WORKING PAPER

Unpacking the Forward Guidance Puzzle

NORGES BANK RESEARCH

11 | 2023

AUTHORS:
SEHYOUN AHN

NORGES BANK

NORGES BANK
WORKING PAPER
11 | 2023

UNPACKING THE FORWARD
GUIDANCE PUZZLE

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles på e-post: servicesenter@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail: servicesenter@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-8143 (online)
ISBN 978-82-8379-291-1 (online)

Unpacking the Forward Guidance Puzzle

SeHyoun Ahn*

June 25, 2023

Abstract

I prove that in any linearized general equilibrium model with existence and uniqueness of equilibrium, any present response to a future shock always becomes weaker the further away the shock is. There is no forward guidance puzzle in any sensible general equilibrium model *if* forward guidance is a one-time shock in the future. The difference from the literature arises from the underappreciated preceding peg of interest rate. The standard forward guidance exercise, of a drop in the interest rate in the future with the preceding peg of interest rate, is a series of explosive deviations disguised as a small, one-time shock in the future. I unpack the standard forward guidance exercise in different ways, and show that there is nothing puzzling about the forward guidance puzzle.

*This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of the Norges Bank. I would like to thank Thor Andreas Thorvaldsen Aursland, Paul Ho, Martin Blomhoff Holm, Greg Kaplan, Mathis Mæhlum, Gisle Natvik, Frank Smets, and participants of Norges Bank Spring Institute for their valuable comments.
Contact: sehyoun.ahn@norges-bank.no

1 Introduction

It is commonly believed that the standard DSGE based policy models at many central banks give unreasonable predictions to future changes. This belief arose from the forward guidance puzzle. The forward guidance puzzle can be summarized as "consumption and inflation respond stronger to a future interest rate change than a present one in the standard New Keynesian (NK) models."[1] This is completely unintuitive, and a vast literature tried to salvage our models from this puzzle. However, as will be shown below, this simplified summary is too strong compared to the actual exercise underlying the puzzle.

The standard exercise for the forward guidance in the literature is to compute the impulse response function to an interest rate drop at time T while holding the real rates fixed at the steady-state level of interest rate prior to time T (see figure 2a). I will refer to this standard exercise as the *pegged* forward guidance exercise to highlight the preceding peg. In the pegged forward guidance exercise, the forward guidance is a promise of the entire path, not just the small drop in the future. Hence, the path of the promised interest rate implicitly defines a series of monetary policy deviations/shocks. Del Negro et al. (2023) and McKay et al. (2016) mention these implicitly defined monetary policy shock*s* (plural), but do not explicitly show them. Most do not even mention them. We will unpack the pegged forward guidance puzzle by looking at one policy shock at a time instead.

Our exercise is a one-time shock in the future while allowing the endogenous determination of the interest rates prior to the shock. I will refer to this exercise as the *endogenous* forward guidance exercise to highlight the endogenous dynamics prior to the shock. For example, in a model where the central bank follows a Taylor rule, the central bank will continue to follow their usual Taylor rule prior to the promised drop in the interest rate in the future. By abstracting away the preceding peg, simplified summaries in the literature suggest that there is an *endogenous* forward guidance puzzle. Is there an endogenous puzzle? The answer is no. In fact, a more general result can be mathematically proven for endogenous exercises. I prove that in a linearized general equilibrium model with existence and uniqueness of equilibrium,[2] any present response to a one-time shock in the future always becomes weaker the further away the shock is. Most, if not all, policy models used at central banks have existence and uniqueness of the equilibrium, so policy models can not have the endogenous forward guidance puzzle including the model used in Del Negro et al. (2023).

Without the endogenous forward guidance puzzle, why do our models exhibit the pegged forward guidance puzzle? The simple answer is that the series of implicitly defined monetary policy shocks grows exponentially in the anticipation horizon (figure 2b). We can see this intuitively as well. With consumption-smoothing households, households increase their consumption prior to the shock realization. Hence, the stabilizing central bank increases in-

[1] For example, the forward guidance puzzle has been described as "this promise [1 percentage point lower real interest rate for a single quarter at some point in the future] has a [stronger] impact on inflation when the promise pertains to interest rates five years in the future than when it pertains to the current interest rate" (McKay et al., 2016), as "shocks to very distant rates have a very powerful impact on today's consumption and inflation" (Gabaix, 2020), and as "…committing to low rates at t=T (or later). The baseline New Keynesian model predicts that the effectiveness of this kind of forward guidance increases with T and explodes as $T \to \infty$" (Angeletos and Lian, 2018).

[2] As we do not work with models without solutions, I take "existence" and "of equilibrium" implicitly, and only highlight the uniqueness for simplicity of discourse.

terest rate prior to the shock realization to stem the overheating economy (figure 1c). With the preceding peg of zero interest rates, the intended increase in interest rate is cancelled with a larger, negative monetary policy shock one period before the shock realization. This process repeats to the present. With longer forward guidance horizons, there is more time for this amplification to repeat. Hence, the explosive response in inflation comes from the explosive monetary policy shock in the present, not in the future. The policy under the pegged exercise is the entire interest rate path including today's, and the exercise should not be summarized to as if it is just a small change in the future.

Further, the preceding peg can not be argued to model the zero-lower bound. The preceding interest peg applies an upper bound, not lower. The pegged forward guidance puzzle states that output and inflation explode today. When inflation and output explode, does the zero-lower bound still bind? The intuition can also be seen by considering the case where the zero-lower bound applies with the correct sign: the zero-lower bound amplifies a contractionary monetary policy shock in the future by preventing the central bank from cutting the interest rates prior to the promised hike. Hence, the argument that the pegged forward guidance puzzle results from the zero-lower bound relies on a misapplied inequality.

Further, the explosive series of monetary policy shocks also highlights a problem with interpreting the entire "keep at zero longer" to be only policy commitment. If the central bank drops the interest rate by 200 basis points *today*, would the market infer a rogue central bank or an economic catastrophe currently unknown to the market? The explosive monetary policy shocks force the former interpretation. Once allowed to infer other causes, agents would never infer the explosive policy shocks even from the standard exercise. I show that just with productivity shocks in the three equation New Keynesian model, agents infer that the central bank is responding to a poor economic outlook and cut consumption today in response. We can see the main intuition from a simpler exercise. Instead of assuming that the path of interest rates implicitly defines only monetary policy deviations, we can allow few periods where the path implicitly defines some other shocks. The pegged forward guidance puzzle disappears allowing just one period of non-monetary policy shock. The pegged forward guidance puzzle is a knife-edge result.

Finally, I show that there is also a problem with the common research strategy employed in the literature to attenuate the strength of forward guidance. The effective discount rate of households is the difference between their discount rate and the real interest rate. In the standard forward guidance exercise, where the interest rate path is fixed, the effective discount rate is solely determined by the discount rate of households. This is why a common strategy has been to introduce less forward-looking agents. However, this simplification no longer holds when the upper bound is removed. To this end, I first show how to use the automatic differentiation on matrix decomposition required for the linear solution method in economics. This allows the calculation of the parameter dependence of the decay rate of the future automatically. When applied to the three equation New Keynesian model, the calculation shows that the discount rate of households[3] has only a weak impact on the decay rate of the future, and that the impact can have either sign depending on other parameter values. This is because the real interest rate equals the discount rate of households in steady-

[3]There is the *discount* rate of households and the *decay* rate of the impact of a future shock in a general equilibrium model. I will use these two definitions separately to reduce a potential confusion.

state, i.e., general equilibrium effect. This is a counter-example to the general research strategy of introducing less forward-looking agents. It is not automatic that the higher discount rate of less forward-looking agents survives the general equilibrium determination of interest rates.

1.1 Related Literature

My main theorem is very general, and applies to all general equilibrium models with uniqueness. Hence, my paper is related to all papers studying the impact of anticipated shocks. One such strand is the literature on news started by Beaudry and Portier (2006, 2007). My proof shows that the impact of any news gets weaker the further away the news is. I also give an alternate, numerically stable method to compute the solution of models with anticipated shock than the augmented state method used in the news literature (Schmitt-Grohé and Uribe, 2012).

The most closely related is the literature on the forward guidance puzzle. After Del Negro et al. (2023) noted the puzzling behavior, many papers have proposed resolutions of the forward guidance puzzle. To list a limited few from a vast literature, papers use different creative ideas like heterogeneity McKay et al. (2016); Kaplan et al. (2018); Hagedorn et al. (2019); Bilbiie (2020), behavioral Gabaix (2020); Angeletos and Lian (2018); Farhi and Werning (2019); García-Schmidt and Woodford (2019); Woodford (2019); McKay et al. (2017), and credibility Bodenstein et al. (2012); Andrade et al. (2019); Campbell et al. (2019). My analysis highlights that we need to clarify exercises with and without the preceding peg, and notes that the higher discount rate of less forward-looking agents might not survive the general equilibrium determination of interest rates.

The only paper that is very similar to mine in arguing that there might be no forward guidance puzzle is Maliar and Taylor (2019). They conduct the same endogenous forward guidance exercise with the three-equation New Keynesian model. However, they failed to note that the direction of the preceding peg is an upper-bound and argues that their analysis shows that forward guidance is ineffective away from the zero-lower bound. As can be seen from my analysis, this is not the most general conclusion possible from their exercise. Also, my main theorem gives the theoretical support to their hypothesis of generality of their exercise with the three equation NK model.

Lastly, as a part of my calculations, I also show how to automatically differentiate the linear solution method in economics contributing to the literature on automatic differentiation (Childers et al., 2022).[4] However, the automatic differentiation automates the computation of any derivatives of economic interests, and hence, is applicable to general economic problems.

2 Main Theorem

This section proves my main theorem on the impossibility of the endogenous forward guidance puzzle. Readers only interested in the forward guidance puzzle can safely skip this

[4]Childers et al. (2022) notes that it might be impossible to automatically differentiate the necessary matrix decomposition. It is possible, and the problem is addressed in appendix G.

section taking the following intuitive summary of theorem 1 as proven.

Theorem 1 (Intuitive Summary)**.** *In general equilibrium models with existence and uniqueness of the equilibrium, any present response to a one-time shock in the future becomes weaker the further away the shock is, and the inverse of the explosive eigenvalues*[5] *of the linearized system give the rate of decay of the impact of the future.*

2.1 Preliminary Calculations

This subsection introduces notations with some preliminary calculations, but the technical derivations can be safely skipped taking equation (4) for granted.

Suppose the linearized system of equations for a general equilibrium model is given by[6]

$$\begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \Phi \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi \end{bmatrix} \varepsilon_{t+1} =: \begin{bmatrix} \Phi_{vv} & \Phi_{vg} \\ \Phi_{gv} & \Phi_{gg} \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi \end{bmatrix} \varepsilon_{t+1} \tag{1}$$

where V denotes the forward-looking variables, g the pre-determined variables, and ε_{t+1} shocks satisfying $\mathbb{E}_t[\varepsilon_{t+1}] = 0$. This form is taken only for the simplicity of the resulting expressions, and the calculations generalize to a more general form like the one given in Sims (2002). Appendix H shows how to convert a system in the form of Sims (2002) into equation (1). Finally, let F be the solution (decision rule) of the system,[7] i.e., $\begin{bmatrix} F \\ I \end{bmatrix}$ forms the basis vectors for the stable subspace of Φ.

