American Economic Journal: Macroeconomics 2019, 11(4): 310–345
https://doi.org/10.1257/mac.20170335

Optimal Forward Guidance[†]

By Florin O. Bilbiie*

Optimal forward guidance is the simple policy of keeping interest rates low for some optimally determined number of periods after the liquidity trap ends and moving to normal-times optimal policy thereafter. I solve for the optimal duration in closed form in a new Keynesian model and show that it is close to fully optimal Ramsey policy. The simple rule "announce a duration of half of the trap's duration times the disruption" is a good approximation, including in a medium-scale dynamic stochastic general equilibrium (*DSGE*) *model. By anchoring expectations of Delphic agents* (*who mistake commitment for bad news*), *the simple rule is also often welfare-preferable to Odyssean commitment.* (*JEL* D84, E12, E43, E52, E56)

What is left for a central bank to do in a liquidity trap when it is out of standard ammunition? This once exotic theoretical question has been haunting the developed world for the past decades: Japan for the last 22 years and counting, the United States and Europe since the financial crisis of 2008, and the vast majority of OECD countries today. One of the remaining policy options is "forward guidance" (FG for short), the promise to keep interest rates low even after normal conditions have resumed, when increasing rates would be the response warranted by optimal policy (an exhaustive discussion of other policy options at the zero lower bound (ZLB) can be found in Woodford 2012). Several central banks have been doing FG in various forms, starting with the Bank of Japan since 1999 and the US Federal Reserve since 2003, as reviewed in great detail elsewhere.[1]

* Department of Economics, University of Lausanne, Internef, Quartier Chamberonne, CH-1015, Lausanne, Switzerland and the Center for Economic and Policy Research (CEPR) (email: florin.bilbiie@unil.ch). Giorgio Primiceri was coeditor for this article. I am grateful to the three anonymous referees, Marco Bassetto, Marco Del Negro, Martin Ellison, Gaetano Gaballo, Jordi Galí, Alejandro Justiniano, Leonardo Melosi, Taisuke Nakata, Anton Nakov, Sebastian Schmidt, Pedro Teles, Mirko Wiederholt, Michael Woodford, and participants in several seminars and conferences for helpful comments and discussions, and to Ezgi Özsöğüt for superb research assistance. I am particularly grateful to Taisuke Nakata and Paul Yoo for providing Fortran code and helping to produce a figure for Ramsey policy in the stochastic case. I gratefully acknowledge without implicating the support of Banque de France via the eponymous Chair at PSE, and of Institut Universitaire de France. I also thank NYU and CREi-UPF Barcelona for hospitality during part of writing this paper.

† Go to https://doi.org/10.1257/mac.20170335 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] Campbell et al. (2016) reviews much of the literature and provides evidence that Federal Reserve FG was counterproductive before 2011 but had expansionary effects thereafter; the authors associate that with a change in policy strategy, namely to FG becoming "Odyssean,"a term explained and discussed in more detail below. Williams (2013) discusses at least three varieties of FG used by the Fed since 2003, and Filardo and Hofman (2014) provides a comprehensive discussion of the international evidence.

A large and important theoretical literature developed to study optimal interest rate policy in a liquidity trap. Keeping interest rates low after the trap ends is one feature of optimal policy in a variety of environments, an insight by now well known, originally due to the seminal work of Eggertsson and Woodford (2003); Jung, Teranishi, and Watanabe (2005); Adam and Billi (2006, 2007); and Nakov (2008).[2] But the optimal policies found in the (expanding) literature that I review in detail below consist of complicated state-contingent plans (for an illustration, see figures 3 and 4 in Eggertsson and Woodford 2003); this complexity makes optimal policy hard to communicate to the public, thus undermining its credibility. Such considerations have led the same authors to propose price-level targeting as a good approximation to optimal policy. Despite this policy's appeal and its being extensively studied in this context for 15 years, it has never been implemented anywhere, perhaps because it requires changing the law (the central bank status) or at least the policy regime, or perhaps because such a policy would also be hard to communicate, for in engineering language it amounts to explaining proportional-integral-derivative (PID) control.

I should check that

I propose a simpler (restricted notion of) optimal policy under a zero lower bound constraint that I call "optimal forward guidance" (OFG for short). This consists of committing to a simpler policy of keeping the interest rate at zero for a *certain number of periods* after the trap and, when that period expires, switching directly to the optimal policy of "normal times." The policy choice variable then becomes the number of periods of extra accommodation, which I solve for in *closed form* in a simple new Keynesian (NK) model. To the best of my knowledge, no closed-form solution for the optimal duration of FG exists in the literature that I scoured, despite the explosion of studies on the topic in the last two decades. The insight is that the solution thus obtained is very close to the Ramsey-optimal policy (that I also solve for comparison), while being much simpler and thus easier to communicate.

The question, "how long should central banks keep interest rates at zero beyond the end of the trap?" is important for welfare and policy. The closed-form answer to it that this paper provides unveils how OFG duration depends on structural features (deep parameters) of the economy. Furthermore, such closed-form solutions can be used to derive *simple rules* that approximate, for practical and operational purposes, what are often complicated and opaque optimal policies.[3]

The main novel contribution of this paper is thus to reformulate the simpler problem of choosing optimal FG duration, and provide a *closed-form solution* as an explicit function of the duration of the liquidity trap, its severity (the size of the financial disruption), and the slopes of aggregate demand and supply. The solution relies chiefly on a stochastic way to model FG, which delivers a closed-form equilibrium—itself novel, as far as I know—as a function of the

[2] The literature analyzing the consequences of the ZLB in formal NK models has been pioneered by, e.g., Fuhrer and Madigan (1997), Wolman (1998), Orphanides and Wieland (1998), and Krugman (1998) who discusses the virtues of lowering long rates by keeping interest rates low in the future and creating expected inflation; see also the discussion by Rogoff (1998) in the same volume.

[3] In the realm of monetary policy in normal times, a close parallel would consist of the analysis of simple Taylor rules and inflation targeting, in which the feedback between closed-form solutions, theory, and policy practice based on simple rules is obvious.

expected trap duration. The analysis is similar in spirit to Woodford's (2011) analysis of optimal government spending in a liquidity trap (LT) in that it restrains attention to a constrained optimum. My modeling of FG stochastically is related to Woodford's analysis of future spending expansions and their impact on the spending multiplier.

The optimal FG duration strikes the intertemporal balance between two forces. The first is expansionary: future cuts in interest rates are expansionary today, which improves welfare at the ZLB for the usual reasons emphasized by the seminal work of Eggertsson and Woodford (2003). Yet a second force reduces lifetime welfare: when the future becomes the present, low interest rates generate positive consumption gaps (and inflation), which is inefficient. This welfare cost is made of two components. First, in the FG state, (inflation and) consumption volatility are inefficient (as already discussed by Eggertsson and Woodford 2003, 178). Second, longer FG duration also mechanically implies a longer duration for which the welfare cost is incurred. I will concentrate mostly on the first component of the cost, which provides an upper bound on the optimal FG, but also show that adding the second channel leaves the conclusion regarding optimal FG essentially unchanged.

I then propose an operational policy prescription rooted in the welfare analysis: a *simple rule* for FG relating the "instrument" (the duration of low interest rates) to the duration of the liquidity trap (LT) and the magnitude of the financial disruption causing it. The latter is given by the ratio of two average values of the natural rate of interest: the absolute value during the trap (when it is negative) and in normal times. The central bank can announce that *once the trap is over*, it will keep interest rates at zero for a given duration of

$$\text{Half of (LT duration} \times \text{disruption).}$$

To be clear, at the time of the announcement, the central bank *only* commits to the rule itself; the *precise duration* of extra accommodation is only calculated once the trap is *over*, *looking back* and observing the duration and size of the disruption. To give an example, suppose a central bank announces this policy in a liquidity trap triggered by an increase in spreads. Suppose that after four years the liquidity trap subsides: the natural interest rate returns positive to its historical average of, say, 1 percent. The bank measures that on average over the (just past) trap period, the natural rate was minus 1 percent, so that the disruption is equal to 1. The central bank then announces that the trap ended and the extra accommodation period starts, keeps the interest rate at 0 for $(0.5 \times 4 \ =)$ 2 years, and finally moves directly to "business as usual," i.e., to the optimal policy of normal times. Ha!

A policy like the one just described is desirable in terms of credibility and communication, for it allows bypassing the commitment problem inherent to FG in certain information environments such as those emphasized by Andrade et al. (forthcoming), Bassetto (2016), or Wiederholt (2015). The simple rule allows for getting the best of two worlds, combining the benefits of state-contingent and time-dependent (calendar-based) FG as follows. The duration and depth of the trap can (almost tautologically) be measured and verified ex post, once the trap ended. Thus, announcing that "FG starts when the trap ends, and for a duration that is such and such function of the trap duration and depth," is an effective way to communicate what may otherwise be a self-defeating commitment. In particular, if some agents regard such commitment, to follow the distinction introduced by Campbell et al. (2012) and discussed below, in a "*Delphic*" way (inferring something about fundamentals), "*Odyssean*" FG can have perverse effects. Andrade et al. (forthcoming) illustrates this point numerically in the context of the model with heterogeneous beliefs, and I show it here analytically. Yet as Bassetto has shown quite generally, Delphic FG can be very useful in this context to anchor the expectations of Delphic agents. The simple rule that I propose is one such example of Delphic (communication) FG, and is also a good approximation to Odyssean (commitment) FG.

Insofar as the simple NK model used here constitutes the core of more complex NK models, the forces highlighted in my analysis are also likely to be at play in those more complicated versions; yet it is unclear whether other forces, absent from the simple framework used by this paper and the rest of the literature, may or may not overturn the main policy conclusions. To address this concern and as a final exercise, I extend the analysis to study OFG and the simple FG rule in an empirically relevant medium-scale DSGE model that follows Justiniano, Primiceri, and Tambalotti (2013). In a version of that model, I calculate OFG numerically and compare it with the simple FG rule by computing the welfare cost of the latter relative to the former. While the simple rule is not as close to OFG as in the simple three-equation model, it remains a good approximation to this (restricted notion of) optimal policy. More importantly, it can provide large welfare benefits relative to "doing nothing" (sticking to strict inflation targeting)—especially when the adverse shock is large and long-lived, the welfare cost of strict inflation targeting (SIT) is one magnitude larger than that of the simple rule. Conversely, the cost of following the simple rule when it is optimal to do nothing (for small and short-lived shocks) is asymmetrically small.

Related Literature.—A vast literature studied optimal policy under a ZLB constraint, thus providing an implicit, but *not closed-form explicit*, answer to the question that I study too. In what follows, I review briefly the main contributions. The seminal papers were written in the early 2000s, as Japan was entering its liquidity trap that, more than two decades after, still persists. Eggertsson and Woodford (2003) and Jung, Teranishi, and Watanabe (2005) were the first to derive analytical conditions for Ramsey-optimal policy and show that it is optimal to keep nominal interest rates at 0 for longer than implied by the duration of the trap. The former paper (Eggertsson and Woodford 2003) analyzed the case where the shock causing the LT follows a two-state Markov chain, which thereafter became the norm for tractable analysis of stochastic ZLB equilibria. Among other contributions, it discussed the trade-offs between current and future volatility faced by a central bank in a liquidity trap, and studied the virtues of a simple policy of price-level targeting as an approximation to the Ramsey policy. The latter paper (Jung, Teranishi, and Watanabe 2005) analyzed a setup in which the shock has known duration (perfect foresight) and showed analytically that commitment implies a longer duration of zero-interest policy than discretion—see their equation (35).[4] Adam and Billi (2006, 2007) and Nakov (2008) extended this to a fully stochastic setup, where shocks follow autoregressive processes and the ZLB is an occasionally binding constraint. They emphasized the large welfare gains from policy commitment compared to discretion; "forward guidance" is inherently part of the optimal commitment policy in this framework too. Nakov (2008) also studied simpler policies such as price-level targeting and argued that, with large deflationary shocks, simple rules fare much worse than optimal policy. Levin et al. (2010) cautions against the potential welfare losses that can result from FG in a large-shock environment.

More recently, Nakata (2016) and Schmidt (2013) extended Ramsey-optimal policy at the ZLB to the joint analysis of monetary and fiscal policy, under both commitment and discretion; among other results, they find that uncertainty creates more scope for stimulus at the ZLB. Several papers explore the implications of the risk of the ZLB binding again in the future; then, the central bank has an incentive to honor its promises, unlike in a setup where the ZLB is a one-off event. FG is thus naturally "sustainable"—if the duration is not too large (Walsh 2017) and if the shocks are frequent enough (Nakata 2018). In a similar environment, Nakata and Schmidt (forthcoming) identify a deflationary bias outside the ZLB brought about by the anticipation of future ZLB risk; appointing a conservative central bank à la Rogoff (1985) alleviates this, improving welfare. I abstract from all such considerations here, although I believe them to be important for policymaking.

A series of contributions look at forward guidance in environments with information frictions, all of which are relevant for my analysis of the virtues of a simple rule. Andrade et al. (forthcoming) provides empirical evidence that agents believe the central bank's announcements (they agree on the interest rate projections). But agents disagree on the implications of those policies on macroeconomic outcomes. The authors build a model of heterogeneous beliefs, in which a fraction of agents take FG to be "Delphic," thus inferring from the central bank's announcement of low interest rates that circumstances will stay gloomy. The other agents take FG to be "Odyssean," i.e., believing the central bank. I use the same framework to look at optimal FG and review this paper's findings in more detail in due course. Wiederholt (2015) builds a model with dispersed information; a fraction of households update their inflation expectations while the rest do not. The latter group does not respond to FG, but the former does. On the one hand, their expected inflation rises as they believe the FG announcement. On the other, they also infer from today's commitment that the bad state will persist, which reduces inflation expectations. Both features—some agents do not update, and those who do become in part more pessimistic as a consequence of FG—go in the same direction, making the overall effect of FG unambiguously lower than in a perfect-information setup. Bassetto (2016) provides further foundations for the benefits of (Delphic) FG as a communication device, in a strategic game with cheap talk from the central bank. Without private information, Odyssean FG ?

