Resolving New Keynesian Puzzles

Maria Eskelinen* Christopher G. Gibbs† Nigel McClung‡

March 6, 2024

Abstract

New Keynesian models generate puzzles when confronted with the zero lower bound (ZLB) on nominal interest rates (e.g. the forward guidance puzzle or the paradox of flexibility). We show that these puzzles are absent in simple and medium-scale models when monetary policy approximates optimal policy, even loosely. The standard approach to modeling monetary policy at the ZLB does not approximate the policy a rational inflation targeting central bank would choose at the ZLB. It is this disconnect that is responsible for the puzzles. The puzzles, therefore, are best thought of as the plausible predictions of implausible monetary policy rather than implausible predictions to plausible monetary policy. We show how to write monetary policy rules that capture the same policy objective with and without the ZLB.

*University of Oxford

†School of Economics, The University of Sydney, Camperdown, NSW 2006. Email: christopher.gibbs@sydney.edu.au. I acknowledge support from ARC Discovery Project Grant 210101204. We thank Scott French, Jonathan Payne, Michael Boutros, Callum Ryan, and Mariano Kulish for helpful comments. We thank participants of the MRG Workshop hosted by UNSW Sydney and the OzMac Workshop hosted by Monash University for comments.

‡Bank of Finland. The views expressed in this paper do not necessarily reflect those of the Bank of Finland.

1 Introduction

While the New Keynesian (NK) model provides an elegant theory of monetary policy during the Great Moderation, when confronted by the zero lower bound (ZLB) on nominal interest rates, it has yielded a litany of puzzles and paradoxes. The most well-known is the so-called forward guidance puzzle, which is the prediction that credible promises to hold interest rates at zero percent generate significant stimulus. So much stimulus that forecasting actual policies enacted by the Federal Reserve, European Central Bank, or the Bank of Japan causes otherwise useful models to *explode.*

To illustrate what we mean by explode, consider the plight of economists in August of 2011 attempting to provide forecasts based on the following FOMC statement:

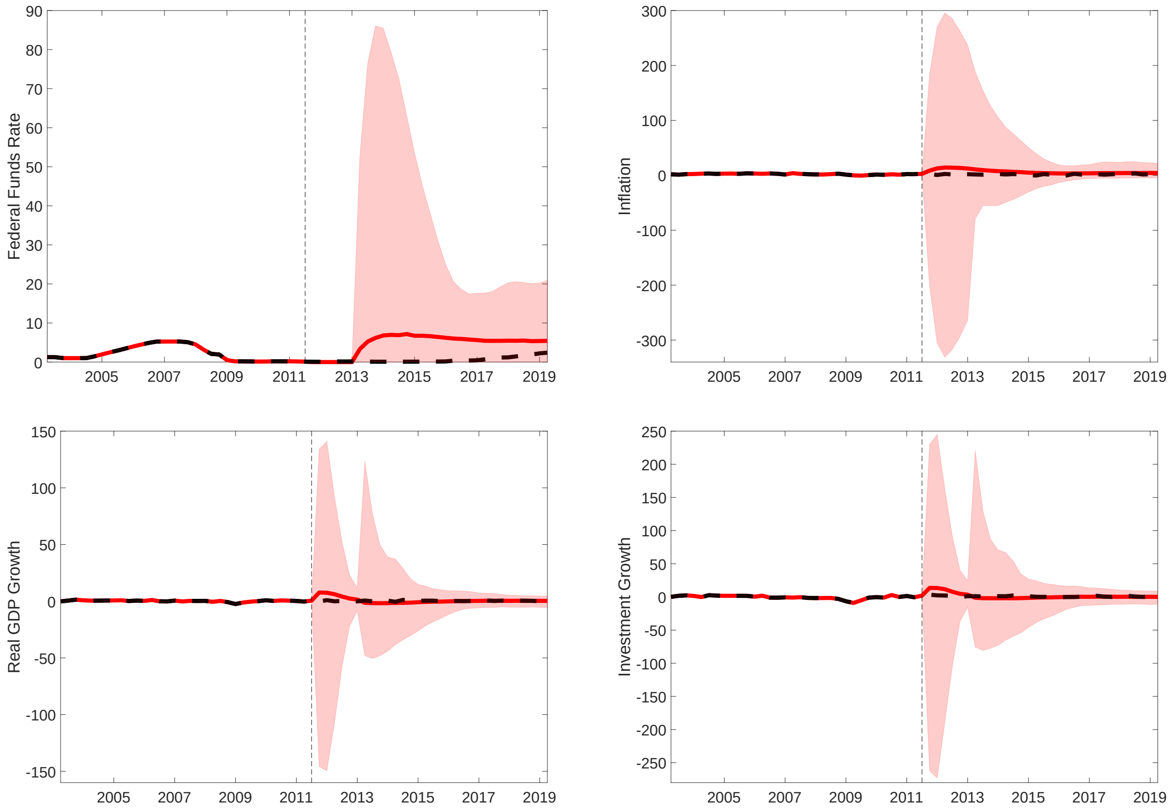
> *The Committee currently anticipates that economic conditions ... are likely to warrant exceptionally low levels for the federal funds rate at least through mid-2013.*
>
> \- FOMC Statement August, 9^{th} 2011

Figure 1 shows what forecasts of seven quarters of zero interest policy are predicted to do for inflation, real GDP growth, investment, and the Federal Funds rate using the model of Smets and Wouters (2007) estimated on data ending in 2004.[1] With forecasts like these, it is not surprising that some of the first papers documenting the forward guidance puzzle came from within the Federal Reserve System. For example, the work by Del Negro, Giannoni and Patterson (2012) or Carlstrom, Fuerst and Paustian (2015), which to our knowledge are among the first papers to study the issues in Figure 1.

In addition, there are at least three other ZLB NK puzzles widely studied in the literature: the fiscal multiplier puzzle – excessive responses to announced fiscal policy changes (e.g., Farhi and Werning, 2016), the paradox of toil – contractionary effects for *positive* productivity or labor supply shocks (e.g., Eggertsson, 2010), and the paradox of the flexibility – larger effects

[1] The forecasts are constructed using Smets and Wouter's original replication files and data (which ends in 2004) to estimate the posterior distribution of the structural parameters. We then follow Cagliarini and Kulish (2013), Jones (2017), Kulish, Morley and Robinson (2017), and Kulish and Pagan (2017) to generate forecasts by sampling from the estimated posterior distribution of the parameters while enforcing a fully credible and known policy of zero interest rates for seven quarters. After seven quarters, monetary policy is governed by an occasionally binding constraints algorithm, where the policy rate is the maximum of zero or the interest rate implied by the model's interest rate rule.

Figure 1: Calendar-based forward guidance and the forward guidance puzzle



Notes: Out-of-sample Q/Q forecasts of the August, 9^{th} 2011 Federal Reserve calendar-based forward guidance promise using the model of Smets and Wouters (2007). The dashed line is the actual data. The solid line is the median forecast. The shaded region represents the 95th and 5th percentiles of forecasts. Inflation and the Federal Funds Rates are annualized.

of monetary policy when prices are *more* flexible (e.g., Eggertsson and Krugman, 2012 or Kiley, 2016). Like with the forward guidance puzzle, economic intuition and often empirical evidence is starkly at odds with these predictions.

We show that the NK puzzles have a single origin: a disconnect between the policy objectives of the modeler and the modeled policymaker. The convention in the NK literature is to capture monetary policy using a reduced form Taylor (1993) type rule. These rules are justified on the grounds that they can approximate optimal policy under discretion or commitment – depending on the specification – and much research has been done to establish this connection without considering the ZLB (see, for example, Woodford, 2001 or Woodford, 2003b). We show, however, that when these rules are adapted to the ZLB they no longer approximate optimal policy – or any rational policy for that matter – of an optimizing central bank. Replacing these rules with ones that approximates, even loosely, the actual choices an optimizing central bank would make at the ZLB can significantly reduce or eliminate the NK puzzles.

Eggertsson and Woodford (2003) show that optimal monetary policy at the ZLB requires history dependence. A policymaker should commit to make-up for current misses from their targets in the future because of their inability to sufficiently lower the policy rate today. That commitment can imply interest rates that remain at zero for longer than they would in the absence of any commitment. Zero interest rate policy, therefore, is a history dependent commitment. Interest rates remain low, or not, depending on what has occurred in the past.

History dependent commitments are absent in the standard approach to modeling the ZLB constraint. Typically, monetary policy is described in NK models by a rule of the form:

$$i_t = (1-\rho_i)\bar{r} + \rho_i i_{t-1} + (1-\rho_i)(\phi_\pi \pi_t + \phi_y y_t), \tag{1}$$

where i_t is the policy rate, π_t is inflation (assuming a zero percent inflation target), y_t is a measure of real activity (usually the output gap), and $\bar{r}$ is the steady state real interest. The parameter ϕ_π and ϕ_y captures the responsiveness of the central bank to current inflation and real activity, respectively, while ρ_i captures a preference for gradual policy adjustment.[2] History dependence is communicated here through $0 < \rho_i < 1$, which encodes past movements in inflation and output shaping expectations of future policy.

To incorporate the ZLB, most researchers – including us in Figure 1 – retain Rule (1) and simply consider the following modification:

$$i_t = \max\left\{0, (1-\rho_i)\bar{r} + \rho_i i_{t-1} + (1-\rho_i)(\phi_\pi \pi_t + \phi_y y_t)\right\}. \tag{2}$$

This rule removes any history dependence to policy. Lagged interest rates no longer capture past economic conditions once the ZLB is hit. Regardless of what occurs at the ZLB under Rule (2), the monetary policymaker credibly commits to not respond to it. Bygones are bygones in the most strict sense when a model is closed with an interest rate rule like (2).

The lack of history dependence found in Rule (2) is counterfactual to how actual central banks behave at the ZLB. As evidenced by the earlier quote from the FOMC, policy is described as dependent on the evolution of economic data. It is also rejected by NK models when

[2]The discussion here is based on the study of structural monetary models that are solved by standard first-order approximations using piece-wise solutions to capture the non-linearity at the ZLB. This assumption captures nearly the universe of models that are used in central banks and the simple environments in which the forward guidance policy is usually studied. Eggertsson and Singh (2019) shows considering the non-linear environments does not eliminate the puzzles. In other words, the NK puzzles are not an artifact of approximation method.

estimated on data outside of ZLB episodes. For example, both Smets and Wouters (2007) and Del Negro, Eusepi, Giannoni, Sbordone, Tambalotti, Cocci, Hasegawa and Linder (2013) find ρ_i to be around 0.8, which indicates significant history dependence to policy from the perspective of the agents that inhabit the model. Therefore, from the agents' perspective, the ZLB under Rule (2) is not just a constraint on the instrument of policy, it is a policy regime change.

