAS), and belief distortions (BD).

A.1.1 Estimating IRFs from Belief Distortion Shocks

In our baseline methodology, our IRFs to the structural belief distortion shock are estimated using the coefficients from model 1. Let $\phi(k)$ be the IRF at horizon $k \geq 0$. Then,

$$\phi_s(k) = \sum_{j=1}^{J} \mathbb{I}_{k \geq j} B_j^s \phi_s(k-j) \tag{8}$$

where $\mathbb{I}_{k \geq j} = 1$ if $k \geq j$ and 0 otherwise. The impact effect of all structural shocks is $\phi_s(0) = A^s \epsilon_t^s$ where A^s is the $n \times n$ impact matrix and ϵ_t^s is a column vector of n shocks. By construction, the first shock is the structural belief distortion shock. To identify the impact matrix we follow Adams and Barrett (2024). That is, CPI forecasts are ordered first in our VAR and we impose the assumption that only a structural belief distortion shock can cause contemporaneous belief distortions.

We scale the IRFs to show the effect of a 1 p.p. positive belief distortion shock. Since the structural methodology does not explicitly include a belief distortion variable, we generate the belief distortion response as

$$\tilde{\phi}_{s,bd}(k) = \phi_{s,f}(k) - \phi_{s,CPI}(k+12)$$

where, $\phi_{s,f}$ is the first row of ϕ_s corresponding to the response of 1-year-ahead implied CPI forecasts and $\phi_{s,CPI}$ is the row of ϕ_s corresponding to the dynamic response of CPI. The IRF $\phi_{s,CPI}(k+12)$ also represents the 12-month-ahead rational expectation, so subtracting it from the current survey forecast gives the belief distortion.

Similarly, when we use the reduced-form methodology, our IRFs are estimated using the coefficients from model 2. In this case:

$$\phi_r(k) = \sum_{j=1}^{J} \mathbb{I}_{k \geq j} B_j^r \phi_r(k-j) \tag{9}$$

and $\phi_r(0) = A^r \epsilon_t^r = w_t^r$ where w_t^r is a column vector of n reduced-form residuals; the first element of w_t^r is the reduced-form belief distortion shock. Again, we scale the IRFs to show the effect of a 1 p.p. positive belief distortion shock.

A.1.2 Estimating IRFs from Monetary Policy Shocks

We derive the IRFs from monetary policy shocks using a two-stage least squares regression as proposed in the SVAR-IV literature (Stock and Watson, 2012; Mertens and Ravn, 2013; Gertler and Karadi, 2015). First, we derive the reduced-form residuals (either from model 1 or 2), w_t using the full sample.

In our baseline model, we use the Swanson (2023) orthogonal shock series as instruments for the target rate shock (FFR), forward guidance (FG) and large scale asset purchases (LSAP), aggregated at the monthly level. To ensure that the lead-lag exogeneity condition in Stock and Watson (2018) is not violated, we regress each of these instruments on 1 lag of the variables included in the VAR model and use the residuals as our instruments.

The first observation of the three instruments is February 1988 and the last is December 2023. We only merge the reduced-form residuals in this time interval. In between this time period, on months when there was no monetary policy announcement, we assign a zero value on all three instruments.

We then regress the residuals on the three monetary policy instruments. That is, for each individual response variable i, we run an OLS regression of the form:

$$w_{i,t} = \alpha_i + \omega_i m_t + \eta_{i,t} \tag{10}$$

where $m_t = \{FFR_t, FG_t, LSAP_t\}'$ and ω_i is the $i-th$ row of an impact matrix ω. We scale these shocks such that each instrument causes the 2Y treasury yields to increase by 1 p.p. on impact; that is $\omega_4 = \{1, 1, 1\}$.

It follows that the IRFs of the endogenous variables in the VAR model to monetary policy shocks are given by:

$$\phi_m(k) = \sum_{j=1}^{J} \mathbb{I}_{k \geq j} B_j^s \phi_m(k-j) \tag{11}$$

where $\mathbb{I}_{k \geq j} = 1$ if $k \geq j$ and 0 otherwise, and $\phi_m(0) = \omega \epsilon_t^m$. Here, ϵ_t^m is a column vector of m policy-type shocks.

A.2 Calculating Optimal Policy

This appendix describes how we use least squares to calculate the optimal policy rule from estimated impulse response functions.

First, encode the unemployment and inflation impulse responses to a belief distortion shock s as vectors $\vec{\phi}_{s,u}$ and $\vec{\phi}_{s,\pi}$, and stack them into the single vector $\vec{\phi}_m$. Do the same with the impulse responses to policy shock m, stacking $\vec{\phi}_{m,u}$ and $\vec{\phi}_{m,\pi}$ into the single vector $\vec{\phi}_m$.

