Output Hysteresis and Optimal Monetary Policy*

Vaishali Garga[a,1], Sanjay R. Singh[b,1]

[a] *Research Department, Federal Reserve Bank of Boston, MA 02210*
[b] *Department of Economics, University of California, Davis, CA 95616*

Abstract

We derive a fully quadratic approximation to welfare under endogenous growth and study optimal monetary policy. Away from the ZLB, optimal commitment policy sets interest rates to eliminate output hysteresis. A strict inflation targeting rule implements the optimal policy. At the ZLB, strict inflation targeting is sub-optimal and admits output hysteresis, defined as a permanent loss in potential output. A new policy rule that targets output hysteresis returns the output to the pre-shock trend and approximates the welfare gains under optimal commitment policy. A central bank unable to commit to future policy actions suffers from *hysteresis bias*: it does not offset past losses in potential output.

Keywords: Output Hysteresis, Optimal Monetary Policy, Zero Lower Bound

JEL: E31, E43, E52, O31, O43

*We owe invaluable debt to Gauti Eggertsson for continuous advice and extensive comments on various versions of this paper. We thank Alexei Abrahams, Francesco Bianchi, James Cloyne, Òscar Jordà, David Glancy, Mark Gertler, Violeta Gutkowski, David Hémous, Peter Howitt, Todd Keister, Mrityunjay Kothari, Pascal Michaillat, Neil Mehrotra, Elizabeth Murry, Matthias Paustian, Alan Taylor, Carl Walsh, David Weil, seminar participants at various institutions and conferences, Yuriy Gorodnichenko, and anonymous referees for helpful comments and suggestions. The views expressed herein are those of the authors and do not indicate concurrence by the Federal Reserve Bank of Boston or the Federal Reserve System. All errors are our own.

Email addresses: `vaishali.garga@bos.frb.org` (Vaishali Garga), `sjrsingh@ucdavis.edu` (Sanjay R. Singh)

1. Introduction

In the aftermath of the Great Recession, the US economy has experienced its slowest post-recession recovery since World War II. Twelve years in, the real GDP is still approximately 15 percent below its pre-recession trend level (Figure 1). One of the primary drivers of this output shortfall has been a slowdown in productivity growth. Decker et al. (2014) show that the recession accelerated the slowdown in startup entry, which is a significant channel for total factor productivity (TFP) growth. Similarly, investment in research and development (R&D), considered to be another important contributor to TFP growth, has fallen considerably during the last recession. These observations underscore concerns raised by several policymakers including Chair Yellen that episodes of slack in aggregate demand could affect the productive potential of an economy.[1]

The standard theoretical treatment of monetary policy is largely silent on the interaction of monetary policy with the economy's productive potential.[2] In this paper, we construct a model in which there is such an interaction. We embed a model of Schumpeterian growth, along the lines of Aghion and Howitt (1992) and Grossman and Helpman (1991), in a New Keynesian (NK) setting. A contraction in aggregate demand reduces the incentives for firms to invest in R&D, which leads to lower innovation. This results in an endogenous slowdown in TFP growth, which accumulates into a persistent output gap. Following a recession, unemployment returns to its natural rate while output remains below its pre-recession trend level. In this framework, monetary policy can affect the long-run potential output. This is in contrast to the traditional NK models which do not incorporate endogenous productivity

[1] Chair Janet Yellen (2015) noted that "... a portion of the relatively weak productivity growth... may be the result of the recession itself... In particular, investment in research and development has been relatively weak... Federal Reserve actions to strengthen the recovery may not only help bring our economy back to its productive potential, but it may also support the growth of productivity and living standards over the longer run."

[2] There is a recent synchronous literature that explores these interactions, including Anzoategui, Comin, Gertler and Martinez (2019), Bianchi, Kung and Morales (2019) and Benigno and Fornaro (2018). Ours is the first paper to analyze the interaction of optimal monetary policy at the ZLB, aggregate demand, and endogenous growth. We discuss this at length later in this section.

growth, and thus, incorrectly predict that output will recover to its pre-recession trend level.

Using this framework, we ask whether it is optimal for monetary policy to engineer a recovery back to the pre-recession trend level. Optimal policy analysis is the focus and main contribution of this paper. In order to analyze normative implications for the conduct of monetary policy, we derive a closed-form linear-quadratic approximation of the representative agent's lifetime utility function. This expression generalizes the approximation derived by Benigno and Woodford (2004) to the endogenous growth environment and nests exogenous growth as a special case. In particular, we decompose the stabilization objectives of the social planner into three key market distortions: a wage inflation gap, a labor efficiency gap and a productivity growth rate gap. Of these, the productivity growth rate gap is novel to the endogenous growth framework and provides an additional rationale for stabilization of short-run fluctuations.

We use this framework to study an economy hit with a temporary shortfall in demand. While our quadratic approximation is general, we focus the discussion on liquidity demand and monetary policy shocks because the model exhibits "divine coincidence" under these shocks. This coincidence implies that monetary policy can completely negate these shocks and maintain the economy at the first-best level. One implication of this property is that while the natural rate of interest, *r-star*, is exogenous, the level of potential output becomes an endogenous object. Hence, these two shocks allow us to tractably study monetary policy under endogenous growth. In this environment, we define *output hysteresis* as the gap between actual output and its initial deterministic trend level. We obtain the following three sets of results.

First, away from the ZLB, an optimizing policymaker with the ability to commit to future policy actions (optimal commitment policy) sets interest rates to offset the permanent output gap. A textbook prescription of the strict inflation targeting rule implements the optimal policy. Although the strict inflation targeting rule implements optimal policy away from the ZLB, it is unable to stabilize aggregate demand when the ZLB becomes a binding constraint.

As a result, the strict inflation targeting rule admits output hysteresis after a ZLB episode. On the other hand, policy rules exist which, if credibly communicated to the public, could prevent output hysteresis whether or not the ZLB is binding. One such rule is a *strict output hysteresis targeting rule*, whereby the central bank targets zero output hysteresis. This rule signals the central bank's ex-ante commitment to running a high-pressure economy in the future when there is no slack in employment. Thus, we find that output hysteresis is contingent on the monetary policy specification.

While the strict output hysteresis targeting rule can eliminate output hysteresis, it raises the question of whether it is desirable to run a high-pressure economy. Our second set of results speak to this concern. At the ZLB, the optimal policy response is to credibly commit to keeping future interest rates low in order to incentivize a recovery close to the pre-recession trend level. A strict output hysteresis targeting rule eliminates all the persistent effects resulting from constrained monetary policy, and closely replicates the welfare gains achieved under optimal commitment policy for a feasible range of parameters.

Third, and most importantly, we uncover a new dynamic inconsistency problem. A policymaker unable to commit to future policy actions (discretionary policy) does not find in its interest to undo permanent output gaps following a ZLB episode. This means that it is suboptimal for policy to be redesigned ex-post in order to offset the existing output hysteresis. We label this as the *hysteresis bias* of a discretionary policymaker. This dynamic inconsistency problem complements our first finding that hysteresis is a consequence of a central bank's policy constraints, in particular its inability to credibly commit to future policy actions, and not due to inept or irrational behavior on part of the central bank.

Our paper is closely related to the recent work of Anzoategui et al. (2019), Benigno and Fornaro (2018), Bianchi et al. (2019), Garcia-Macia (2015), Guerron-Quintana and Jinnai (2019), Moran and Queraltó (2018) and Queraltó (2019), all of whom integrate endogenous growth into a standard business cycle framework. Among these papers, our framework is most similar to that of Benigno and Fornaro (2018), who identify the possibility of an economy entering a phase characterized by a persistent liquidity trap and low TFP growth due to pessimistic expectations. We complement their elegant analysis by studying optimal monetary policy in response to shocks to economic fundamentals, while Benigno and Fornaro (2018) study the possibility that the economy is trapped in the ZLB equilibrium. To our best knowledge, ours is the first paper to analyze the desirability of admitting permanent output gaps in the presence of severe demand shortfalls, which is a particularly relevant consideration once the ZLB is binding. The analytical result on hysteresis bias is new to the literature and has important implications for central bank policy.[3]

We contribute to the optimal monetary policy literature by providing an analytically tractable generalization of the textbook optimal policy problem with nominal rigidities (Woodford 2003, Benigno and Woodford 2004). Recently, a number of papers have explored the implications for optimal monetary policy in a hysteresis-prone environment. Acharya et al. (2018) study an environment with permanent skill-loss resulting from temporary unemployment at the ZLB, while Galí (2016) works with an insider-outsider model of labor markets as in Blanchard and Summers (1986) —see also Erceg and Levin (2014), and Farmer (2012). In an endogenous TFP growth setting, away from the ZLB, Annicchiarico and Pelloni (2016) study Ramsey policy, and Ikeda and Kurozumi (2014) study the use of simple operational rules. We complement these various analyses by allowing contractions in demand to negatively affect long-run supply via endogenous productivity growth.

Our paper also adds to the Hansen/Summers secular stagnation literature (see also Eggertsson and Mehrotra 2015; Garga 2019). While we do not analyze permanent recessions, we formalize how demand-side and supply-side secular stagnation are related. In our setting, a temporary shock to r^* propagates through a slowdown in TFP growth to generate a permanent effect on the level of output. Our paper demonstrates that secular stagnation may be a consequence of policy constraints, in particular, the lack of central bank credibility.[4]

[3] Stadler (1990) and Fatas (2000) are important precursors to this recent literature.

[4] We refer the reader to Eggertsson and Egiev (2019) for a very detailed review of the fundamentals-driven liquidity trap literature.

2. A New Keynesian Model with Endogenous Growth

We integrate a textbook model of endogenous growth into a NK environment. There are six main agents in our model—households, wage unions, firms, entrepreneurs, the fiscal authority, and the central bank—described below.