Now, consider the planning problem for an anticipated shock $\varepsilon_{k|0}$ that will be realized at time k, but known at time 0. Given the linear form, we can substitute forward to get[8]

$$\begin{bmatrix} V_k \\ g_k \end{bmatrix} = \begin{bmatrix} \Phi_{vv} & \Phi_{vg} \\ \Phi_{gv} & \Phi_{gg} \end{bmatrix}^k \begin{bmatrix} V_0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_{k|0} \end{bmatrix}$$

Once the shock realizes, the system must be back on the stable subspace to satisfy the transversality condition.[9] Hence, we have

$$(\Phi^k)_{vv} V_0 = V_k = F g_k + F \varepsilon_{k|0} = F(\Phi^k)_{gv} V_0 + F \varepsilon_{k|0}$$

resulting in

$$V_0 = \left[(\Phi^k)_{vv} - F(\Phi^k)_{gv}\right]^{-1} F \varepsilon_{k|0} \tag{2}$$

We can further simplify the expression using the following relationship:

[5] eigenvalues outside the unit circle

[6] Expectation operators are dropped for simplicity of notation. Expectation operators will just remove the shock-terms (certainty equivalence).

[7] Though some linear solution methods, e.g., Sims (2002), do not return F, one can use the Schur decomposition computed as a part of the solution method to compute F.

[8] The ε_{t+k} terms drop from $\mathbb{E}_{t+k}[\varepsilon_{t+k}] = 0$.

[9] Alternately, solving with the perfect foresight solution method would result in the same expression as the following for this linear model.

Lemma 1.

$$(\Phi^k)_{vv} - F(\Phi^k)_{gv} = (\Phi_{vv} - F\Phi_{gv})^k \tag{3}$$

Proof. Refer to Appendix A for the proof. □

Plugging equation (3) into equation (2) gives

$$V_0 = (\Phi_{vv} - F\Phi_{gv})^{-k} F\varepsilon_{k|0} \tag{4}$$

The mathematical intuition behind the calculations is the invariance of the stable subspace. Lemma 1 shows that the component in the orthogonal complement of the stable subspace evolves by the dynamics projected on to the orthogonal complement of the stable subspace. The translation of this mathematical intuition to an economic intuition can be given as: the difference between the planned behavior with and without anticipated shocks remains orthogonal to the planned behavior without anticipated shocks under the system dynamics.[10] Refer to Appendix A for further details.

2.2 The Decay of Future Shocks

As the initial response satisfies equation (4), it is sufficient to analyze the eigenvalues of $(\Phi_{vv} - F\Phi_{gv})$ to understand amplification or decaying of present responses to a future shock in the anticipation horizon. The present response to a future shock will be decayed for eigenvalues outside the unit circle, i.e., eigenvalues with the absolute value larger than 1, and amplified for those inside.[11] In the literature, we are given the impression that it is possible to have eigenvalues of $(\Phi_{vv} - F\Phi_{gv})$ inside the unit circle, i.e., to have present responses become stronger in the anticipation horizon. However, this is not possible in a general equilibrium model with uniqueness:

Theorem 1. *For a (linearized) general equilibrium model of the form equation* (1)*, if the equilibrium exists and is unique,*[12] *then eigenvalues of* $(\Phi_{vv} - F\Phi_{gv})$ *are equal to the eigenvalues of* Φ *outside the unit circle, i.e., explosive roots.*

Proof. Refer to Appendix A for the proof. □

In partial equilibrium, no market mechanism guarantees that the impact of a future shock decays. For example, if the real interest rate is set equal to the discount rate of the household, the impact of a future shock does not decay. However, in a general equilibrium model with endogenous dynamics leading up to the future shock, market mechanisms guarantee that such interest rate paths are not possible.

Before proceeding to the forward guidance analyses, Theorem 1 also gives a decomposition based computation method with numerical stability even with rounding errors. Without

[10] This statement sounds true by definition. This is an artifact of the invariance. The invariance is why we know that the quantity can be defined at all. See footnote 25 for a similar example.

[11] The multiplication by the inverse flips the explosive/stable relationship of the eigenvalues.

[12] for all possible initial condition g_0. It is possible for solution to exist with respect to different shocks, but not for all initial g_0, so a strong form of uniqueness, but this is a more standard definition. I only know of Sims (2002) that uses the weaker definition.

numerical stability, explosive responses can result from accumulation of rounding errors. The numerically stable method further has a linear scale in the anticipation horizon compared to the cubic scale of the standard augmented-state method used in the literature to solve models with anticipated/news shocks (See Del Negro et al. (2023) and Schmitt-Grohé and Uribe (2012) for example).[13] Hence, one should use the decomposition based method if the anticipation horizon is long. Details are relegated to Appendix C.

3 Forward Guidance

Theorem 1 implies that any present response to a one-time monetary policy shock cannot become stronger in anticipation horizon in any general equilibrium model with existence and uniqueness of equilibrium. Hence, there is no *endogenous forward guidance puzzle*. To clarify and reconcile the differences between the endogenous and the pegged exercise, I use the three-equation New Keynesian model of a version from the textbook of Herbst and Schorfheide (2015)[14] with the equations given by

$$dy_t = \frac{1}{\tau}(i_t - \pi_t) \tag{5}$$

$$d\pi_t = \rho\pi_t - \frac{\tau(\varepsilon - 1)}{\theta_p} y_t \tag{6}$$

$$di_t = \rho_R(\phi_\pi \pi_t + \phi_y y_t - i_t)dt + \sigma_R dW_{R,t} \tag{7}$$

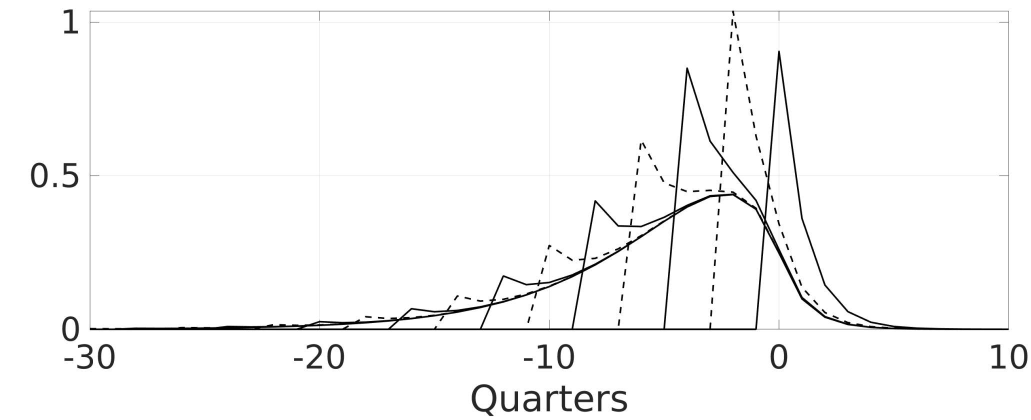
The analyses in main text use parameter values from the posterior mode of the Bayesian estimation replicating Herbst and Schorfheide (2015). The figures with parameter uncertainty bands in appendix F confirm that the analyses in main text hold for other parameters in the Bayesian posterior. To reiterate, theorem 1 is general and applies to all sensible model, and a textbook model is chosen only for illustrations.

3.1 Endogenous Forward Guidance Exercise

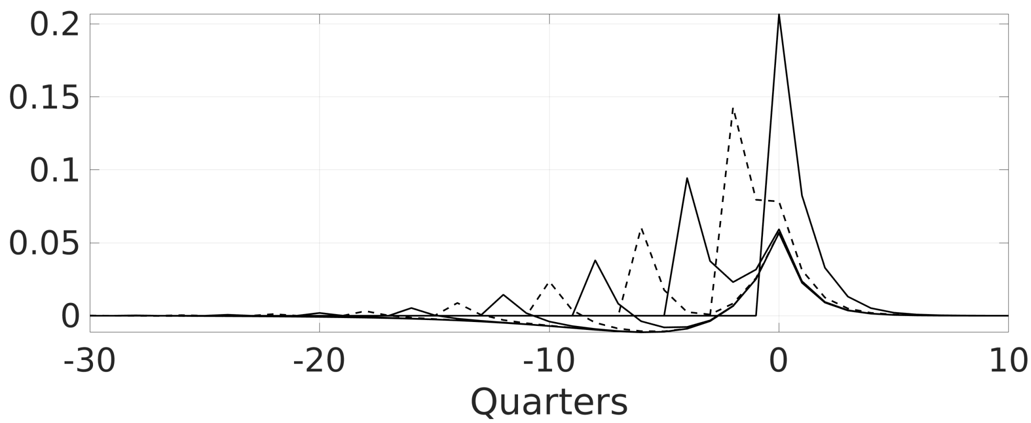
Figure 1 shows the responses to a one-time monetary policy shock with different anticipation horizons. Responses do not explode. In fact, there is almost no response in any of the variables to a shock further than 6 years in the future. Figure 2 of Del Negro et al. (2023)

[13] In Appendix D, I prove that the augmented-state method gives the same solution as my method, but could not prove that the augmented-state method is numerically stable in general. Numerical experiments for the three-equation NK model confirms that both methods give the same answer for all horizons, but numerical experiments are only suggestive of the stability of the augmented state method, not definitive. When in doubt, it is safer to use the decomposition based method with proven stability. See Appendix D for further technical details.

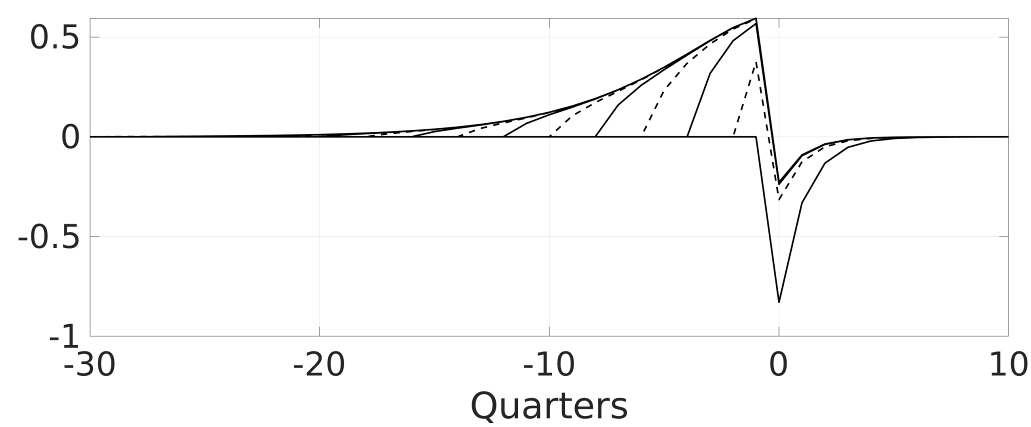
[14] but converted to continuous time (See Appendix F for the precise model description). There is no fundamental difference between continuous and discrete time approach for my analysis. I recycle the replication code of Herbst and Schorfheide (2015) written to test my codebase in continuous time (See Achdou et al. (2022) for why). I prove the same results in continuous time in Appendix B. We can also see from Maliar and Taylor (2019) that the response to a one-time shock is the same in discrete time.



(a) Inflation



(b) GDP



(c) Interest Rate

Figure 1: Impulse response functions to one-time MP shock at time 0 known a different number of periods ahead. Solid lines show anticipation horizons of (0, 8, 16, ...) quarters, and dashed lines (4, 12, 20, ...) quarters.

only shows 15 quarters, so it is unclear whether the responses do go to zero or not. But mathematically, the responses must go to zero.[15]

Beyond the decay, the responses are also intuitive. Considering any line for a fixed anticipation horizon, Figure 1b shows that consumption-smoothing households respond to the future by increasing consumption before the shock realizes, leading to a boom prior to the shock realization. Inflation follows a similar increase as shown in figure 1a. The central bank responds to the increase in inflation and output preceding the expansionary shock by *increasing* the interest rate, as shown in figure 1c. The increase in interest rate is enough to stem the overheating economy.

Others have noted this response first. Del Negro et al. (2023) describes this behavior in the introduction of the 2023 draft, and Maliar and Taylor (2019) show that the same figures with a closed form solution. Maliar and Taylor (2019) argue that the exercise shows that the forward guidance is irrelevant when the zero-lower bound does not bind. However, the interest responses in figure 1c show that the zero-lower bound is irrelevant. The zero-lower bound prevents the central bank from lowering the interest rate, not from increasing it. Mechanically zeroing out the interest rate preceding the shock is equivalent to a zero-*upper* bound!