To understand why Rule (2) represents *policy* regime change, consider an equivalent representation of Rule (2):

$$i_t = \max\left\{0, \bar{r} + \phi_\pi \omega_t^\pi + \phi_y \omega_t^y\right\}, \tag{3}$$

where

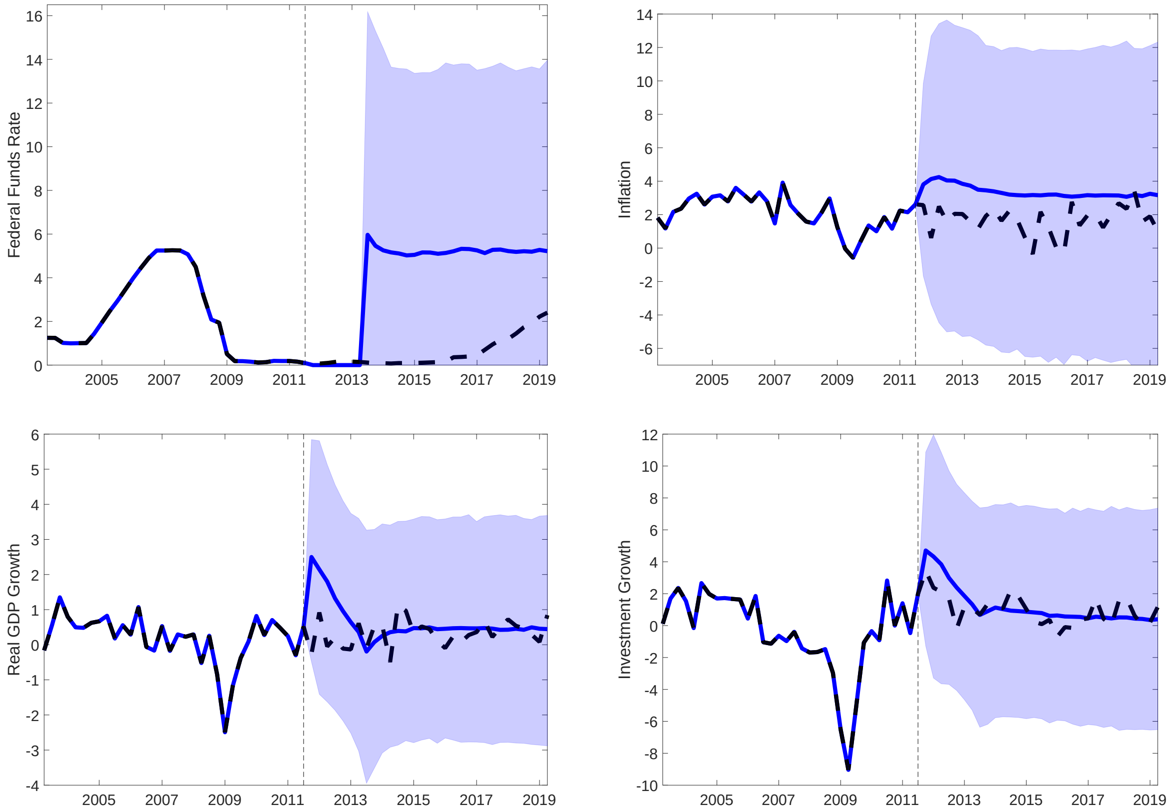
$$\begin{aligned} \omega_t^\pi &= \omega_{t-1}^\pi + (1-\rho_i)(\pi_t - \omega_{t-1}^\pi) \\ \omega_t^y &= \omega_{t-1}^y + (1-\rho_i)(y_t - \omega_{t-1}^y) \end{aligned}$$

are weighted averages of past inflation and output. Models with either Rule (2) or Rule (3) implement the same equilibrium outcomes in standard linearized economies when the ZLB never binds.[3] However, Rules (3) makes clear the degree of history dependence that is implicitly assumed in Rule (1). The central bank responds to geometric averages of past inflation and output. High weight is placed on observations that occurred far in the past when ρ_i is large. Quite naturally, the history that policymakers are responding to is that of past inflation and output, consistent with real world statutory mandates.

Rule (2) and Rule (3) make very different predictions for policy at the ZLB. Under Rule (2), exogenously setting $i_t = 0$ deletes the central bank and its objectives from the model. In addition, upon lift off of the interest rate, Rule (2) does not return to approximating an optimizing central bank. Instead, the central bank down-weights its response to current output and inflation by doggedly refusing to raise interest rates fast enough. Under Rule (3), however, exogenously setting $i_t = 0$ does not change what the central bank fundamentally cares about in the model. It simply implements exactly what the ZLB represents: a constraint on the choice

[3] The equivalence holds for the rational solution of a first-order approximation of the economy in log-deviation from steady state form when initial conditions are the same, $\bar{r}$ is constant, and there is no monetary policy shock. Monetary policy shocks obviously propagate differently in equilibrium between these two rules. Likewise, if $\bar{r}$ tracks the natural rate shock, then the rules implement different policies, which is discussed in detail in Section 2.

Figure 2: Calendar-based forward guidance without the forward guidance puzzle



Notes: Out-of-sample Q/Q forecasts of the August, 9^{th} 2011 Federal Reserve calendar-based forward guidance promise using the model of Smets and Wouters (2007) with the interest rate rule written in the form of (3). The dashed line is the actual data. The solid line is the median forecast. The shaded region represents the 95th and 5th percentiles of forecasts. Inflation and the Federal Funds Rates are annualized.

of the instrument of policy. The central bank and the private sector do not lose their ability to track the evolution of output and inflation during the period interest rates are constrained; nor does the private sector lose their ability to formulate expectations about how policymakers will respond to current misses of their targets in the future.

Figure 2 shows how the out-of-sample forecasts from Figure 1 change when we transform the interest rate rule of the Smets and Wouters model to the form of Rule (3).[4] The forecasts no longer explode. The responses are large, but quite sensible. Forward guidance here provides a significant boost to the economy, larger than what occurred on average, but not unreasonable given that the policy is assumed perfectly credible in the forecasts. We show that for empirically relevant exercises like this one, the New Keynesian puzzles are more-or-less absent at the posterior estimates of the model parameters applied to Rule (3). However, for extremely long expected ZLB episodes some puzzles remain. To remove all puzzles, a central bank must follow a rule that more than makes-up for past misses in inflation from the target at the ZLB, which

[4]It is straightforward to verify that it implements the same rational solution when the ZLB is not considered. The exact formulation of the rule we use is given in Section 4.

is what Eggertsson and Woodford (2003) show characterizes optimal policy.

We establish the disconnect between monetary policy rules and approximating optimal policy at the ZLB by revisiting standard results from the NK literature. We show that when the ZLB is not imposed, the unconditional optimal target criteria proposed by Blake (2001) and Jensen and McCallum (2002) may be approximately implemented by a rule like Rule (1). However, its approximation to optimal policy does not extend to the ZLB case. Here we follow Eggertsson and Woodford (2003) and derive optimal policy to a real interest rate shock that causes a one-time bind of the ZLB for an unknown duration. From the optimal policy solution, we recover the optimal state-contingent forward guidance policy. We pair that policy with either Rule (2) or Rule (3) and compare outcomes. Rule (3) approximates the optimal policy outcomes quite well. In contrast, under Rule (2), the usual forward guidance puzzle results are observed. Therefore, only Rule (3) approximates optimal policy with and without the ZLB.

To quantify the role history dependence plays in ameliorating the puzzles, we derive closed-formed solutions to the standard NK model under a general policy rule that responds to a weighted average of all past inflation. We show that the severity of the puzzles is a function of how much or how little of the *right* history dependence is assumed. The severity of the puzzles is decreasing as function of how much weight the central bank places on past inflation outcomes. Policy must be more history dependent than a price level targeting regime to eliminate the puzzles. In other words, policy must follow a rule that more than makes up for past misses at the ZLB – in accordance with the optimal policy of Eggertsson and Woodford (2003) – to eliminate all puzzles.

We believe there are three important takeaways. First, the representative agent NK (RANK) model has fewer flaws capturing ZLB episodes than is currently recognized conditional on modeling monetary policy with the appropriate history dependence. This is a point that other researchers have only narrowly recognized when exploring shadow interest rates measures proposed by Krippner (2013) and Wu and Xia (2016) in place of a policy rate with a binding ZLB. For example, Hills and Nakata (2018) and Bonciani and Oh (2023) show discrepancies between shadow rates and Rule (2) can explain the fiscal forward guidance puzzle and the paradox of flexibility.[5]

[5] Although, other issues with using shadow rates remain in fully capturing the ZLB as explained in Krippner (2020).

Second, even without the NK puzzles, the RANK model predicts incredibly powerful general equilibrium effects. For example, the optimal policy recommendation of Eggertsson and Woodford (2003), which show that a central bank can almost perfectly stabilize the economy at the ZLB for an arbitrarily large shock, does not rely on, or is an example of, the forward guidance puzzle. Therefore, the large literature that has developed to eliminate the puzzles by dampening general equilibrium effects of expectations through bounded rationality as in Angeletos and Lian (2018), Farhi and Werning (2019), García-Schmidt and Woodford (2019), Gabaix (2020), Eusepi, Gibbs and Preston (2022), Evans, Gibbs and McGough (2022) still have an important role to play in explaining economic dynamics and in the design of policy.

Finally, our results echo the concerns raised by Brassil, Ryan and Yadav (2023) on the misuse of ad hoc policy rules and we argue for caution when modifying interest rate rules in isolation. It has become standard to ignore optimizing behavior by policymakers in quantitative structural modeling with researchers studying a variety of modifications to policy rules and describing them as capturing real world policy objectives when in fact there may be no such connection justified in the model.

1.1 Related Literature

The New Keynesian puzzles literature is large and has two main branches. The most significant branch from a quantitative and policy perspective is on the forward guidance puzzle. The seminal papers include Del Negro et al. (2012) and Carlstrom et al. (2015), which both illustrate the problem – explosive responses of output and inflation to promised pegs of the interest rate – and posit solutions. The latter show that sticky information can reduce or eliminate puzzles while the former argues that the credibility of the policy at longer horizons is implausible. In fact, Del Negro et al. (2012)'s point on credibility of the policy makes a similar argument to ours. The implications of the forward guidance thought experiment imply future movements of interest rates that are implausible. We quantify their insight and connect it to optimal policy.

Many papers have sought to eliminate the forward guidance puzzle by changing primitive assumptions of the NK model. For example, Del Negro, Giannoni and Patterson (2023) makes use of a perpetual youth model to micro found a reduction in the horizon of agents' expectations delivering the same ameliorative effect on forward guidance as imperfect credibility. Further refinements on using credibility to resolve the puzzle are studied by Haberis, Harrison and

Waldron (2019) and Gibbs and McClung (2023). They both show that partial credibility of holding rates at zero resolves the forward guidance puzzle. Andrade, Gaballo, Mengus and Mojon (2019) study the case where some agents interpret forward guidance as worse economic conditions rather than a credible promise of stimulus, which also ameliorates the effect of such policies.

There are many studies on information frictions or bounded rationality as a way of resolving the puzzle such as Kiley (2016) using sticky information; Angeletos and Lian (2018), Farhi and Werning (2019), and Evans et al. (2022) using level-k reasoning; Gabaix (2020) myopia; and Eusepi et al. (2022) who use adaptive learning. Each these bounded rationality papers makes use of the fact that the significant impacts of forward guidance come through the general equilibrium effects of expectations. Bounded rationality lowers these effects and may eliminate the puzzles.

Another sub-strand of this literature is incomplete markets. McKay, Nakamura and Steinsson (2016) and Hagedorn, Luo, Manovskii and Mitman (2019) show that incomplete markets in a heterogeneous agent NK model can resolve the puzzle. Eggertsson, Mehrotra and Robbins (2019) show that the puzzle is absent in overlapping generation models with debt constraints. However, Farhi and Werning (2019) and Bilbiie (2020) show that incomplete market is not a robust way to eliminate the forward guidance puzzle.

In addition to these behavioral approaches, Cochrane (2017, 2023), McClung (2021), Gibbs and McClung (2023), and Diba and Loisel (2021) show that alternative monetary and fiscal policy frameworks also can eliminate the puzzles. The last paper listed shows that the economy is puzzle free under an interest rate peg when monetary policy is conducted via money supply rules. The remaining papers show that closing a model with active fiscal policy as under the Fiscal Theory of the Price Level provides a puzzle free equilibrium in the NK model. Our paper fits into this strand of the literature. The policies studied by these authors resolve the forward guidance puzzle because the relevant history dependence of policy is maintained during the period of constrained interest rates.

The second smaller branch of this literature focuses on the other puzzles: fiscal multipliers, toil, and flexibility. The fiscal multiplier puzzle is essentially the same as the forward guidance puzzle in that anticipated fiscal policy has implausibly large effects. Hills and Nakata (2018) offer a resolution for this puzzle that involves tracking a shadow rate at the ZLB, which encodes the correct history dependence. Eggertsson (2010) identified the paradox of toil as the result that negative supply shocks are expansionary at the ZLB. The paradox of flexibility is defined in Eggertsson and Krugman (2012) and is the result that reductions in price stickiness increases the relative strength of all the other puzzles at the ZLB, i.e., in the face of current and anticipated shocks, the ZLB constraint is even more disruptive in the model as price flexibility increases. Bonciani and Oh (2023) note that a shadow rate here to can eliminate the flexibility puzzle. These puzzles are often, but not always, subordinate to the forward guidance puzzle. If you eliminate the forward guidance puzzle, then typically that removes the other puzzles as well but not vice versa. In addition, Wieland (2019) provide an empirical test for the paradox of toil. No empirical support is found for this puzzle, which is consistent with our argument that monetary policy in reality is not described well by Rule (2).

2 Inertial Interest Rate Rules and approximating optimal policy

We revisit optimal policy in the standard New Keynesian environment under the timeless perspective. We show that optimal commitment policy in the NK model may be approximated in the absence of the ZLB by either an interest rate rule like (1), the standard inertial rule, or by a rule like (3), which responds to weighted averages of past inflation and output, which we call a weighted average rule. The approximations, however, diverge greatly at the ZLB.