Households and Wage Setting: Each of a continuum of monopolistically competitive households (indexed on the unit interval) supplies a differentiated labor service to the production sector. There is perfect consumption risk-sharing within the household. Household utility is derived from consuming a final consumption good, (disutility) from supplying labor, and from holding a risk-free bond:

$$\mathbb{E}_t \Sigma_{s=0}^{\infty} \beta^j \left[\log(C_{t+s}) - \frac{\omega}{1+\nu} \int_0^1 L_{t+s}(j)^{1+\nu} dj + \xi_t \frac{B_{t+1}}{P_t} \right],$$

where $\nu > 0$ is the inverse Frisch elasticity of labor supply, $\omega > 0$ is a parameter that pins down the steady-state level of hours, and $\beta \in (0,1)$ is the discount factor. ξ_t is a liquidity demand shock. It represents a "purely intertemporal" shock (Eggertsson 2008) which allows us to maintain *divine coincidence*. A central bank following optimal commitment policy does not face a trade-off in stabilizing output and inflation fluctuations arising from this shock.[5]

Labor income $W_t L_t$ is subsidized at a fixed rate τ^w. Households own an equal share of all firms, and receive Γ_t dividends from profits, pay taxes τ^b on their incomes from riskless bonds, and receive a lump-sum government transfer T_t. The household budget constraint in period t states that consumption expenditure plus asset accumulation must equal disposable income:

$$P_t C_t + B_{t+1} = (1-\tau^b) B_t (1+i_t) + (1+\tau^w) W_t L_t + \Gamma_t + T_t. \tag{1}$$

The stochastic discount factor by which financial markets discount nominal income in

[5] We assume that the household cannot issue any risk-free debt B_{t+1}. See also Cuba-Borda and Singh (2019) for references on bonds in utility.

period $t+1$ is given by $Q_{t,t+1} = \beta \frac{C_{t+1}^{-1}}{C_t^{-1}} \frac{P_t}{P_{t+1}}$. The household does not choose hours directly. Rather each type of worker is represented by a wage union that sets wages on a staggered basis. Consequently the household supplies labor at the posted wages as demanded by firms. Wage setting follows the modeling of Erceg et al. (2000). Perfectly competitive labor agencies combine j type labor services into a homogeneous labor composite L_t according to a Dixit-Stiglitz aggregator $L_t = \left[\int_0^1 L_t(j)^{\frac{1}{1+\lambda_{w,t}}} dj\right]^{1+\lambda_{w,t}}$, where $\lambda_{w,t} > 0$ is the (time-varying) nominal wage markup. Labor unions representing workers of type j set wages (with indexation) on a staggered basis following Calvo (1983), taking as given the demand for their specific labor input: $L_t(j) = \left(\frac{W_t(j)}{W_t}\right)^{-\frac{1+\lambda_{w,t}}{\lambda_{w,t}}} L_t$, where $W_t = \left[\int_0^1 W_t(j)^{\frac{-1}{\lambda_{w,t}}} dj\right]^{-\lambda_{w,t}}$. In particular, with probability $1-\theta_w$, the type-j union is allowed to re-optimize its wage contract and it chooses W_t^* to minimize the disutility of working for laborer of type j, taking into account the probability that it will not get to reset wage in the future. If a union is not allowed to optimize its wage rate, it indexes the wage to the steady state wage inflation rate, $\bar{\Pi}^w$. Workers supply whatever amount of labor is demanded at the posted wage. By the law of large numbers, the probability of resetting the nominal wage corresponds to the fraction of types who actually change their wage. Consequently, the nominal wage evolves as:

$$W_t^{\frac{-1}{\lambda_{w,t}}} = (1-\theta_w)\, W_t^{*\,\frac{-1}{\lambda_{w,t}}} + \theta_w (W_{t-1}\bar{\Pi}^w)^{\frac{-1}{\lambda_{w,t}}}. \tag{2}$$

Production: On the production side, we use a discrete time version of the Schumpeterian growth model (Aghion and Howitt, 2008, Ch. 4). The final consumption good is produced by perfectly competitive firms using a homogeneous labor composite supplied by the wage union and a CES composite of intermediate goods weighted by their productivity:[6] $Y_t^G = M_t^{1-\alpha} L_t^{1-\alpha} \int_0^1 A_{it}^{1-\alpha} x_{it}^{\alpha} di$, where each x_{it} is the flow of intermediate product i used at time t, the productivity parameter, A_{it} reflects the quality of that product, and M_t is the stationary

[6]We denote gross output by Y_t^G, to keep it distinct from Y_t (defined shortly after), which we refer to as the GDP analog of our model.

(aggregate) productivity shock. The firms choose L_t and $\{x_{it}\}_{i\in[0,1]}$ to maximize profits, taking as given both the wage index W_t and the prices of the intermediate goods $\{p_{it}\}_{i\in[0,1]}$.

There is a continuum of intermediate goods indexed by $i \in [0,1]$, each of which is produced by a sector-specific monopolist. The monopolist uses one unit of the final good to produce one unit of her own good. Each monopolist faces a marginal cost of P_t. Each intermediate monopolist sets prices flexibly to maximize her firm's profits, taking as given the final sector's demand for its product. In particular, she solves for

$$\max_{p_{it}} (1-\tau^p)p_{it}x_{it} - P_t x_{it} \quad s.t. \quad \frac{p_{it}}{P_t} = \alpha M_t^{1-\alpha} L_t^{1-\alpha} A_{it}^{1-\alpha} x_{it}^{\alpha-1}, \tag{3}$$

where τ^p is a sales tax/subsidy imposed on the monopoly price. Further, we assume that there is a competitive fringe in every sector which can produce the intermediate good with quality $\frac{A_{it}}{\gamma}$, where $\gamma > 1$ is the step-size of innovation and captures the quality distance between the frontier and laggard firms within a sector. As a result, the intermediate monopolist cannot charge a price higher than $p_{it} = \gamma^{1-\alpha} P_t$. In equilibrium, the monopolist charges a price given by $p_{it} = \zeta P_t \equiv \min\left(\gamma^{1-\alpha}, \frac{1}{(1-\tau^p)\alpha}\right) P_t$.[7] Note that the intermediate firm's profits are linear in the labor demanded by the final good's firm and its own productivity.[8] Higher own productivity enables the firm to capture a larger share of the demand for the final good. Profits are given by $\Gamma_t(A_{it}) = \chi^m P_t M_t L_t A_{it}$, where $\chi^m = ((1-\tau^p)\zeta - 1)\left(\frac{\alpha}{\zeta}\right)^{\frac{1}{1-\alpha}}$.[9]

R&D Entrepreneurs: There is a single entrepreneur in each sector who invests $RD(z_{it})A_{it}$ of final good in research and development in period t, where $RD' > 0$, $RD'' > 0$.[10] The dependence on productivity A_{it} is assumed for stationarity. With probability z_{it}, she is suc-

[7] In Schumpeterian models with non-drastic innovations, a limit pricing assumption of this form is commonly made. We refer the reader to appendix of Aghion, Akcigit and Howitt (2014) and appendix to chapter 7 of Barro and Sala-i Martin (2004) for detailed derivations. Later, we will assume away time-varying taxes or subsidies, and log-linearize the model around the efficient steady state. This *min* operator will then simplify into an equality. We thank an anonymous referee for suggesting that we clarify this point.

[8] Such linearity is central to various endogenous growth models (Jones, 2005).

[9] We are graetful to Fabian Winkler for pointing out a typo in specification of χ^m in an earlier draft.

[10] We follow Aghion et al. (2014) in this discrete time analog of their classic Schumpeterian model, but extend it to allow for a more general innovation production function that allows decreasing returns to R&D.

cessful in making a process improvement. The productivity in sector i goes up by a factor of $\gamma > 1$ (step-size of innovation) and she gets the monopoly rights (patent) over production of the intermediate good in the following period. If she fails to innovate, the incumbent monopolist continues to produce with productivity A_{it} until replaced by a successful entrant. Following Acemoglu and Akcigit (2012) and Benigno and Fornaro (2018), we further assume that the incumbent monopolist's patent may expire with an exogenous probability η.

Specifically, we assume that $RD(z_{it}) = \delta z_{it}^{\varrho}$, where $\delta > 0$ and and $\varrho > 1$ is the inverse elasticity of innovation intensity to R&D expenses. A research subsidy τ^r is provided by the government to the entrepreneur. The entrepreneur in every sector chooses z_{it} to maximize her expected discounted profits (from the patent):

$$\max_{z_{it}\in[0,1]} \{z_{it}\mathbb{E}_t Q_{t,t+1} V_{t+1}(\gamma A_{it}) - (1-\tau^r)P_t RD(z_{it})A_{it}\}, \tag{4}$$

where the value of the patent is given by $V_t = \Gamma_t + (1 - z_{it} - \eta)\mathbb{E}_t Q_{t,t+1} V_{t+1}$ and $z_{it} + \eta \leq 1$. The value function is linear in productivity (see Appendix A). Writing the normalized value function as $\tilde{V}_{it} \equiv \frac{V_{it}}{P_t A_{it}}$ and focusing on the symmetric equilibrium, we solve for the interior solution, where $z_t > 0$:

$$\varrho z_t^{\varrho-1} = \beta \mathbb{E}_t \left(\frac{C_{t+1}}{C_t}\right)^{-1} \frac{\gamma \tilde{V}_{t+1}}{(1-\tau_t^r)\delta}. \tag{5}$$

According to equation (5), the entrepreneur chooses the innovation intensity so that the discounted marginal revenue of an additional unit of innovation intensity is equal to the marginal cost of this unit. An increase in demand for the final good increases the value of obtaining the patent: for a given cross-sectional distribution of productivities, an increase in demand for the final good requires higher quantities of intermediate goods to fulfill that demand. Since a monopolist's profits are increasing in the quality of its product, she can capture a higher share of the increased market with a successful innovation.