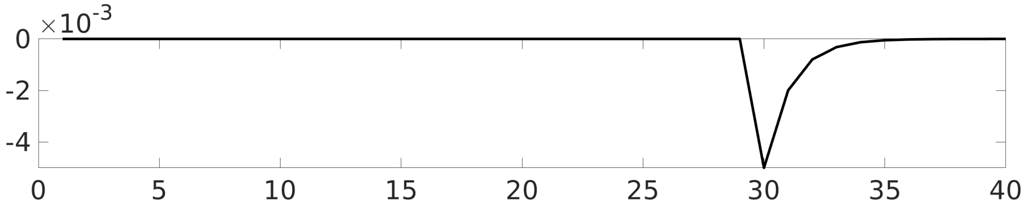
The direction of the upper and lower bound can be seen more clearly with an exercise where the zero-lower bound binds with the correct sign. With a promise of an interest rate *increase* in the future, the central bank tries to fight the preceding contractions (vertical reflection of figure 1b) by decreasing the interest rates (vertical reflection of figure 1c). The zero-lower bound prevents the central bank from decreasing the interest rates. Hence, the zero-lower bound amplifies a contractionary MP shock in the future by preventing the central bank from cutting the interest rates prior to the hike. The logic can be inverted for the forward guidance of an expansionary promise. The zero-*upper* bound amplifies an *expansionary* MP shock in the future by preventing the central bank from *increasing* the interest rates prior to the *cut*. We can already guess what will happen with this upper bound.

Comparing across different horizons, figure 1a shows that for inflation, the peak response increases in the anticipation horizon for horizons less than 3 quarters, but decreases after.[16] The initial increase does not contradict theorem 1. In higher dimensions, decay does not guarantee monotonic decrease to zero in all variables, but an eventual decrease to zero.[17] This might be considered a forward guidance puzzle in inflation, but it is much weaker than the impression made by the simplified summary of the forward guidance puzzle. The picture is simpler for GDP. Figure 1b shows that the peak response in output monotonically decreases in horizon of the shock. With consumption-smoothing, households can smooth out the response more with a longer preparation time than shorter, resulting in a lower peak response. In short, agents in the standard NK models respond intuitively to a one-time monetary policy shock in the future without the preceding zero-*upper* bound, and the endogenous forward guidance puzzle does not exist.

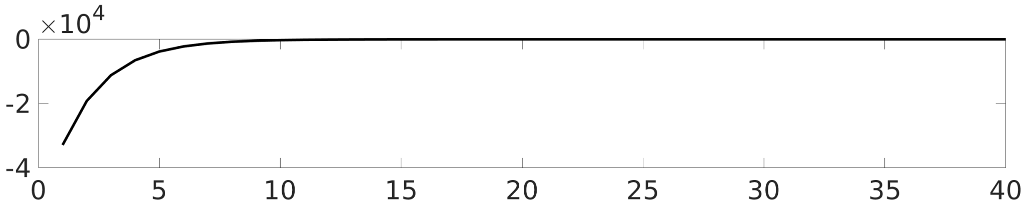
[15] An explosive behavior can numerically arise as an accumulation of rounding-errors. See appendices C and D for details.

[16] This is also noted in the introduction of 2023 version of Del Negro et al. (2023).

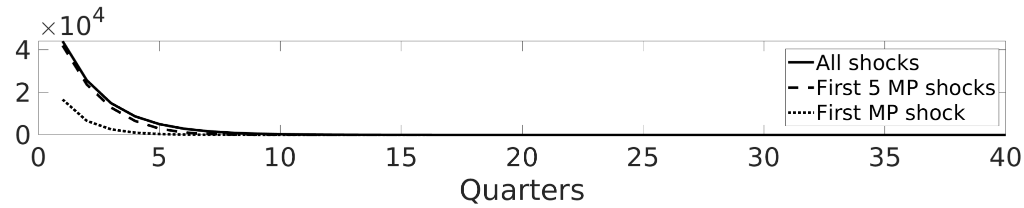
[17] It is monotonic in a basis representation based on eigenvectors.



(a) Interest Rate



(b) MP shocks to generate interest rate path above



(c) Inflation responses keeping different number of MP shocks

Figure 2: Future MP shock with preceding zero peg

3.2 Pegged Forward Guidance Puzzle

In the standard exercise of forward guidance in the literature, researchers took a literal translation of the "keep at zero for longer" by pegging the interest rate to zero before the drop in interest rate in the future, as shown in figure 2a. This is the one we are calling the pegged forward guidance puzzle. The entire path of interest rate implicitly defines monetary policy shocks up to the time of the drop in interest rate. Only Del Negro et al. (2023) and McKay et al. (2016) explicitly mention the implicitly defined shocks, but this type of policy deviations from the central bank's normal behaviors exist under all pegged forward guidance exercise. Now, how do these implicitly defined policy shocks look like? Figure 2b shows the monetary policy shocks necessary to generate the interest rate path. Required shocks are explosive.

The intuition for the exponential growth in shocks can be seen from figure 1c. The central bank increases interest rate prior to the promised cut to stem the overheating economy as described in the last section. The interest rate is larger in magnitude prior to the promised cut than the magnitude after realization. For the peg of zero, a larger negative expansionary shock is necessary to cancel out the increase one period prior to the promised drop. This process repeats till present, and results in the series of exploding monetary policy shocks. In fact, the growth in monetary policy shocks becomes exactly exponential for long enough horizons.

Given the size of the required monetary policy deviations from the normal operations of the central bank, this policy promise is unlikely to have any credibility for any rational households, and section 4 shows that agents interpret the promise differently if allowed. However, as agents believe the policy promise as credible, agents respond as they are modelled by explosively increasing inflation and consumption. We can see more clearly that the explosive response in inflation does not result from explosive responses to a distant shock by studying the first few monetary policy shocks. Figure 2c shows the response in inflation with just the first and first 5 monetary policy shocks given in figure 2b. The first shock already leads to the explosive response, and the first five shocks recover the entire response. Agents are *not* responding strongly to a future shock. Agents react strongly to the explosive shocks brought to the *present* by the preceding peg. An interest rate drop further in the future corresponds to a stronger expansionary stance of the central bank. It is not puzzling that agents react more to a larger shock.

To avoid a potential presentism, one historical context from 2012 made the exercise with the preceding peg seem more natural. Campbell et al. (2012) made a useful distinction of the forward guidance into Delphic (information content) and Odyssean (commitment) components. Along with the distinction came a powerful description of Odysseus bound to the mast, keeping the policy promise despite the rising inflation. The exploding inflation seemingly highlighted the Odyssean nature of the exercise. However, the Odyssean distinction should only be about the time-inconsistency of the policy commitment, and does not automatically imply the upper bound. Even a one-period policy commitment in the future is time-inconsistent (Odyssean) as the central bank will want to renege on the promise once the future arrives. The natural flow of the literature made all of us overlook, but the standard pegged exercise is not just Odyssean, but *heroically* Odyssean.

4 Unintentionally Delphic, Odyssean Forward Guidance

Even if the standard exercise with the zero-upper bound is the correct translation of "keep zero for longer," is it correct to assume that agents would infer that the entire path of zero interest rate is caused solely by the central bank with explosive deviations? The problem with this formulation is that rational agents will reason through why the central bank promises the zero interest rate. Even if the central bank wants to make an Odyssean-only forward guidance announcement, an unintentional Delphic forward guidance follows along. The intuition can be seen more clearly without the forward guidance. If the central bank suddenly cuts the policy rate by 200 basis points *today*, would the market infer a rogue central bank or an economic catastrophe currently unknown to themselves? When the size of the monetary policy shock is within the regular operations, ignoring the information update of agents remains a good approximation. However, with the exploding monetary policy shocks, ignoring the information content becomes an exponentially worse approximation for the forward guidance exercises.

4.1 Main Intuition

The next subsection shows a more systematic approach, but the main intuition can be seen from a simpler exercise. In the literature, the pegged path of interest rates until time k

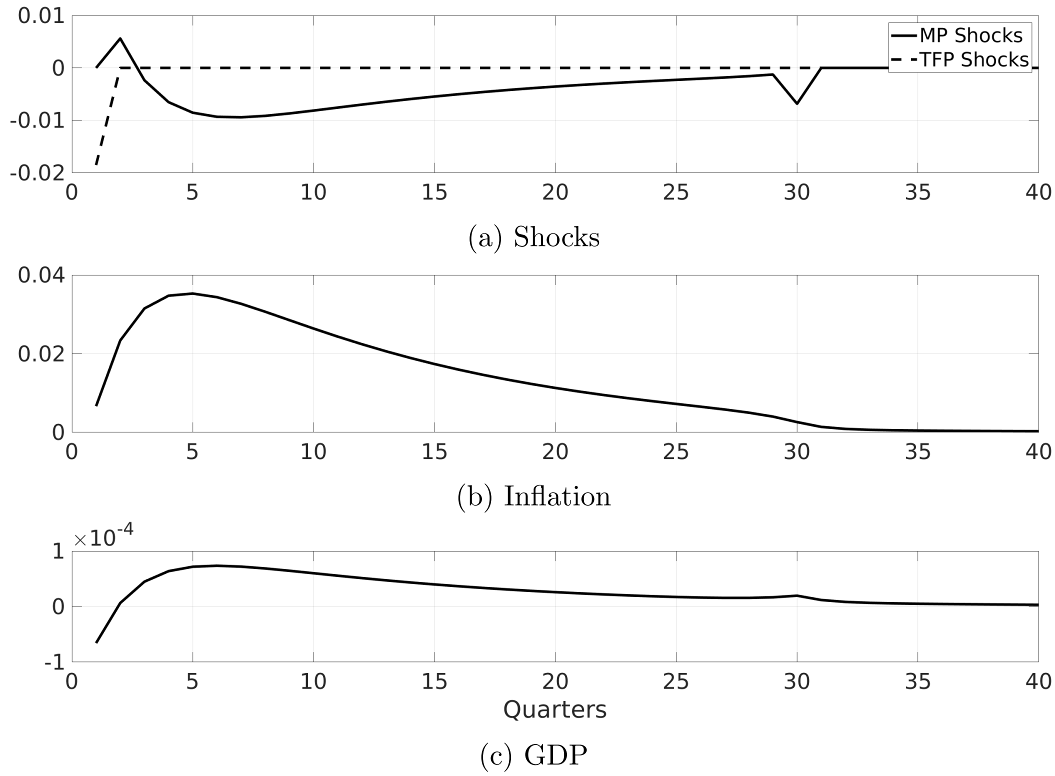


Figure 3: Pegged forward guidance exercise where the path of interest rates implicitly define $\varepsilon_{TFP,1}$ and $\{\varepsilon_{mp,t}\}_{t=2,...,30}$ (instead of $\{\varepsilon_{mp,t}\}_{t=1,...,30}$ underlying figure 2).

defines the monetary policy shocks, $\{\varepsilon_{mp,t}\}_{t=1,...,k}$. This implicitly makes the assumption that the entire path of interest rate is caused only by the central bank. This assumption can be weakened considering an alternate scenario. Supposed that the path of interest rates implicitly determines productivity shocks[18] to time l and monetary policy shocks from time $l+1$ to time k, i.e., $\{\varepsilon_{TFP,t}\}_{t=1,...,l} \cup \{\varepsilon_{mp,t}\}_{t=l+1,...,k}$. In this setup, the central bank uses $(k-l)$ period of forward guidance l period in the future to counter the poor economy till period l. This exercise can be interpreted as an interplay of the Delphic and Odyssean component of the announcement. Agents infer from the forward guidance announcement that the economy will do poorly for the next l periods from bad productivity (Delphic), and that the central bank will keep the rates low for $(k - l)$ periods longer from period l (Odyssean).