[4] Jung, Teranishi, and Watanabe's (2005) results also hold in continuous time, as shown in an otherwise identical perfect-foresight setup by Werning (2012), which then extends the analysis to optimal fiscal policy at the ZLB.

(meant purely as a commitment device) is redundant in this model. But with private information, communication can improve welfare and (Delphic) FG emerges naturally as such a communication strategy; this further interacts with the Odyssean FG's commitment and credibility dimensions. I draw on these contributions to justify the benefits of a simple rule.[5]

I. Forward Guidance in a Liquidity Trap: A Stochastic Approach

The baseline framework is the standard new Keynesian model, the details of which are skipped since they are readily available, e.g., in the textbooks of Woodford (2003) or Galí (2008). The economy is described by standard aggregate demand and supply equations:

$$c_t = E_t c_{t+1} - \sigma\big(i_t - E_t \pi_{t+1} - \rho_t\big), \tag{1}$$

$$\pi_t = \beta_e E_t \pi_{t+1} + \kappa c_t. \tag{2}$$

The first equation is the "investment-savings (IS) curve," where i_t is the nominal interest rate in levels, ρ_t is an exogenous disturbance that moves the natural interest rate, and σ is the elasticity of intertemporal substitution. The second equation is a standard "new Keynesian Phillips curve" coming from a forward-looking pricing model à la the Calvo-Yun or Rotemberg model with slope $\kappa = \psi\big(\varphi + \sigma^{-1}\big)$, where $\psi = (1 - \zeta)(1 - \beta_e \zeta)/\zeta$ with ζ being the probability to be unable to change one's price in the Calvo-Yun model, φ is the inverse constant-consumption elasticity of labor supply, and β_e is the discount factor of price setters that may be different than the households'.

I first derive optimal FG under an additional simplifying assumption, and then show that optimal FG remains essentially unchanged once I relax it. Namely, consider a simpler aggregate supply curve, the contemporaneous Phillips curve:

$$\pi_t = \kappa c_t. \tag{3}$$

This is microfounded in Bilbiie (2018) based on a Rotemberg pricing scheme whereby firms' reference prices equal yesterday's aggregate price index.[6] In a Calvo-Yun framework, (3) instead arises if in each period a fraction of firms re-optimizes its price freely, while the remaining fraction keeps its price fixed; thus, price setters do not take into account the forward-looking nature of their decision, in that they do not recognize that today's reset price prevails with some probability in future periods. The latter setup amounts to assuming $\beta_e = 0$ in the pricing

[5] My recommendation for a simple rule is related to the analysis of optimal monetary policy delegation, Rogoff (1985) and Walsh (1995) being the classic references in the Barro–Gordon setup; Vestin (2006), Svensson and Woodford (2004), Walsh (2003), and Bilbiie (2014) study optimal delegation in the NK model.

[6] That paper also solves for OFG (along with deriving in closed-form determinacy properties, conditions to solve the FG puzzle, and other monetary policy implications including in LTs) in an analytical heterogeneous-agent NK model; see also Bilbiie (forthcoming) for an interpretation of the latter class of models through the prism of a NK cross, and Bilbiie (2008) for an early example of such an analytical two-agent (TANK) model.

decision *only* and leads to $\kappa = (\varphi + \sigma^{-1})(1 - \zeta)/\zeta$, with ζ now interpreted as the ratio of the shares of flexible to fixed prices; notice that firms' shares are held by a myopic mutual fund that solves the pricing decision and distributes profits to the households. I relax this assumption (consider the standard (2) with $\beta_e > 0$: firms owned by the households) in Section IIA to show that the optimal FG remains essentially unchanged. The upper-left panel of Figure 2 therein shows that optimal FG varies very little with the discount factor of firms (the derivations are in online Appendix A4).

natural rate

A. *The Liquidity Trap*

I follow the seminal paper by Eggertsson and Woodford (2003) to model the zero lower bound. Namely, ρ_t follows a Markov chain with two states. The first is the steady state denoted by S, with $\rho_t = \rho$, and is absorbing: once in it, there is a probability of 1 of staying. The other state is transitory and denoted by L: $\rho_t = \rho_L < 0$ with persistence probability p (conditional upon starting in state L, the probability that $\rho_t = \rho_L$ is p, while the probability that $\rho_t = \rho$ is $1 - p$).

At time t, consider a negative realization of $\rho_t = \rho_L < 0$, meant to capture in this reduced-form model an increase in spreads as in Woodford (2011) and Cúrdia and Woodford (2010). To simplify even further, I assume that the monetary authority tracks the natural interest rate of this economy ρ_t whenever feasible, meaning $i_t = \max(\rho_t, 0)$. It follows that the ZLB will bind when $\rho_t = \rho_L < 0$, while the flexible-price efficient equilibrium will be achieved whenever $\rho_t = \rho$.

Since the shock is unexpected, we can solve the model in the ZLB state, denoting by subscript L the time-invariant equilibrium values of inflation and consumption therein:

$$c_L = \frac{\sigma}{1 - p\nu}\rho_L; \quad \pi_L = \kappa c_L, \tag{4}$$

where

$$\nu \equiv 1 + \sigma\kappa \geq 1$$

is a composite parameter capturing the response of consumption in a liquidity trap to *news* about future income/consumption. It is larger than 1, for an increase in future income is associated with inflation and, at the lower bound, with a fall in real interest rates, which is further expansionary.[7] This effect is stronger the higher the elasticity of intertemporal substitution and the larger the slope of the Phillips curve (the more flexible are prices). Naturally, $\nu = 1$ when aggregate supply is vertical (prices are fixed $\kappa = 0$).

[7] In a "regular" equilibrium whereby the zero bound does not bind and monetary policy follows an active interest rate rule, this parameter is obviously less than one (because news about future income bring about higher real rates, through higher inflation); in particular, $\nu = 1 - \sigma\kappa(\phi - 1)$, where $\phi > 1$ is the response of nominal interest rates to expected inflation.

The LT equilibrium features deflation and a recession as long as

$$p < \nu^{-1}. \tag{5}$$

The restriction is needed to ensure existence of finite fundamental LT equilibria.[8] A complementary condition is instead necessary for the existence of self-fulfilling, sunspot-driven LT equilibria of the type first studied by Benhabib, Schmitt-Grohé, and Uribe (2001, 2002), delivering $c_s = \sigma\rho/(1 - p_s\nu)$: namely, $p_s > \nu^{-1}$ with p_s the persistence of the sunspot; see Mertens and Ravn (2014) for a clear exposition in the context of fiscal multipliers. One can easily apply the apparatus developed below to solve for optimal FG in such sunspot equilibria too. However, as Schmitt-Grohé and Uribe (2017) argues, the optimal monetary policy in such an equilibrium in fact implies *increasing* interest rates (policy which, in a sunspot equilibrium, is expansionary). Throughout this paper, I therefore confine attention to fundamental LT equilibria whereby constraint (5) holds.

B. *Forward Guidance*

In this paper, I advance a (to the best of my knowledge) novel way to model FG that allows for transparent closed-form solutions; namely, I model FG stochastically through a Markov chain, as a state of the world with a probability distribution, as follows. Recall that the (stochastic) expected duration of the LT is $T_L = (1-p)^{-1}$, the stopping time of the Markov chain. I assume that after this time T_L, the central bank commits to keeping the interest rate at 0 while $\rho_t = \rho > 0$, with probability q. Denote this state by F, and let $T_F = (1-q)^{-1}$ denote the expected duration of FG. Appendix A reviews the main properties of the Markov chain implied by this structure: there are three states: liquidity trap L ($i_t = 0$ and $\rho_t = \rho_L$), forward guidance F ($i_t = 0$ and $\rho_t = \rho$), and steady state S ($i_t = \rho_t = \rho$), of which the last one is absorbing. The probability to transition from L to L is, as before, p, and from L to F is $(1-p)q$. The persistence of state F is q, and the probability to move back to steady state from F is hence $1-q$. Notice that this is related to Woodford's (2011) analysis of the future fiscal stimulus's multiplier.

Under this stochastic structure, expectations are determined by

$$\begin{aligned} E_t c_{t+1} &= p c_L + (1-p) q c_F, \\ E_t \pi_{t+1} &= p \pi_L + (1-p) q \pi_F. \end{aligned} \tag{6}$$

Evaluating the IS (1) and Phillips (3) curves during state F, solving for the time-invariant equilibrium during that state (once the trap is over and during the

[8] An additional restriction rules out "starvation" $p < (1 + \sigma\rho_L)\nu^{-1} < \nu^{-1}$; it is needed to ensure that consumption stays positive, $C_L > 0$ in levels or $c_L > -1$ in deviations. This restriction is in fact tighter than (5), thus ruling out explosive paths.

FG period), and using it to solve for the time-invariant equilibrium during the L state delivers Proposition 1.

PROPOSITION 1: *Equilibrium consumption and inflation during the forward guidance state F and the liquidity trap state L, respectively, are given by*

$$c_F = \frac{\sigma}{1 - q\nu}\rho, \tag{7}$$

$$c_L = \frac{1-p}{1-p\nu}\frac{q\nu}{1-q\nu}\sigma\rho + \frac{\sigma\rho_L}{1-p\nu}, \tag{8}$$

and $\pi_F = \kappa c_F, \pi_L = \kappa c_L$.

As already emphasized by the literature numerically, keeping the interest rates low creates an expansion in the "future" (in the F state). This is true as long as the persistence probability of the FG state satisfies the same restriction as the persistence of the trap itself, namely

$$q < \nu^{-1}. \tag{9}$$

The future expansion is increasing in the degree of forward guidance q:

$$\frac{dc_F}{dq} = \frac{\sigma\nu}{(1-q\nu)^2}\rho = \frac{\nu}{1-q\nu}c_F > 0.$$

In addition to creating a future expansion, FG also mitigates the effect of the LT by bringing about an expansionary force today. In particular, in the L state we now have (evaluating the model (1)–(3) taking into account the zero lower bound and expectation formation as defined in (6)):

$$c_L = \frac{(1-p)q\nu}{1-p\nu}c_F + \frac{\sigma}{1-p\nu}\rho_L. \tag{10}$$

Replacing the equilibrium expression in the F state here, we obtain the equilibrium solution during the liquidity trap under forward guidance c_L. More FG leads to higher consumption (and hence inflation) during the trap,

$$\frac{dc_L}{dq} = \frac{\nu(1-p)}{1-p\nu}\frac{1}{1-q\nu}c_F > 0,$$

restricting attention to the standard region, whereby $q > 1/\nu$. Indeed, as the probability approaches this bifurcation point, the effect of FG becomes explosive. Carlstrom, Fuerst, and Paustian (2015) and Kiley (2016) noticed this in the context of the NK model with *deterministic* shocks (known duration) too; the latter paper shows that sticky-information models are free of such behavior and other paradoxical results occurring in sticky-price models.

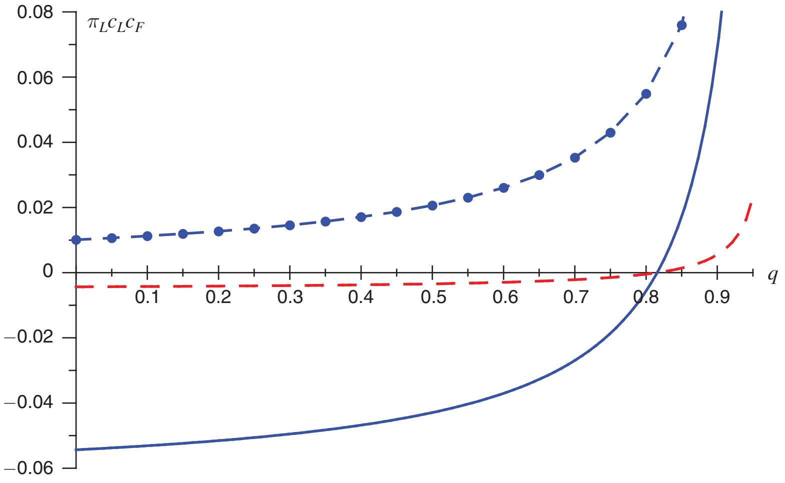

Figure 1

Note: Consumption during LT (blue solid line) and FG (blue dash-dot line) and inflation (red dashed line) during LT as a function of FG persistence.