Aggregation and Market Clearing: The aggregate behavior of the economy depends on the aggregate productivity index, defined as $A_t = \int_0^1 A_{it} di$. Because of the linear production function, we can aggregate the firm-level variables to form aggregate composites. Specifically, $RD_t = \int RD_{it}di$ is the total R&D expenditure and $X_t = \int X_{it}di$ is the aggregate intermediate good produced in the economy. We can rewrite the aggregate output and the nominal wage purely in the form of aggregates as well. The growth rate of output in the economy is equal to the growth rate of aggregate productivity $g_{t+1} = \frac{A_{t+1}-A_t}{A_t}$. In any period, innovations occur in z_t sectors, while $1 - z_t$ sectors use the previous period's production technology. Aggregating across all the sectors, we get the following equation governing the dynamics of aggregate productivity:

$$A_{t+1} = A_t + z_t(\gamma - 1)A_t \implies g_{t+1} = z_t\,(\gamma - 1)\,. \tag{6}$$

This means that the growth rate of the economy in period $t + 1$ is determined in period t and equals the number of innovating sectors multiplied by the step-size of innovation. The number of innovating sectors z_t may be interpreted as new entrants since the incumbent firms do not undertake R&D investment in our model. The final output produced in the economy is used for consumption, research, and the production of intermediate goods: $Y_t^G = C_t + RD_t + X_t$. Henceforth, we define $Y_t^G - X_t = (1 - \frac{\alpha}{\zeta})Y_t^G \equiv Y_t$ as GDP.

From equations (5) and (6), note that a percent change in innovation investment translates into $\frac{\bar{g}}{\varrho(1+\bar{g})}$ percent change in the gross productivity growth rate, where $\frac{1}{\varrho}$ is the elasticity of innovation intensity, and ϱ is assumed to be greater than 1 following the innovation literature (see Acemoglu and Akcigit 2012). The quantitative importance of endogenous growth depends on the value of the parameter ϱ.

Fiscal and Monetary Policy: To close the model, we assume a net zero supply of risk-free bonds: $B_t = 0$. The government's budget is balanced every period, so the total lump-sum transfers are equal to the sum of intermediate-good, labor, and research taxes: $P_tT_t = \tau^p \int_0^1 p_{it}x_{it}di + \tau^r P_t RD_t + \tau^w \int_0^1 W_t(h)L_t(h)dh$. We assume that an independent central bank sets the nominal interest rate on the risk-free government bonds. While solving for optimal monetary policy in the following section, we will often compare the equilibrium to the one obtained when the central bank follows a Taylor rule in setting the economy's nominal interest rate:

$$1 + i_t = \max\left(1, (1 + i_{ss})\left(\frac{\Pi_{W,t}}{\bar{\Pi}_W}\right)^{\phi_\pi}\left(\frac{L_t}{\bar{L}}\right)^{\phi_y}\varepsilon_t^i\right); \quad \phi_\pi > 1, \phi_y \geq 0, \tag{7}$$

where, ε_t^i represents a monetary policy shock. According to this Taylor rule, the nominal interest rate is set in order to target deviations of wage inflation and employment from their respective steady-state targets, as long as the implied nominal interest rate is non-negative.

2.1. Equilibrium

We formally define the economy's competitive equilibrium in Appendix A. In order to arrive at a stationary system of equations, we normalize the equilibrium equations by dividing the non-stationary variables such as consumption, output, and real wage, by the level of productivity. This allows us to solve for the balanced growth path (BGP) of the stationary competitive equilibrium. Given an initial level of TFP and the law of motion for TFP, we can recover the non-stationary equilibrium in which the non-stationary variables grow at a constant rate given by the BGP growth rate.

We find the BGP by imposing restrictions on the parameters such that the steady state satisfies a) $z \in (0, 1-\eta)$, b) consumption is non-negative, and c) the nominal interest rate is non-negative. In our numerical simulations, we verify that the innovation probability is bounded, that is, $z_t \in (0, 1-\eta)$.

Equilibrium Concepts and Policy Instruments

We define the efficient BGP as the one in which the welfare of the representative household is maximized subject to the production technology of the final consumption good, the law of motion for TFP, and the economy's resource constraint, for a given initial TFP level. The complete system of equations is provided in Appendix D.

The BGP of the competitive equilibrium allocation is inefficient due to three static distortions in our setup: (i) monopoly power in each intermediate goods sector, (ii) monopolistic competition in the labor market, and (iii) inter-temporal research externalities. While the first two distortions are common in the business cycle literature, the third distortion is specific to the endogenous growth model. On one hand, the entrepreneur is unable to reap all the benefits of her innovation because she gets replaced with positive probability by a new entrant or due to exogenous patent expiration. This makes her under-invest in R&D. On the other hand, an entrant replaces the incumbent in order to profit from the full step-size of the innovation, rather than the incremental gain in knowledge, so this incentivizes the entrepreneur to over-invest in R&D. Private R&D investment can be higher or lower than the efficient allocation on account of these two opposing forces. These steady-state distortions imply that in the absence of relevant fiscal instruments, monetary policy could affect the growth rate of output in the long-run.

Proposition 1 (BGP Efficiency)**.** *Assuming the policymaker has access to non-distortionary lump-sum taxes, the BGP of the competitive equilibrium can be made efficient using the following three taxes: a) a sales subsidy, $\tau^p = 1 - \frac{1}{\alpha}$, b) a labor tax, $\tau^w = \frac{1-\lambda_w}{\lambda_w}$, and c) a research tax, $\tau^r = 1 - \left[\left(\frac{\gamma l^* \frac{1-\alpha}{\alpha} \alpha^{\frac{\alpha}{1-\alpha}}}{1+g-\beta(1-z^*-\eta)}\right)\left(\frac{(1-\beta)(1+g^*)}{(\gamma-1)c^*}\right)\right]$; where terms with $*$ denote the efficient steady-state values.*

As shown by Woodford (2003) and Benigno and Woodford (2004), the linear-quadratic approximation to the social welfare function around the non-stochastic efficient steady state is justified if there are no distortions under price stability. We follow the monetary economics literature and make the following assumption in our analysis:

Assumption 1. *The fiscal authority provides the set of constant subsidies described in Proposition 1, such that the steady state of the stationary competitive equilibrium is efficient.*

This assumption implies that the average productivity growth rate is optimal and independent of monetary policy. The idea is to disassociate the welfare losses arising from fluctuations in the growth rate from those arising from suboptimal growth occurring solely due to monopoly distortions and research externalities. We log-linearize the stationary competitive equilibrium around the efficient steady state and define the resulting equilibrium as an approximate equilibrium (the formal definition is provided in the appendix). Henceforth, we assume that the (normalized) economy is in the efficient steady state at the beginning of time, $t = 0$.

In this economy, the first-best allocation is the competitive equilibrium allocation under flexible wages such that the fiscal authority utilizes (non-distortionary) time-varying taxes in order to maximize the representative agent's welfare. The natural-rate allocation (interchangeably the flexible-wage allocation) is the competitive equilibrium allocation under flexible wages such that the fiscal authority provides (non-distortionary) constant tax instruments, as outlined in Proposition 1. The sticky-wage allocation is the competitive equilibrium allocation under staggered (nominal) wages such that the fiscal authority provides (non-distortionary) constant tax instruments, as outlined in Proposition 1. We refer the reader to Appendix D.9.1, D.9.2, and D.9.3 for a formal definition of these equilibrium concepts.[11]

Under liquidity demand and monetary policy shocks, we obtain the following proposition:

Proposition 2. *The natural-rate allocation coincides with the first-best allocation under liquidity demand and monetary policy shocks in a stationary equilibrium.*

Proposition 2 implies that the representative agent's welfare is maximized if the policymaker can replicate the natural-rate allocation. Note that this outcome is possible if the policymaker has access to time-varying tax instruments (see, for example, Correia, Farhi, Nicolini

[11]When we refer to the natural-rate and the first-best allocations, we use the time-0 concept of flexible prices. According to this concept, the relevant state variable is the one under a counterfactual path where prices and wages had been flexible since the beginning of time. We defer a full discussion of time-0 and time-t flexible prices to Appendix D.9.

and Teles 2013). We assume that the policymaker does not have access to these time-varying fiscal instruments: the fiscal authority satisfies Assumption 1, and adjusts lump-sum taxes every period to balance the budget. The central bank sets the nominal interest rate i_t on the risk-free (nominal) bond B_t subject to the ZLB constraint:

$$i_t \geq 0 \quad \forall t. \tag{8}$$

The nominal interest rate is the central bank's only policy instrument.

Note that the *divine coincidence* property also implies that the natural rate of interest, r^*, is exogenous even in the presence of endogenous growth. This property helps isolate the role of monetary policy. Whether potential output is endogenous or not, depends on the precise definition of price/wage flexibility in the presence of a pre-determined state variable. We define potential output as time-t potential output, that is, the level of output that would occur if price and nominal wages are set flexibly in the current period and future periods, taking as given the evolution of the state variable (Woodford 2003, Ch. 5). We refer the reader to appendix D.9 for a formal discussion of alternate concepts of price-flexibility under endogenous growth.

Calibration: For illustration purposes, we calibrate the distorted steady state of the model with the parameters reported in Table 1, using quarterly time periods. There are nine parameters. We calibrate three parameters using values standard in the NK literature. The discount factor β equals 0.99. Preferences are logarithmic in consumption and the inverse Frisch elasticity ν is set at 2. The wage adjustment probability is set such that wages are reset once every four quarters and the BGP wage markup is 10%. We choose the three innovation parameters: step-size of innovation γ, (inverse of the) innovation elasticity ϱ, and the cost parameter in R&D investment δ to match (i) a 2% ratio of corporate R&D spending to GDP, a value often considered as the benchmark in endogenous growth models, (ii) a 3.6% per year average probability of an innovation in a given sector (Howitt, 2000), and (iii) a 2%

annual steady-state growth rate. The parameter η is set such that the (annual) probability of that the firm's patent will expire, $\eta + z$, is 15%, which is also the rate of depreciation of R&D stock estimated by the Bureau of Labor Statistics (see also Benigno and Fornaro 2018).As noted earlier, we solve the model around the efficient BGP that is consistent with these parameters. We set the labor share $1 - \alpha$ to 0.5 such that the growth rate in the efficient BGP is six times that in the distorted BGP, which is within the range of estimates of Jones and Williams (1998).[12] In Appendix C, we show show the impulse responses under the assumption of AR(1) process for shocks. These are similar to what the recent literature on endogenous growth in DSGE models has found. In the interest of space, we proceed to optimal policy analysis, which is the focus of this paper.