Figure 3 shows the exercise with $l = 1$ and $k = 30$. Hence, the central bank runs a 29 period (Odyssean) forward guidance from next period to counter the productivity shock this period. Figure 3b shows that the inflation response does not explode. In comparison, the scale of inflation response is 10^4 with $l = 0$. Figure 3c shows that the GDP response is also of the correct scale. GDP response is negative initially since the interest rates path of zero is given by a negative TFP shock that future monetary policy shocks attempt to undo. Surprisingly, even with $l = 1$, the pegged forward guidance puzzle disappears.

This result is also robust to different values of l as long as $l \neq 0$. Figure 4 shows the

[18]I use productivity shocks because I used a textbook three-equations NK model. In richer models, other shocks, e.g., demand shocks, can be used.

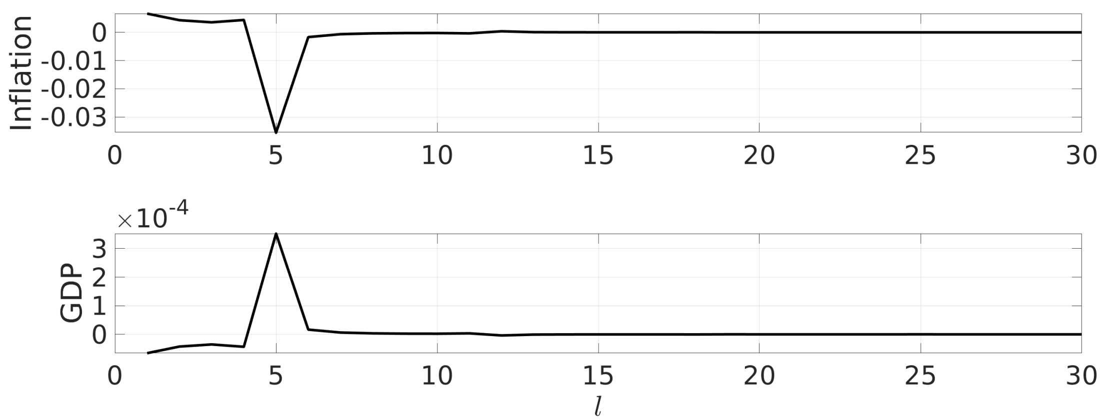


Figure 4: Initial response to different values of l.

initial inflation responses to different values of l. The figure shows that inflation does not explode for any value of $l > 0$. In fact, the result for $l = 0$ could not be shown in the same figure given the scale at 10^4. The pegged forward guidance puzzle with $l = 0$ is a knife-edge result. As soon as agents are allowed to infer any non-monetary policy cause for the interest rate path, the pegged forward guidance puzzle disappears.

The choice of l where the zero-interest rate is caused by the economy or the central bank is arbitrary. Also, it is possible that both productivity shocks and monetary policy shocks are active in a given period. The following systematic generalizations allow for those calculations.

4.2 Systematic Analysis

The systematic way to convert the main intuition is the computation of the conditional probability update from the policy announcement. This is implicitly the calculation underlying the conversion of the interest rate path to the implicit series of monetary policy shocks as well. The key difference is that other shocks than just the monetary policy shocks are necessary to allow agents to infer other reasons of the zero interest rates. As the conditional probability calculations are standard, technical details are relegated to appendix E. In simple notations, agents update the probability distribution of future shocks conditional on the announcement

$$\varepsilon_{MP;l}|\text{interest path} \qquad \varepsilon_{others;l}|\text{interest path}$$

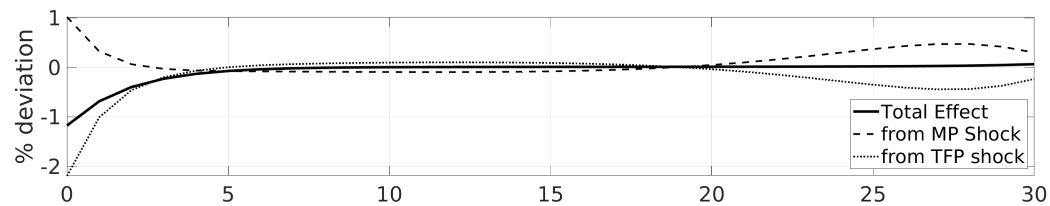
for $l = 0, 1, \ldots, k$ upon policy announcement and adjust their planning problem accordingly.[19] The type of other shocks depends on the specific model under consideration. Dif-

[19] In this formulation, the standard exercise corresponds to imposing

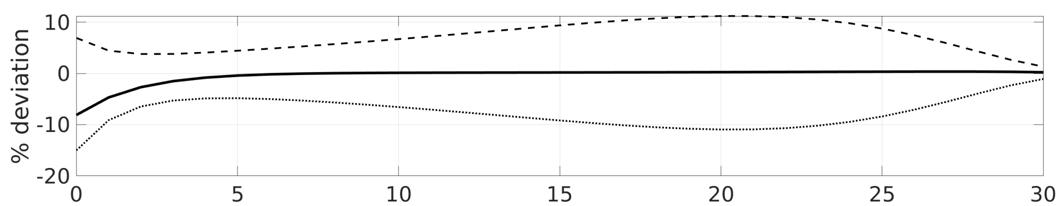
$$\varepsilon_{others;l}|\text{announcement} = \varepsilon_{others;l}, \forall l$$

and hence,

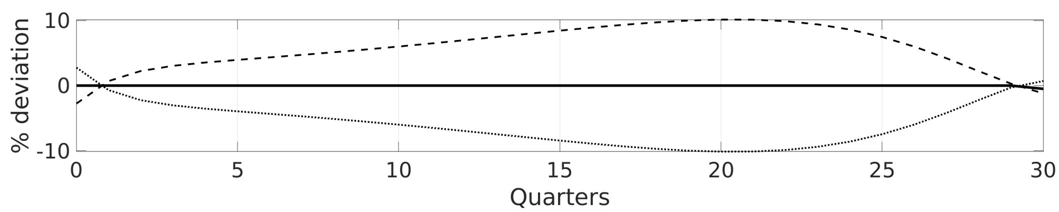
$$\mathbb{E}[\varepsilon_{others;l}|\text{announcement}] = \mathbb{E}[\varepsilon_{others;l}] = 0, \forall l.$$



(a) Output



(b) Inflation



(c) Interest rate

Figure 5: Conditional probability based forward guidance exercise when the central bank announces stimulation in the future with preceding zero lower bound.

ferent shocks can be introduced if necessary, but most DSGE models already have more shocks than the monetary policy shock. For example, the three equation NK model has TFP shocks. With the additional shocks, households rationally infer how much of the zero interest rate is caused by MP and TFP shocks.

I present the result of the updated conditional expectation in figure 5. First, figure 5b shows that there is no explosive inflation. This is because agents are allowed to infer other shocks than just the MP shocks. Figure 5c shows that agents infer that the MP shock and TFP shock off-set each other to result in the zero interest rate, i.e., households infer that the central bank plans to react to bad economic situations in the future. Though they are allowed to infer the explosive monetary policy shocks, rational agents would never infer that series of shocks from the standard NK model. Output, given in figure 5a, shows this more clearly. Households infer negative TFP shocks in the future from the interest rate path announcement. Hence, in period 0, where the action is actually realized,[20] agents drop consumption, not increase.

As an aside, this exercise also suggests an alternate empirical strategy than those currently used in the literature. Following the distinction from Campbell et al. (2012), subsequent literature (Del Negro et al., 2023; D'Amico and King, 2015; Andrade and Ferroni, 2021; Zlobins, 2021) have identified the "Odyssean" component of the forward guidance using sign restrictions, and the predictions of the standard exercise have been compared to the Odyssean component. However, theoretical models also have a prediction on the information content (Delphic) component. Further empirical analyses can benefit from handling both the commitment and information update at the same time.

To get back to the main analysis, even the calculation in this section is not the correct exercise of the reality. As agents already know about the poor economy prior to the policy announcement, the entire interest rate path should not be treated as news.[21] Further theoretical and empirical research on the information content of Odyssean policy announcements are needed. However, the analysis in this section highlights that even a simple extension of being true to the other shocks already in the model is sufficient to remove the pegged forward guidance puzzle. Agents should be allowed to think at higher-order and, hence, to infer the underlying causes of the central bank announcement. It is not puzzling to get a strong expansionary response when we forced agents to believe the entire low interest rate to be caused just by a massive expansionary policy.

5 Semi-elasticity of Decay Rate on Parameters

In this section, I show how to compute the parameter dependence of the decay rate of the future. The calculation shows which parameters attenuate the strength of endogenous forward guidance of a one-time shock in the future. Since there is no forward guidance puzzle endogenous or pegged, this might not be of any interest. However, the calculation is generally applicable to any GE models with uniqueness. Also, the calculation produces a

This would be a valid interpretation if agents believe that they have equal information as the central bank.

[20] The rest of the paths are planned/expected actions.

[21] For example, running the Kalman filter to get the initial states is a better approximation than assuming no prior information.

counter-example to a general research strategy in the forward guidance puzzle literature. A common strategy is to make agents less forward looking (See section 1.1). With the standard exercise with the interest rate pegged, the effective discount rate is determined only by the discount rate of households. Hence, the partial equilibrium logic generalizes automatically to the standard exercise with the preceding peg. Without the peg, the general equilibrium determination of the interest rates makes this simplification no longer automatic.

To get back to the analysis of the parameter dependence, theorem 1 shows that the explosive eigenvalues of Φ[22] determine the decay rate of the future. Hence, let λ be the smallest explosive eigenvalue of Φ.[23] Then, for a given parameter θ, we can study the semi-elasticity

$$\frac{\partial \lambda}{\partial \log(\theta)}$$

to understand the impact of θ on the rate of decay. In theory, we can use any automatic differentiation toolbox to compute this semi-elasticity, but I have not seen the Schur decomposition included in any. I give the necessary calculations to automatically differentiate the Schur decomposition in the appendix G,[24] and one can refer to Ahn (2016) for the open source codes.

Table 1 gives the semi-elasticities of the decay rate from the three equation NK model with respect to different parameters. First row shows that the Taylor-rule coefficient on inflation has the biggest impact on the decay rate. A 1% increase will lead to the decay rate increase of 0.02, or the half-life decrease of 0.15 quarters. This is a sizeable impact given the baseline half-life of 2.64 quarters. This corroborates the result found in Maliar and Taylor (2019) that the Taylor rule parameters have a strong impact on the model dynamics, but we can directly compute the impact of parameters without closed form solutions. Other consumption smoothing and adjustment cost parameters have intuitively correct impact on the decay rate of the future.

We, however, get a somewhat puzzling response for the decay rate of households. The semi-elasticity is small and of the wrong sign, i.e., an increase in the discount rate of households do not necessarily lead to faster decay of the future shocks. Though seemingly unintuitive, this is due to the GE effects. First, the effective discount rate for households is the difference between their own discount rate and the real interest rate. The equilibrium condition forces the real interest rate to equal the discount rate of households in steady-state, leading to a muted response. Second, the discount rate of households interacts with the monetary policy rule. The monetary policy rule is a rule without discounting, unlike households. Hence, the general equilibrium determination of the interest rate can lead to the decay rate of the future to go in the wrong direction from the partial equilibrium logic.

[22]For people who have skipped section 2, Φ corresponds to the dynamics matrix of the linearized DSGE model. See equation (1).

[23]For a given shock, it is possible that the shock does not load on all explosive eigenvalues. Hence, λ should be the smallest explosive eigenvalue with a non-zero loading.