2.1 Classic optimal policy in the NK model and inertial rules

Consider the standard optimal monetary policy problem with commitment from the timeless perspective described by Clarida, Gali and Gertler (1999) or Woodford (2003a). The central bank seeks to maximize household welfare by committing to a policy to offset shocks to the real interest rate and to markups. The central bank seeks to minimize

$$\min_{\pi_t, y_t, i_t} \left\{ \frac{1}{2} E_t \sum_{T=t}^{\infty} \beta^{T-t} (\pi_t^2 + \alpha y_t^2) \right\}, \tag{4}$$

which is a quadratic approximation to household utility where α is the weight given to variation in output (y_t) relative to inflation (π_t). Alternatively, we may view the loss function as that of an inflation targeting central bank with a dual mandate. The central bank seeks to minimize deviation in inflation from a target ($\bar{\pi} = 0$ in this case) while considering the output costs when implementing policy. We assert that most modelers roughly have in mind an objective of the form of (4) whenever they write down a reduced form monetary policy rule. The tacit assumption in the literature is that Taylor-type monetary policy rules approximately implement policy that minimizes (4).

That assumption is mostly true when the ZLB is not present. Woodford (2001) shows this for non-inertial interest rate rules. Woodford (2003b) shows it is true of inertial interest rate rules that respond to lagged interest rates. It is not true, however, when the ZLB is considered. To understand why, consider the optimal policy problem of a central bank that seeks to minimize (4) taking as given the first order conditions for household's and firm's decisions:

$$y_t = E_t y_{t+1} - \frac{1}{\sigma}(i_t - E_t \pi_{t+1} - r_t^n)) \tag{5}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa y_t + \mu_t \tag{6}$$

where Equations (5) and (6) are the standard NK IS and Phillips curves log-linearized around a zero inflation steady state, i_t is the policy rate, σ^{-1} is the intertemporal elasticity of substitution, r_t^n is a potentially autoregressive real interest rate shock with mean $\bar{r}$, β measures the rate of time preference, κ is composite parameter capturing the degree of price rigidity, and μ_t is a potentially autoregressive cost push shock. When the ZLB is considered as a possible constraint, the solution must also satisfy $i_t \geq 0$.

2.2 Approximating policy without the ZLB

The optimal target criterion that minimizes the central bank's loss function (4) from the timeless perspective, ignoring the ZLB constraint, is

$$y_t - y_{t-1} = -\frac{\kappa}{\alpha}\pi_t. \tag{7}$$

However, it is illustrative to study a refinement to the target criteria: the *unconditional* optimal target criteria proposed by Jensen and McCallum (2002) and Blake (2001). The former shows numerically and the latter analytically that welfare on average is improved when time discounting by the central bank is ignored resulting in the targeting criteria:

$$y_t - \beta y_{t-1} = -\frac{\kappa}{\alpha}\pi_t. \tag{8}$$

This criterion is convenient because we can write this as

$$(1 - \beta L)y_t = -\frac{\kappa}{\alpha}\pi_t$$

where L is the lag operator. The inverse of $(1 - \beta L)$ exists provided $|\beta| < 1$. Therefore, we can write

$$y_t = -\frac{\kappa}{\alpha}\frac{\pi_t}{1 - \beta L}.$$

Proposition 1: *The optimal target criterion (8) may be implemented by either of the following interest rate rules*

$$\textit{Optimal Rule 1} \quad \begin{cases} i_t &= \frac{\sigma\kappa}{\alpha(1-\beta)}\omega_t^\pi + \sigma E_t y_{t+1} + E_t \pi_{t+1} + r_t^n \\ \omega_t^\pi &= \omega_{t-1}^\pi + (1-\beta)(\pi_t - \omega_{t-1}^\pi) \end{cases} \tag{9}$$

$$\textit{Optimal Rule 2} \quad i_t = \beta i_{t-1} + \frac{\sigma\kappa}{\alpha}\pi_t + (1 - \beta L)\left(\sigma E_t y_{t+1} + E_t \pi_{t+1} + r_t^n\right) \tag{10}$$

From Proposition 1, there are two things to note. First, optimal policy may be implemented with either a rule that includes lagged interest rates like Rule (1), or one that includes a weighted average of past inflation like Rule (3). Either formulation implements the same policy when the ZLB is not present. Second, history dependence is more complicated in the second formulation (10) than in the first (9). To properly implement optimal policy, the lagged interest rate is not a sufficient statistic for history. Past forecasts and shocks are also required. In contrast, past inflation is all that is required to implement the rule (9).

Consider what this means for approximating optimal policy with a rule like (1):

$$i_t = (1 - \rho_i)\bar{r} + \rho_i i_{t-1} + (1 - \rho_i)(\phi_\pi \pi_t + \phi_y y_t),$$

versus a rule like (3):

$$
\begin{aligned}
i_t &= \bar{r} + \phi_\pi \omega_t^\pi + \phi_y \omega_t^y \\
\omega_t^z &= \omega_{t-1}^z + (1-\rho_i)(\pi_t - \omega_{t-1}^z) \text{ for } z = \pi \text{ or } y
\end{aligned}
$$

when the central bank is faced only with demand shocks. Both rules fail to deliver the Blanchard and Galí, 2007's Divine Coincidence, i.e., perfectly stabilizing output and inflation in response to r_t^n. Each rule only responds to r_t^n indirectly. However, only (3) offers a simple modification that engineers this result. Replace $\bar{r}$ with r_t^n in the weighted average inflation rule and it more closely approximate (9). However, do this substitution in the second rule (1) and it may actually worsen the approximation to optimal policy if r_t^n is persistent.

Why does the approximation to optimal policy worsen for Rule (1) when $\bar{r}$ is replace with r_t^n? Because the lagged interest rate encodes the wrong history dependence when policy responds to r_t^n. The additional lag of the shock and expectations in (10) are there to undo the wrong history dependence that is encoded in lagged interest rates when responding to demand shocks.

2.3 Approximating policy with the ZLB

The natural question here is if lagged interest rates are an issue, then why do the puzzles occur even when the policy rule includes no history dependence? Such as when a rule like

$$
i_t = \bar{r} + \phi_\pi \pi_t + \phi_y y_t
$$

is considered. The answer is that history dependence is an integral part of any sensible policy at the ZLB. The wrong history dependence (i.e. lags of i in the rule) or the omission of history dependence to policy is the problem.

Eggertsson and Woodford (2003) show that the optimal target criterion when $i_t \geq 0$ is a time-varying price level target:

$$
\begin{aligned}
\tilde{p}_t &= p_t^* \\
\tilde{p}_t &= p_t + \frac{\alpha}{\kappa} x_t \\
p_{t+1}^* &= p_t^* + \beta^{-1}\left((1+\kappa\sigma^{-1}) - L\right)(p_t^* - \tilde{p}_t),
\end{aligned}
$$

where $\tilde{p}_t$ adjusts according to past misses in that target $(p_t^* - \tilde{p}_t)$ and L is again the lag operator. This criterion requires the central bank to be extremely history dependent. The central bank should not just be a price level targeter by perfectly making up for past misses. It should promise to permanently overshoot on the price level in response to shocks that cause the ZLB to bind – that is, policy should do *more* than make-up for past misses.

Implementing or approximating the optimal target criterion using an interest rate rule requires forward guidance. A central bank that follows Rule (2) or Rule (3) may announce how long they intend to hold interest rates at zero in response to a shock. But that promise is explicitly paired with a promise to return to an interest rule that includes how policy will adjust to what occurred during the ZLB episode. Rule (2) of course promises to not respond at all to what has occurred during the ZLB episode.

To illustrate the impact of either the wrong or an omission of history dependence to policy, we proceed in two steps. First, we replicate the ZLB thought experiment explored by Eggertsson and Woodford (2003) of optimally responding to a real interest rate shock of uncertain duration that causes the ZLB to bind. The optimal policy to this shock implies a state-contingent forward guidance promise, where the number of quarters of zero interest policy is indexed by the expected duration of the shock. Second, we use the optimal state-contingent forward guidance promise to simulate outcomes to the same shock under a promise that policy returns to either Rule (2) or Rule (3), rather than optimal policy. We then compare the equilibrium outcomes.

The real interest rate shock follows a two-state reducible Markov process. In period one, the shock occurs and $r_t^n = r_S < 0$. We call this state S. The shock remains in effect in each period with probability $(1-\delta)$. With complementary probability δ, the real interest rate returns to steady state, $r_t^n = r_N = \bar{r} > 0$. We call this state N. The central bank and the private sector understand the shock process. The central bank responds in the period the shock occurs with a state-contingent forward guidance promise, where for every possible duration of the shock, $\tau = 1, 2, 3, ...$, the central bank provides a promised duration of additional periods of zero interest rate policy in state N, e.g., $k_\tau = \{1, 2, 2, 3, 3, 4, ...\}$, which generate the appropriate initial conditions for optimal policy forever after.

Figure 3 shows the equilibrium outcomes under optimal policy compared to the approximations using either Rule (2) or Rule (3) for realizations of the shocks lasting one through ten quarters. For comparison purposes, we have highlighted the paths associated with a shock lasting four quarters. We use the same calibration for Eggertsson and Woodford (2003) with $\beta = 0.99$, $\sigma = 2$, $\kappa = 0.02$, $r_N = 0.1$, $r_S = -0.005$ and $\delta = 0.1$.[6] For the policy rules, we assume $\rho_i = 0.8$, $\phi_\pi = 1.5$, and $\phi_y = 0.5$. Lastly, we set $\alpha = \kappa/(\sigma\phi_\pi)$ so that the weight placed on stabilizing inflation relative to output is comparable to the standard interest rate rule coefficients chosen.[7] By construction, the forward guidance policy is the same under all three specifications. The only difference is the policy pursued after the interest rate lifts off from zero.

The equivalence of Rule (2) and Rule (3) is broken. The two rules generate extremely different equilibrium dynamics and welfare outcomes in response to the same shock and forward guidance policy. The weighted average rule continues to approximate optimal policy, while the inertial rule exhibits the forward guidance puzzle.

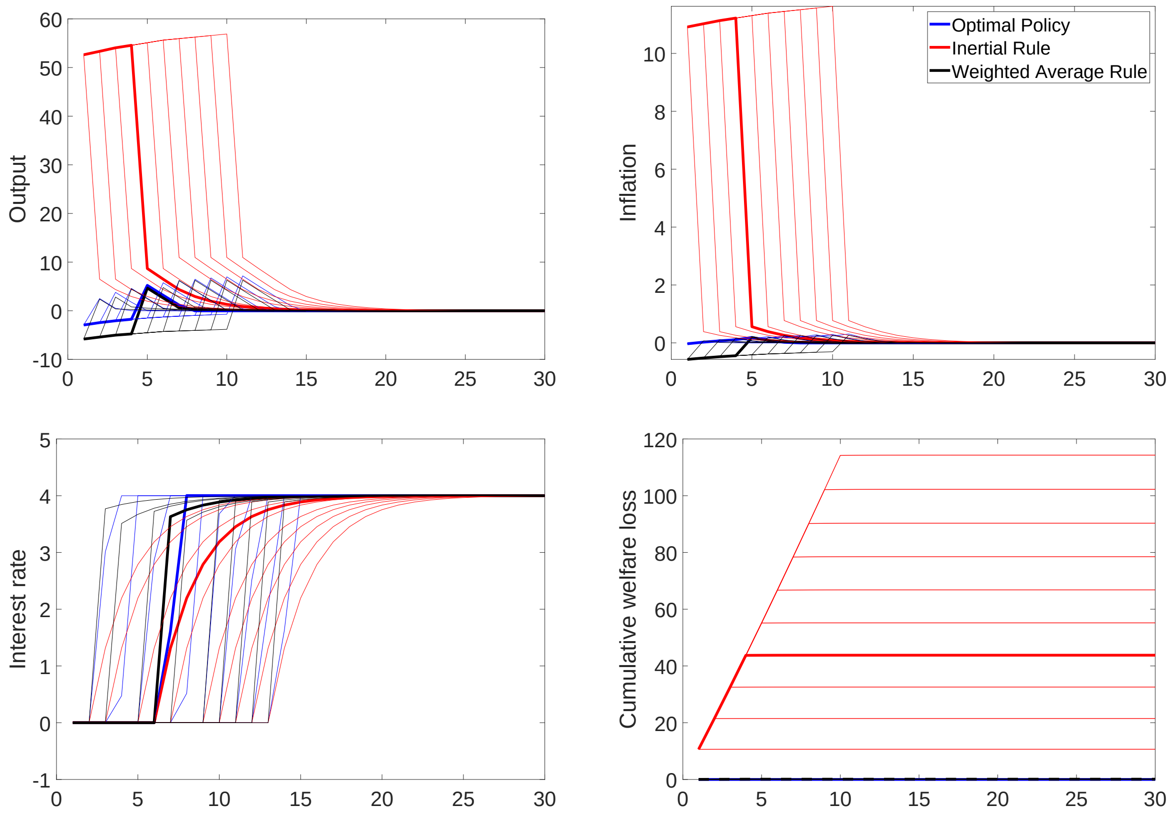
To understand the economics of the different outcomes, compare the highlighted paths of the interest rate for a shock that lasts four quarters in the bottom left panel of Figure 3. Optimal policy calls for a return to the neutral rate two quarters after liftoff. The weighted average rule returns policy to neutral in about eight quarters after liftoff. The inertial rule, however, does not return policy to neutral for more than 24 quarters (six years) after liftoff. Moreover, this policy is pursued despite the fact it is known with certainty that no further shocks will occur. Policymakers are promising to systematically make errors for years because policy is dependent on past interest rate realizations rather than past inflation and output realizations. Policymakers here have explicitly abandoned their dual mandate.