3. Optimal Monetary Policy

We derive a closed-form quadratic approximation of the household's utility function, and highlight three main results. One, away from the ZLB, optimal commitment policy does not involve permanent losses in output, and is implementable with the strict inflation targeting rule. Two, at the ZLB, optimal commitment policy returns the economy *close* to the pre-shock trend level by keeping interest rates lower in the future once the ZLB is no longer binding. Three, at the ZLB, optimal discretionary policy involves excessive output hysteresis relative to the optimal commitment policy. We label this as the *hysteresis bias* of the central bank. The central bank's lack of credibility tools is sufficient to generate output hysteresis. Numerically, we show that a novel strict output hysteresis targeting policy closely replicates optimal commitment policy, thereby implying significant welfare gains over optimal

[12] This additional parameter affects the present discounted value of owning a patent. Instead of the probability of survival being $1 - z$, it is $1 - z - \eta$. The probability of innovation success z may also be interpreted as the firm entry rate, consistent with the "creative destruction" literature. In alternate calibrations, we found that the results are not significantly altered. We experimented with matching z to an establishment turnover rate of 24%, or the job turnover rate of 32% from Business Dynamism Statistics (1977–2007). In another calibration, we fixed the innovation step-size to 1.06, following Acemoglu and Akcigit (2012).

discretionary policy. This is true for a range of values for the key parameter ϱ, which regulates the innovation sensitivity to R&D investment.

3.1. Quadratic Approximation of Welfare

One primary contribution of our paper is that we derive a quadratic approximation of the representative household's welfare under endogenous growth. This approximation generalizes the quadratic objective derived by Benigno and Woodford (2004) to an endogenous growth setting, and enables us to solve for the optimal policy in a tractable manner.

Proposition 3. *Assume that the economy is at the efficient steady state at time $t = 0$, with initial productivity level A_0. Under the sticky-wage allocation, the quadratic approximation to the representative agent's lifetime utility function $\mathbb{W}_0$ around the non-stochastic efficient steady state is given by*

$$\frac{\mathbb{W}_0 - \mathbb{W}_0^*}{U_{c_{ss}} y_{ss}} = -\frac{1}{2}\sum_{t=0}^{\infty}\beta^t \left[\lambda_y \underbrace{\left(\hat{y}_t - \frac{\beta}{1-\beta}\frac{1}{\nu+\frac{y}{c}}\hat{g}_{t+1}\right)^2}_{(i)} + \lambda_g \underbrace{\hat{g}_{t+1}^2}_{(ii)} + \lambda_\pi \underbrace{(\hat{\pi}_t^w)^2}_{(iii)}\right] + \mathcal{O}(||\hat{\xi}_t, \hat{\varepsilon}_t^i||^3) + t.i.p. \tag{9}$$

(i) : labor efficiency gap, (ii): productivity growth rate gap, and (iii): wage inflation gap,

where $\lambda_y = \left(\nu + \frac{y}{c}\right) > 0$, $\lambda_g = \frac{c}{y}\frac{\beta}{1-\beta}\left[\frac{\nu}{\nu+\frac{y}{c}}\frac{\beta}{1-\beta} + [(\varrho-1)\eta_g + 1]\right] > 0$, $\lambda_\pi = \frac{1+\lambda_w}{\lambda_w}\frac{1}{\kappa_w} > 0$, $\kappa_w \equiv \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+\nu\left(1+\frac{1}{\lambda_w}\right))} > 0$, $\eta_g = \frac{1+g}{g} > 1$, and t.i.p. stands for "terms independent of policy". $\mathbb{W}^*$ denotes welfare under the (time-0) first-best allocation. The approximation is scaled by the constant $U_{c_{ss}} y_{ss} = \frac{y_{ss}}{c_{ss}}$ (evaluated at the efficient steady state).

This approximation is composed of three gaps: (i) the labor efficiency gap, (ii) the productivity growth rate gap, and (iii) the wage inflation gap. These are the stabilization goals for a planner who wants to maximize social welfare.

The first and the third terms are standard in a textbook NK model (Galí (2015)). The first term, the labor efficiency gap, is the difference between the marginal product of labor and the marginal rate of substitution between consumption and leisure for the representative household; $(i) = mrs_t - mpn_t$, where these terms denote deviations from the respective steady-state values, and mpn_t corresponds to the (productivity-adjusted) real wage. This labor efficiency gap captures the time-varying wedge in the household's disutility from supplying labor at a pre-set nominal wage. The third term, the wage inflation gap, describes the loss in efficiency resulting from the dispersion in wages across members of the household. Under flexible wages, both the labor inefficiency gap and the wage inflation gap equal zero.

The second term, the productivity growth rate gap, is the new stabilization goal due to endogenous productivity growth. Current investment in R&D contributes to a persistent increase in the level of productivity. These inter-temporal spillovers of R&D investment may not be internalized by private agents and may result in too high or too low a response from this investment relative to the first-best allocation. Starting from a productivity level A_0, the growth rate gap in equation (9) captures the suboptimality of deviations from the first-best level of productivity given by $A_t^* = A_0(1+g_{ss})^t$ for all $t > 0$. Under nominal (wage) rigidities, demand shocks may induce this permanent output gap, thereby leaving the agent permanently worse off. This gap disappears under the exogenous growth assumption, and the quadratic approximation simplifies to the setting in the textbook treatment of Galí (2015, Ch. 4).

3.2. Optimal Policy Away from the Zero Lower Bound

Optimal monetary policy away from the ZLB involves setting the nominal interest rate in order to perfectly stabilize output and productivity along the first-best allocation.

Proposition 4 (Optimal Policy away from the ZLB)**.** *Given a process for liquidity demand and monetary policy shocks, optimal monetary policy under a sticky-wage allocation tracks the natural rate of interest when the ZLB constraint is slack.*

From Proposition 2, we know that the natural-rate allocation coincides with the first-best allocation. Under a sticky-wage allocation, setting the nominal interest rate to track the natural interest rate implements the natural-rate allocation, thereby replicating the first-best allocation.

Corollary 1. *When the ZLB is slack, the time series of output under optimal policy is a trend stationary process (integrated of order zero), that is,* $\log Y_t = a + b * t$*, where* $a = \log Y_0$ *is the initial level of output, and* $b = \log(1 + g_{ss})$ *is the steady-state productivity growth rate.*

Thus, away from the ZLB, permanent output gaps are undesirable in response to temporary demand shocks. Furthermore, we can derive the deviations in the levels of productivity and output under a standard Taylor rule (equation 7) from the respective natural-rate levels, assuming local determinacy, as follows:

$$\log A_t - \log A_t^e = \sum_{s=0}^{t-1} \psi_g^i \epsilon_s^i; \quad \log Y_t - \log Y_t^e = \hat{y}_t + \sum_{s=0}^{t-1} \psi_g^i \epsilon_s^i,$$

where $\psi_g^i > 0$ (the detailed expression is shown in Appendix C) and ϵ_t^i is the liquidity demand shock or the monetary policy shock at time t. We refer to the permanent deviation in output from the natural-rate benchmark as *output hysteresis* (or as permanent output gap). The following proposition generalizes the standard NK model result to an endogenous growth environment:

Proposition 5 (Output Hysteresis)**.** *Given the standard monetary policy rule (equation 7) and a slack ZLB constraint, transitory (modeled as AR(1) process) liquidity demand shocks or monetary policy shocks induce output hysteresis if and only if monetary policy does not follow a strict targeting rule, i.e.* $Y_T \neq Y_T^e \iff \{\phi_\pi, \phi_y > 0 : \phi_\pi \not\to \infty \cup \phi_y \not\to \infty\}$*, where* $1 < T < \infty$ *such that* $y_T = y$ *(steady state value) and* $y_T \equiv \frac{Y_T}{A_T}$ *is the normalized (or stochastically detrended) output.*

Permanent output gaps emerge as a consequence of the standard monetary policy specification assumed in equation (7). Normalized output exhibits a monotonic response to shocks which approaches zero as the shocks die out. The sum of the productivity growth rate deviations from the steady state cumulate to the output hysteresis, denoted henceforth by $h_t \equiv \sum_{s=1}^{t} \hat{g}_s = \hat{g}_t + h_{t-1}$. If the monetary policy follows a strict inflation targeting rule, this output hysteresis does not emerge. Setting the nominal interest rate so as to track the natural interest rate leads to perfect stabilization of the economy. However, it may not be possible for the central bank to implement this optimal policy due to a binding ZLB constraint. This inability to perfectly track the natural interest rate gives rise to permanent supply side deviations. This implication also formalizes the concept of *Inverse Say's Law* (Summers 2015).

3.3. Optimal Policy at the Zero Lower Bound

A policy rule that perfectly stabilizes the economy when the nominal interest rate is away from the ZLB may fail to do so when monetary policy is constrained by the ZLB. Output hysteresis can arise even with policies that are optimal away from the ZLB.

We follow Eggertsson and Woodford (2003) and assume a two-state Markov chain for the natural interest rate ($\hat{r}_t^n$).[13] The economy unexpectedly hits the ZLB in period 1; that is, the nominal interest rate consistent with target inflation breaches a policy lower bound constraint, $r_t^n < i^{LB}$ (assume $i^{LB} = 0$): $\hat{r}_t^n = \hat{r}_S < 0 \;\forall\; 1 \leq t < T^e$ (Assumption A1a). With probability μ the economy continues to stay in the low state, and with $1-\mu$ probability the shock returns to the absorbing target-inflation steady state. We assume that the economy is back at this steady state after a stochastic but finite time $T^e < \infty$: $\hat{r}_t^n = (1-\beta) > 0 \;\forall\; t \geq T^e$ (Assumption A1b).

Further, we assume restrictions on parameters such that the equilibrium is locally determinate around the deflation steady state (Assumption A2). We calibrate the expected duration of the ZLB as $\frac{1}{1-\mu} = 3.7$ quarters or about 11 months (following Swanson and Williams 2014), and the natural interest rate at $\hat{r}_S = -1.43\%$ (annual). This calibration is

[13] In the notation of our framework, $\hat{r}_t^n = -\xi_t + (1-\beta)$. $\xi > 1-\beta$ makes the ZLB binding.

chosen to target a 7.5% drop in (normalized) output and 1% drop in nominal wage inflation relative to the target steady state in order to replicate the average drop in output and inflation during the Great Recession (following Eggertsson et al. 2020). The central bank is assumed to follow the strict inflation targeting rule.