[24]This allows automatic differentiation of the linear solution methods in economics. This complements Childers et al. (2022) who bypass the automatic differentiation of Schur decomposition. Bypassing Schur decomposition will be faster for likelihood evaluation, but loses the ability to analyze intermediate quantities like eigenvalues and the applicability to continuous time models. See appendix G for more details.

In fact, though the magnitude of semi-elasticity on the discount rate of households remains small for all parameters, the sign depends on other model parameters. In other words, less forward-looking households do not automatically imply attenuation of the forward guidance without the preceding peg.

Table 1: Semi-elasticity of decay rate on parameters.

Description	$\frac{\partial \lambda}{\partial \log(\theta)}$	$\frac{\partial(\text{half-life})}{\partial \log(\theta)}$	Parameter
Taylor-rule coefficient on inflation	1.99	−15.4	ϕ_π
Rate of nominal interest rate adjustment	0.93	−7.23	ρ_R
Elasticity of substitution	−0.23	1.79	ε
Inflation adjustment cost	0.22	−1.74	θ_p
CRRA elasticity	0.14	1.09	τ
Taylor-rule coefficient on output gap	−0.14	−1.09	ϕ_y
Discount rate of households	−0.0012	0.0095	ρ

Notes:
[1] Values are converted to discrete time equivalent from continuous time model ($\lambda_{\text{dt}} = e^{\lambda_{\text{ct}}}$) as for main text is in discrete time.
[2] half-life $= \frac{\log(2)}{\lambda_{\text{ct}}}$ (in quarters)
[3] Baseline values from the posterior mode: $\lambda = 1.3$, and half-life $= 2.64$ quarters
[4] See appendix F for the precise definition of the parameters.

6 Conclusion

I have proven that existence and uniqueness of equilibrium in general equilibrium models are sufficient to guarantee that any present response to a one-time shock in the future is decayed in the anticipation horizon. There is no *endogenous* forward guidance puzzle. The simplified summary of the pegged forward guidance puzzle was too strong. Since the forward guidance is an entire path of interest rates with the pegged exercise, a longer path implies a larger cumulative policy deviations. Agents respond reasonably to this explosive shock today, not explosively to a small change in the future. The pegged exercise applies an upper bound, not lower. The pegged forward guidance puzzle is also a knife-edge result that disappears as soon as agents can infer that poor economic conditions underlying the forward guidance of the central bank. There is nothing puzzling about the pegged forward guidance puzzle. However, the forward guidance puzzle can remain in spirit. The decay rate of the future in our models can still be too low, resulting in stronger responses in our models to a future shock than in the data. It is time to study whether there is this quantitative endogenous forward guidance puzzle instead of resolving the pegged forward guidance puzzle.

References

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2022). Income and wealth distribution in macroeconomics: A continuous-time approach. *The review of economic studies*, 89(1):45–86.

Ahn, S. (2016). Automatic differentiation for matlab (github repo). `https://github.com/sehyoun/MATLABAutoDiff/`.

Ahn, S., Kaplan, G., Moll, B., Winberry, T., and Wolf, C. (2018). When inequality matters for macro and macro matters for inequality. *NBER macroeconomics annual*, 32(1):1–75.

Andrade, P. and Ferroni, F. (2021). Delphic and odyssean monetary policy shocks: Evidence from the euro area. *Journal of Monetary Economics*, 117:816–832.

Andrade, P., Gaballo, G., Mengus, E., and Mojon, B. (2019). Forward guidance and heterogeneous beliefs. *American Economic Journal: Macroeconomics*, 11(3):1–29.

Angeletos, G.-M. and Lian, C. (2018). Forward guidance without common knowledge. *American Economic Review*, 108(9):2477–2512.

Beaudry, P. and Portier, F. (2006). Stock prices, news, and economic fluctuations. *American Economic Review*, 96(4):1293–1307.

Beaudry, P. and Portier, F. (2007). When can changes in expectations cause business cycle fluctuations in neo-classical settings? *Journal of Economic Theory*, 135(1):458–477.

Bilbiie, F. O. (2020). The new keynesian cross. *Journal of Monetary Economics*, 114:90–108.

Bodenstein, M., Hebden, J., and Nunes, R. (2012). Imperfect credibility and the zero lower bound. *Journal of Monetary Economics*, 59(2):135–149.

Campbell, J. R., Evans, C. L., Fisher, J. D., Justiniano, A., Calomiris, C. W., and Woodford, M. (2012). Macroeconomic effects of federal reserve forward guidance [with comments and discussion]. *Brookings papers on economic activity*, pages 1–80.

Campbell, J. R., Ferroni, F., Fisher, J. D., and Melosi, L. (2019). The limits of forward guidance. *Journal of monetary economics*, 108:118–134.

Childers, D., Fernández-Villaverde, J., Perla, J., Rackauckas, C., and Wu, P. (2022). Differentiable state-space models and hamiltonian monte carlo estimation. Technical report, National Bureau of Economic Research.

D'Amico, S. and King, T. B. (2015). What does anticipated monetary policy do?

Davies, P. I. and Higham, N. J. (2003). A schur-parlett algorithm for computing matrix functions. *SIAM Journal on Matrix Analysis and Applications*, 25(2):464–485.

Del Negro, M., Giannoni, M. P., and Patterson, C. (2023). The forward guidance puzzle. *Journal of Political Economy Macroeconomics*, 1(1):000–000.

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*, volume 38. OUP Oxford.

Farhi, E. and Werning, I. (2019). Monetary policy, bounded rationality, and incomplete markets. *American Economic Review*, 109(11):3887–3928.

Gabaix, X. (2020). A behavioral new keynesian model. *American Economic Review*, 110(8):2271–2327.

García-Schmidt, M. and Woodford, M. (2019). Are low interest rates deflationary? a paradox of perfect-foresight analysis. *American Economic Review*, 109(1):86–120.

Hagedorn, M., Luo, J., Manovskii, I., and Mitman, K. (2019). Forward guidance. *Journal of Monetary Economics*, 102:1–23.

Herbst, E. P. and Schorfheide, F. (2015). *Bayesian estimation of DSGE models*. Princeton University Press.

Kaplan, G., Moll, B., and Violante, G. L. (2018). Monetary policy according to hank. *American Economic Review*, 108(3):697–743.

Maliar, L. and Taylor, J. B. (2019). Forward guidance: Is it useful away from the lower bound? Technical report, National Bureau of Economic Research.

McKay, A., Nakamura, E., and Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, 106(10):3133–3158.

McKay, A., Nakamura, E., and Steinsson, J. (2017). The discounted euler equation: A note. *Economica*, 84(336):820–831.

Schmitt-Grohé, S. and Uribe, M. (2012). What's news in business cycles. *Econometrica*, 80(6):2733–2764.

Sims, C. A. (2002). Solving linear rational expectations models. *Computational economics*, 20(1-2):1.

Woodford, M. (2019). Monetary policy analysis when planning horizons are finite. *NBER macroeconomics annual*, 33(1):1–50.

Zlobins, A. (2021). Macroeconomic effects of the ecb's forward guidance. *Empirical Economics*, 61:2587–2611.

A Proofs

Lemma 1.

$$(\Phi^k)_{vv} - F(\Phi^k)_{gv} = (\Phi_{vv} - F\Phi_{gv})^k$$

Proof. For any g, $\begin{bmatrix} Fg \\ g \end{bmatrix}$ is on the stable manifold by definition of F. Hence,

$$\Phi \begin{bmatrix} Fg \\ g \end{bmatrix} = \begin{bmatrix} \Phi_{vv}Fg + \Phi_{vg} \\ \Phi_{gv}Fg + \Phi_{gg} \end{bmatrix}$$

is also on the stable manifold by the invariance of the stable manifold.[25] Using the definition of F again,

$$\Phi_{vv}Fg + \Phi_{vg}g = F\Phi_{gv}Fg + F\Phi_{gg}g$$

holds for all g. Hence, we have

$$\Phi_{vv}F + \Phi_{vg} = F\Phi_{gv}F + F\Phi_{gg} \tag{8}$$

Also, from definition/multiplication, we have

$$\begin{bmatrix} (\Phi^{k+1})_{vv} \\ (\Phi^{k+1})_{gv} \end{bmatrix} = \begin{bmatrix} \Phi_{vv} & \Phi_{vg} \\ \Phi_{vg} & \Phi_{gg} \end{bmatrix} \begin{bmatrix} (\Phi^{k})_{vv} \\ (\Phi^{k})_{gv} \end{bmatrix} = \begin{bmatrix} \Phi_{vv}(\Phi^{k})_{vv} + \Phi_{vg}(\Phi^{k})_{gv} \\ \Phi_{gv}(\Phi^{k})_{vv} + \Phi_{gg}(\Phi^{k})_{gv} \end{bmatrix}$$

Finally, we can prove our relationship. Consider an induction step for a given k,

$$\begin{aligned} (\Phi^{k+1})_{vv} - F(\Phi^{k+1})_{gv} &= \left[\Phi_{vv}(\Phi^{k})_{vv} + \Phi_{vg}(\Phi^{k})_{gv}\right] - F\left[\Phi_{gv}(\Phi^{k})_{vv} + \Phi_{gg}(\Phi^{k})_{gv}\right] \\ &= (\Phi_{vv} - F\Phi_{gv})(\Phi^{k})_{vv} - (F\Phi_{gg} - \Phi_{vg})(\Phi^{k})_{gv} \\ &= (\Phi_{vv} - F\Phi_{gv})(\Phi^{k})_{vv} - (\Phi_{vv} - F\Phi_{gv})F(\Phi^{k})_{gv} \\ &= (\Phi_{vv} - F\Phi_{gv})^{k+1} \end{aligned}$$

where the second equality comes from substituting in the invariance condition of equation (8), and the last equality from induction. The base case is by definition, and this concludes the proof. □

Theorem 1. *For a (linearized) GE model with the dynamics given by*

$$\begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} \Phi_{vv} & \Phi_{vg} \\ \Phi_{gv} & \Phi_{gg} \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi \end{bmatrix} \varepsilon_t$$

if the solution exists and is unique, then eigenvalues of $(\Phi_{vv} - F\Phi_{gv})$ are equal to the eigenvalues of Φ outside the unit circle, i.e., explosive roots.

Proof. Denote a Jordan normal form[26] of

$$\begin{bmatrix} \Phi_{vv} & \Phi_{vg} \\ \Phi_{gv} & \Phi_{gg} \end{bmatrix}$$

[25] Alternately, the intuitive statement is that if the jump variables, V, are determined by the solution (decision rule) F, then the values in the next period need to satisfy the decision rule relationship as well. The fact that such a rule F exists comes from the invariance of the stable manifold, not the other way.

[26] Computation of the Jordan normal form is not numerically stable, so the Jordan normal form is only used for the proof.

by

$$\begin{bmatrix}\Phi_{vv} & \Phi_{vg}\\ \Phi_{gv} & \Phi_{gg}\end{bmatrix} = P\begin{bmatrix}T_u & 0\\ 0 & T_s\end{bmatrix}P^{-1} \tag{9}$$

$$P =: \begin{bmatrix}P_{vu} & P_{vs}\\ P_{gu} & P_{gs}\end{bmatrix} \tag{10}$$

where the eigenvalues are ordered such that T_u contains the explosive eigenvalues, and T_s stable.[27] We get from the existence and uniqueness requirement that we have the correct number of stable/explosive eigenvalues for T_u and T_s. Further, we know that we get the 0 in the upper-right of the (block) Jordan normal form as T_u and T_s cannot share eigenvalues.