This model feature is also related to what Del Negro, Giannoni, and Patterson (2012) dubbed the *forward guidance puzzle*. The way to understand the puzzle here is that not only is the effect of FG eventually explosive in its own persistence q, the FG puzzle as defined by the authors is more about a second model property—the power of FG itself (its expansionary effect during the trap dc_L/dq) is increasing with the persistence of the trap p—the later FG takes place, the higher its power. In my analytical framework, this cross derivative is easily calculated:

$$\frac{d^2 c_L}{dq dp} = (\nu - 1)\frac{\nu c_F}{(1 - p\nu)^2(1 - q\nu)} \geq 0.$$

Intuitively, the higher the composite parameter capturing the effect of news ν, the more acute the FG puzzle is (with the effect becoming at best invariant to the horizon in the limit case of fixed prices $\nu = 1$).[9]

Figure 1 plots consumption in the liquidity trap and in the FG state, along with (annualized) inflation in the liquidity trap, as a function of the FG probability q for the parameter region defined by (9). The illustrative parametrization used in the

[9] Several solutions to the puzzle exist, relying on some form of "discounting" in the Euler equation. Examples include Del Negro, Giannoni, and Patterson (2012) (using a perpetual-youth model); García-Schmidt and Woodford (2019), which uses a model in which agents follow a specific cognitive process to form expectations, leading to the notion of "reflective equilibrium"; and McKay, Nakamura, and Steinsson (2016), which uses an incomplete-markets model. Other HANK-type models that feature aggregate-demand amplification along the lines of Bilbiie (2008), imply Euler equation *compounding* instead of discounting, and an *aggravation* of the FG puzzle; see, e.g., Bilbiie (2018, forthcoming).

figure has $\beta = 0.99$, $\psi = 0.01$, $\sigma = 1$, $\varphi = 1$, $p = 0.8$, and a spread shock of $\rho_L = -0.01$, i.e., 4 percent per annum. This delivers a recession of 5.5 percent and annualized inflation of 1 percent in the absence of FG $(q = 0)$.[10] Notice that FG $q = 0.815$ closes both the consumption gap and inflation, illustrating a more general result emphasized in the following proposition.

PROPOSITION 2. *The duration of FG that stabilizes the economy perfectly* (*closes the gap and delivers zero inflation*) *is determined by the persistence probability*

$$q^0 = \frac{1}{\nu}\frac{\Delta_L}{1 - p + \Delta_L},$$

where

$$\Delta_L \equiv \frac{-\rho_L}{\rho} > 0.$$

The term Δ_L is a new parameter capturing the relative size of the *financial disruption* causing the ZLB episode (and recession) and is one of the key determinants of optimal FG. Cúrdia and Woodford (2009) outlines a model with credit frictions in which this disruption occurs endogenously as a spread shock. Clearly, q^0 defined by Proposition 2 is *not* the optimal horizon of FG. Nor is FG necessarily pushed forward into the future as much as possible (which, in this stochastic case, means "close to the asymptote"). The reason why it is not optimal to use FG to perfectly close the gap in the trap has to do with the welfare costs of FG, which are in fact incurred in the future, in the F state.

Optimal policy consists of FG that strikes a balance between these two opposing forces, as we will see next. First, I solve for what I call "optimal forward guidance," taking as constraints the (time-invariant) equilibrium conditions and, given a policy of keeping interest rates low beyond the end of the trap with a certain probability q, I solve for the probability q that maximizes welfare, which determines the expected duration of FG. Then, I solve for the full Ramsey-optimal policy: the optimal path of interest rates that maximizes welfare, which in equilibrium will imply FG—interest rates are at zero beyond the end of the trap for a certain duration, as shown in the literature reviewed in the introduction. As we shall see, the two notions of optimal policy imply FG durations that are very close for standard parameterizations, the advantage of the former being that it delivers a closed-form solution.

II. Optimal Forward Guidance

It is well known (Woodford 2003, chapter 6; see also Woodford 2011 for a ZLB application) that in this economy the welfare function can be represented as a quadratic loss function under certain conditions that are fulfilled here (an optimal

[10] Bianchi and Melosi (2017) shows that accounting for *policy* uncertainty (about how public debt will be stabilized) helps explain the lack of deflation characterizing the US data post-2008 (which contradicts the prediction of the simple NK model).

subsidy makes the steady-state efficient); namely, a benevolent central bank will minimize

$$E_0\sum_{t=0}^{\infty}\beta^t L_t = \frac{1}{2}E_0\sum_{t=0}^{\infty}\beta^t\left(\pi_t^2 + \chi c_t^2\right), \tag{11}$$

where χ is the welfare-based relative weight attached to stabilizing real activity, equal in the baseline model to κ/ε (with ε the elasticity of substitution between individual goods).

Since the equilibrium solution is time invariant in each of the three states, the per period loss function is, for any state, $j = \{L, F, S\}$: $\pi_j^2 + \chi c_j^2 = \left(\chi + \kappa^2\right)c_j^2$. Recall that in state S the economy is back to steady state, so the loss there is 0. Appendix A uses the Markov chain structure to derive in detail the appropriate lifetime welfare objective, which is of the form

$$W = \frac{1}{1-\beta p}\frac{1}{2}\left[c_L^2 + \omega(q)\,c_F^2\right],$$

where $\omega(q)$ is the appropriate discount factor for the FG state, given the Markov chain structure,

$$\omega(q) = \frac{1-\beta p + \beta(1-p)q}{1-\beta q}. \tag{12}$$

In particular, the optimal weight counts for the times the process spends in state F when starting from F (given by $(1-\beta q)^{-1}$), as well as for all the times spent in time F when starting from L, before being absorbed into S (given by $\beta(1-p)q/((1-\beta p)(1-\beta q))$. Notice that $\omega(q)$ is increasing in q. The longer the economy spends in F, the larger the total welfare cost of consumption variability in that state.

The central bank chooses FG duration (persistence probability q) by solving

$$\min_q W = \frac{1}{1-\beta p}\frac{1}{2}\left[c_L^2 + \omega(q)c_F^2\right], \tag{13}$$

taking as constraints the equilibrium values c_F and c_L given, respectively, by (7) and (8). The first-order condition is

$$c_L\frac{dc_L}{dq} + \omega(q)\,c_F\frac{dc_F}{dq} + \frac{1}{2}\frac{d\omega(q)}{dq}c_F^2 = 0 \tag{14}$$

and has a clear, intuitive interpretation. The first term is the welfare *benefit* of more forward guidance, through remedying the LT-caused recession and hence minimizing consumption volatility in the trap; this is proportional to the level of consumption in the trap. The last two terms are the *total cost* of forward guidance. The former is the direct cost, a future consumption boom being associated with inefficient volatility; the latter is the discounting effect discussed above; the longer the time spent under FG, the larger the cost (which is proportional to consumption volatility in the F state).

I first derive optimal FG in closed form under an additional simplifying assumption, which I relax below to show that results are essentially unchanged.

Namely, consider a welfare function that is simplified even further; the central bank attaches *equal weights* (in the sense of ignoring the extra discounting costs) to the future and present in its intertemporal objective, $\omega(q) = 1$, $\omega'(q) = 0$, hence solving

$$\min_q \frac{1}{2}\left(c_L^2 + c_F^2\right).$$

Strictly speaking, this is only accurate in the limit whereby the central bank (too) has a discount factor of 0, a case in which it nonetheless still cares about volatility in the F state. I relax this assumption and study the case of optimal discounting in Section IIA, showing that optimal FG is essentially unchanged. More generally, this case provides an *upper bound measure* on the amount of optimal FG because it ignores part of the welfare *costs* of FG. Furthermore, since this component of the welfare cost is proportional to squared consumption, it is second order and likely to be small. In this case of treating the present and future symmetrically, the first-order condition simplifies to

$$-c_L \frac{dc_L}{dq} = c_F \frac{dc_F}{dq},$$

which delivers Proposition 3 (Appendix B shows that the sufficient second-order condition also holds).

PROPOSITION 3: *The optimal duration of FG is* $q = 0$ *if* $\Delta_L < (1 - p\nu)^2 / (1 - p)$ *and* $q^* > 0$ *otherwise*, *where*

$$q^* = \frac{1}{\nu}\frac{\Delta_L - \frac{(1-p\nu)^2}{1-p}}{1 - p + \Delta_L}.$$

The closed-form solution obtained here (by virtue of the tractable setup used to model FG) allows novel insights into the magnitude of the optimal duration and its economic determinants, insights that are likely to translate to more complex frameworks that embed this core mechanism. First, *some FG* is optimal $(q^* > 0)$ whenever the size of the disruption is larger than the threshold defined by the proposition, which under the baseline calibration is 0.14. Conversely, if the disruption is smaller than this threshold, it is optimal to *refrain from FG altogether*. The intuition is that when the disruption is small, the welfare cost of the trap, albeit first order, is also small, and so is the benefit of FG. Therefore, the welfare cost of FG, although second order, is enough to prevent an optimizing central bank from doing FG; to the best of my knowledge, this finding is novel. Second, quite evidently, the optimal level of FG is strictly lower than the perfectly stabilizing level of FG $(q^* < q^0)$ since the latter simply ignores the welfare cost of FG in the future.

Optimal FG is shaped by its key determinants as follows. The higher the disruption Δ_L and/or its persistence p, the higher the optimal level of FG $(dq^*/d\Delta_L > 0;\ dq^*/dp > 0)$. The intuition is the same for both parameters; larger shock or larger persistence creates a higher recession, a higher welfare cost of the LT, and hence more of a welfare scope for FG. With the composite parameter $\nu = 1 + \kappa\sigma = 1 + \psi(\varphi\sigma + 1)$, things are different. Note that ν itself is increasing with price flexibility ψ and with intertemporal substitution σ, while it is decreasing with labor elasticity (increasing with φ). The optimal level of FG is increasing with ν $(dq^*/d\nu > 0)$ if

$$\Delta_L < \frac{1 - (p\nu)^2}{1 - p},$$

which under the baseline parameterization is 1.67. More price flexibility calls for more FG, but only when the disruption is not "too large." Given that the steady-state interest rate is around 1 percent, this restriction seems very likely to be verified (it requires that the shock be lower than 6.5 percent per annum).

Robustness: Optimal Discounting and Forward-looking Pricing.—In order to obtain sharp closed-form solutions, I made two simplifying assumptions: equal discounting of present and future and a contemporaneous Phillips curve. I now illustrate that the results are robust to relaxing these assumptions; since a closed-form solution is not feasible any longer, we do this numerically. Take first **optimal discounting**; the optimal duration of FG is determined by solving (14) under (12). Appendix A contains the detailed solution and Figure A1 illustrating that the optimal q in this case (plotted as a function of p and ν) is very close to the equal-weights q^* in Proposition 3.[11]

Turn now to the more general **new Keynesian Phillips curve** with discounting $\beta_e > 0$, (2). Online Appendix A4 derives the full solution under FG, which has the same structure as in the previous, simpler case. Figure 2 plots the optimal FG duration as a function of the degree of forward-looking pricing β_e, the LT persistence p, the slope of the Phillips curve, and the disruption, each for two cases: equal weights $\omega(q) = 1$ and optimal discounting (12).[12] The figure illustrates two main findings. The first (in the upper left panel) is that the degree of optimal FG is *essentially invariant* to the discount factor of firms; the blue solid line is almost a parallel line to the horizontal axis, which shows that optimal FG derived under $\beta_e = 0$ is a good approximation to the general case. The second finding is that optimal FG derived under equal weights and optimal FG under optimal discounting are very similar; the solid blue and dashed red lines are very near to each other.

That the discount factor of a firm's shareholders β_e (the degree of forward-looking pricing) influences very little the degree of optimal FG is one facet of a more general insight. FG in this model is mostly about *aggregate demand* rather than aggregate supply. A related advantage of focusing on the simpler Phillips curve

[11] The main difference concerns the trap persistence. With optimal discounting, the welfare cost of FG now receives a larger weight, and it is only optimal to do FG if things are bad enough (i.e., if the size and/or persistence of the trap are large enough), an illustration q^* being an upper bound on optimal FG. Online Appendix A3 provides a supplementary figure comparing this with Ramsey policy.

[12] The figure makes use of a numerical solution. I show in the Appendix that the solution of the optimal FG problem now results in a (sixth-order) polynomial equation in q that cannot be solved in closed form.

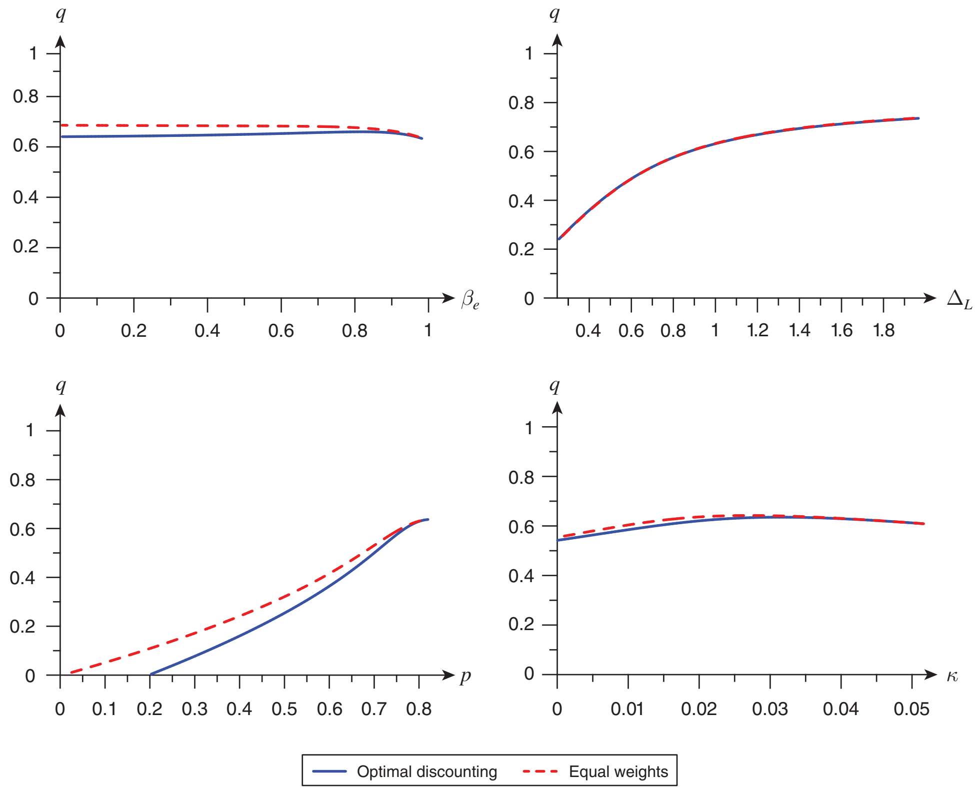

FIGURE 2. OPTIMAL FG PERSISTENCE UNDER NK PHILLIPS CURVE, EQUAL WEIGHTS (RED DASHED LINE) AND OPTIMAL DISCOUNTING (SOLID BLUE LINE)

Note: Functions of β, p, κ, and Δ_L, respectively.