Proposition 6 (Output Hysteresis at the ZLB)**.** *Under the strict inflation targeting rule ($\phi_\pi \to \infty$ in equation 7), a positive shock to liquidity demand or a contractionary monetary policy shock, such that the ZLB is binding for a finite time T^e, results in output hysteresis.*

When the ZLB is binding ($t < T^e$), there is wage deflation and low output along the equilibrium path, and when the ZLB is no longer binding ($t \geq T^e$), the central bank raises the nominal interest rate to the steady-state level. While employment and wage inflation return to their natural-rate levels, the economy's productive potential is permanently lower relative to the counterfactual path in which the ZLB is not binding. Such losses in potential output can be sizable for reasonable durations of a ZLB recession.

Should monetary policy offset these hysteresis effects at the ZLB? To provide an answer, we derive the optimal monetary policy at the ZLB under two regimes.

Optimal Policy under Commitment

We first solve the optimal commitment policy; that is, optimal policy when the central bank can credibly commit to future state-contingent policy actions. We describe the commitment problem and its solution in Appendix E.1. Since the solution to this optimal policy problem does not have a closed-form expression, we solve it numerically for each state-contingent realization of the shock using a shooting algorithm outlined in Eggertsson and Woodford (2003).[14]

The solid red line in Figure 2 shows the optimal commitment equilibrium output, inflation rate, TFP growth rate, and the nominal interest rate under a realization of the shock with

[14] Extensive documentation of this stochastic algorithm is also available in Eggertsson, Egiev, Lin, Platzer and Riva (2020).

ZLB binding for 28 quarters. Under optimal policy, the central bank minimizes total losses in welfare by trading welfare losses during the ZLB against the welfare losses that arise after the ZLB stops binding. By committing to keeping interest rates low upon exit from the ZLB, the central bank creates anticipation of a boom, which lowers the real interest rate during the ZLB episode. Compared to the equilibrium under the strict inflation targeting rule (solid blue line with crosses), optimal commitment policy reduces the on-impact effect of the shock (the drop in wage inflation and output are only 0.09 percent and 3.11 percent, respectively). Upon exit from the ZLB, the central bank keeps the interest rate lower for three additional quarters to follow through with its promise and thus creates a boom in (normalized) output and inflation. Because of procyclicality of investment in innovation, the TFP growth rate overshoots its steady-state level, thereby returning output *close* to its pre-recession trend level (the output hysteresis is only -0.74 percent compared to -4.05 percent under the strict inflation targeting rule).

Under optimal commitment, the policymaker trades off positive output hysteresis against higher wage dispersion inefficiency upon exiting from the ZLB. In other words, the ZLB introduces a short-run versus long-run trade-off for the central bank, even when there are no initial steady state distortions. If the policy objective puts a higher weight on growth rate stabilization relative to the "true" welfare weight, λ_g in equation (9), output hysteresis can be fully eliminated (as shown in row 3 of Table 2), but this comes at the expense of a commitment to accommodating even higher inflation upon exit from the ZLB.

Optimal Policy under Discretion: Hysteresis Bias

We now analyze optimal monetary policy when the policymaker is unable to (ex-ante) commit to future state-contingent policy actions. Such a policy equilibrium is referred to as discretionary, time-consistent, or Markov Perfect Equilibrium (MPE, defined in Maskin and Tirole 2001). We describe the problem formally in Appendix E.1, and obtain the following proposition:

Proposition 7 (Optimal Discretionary Policy at the ZLB)**.** *Given Assumptions A1 and A2 and an initial level of productivity* A_0*, the MPE is characterized by:*

$$\text{for } 1 < t < T^e, \quad \hat{y}_t = \psi_y r_S^n < 0;\ \hat{\pi}_t^w = \psi_p r_S^n < 0;\ \hat{g}_t = \psi_g r_S^n < 0;\ \log A_{t+1} = \log A_t + \psi_g r_S^n$$

$$\text{and for } t \geq T^e, \quad \hat{y}_t = \hat{\pi}_t^w = \hat{g}_t = 0; \log A_{t+1} = \log A_{t+1}^* + (T^e - 1)\psi_g r_S^n < \log A_{t+1}^*$$

where $\psi_y = \frac{(1-\beta\mu)\eta_C^{-1}}{(1-\beta\mu)(1-\mu)-\kappa_w(\nu+\eta_C)\mu\eta_C^{-1}} > 0$*,* $\psi_p = \frac{\kappa_w(\nu+\eta_C)}{1-\mu\beta}\psi_y > 0$*, and* $\psi_g = \frac{1-\frac{c}{y}\eta_C}{\frac{\mathbb{R}}{y}\varrho\eta_g}\psi_y > 0$*.* A_{t+1}^* *is the (time-0) first-best output at time* $t+1$*; and* $\log A_1 = \log A_0 + \log(1+g_{ss})$*.*

Under MPE, the path of the interest rate is such that the economy returns to the (normalized) steady state as soon as the shock abates (at time T^e). Since the central bank cannot credibly promise to maintain low interest rates in the future (after time T^e), the ZLB period exhibits excessive deflation and below-potential output relative to the optimal commitment equilibrium. This result was identified as *deflation bias* of optimal discretionary monetary policy by Eggertsson (2006). We identify a new dynamic inconsistency result in the endogenous growth setup. After the ZLB episode ends, the policymaker does not offset the difference in the level of productivity from the first-best allocation. MPE, thus, admits a unit root in the time-series of productivity and hence, output. We label this result as the *hysteresis bias* of optimal discretionary monetary policy, which is novel to our framework.

The hysteresis bias emerges despite the level of productivity being an endogenous state variable. The efficiency of resource allocation in the normalized economy is independent of the level of productivity. As soon as the central bank is able to set the normalized variables to their steady-state values, it does so. Past deviations of growth rate enter the welfare-loss as additive inefficiencies that do not influence the decisions of the policymaker optimizing at time $t \geq T^e$.

The hysteresis bias strengthens the result from Proposition 6 that output hysteresis is an artifact of policy constraints faced by the central bank and does not arise because of irrational or inept behavior on part of the central bank. An absence of commitment credibility generates a permanent output shortfall. If the central bank could credibly commit to being irresponsible, à la Krugman (1998), it could not only reduce the deflation experienced during ZLB periods, but also minimize the permanent output gap.[15] This raises the stakes for optimal commitment policy: the central bank must credibly communicate this policy to the public ex-ante.

Comparison with Policy under Exogenous Growth

How does optimal commitment policy compare to its counterpart in the textbook exogenous growth environment?[16] Figure 3 compares the evolution of the nominal interest rate, output, and wage inflation under endogenous growth (solid blue line with crosses) versus exogenous growth (dashed red line). The optimal policy under exogenous growth does not allow the central bank to accommodate as high an inflation rate after a ZLB episode as the optimal policy under endogenous growth allows. Overall, the paths of economic variables are similar across the two scenarios. This is because the key problem in the endogenous growth environment, as in the exogenous growth environment, is deficient aggregate demand. Since R&D investment is pro-cyclical under liquidity demand shocks, stabilizing inflation stabilizes aggregate output and hence R&D investment. The main implication of this analysis is that while the optimal commitment policy prescription under endogenous growth may not be significantly different from the exogenous growth environment, the cost of not adhering to optimal commitment rules is elevated because of the possibility of permanent output gaps.

Alternative Policy Rules at the ZLB

Eggertsson and Woodford (2003) have underscored the complex nature of the optimal com-

[15] A central bank can use commitment tools in an environment where Modigliani-Miller theorem breaks down. We refer the reader to Eggertsson (2006) and Bhattarai, Eggertsson and Gafarov (2019) for examples of modeling commitment policy tools. Studying implications for the use of unconventional policy in the hysteresis environment is an important agenda for future research.

[16] Another relevant comparison to consider is with a policymaker who does not internalize that she can influence the productivity growth rate in an endogenous growth environment. We refer the interested reader to Garga and Singh (2019) for this comparison.

mitment policy: it may not be feasible to properly communicate the policy stance to the public even if full credibility can be achieved. The optimal discretionary policy, on the other hand, suffers from hysteresis bias as it does not offset past inefficiencies. In this regard, alternate policy rules that have a built-in commitment to reverse past policy mistakes assume importance. We discuss two such rules in this section.

The first rule is the *output hysteresis targeting* rule, where the central bank targets the history of productivity growth rate deviations resulting from current and past demand shocks. Specifically, this hysteresis-augmented Taylor rule incorporates an additional target of the cumulative sum of all deviations in productivity growth rate resulting from the history of shocks until time t: $\hat{i}_t = \max\left(-\frac{\bar{i}}{1+\bar{i}}, \phi_\pi \hat{\pi}_t^w + \phi_y \hat{L}_t + \phi_h h_{t+1} + \hat{\varepsilon}_t^i\right)$, where $h_{t+1} \equiv \sum_{s=1}^{t+1} \hat{g}_s = 0$. When $\phi_h \to \infty$, we label the rule as the *strict output hysteresis targeting* (SOHT) rule. The second rule is the nominal wage level targeting (NWLT) rule, where the central bank ex-ante announces its intention to set interest rates in order to attain a particular level w^* for the normalized output (y_t) adjusted nominal wages w_t^n: $w_t^n + \lambda y_t = w^*$; where $\lambda \equiv \frac{1+\lambda_w}{\lambda_w}$.

Figure 4 plots the paths of nominal interest rate, output, and wage inflation under the SOHT (dashed-dotted red line with circles) and NWLT (dash-dotted green line with stars) rules against the optimal commitment policy (solid blue line with crosses) and optimal discretionary policy (solid red line) rules for a 28-quarter realization of the ZLB from the assumed two-state Markov chain. As with optimal commitment, these rules prescribe a lower-for-longer interest rate path (relative to optimal discretion). Consequently, the central bank is willing to accommodate higher wage inflation upon exit from the ZLB.[17] The in-built forward guidance in these rules, through higher expected inflation, leads to a reduction in the real interest rate during the ZLB. This explains a lower drop in inflation and normalized output on impact.