From existence and uniqueness, we get that P_{gs} is invertible, and we can solve for F with our Jordan normal form with the invertibility. With this Jordan normal form, we have that $\begin{bmatrix}P_{vs}\\ P_{gs}\end{bmatrix}$ form the basis vectors for the stable subspace. Hence, we have $\begin{bmatrix}P_{vs}P_{gs}^{-1}\\ I\end{bmatrix}$ also forms the basis vectors for the stable subspace, i.e.,

$$F := P_{vs}P_{gs}^{-1}.$$

Invertibility of P_{gs} further gives us the block matrix inverse formula of

$$\begin{aligned}
P^{-1} &= \begin{bmatrix}\widetilde{P}_{vu} & \widetilde{P}_{vs}\\ \widetilde{P}_{gu} & \widetilde{P}_{gs}\end{bmatrix}\\
\widetilde{P}_{vu} &:= \left(P_{vu} - P_{vs}P_{gs}^{-1}P_{gu}\right)^{-1}\\
\widetilde{P}_{vs} &:= -\left(P_{vu} - P_{vs}P_{gs}^{-1}P_{gu}\right)^{-1}P_{vs}P_{gs}^{-1}\\
\widetilde{P}_{gu} &:= -P_{gs}^{-1}P_{gu}\left(P_{vu} - P_{vs}P_{gs}^{-1}P_{gu}\right)^{-1}\\
\widetilde{P}_{gs} &:= P_{gs}^{-1} + P_{gs}^{-1}P_{gu}\left(P_{vu} - P_{vs}P_{gs}^{-1}P_{gu}\right)^{-1}P_{vs}P_{gs}^{-1}
\end{aligned}$$

Multiplying out equation (9) with our notation, we get

$$\begin{aligned}
\Phi_{vv} - F\Phi_{gv} &= \left(P_{vu}T_u\widetilde{P}_{vu} + P_{vs}T_s\widetilde{P}_{gu}\right) - F\left(P_{gu}T_u\widetilde{P}_{vu} + P_{gs}T_s\widetilde{P}_{gu}\right)\\
&= (P_{vu} - P_{vs}P_{gs}^{-1}P_{gu})T_u\widetilde{P}_{vu}\\
&= (P_{vu} - P_{vs}P_{gs}^{-1}P_{gu})T_u(P_{vu} - P_{vs}P_{gs}^{-1}P_{gu})^{-1}
\end{aligned}$$

The last line is a Jordan normal form of $(\Phi_{vv} - F\Phi_{gv})$, and hence, the eigenvalues of $(\Phi_{vv} - F\Phi_{gv})$ are precisely the explosive eigenvalues of the original system. □

[27] One can trivially reorder any Jordan normal form to get the required ordering of eigenvalues.

B Continuous Time Formulation

Suppose our linearized system of equations are given by

$$\begin{aligned}\begin{bmatrix} dV \\ dg \end{bmatrix} &= A \begin{bmatrix} V \\ g \end{bmatrix} + \begin{bmatrix} 0 \\ \Gamma \end{bmatrix} dW_t \\ &=: \begin{bmatrix} A_{vv} & A_{vg} \\ A_{gv} & A_{gg} \end{bmatrix} \begin{bmatrix} V \\ g \end{bmatrix} + \begin{bmatrix} 0 \\ \Gamma \end{bmatrix} dW_t \end{aligned} \tag{11}$$

We now consider the planning problem for the information about a shock ε_T that will be realized in the future. Then given the linear form, the solution is given by

$$\begin{bmatrix} V_T \\ g_T \end{bmatrix} = exp\left(T \begin{bmatrix} A_{vv} & A_{vg} \\ A_{gv} & A_{gg} \end{bmatrix} \right) \begin{bmatrix} V_0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_T \end{bmatrix}$$

For notational simplicity, define

$$\begin{bmatrix} \Psi(t)_{vv} & \Psi(t)_{vg} \\ \Psi(t)_{gv} & \Psi(t)_{gg} \end{bmatrix} := exp\left(t \begin{bmatrix} A_{vv} & A_{vg} \\ A_{gv} & A_{gg} \end{bmatrix} \right)$$

Let F be our solution (decision rule) to the linearized system, i.e., $\begin{bmatrix} F \\ I \end{bmatrix}$ is basis vectors of the stable manifold. Then the invariance of the stable manifold gives the condition

$$A_{vv}F + A_{vg} = FA_{gv}F + FA_{gg}. \tag{12}$$

Once the shock realizes, the system needs to be back on the stable manifold to satisfy the transversality condition. Hence, we have

$$\Psi(T)_{vv}V_0 = V_T = Fg_T = F\Psi(T)_{gv}V_0 + F\varepsilon_T$$

which we can solve for V_0 and get

$$V_0 = \left(\Psi(T)_{vv} - F\Psi(T)_{gv}\right)^{-1} F\varepsilon_T. \tag{13}$$

We can further simplify the expression using the following relationship

Lemma 2.

$$\left(\Psi(t)_{vv} - F\Psi(t)_{gv}\right) = e^{t(A_{vv} - FA_{gv})}$$

Proof. This is because

$$\left(\Psi(t)_{vv} - F\Psi(t)_{gv}\right) = \begin{bmatrix} I & -F \end{bmatrix} \begin{bmatrix} \Psi(t)_{vv} & \Psi(t)_{vg} \\ \Psi(t)_{gv} & \Psi(t)_{gg} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

Therefore,

$$\begin{aligned}\frac{d}{dt}\left(\Psi(t)_{vv} - F\Psi(t)_{gv}\right) &= \begin{bmatrix} I & -F \end{bmatrix} \begin{bmatrix} A_{vv} & A_{vg} \\ A_{gv} & A_{gg} \end{bmatrix} \begin{bmatrix} \Psi(t)_{vv} & \Psi(t)_{vg} \\ \Psi(t)_{gv} & \Psi(t)_{gg} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} A_{vv} - FA_{gv} & A_{vg} - FA_{gg} \end{bmatrix} \begin{bmatrix} \Psi(t)_{vv} \\ \Psi(t)_{gv} \end{bmatrix} \\ &= (A_{vv} - FA_{gv})\Psi(t)_{vv} + (A_{vg} - FA_{gg})\Psi(t)_{gv} \\ &= (A_{vv} - FA_{gv})\Psi(t)_{vv} - (A_{vv} - FA_{gv})F\Psi(t)_{gv} \\ &= (A_{vv} - FA_{gv})(\Psi(t)_{vv} - F\Psi(t)_{gv})\end{aligned}$$

where the second to last equality comes from the invariance condition given in equation (12). As our expression satisfies solving the (linear differential equation), we get

$$\Psi(t)_{vv} - F\Psi(t)_{gv} = e^{t(A_{vv} - FA_{gv})}$$

□

Hence, we get

$$V_0 = e^{-T(A_{vv} - FA_{gv})} F\varepsilon_T$$

Now, the main theorem is

Theorem 2. *For a (linearized) GE model with the dynamics given by*

$$\begin{bmatrix} dV_t \\ dg_t \end{bmatrix} = A \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Gamma \end{bmatrix} dW_t$$

where the solution exists and is unique. The eigenvalues of $(A_{vv} - FA_{gv})$ *are equal to the eigenvalues with positive real part, i.e., explosive roots, of* A*.*

Proof. The proof for discrete time applies to continuous time with A's in place of Φ's. □

C Numerically Stable Computation

I compute things in continuous time, but the derivation for discrete time follows the same logic. One can just use equations (16) and (17) for discrete time.

In theory, given that we can compute the initial V_0, we can compute the entire dynamics using $\Psi(t)$. However, this would require exponentiating a matrix with both positive and negative eigenvalues. That will lead to an (exponential) accumulation of rounding errors independent of the direction of computation. Fortunately, we can take a nice decomposition to make all parts of the computation numerically stable. To do so, define

$$V_t^{\perp} := V_t - Fg_t$$

where $V_t^\perp$ is the decision variables not on the stable subspace. Then we have

$$\begin{aligned}
\dot{V}_t^\perp &= \dot{V}_t - F\dot{g}_t \\
&= A_{vv}V_t + A_{vg}g_t - FA_{gv}V_t - FA_{gg}g_t \\
&= A_{vv}V_t^\perp + A_{vv}Fg_t + A_{vg}g_t - FA_{gv}V_t^\perp - FA_{gv}Fg_t - FA_{gg}g_t \\
&= (A_{vv} - FA_{gv})V_t^\perp \\
\dot{g}_t &= A_{gv}V_t + A_{gg}g_t \\
&= A_{gv}V_t^\perp + A_{gv}Fg_t + A_{gg}g_t \\
&= (A_{gv}F + A_{gg})g_t + A_{gv}V_t^\perp
\end{aligned}$$

Stacking the expressions, we get

$$\begin{bmatrix} \dot{g}_t \\ \dot{V}_t^\perp \end{bmatrix} = \begin{bmatrix} A_{gg} + A_{gv}F & A_{gv} \\ 0 & (A_{vv} - FA_{gv}) \end{bmatrix} \begin{bmatrix} g_t \\ V_t^\perp \end{bmatrix}$$

Giving us,

$$\begin{bmatrix} g_{t+\Delta t} \\ V_{t+\Delta t}^\perp \end{bmatrix} = exp\left(\Delta t \begin{bmatrix} A_{gg} + A_{gv}F & A_{gv} \\ 0 & (A_{vv} - FA_{gv}) \end{bmatrix}\right) \begin{bmatrix} g_t \\ V_t^\perp \end{bmatrix}$$

Using Parlett's recurrence,[28] we get

$$\begin{bmatrix} g_{t+\Delta t} \\ V_{t+\Delta t}^\perp \end{bmatrix} = \begin{bmatrix} e^{\Delta t(A_{gg} + A_{gv}F)} & F_{12} \\ 0 & e^{\Delta t(A_{vv} - FA_{gv})} \end{bmatrix} \begin{bmatrix} g_t \\ V_t^\perp \end{bmatrix}$$

where F_{12} is found by solving a Sylvester equation:

$$\begin{aligned}
(A_{gg} + A_{gv}F)F_{12} - F_{12}(A_{vv} - FA_{gv}) \\
= e^{\Delta t(A_{gg} + A_{gv}F)}A_{gv} - A_{gv}e^{\Delta t(A_{vv} - FA_{gv})}
\end{aligned}$$

However, as $A_{vv} - FA_{gv}$ has positive eigenvalues, we do not want to compute positive exponential for numerical stability. Hence, we solve an adjusted equation of

$$\begin{aligned}
&(A_{gg} + A_{gv}F)\widetilde{F}_{12} - \widetilde{F}_{12}(A_{vv} - FA_{gv}) = e^{\Delta t(A_{gg} + A_{gv}F)}A_{gv}e^{-\Delta t(A_{vv} - FA_{gv})} - A_{gv} \\
&\widetilde{F}_{12} := F_{12}e^{-\Delta t(A_{vv} - FA_{gv})}
\end{aligned}$$

resulting in a numerically stable recurrence of

$$g_{t+\Delta t} = e^{\Delta t(A_{gg} + A_{gv}F)}g_t + \widetilde{F}_{12}V_{t+\Delta t}^\perp \tag{14}$$

$$v_{t+\Delta t}^\perp = e^{-(T-t-\Delta t)(A_{vv} - FA_{gv})}F\varepsilon_T. \tag{15}$$

Similar calculations in discrete time give (numerically stable) relationship of

$$g_{l+1} = (\Phi_{gg} + \Phi_{gv}F)g_l + \Phi_{gv}V_l^\perp \tag{16}$$

$$V_l^\perp = (\Phi_{vv} - F\Phi_{gv})^{-(k-l)}F\varepsilon_k. \tag{17}$$

[28]See Davies and Higham (2003) for reference.