Define the welfare matrix by

$$W \equiv \begin{pmatrix} \lambda I & 0 \\ 0 & (1-\lambda)I \end{pmatrix}$$

where I is the identity matrix with dimensions matching the length of $\vec{\phi}_{r,u}$ and $\vec{\phi}_{r,\pi}$. The W matrix is useful, because $\vec{\phi}_r' W \vec{\phi}_r$ gives the welfare loss in equation (6) as estimated by the VAR without any counterfactual policy.

For a policy coefficient ψ, the counterfactual stacked IRF vector is $\vec{\phi}_r + \vec{\phi}_m \psi$. Thus the objective is to solve

$$\min_{\psi} \left(\vec{\phi}_r + \vec{\phi}_m \psi\right)' W \left(\vec{\phi}_r + \vec{\phi}_m \psi\right)$$

This can be rewritten as a least squares problem:

$$\min_{\psi} \left(W^{\frac{1}{2}}\vec{\phi}_r + W^{\frac{1}{2}}\vec{\phi}_m \psi\right)' \left(W^{\frac{1}{2}}\vec{\phi}_r + W^{\frac{1}{2}}\vec{\phi}_m \psi\right)$$

Finally, the least squares problem is solved by projecting $W^{\frac{1}{2}}\vec{\phi}_r$ onto $-W^{\frac{1}{2}}\vec{\phi}_m$. Thus the optimal policy rule ψ is given by

$$\arg\min_{\psi} \mathcal{W}_w(\psi) = -(\vec{\phi}_m' W \vec{\phi}_m)^{-1} \vec{\phi}_m' W \vec{\phi}_r$$

Having estimated the IRFs for belief distortions (see equations 8, 9) and monetary policy shocks (see equation 11) and the optimal response coefficients, ψ, we generate the counterfactual IRF by equation (7).

A.3 Bootstrapped Standard Errors

We run a residual bootstrap approach (Runkle, 1987) to generate confidence bands (see plots) and standard errors (see tables). The bootstrap consists of 1,000 replications. Let y_t denote the vector of variables from the original sample. We use y_t to run the VAR model of 1 (or 2) and retain the estimated coefficients $\hat{B}_j \ \forall j \in \{1, ...J\}$ (including the intercept) and the reduced-form residuals $\hat{w}_t$. We also use these coefficients and the original-sample variables to compute the fitted values $\hat{y}_t$. For each replication b, we randomly resample the residuals with replacement and add them to the fitted values of the original sample to create artificial response variables:

$$y_t^b = \hat{y}_t + \hat{w}_k \tag{12}$$

where $k \in \{1, ..., T\}$ is an index denoting the randomly selected time from the sample. We then run the following regression:

$$y_t^b = \sum_{j=1}^{J} \hat{B}_j^b y_{t-j} + \hat{w}_t^b \tag{13}$$

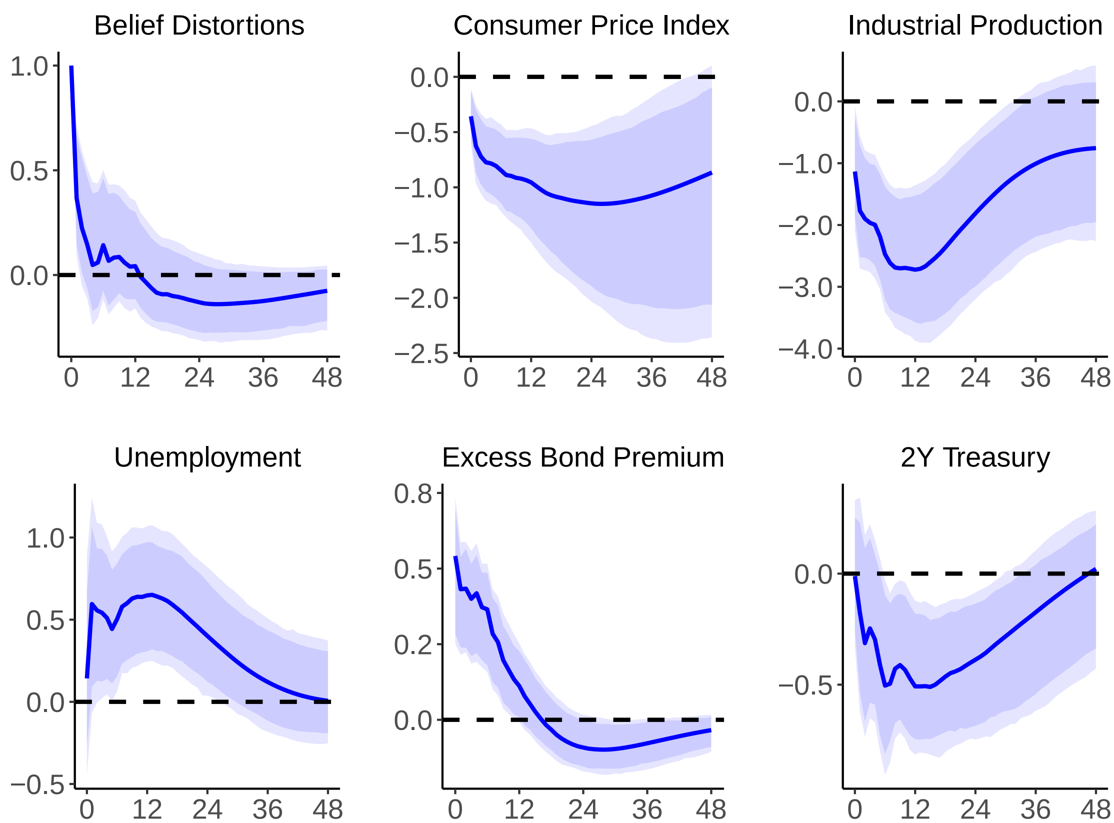
Notice that the regressors remain unchanged in all bootstrap replications and are determined by the original sample. We collect all coefficients from the replication and then move on to the next replication where we draw a new random sample of residuals and repeat the aforementioned process.