In Table 2, we compare the permanent output gaps and the relative welfare losses obtained under various rules. Welfare loss is reported as a percentage of the consumption equivalent

[17]Away from the ZLB, both NWLT and SOHT implement optimal commitment policy.

welfare loss under discretionary policy.[18] We also display numerical results obtained under a nominal GDP targeting rule. Both NWLT and SOHT rules imply significant welfare gains and smaller permanent output gaps relative to the MPE. In fact, the SOHT rule, by definition, completely eliminates the permanent output gap, thereby closely replicating the (relative) welfare gains achieved under the optimal commitment policy. Under the NWLT rule, there is a permanent output gap of -2.26 percent. Compared with the NWLT rule, the SOHT rule requires the central bank to be more tolerant of higher wage inflation upon exiting from the ZLB.

We believe that the SOHT rule may offer an advantage in communication over the NWLT rule. A central bank's commitment to keeping the interest rate lower until output has been restored to the pre-shock trend level is arguably more readily observable by the public, assuming that achieving credibility is not a constraint for the central bank. Such a policy of hysteresis targeting is equivalent to a real GDP targeting rule, since in our model, $\log Y_t - \log Y_t^e = h_t$. However, the SOHT rule comes with an operational shortcoming. Hysteresis targeting requires knowledge of the counterfactual output trend that would obtain had nominal rigidities been absent since the economy began (that is, time-0 potential output). This is because liquidity demand shocks that push the economy to the ZLB do not affect the time-0 potential output in our model (see Proposition 2). Commonly used real-time estimates of potential output based on statistical filters, however, do not provide a reliable estimate of the time-0 potential output. In Appendix E.4, we show that these real-time estimates correspond more closely to the time-t potential output rather than the time-0 potential output in our model. One way around this practical problem comes from Coibion, Gorodnichenko and Ulate (2018), in which the authors estimate the economy's output gap assuming that only supply shocks affect output in the long-run. They estimate this output gap to be at least seven percentage points in 2017:Q1 (relative to 2007:Q1). The output gap,

[18] In the baseline calibration, consumption equivalent welfare loss under discretion is equal to 0.0048% of steady state consumption.

measured in this way, provides a measure of the gap between the time-0 potential output and the actual output, which is needed to make the SOHT rule operational in our model.

Table 3 shows a similar comparison of various policy rules against the optimal commitment policy for a range of numbers for the innovation elasticity, measured as the inverse of parameter ϱ. For various values of ϱ, we vary δ (the R&D cost parameter) to maintain a fixed efficient-BGP growth rate of 12%. Furthermore, we recalibrate the probability of escape from ZLB $(1-\mu)$ and the natural rate of interest $r_S^n < 0$ so as to keep (normalized) output drop and inflation drop under discretion fixed at -7.5 percent and -1 percent, respectively. Keeping the output drop and inflation drop fixed under discretion is useful to compare policies under different values of ϱ. There are two key takeaways from this table.

One, various inertial rules offer welfare gains over discretionary policy for a wide range of parameters.[19] The SOHT rule consistently approximates the (relative) welfare gains achieved under optimal commitment policy. The NWLT rule, with a built-in commitment to keep interest rates lower for longer, like the SOHT rule, also allows stabilization of the economy at the ZLB. These results are consistent with the literature, on policy rules with mismeasurement in potential output, that also finds superior performance of price-level targeting rules (see for e.g. Gorodnichenko and Shapiro, 2007).[20]

Two, the quantitative magnitude of the output hysteresis depends on the elasticity of the innovation intensity. A lower value for ϱ allows the model to generate large changes in the productivity growth rate and hence the level of GDP. The innovation literature (see for e.g. Acemoglu and Akcigit, 2012) often considers estimates for ϱ over a relatively wide range $\in (1.3, 10)$. In the business cycle literature, Anzoategui et al. (2019) estimate $\varrho \in (2.50, 2.90)$.

[19] There need not be a monotonic relationship between (relative) welfare loss and output hysteresis because of an additional penalty due to inflation variability across policy rules. We can see this by comparing rows 2 and 3 of Table 2. If the policy objective puts a higher weight on growth rate stabilization relative to the "true" welfare weight, λ_g in equation (9), output hysteresis can be fully eliminated, but this comes at the expense of a commitment to accommodating even higher inflation upon exit from the ZLB and hence higher relative welfare loss. We thank an anonymous referee for advising us to clarify this important point.

[20] In Appendix H, we analyze optimal policy away from the ZLB in response to discount rate, stationary TFP, and wage markup shocks. There too we find that the NWLT rule improves welfare compared to the strict inflation targeting rule consistently across the variety of shocks considered.

Estimates of cyclical sensitivity of TFP, R&D investment, and firm entry, respectively, can be used to infer bounds on ϱ. Our own assessment, from comparing estimated impulse response functions to identified monetary policy shocks in the data to those in a medium-scale DSGE model, is that estimates of ϱ lie in the range between one and three depending on the interpretation applied to the creative-destruction mechanism. In the interest of space and given the focus of this paper, we briefly discuss some details of our quantitative exercise next and relegate a more formal discussion to Appendix I.

Using Jordà (2005) local projections and external instruments, we find that following a contractionary monetary policy shock, scaled to generate a 100 basis points increase in federal funds rate on impact, the utilization-adjusted TFP falls by 0.6% three years out.[21] Corporate R&D investment, measured from the Compustat data, is not responsive enough to explain this estimated TFP response. A calibrated medium-scale DSGE model suggests that an estimate of 3 for ϱ is consistent with the response of corporate R&D, while an estimate closer to 1 seems to better fit the response of TFP. Responsiveness of firm entry indicators (new incorporations/net establishment births) suggests that creative destruction from firm entry may reconcile this low estimate of ϱ consistent with the estimated TFP response. Our short exercise implies that future research using richer firm-dynamic models, that incorporate firm entry and R&D investment as distinct drivers of TFP growth, is needed to quantify the contribution of various forces of innovation in explaining the estimated TFP response (see Bilbiie, Ghironi and Melitz, 2008; Mehrotra and Sergeyev, 2020).

[21] We use narrative monetary policy surprises from Wieland and Yang (2016) based on Romer and Romer (2004)'s methodology and high-frequency surprises from Gorodnichenko and Weber (2016). Several recent papers have emphasized similar results. Jordà, Singh and Taylor (2020) use panel data for seventeen advanced countries over 1890–2015 to provide causal evidence for the persistent effects of monetary policy. Moran and Queraltó (2018) also provide empirical evidence in support of the endogenous TFP growth mechanism for such persistent effects. Ridder (2017) finds evidence of contraction in R&D during the Great Recession. Meier and Reinelt (2019) emphasize a markup dispersion channel to explain the estimated TFP response to monetary policy shocks.

4. Conclusion

This paper solves for optimal monetary policy in a NK model with endogenous Schumpterian growth. We formalize a new dynamic inconsistency result whereby output hysteresis is a consequence of the central bank's lack of commitment credibility tools. Studying unconventional policy that can alleviate such commitment concerns in hysteresis-prone environments is a promising agenda for future research.

While our main analysis carves out a role for monetary policy, output hysteresis can also be avoided with the use of appropriate fiscal policy tools as argued by DeLong and Summers (2012) and Fatás and Summers (2015). Through the lens of our model, there are two main implications regarding the design of fiscal policy. One, as we show in Appendix D.6, a policymaker with access to a set of time-varying tax instruments can replicate the first-best allocation, thereby fully stabilizing the economy even at the ZLB. We had assumed away the use of such policy instruments in our analysis of optimal monetary policy. Two, as we show in Appendix F, timely, temporary, and targeted fiscal policy interventions in the form of R&D investment credits that are implemented during a ZLB episode are expansionary in the short run as they increase employment and inflation, as well as in the long run as they permanently increase the level of output. These results suggest that besides high fiscal multipliers at the ZLB, fiscal stimulus can have persistent effects on living standards. In this paper, we focused exclusively on studying output hysteresis in liquidity trap episodes driven by adverse fundamentals. Analyzing stabilization policies that are robust across fundamentals-driven as well as expectations-driven liquidity traps is an important topic for future research.

References

Acemoglu, D., Akcigit, U., 2012. Intellectual property rights policy, competition and innovation. Journal of the European Economic Association 10, 1–42.

Acharya, S., Bengui, J., Dogra, K., Wee, S., 2018. Slow recoveries and unemployment traps: Monetary policy in a time of hysteresis. FRBNY Staff Report .

Aghion, P., Akcigit, U., Howitt, P., 2014. The schumpeterian model. URL: http://www.brown.edu/Departments/Economics/Faculty/Peter_Howitt/2070-2015/Ch5-Schumpeter_140521.pdf. lecture Notes, Chapter 5.

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica 60.

Aghion, P., Howitt, P.W., 2008. The Economics of Growth. MIT Press.

Annicchiarico, B., Pelloni, A., 2016. Innovation, growth and optimal monetary policy .

Anzoategui, D., Comin, D., Gertler, M., Martinez, J., 2019. Endogenous technology adoption and r&d as sources of business cycle persistence. American Economic Journal: Macroeconomics .

Barro, R., Sala-i Martin, X., 2004. Economic Growth. The MIT Press, MIT Press.

Benigno, G., Fornaro, L., 2018. Stagnation traps. The Review of Economic Studies 85, 1425–1470.

Benigno, P., Woodford, M., 2004. Optimal monetary and fiscal policy: A linear-quadratic approach, in: NBER Macroeconomics Annual 2003, Volume 18. The MIT Press, pp. 271–364.

Bhattarai, S., Eggertsson, G.B., Gafarov, B., 2019. Time consistency and the duration of government debt: A signalling theory of quantitative easing. Technical Report.

Bianchi, F., Kung, H., Morales, G., 2019. Growth, slowdowns, and recoveries. Journal of Monetary Economics .

Bilbiie, F.O., Ghironi, F., Melitz, M.J., 2008. Monetary policy and business cycles with endogenous entry and product variety, in: NBER Macroeconomics Annual 2007, Volume 22, pp. 299–353.

Blanchard, O.J., Summers, L.H., 1986. Hysteresis and the european unemployment problem, in: NBER Macroeconomics Annual 1986, Volume 1. MIT Press, pp. 15–90.

Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. JME .