D Solving with Augmented States

Suppose that we have the discrete-time formulation

$$\begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} \Phi_{vv} & \Phi_{vg} \\ \Phi_{gv} & \Phi_{gg} \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi_g \end{bmatrix} \varepsilon_{t+1}$$

The standard augmented-state method allow for anticipated shocks to this system of equations by extending the states with

$$\begin{bmatrix} V_{t+1} \\ g_{t+1} \\ v_{t+2|t+1} \\ v_{t+3|t+1} \\ \vdots \\ v_{t+k|t+1} \\ v_{t+k+1|t+1} \end{bmatrix} = \begin{bmatrix} \Phi_{vv} & \Phi_{vg} & 0 & 0 & \cdots & 0 & 0 \\ \Phi_{gv} & \Phi_{gg} & \Psi_g & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & I & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & \cdots & 0 & I \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} V_t \\ g_t \\ v_{t+1|t} \\ v_{t+2|t} \\ \vdots \\ v_{t+k-1|t} \\ v_{t+k|t} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ \Psi_g & 0 & 0 & \cdots & 0 & 0 \\ 0 & I & 0 & \ddots & 0 & 0 \\ 0 & 0 & I & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & I & 0 \\ 0 & 0 & 0 & \cdots & 0 & I \end{bmatrix} \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+2|t+1} \\ \varepsilon_{t+3|t+1} \\ \vdots \\ \varepsilon_{t+k|t+1} \\ \varepsilon_{t+k+1|t+1} \end{bmatrix} \tag{18}$$

Let the decision rule given by

$$V_t = F g_t + \sum_{l=1}^{k} H_l v_{t+l|t}$$

When we consider what happens to any given value $\tilde{v}_{v+l|t}$, then equation (18) becomes,

$$\begin{bmatrix} V_{t+1} \\ g_{t+1} \\ v_{t+2|t+1} \\ \vdots \\ v_{t+l-1|t+1} \\ v_{t+l|t+1} \\ v_{t+l+1|t+1} \\ \vdots \\ v_{t+k+1|t+1} \end{bmatrix} = \begin{bmatrix} \Phi_{vv} H_l \tilde{v}_{t+l|t} \\ \Phi_{vg} H_l \tilde{v}_{t+l|t} \\ 0 \\ \vdots \\ 0 \\ \tilde{v}_{t+l|t} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Hence, the invariant subspace condition becomes

$$\Phi_{vv} H_l \tilde{v}_{t+l|t} = F \Phi_{gv} H_l \tilde{v}_{t+l|t} + H_{l-1} \tilde{v}_{t+l|t}.$$

Since this should hold for all $\tilde{v}_{t+l|t}$, we have

$$\Phi_{vv} H_l = F \Phi_{gv} H_l + H_{l-1}$$

resulting in

$$H_l = (\Phi_{vv} - F\Phi_{gv})^{-1} H_{l-1}$$

The same calculation for $l = 1$ gives

$$H_1 = (\Phi_{vv} - F\Phi_{gv})^{-1} F \Psi_g$$

Hence,

$$H_l = (\Phi_{vv} - F\Phi_{gv})^{-l} F \Psi_g,$$

which is exactly the expression we have found in equation (4).

The actual solution F and H_l's are not computed in this manner, but with a Schur decomposition. Without the finite-algebra/rounding errors, results will be exactly equal to the solution found using equation (4). However, it is hard to prove the numerical stability property of the solution method with Schur decomposition as a part of the computation.[29] Though I failed to prove the stability mathematically, numerical experiments on the three equation model confirms the stability giving exactly the same solution as the decomposition based calculation.

A bigger difference between the two methods comes from the scale of computation. The augmented-states method scale at $O((n + n \times k)^3)$ where k is the anticipation length, and n is the dimensions of the state. In comparison, my method scales linearly in k as we only need to compute one matrix multiplication per length of anticipation, i.e., $O(n^3 k)$. Intuitively, the augmented-states method wastes a lot of time operating on zeros as the Schur decomposition cannot take advantage of the sparsity structures.

In practice, one can just use the augmented-states method to handle anticipated shocks for short horizons allowing one to use pre-existing codebase, but should implement equations (14) and (15) for long horizons.

E Conditional Probability

E.1 Conditional Distribution of Normal Distribution

For our linear model with the given shock structure, the resulting distribution is Gaussian, and the computation of the conditional probability is equivalent to computing the conditional distribution of a normal distribution. This is standard and only presented for completeness.

Suppose we have exogenous shocks and states with uncertainty collected as x

$$x \sim N(\mu_x, \Sigma_x)$$

[29] Though Schur decomposition is numerically stable with respect to perturbations of the values of the dynamics matrix, we require the numerical stability in the horizon of anticipation instead. Hence, stability does not generalize automatically, but it is most likely.

with observation given as

$$y = Hx$$

If we stack x and y, then we get that

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ H\mu_x \end{bmatrix}, \begin{bmatrix} \Sigma_x & \Sigma_x H' \\ H\Sigma_x & H\Sigma_x H' \end{bmatrix}\right)$$

Hence, we can compute the conditional probability given an observation of $y = a$ as

$$(x|y = a) \sim N\left(\overline{\mu}, \overline{\Sigma}\right)$$

$$\overline{\mu} = \mu_x + \Sigma_x H'(H\Sigma_x H')^{-1}(a - H\mu_x) \tag{19}$$

$$\overline{\Sigma} = \Sigma_x - \Sigma_x H'(H\Sigma_x H')^{-1} H\Sigma_x \tag{20}$$

One can refer to section 4.13 of Durbin and Koopman (2012) to see that this is an equivalent formulation for Kalman smoothing.

Further, if we are interested in $z = H_z x$ instead, then we have

$$(z|y = a) \sim N\left(H_z\overline{\mu}, H_z\overline{\Sigma}H_z'\right)$$

E.2 Application to our problem

We will study the information update when the central bank communicates that $r_k = -0.005$ with $r_i = 0$ for $i < k$. There can be many updating methods, but consider for the exercise that agents take this as the true realization of the future, and update their belief about the distribution of the future shocks.

In our notation, we have

$$x = \begin{bmatrix} g_0 \\ \varepsilon_{1|0;\text{z}} \\ \vdots \\ \varepsilon_{k|0;\text{z}} \\ \varepsilon_{1|0;\text{mp}} \\ \vdots \\ \varepsilon_{k|0;\text{mp}} \\ \varepsilon_{1|0;\text{g}} \\ \vdots \\ \varepsilon_{k|0;\text{g}} \end{bmatrix} \sim N(\mu_x, \Sigma_x) = N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \text{diag}\left(\begin{bmatrix} P \\ \Sigma_z \\ \vdots \\ \Sigma_z \\ \Sigma_{mp} \\ \vdots \\ \Sigma_{mp} \\ \Sigma_g \\ \vdots \\ \Sigma_g \end{bmatrix}\right)\right)$$

where P denotes the ergodic distribution of g, Σ's the variances of the shocks, and diag block-diagonal matrix, and

$$y = \begin{bmatrix} r_{1|0} \\ \vdots \\ r_{k|0} \\ r_{k|0} \end{bmatrix} \qquad a = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -0.005 \end{bmatrix}$$

We have used the Σ's corresponding to the variances from the estimation, but different variances can be taken for the future shocks. Similarly, if we were to run the Kalman filter, the distribution g_0 would be given by the filtered state instead of $N(0, P)$ corresponding to the ergodic distribution.

Finally, we can compute the IRFs of $\varepsilon_{i|0;shock}$ by using equations (16) and (17) to get $\begin{bmatrix} V_{i|0} \\ g_{i|0} \end{bmatrix}$, and use the observation matrix H_r (such that $r_i = H_r \begin{bmatrix} V_i \\ g_i \end{bmatrix}$) to fill in H. Given the μ_x, Σ_x, a, and H, we can compute the conditional distribution, $N(\overline{\mu}, \overline{\Sigma})$, using equations (19) and (20).

F Three-Equation Model with Calibration

For my experiments, I use the 3-equation NK model in Herbst and Schorfheide (2015) converted to continuous time. The equations are:

$$
\begin{aligned}
&d\lambda_t = (\rho + \gamma + z_t - (i_t - \pi_t))\,\lambda_t \\
&d\lambda_{\pi,t} = (i_t - \pi_t - \gamma - z_t - (\pi_t - \overline{\pi}))\lambda_{\pi,t} - \varepsilon\left[\xi_H\,(c_t)^{\tau} - \frac{\varepsilon - 1}{\varepsilon}\right] y_t \\
&di_t = \rho_R\,(i^* + \phi_\pi(\pi_t - \pi^*) + \phi_y(\log(y_t) - \log(y_t^*)) - i_t)\,dt + \sigma_R dW_{R,t} \\
&dz_t = -\rho z_t dt + \sigma_z dW_{z,t} \\
&dg_t = \rho_g(g - g_t)dt + \sigma_g dW_{g,t} \\
&\zeta_t = 1 - \frac{1}{e^{g_t}} \\
&(c_t)^{-\tau} = \lambda_t \\
&\left(1 - \zeta_t - \frac{\theta_p}{2}\pi_t^2\right) y_t = c_t \\
&\pi_t = \frac{\lambda_{\pi,t}}{\theta_p y_t} \\
&(1 - \zeta_t) y_t^* = \left(\xi_H \frac{\varepsilon}{\varepsilon - 1}\right)^{-\frac{1}{\tau}}
\end{aligned}
$$

where λ_t is Lagrange multiplier for household budget constraint, $\lambda_{\pi,t}$ Lagrange multiplier for price adjustment firm, i_t interest rate, π_t inflation, y_t GDP, z_t productivity, g_t government expenditure, and c_t consumption. We can manually substitute variables, e.g., λ's, to match the expressions found in Herbst and Schorfheide (2015), but I leave the expressions to be closer to my implementation.[30]

I have replicated the Bayesian estimation of Herbst and Schorfheide (2015), and figure 6 shows that the estimated values are the same for parameters with equivalent definitions in discrete-time and continuous time. However, there are parameters that do not have an exact analogy, e.g., continuous time consumption smoothing against the quarterly consumption

[30] Programs will automatically handle the substitutions for us after automatic differentiation.

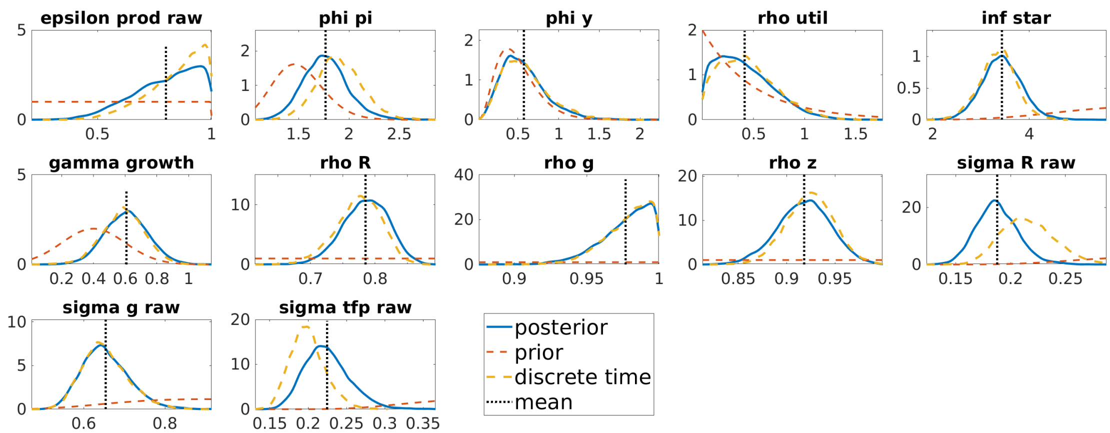


Figure 6: Posterior from Bayesian Estimation

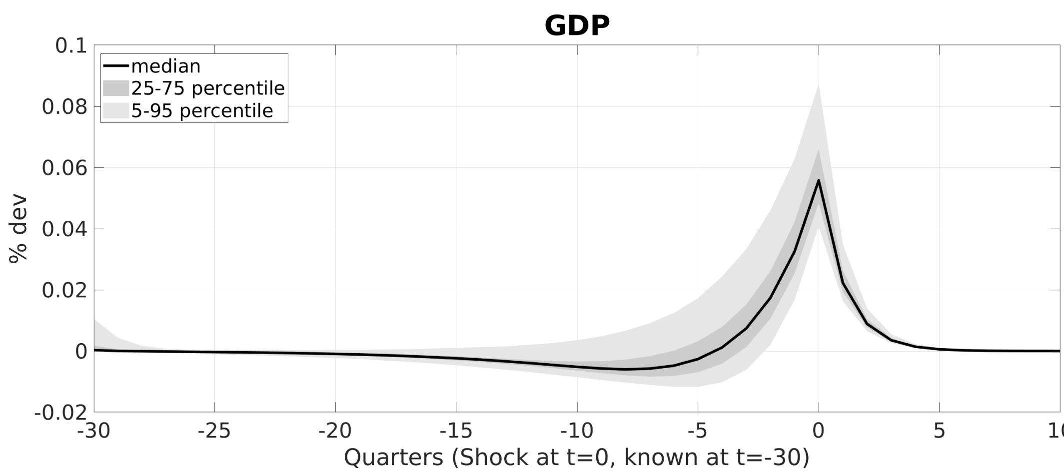


Figure 7: Output response

smoothing parameter. For those parameters, the posterior is not exactly the same, but is similar in shape.