Figure 4 further illustrates the role of that history dependence plays. Here we plot the time zero expected paths for output under the inertial and weighted average rules for different values of ρ_i. We simulate the model for a large number of realizations of the shock and then weight the individual outcomes by the probability with which they occur to summarize the expected outcomes of the different policies. The solid blue line shows the outcomes under optimal policy. The dashed black line that is closest to it corresponds to the weighted average

[6]We follow the solution method described by Eusepi et al. (2022), which is found in their online appendix. The method is based on the original solution algorithm proposed by Eggertsson and Woodford (2003) and described more recently in Eggertsson, Egiev, Lin, Platzer and Riva (2021).

[7]The same conclusions holds between the standard rules and optimal policy if we set $\phi_\pi = \kappa/(\sigma\alpha)$ and α at its welfare theoretic value. The welfare theoretic value of α is $\alpha = \theta/\kappa$ with $\theta = 7.87$.

Figure 3: Standard rules versus optimal rules without the ZLB



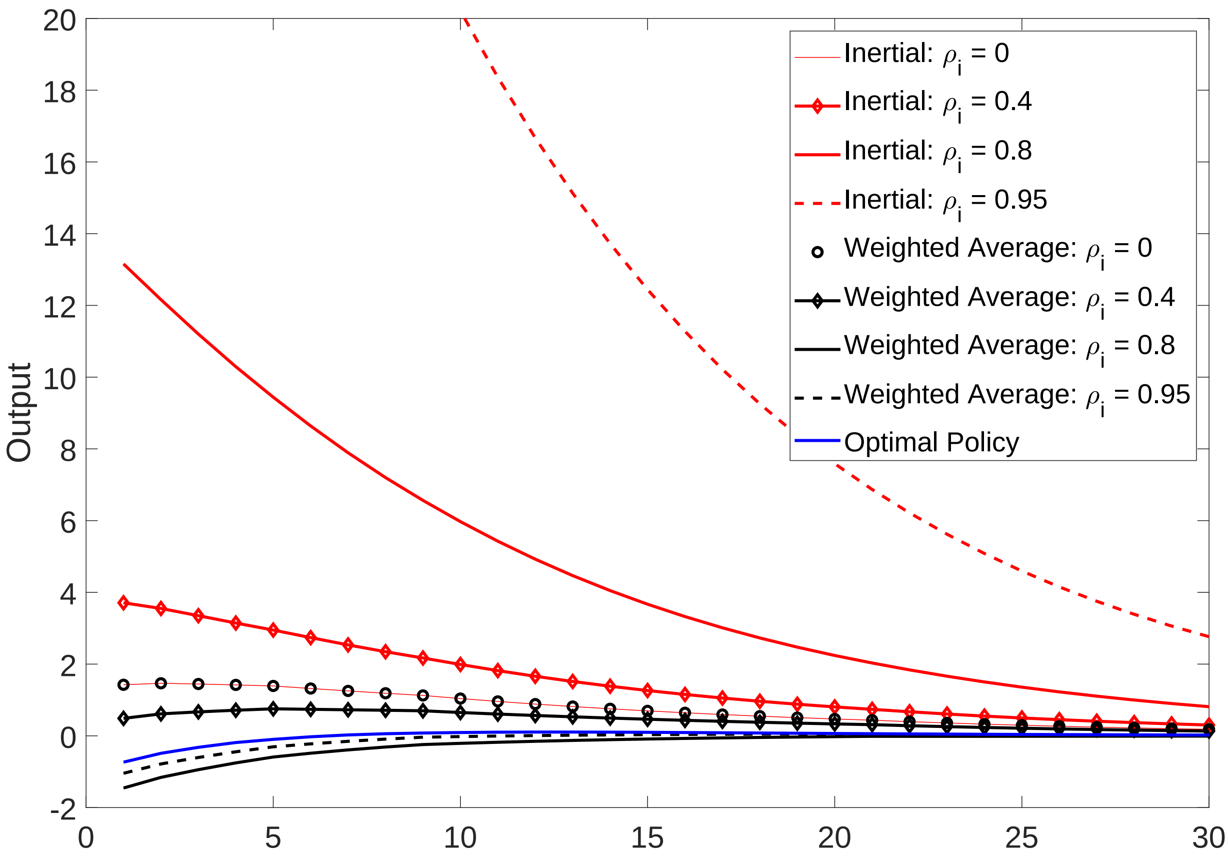
Notes: Each line corresponds to a different realization of the Markov shock process. The outcomes for a shock lasting four quarters are highlighted to make comparisons across specifications easier. Inflation and interest rates are expressed in annual terms.

rule with $\rho_i = 0.95$. As we decrease ρ_i, the approximation to optimal policy deteriorates.

The opposite relationship between ρ_i and approximating optimal policy occurs for the inertial rule with a lag interest rate. Less history dependence - ρ_i closer to zero - generates outcomes closer to optimal policy. When $\rho_i = 0$, the weighted average rule and the inertial rule are the same. Inertial rules encode the wrong history dependence at the ZLB delivering better outcomes with *less* history dependence.

A further implication of Figure 4 is that the stabilization outcomes that are possible under optimal policy are not due to the forward guidance puzzle. Gibbs and McClung (2023) provide a sufficient condition to rule out the forward guidance puzzle in an equilibrium. It is straightforward to numerically verify that optimal policy satisfies it under standard calibrations. In fact, the power of forward guidance to stabilize the economy under optimal policy here does not rely on any NK puzzle. Why this is true is made clear in the next section.

Figure 4: Approximating optimal policy as a function of history dependence



Notes: Each line corresponds to the time zero expectation of the path of output. We simulate the model for many realizations of the shock and then weight the individual outcomes by the probability with which they occur.

3 Resolving New Keynesian Puzzle

We now turn to quantifying how much of the *right* history dependence is required to eliminate the NK puzzles completely. We do so by generalizing the definitions of the NK puzzle put forward by Diba and Loisel (2021) and analyzing the consequence of following a weighted average rule with different values for ρ_i.

To accommodate Diba and Loisel (2021) definitions, we modify the NK model defined by equations (5) and (6) in several ways. First, we remove the cost push shock from the model and replace it with a supply shock that represents variation in marginal cost from changes to labor supply (a_t). This shock allows us to explore the paradox of toil. Second, we add a government spending shock that may affect both supply and demand in the economy (g_t). This shock allows us to explore the fiscal multiplier puzzle. Finally, we assume that all shocks are i.i.d. without exogenous persistence. The NK model is now

$$y_t = E_t y_{t+1} - \sigma^{-1}\left(i_t - E_t\pi_{t+1} - r_t^n\right) + g_t - E_t g_{t+1} \tag{11}$$

$$\pi_t = \beta E_t\pi_{t+1} + \kappa\left(y_t - \delta_g g_t - a_t\right). \tag{12}$$

The NK puzzles are defined by way of a thought of experiment. Contemplate the effect of an anticipated shock that is known at time $t = T$ but occurs in time $t = T^* > T$ under the following monetary policy regime:

$$i_t = \begin{cases} \bar{i} + \phi\pi_t & \text{for } t = T, T+1, ..., T^* \\ \bar{i} + \phi^*\pi_t & \text{for } t > T^*, \end{cases} \tag{13}$$

where $0 \leq \phi < 1$ and $\phi^* > 1$. Then, ask what happens to the effect of the shock in time T if the same shock occurs at an even later date, i.e., as $\Delta_p = T^* - T > 0$ increases.

We can derive a closed-form solution for the equilibrium effects of shocks by using the Phillips curve to eliminate y_t in the IS curve. When $t < T^*$, the model may be expressed as

$$\left(\beta L^{-2} - (\beta + 1 + \frac{\kappa}{\sigma})L^{-1} + \left(1 + \frac{\kappa\phi}{\sigma}\right)\right)\pi_t = X_t$$

$$X_t \equiv -\frac{\kappa}{\sigma}(\bar{i} - r_t^n) - \kappa(a_t - E_t a_{t+1}) + \kappa(1 - \delta_g)\left(g_t - E_t g_{t+1}\right), \tag{14}$$

where L is the lag operator. Factoring the lag polynomial we have

$$(L^{-1} - \lambda_1)(L^{-1} - \lambda_2)\pi_t = X_t$$

where the eigenvalues for the economically relevant parameters satisfy $0 \leq \lambda_1 < 1 < \lambda_2$. Using the method of partial fractions, we can write this as

$$\pi_t \;=\; \frac{1}{\lambda_2 - \lambda_1} E_t \left[\frac{\lambda_1^{-1}}{1 - (\lambda_1 L)^{-1}} - \frac{\lambda_2^{-1}}{1 - (\lambda_2 L)^{-1}} \right] X_t.$$

Finally, under the assumed monetary policy and shocks processes it follows that $E_t \pi_{T+T^*+j} = 0$ for all $j > 0$, which provide the necessary limit conditions to construct an unique rational expectation equilibrium by solving the model forward in time:

$$\begin{aligned}
\pi_t &= \frac{1}{\lambda_2 - \lambda_1} E_t \left[\sum_{T=t}^{T^*} \left(\left(\frac{1}{\lambda_1} \right)^{T-t+1} - \left(\frac{1}{\lambda_2} \right)^{T-t+1} \right) X_T \right] \\
y_t &= \frac{1}{\kappa(\lambda_2 - \lambda_1)} E_t \left[\sum_{T=t+1}^{T^*} \left((1 - \beta\lambda_1) \left(\frac{1}{\lambda_1} \right)^{T-t+1} - (1 - \beta\lambda_2) \left(\frac{1}{\lambda_2} \right)^{T-t+1} \right) X_T \right] \\
&+ \frac{\lambda_1^{-1} - \lambda_2^{-1}}{\kappa\left(\lambda_2 - \lambda_1\right)} X_t + \delta_g g_t + a_t
\end{aligned}$$

Diba and Loisel (2021) define The NK puzzles by studying the properties of dynamic multipliers for a shock as $\Delta_p \to \infty$,

$$\lim_{\Delta_p \to +\infty} \partial z_t / \partial X_{m,t+\Delta_p} = \infty \text{ where } z \in \{\pi, x\} \text{ and } X_m \in \{i^*, g, a\}$$

Definition 1 (forward guidance puzzle) *When the policy rate is expected to be set passively during the next $\Delta_p > 0$ periods, the response of current inflation and output to an expected policy-rate shock $i^* \neq \bar{i}$, Δ_p periods ahead, goes to infinity with Δ_p i.e.*

$$\lim_{\Delta_p \to +\infty} \partial z_t / \partial i_{t+\Delta_p} = \infty \textit{ where } z \in \{\pi, x\}$$

.

Definition 2 (fiscal multiplier puzzle) *When the policy rate is expected to be set passively during the next $\Delta_p > 0$ periods, the response of current inflation and output to an expected expansionary government spending shock, $g_{t+\Delta_p} > 0$, Δ_p periods ahead, goes to positive infinity with Δ_p, i.e.,*

$$\lim_{\Delta_p \to +\infty} \partial z_t / \partial g_{t+\Delta_p} = \infty \textit{ where } z \in \{\pi, x\}$$

.

Definition 3 (paradox of toil) *When the policy rate is expected to be set passively during the next $\Delta_p > 0$ periods, the response of current output to a positive supply shock, $a_{t+\Delta_p} > 0$, Δ_p periods ahead, is weakly contractionary with Δ_p, i.e.,*

$$\lim_{\Delta_p \to +\infty} \partial x_t / \partial a_{t+\Delta_p} \leq 0$$

.