(3) is that it isolates those benefits of history-dependent policy that are due *exclusively* to the possibility of a binding lower bound, whereas the existing analysis using the forward-looking (2) mixes those with the standard benefits that are not ZLB related (thoroughly reviewed in Woodford 2003, chapter 7).

In the remainder of the paper, I will therefore work with the simplest case $\beta_e = 0$, i.e., with (3), which allows for obtaining transparent closed-form solutions; but we can now be reassured that the results derived below are robust to considering the more general (2), as can be easily checked numerically in a manner similar to the one just used.

III. Optimal Forward Guidance and Ramsey-Optimal Policy

The notion of OFG proposed here is a rather *restricted* notion of optimality—it confines attention to policies that imply a zero interest rate for a certain duration, and an immediate exit thereafter. How different is this from the *unrestricted*, *fully* *optimal* Ramsey policy with a lower bound pioneered by Eggertsson and Woodford (2003); Jung, Teranishi, and Watanabe (2005); Adam and Billi (2006); Nakov (2008); and reviewed in the introduction? The answer, as it turns out, is not much. I first prove an *equivalence proposition*: under perfect foresight (and maintaining throughout the supply side described by (3) for simplicity), optimal FG is equivalent to Ramsey-optimal policy. I then show that, under uncertainty, the ex ante, expected optimal durations under OFG and Ramsey are also very close, as a function of the expected duration of the shock.

A. *OFG and Ramsey-Optimal Policy under Perfect Foresight*

In this section, I extend the results on OFG to the deterministic setting studied originally by Jung, Teranishi, and Watanabe (2005) and revisited thereafter by others: the discount factor shock lasts a known number of periods $\rho_t = \rho_L < 0$ for $t = 0, T-1$, and $\rho_t = \rho$ for $t \geq T$. The result is summarized in the following proposition, whose proof (which seems instructive in its own right) follows after.

PROPOSITION 4: *In the simple NK model under perfect foresight, OFG and Ramsey policy are equivalent.*

FG in the perfect-foresight setup takes the form of a pre-announced number of periods with zero interest rates T^*: $i_t = 0$ and $\rho_t = \rho$ for $t = T, \ldots, T+T^*-1$ and $i_t = \rho$ for $t \geq T+T^*$. **Optimal FG** is found by first solving for the equilibrium for a *given FG* duration T^*, and considering a central bank that chooses the welfare-maximizing T^*. We first solve for the equilibrium during the FG period and then during the LT period. During FG, i.e., between T and $T+T^*$, we have (using the boundary condition $c_{T+T^*} = 0$) for j from 0 to T^*-1,

$$c_{T+j} = \left(\nu^{T^*-j} - 1\right)\kappa^{-1}\rho. \tag{15}$$

Evidently, more FG (higher T^*) leads to a higher future expansion $dc_{T+j}/dT^* = \nu^{T^*-j}\kappa^{-1}\rho\ln\nu$. Given this future solution, which at $j = 0$ gives the boundary condition for c_T, the equilibrium during the LT (for t such that $0 \leq j < T$) is readily obtained as

$$\begin{aligned} c_j &= \nu^{T-j}c_T + \left(\nu^{T-j} - 1\right)\kappa^{-1}\rho_L \\ &= \kappa^{-1}\rho\left(\nu^{T-j}\left(\nu^{T^*} - 1\right) - \left(\nu^{T-j} - 1\right)\Delta_L\right). \end{aligned} \tag{16}$$

More FG (higher T^*) also leads to a higher expansion today $dc_j/dT^* = \nu^{T+T^*-j}\kappa^{-1}\rho\ln\nu$. The derivative itself is increasing in T, which illustrates the "FG puzzle" discussed above (the further FG is pushed into the future, the higher its effect). *Optimal FG* under perfect foresight is found by differentiating the intertemporal objective function (11) with respect to T^*, subject to the constraints (15) and (16); the first-order condition is[13]

$$\sum_{j=0}^{T-1} \beta^j c_j \frac{dc_j}{dT^*} + \sum_{j=0}^{T^*-1} \beta^{T+j} c_{T+j} \frac{dc_{T+j}}{dT^*} = 0.$$

Replacing and calculating the sum, the optimal FG duration T^* is a solution to

$$(17) \qquad \Delta_L\left(\nu^T \frac{1-\left(\beta\nu^{-2}\right)^T}{1-\beta\nu^{-2}} - \frac{1-\left(\beta\nu^{-1}\right)^T}{1-\beta\nu^{-1}}\right)$$
$$= \left(\nu^{T^*}-1\right)\nu^T \frac{1-\left(\beta\nu^{-2}\right)^T}{1-\beta\nu^{-2}}$$
$$+ \nu^{T^*}\left(\beta\nu^{-1}\right)^T \frac{1-\left(\beta\nu^{-2}\right)^{T^*}}{1-\beta\nu^{-2}} - \left(\beta\nu^{-1}\right)^T \frac{1-\left(\beta\nu^{-1}\right)^{T^*}}{1-\beta\nu^{-1}}.$$

Now consider **Ramsey policy** in this deterministic setting. This is a simplified version of Jung et al. (2005), using the same Lagrangian method as theirs to solve the problem

$$\min_{c_t} E_0 \sum_{t=0}^{\infty} \beta^t \left[c_t^2 + 2\phi_t(c_t - \nu c_{t+1} - \sigma\rho_t)\right]$$

subject to

$$(18) \qquad c_t \leq \nu E_t c_{t+1} + \sigma\rho_t,$$

where the quadratic loss (under (3)) per period is still $\left(\kappa^2+\chi\right)c_t^2$, but I rescaled the Lagrange multiplier of the constraint ϕ_t accordingly. To obtain the inequality constraint on the control c_t, I combined (1), (3), and the lower bound $i_t \geq 0$. The first-order conditions are

$$(19) \qquad c_t + \phi_t - \beta^{-1}\nu\phi_{t-1} = 0,$$
$$(c_t - \nu E_t c_{t+1} - \sigma\rho_t)\phi_t = 0, \quad \text{and} \quad \phi_t \geq 0,$$

where ϕ_t is the costate Lagrange multiplier on the constraint, with given initial value $\phi_{-1} = 0$. The second line in (19), together with the constraint, summarizes the Kuhn-Tucker conditions.

Optimal commitment policy implies that the ZLB constraint binds ($\phi_t > 0$ and $i_t = 0$) for longer than the duration of the financial shock, as already emphasized by all the papers cited above.[14] For the purpose of comparison with

[13] Strictly speaking, this is correct when time is continuous and Leibniz's rule can be applied. In writing the first-order condition, I make use of a limit argument and ignore the effect of changing T^* on the objective function through changing the limit of summation; that term is propositional to $c_{T+T^*}^2$, which is 0.

[14] In particular, equation (35) in Jung et al. (2001) reads $0 \leq T^d \leq T^c < \infty$; with known shock duration, the time of zero interest rates under discretion T^d is weakly positive and lower than the time of zero interest under

the notion of "optimal FG" introduced above, I solve for the optimal duration of FG as implied by Ramsey policy and call it T^{*R}, as a function of the duration of the shock T. Unlike the papers mentioned above, and again by virtue of our simplifying assumptions, we can solve for the path of Lagrange multipliers ϕ_t in closed form. Appendix B describes the solution in detail.

Here, since it is so simple yet captures transparently a key mechanism of the literature reviewed above, it is useful to describe the main idea. A policy of zero interest rates is defined implicitly by a binding constraint, i.e., $\phi_t > 0$; we can therefore combine the second equation (the binding constraint) in (19) with the first and obtain a second-order stochastic difference equation. With our simple model, this equation is very easy to solve; provided that the roots are on the right sides of the unit circle (which they are; as I show in the Appendix, the roots are $\nu^{-1} < 1$ and $\beta^{-1}\nu > 1$), a unique equilibrium requires two boundary conditions. The first is the initial condition, $\phi_{-1} = 0$, and the second is a boundary condition that implicitly defines the stopping time, i.e., the expected duration of zero interest rates implied by Ramsey policy T^{*R}. Thus, we find T^{*R} by solving the boundary condition

$$\phi_{T+T^{*R}-1} = 0, \tag{20}$$

where T is the duration of the shock. In Appendix B, I show that this delivers a (nonlinear) equation linking T^{*R} and all the model parameters including the shock duration of the form $\tilde{\phi}(T^{*R}) = \hat{\phi}(T^{*R}, T)$ (the functions are presented in the Appendix). Note that $\tilde{\phi}(\cdot)$ is an increasing function of the FG duration and gives the value of relaxing the constraint contingent upon the shock taking the value ρ_L, value calculated at the stopping time T. At the same moment T, $\hat{\phi}(\cdot)$ gives the value of relaxing the constraint during FG, i.e., once the shock stopped hitting but interest rates are still at zero. The optimal duration of FG is found at the intersection of the two.

In the Appendix, I show that this leads to *exactly* the same equation as when choosing the FG duration directly, (17), thus proving Proposition 4.

B. *Ramsey and OFG under Stochastic Duration*

Matters are slightly different with Markov shocks, as in Eggertsson and Woodford (2003). Ramsey-optimal policy solves exactly the same problem as above—namely (19)—but between 0 and $T-1$ there is now uncertainty, and T is itself a random variable with expected value $T_L = E(T) = (1-p)^{-1}$. In Appendix B, I show, in a manner similar to the one used above under perfect foresight, how to derive a measure of the *ex ante*, *expected FG duration* under Ramsey policy, as a function of the *expected* duration of the shock. I compare the Ramsey-implied FG duration T_F^R with the optimal FG duration found

commitment T^c, which is itself finite. Hasui, Sugo, and Teranishi (2016) suggests, intriguingly, that in an economy with inflation, persistence optimal policy is characterized by front-loading, or early tightening, rather than forward guidance; rates should be increased before the trap ends.

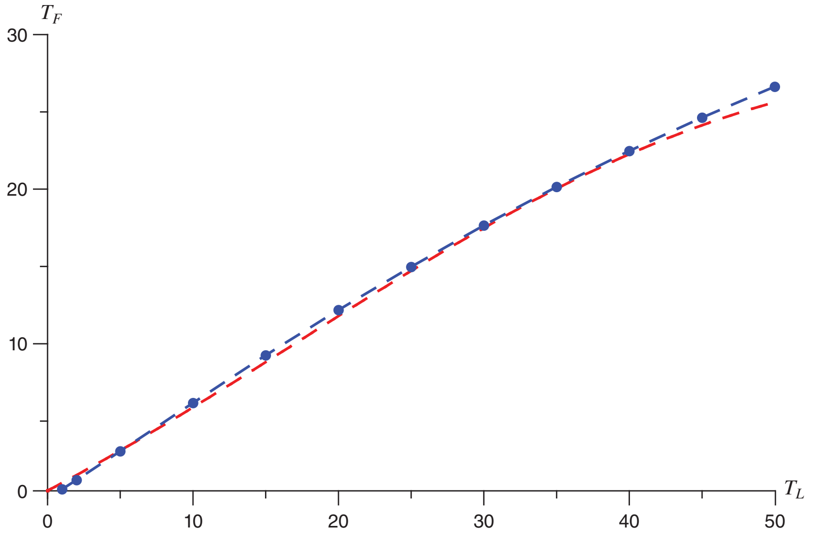

FIGURE 3

Note: FG duration implied by Ramsey policy (blue dot-dash line), along with optimal FG duration under equal weights (red dash line).

above in closed form, $(1-q^*)^{-1}$, for the baseline parameter values, in Figure 3; the measures are remarkably close to each other.[15] This extends to other calibrations, and also holds for the optimal-discounting optimal FG (see Figure A1 in the Appendix). The bottom line is that the simpler notion of optimal FG proposed in this paper gives very similar conclusions to those arising in a full Ramsey-optimal monetary policy analysis, such as those previously used to argue for forward guidance on welfare grounds.

IV. The Best of Both Worlds: A Simple Rule for Forward Guidance

Both of the optimal FG concepts derived above assume, and in fact rely upon, commitment. But adopting an optimal policy rule and explaining its determinants may be very hard to communicate in practice, even in this simple model and even for the simplest notion of optimal policy introduced here and summarized by q^*. This may indeed aggravate the credibility problem inherent to this type of policy commitment, an issue likely to be amplified in more complex and realistic models where the determinants of optimal FG will certainly be less transparent. The next section proposes a *simple rule* for FG to address such issues.

[15] One caveat to note is that of course, since the duration is stochastic, the realization $T = T_L$ has itself a probability $\Pr(T = T_L) = p^{T_L}(1-p)$; it is a different exercise to calculate actual FG duration as a function of actual shock duration for a given probability, and thus taking into account that durations have a probability distribution $\Pr(T = x) = p^x(1-p)$. In online Appendix B, I report Figure B1 plotting this for different probabilities, produced by Taisuke Nakata and Paul Yoo using Fortran code. While this figure is informative, it is not the right standard of comparison with OFG, which, as defined here, is not calculated as such; the "fair" comparison of comparable objects is in Figure 3.