Finally, figures 7 to 9 shows the impulse response functions to future shocks, and confirms that the posterior mode value shown in the main text is typical in responses.

G Automatic Differentiation of Schur Decomposition

A Schur decomposition is defined as a pair (U, T) that satisfies [31]

$$A = UTU' \tag{21}$$

$$U'U = I \tag{22}$$

$$T_{kl} = 0, \text{for } k < l \tag{23}$$

[31] I give expressions for the complex Schur decomposition, so that T is strictly triangular, but the extension to real Schur decomposition follows the same logic.

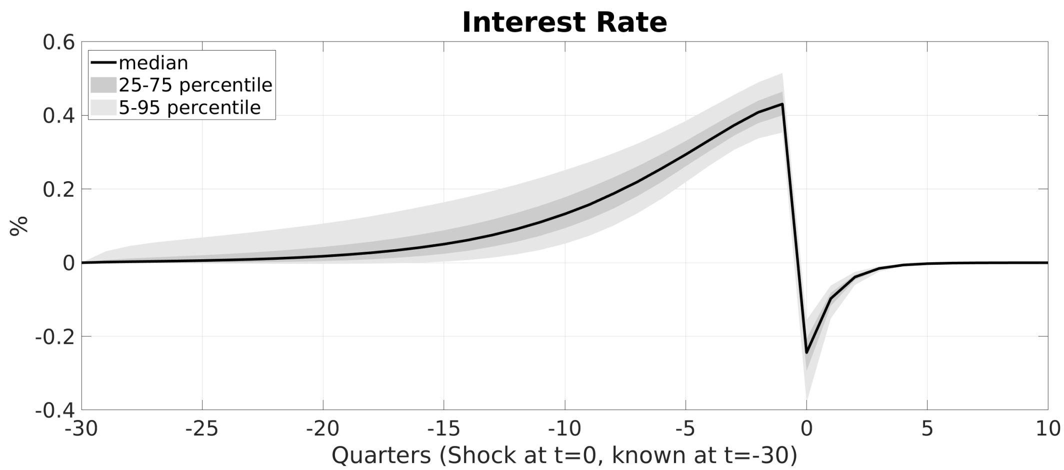


Figure 8: Interest rate response

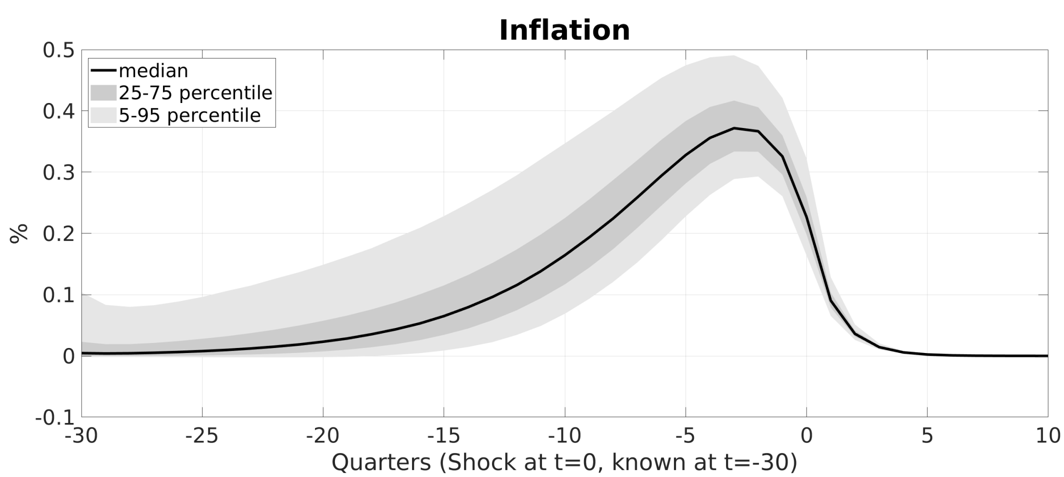


Figure 9: Inflation response

The required conditions/equations are (infinitely) differentiable, so we should be able to apply the automatic differentiation. However, a blind application will fail because Schur decomposition is only unique up to rotations. This is precisely why we can reorder the eigenvalues. Hence, the application of an automatic differentiation toolbox on `<schur>` function will differentiate the "random selection" for one of the Schur decomposition. This results in an ill-behaved derivative. However, we actually need *one* decomposition that satisfies the Schur decomposition requirement and the order requirement of eigenvalues, and we can just differentiate around the particular decomposition, $\overline{U}$ and $\overline{T}$. As equations (21) and (22) are smooth, perturbed T will also satisfy the correct ordering of eigenvalues required for linear solution methods in economics.[32]

For the actual differentiation, the product rule gives:

$$\frac{\partial A_{ij}}{\partial \theta} = \sum_k \sum_l \left[\frac{\partial U_{ik}}{\partial \theta} \overline{T}_{kl} \overline{U}_{jl} + \overline{U}_{ik} \frac{\partial T_{kl}}{\partial \theta} \overline{U}_{jl} + \overline{U}_{ik} \overline{T}_{kl} \frac{\partial U_{jl}}{\partial \theta} \right] \tag{24}$$

$$\sum_k \left[\frac{\partial U_{ki}}{\partial \theta} \overline{U}_{kj} + \overline{U}_{ik} \frac{\partial U_{kj}}{\partial \theta} \right] = 0 \tag{25}$$

$$\frac{\partial T_{kl}}{\partial \theta} = 0, \text{for } k < l \tag{26}$$

Equations (24) to (26) are linear in $\frac{\partial U}{\partial \theta}$ and $\frac{\partial T}{\partial \theta}$, so we can directly solve for $\frac{\partial U}{\partial \theta}$ and $\frac{\partial T}{\partial \theta}$. The same calculation applies to the generalized Schur(QZ)-decomposition and to most matrix decomposition.[33]

For economists, the problem showed up as "the failure of the blind application of automatic differentiation to `<schur>` function," but this is actually a common problem with the automatic differentiation. One can recycle codes when the implementations under the hood matches the mathematical concept, but there are instances where the underlying implementation differs.[34] For a simpler example, consider solving a non-linear function $f(x) = 0$ for x. If the solver uses an iterative method to find x, we should not attach the automatic differentiation to the solver, and differentiate after x^* is found instead. Hence, one should not preemptively conclude the inapplicability of automatic differentiation from a failure of an automatic differentiation toolbox.

[32] Hence, we do not have to worry about differentiation reordering contrary to comments in Childers et al. (2022).

[33] One should take care to understand the requirements of the decomposition, similar to non-uniqueness problem with the Schur decomposition.

[34] A more technical problem is that as most numerically intense operations are outsourced to pre-existing codebase in a faster language, e.g., Fortran, and even Julia is not immune to this problem. Automatic differentiation would fail with these functions even though the operation is perfectly differentiable, and in these instances one cannot recycle the same code for automatic differentiation.

H General Form

Consider a more general linearized system of equations of the form similar to Sims (2002)[35]

$$\Gamma_0 \begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \Gamma_1 \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi_g \end{bmatrix} \varepsilon_{t+1}$$

where Γ_0 is an $n \times n$ matrix. Usually, Γ_0 is "inverted" jointly using the generalized Schur decomposition. This is because Γ_0 can be non-invertible, and there is usually no reason to take additional computational burden to handle Γ_0 separately. However, it is still possible to handle Γ_0 first, and take a regular Schur decomposition after.

Suppose rank of $\Gamma_0 = m < n$, and let U, D, V the singular value decomposition of Γ_0, i.e.,

$$\Gamma_0 = UDV^*$$
$$U^*U = I$$
$$V^*V = I$$

Then since D will have non-zero entry for the first m diagonal entries from the rank of Γ_0, we have

$$\begin{bmatrix} [DV^*]_m \\ 0 \end{bmatrix} \begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} [U^*\Gamma_1]_m \\ [U^*\Gamma_1]_{\neg m} \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} [U^*\Psi]_m \\ [U^*\Psi]_{\neg m} \end{bmatrix} \varepsilon_{t+1}$$

where $[\cdot]_m$ denotes the first m rows of the matrix and $[\cdot]_{\neg m}$ the remaining rows. Since ε_{t+1} cannot affect variables in period T, $[U^*\Psi]_{\neg m} = 0$ resulting in

$$\begin{bmatrix} [DV^*]_m \\ 0 \end{bmatrix} \begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} [U^*\Gamma_1]_m \\ [U^*\Gamma_1]_{\neg m} \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} [U^*\Psi]_m \\ 0 \end{bmatrix} \varepsilon_{t+1}$$

Time can be shifted once, to get

$$\begin{bmatrix} [DV^*]_m \\ [U^*\Gamma_1]_{\neg m} \end{bmatrix} \begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} [U^*\Gamma_1]_m \\ 0 \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} [U^*\Psi]_m \\ 0 \end{bmatrix} \varepsilon_{t+1}$$

Now, the matrix on the left-hand side is full-rank. Suppose for contradiction that the matrix is not full-rank. Then the same process as above will result in rows will all zeros. This means that there are fewer equations than the number of variables, i.e., a contradiction.

Therefore, we have

$$\begin{aligned} \begin{bmatrix} V_{t+1} \\ g_{t+1} \end{bmatrix} &= \begin{bmatrix} [DV^*]_m \\ [U^*\Gamma_1]_{\neg m} \end{bmatrix}^{-1} \begin{bmatrix} [U^*\Gamma_1]_m \\ 0 \end{bmatrix} \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} [DV^*]_m \\ [U^*\Gamma_1]_{\neg m} \end{bmatrix}^{-1} \begin{bmatrix} [U^*\Psi]_m \\ 0 \end{bmatrix} \varepsilon_{t+1} \\ &=: \Phi \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} \Psi_V \\ \Psi \end{bmatrix} \varepsilon_{t+1} = \Phi \begin{bmatrix} V_t \\ g_t \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi \end{bmatrix} \varepsilon_{t+1} \end{aligned}$$

[35] I do not carry the η terms in the equation form of Sims (2002) because I separate variables into the forward-looking, V_t, and the backward-looking, g_t, variables explicitly. V_t will jump to guarantee that the dynamics stays on the stable manifold, similar to η's in Sims (2002). It is possible to convert the Sims's form with η_t into our form using a similar calculation.

where the last equality is because V_t's are forward-looking variables, and Ψ_V will not affect V_{t+1} as the variable will "jump" to satisfy the transversality condition.

Handling of Γ_0 is more natural in continuous time and is addressed in Ahn et al. (2018), but it is faster just to work out the algebra directly.