Coibion, O., Gorodnichenko, Y., Ulate, M., 2018. The cyclical sensitivity in estimates of potential output. Brookings Papers on Economic Activity , 343.

Correia, I., Farhi, E., Nicolini, J.P., Teles, P., 2013. Unconventional fiscal policy at the zero bound. American Economic Review 103, 1172–1211.

Cuba-Borda, P., Singh, S.R., 2019. Understanding persistent stagnation. FRB IFDP 1243 .

Decker, R., Haltiwanger, J., Jarmin, R.S., Miranda, J., 2014. The secular decline in business dynamism in the us .

DeLong, B.J., Summers, L.H., 2012. Fiscal policy in a depressed economy. Brookings Papers on Economic Activity 1, 233–297.

Eggertsson, G.B., 2006. The deflation bias and committing to being irresponsible. Journal of Money, Credit and Banking 38, 283–321.

Eggertsson, G.B., 2008. Great expectations and the end of the depression. American Economic Review 98, 1476–1516.

Eggertsson, G.B., Egiev, S.K., 2019. Fundamental driven liquidity traps: A unified theory of the great depression and the great recession .

Eggertsson, G.B., Egiev, S.K., Lin, A., Platzer, J., Riva, L., 2020. A Toolkit for Solving Models with a Lower Bound on Interest Rates of Stochastic Duration. Technical Report. Brown University.

Eggertsson, G.B., Mehrotra, N.R., 2015. A model of secular stagnation .

Eggertsson, G.B., Woodford, M., 2003. Zero bound on interest rates and optimal monetary policy. Brookings Papers on Economic Activity 2003, 139–233.

Erceg, C.J., Henderson, D.W., Levin, A.T., 2000. Optimal monetary policy with staggered wage and price contracts. Journal of Monetary Economics 46, 281–313.

Erceg, C.J., Levin, A.T., 2014. Labor force participation and monetary policy in the wake of the great recession. Journal of Money, Credit and Banking 46, 3–49.

Farmer, R.E., 2012. Confidence, crashes and animal spirits. The Economic Journal 122, 155–172.

Fatas, A., 2000. Do business cycles cast long shadows? short-run persistence and economic growth. Journal of Economic Growth 5, 147–162.

Fatás, A., Summers, L.H., 2015. The permanent effects of fiscal consolidations .

Galí, J., 2015. Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications. Princeton University Press.

Galí, J., 2016. Hysteresis and the european unemployment problem revisited .

Garcia-Macia, D., 2015. The financing of ideas and the great deviation.

Garga, V., 2019. Fiscal Expansions in Secular Stagnation: What if it isn't Secular? Technical Report. Brown University.

Garga, V., Singh, S.R., 2019. Output hysteresis and optimal monetary policy. Federal Reserve Bank of Boston Working Paper 19-19 .

Gorodnichenko, Y., Shapiro, M.D., 2007. Monetary policy when potential output is uncertain: Understanding the growth gamble of the 1990s. Journal of Monetary Economics 54, 1132–1162.

Gorodnichenko, Y., Weber, M., 2016. Are sticky prices costly? evidence from the stock market. American Economic Review 106, 165–199.

Grossman, G.M., Helpman, E., 1991. Quality ladders in the theory of growth. The Review of Economic Studies 58, 43–61.

Guerron-Quintana, P.A., Jinnai, R., 2019. Financial frictions, trends, and the great recession. Quantitative Economics 10, 735–773.

Hansen, A.H., 1939. Economic progress and declining population growth. American Economic Review 29, 1–15.

Howitt, P., 2000. Endogenous growth and cross-country income differences. American Economic Review , 829–846.

Ikeda, D., Kurozumi, T., 2014. Post-Crisis Slow Recovery and Monetary Policy. Technical Report.

Jones, C.I., 2005. Growth and ideas. Handbook of economic growth vol 1 .

Jones, C.I., Williams, J.C., 1998. Measuring the social return to r&d. Quarterly Journal of Economics , 1119–1135.

Jordà, Ò., 2005. Estimation and inference of impulse responses local projections. The American Economic Review 95, 161–182.

Jordà, Ò., Singh, S.R., Taylor, A.M., 2020. The long-run effects of monetary policy. NBER working paper no. 26666 .

Krugman, P.R., 1998. It's Baaack: Japan's Slump and the Return of the Liquidity Trap. Brookings Papers on Economic Activity 2, 137–205.

Maskin, E., Tirole, J., 2001. Markov perfect equilibrium: I. observable actions. Journal of Economic Theory 100, 191–219.

Mehrotra, N., Sergeyev, D., 2020. Financial shocks, firm credit and the great recession. Journal of Monetary Economics .

Meier, M., Reinelt, T., 2019. Monetary policy, markup dispersion and aggregate tfp .

Moran, P., Queraltó, A., 2018. Innovation, productivity, and monetary policy. Journal of Monetary Economics 93, 24–41.

Queraltó, A., 2019. A model of slow recoveries from financial crises. Journal of Monetary Economics .

Ridder, M.d., 2017. Investment in Productivity and the Long-Run Effect of Financial Crisis on Output. Technical Report. Cambridge-INET Working Paper Series.

Romer, C.D., Romer, D.H., 2004. A new measure of monetary shocks: Derivation and implications. The American Economic Review .

Stadler, G.W., 1990. Business cycle models with endogenous technology. American Economic Review , 763–778.

Summers, L.H., 2015. Demand side secular stagnation. American Economic Review 105, 60–65.

Swanson, E.T., Williams, J.C., 2014. Measuring the effect of the zero lower bound on medium- and longer-term interest rates. American Economic Review 104, 3154–85.

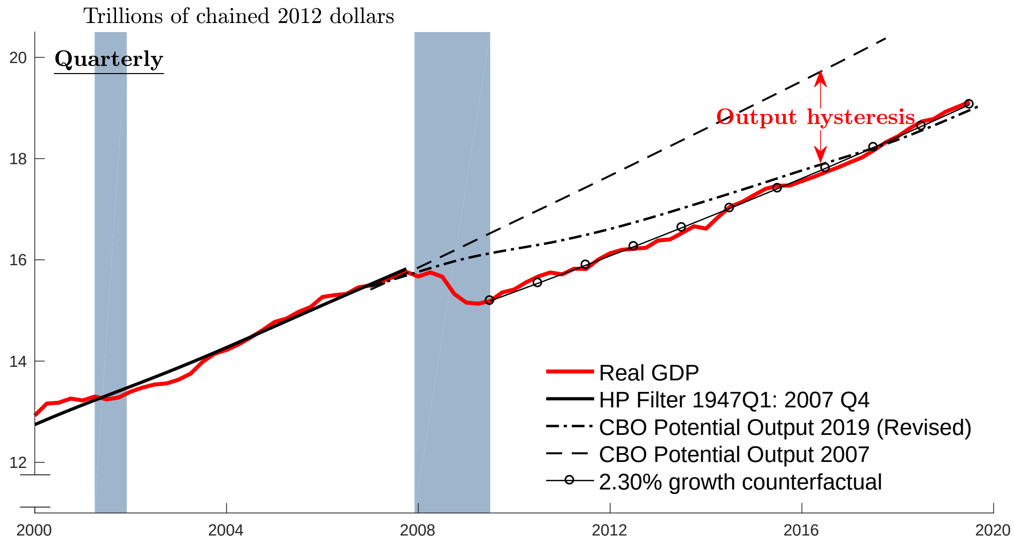
Wieland, J.F., Yang, M.J., 2016. Financial dampening. Technical Report. National Bureau of Economic Research.

Woodford, M., 2003. Interest and prices: Foundations of a theory of monetary policy. Princeton University Press Princeton, NJ.

Yellen, J.L., 2015. The outlook for the economy. Speech at the Providence Chamber of Commerce, Providence, Rhode Island, May .

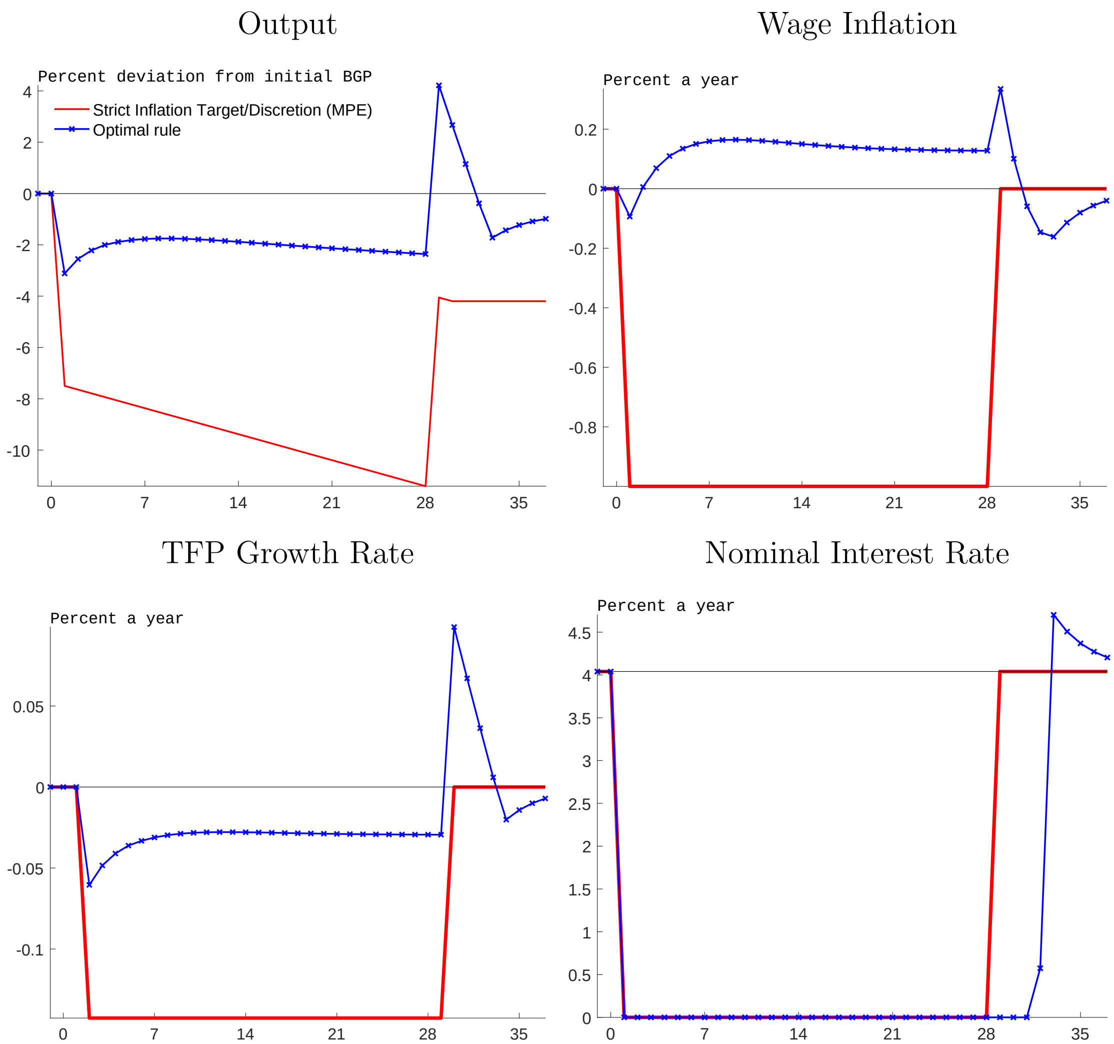
Figures

Figure 1: Real Gross Domestic Product (GDP)