Finally, we use $\hat{B}_j^b \ \forall j \in \{1, ...J\}, \ \forall b \in \{1, ..., 1000\}$ to replicate all IRFs 1000 times and therefore estimate 1,000 vectors of optimal responses to belief distortion shocks, each of whom we denote by ψ^b.

When replicating the IRFs to monetary policy shocks, we follow Swanson (2023) and keep the ω impact matrix fixed throughout all replications to avoid introducing noise from the first-stage estimation of the impact effects of monetary policy shocks. Assessing the uncertainty of the effects of monetary policy shocks is not the focus of this paper. In contrast, the impact of our newly estimated belief distortion shocks does feature bootstrapped uncertainty.

B Additional Plots

This appendix includes plots with further detail, complementing those in the main text.

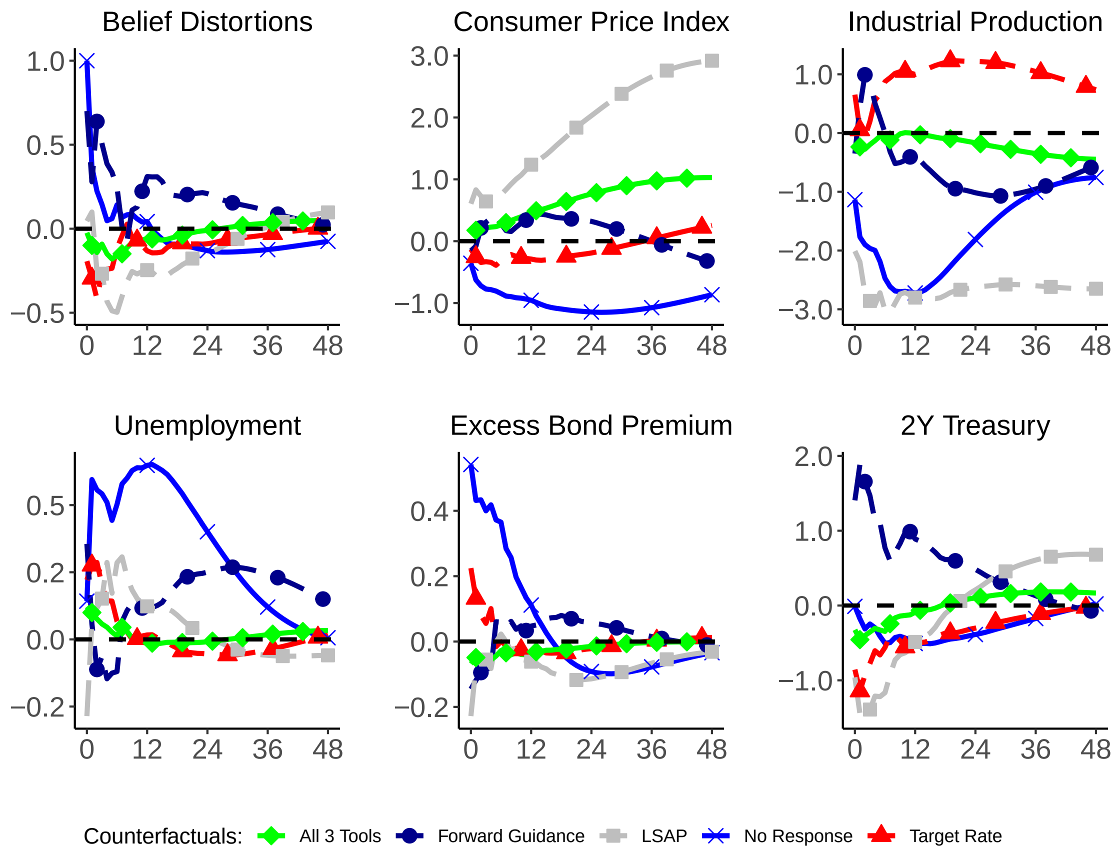


Notes: The figure shows all impulse responses to a positive structural belief distortion shock that causes 1-year-ahead CPI reported forecasts to depart from their rational expectations by 1 percentage point. Bootstrapped confidence intervals on 1,000 bootstrap replications are reported at the 95% and 90% level.

Figure 4: All Impulse Responses to Structural Belief Distortion Shocks

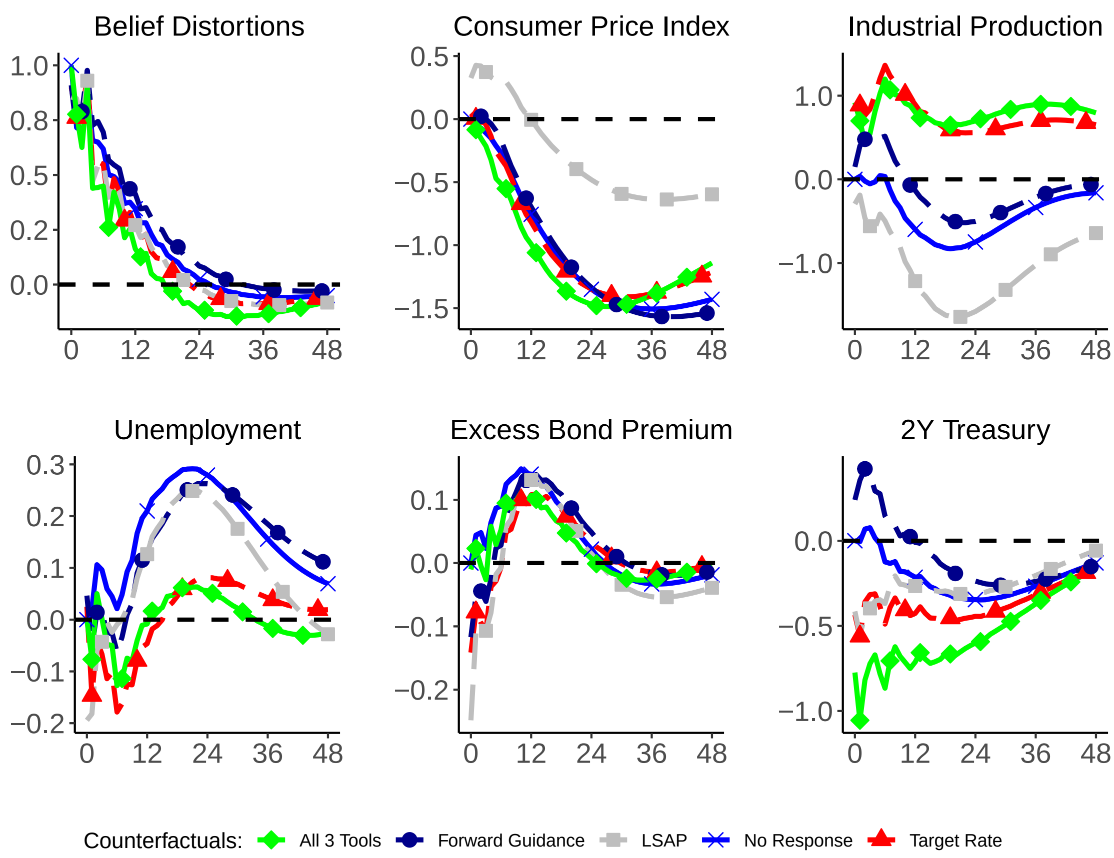
Notes: The figure shows all impulse responses to a positive reduced-form belief distortion shock that causes 1-year-ahead CPI reported forecasts to depart from their rational expectations by 1 percentage point. Bootstrapped confidence intervals on 1,000 bootstrap replications are reported at the 95% and 90% level.