A. *A Simple FG Rule in the Baseline NK Model*

The simple rule whose virtues I analyze here combines the advantages of date-based (or time-dependent) and state-contingent FG; in that sense, it allows for achieving the best of both worlds. It works as follows. The central bank first commits to (and communicates to the public) that it will keep interest rates low *after* the trap is over (once the natural rate returns positive) for a duration of

$$T_F^S = \frac{1}{2}\Delta_L T_L, \tag{21}$$

where the superscript S denotes "simple," and we recall that $T_L = (1-p)^{-1}$ is the expected duration of the trap and Δ_L.

The key element is that at the time of the announcement (i.e., commitment and communication of the policy), Δ_L and T_L do not need to be known or measured ex ante. The central bank need only commit to the rule itself, without providing any estimate of these objects. Ex post, once the trap is over, the duration of the trap is (tautologically) observable, and the size of the disruption can be measured. From this moment onward, looking back, the central bank thus computes the precise value of the number of periods for which it will keep interest rates at zero T_F^S and communicates this to the public.

In central bank policy language, is the simple rule date based or state contingent? Both. It is date based because it consists of announcing a duration. It is state contingent because it is triggered once a (verifiable) state takes a certain value, i.e., when the natural interest rate returns positive.

The precise expression (21) is obtained using the optimal FG closed-form solution in Proposition 3 for the case with Markov shocks. Note that T_F^S is (almost) the expected duration $(1-q_S)^{-1}$ corresponding to a probability given by $q_S = 1 - 2\left(1 + \frac{\Delta_L}{1-p}\right)^{-1}$.[16] The "almost" qualifier stands for a minor, half-period adjustment that is just a quibble.[17]

The simple rule (21) captures intuitively the welfare-based determinants of optimal FG: the disruption and the duration of the trap, Δ_L and T_L. Figure 4 plots this simple-rule duration together with the optimal durations analyzed above (which are the same as in the previous figure). The simple rule does remarkably well at approximating optimal policy; the maximum difference under this calibration is of almost two quarters when the trap lasts eight to ten years. In online Appendix C, I

[16] This relies on an approximation of q^* around the point where $\nu = 1$, which is not too farfetched in practice since prices are very sticky on aggregate (the estimated slope of the Phillips curve is in the 0.01–0.05 range). This is virtually indistinguishable from the optimal probability q^* for the baseline calibration and calibrations within that range.

[17] Strictly speaking, the duration corresponding to q_s is $\frac{1}{2}(1 + \Delta_L T_L)$, while the T_F^S in (21) is the expected duration corresponding to the probability $q = 1 - 2\frac{1-p}{\Delta_L}$. This probability is the limit (as $\nu \to 1$) of the optimal probability in a slightly different model in which the probability of transitioning from L to F is $1-p$ (instead of $(1-p)q$) and otherwise the transition matrix is identical (q is still the persistence of FG). In other words, forward guidance in that model is guaranteed to last at least one period, which explains the "plus one." The difference is, however, minor, and I have chosen the expression in text because it is simpler.

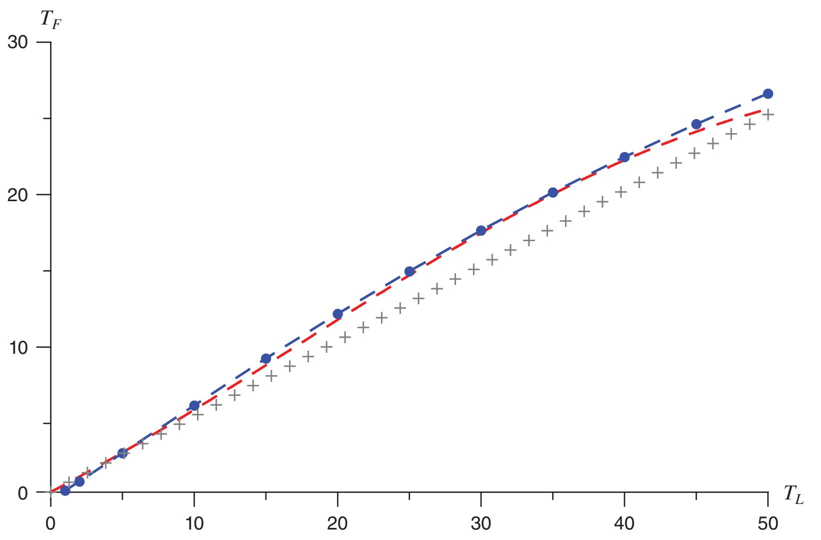

FIGURE 4

Note: FG duration implied by the simple rule (crosses), Ramsey policy (blue dot-dash line), and OFG under equal weights (red dash line), as a function of trap duration.

TABLE 1—LOSS RELATIVE TO OFG

	Markov		Perfect foresight	
	$p = 0.9$	$p = 0.975$	$T = 10$	$T = 40$
$\frac{L_{SIT}}{L_{OFG}}$	2.057	20.833	3.333	7.246
$\frac{L_S}{L_{OFG}}$	1.01	1.416	1.00	1.021

similarly derive a simple rule for the case of perfect foresight, known shock duration. The expression is still closed form but somewhat more involved, and it is even closer to optimal FG.

How does the simple FG rule fare in terms of welfare? I start by evaluating this in the representative-agent model, an endeavour similar in spirit to Eggertsson and Woodford's comparison of price-level targeting and SIT. I evaluate the loss (11) under the simple rule relative to optimal policy (OFG), and compare it to the relative loss under SIT (no FG). For the Markov case, I replace the equilibrium outcomes c_L and c_F found in Proposition 1 evaluated at $q = q^*, q_S$ and 0, respectively. For the perfect-foresight case, I replace the equilibrium expressions (15) and (16), where the extra accommodation is, respectively, T^*, T^*_S, and 0 (for $T = 10$, we have $T^* = T^*_S = 4$, while for $T = 40$, $T^* = 16$ and $T^*_S = 17$). Table 1 summarizes the results for two trap durations: 10 periods and 40 periods (in the Markov case, these are the expected durations given the probabilities).

The simple rule delivers a welfare level that is very close to optimal policy, especially relative to the suboptimal policy of strict inflation targeting. Its key advantage, however, comes from its communication virtues. A central bank wanting to implement FG that is close to optimal in this model can communicate T_F^S to the public *even if* it *does not know ex ante* the length of the liquidity trap T_L. It is enough that the central bank states, "whatever the length of the trap, the nominal interest rate will stay at its effective lower bound for T_F extra periods, where T_F is defined *ex post* (once T_L is observable and Δ_L measurable) as half of one plus the product of the trap's duration and disruption." This alleviates an important communication loophole of FG that I underline below.

The simple rule delivers the best of both worlds, in the following sense. Since it is specified in terms of durations, it is *time dependent*, or *date based*. But since it specifies FG duration as a function of trap duration (which can be observed and verified ex post), it exploits the advantage of a state-contingent rule: it is a way for the central bank to communicate, to signal something about an unobservable state. Evidently, in order to substantiate this point, we need a model with information imperfections, our next topic.

B. *When a Simple Rule Is Better Than Optimal* (*Odyssean*) *Policy: Delphic Guidance for Delphic Agents*

The advantages pertaining to credibility and communication of simple rules apply with particular force in models where the commitment problem is modeled explicitly as related to information frictions. I apply the previous analysis to illustrate this point in such a model, based on Andrade et al. (forthcoming). The analysis can be extended to different environments such as, e.g., Wiederholt's (2015) dispersed information, Bassetto's (2016) analysis of cheap talk and reputation, or Melosi's (2017) model with a "signaling channel." The main benefit of a simple rule in a model with Delphic agents (who take FG to mean communication, rather than commitment) is that it makes FG itself Delphic; it anchors agents' expectations, whereas FG in the form of commitment to an optimal rule becomes self-defeating if some agents use the central bank's announcements to infer a state. As we will see, in such a setup, optimal policy, calculated here in closed form as the optimal duration of (Odyssean, commitment) FG, may even require refraining from FG altogether, when enough agents are Delphic.

Andrade et al. (forthcoming) models the distinction (introduced by Campbell et al. 2012) between Odyssean (commitment) versus Delphic (communicating fundamentals) FG as follows. A fraction of "optimistic" agents perceive FG as Odyssean; they believe that the central bank will keep interest rates low for the announced number of periods after the ZLB stops binding. The remainder, "pessimistic" agents, perceive FG in a Delphic way; they believe that the central bank's commitment in fact signals something about the fundamentals. Namely, pessimists believe that the ZLB will bind for exactly as many periods as the central bank promises to do FG. The paper provides empirical support for this assumption (that agents agree on interest rates but disagree on macroeconomic aggregates) and then shows numerically that the optimal duration of FG depends on the share of pessimistic agents in a nonlinear way; more pessimistic agents call for more FG up to a threshold, beyond which the contractionary effect of FG through pessimistic agents' beliefs dominates and less FG is called for.

To this analysis, I add one layer that allows for closed-form solutions. I model beliefs about probabilities rather than durations. Assume that a fraction α of agents are pessimistic and perceive FG in a Delphic way: they think that in state F, which still occurs with probability $q(1-p)$, the value of the discount factor shock is ρ_L. The remainder fraction of $1-\alpha$ optimistic agents perceive FG in an Odyssean way and expect that in state F the value of the discount factor shock is ρ. This is the only dimension of disagreement, and otherwise the model is identical to the benchmark model studied previously.

When state F materializes, pessimistic agents update their beliefs (ex post, both agents know the true shock); therefore, the equilibrium during the F state is identical for both agents and hence the same as under homogenous beliefs: $c_F^o = c_F^m = c_F = \sigma\rho/(1-q\nu)$, denoting type $j = o,m$ (optimists and pessimists, respectively). Where heterogenous beliefs make a difference is for the L state, through expectations; ex ante pessimists are wrong, and optimists know this. Expectations are therefore (similar equations hold for inflation expectations)

$$E_t^o c_{t+1}^o = pc_L^o + (1-p)qc_F,$$

$$E_t^m c_{t+1}^m = pc_L^m + (1-p)qc_F^{mw},$$

where $c_F^{mw} = \sigma\rho_L/(1-q\nu)$ is the value (wrongly) expected by pessimists ex ante. Solving the model under these beliefs, we obtain aggregate consumption during the trap for each type:

$$c_L^j = \frac{q(1-p)\nu}{1-p\nu}c_F^j + \frac{\sigma}{1-p\nu}\rho_L,$$

where for optimists $c_F^o = c_F = \sigma\rho/(1-q\nu)$, while for pessimists $c_F^{mw} = \sigma\rho_L/(1-q\nu)$.

Aggregate consumption during the trap is therefore (with the same Δ_L notation):

$$c_L = \frac{q(1-p)\nu}{1-p\nu}\frac{\sigma\rho}{1-q\nu}(1-\alpha(1+\Delta_L)) + \frac{\sigma}{1-p\nu}\rho_L.$$

The heterogenous-beliefs channel (Andrade et al. forthcoming) proportionally weakens FG power; dc_L/dq is now scaled down by a factor $1-\alpha(1+\Delta_L)$. Indeed, FG can have perverse *contractionary* effects with enough pessimists $\alpha > (1+\Delta_L)^{-1}$. The threshold is lower the larger the disruption Δ_L (and hence the larger the pessimists' misperception).

Optimal policy consists of balancing two forces. On the one hand, FG being less effective during the trap (because it may be misinterpreted) means *more* of it is needed to achieve the same outcome; this generates *more* scope for FG *ceteris paribus*, up to the point where FG becomes indeed contractionary. On the other hand, the effect of heterogeneous beliefs on the benefits and costs of FG is asymmetric; the welfare cost of inefficient volatility during the F state is independent of α. The trade-off is resolved (for the case of equal discounting) as described in the closed-form expression for the heterogeneous-beliefs OFG duration:

$$q_{HB}^{*} = \begin{cases} 0 & \text{if } \alpha > (1+\Delta_L)^{-1}\left(1 - \dfrac{(1-p\nu)^2}{(1-p)\Delta_L}\right) \\ \dfrac{1}{\nu}\dfrac{\Delta_L - \frac{(1-p\nu)^2}{(1-p)(1-\alpha(1+\Delta_L))}}{(1-p)(1-\alpha(1+\Delta_L)) + \Delta_L} > 0 & \text{otherwise} \end{cases}.$$

Notice, first, that it is optimal to *refrain* from FG for values of α that are lower than the threshold $(1+\Delta_L)^{-1}$ making FG contractionary. The degree of OFG q_{HB}^{*} is *increasing* with the degree of information imperfections α (the expansionary effect of heterogeneous beliefs dominates) as long as α is lower than a certain threshold. Beyond the threshold, the welfare cost of excess volatility once the trap is over—cost that is unaffected by heterogeneous beliefs—dominates instead. Intuitively, the expansionary channel prevails when the scope for FG is strong to start with, i.e., when prices are flexible enough, when there is enough intertemporal substitution (both of which translate into a higher ν), when the disruption is large, and when the trap is persistent.[18]

The foregoing assumes that the central bank chooses an optimal policy within the class of *commitment*, Odyssean policy options (the duration of which commitment is a function of the share of Delphic agents). But the central bank can do better by recognizing the information friction (that some agents are Delphic), and using instead a *communication* policy. Bassetto (2016) shows, in a more general context, that Odyssean FG is redundant in an environment with informational asymmetries, while Delphic FG allows for achieving better equilibria. Our simple rule (21) has this flavor, as it blends the advantages of state-contingent and date-based FG. The former, because it refers to something that all agents can agree on (the end of the trap, ex post). The latter, because it is simpler to communicate a *date* at which interest rates shall turn positive, rather than a complicated mapping into the evolution of some state variables. The simple rule allows for achieving a better equilibrium in this model because it resolves the information asymmetry; it implicitly consists of Delphic FG, which is what Delphic agents need. And it approximates well Odyssean FG, which is what Odyssean agents need.