Definition 4 (paradox of flexibility) *When the policy rate is expected to be set passively during the next $\Delta_p > 0$ periods, the response of current inflation and output to an expected shock Δ_p periods ahead goes to positive or negative infinity as κ goes to infinity, i.e.,*

$$\lim_{\kappa \to +\infty} \partial z_t / \partial v_{t+\Delta_p} = \pm\infty \textit{ where } z \in \{\pi, x\} \textit{ and } v = \{i^*, g, a\}$$

.

Three of the four puzzles are direct consequence of the properties of λ_1. The forward guidance puzzle and the fiscal multiplier puzzles are caused by the fact that $\lambda_1 < 1$, which makes $\lambda_1^{-\Delta_p}$ grow without bound as Δ_p increases. The paradox of flexibility occurs because $\lambda_1 \to 0$ as $\kappa \to \infty$. The remaining puzzle occurs because the properties of X_t with a negative sign always appearing in front of the anticipated supply shocks in equilibrium.

3.1 The Puzzles under weighted average policy rule

Consider the NK puzzle thought experiment under the following monetary policy

$$i_t = \begin{cases} \bar{i} + \phi\pi_t & \text{for } t = T, T+1, ..., T^* \\ \bar{i} + \phi^*\omega_t & \text{for } t > T^*, \end{cases} \quad (15)$$

$$\omega_t^\pi = \begin{cases} \rho\omega_{t-1} + \pi_t & \text{for } t = T, T+1, ..., T^* \\ \rho^*\omega_{t-1} + \pi_t & \text{for } t > T^*. \end{cases} \quad (16)$$

The idea is that agents understand that past inflation outcomes matter for policy in the future even if those past outcomes do not currently matter for the setting of the policy rate. The proposed policy nests the ZLB experiment when $\bar{i} = -\bar{r}$ and $\phi = 0$. The parameter ρ captures the degree policy responds to the outcomes in the first regime when in the second regime. It can approximate more-than-make-up policy required by optimal commitment at the ZLB by setting $\rho > 1$.

The question of interest is: how large must ρ be to remove the puzzle? For transparency and tractability, we answer this question with two special cases, which generalize.

3.2 Case 1: $\beta = 0$

Let $\beta = 0$ and assume the economy is in steady state at time $t = T - 1$ such that $\pi_{T-1} = \omega_{T-1} = 0$. In time $t = T$, a single shock is anticipated to occur in period $T^* > T$. When $t \geq T^* + 1$, the economy evolves as

$$\begin{aligned} \pi_t &= \frac{(\kappa+\sigma)}{\sigma}\pi_{t+1} - \frac{\kappa\phi^*}{\sigma}\omega_t + X_t \\ \omega_t &= \rho^*\omega_{t-1} + \pi_t \end{aligned}$$

where X_t is defined the same as in equation (14). We call this the terminal regime. When $T < t \leq T^*$, the economy evolves according

$$\begin{aligned} \psi\pi_t &= \pi_{t+1} + \frac{1}{1+\kappa\sigma^{-1}}X_t \\ \omega_t &= \rho\omega_{t-1} + \pi_t \end{aligned}$$

where $\psi = \frac{1+\kappa\sigma^{-1}\phi}{1+\kappa\sigma^{-1}} \in [0,1)$ and ω_t is decoupled from inflation except to record it history during the passive policy period.

In the terminal regime, the minimum state variable solution for inflation is given by

$$\pi_t = -\Omega^* \omega_{t-1} + C^* X_t$$

where $\Omega^* > 0$ and C^* are functions of ϕ^* and ρ^* and where for convenience we factor out a negative one from Ω^*.[8] The perfect foresight expectation of inflation for time T^*+1 formed in time T is given by

$$E_T \pi_{T^*+1} = -E_T \Omega^* \omega_{T^*} = -E_T \Omega^* \sum_{j=0}^{\Delta_p} \rho^j \pi_{T^*-j},$$

where the last equality holds because in time T^* we are in the passive regime. Using this expectation, we solve for the perfect foresight path of inflation given the shock and policy. Starting with time $t = T^*$,

$$\psi \pi_{T^*} = -\Omega^* \sum_{j=0}^{\Delta_p} \rho^j \pi_{T^*-j} + \frac{1}{1+\kappa\sigma^{-1}} X_{T^*}.$$

Solving for π_{T^*}, we have

$$\pi_{T^*} = -(\psi + \Omega^*)^{-1} \left(\Omega^* \sum_{j=1}^{\Delta_p} \rho^j \pi_{T^*-j} - \frac{1}{1+\kappa\sigma^{-1}} X_{T^*} \right).$$

Repeating the process for π_{T^*-1}, we have

$$\psi \pi_{T^*-1} = -(\psi + \Omega^*)^{-1} \left(\Omega^* \sum_{j=1}^{\Delta_p} \rho^j \pi_{T^*-j} - \frac{1}{1+\kappa\sigma^{-1}} X_{T^*} \right) + \frac{1}{1+\kappa\sigma^{-1}} X_{T^*-1}.$$

Solving for π_{T^*-1}, we have

$$\pi_{T^*-1} = -(\psi(\psi + \Omega^*) + \Omega^* \rho)^{-1} \left(\Omega^* \sum_{j=2}^{\Delta_p} \rho^j \pi_{T^*-j} - \frac{X_{T^*} + (\psi + \Omega^*) X_{T^*-1}}{1+\kappa\sigma^{-1}} \right),$$

where X_t shows up a second time because of the expectations of the shock within the definition

[8] We provide the closed-form solution for Ω^* and C^* in the appendix.

of X_t. Working backwards in this way until time T yields

$$\pi_T = \Phi(\Delta_p)^{-1}\frac{X_{T^*} + (\psi + \Omega^*)X_{T^*-1}}{1 + \kappa\sigma^{-1}} \tag{17}$$

where

$$\Phi(K) = \psi^{K+1} + \Omega^*(\psi^K + \rho\psi^{K-1} + \rho^2\psi^{K-2} + \ldots + \rho^{K-1}\psi + \rho^K). \tag{18}$$

The dynamic multiplier of interest for any shock X_m is therefore

$$\frac{\partial \pi_T}{\partial X_{m,t+\Delta_p}} = \frac{C_m}{1 + \kappa\sigma^{-1}}\Phi(\Delta_p)^{-1},$$

where C_m is the appropriate constant from the mth shock.

Proposition 2: *The NK model (11), (12), (15), and (16) with* $\beta = 0$, $\phi^* > 1$, $0 \leq \phi < 1$, *and* $0 < \rho^* < 1$ *has the following properties*

1. *The equilibrium does not exhibit the forward guidance puzzle, fiscal multiplier puzzle, or paradox of flexibility if* $\rho > 1$.

2. *The equilibrium exhibits the forward guidance and fiscal multiplier puzzles for* $0 \leq \rho < 1$, *and* $\rho^* \neq \bar{\rho}$ *where* $\bar{\rho} \in [0, 1)$ *is defined in the appendix. If* $\rho^* = \bar{\rho}$ *then the equilibrium exhibits the forward guidance puzzle but does not exhibit the fiscal multiplier puzzle.*

3. *The magnitude of the inflation/output response to anticipated monetary or fiscal policy shocks is decreasing in* ρ.

4. *The equilibrium exhibits the paradox of toil if and if only* ρ^* *is sufficiently small such that* $\rho^* \leq \bar{\rho}$ *where* $\bar{\rho} \in [0, 1)$.

5. *The equilibrium exhibits the paradox of flexibility if and only if* $\rho = 0$.

The limit of $\Phi(K)$ as $K \to \infty$ is either 0 or ∞ depending on the values of ρ. If $0 \leq \rho < 1$, then $\Phi(K)$ goes to zero. If $\rho > 1$, then the limit is infinity. And if $\rho = 1$, then the limit is $(1-\psi)^{-1}\Omega^*$. Therefore, the forward guidance and fiscal multiplier puzzles cannot be completely eliminated by a weighted average rule. It requires make-up policy as prescribed by optimal commitment policy. Importantly, though, the size of the multiplier is decreasing in ρ, which is why the forward guidance puzzle is so diminished in our quantitative example in the Introduction when ρ was large but less than one.

The paradox of toil result depends on the terminal solution Ω^*. The current impact inherits the anticipated effect of Ω^* from when the passive regime ends. Therefore, it implies the *correct* expected sign for the shock. Because monetary policy is history dependent, the correct sign of the shock propagates backwards. If ρ^* is large enough, the sign of the multiplier on the supply shock flips to the correct sign. Interestingly, the same threshold for ρ^* matters for the qualitative effects of the fiscal shock. If $\rho^* < \bar{\rho}$ ($\rho^* > \bar{\rho}$) then the anticipated monetary policy shock raises (lowers) output. However, if $\rho = \bar{\rho}$ exactly then the government spending shock has no effect on current output for any Δ_p. Hence, the fiscal multiplier puzzle is absent in this special case.

Finally, for the paradox of flexibility, history dependent policy changes the relationship of price flexibility with economic outcomes in two ways. First, the dynamic multipliers have a κ in the denominator. As price flexibility increases, its effect on the shocks is balanced when κ also appears in the numerator through X_t. Second, price flexibility's effect through the terminal condition is finite. Taking the limit of $-\Omega^*$ as $\kappa \to \infty$ yields ρ^*.

3.3 Case 2: $0 < \beta < 1$

We now return to the full model with $0 < \beta < 1$. The solution method used previously does not yield useful closed form solutions here. We, therefore, consider a special case that is inspired both by Cochrane (2017) and Gibbs and McClung (2023). We assume that $0 \leq \phi < 1$ but $\phi^* = 0$, while all else remains the same. We then assume that agents coordinate their expectations around the history that the central bank is tracking, ω_t. This is akin to a sunspot equilibrium and is illustrative of what history dependence does in a determinate equilibrium when the central bank does credibly respond to the history.

We can write the model in terms of just inflation

$$\left(\beta L^{-2} - (\beta + 1 + \frac{\kappa}{\sigma})L^{-1} + \left(1 + \frac{\kappa\phi}{\sigma}\right)\right)\pi_t = X_t. \tag{19}$$

Then, using the equation for ω_t, we express

$$\pi_t = (1 - \rho L)\omega_t.$$

We can use the above to eliminate π_t from (19) to arrive at representation of the model in terms of ω_t, the lone endogenous state variable:

$$\left(\beta L^{-2} - (\beta + 1 + \frac{\kappa}{\sigma})L^{-1} + 1\right)(1 - \rho L)\omega_t = X_t.$$

Factoring the lag polynomial, we can write the model as

$$(L^{-1} - \lambda_1)(L^{-1} - \lambda_2)(1 - \rho L)\omega_t^{\pi} = X_t. \tag{20}$$

The expression above has three roots: $0 < \lambda_1 < 1 < \lambda_2$ and ρ. The last root is a free parameter. If we set it above one, then we have two roots outside the unit circle and one root inside the unit circle, which satisfies the Blanchard-Kahn conditions without any further assumption required to construct a unique rational expectations equilibrium.[9]

The effect in time $t = T$ of a one-time anticipated shock (i^*, a^*, or g^*) that is known in T and occurring in time T^* is

$$\begin{aligned}
\pi_T &= \frac{1}{\rho - \lambda_2}\sum_{k=1}^{2}\left(\lambda_2^{-\Delta_p - k} - \rho^{-\Delta_p - k}\right)X_{T^*+1-k} \\
y_T &= \frac{1-\beta}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right)\frac{X_{T^*}}{\kappa} + \frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 2} - \rho^{-\Delta_p - 2}\right)\frac{X_{T^*-1}}{\kappa}
\end{aligned}$$

Proposition 3: *The NK model (11), (12), (15), and (16) with $\rho = \rho^* \geq 0$, $0 \leq \phi < 1$ and $\phi^* = 0$ has the following properties*

1. *The equilibrium is puzzle free if $\rho > 1$.*
2. *The magnitude of the inflation/output response to anticipated monetary or fiscal policy shocks is decreasing as ρ increases.*

[9] This equilibrium is similar to those constructed using the approach of Bianchi and Nicolò (2021).

3. *The equilibrium exhibits forward guidance puzzle, the fiscal multiplier puzzle, the paradox of toil, and the paradox of flexibility if* $\rho = 0$.

Because we are solving forward with two unstable roots, the impact of the policies is decreasing with Δ_p. The promise to more-than-make-up for past misses provides the dampening to expectations that keeps expectations in check.

It is the lack of history dependence when policy is set passively or pegged that generates the puzzles. Once history dependence is introduced, the puzzles begin to weaken. When history dependence is large enough, the puzzles are gone.