Figure 5: All Impulse Responses to Reduced-form Belief Distortion Shocks



Notes: The figure shows the effects of belief distortion shocks under five scenarios: no monetary policy reaction, optimal reaction with the target rate, optimal reaction with forward guidance, optimal reaction with LSAP, and optimal reaction with all three monetary tools simultaneously.

Figure 6: Counterfactual Monetary Policy Response to Structural Belief Distortions (Detailed)



Notes: The figure shows the effects of reduced-form belief distortion shocks under different scenarios of monetary policy responses.

Figure 7: Counterfactual Monetary Policy Response to Reduced-form Belief Distortions (Detailed)

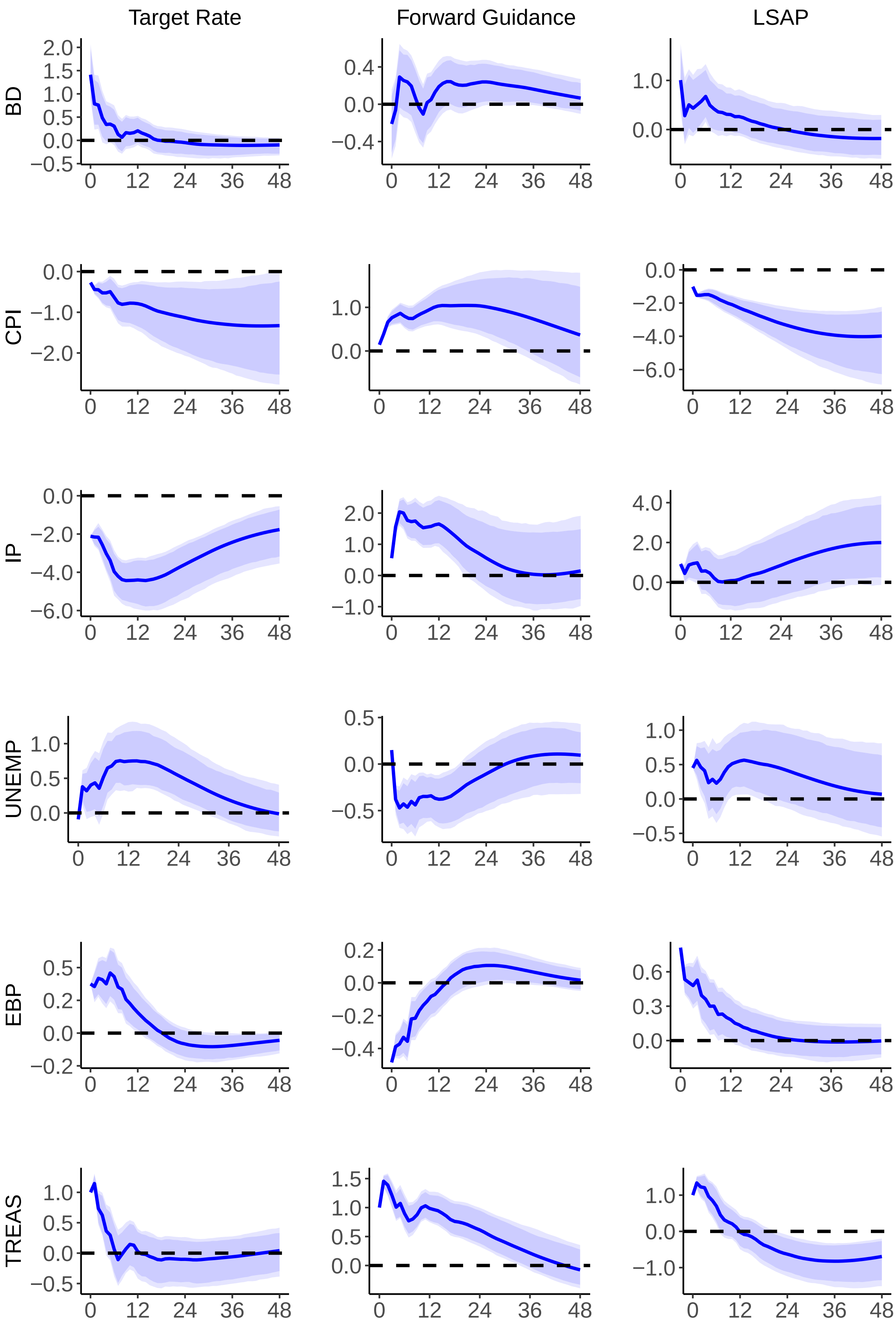


Figure 8: Impulse Responses to Monetary Policy Shocks (Detailed)