In fact, in this model, a simple rule can lead to *higher welfare* than optimal Odyssean commitment if enough agents are pessimistic (Delphic). Suppose the central bank announces the simple rule T_F^S, thus anchoring the expectations of Delphic agents who at time $T < T_F^S$ can verify the state $\rho_t = \rho$. The

[18] The exact expression of the threshold is $(1+\Delta_L)^{-1}\left(1 - \frac{(1-p\nu)^2 + (1-p\nu)\sqrt{(1-p\nu)^2+\Delta_L^2}}{(1-p)\Delta_L}\right) < (1+\Delta_L)^{-1}$. Notice that the derivative of q_{HB}^{*} at $\alpha = 0$ is positive if and only if $\frac{\Delta_L}{1-p} > 2\left(\frac{(1-p\nu)^2}{(1-p)^2} - 1\right)^{-1}$.

ensuing equilibrium is therefore the one of homogeneous beliefs, but under the (suboptimal) simple rule. How does welfare under this policy compare with the optimal Odyssean FG derived above, taking as a constraint that some agents are Delphic? Beyond a threshold share of pessimistic (Delphic) agents, the simple rule leads to higher welfare than the optimal Odyssean FG policy. I evaluate this threshold numerically and show that it is small in both the stochastic and deterministic cases.

The loss under the optimal Odyssean FG policy in the stochastic case q_{HB}^* is proportional to $c_L^2(q_{HB}^*(\alpha),\alpha) + c_F^2(q_{HB}^*(\alpha))$, where $c_L(q_{HB}^*(\alpha),\alpha)$ and $c_F(q_{HB}^*(\alpha))$ are the equilibrium values derived above, evaluated at the point where FG is set optimally. The loss under the simple rule is instead invariant to α and given by $c_L^2(q_S) + c_F^2(q_S)$, where q_S is the expression used earlier for (21). It can be easily shown that there exists a threshold share of pessimistic agents $\bar{\alpha}$ beyond which the simple rule dominates Odyssean guidance from a welfare perspective: quantitatively, the share is very small under the standard calibration, namely $\bar{\alpha} = 0.001$ (one in a thousand persons). The threshold share increases but still remains small if we increase the trap persistence ($\bar{\alpha} = 0.005$ when $p = 0.9$ and $\bar{\alpha} = 0.08$ when $p = 0.975$) or the elasticity to news ($\bar{\alpha} = 0.026$ if $\nu = 1.1$ and $\bar{\alpha} = 0.1$ if $\nu = 1.2$).

The deterministic setup is more natural to model the implementation of the simple rule with heterogeneous beliefs. In online Appendix D, I outline the model solution when beliefs are about *duration* (rather than probabilities), as in Andrade et al. (forthcoming). I solve for the optimal FG duration (which, as we know from Proposition 4, coincides with Ramsey-optimal policy) and evaluate welfare numerically under this optimal Odyssean policy under the baseline calibration and with a large duration, $T = 40$. The share of pessimists beyond which it becomes welfare preferable to switch to the simple rule instead of following the optimal Odyssean commitment is a mere $\alpha = 0.015$. Thus, under this admittedly simple calibration, if more than 1.5 percent of the population is pessimistic, it is more preferable, from an aggregate-welfare viewpoint, to follow the simple rule than the optimal Odyssean FG that internalizes the behavior of Delphic agents. Put differently, if 40 percent of agents are Delphic, the welfare loss under Odyssean FG is *almost five (!) times higher* than under the simple rule (see Table D1 in online Appendix D). The policy prescription thus seems to be to "keep it simple."

Read till here.

V. OFG and the Simple Rule in a Medium-Scale DSGE Model

The foregoing analysis has used the simplest three-equation NK model to obtain an analytical solution to the optimal policy problem with a zero lower bound, and a simple rule that approximates that optimal policy reasonably well. Yet policy analysis at central banks (including optimal policy) is most often formulated in the context of empirically relevant, richer, medium-scale DSGE models that—building on the pioneering contributions of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007)—include a variety of realistic frictions.

Yet somewhat paradoxically this has not, to the best of my knowledge, been extended or applied to the analysis of optimal monetary (and fiscal) policy with a zero lower bound. Indeed, *virtually all* the literature approaching this topic used the simple three-equation NK model or small deviations from it (see, e.g., Nakata 2017 for Ramsey policy in a nonlinear Calvo model with nominal public debt). One exception is Bilbiie, Monacelli, and Perotti (forthcoming), solving numerically for "*optimal*" *government spending* (for a given monetary policy) in a state-of-the-art medium-scale DSGE model with real and nominal rigidities, following closely Justiniano, Primiceri, and Tambalotti (2013, hereinafter JPT) and applying it to the analysis of fiscal stimulus during the Great Recession.

Computing the fully optimal Ramsey policy with a ZLB constraint is a very challenging numerical task in this medium-scale model that features, as we shall briefly see, six state variables. I have not yet been able to solve this challenging problem but am currently working on it in related follow-up work. For this reason, this paper only takes one step in this direction. Rather than compute the full Ramsey policy, I exploit the perfect-foresight nature of the solution and compute the medium-scale DSGE equivalent of what we defined above as "optimal forward guidance OFG." That is, I compute the welfare-maximizing number of periods for which the nominal interest rate should be kept at zero beyond the end of the liquidity trap (as before, this is the "optimal policy" but within the restricted class of policies whereby the interest rate is either zero or equal to the normal-times discount rate).

Namely, I solve for OFG—the duration of FG that maximizes welfare, given a duration of the ZLB—in a slightly modified version of the JPT model, which I solve nonlinearly subject to a ZLB constraint using complementarity methods available in Dynare (Adjemian et al. 2011). Then I compute the welfare cost of SIT (i.e., doing nothing) relative to OFG policy and compare this with the welfare cost of following the simple FG rule (relative to OFG).[19]

The JPT model that I use for this exercise extends the three-equation NK model hitherto considered by introducing habits in consumption, investment in physical capital subject to quadratic adjustment costs, sticky nominal wages, and indexation of prices and wages. Differently from JPT, I model nominal rigidity by assuming quadratic costs of adjusting prices and wages à la Rotemberg rather than Calvo—the reason being that the former delivers a compact nonlinear Phillips curve that makes the nonlinear solution employed here easier (whereas with the latter, the dispersions of prices and wages appear as endogenous state variables, augmenting an already large state space).

The description of the complete model is relegated to online Appendix E for brevity. This includes the calibration in Table E1, which follows closely the estimates of JPT, where available, with two differences. First, the price and wage stickiness parameters have no immediate mapping with the Calvo framework used in JPT, so I choose a calibration that follows their relative estimates (prices are

[19] In ongoing follow-up work, we compute Ramsey-optimal policy in this model under a ZLB constraint, a phenomenal technological challenge for this rich model with six endogenous variables with large enough shocks (and long enough duration) to replicate the observed Great Recession.

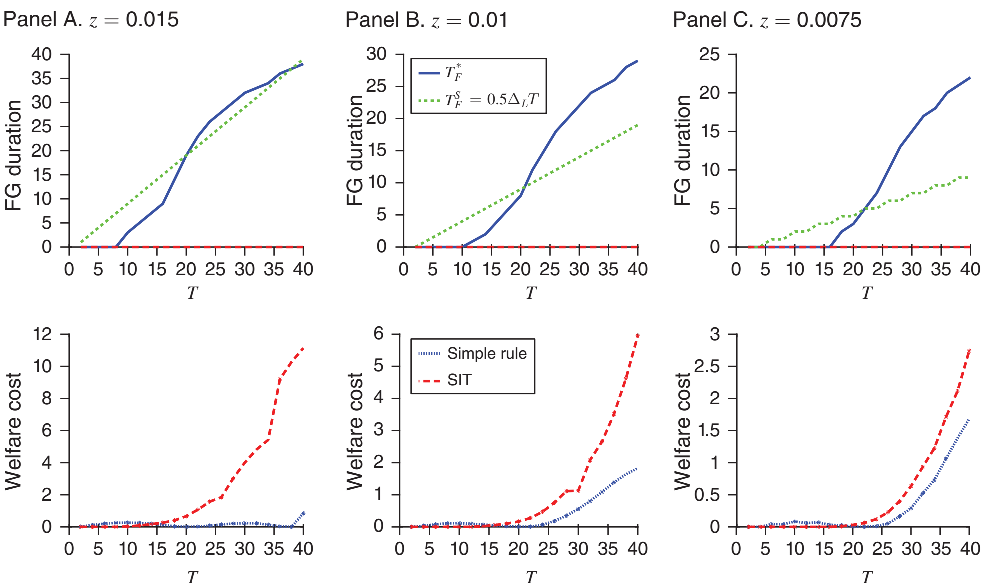

FIGURE 5. OFG IN DSGE

stickier than wages) and conduct two robustness checks below. The second modification is required by the very purpose of the paper. JPT estimates a monetary policy rule here; because I look at (a restricted notion of) an optimal rule, I replace that with a rule consistent with the one used for the simpler model, whereby the central bank chooses the duration of zero interest after the ZLB ends (and else tracks the discount rate). This greatly simplifies the solution and allows modeling FG exactly as before.

With this calibration, I solve the full nonlinear model in online Appendix E subject to the ZLB constraint for a given size z and duration T_L of the shock z_t, and a given FG duration T_F. As a baseline shock size I use $z_t = z = 0.01$, under which the model fares well at reproducing the dynamics of key macroeconomic aggregates during the Great Recession for the observed shock duration T_L (see Bilbiie, Monacelli, and Perotti forthcoming, Figure 2). I then conduct robustness checks for smaller and larger shocks.

As part of the solution, I calculate lifetime welfare of the representative agent $\mathcal{V}(F;\cdot)$. OFG is the welfare-maximizing value of T_F; that is, the value T_F^* that solves

$$T_F^* = \operatorname{argmax} \mathcal{V}(T_F).$$

The upper row of Figure 5 plots this OFG duration (blue solid line) along with the simple-rule duration computed as before $T_F^S = 0.5 \times \Delta_L \times T_L$ (green dashed line), where Δ_L is calculated in the Appendix, while the lower row plots the welfare cost (in equivalent consumption percentage variation, relative to OFG $T_F = T_F^*$) of the simple rule $T_F = T_F^S$ (blue dotted lines) as compared with adopting SIT $T_F = 0$ (dashed red lines). That is the number a such that the household is indifferent, from the standpoint of lifetime utility, between the allocation with OFG and the allocation with FG, for $T_F = T_F^S, 0$, in turn:

$$\mathcal{V}[(1+a)\, C_t; T_F] = \mathcal{V}(\cdot\,; T_F^*).$$

Each row contains three panels for three shock sizes: the baseline $z = 0.01$ (middle panel), as well as a large shock $z = 0.015$ (left) and a small shock $z = 0.0075$ (right).

Our analytical results need to be nuanced in the quantitative model. For one, the numerical DSGE results confirm that it is optimal for the central bank to do FG (any at all) only if the LT duration and shock size are large enough. But this finding is more pronounced in the quantitative model. FG becomes welfare-increasing only beyond an LT duration of 8 to 16 quarters (depending on shock size). FG is thus not necessarily the optimal policy prescription in the presence of *inertia*, a feature of the data (Hasui, Sugo, and Teranishi 2016 provides an illustrative example substantiating this point with a small three-equation model with AR(1) in output). This reinforces the case for solving for the full Ramsey policy in a quantitative DSGE model with a liquidity trap that I pursue in current work.

Once FG is optimal (when the shock and/or its duration are large enough), the simple-rule FG is a quantitatively reasonable approximation to the optimal FG duration as illustrated by the top row of Figure 5; the two FG durations are close in particular in the case of large shocks and only diverge significantly for the case of small shocks. Yet the latter case is the one where the welfare cost of "getting it wrong" is relatively small. As apparent from the lower row, it is at most 1.5 percent when the LT duration is 40 quarters, whereas the cost of "doing nothing" is, while still rather small, twice as large. With larger shocks, the welfare cost of the simple rule stays small, whereas the cost of doing nothing (SIT) increases significantly, dwarfing the former (it is three times as large under the baseline, and ten times as large for the "large shock" case).

Figure E1 in online Appendix E presents robustness checks for two alternative calibrations (for the baseline shock size): more flexible prices and more sticky wages. When prices are more flexible, as expected from our analytical results, more FG is optimal (the OFG and simple-rule durations diverge). That is both because the power of FG is amplified (more future inflation means a higher expansion today through the fall in real rates at the ZLB) and because the welfare cost of FG is larger, thus making it more desirable, because the deflation-driven recession is now larger to start with. Consequently, the welfare costs of both doing nothing and of the simple rule are now larger, although the former is still larger than the latter, especially at high LT durations. With more sticky wages, the opposite is true; the simple rule becomes a good approximation to OFG at high duration when its welfare costs also become negligible, whereas the cost of doing nothing is still large.

VI. Conclusions

This paper proposes a new way to think about optimal policy in a liquidity trap that is simple and easy to communicate. This policy, that I call "optimal forward guidance," consists of committing to keeping the interest rate at zero for a number of periods once the trap is over, and moving directly to the "business-as-usual" optimal policy of normal times thereafter. The number of periods (OFG duration) is the relevant policy choice and is determined by maximizing welfare. I solve for it in closed form and then propose a "simple rule" for determining the optimal duration. While this simpler policy is more restricted than the full-fledged Ramsey-optimal policy extensively analyzed by others, I show that in the baseline new Keynesian model, it is very close to it. Under perfect foresight, the two notions in fact coincide.