4 Quantitative Implications

We return to the model of Smets and Wouters (2007) to investigate the quantitative implications of our finding. We quantify how differently anticipated shocks transmit under an interest rate peg when the policy rule is written in weighted average form. Figures 1 and 2 in the Introduction previewed these results. Shocks transmit differently under the two formulations of policy.

We take the posterior distribution from the original Smets and Wouters (2007) paper as given. We do not re-estimate the model. We want to hold everything fixed apart from the form of the policy rule. We convert the monetary policy rule in the model given by

$$i_t = \rho_i i_{t-1} + (1-\rho_i)\left(\phi_\pi \pi_t + \phi_x x_t\right) + \phi_{dx}(x_t - x_{t-1}) + \epsilon_{r,t} \tag{21}$$

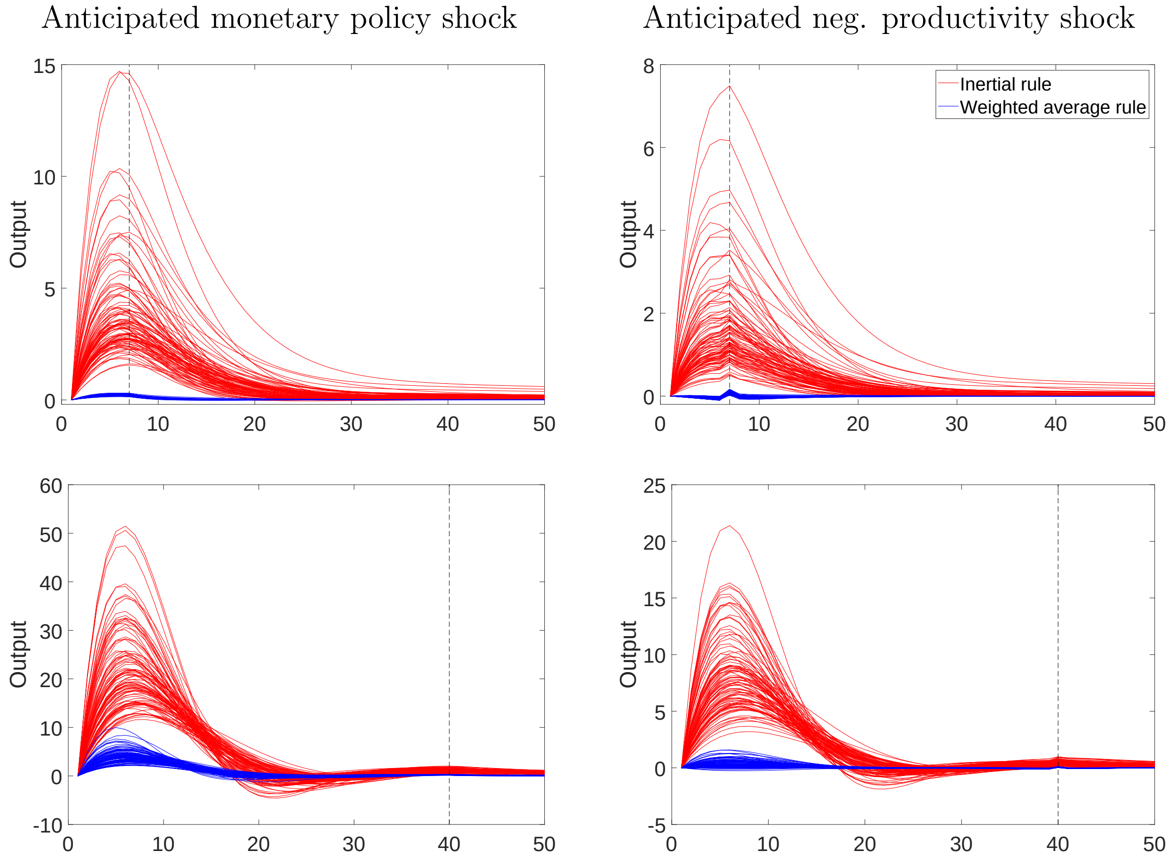
to a weighted average form given by

$$\begin{aligned} i_t &= \phi_\pi \omega_t^\pi + \phi_x \omega_t^x + \phi_{dx}\omega_t^{dx} + \omega_t^{\epsilon_r} \qquad (22)\\ \omega_t^\pi &= \omega_{t-1}^\pi + (1-\rho_i)(\pi_t - \omega_{t-1}^\pi)\\ \omega_t^x &= \omega_{t-1}^x + (1-\rho_i)(x_t - \omega_{t-1}^x)\\ \omega_t^{dx} &= \omega_{t-1}^{dx} + \rho_i(x_t - x_{t-1})\\ \omega_t^{\epsilon_r} &= \omega_{t-1}^{\epsilon_r} + \rho_i \epsilon_{r,t} \end{aligned}$$

where x_t is the output gap. These are the same rule used to create Figures 1 and 2.

Figure 5 plots the impulse response functions for two shocks using draws from the model's

Figure 5: An inertial rule vs a weighted average rule in the Smets and Wouters model



Notes: The black dashed line indicates when the shock occurs. Each solid line corresponds to the impulse response implied by one draw from the posterior distribution of parameters for a one-time 100-basis point monetary policy shock anticipated to occur seven or forty quarters in the future under an interest rate peg (left) or a one-time one standard deviation negative productivity shock anticipated to occur seven or forty quarters in the future under an interest rate peg (right). In period 8 (41) policy returns to either the inertial rule or the weighted average rule. Neither shock has persistence.

parameter posterior distribution. The first shock is an anticipated one-off 100-basis point monetary policy shock that occurs either seven (top) or forty (bottom) quarters in the future under an interest rate peg. The second shock is an anticipated one-off one standard deviation productivity shock that occurs either seven (top) or forty (bottom) quarters in the future under an interest rate peg. After the shocks occur, the policy rule becomes active again and it is parameterized according to the posterior draw. We show only the responses of output here for convenience. The other endogenous variables behave similarly to the shocks.

The inertial rule responses shown in red display the forward guidance puzzle, the paradox of toil, and the paradox of flexibility. Evidence of the latter puzzle is the significant variance in the magnitude of shock responses to different draws from the posterior. Much of this variation is explained by changes in wage and price stickiness with different posterior draws.

The weighted average rule responses shown in blue behave very differently compared to the inertial rule responses. The history dependence in the weighted average rule is insufficient to eliminate the puzzles since the posterior estimate of $\rho_i = 0.81$ with and 95% HPD interval of 0.77 to 0.85, consistent with Propositions 1 and 2 in the simple model. This is visible by comparing the top and bottom rows of plots in Figure 5. The response of output to the anticipated shocks increases in both cases when the shock is expected to occur farther into the future. However, the history dependence significantly dampens the effects of the puzzles such that they are much less relevant for policy at business cycle frequencies.

5 Conclusion

New Keynesian models of monetary policy are known to generate puzzling predictions for output and inflation in response to anticipated shocks when nominal interest rates are pegged or constrained by the zero lower bound. We show that all New Keynesian puzzles and paradoxes are a result of assuming that monetary policy has either the wrong or no history dependence. Adding an empirically plausible amount of history dependence to monetary policy over inflation and output during ZLB episodes significantly mitigate the New Keynesian puzzles or can eliminate them entirely.

Appendix For Online Publication

A1 Weighted average representations

We use lag operators to convert rules with interest rate smoothing to rule that are weighted averages of past output and inflation, where $LX_t = X_{t-1}$ and $L^n X_t = X_{t-n}$ for $n = ..., -2, -1, 0, 1, 2, ...$ Following Sargent (2009), we consider polynomials in the lag operator

$$A(L) = a_0 + a_1 L + a_2 L^2 + ... = \sum_{j=0}^{\infty} a_j L^j$$

where the a_j's are constants and

$$A(L)X_t = (a_0 + a_1 L + a_2 L^2 + ...)X_t = \sum_{j=0}^{\infty} a_j X_{t-j}.$$

We exploit the following relationship

$$A(L) = \frac{1}{1 - \lambda L} = (1 + \lambda L + \lambda^2 L^2 + ...) = \sum_{j=0}^{\infty} \lambda^j,$$

where if $|\lambda| < 1$

$$\sum_{j=0}^{\infty} \lambda^j = \frac{1}{1 - \lambda}.$$

Applying these operations to a Taylor-type rule with interest rate smoothing we can derive equation (**??**):

$$\begin{aligned}
i_t - \rho_i i_{t-1} &= (1-\rho_i)\bar{r} + (1-\rho_i)(\phi_\pi \pi_t + \phi_y y_t) \\
(1-\rho_i L) i_t &= (1-\rho_i)\bar{r} + (1-\rho_i)(\phi_\pi \pi_t + \phi_y y_t) \\
i_t &= \frac{1}{1-\rho_i L}\left((1-\rho_i)\bar{r} + (1-\rho_i)(\phi_\pi \pi_t + \phi_y y_t)\right) \\
&= (1-\rho_i)\sum_{j=0}^{\infty} \rho_i^j \bar{r} + (1-\rho_i)\left(\sum_{j=0}^{\infty} \rho_i^j \left(\phi_\pi \pi_{t-j} + \phi_y y_{t-j}\right)\right) \\
&= (1-\rho_i)\frac{1}{1-\rho_i}\bar{r} + (1-\rho_i)\left(\sum_{j=0}^{\infty} \rho_i^j \left(\phi_\pi \pi_{t-j} + \phi_y y_{t-j}\right)\right) \\
&= \bar{r} + (1-\rho_i)\sum_{j=0}^{\infty} \rho_i^j \left(\phi_\pi \pi_{t-j} + \phi_y y_{t-j}\right).
\end{aligned}$$

We arrive at the representation of rule with auxiliary variables ω_t^π and ω_t^y by way of the following calculations:

$$\begin{aligned}
\omega_t^z &= \omega_{t-1}^z + (1-\rho_i)(z_t - \omega_{t-1}^z) \\
\omega_t^z &= \rho_i \omega_{t-1}^z + (1-\rho_i) z_t \\
\omega_t^z - \rho_i \omega_{t-1}^z &= (1-\rho_i) z_t \\
(1-\rho_i L)\omega_t^z &= (1-\rho_i) z_t \\
\omega_t^z &= \frac{(1-\rho_i)}{1-\rho_i L} z_t = (1-\rho_i)\sum_{j=0}^{\infty} \rho_i^j z_{t-j}
\end{aligned} \tag{A1}$$

Applying this representation for $z = y, \pi$ yields equation (3).

A2 Proofs of propositions

Proposition 1: The unconditional optimal targeting criteria is

$$y_t = -\frac{\kappa}{\alpha}\frac{\pi_t}{1-\beta L}.$$

Using the IS equation, we can write this as

$$E_t y_{t+1} - \frac{1}{\sigma}(i_t - E_t \pi_{t+1} - r_t^n) = -\frac{\kappa}{\alpha}\frac{\pi_t}{1-\beta L}.$$

Solving for i_t we have

$$i_t = \frac{\sigma\kappa}{\alpha}\frac{\pi_t}{1-\beta L} + \sigma E_t y_{t+1} + E_t \pi_{t+1} + r_t^n.$$

The first representation of the optimal rule comes from using (A1) where

$$\begin{aligned} (1-\beta)\frac{\pi_t}{1-\beta L} &= \omega_t^\pi \\ \omega_t^\pi &= \omega_{t-1}^\pi + (1-\beta)(\pi_t - \omega_{t-1}^\pi). \end{aligned}$$

The second representation of the optimal rule comes from multiplying both sides of the equation by $(1-\beta L)$:

$$\begin{aligned} (1-\beta L)i_t &= (1-\beta L)\left(\frac{\sigma\kappa}{\alpha}\frac{\pi_t}{1-\beta L} + \sigma E_t y_{t+1} + E_t \pi_{t+1} + r_t^n\right) \\ i_t - \beta i_{t-1} &= \frac{\sigma\kappa}{\alpha}\pi_t + (1-\beta L)\left(\sigma E_t y_{t+1} + E_t \pi_{t+1} + r_t^n\right). \end{aligned}$$

Proposition 2: Consider first the calculation of the perfect foresight solution in the terminal regime. We write the reduced form model in matrix form as

$$\begin{pmatrix} 1 & \frac{\kappa\phi^*}{\sigma} \\ -1 & 1 \end{pmatrix}\begin{pmatrix} \pi_t \\ \omega_t^\pi \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \rho^* \end{pmatrix}\begin{pmatrix} \pi_{t-1} \\ \omega_{t-1}^\pi \end{pmatrix} + \begin{pmatrix} \frac{\sigma+\kappa}{\sigma} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} E_t\pi_{t+1} \\ E_t\omega_{t+1}^\pi \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} X_t$$