To solve in closed form for the OFG duration, I first model FG stochastically, as a state of the world with an attached persistence probability that fully determines its expected duration. The optimal duration depends in a very intuitive way on the duration of the liquidity trap, its severity, and a composite parameter capturing the effect of news on aggregate activity (this parameter instead is a function of the slopes of aggregate demand and supply).

A *simple rule for FG* approximates the optimal duration in the baseline model and can be used to communicate this policy. In the spirit of interest rate rules, it consists of explicitly announcing an FG duration that is *half of the trap's duration times the financial disruption* (the interest rate spreads inducing negative natural interest). The simple rule FG duration is very close to that implied by optimal FG and Ramsey policy, even when the liquidity trap is very long-lived, which reinforces the desirability of this simple rule. Other arguments for adopting such a rule pertain to its being easier to communicate and more transparent, hence increasing its credibility (see Woodford 2012 for a discussion). They apply to any rule-based regime, but they do so more forcefully here, especially at long trap durations, for it is conceivably harder to commit ex ante to a complicated policy that is supposed to happen in (say) ten years' time.

Committing to the simple rule has a further credibility and communication advantage that is specific to FG and is related to more general arguments put forth by Bassetto (2016)—it can allow bypassing the commitment problem inherent to optimal FG with imperfect information. I show this in a simple model of heterogeneous beliefs based on Andrade et al. (forthcoming), but conjecture that it will hold in any framework where some agents use a policy announcement to make an inference about the future state, as in Wiederholt (2015). The simple rule blends the best of two worlds, state-contingent and time-dependent FG. It is a good approximation to commitment (Odyssean) FG, and serves as communication (Delphic) FG for Delphic agents who may otherwise compromise FG policy. Indeed, if enough agents are Delphic, the simple rule delivers higher welfare than optimal Odyssean policy; under a standard calibration, the share of Delphic agents that makes the simple rule welfare dominating is very small.

The simple rule remains a good approximation to OFG in another extension of the baseline model—a considerably richer, empirically realistic medium-scale DSGE model. I compute numerically the durations corresponding to OFG and to the simple rule in a version of JPT's model featuring consumption habits, investment adjustment cost, and stickiness and indexation of both prices and wages. I also evaluate the welfare cost of following the simple rule rather than OFG and compare it with the cost of doing nothing (sticking to strict inflation targeting). Relative to the three-equation model, the results need to be somewhat qualified; some FG is optimal only if the recession is "bad enough" (large shock and/or high duration), but the welfare cost of following the simple rule, and doing some FG when it would be optimal to abstain from it, is very small across several calibrations. The simple rule and OFG give divergent prescriptions if the disruption is low and the shock long-lived; but in these instances, the welfare cost of the simple rule is still small (even at high durations), and smaller than that of SIT. Most importantly the welfare cost of the simple rule stays small for larger shocks, while it is one order of magnitude more costly to stick to SIT. All in all, these results suggest that the simple rule is a good enough policy in this more realistic model; a precise metric for this requires a full-fledged Ramsey-optimal policy exercise in this large nonlinear model with a ZLB, an exercise beyond the scope of this paper that I pursue in ongoing research.

"Communication" has been a key word in describing the benefits of the simple rule that approximates the optimal FG commitment. To implement this rule in practice, it is therefore of the essence to communicate clearly its inputs T_L and Δ_L and how they will be measured. This raises empirical issues well beyond the scope of this paper, but we can suggest some possible avenues. The trap duration T_L's data counterpart corresponds to the duration of growth slowdown, defined for instance as output gap deviations with and without a financial crisis (along the lines of Schularick and Taylor 2012) or as deleveraging periods (along the lines of Koo 2008). The financial disruption that governs the depth of the trap Δ_L can be approximated by following recent estimates of the natural interest rate (e.g., Laubach and Williams 2015) or estimating a structural model where the spread can be estimated directly (à la Cúrdia and Woodford 2009, 2010). Either way, it suggests that more research is needed to inform central banks on what would then become two key policy inputs.

Appendix A: Optimal FG and a Markov Chain Representation

Under the Markov chain assumed in text, there are three states for the vector (ρ_t, i_t): $L = (\rho_L, 0)$, $F = (\rho, 0)$, and $S = (\rho, \rho)$, where the last is an absorbing state. The probability of transition into F starting from L is $q(1-p)$, and starting from F itself is q. The expected duration of FG is hence $(1-q)^{-1}$. The transition matrix is, for states L, F, S:

$$P = \begin{pmatrix} p & q(1-p) & (1-p)(1-q) \\ 0 & q & 1-q \\ 0 & 0 & 1 \end{pmatrix}.$$

Of particular interest is the sub-matrix corresponding to transient states $P_1 = \begin{pmatrix} p & q(1-p) \\ 0 & q \end{pmatrix}$; its powers t indicate the probability of being in a

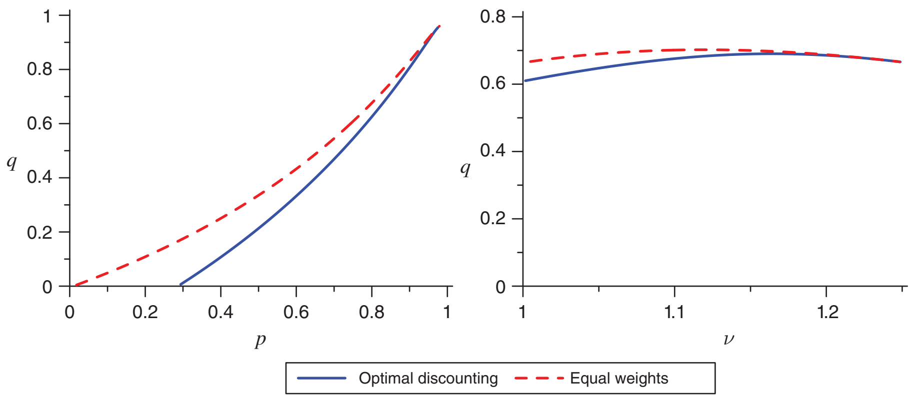

Figure A1

Note: Optimal FG persistence for equal weights (red dashed line) and optimal discounting (solid blue line), as a function of LT persistene (left panel) and news elasticity (right panel).

given state after t periods, starting from each transient state, respectively, $\begin{pmatrix} p^t & (1-p)q\frac{(p^t-q^t)}{p-q} \\ 0 & q^t \end{pmatrix}$. The *fundamental matrix* of this chain is denoted by N and defined as

$$N = (I - P_1)^{-1} = \begin{pmatrix} \frac{1}{1-p} & \frac{q}{1-q} \\ 0 & \frac{1}{1-q} \end{pmatrix}.$$

It gives the expected number of periods that the chain is in a state given an initial state, where the initial state is counted when staying in the same state. One can then define time to absorption, starting from each state, as $T = Ne = \left(\frac{1}{1-p} + \frac{q}{1-q}, \frac{1}{1-q}\right)'$, where $e \equiv (1,1)$. Note that both absorption *probabilities* are of course 1. To compute **expected discounted lifetime utility**, denote the time-invariant solution for utility in each state as $U = (U_L \; U_F \; U_S)$. Recall that once the absorbing steady state is reached, the flexible-price equilibrium is implemented, so the welfare function in the absorbing state is irrelevant. Therefore, we use $U_1 = (U_L \; U_F)$ to calculate lifetime welfare. Furthermore, the probability of being in state L after t periods is simply p^t, while for state F it is the sum of two conditional probabilities (starting from L and from F itself), namely $q^t + (1-p)q\frac{(p^t-q^t)}{p-q}$. We thus obtain expected present discounted welfare:

$$\frac{1}{1-\beta p} U_L + \left(\frac{1}{1-\beta q} + \frac{\beta(1-p)q}{(1-\beta p)(1-\beta q)}\right) U_F,$$

where the second term in brackets accounts for time spent in F when reaching F conditional upon starting from state L. Defining the relative weight $\omega(q)$ as in text (equation (12)), we can write welfare as $\frac{1}{1-\beta p}[U_L + \omega(q)\,U_F]$. Notice that $\omega'(q) = \beta\frac{1-\beta p + (1-p)}{(1-\beta q)^2} > 0$. The first-order condition becomes

$$\frac{\nu(1-p)^2}{(1-p\nu)^2}\frac{q\nu}{1-q\nu} + \frac{\nu(1-\beta p+\beta(1-p)q)}{(1-q\nu)(1-\beta q)} + \frac{1}{2}\beta\frac{1-\beta p+1-p}{(1-\beta q)^2} = \frac{\nu(1-p)}{(1-p\nu)^2}\Delta_L,$$

whose solution under the baseline parameterization is $q = 0.64$. An analytical solution of this (cubic) equation is not very easy to obtain nor very informative.[20] Figure A1 plots the implied optimal q as a function of p and ν for the otherwise baseline parameterization, along with the corresponding optimal q^* in the equal-weights case. The main difference occurs as a function of the trap persistence. With optimal discounting, since the welfare cost of FG now receives a larger weight, it is only optimal to do FG if things are bad enough, for high enough size and/or persistence of the trap. This illustrates our previous insight that q^* is an upper bound on optimal FG.[21]

Appendix B: Ramsey-Optimal Policy and Forward Guidance

This Appendix deals with the case of perfect foresight and the online Appendix covers Markov shocks. The solution to the Ramsey problem is described by the two conditions in (19); combining them to eliminate consumption c_t, we obtain

$$\left[-\nu E_t\phi_{t+1} + \left(1+\beta^{-1}\nu^2\right)\phi_t - \beta^{-1}\nu\phi_{t-1} + \sigma\rho_t\right]\phi_t \geq 0.$$

Call T the stopping time of the exogenous shock (under Markov shocks, this is a stochastic variable with expected value $(1-p)^{-1}$) and T_F^R the (unknown, to be determined) number of periods for which the ZLB binds, determined implicitly by the boundary condition $\phi_{T+T_F^R-1} = 0$. First notice that once ZLB stopped binding (for any $t \geq T + T_F^R$), the economy is back at steady state, $i_t = \rho$, $c_t = 0$, $\phi_t = 0$. Therefore, we only need to solve for ϕ_t when ZLB binds, for $t \leq T + T_F^R - 1$, a case in which we have the second-order difference equation

$$\nu E_t\phi_{t+1} - \left(1+\beta^{-1}\nu^2\right)\phi_t + \beta^{-1}\nu\phi_{t-1} = \sigma\rho_t,$$

or written with lag operators (recall $L^{-j}x_t = E_t x_{t+j}$):

$$\left[L^{-2} - \left(\nu^{-1}+\beta^{-1}\nu\right)L + \beta^{-1}\right]\phi_{t-1} = \sigma\nu^{-1}\rho_t.$$

[20] The proof of uniqueness in the equal-discounting case, combined with the observation that $\omega''(q) > 0$ ($\omega(q)$ is convex, suggests that there is a unique equilibrium to this problem satisfying the constraints that q is a probability and $q < 1/\nu$ (numerical simulations confirm this).

[21] Figure A1 in online Appendix A3 plots illustrate that the OFG duration, the Ramsey implied duration, and the equal-weights OFG duration are all very close to each other.

The eigenvalues being obvious $\beta^{-1}\nu > 1$ and $\nu^{-1} < 1$, we can solve by factorizing as $(L^{-1} - \beta^{-1}\nu)(L^{-1} - \nu^{-1})\phi_{t-1} = \sigma\nu^{-1}\rho_t$. This delivers the backward-forward solution

$$\phi_t = \nu^{-1}\phi_{t-1} - \sigma\beta\nu^{-2}E_t \sum_{j=0}^{T+T_F^R-1} (\beta\nu^{-1})^j \rho_{t+j}, \tag{B1}$$

where the forward summation goes on only as long as the solution applies, i.e., up to $T + T_F^R - 1$, where T_F^R is unknown. Consider first what happens *after the shock has been absorbed*, i.e., between the (possibly stochastic) T and $T + T_F^R - 1$, whereby $E_t\rho_{t+j} = \rho$. Solving for $\phi_{T+T_F^R-1}$, we obtain

$$\phi_{T+T_F^R-1} = \nu^{-T_F^R}\phi_{T-1} - \frac{\sigma\beta\nu^{-2}\rho}{1-\beta\nu^{-1}}\left(\frac{1-\nu^{-T_F^R}}{1-\nu^{-1}} - \beta\nu^{-1}\frac{1-(\beta\nu^{-2})^{T_F^R}}{1-\beta\nu^{-2}}\right),$$

which combined with the boundary condition (20) ($\phi_{T+T_F^R-1} = 0$) delivers the Lagrange multiplier in the moment of absorption:

$$\frac{1-\beta\nu^{-1}}{\sigma\beta\nu^{-2}\rho}\phi_{T-1} = \frac{\nu^{T_F^R}-1}{1-\nu^{-1}} - \beta\nu^{-1}\frac{\nu^{T_F^R}-(\beta\nu^{-1})^{T_F^R}}{1-\beta\nu^{-2}} \equiv \tilde{\phi}(T_F^R). \tag{B2}$$

This defines an increasing function $\tilde{\phi}(T_F^R)$ that is independent of the exogenous random stopping time T. The above solution from T onward applies under perfect foresight, where we called the stopping time $T_F^R = T^{*R}$. To solve the problem between 0 and T, note that we know with certainty that the discount factor will be ρ from T onward and $\rho_L < 0$ before, i.e.,

$$\begin{aligned}\phi_t &= \nu^{-1}\phi_{t-1} - \sigma\beta\nu^{-2}\sum_{j=t}^{T-1}(\beta\nu^{-1})^{j-t}\rho_L - \sigma\beta\nu^{-2}\sum_{j=T}^{T+T^{*R}-1}(\beta\nu^{-1})^{j-t}\rho \\ &= \nu^{-1}\phi_{t-1} - \sigma\beta\nu^{-2}\frac{1-(\beta\nu^{-1})^{T-t}}{1-\beta\nu^{-1}}\rho_L - \sigma\beta\nu^{-2}(\beta\nu^{-1})^{T-t}\frac{1-(\beta\nu^{-1})^{T^{*R}}}{1-\beta\nu^{-1}}\rho.\end{aligned}$$

Solving this backwards delivers $\phi_t = \nu^{-(t+1)}\phi_{-1} + \sum_{i=0}^{t}\nu^{-i}X_{t-i}$, with $X_t \equiv -\sigma\beta\nu^{-2}\rho\left(\frac{1-(\beta\nu^{-1})^{T-t}}{1-\beta\nu^{-1}}\Delta_L - (\beta\nu^{-1})^{T-t}\frac{1-(\beta\nu^{-1})^{T^{*R}}}{1-\beta\nu^{-1}}\right)$ and further,

$$\begin{aligned}\phi_t = \sigma\beta\nu^{-2}\rho\Bigg(&\frac{\Delta_L}{1-\beta\nu^{-1}}\sum_{i=0}^{t}\nu^{-i}\left(1-(\beta\nu^{-1})^{T-t+i}\right) \\ &- \frac{1-(\beta\nu^{-1})^{T^{*R}}}{1-\beta\nu^{-1}}\sum_{i=0}^{t}\nu^{-i}(\beta\nu^{-1})^{T-t+i}\Bigg).\end{aligned}$$

Evaluating at the last period of negative shock $t = T-1$, we obtain

$$\phi_{T-1} = \frac{\sigma\beta\nu^{-2}\rho}{1-\beta\nu^{-1}}\Bigg(\Delta_L\frac{1-\nu^{-T}}{1-\nu^{-1}} - \Delta_L\big(\beta\nu^{-1}\big)\frac{1-\big(\beta\nu^{-2}\big)^T}{1-\beta\nu^{-2}}$$

$$-\Big(1-\big(\beta\nu^{-1}\big)^{T^{*R}}\Big)\big(\beta\nu^{-1}\big)\frac{1-\big(\beta\nu^{-2}\big)^T}{1-\beta\nu^{-2}}\Bigg).$$

The optimal stopping time is thus a solution to

$$\Delta_L\left(\frac{1-\nu^{-T}}{1-\nu^{-1}} - \beta\nu^{-1}\frac{1-\big(\beta\nu^{-2}\big)^T}{1-\beta\nu^{-2}}\right)$$

$$= \Big(1-\big(\beta\nu^{-1}\big)^{T^{*R}}\Big)\beta\nu^{-1}\frac{1-\big(\beta\nu^{-2}\big)^T}{1-\beta\nu^{-2}} + \frac{\nu^{T^{*R}}-1}{1-\nu^{-1}} - \beta\nu^{-1}\frac{\nu^{T^{*R}}-\big(\beta\nu^{-1}\big)^{T^{*R}}}{1-\beta\nu^{-2}}.$$

Using straightforward algebra, multiplying this with ν^T delivers exactly the same equation as that determining OFG duration (17), thus proving the proposition.

REFERENCES

Adam, Klaus, and Roberto M. Billi. 2006. "Optimal Monetary Policy under Commitment with a Zero Bound on Nominal Interest Rates." *Journal of Money, Credit and Banking* 38 (7): 1877–1905.

Adam, Klaus, and Roberto M. Billi. 2007. "Discretionary Monetary Policy and the Zero Lower Bound on Nominal Interest Rates." *Journal of Monetary Economics* 54 (3): 728–52.

Adjemian, Stéphane, Houtan Bastani, Michel Juillard, Fréderic Karamé, Junior Maih, Ferhat Mihoubi, George Perendia, et al. 2011. "Dynare: Reference Manual Version 4." CEPREMAP Dynare Working Paper 1.

Andrade, Philippe, Gaetano Gaballo, Eric Mengus, and Benoît Mojon. Forthcoming. "Forward Guidance with Heterogenous Beliefs." *American Economic Journal: Macroeconomics.*

Bassetto, Marco. 2019. "Forward Guidance: Communication, Commitment, or Both?" Federal Reserve Bank of Chicago Working Paper 2019–05.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2001. "The Perils of Taylor Rules." *Journal of Economic Theory* 96 (1–2): 40–69.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2002. "Avoiding Liquidity Traps." *Journal of Political Economy* 110 (5): 535–63

Bianchi, Francesco, and Leonardo Melosi. 2017. "Escaping the Great Recession." *American Economic Review* 107 (4): 1030–58.

Bilbiie, Florin O. 2008. "Limited Asset Market Participation, Monetary Policy and (Inverted) Aggregate Demand Logic." *Journal of Economic Theory* 140 (1): 162–96.

Bilbiie, Florin O. 2014. "Delegating Optimal Monetary Policy Inertia." *Journal of Economic Dynamics and Control* 48: 63–78.

Bilbiie, Florin Ovidiu. 2018. "Monetary Policy and Heterogeneity: An Analytical Framework." Centre for Economic Policy Research (CEPR) Discussion Paper 12601.

Bilbiie, Florin O. Forthcoming. "The New Keynesian Cross." *Journal of Monetary Economics.*

Bilbiie, Florin O., Tommaso Monacelli, and Roberto Perotti. Forthcoming. "Is Government Spending at the Zero Lower Bound Desirable?" *American Economic Journal: Macroeconomics.*

Campbell, Jeffrey R., Charles L. Evans, Jonas D.M. Fisher, and Alejandro Justiniano. 2012. "Macroeconomic Effects of Federal Reserve Forward Guidance." *Brookings Papers on Economic Activity* 42 (1): 1–80.

Campbell, Jeffrey R., Jonas D. M. Fisher, Alejandro Justiniano, and Leonardo Melosi. 2016. "Forward Guidance and Macroeconomic Outcomes since the Financial Crisis." Federal Reserve Bank of Chicago Working Paper 2016-07.

Carlstrom, Charles T., Timothy S. Fuerst, and Matthias Paustian. 2015. "Inflation and Output in New Keynesian Models with a Transient Interest Rate Peg." *Journal of Monetary Economics* 76: 230–43.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1): 1–45.

Cúrdia, Vasco, and Michael Woodford. 2009. "Credit Frictions and Optimal Monetary Policy." Bank for International Settlements (BIS) Working Paper 278.

Cúrdia, Vasco, and Michael Woodford. 2010. "Credit Spreads and Monetary Policy." *Journal of Money, Credit and Banking* 42 (S1): 3–35.

Del Negro, Marco, Marc Giannoni, and Christina Patterson. 2012. "The Forward Guidance Puzzle." Federal Reserve Bank of New York Staff Report 574.

Eggertsson, Gauti B., and Michael Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity* 33 (1): 139–233.

Filardo, Andrew, and Boris Hofmann. 2014. "Forward Guidance at the Zero Lower Bound." *BIS Quarterly Review:* 37–53.

Fuhrer, Jeffrey C., and Brian F. Madigan. 1997. "Monetary Policy When Interest Rates Are Bounded at Zero." *Review of Economics and Statistics* 79 (4): 573–85.

Galí, J. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework.* Princeton: Princeton University Press.

García-Schmidt, Mariana, and Michael Woodford. 2019. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review* 109 (1): 86–120.

Hasui, Kohei, Tomohiro Sugo, and Yuki Teranishi. 2016. "Liquidity Trap and Optimal Monetary Policy Revisited." JSPS Grants-in-Aid for Scientific Research, Understanding Persistent Deflation in Japan Working Paper 079.

Jung, Taehun, Yuki Teranishi, and Tsutomu Watanabe. 2005. "Optimal Monetary Policy at the Zero-Interest-Rate Bound." *Journal of Money, Credit and Banking* 37 (5): 813–35.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2013. "Is There a Trade-Off between Inflation and Output Stabilization?" *American Economic Journal: Macroeconomics* 5 (2): 1–31.

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

Koo, Richard C. 2008. *The Holy Grail of Macroeconomics: Lessons from Japan's Great Recession.* Singapore: John Wiley and Sons.

Krugman, Paul R. 1998. "It's Baaack: Japan's Slump and the Return of the Liquidity Trap." *Brookings Papers on Economic Activity* 28 (2): 137–205.

Laubach, Thomas, and John C. Williams. 2015. "Measuring the Natural Rate of Interest Redux." Federal Reserve Bank of San Francisco Working Paper 2015-16.

Levin, Andrew, David López-Salido, Edward Nelson, and Tack Yun. 2010. "Limitations on the Effectiveness of Forward Guidance at the Zero Lower Bound." *International Journal of Central Banking* 6 (1): 143–89.

McKay, Alisdair, Emi Nakamura, and Jòn Steinsson. 2016. "The Power of Forward Guidance Revisited." *American Economic Review* 106 (10): 3133–58.

Melosi, Leonardo. 2017. "Signaling Effects of Monetary Policy." *Review of Economic Studies* 84 (2): 853–84.

Mertens, Karel R. S. M., and Morten O. Ravn. 2014. "Fiscal Policy in an Expectations-Driven Liquidity Trap." *Review of Economic Studies* 81 (4): 1637–67.

Nakata, Taisuke. 2016. "Optimal Fiscal and Monetary Policy with Occasionally Binding Zero Bound Constraints." *Journal of Economic Dynamics and Control* 73: 220–40.

Nakata, Taisuke. 2017. "Optimal Government Spending at the Zero Lower Bound: A Non-Ricardian Analysis." *Review of Economic Dynamics* 23: 150–69.

Nakata, Taisuke. 2018. "Reputation and Liquidity Traps." *Review of Economic Dynamics* 28: 252–68.

Nakata, Taisuke, and Sebastian Schmidt. Forthcoming. "Conservatism and Liquidity Traps." *Journal of Monetary Economics.*

Nakov, Anton. 2008. "Optimal and Simple Monetary Policy Rules with Zero Floor on the Nominal Interest Rate." *International Journal of Central Banking* 4 (2): 73–127.

Orphanides, Athanasios, and Volker Wieland. 1998. "Price Stability and Monetary Policy Effectiveness When Nominal Interest Rates Are Bounded at Zero." https://www.federalreserve.gov/pubs/feds/1998/199835/199835pap.pdf.

Rogoff, Kenneth. 1985. "The Optimal Degree of Commitment to an Intermediate Monetary Target." *Quarterly Journal of Economics* 100 (4): 1169–89.

Rogoff, Kenneth. 1998. "Comment on Paul Krugman's 'It's Baaack: Japan's Slump and the Return of the Liquidity Trap.'" *Brookings Papers on Economic Activity* 28 (2): 194–99.

Schmidt, Sebastian. 2013. "Optimal Monetary and Fiscal Policy with a Zero Bound on Nominal Interest Rates." *Journal of Money, Credit and Banking* 45 (7): 1335–50.

Schmitt-Grohé, S., and Martín Uribe. 2017. "Liquidity Traps and Jobless Recoveries." *American Economic Journal: Macroeconomics* 9 (1): 165–204.

Schularick, Moritz, and Alan M. Taylor. 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *American Economic Review* 102 (2): 1029–61.

Smets, Frank, and Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3): 586–606.

Svensson, Lars E. O., and Michael Woodford. 2004. "Implementing Optimal Policy through Inflation-Forecast Targeting." In *Inflation-Targeting Debate,* edited by Ben S. Bernanke and Michael Woodford. Chicago: University of Chicago Press.

Vestin, David. 2006. "Price-Level versus Inflation Targeting." *Journal of Monetary Economics* 53 (7): 1361–76.

Walsh, Carl E. 1995. "Optimal Contracts for Central Bankers." *American Economic Review* 85 (1): 150–67.

Walsh, Carl E. 2003. "Speed Limit Policies: The Output Gap and Optimal Monetary Policy." *American Economic Review* 93 (1): 265–78.

Walsh, Carl E. 2017. "Simple Sustainable Forward Guidance at the ELB." https://people.ucsc.edu/~walshc/MyPapers/2017Nov_SustainableForwardGuidance.pdf.

Werning, Iván. 2012. "Managing a Liquidity Trap: Monetary and Fiscal Policy." http://economics.mit.edu/files/7558.

Wiederholt, Mirko. 2015. "Empirical Properties of Inflation Expectations and the Zero Lower Bound." https://www.ecb.europa.eu/pub/conferences/shared/pdf/20151105_challenges/20151105Challenges-6-Wiederholt.pdf.

Williams, John C. 2013. "Will Unconventional Policy Be the New Normal?" FRBSF Economic Letter 2013-29.

Wolman, Alexander L. 1998. "Staggered Price Setting and the Zero Bound on Nominal Interest Rates." *Federal Reserve Bank of Richmond Economic Quarterly* 84 (4).

Woodford, Michael. 2003. *Interest & Prices: Foundations of a Theory of Monetary Policy.* Princeton: Princeton University Press.

Woodford, Michael. 2011. "Simple Analytics of the Government Expenditure Multiplier." *American Economic Journal: Macroeconomics* 3 (1): 1–35.

Woodford, Michael. 2012. "Methods of Policy Accommodation at the Interest-Rate Lower Bound." https://www.kansascityfed.org/publicat/sympos/2012/mw.pdf.