Rearranging matrices we have

$$\begin{pmatrix} \pi_t \\ \omega_t^\pi \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\kappa\rho^*\phi^*}{\sigma+\kappa\phi^*} \\ 0 & \frac{\rho^*\sigma}{\sigma+\kappa\phi^*} \end{pmatrix}\begin{pmatrix} \pi_{t-1} \\ \omega_{t-1}^\pi \end{pmatrix} + \begin{pmatrix} \frac{\kappa+\sigma}{\sigma+\kappa\phi^*} & 0 \\ \frac{\kappa+\sigma}{\sigma+\kappa\phi^*} & 0 \end{pmatrix}\begin{pmatrix} E_t\pi_{t+1} \\ E_t\omega_{t+1}^\pi \end{pmatrix} + \begin{pmatrix} \frac{\sigma}{\sigma+\kappa\phi^*} \\ \frac{\sigma}{\sigma+\kappa\phi^*} \end{pmatrix} X_t$$

Using the method of undetermined coefficients, the solution is

$$\begin{pmatrix} \pi_t \\ \omega_t^\pi \end{pmatrix} = \begin{pmatrix} 0 & -\frac{\rho^*\sigma-\sigma+\sqrt{(\sigma(\rho^*-1)+\kappa(\rho^*-\phi^*))^2+4\kappa(\kappa+\sigma)\rho^*\phi^*}+\kappa\rho^*-\kappa\phi^*}{2(\kappa+\sigma)} \\ 0 & \frac{\rho^*\sigma+\sigma-\sqrt{(\sigma(\rho^*-1)+\kappa(\rho^*-\phi^*))^2+4\kappa(\kappa+\sigma)\rho^*\phi^*}+\kappa\rho^*+\kappa\phi^*}{2(\kappa+\sigma)} \end{pmatrix} \begin{pmatrix} \pi_{t-1} \\ \omega_{t-1}^\pi \end{pmatrix}$$
$$+ \begin{pmatrix} \frac{2\sigma}{\rho^*\sigma+\sigma+\sqrt{(\sigma(\rho^*-1)+\kappa(\rho^*-\phi^*))^2+4\kappa(\kappa+\sigma)\rho^*\phi^*}+\kappa\rho^*+\kappa\phi^*} \\ \frac{2\sigma}{\rho^*\sigma+\sigma+\sqrt{(\sigma(\rho^*-1)+\kappa(\rho^*-\phi^*))^2+4\kappa(\kappa+\sigma)\rho^*\phi^*}+\kappa\rho^*+\kappa\phi^*} \end{pmatrix} X_t$$

Therefore,

$$\begin{aligned} \Omega^* &:= \frac{\rho^*\sigma-\sigma+\sqrt{\left(\sigma\left(\rho^*-1\right)+\kappa\left(\rho^*-\phi^*\right)\right)^2+4\kappa(\kappa+\sigma)\rho^*\phi^*}+\kappa\rho^*-\kappa\phi^*}{2(\kappa+\sigma)} > 0 \\ C^* &:= \frac{2\sigma}{\rho^*\sigma+\sigma+\sqrt{\left(\sigma\left(\rho^*-1\right)+\kappa\left(\rho^*-\phi^*\right)\right)^2+4\kappa(\kappa+\sigma)\rho^*\phi^*}+\kappa\rho^*+\kappa\phi^*} \end{aligned}$$

The puzzle resolutions follow from the properties of $\Phi(\Delta_p)$. By assumption $0 < \psi < 1$. We can rewrite this expression as

$$\Phi(\Delta_p) = \frac{\Omega^*\left(\rho^{\Delta_p+1}-\psi^{\Delta_p+1}\right)}{\rho-\psi} + \psi^{\Delta_p+1}$$

The following limits are easily obtained

1. when $0 < \rho < 1$ then
$$\lim_{\Delta_p\to\infty} \Phi(\Delta_p) = 0 \tag{A2}$$
2. when $\rho = 1$, then
$$\lim_{\Delta_p\to\infty} \Phi(\Delta_p) = (1-\psi)^{-1}\Omega^* \tag{A3}$$
3. when $\rho > 1$, then
$$\lim_{\Delta_p\to\infty} \Phi(\Delta_p) = +\infty \tag{A4}$$
4. when $\rho \geq 0$ and $\phi > 0$, then
$$\lim_{\kappa\to\infty} \Phi(\Delta_p) = \frac{\left(\sqrt{\left(\rho^*+\phi^*\right)^2}+\rho^*-\phi^*\right)\left(\rho^{\Delta_p+1}-\phi^{\Delta_p+1}\right)}{2(\rho-\phi)} + \phi^{\Delta_p+1} \tag{A5}$$

5. when $\rho = 0$ and $\phi = 0$, then

$$\lim_{\kappa \to \infty} \Phi(\Delta_p) = 0 \tag{A6}$$

Proposition 2 part 3 states that the puzzles are decreasing as ρ is increasing. This requires that $\partial \Phi(\Delta_p)/\partial \rho > 0$.

$$\frac{\partial \Phi(\Delta_p)}{\partial \rho} = \frac{\Omega^* \left(\psi \left(\psi^{\Delta_p} - \rho^{\Delta_p}\right) + \Delta_p(\rho - \psi)\rho^{\Delta_p}\right)}{(\rho - \psi)^2}$$

To see this always positive note that when $\rho = 0$, the above is unambiguously positive.

The inflation and output response to an anticipated shock are given by

$$\begin{aligned} \pi_T &= \Phi(\Delta_p)^{-1} \frac{X_{T^*} + (\psi + \Omega^*) X_{T^*-1}}{1 + \kappa\sigma^{-1}} \\ y_T &= \frac{1}{\kappa} \Phi(\Delta_p)^{-1} \frac{X_{T^*} + (\psi + \Omega^*) X_{T^*-1}}{1 + \kappa\sigma^{-1}}. \end{aligned}$$

Interest rate shock: The dynamic multiplier for an interest rate shock is:

$$\partial y_t / \partial i_{t+\Delta_p} = -\frac{1}{\kappa} \Phi(\Delta_p)^{-1} \frac{\kappa\sigma^{-1}}{1 + \kappa\sigma^{-1}}, \; \partial \pi_t / \partial i_{t+\Delta_p} = -\Phi(\Delta_p)^{-1} \frac{\kappa\sigma^{-1}}{1 + \kappa\sigma^{-1}}.$$

When $0 \leq \rho < 1$, then the limit of $\Phi(\Delta_p)$ is given by (A2) and the forward guidance puzzle is present. When $\rho = 1$, the limit of $\Phi(\Delta_p)$ is given by (A3) and forward guidance puzzle is absent. However, there is a weak anti-horizon effect, where the effect of forward guidance doesn't go to zero even when the anticipated shock is infinitely far in the future. When $\rho > 1$, then the limit of $\Phi(\Delta_p)$ is given by (A4) and the forward guidance puzzle is eliminated. The paradox of flexibility only occurs for this shock when $\rho = 0$.

Government spending shock: The dynamic multiplier for a government spending shock is:

$$\partial y_t / \partial g_{t+\Delta_p} = (1 - \delta_g) \Phi(\Delta_p)^{-1} \frac{(1 - \psi - \Omega^*)}{1 + \kappa\sigma^{-1}}, \; \partial \pi_t / \partial g_{t+\Delta_p} = \kappa(1 - \delta_g) \Phi(\Delta_p)^{-1} \frac{(1 - \psi - \Omega^*)}{1 + \kappa\sigma^{-1}}.$$

When $0 \leq \rho < 1$, then the limit of $\Phi(\Delta_p)$ is given by (A2) and the forward guidance puzzle is present, except in the special case where

$$\rho^* = \bar{\rho} := \frac{(1-\phi)\left(\kappa\phi^* + \kappa(1-\phi) + \sigma\right)}{\left(\phi^* - \phi + 1\right)\left(\kappa + \sigma\right)}.$$

where by assumption $0 \leq \phi < 1$ and $\phi^* > 1$. If $\rho^* = \bar{\rho}$ then $\psi + \Omega^* - 1 = 0$ and therefore inflation and output do not respond to the anticipated fiscal shock. The fiscal forward guidance puzzle is eliminated when $\rho \geq 1$. The paradox of flexibility is only present when $\rho = 0$ and $\phi = 0$.

Supply shock: The dynamic multiplier for a supply shock is:

$$\partial y_t / \partial a_{t+\Delta_p} = \Phi(\Delta_p)^{-1} \frac{(\psi + \Omega^* - 1)}{1 + \kappa\sigma^{-1}}, \quad \partial \pi_t / \partial a_{t+\Delta_p} = \kappa\Phi(\Delta_p)^{-1} \frac{(\psi + \Omega^* - 1)}{1 + \kappa\sigma^{-1}}.$$

This take the same form as the government spending shock. The effect of this shock goes to zero as $\Delta_p \to \infty$ when $\rho > 1$ by (A4).

Resolving the paradox of toil requires

$$\partial y_t / \partial a_{t+\Delta_p} = \Phi(\Delta_p)^{-1} \frac{(\psi + \Omega^* - 1)}{1 + \kappa\sigma^{-1}} \geq 0.$$

This expression is positive so long as

$$\psi + \Omega^* > 1.$$

which holds if and only if $\rho^* > \bar{\rho}$.

Proposition 3: We first derive the solution for inflation by solving forward with λ_2 and ρ.

$$\begin{aligned} (L^{-1} - \lambda_1)(L^{-1} - \lambda_2)(1 - \rho L)\omega_t^{\pi} &= X_t \\ (L^{-1} - \lambda_1)(\rho\omega_{t-1} + \pi_t) &= \frac{1}{(L^{-1} - \lambda_2)(1 - \rho L)} X_t \end{aligned}$$

We split the right hand side using partial fractions

$$\frac{1}{(L^{-1}-\lambda_2)(1-\rho L)} = \frac{A}{(L^{-1}-\lambda_2)} + \frac{B}{(1-\rho L)}$$

where when $L = \lambda_2^{-1}$ we have

$$\begin{aligned} 1 &= A(1-\rho L) + B(L^{-1}-\lambda_2) \\ &= A(1-\rho\lambda_2^{-1}) + B(\lambda_2-\lambda_2) \\ &= A\left(\frac{\lambda_2-\rho}{\lambda_2}\right) \\ &... \\ A &= \frac{\lambda_2}{\lambda_2-\rho} \end{aligned}$$

and when $L = \rho^{-1}$ we have

$$\begin{aligned} 1 &= A(1-\rho L) + B(L^{-1}-\lambda_2) \\ &= A(1-\rho\rho^{-1}) + B(\rho-\lambda_2) \\ &= B(\rho-\lambda_2) \\ &... \\ B &= \frac{1}{\rho-\lambda_2} \end{aligned}$$

Then we can write

$$\begin{aligned} (L^{-1}-\lambda_1)(\rho\omega_{t-1}+\pi_t) &= \frac{1}{\rho-\lambda_2}\left(\frac{-\lambda_2}{(L^{-1}-\lambda_2)} + \frac{1}{(1-\rho L)}\right)X_t \\ &= \frac{1}{\rho-\lambda_2}\left(\frac{-\lambda_2}{-\lambda_2(1-(\lambda_2 L)^{-1})} + \frac{1}{-(\rho L)(1-(\rho L)^{-1})}\right)X_t \\ &= \frac{1}{\rho-\lambda_2}\left(\frac{1}{(1-(\lambda_2 L)^{-1})} - \frac{(\rho L)^{-1}}{(1-(\rho L)^{-1})}\right)X_t. \end{aligned}$$

We assume that for $t < T$ the economy is at steady state such that $\pi_{T-j} = \omega_{T-j} = 0$ for $j = 1, 2, \ldots$. With this assumption we can write the above as

$$\begin{aligned}
\rho\omega_{T-1} + \pi_T - \lambda_1(\rho\omega_{T-2} + \pi_{T-1}) &= \frac{1}{\rho - \lambda_2}\left(\frac{1}{(1-(\lambda_2 L)^{-1})} - \frac{(\rho L)^{-1}}{(1-(\rho L)^{-1})}\right) X_{T-1} \\
\pi_T &= \frac{1}{\rho - \lambda_2}\left(\frac{1}{(1-(\lambda_2 L)^{-1})} - \frac{(\rho L)^{-1}}{(1-(\rho L)^{-1})}\right) X_{T-1}
\end{aligned}$$

Imposing the policy (15) and (16) and assuming that shocks are non-zero only in $t = T^*$ such that $X_j = 0$ except when $t = T^*$ or $t = T^* - 1$

$$\begin{aligned}
\pi_T &= \frac{1}{\rho - \lambda_2}\left(\sum_{j=0}^{T^*+1}\left(\frac{1}{\lambda_2}\right)^j X_{T+j-1} - \frac{1}{\rho}\sum_{j=0}^{T^*}\left(\frac{1}{\rho}\right)^j X_{T+j}\right) \\
&= \frac{1}{\rho - \lambda_2}\sum_{k=1}^{2}\left(\lambda_2^{-\Delta_p - k} - \rho^{-\Delta_p - k}\right) X_{T^*+1-k} \\
&= \frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right)\frac{\kappa}{\sigma}(\bar{r} - i^*) \\
&\quad + \left(\frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right) - \frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 2} - \rho^{-\Delta_p - 2}\right)\right)\kappa(1 - \delta_g)g^* \\
&\quad - \left(\frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right) - \frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 2} - \rho^{-\Delta_p - 2}\right)\right)\kappa a^* \\
&\ldots \\
\pi_T &= \frac{1}{\rho - \lambda_2}\left\{\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right)\frac{\kappa}{\sigma}(\bar{r} - i^*) + \left((\lambda_2 - 1)\lambda_2^{-\Delta_p - 2} + (1 - \rho)\rho^{-\Delta_p - 2}\right)\kappa\left((1 - \delta_g)g^* - a^*\right)\right\}
\end{aligned}$$

To solve for the equilibrium output response, we start by rearranging the Phillips curve

$$\kappa y_{t+1} = \pi_t - \beta\pi_{t+1}.$$

Starting from period T, we have

$$\begin{aligned}
\kappa y_T &= \frac{1}{\rho - \lambda_2}\sum_{k=1}^{2}\left(\lambda_2^{-\Delta_p - k} - \rho^{-\Delta_p - k}\right) X_{T^*+1-k} - \beta\frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - k} - \rho^{-\Delta_p - k}\right) X_{T^*} \\
&= \frac{1 - \beta}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right) X_{T^*} + \frac{1}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 2} - \rho^{-\Delta_p - 2}\right) X_{T^*-1} \\
&\ldots \\
y_T &= \frac{1 - \beta}{\rho - \lambda_2}\left(\lambda_2^{-\Delta_p - 1} - \rho^{-\Delta_p - 1}\right)\frac{1}{\sigma}(\bar{r} - i^*) \\
&\quad + \frac{1}{\rho - \lambda_2}\left\{\left(((1 - \beta)\lambda_2 - 1)\lambda_2^{-\Delta_p - 2} + (1 - (1 - \beta)\rho)\rho^{-\Delta_p - 2}\right)\left((1 - \delta_g)g^* - a^*\right)\right\}
\end{aligned}$$

For part 1 of the proposition, we note that $\lambda_2 > 1$ and $\rho > 1$ by assumption. When we send $\Delta_p \to \infty$, the expressions for π_T and y_T converge to zero. Therefore, the forward guidance puzzle and the fiscal multiplier puzzle are not present. For the paradox of toil, we have

$$y_T = \frac{1}{\rho - \lambda_2}\left\{\left(((1-\beta)\lambda_2 - 1)\lambda_2^{-\Delta_p - 2} + (1-(1-\beta)\rho)\rho^{-\Delta_p-2}\right)(-a^*)\right\}$$

Consider the case when $\rho = 1$, then

$$\frac{\partial y_T}{\partial a^*} = \frac{\left(\lambda_2 - \beta\lambda_2 + \beta\lambda_2^{\Delta_p+2} - 1\right)}{\lambda_2 - 1}\lambda_2^{-\Delta_p-2} > 0$$

because $\lambda_2 > 1$. Consider the case when $\rho \to \lambda_2$ from either above or below

$$\lim_{\rho\to\lambda_2}\frac{\partial y_T}{\partial a^*} = \lambda_2^{-\Delta_p-3}\left(2 + \Delta_p - (1-\beta)\lambda_2(1+\Delta_p)\right)$$

A typical value for $\beta = 0.99$. Typical parameterizations of the model imply λ_2 close to one, which means that second term in the parentheses is small in absolute value relative to the first, which means the multiplier is positive. Finally, taking the limit as $\rho \to 0$ of $\frac{\partial y_T}{\partial a^*}$ is zero. Therefore, we conclude that the paradox of toil is absent for all $\rho \geq 1$.

For the paradox of flexibility, we look at λ_2, which is equal to

$$\lambda_2 = \frac{\sqrt{\left(-\beta - \frac{\kappa}{\sigma} - 1\right)^2 - 4\beta\left(\frac{\kappa\phi}{\sigma} + 1\right)} + \beta + \frac{\kappa}{\sigma} + 1}{2\beta}$$

and

$$\lim_{\kappa\to\infty}\lambda_2 = \infty.$$

In addition,

$$\lim_{\kappa\to\infty}\lambda_2^{-1}\kappa = \frac{2\beta\sigma}{\sigma\sqrt{\frac{1}{\sigma^2} + 1}}.$$

so the paradox of flexibility is absent.

References

Andrade, Philippe, Gaetano Gaballo, Eric Mengus, and Benoit Mojon, "Forward guidance and heterogeneous beliefs," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 1–29.

Angeletos, George-Marios and Chen Lian, "Forward guidance without common knowledge," *American Economic Review*, 2018, *108* (9), 2477–2512.

Bianchi, Francesco and Giovanni Nicolò, "A generalized approach to indeterminacy in linear rational expectations models," *Quantitative Economics*, 2021, *12* (3), 843–868.

Bilbiie, Florin O, "The new Keynesian cross," *Journal of Monetary Economics*, 2020, *114*, 90–108.

Blake, Andrew P, "A'timeless perspective'on optimality in forward-looking rational expectations models," Technical Report, Citeseer 2001.

Blanchard, Olivier and Jordi Galí, "Real wage rigidities and the New Keynesian model," *Journal of money, credit and banking*, 2007, *39*, 35–65.

Bonciani, Dario and Joonseok Oh, "Monetary policy inertia and the paradox of flexibility," *Journal of Economic Dynamics and Control*, 2023, *151*, 104668.

Brassil, Anthony, Callum Ryan, and Ani Yadav, "Bias in Central Bank Interest Rate Rule Comparisons," Technical Report, Reserve Bank of Australia 2023.

Cagliarini, Adam and Mariano Kulish, "Solving Linear Rational Expectations Models with Predictable Structural Changes," *Review of Economics and Statistics*, 2013, *95*, 328–336.

Carlstrom, Charles T, Timothy S Fuerst, and Matthias Paustian, "Inflation and output in New Keynesian models with a transient interest rate peg," *Journal of Monetary Economics*, 2015, *76*, 230–243.

Clarida, Richard, Jordi Gali, and Mark Gertler, "The science of monetary policy: a new Keynesian perspective," *Journal of economic literature*, 1999, *37* (4), 1661–1707.

Cochrane, John H, "The new-Keynesian liquidity trap," *Journal of Monetary Economics*, 2017, *92*, 47–63.

_ , *The fiscal theory of the price level*, Princeton University Press, 2023.

Del Negro, Marco, Marc P Giannoni, and Christina Patterson, "The forward guidance puzzle," *FRB of New York Staff Report*, 2012, (574).

_ , Stefano Eusepi, Marc P Giannoni, Argia M Sbordone, Andrea Tambalotti, Matthew Cocci, Raiden Hasegawa, and M Linder, "The FRBNY DSGE model," *FRB of New York Staff Report*, 2013, (647).

Diba, Behzad and Olivier Loisel, "Pegging the interest rate on bank reserves: a resolution of New Keynesian puzzles and paradoxes," *Journal of Monetary Economics*, 2021, *118*, 230–244.

Eggertsson, Gauti B, "The paradox of toil," *FRB of New York Staff Report*, 2010, (433).

_ and Michael Woodford, "Zero bound on interest rates and optimal monetary policy," *Brookings papers on economic activity*, 2003, *2003* (1), 139–233.

_ and Paul Krugman, "Debt, deleveraging, and the liquidity trap: A Fisher-Minsky-Koo approach," *The Quarterly Journal of Economics*, 2012, *127* (3), 1469–1513.

_ and Sanjay R Singh, "Log-linear Approximation versus an Exact Solution at the ZLB in the New Keynesian Model," *Journal of Economic Dynamics and Control*, 2019, *105*, 21–43.

_ , Neil R Mehrotra, and Jacob A Robbins, "A model of secular stagnation: Theory and quantitative evaluation," *American Economic Journal: Macroeconomics*, 2019, *11* (1), 1–48.

_ , Sergey K Egiev, Alessandro Lin, Josef Platzer, and Luca Riva, "A toolkit for solving models with a lower bound on interest rates of stochastic duration," *Review of Economic Dynamics*, 2021, *41*, 121–173.

Eusepi, Stefano, Christopher G Gibbs, and Bruce Preston, "Monetary policy trade-offs at the zero lower bound," *Available at SSRN 4069078*, 2022.

Evans, George W, Christopher G Gibbs, and Bruce McGough, "A Unified Model of Learning to Forecast," Technical Report 2022.

Farhi, Emmanuel and Iván Werning, "Fiscal multipliers: Liquidity traps and currency unions," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 2417–2492.

__ **and** __ , "Monetary policy, bounded rationality, and incomplete markets," *American Economic Review*, 2019, *109* (11), 3887–3928.

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

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

Gibbs, Christopher G and Nigel McClung, "Does my model predict a forward guidance puzzle?," *Review of Economic Dynamics*, 2023.

Haberis, Alex, Richard Harrison, and Matt Waldron, "Uncertain policy promises," *European Economic Review*, 2019, *111*, 459–474.

Hagedorn, Marcus, Jinfeng Luo, Iourii Manovskii, and Kurt Mitman, "Forward guidance," *Journal of Monetary Economics*, 2019, *102*, 1–23.

Hills, Timothy S and Taisuke Nakata, "Fiscal multipliers at the zero lower bound: the role of policy inertia," *Journal of Money, Credit and Banking*, 2018, *50* (1), 155–172.

Jensen, Christian and Bennett T McCallum, "The non-optimality of proposed monetary policy rules under timeless perspective commitment," *Economics Letters*, 2002, *77* (2), 163–168.

Jones, Callum, "Unanticipated shocks and forward guidance at the zlb," *Manuscript (October)*, 2017.

Kiley, Michael T, "Policy paradoxes in the New Keynesian model," *Review of Economic Dynamics*, 2016, *21*, 1–15.

Krippner, Leo, "Measuring the stance of monetary policy in zero lower bound environments," *Economics Letters*, 2013, *118* (1), 135–138.

__ , "A note of caution on shadow rate estimates," *Journal of Money, Credit and Banking*, 2020, *52* (4), 951–962.

Kulish, Mariano and Adrian Pagan, "Estimation and solution of models with expectations and structural changes," *Journal of Applied Econometrics*, 2017.

__ **, James Morley, and Tim Robinson**, "Estimating DSGE models with zero interest rate policy," *Journal of Monetary Economics*, 2017, *88*, 35–49.

McClung, Nigel, "The power of forward guidance and the fiscal theory of the price level," *International Journal of Central Banking*, 2021, *17*, 79–135.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson, "The power of forward guidance revisited," *American Economic Review*, 2016, *106* (10), 3133–58.

Negro, Marco Del, Marc P Giannoni, and Christina Patterson, "The forward guidance puzzle," *Journal of Political Economy Macroeconomics*, 2023, *1* (1), 43–79.

Sargent, Thomas J, *Dynamic macroeconomic theory*, Harvard University Press, 2009.

Smets, Frank and Rafael Wouters, "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American Economic Review*, 2007, *97* (3), 586–606.

Taylor, John B, "Discretion versus policy rules in practice," in "Carnegie-Rochester conference series on public policy," Vol. 39 Elsevier 1993, pp. 195–214.

Wieland, Johannes F, "Are negative supply shocks expansionary at the zero lower bound?," *Journal of Political Economy*, 2019, *127* (3), 973–1007.

Woodford, Michael, "The Taylor rule and optimal monetary policy," *American Economic Review*, 2001, *91* (2), 232–237.

— , *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2003.

— , "Optimal interest-rate smoothing," *The Review of Economic Studies*, 2003, *70* (4), 861–886.

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.