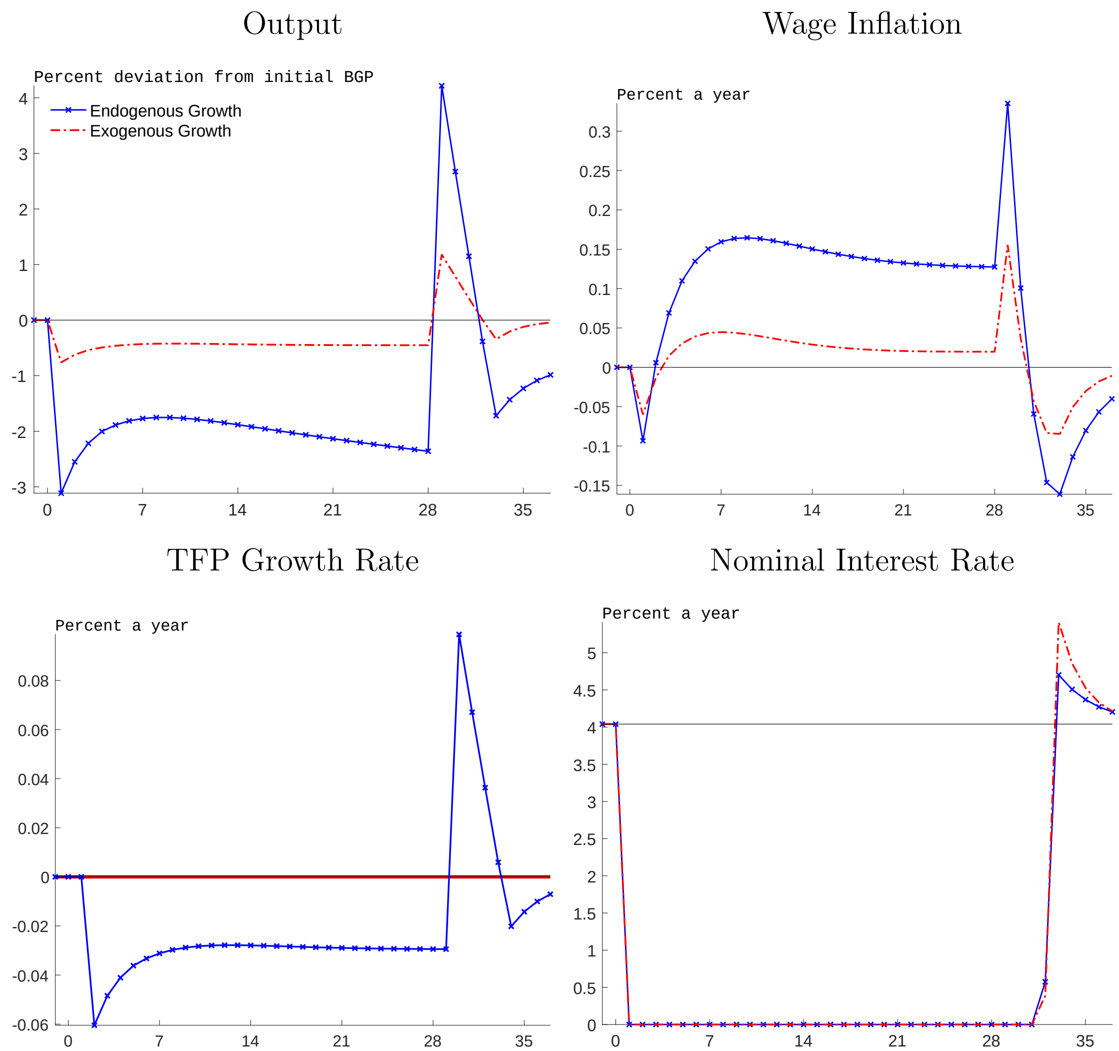
Source: Authors' calculations using quarterly real GDP data from the St. Louis FRED database and 2007 and 2019 potential output taken from the Congressional Budget Office's January 2007 and January 2019 releases.
Note: The trend line up to 2007:Q4 is estimated on quarterly data from 1947:Q1 to 2007:Q4 using the Hodrick-Prescott filter with a smoothing parameter of 1600. The solid black line with circles is constructed using a 2.30 percent annual growth rate starting in 2009Q2. The shaded areas represent recessions dated by the National Bureau of Economic Research.

Figure 2: Optimal Policy at the Zero Lower Bound



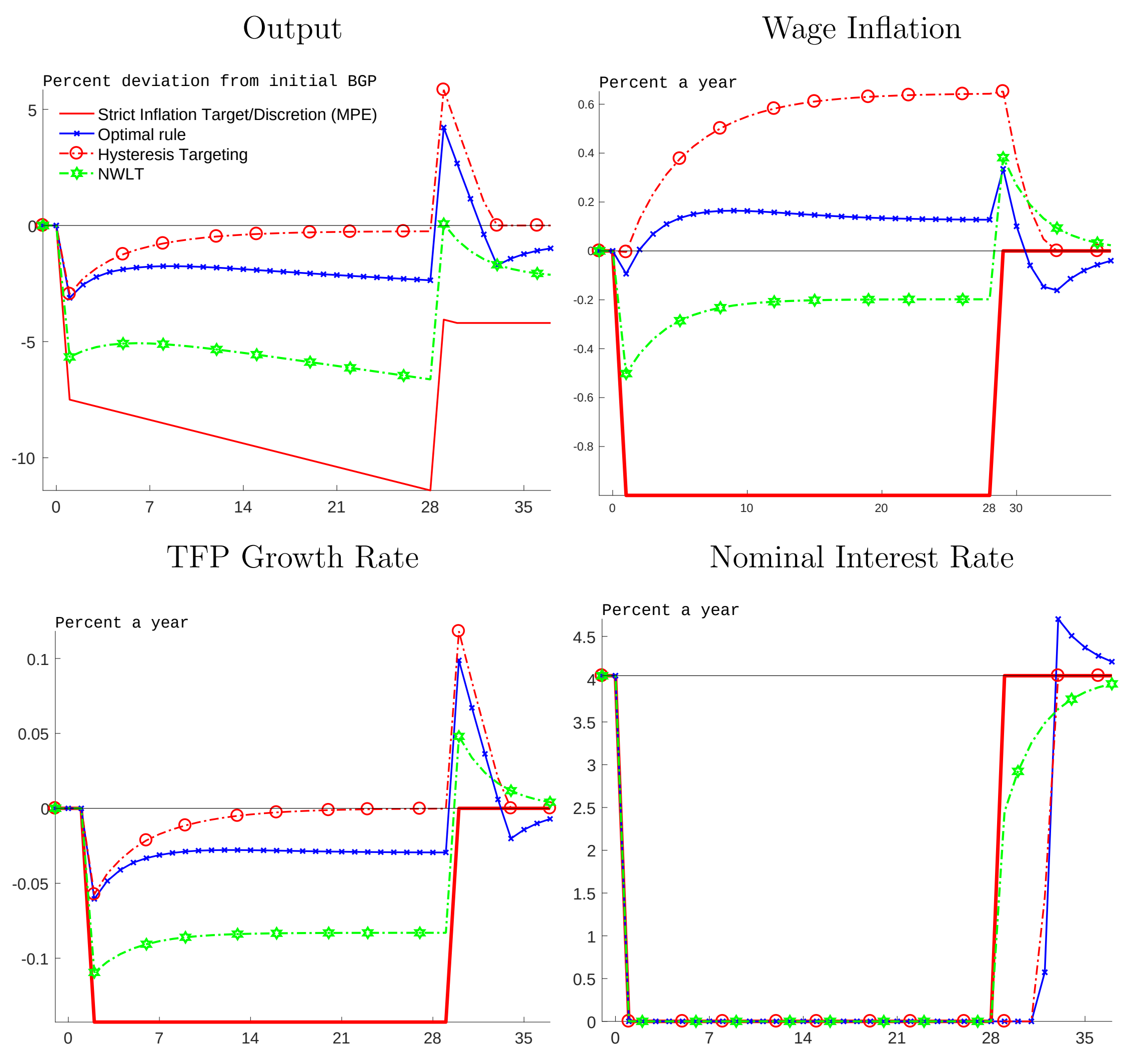
Source: Authors' calculations.
Note: The figure reports one realization of output, wage inflation, the productivity growth rate, and the nominal interest rate from a two-state Markov chain for the natural interest rate under alternate policy equilibria. In period 1, the natural interest rate becomes negative and stays there for 28 quarters, and then returns to the full employment steady state. The realizations under a Taylor rule, Markov-Perfect Equilibrium (or discretionary) optimal policy, and optimal commitment policy are shown. TFP growth rate and wage inflation are plotted as (annualized) percentage deviation from their respective steady states. Output is shown as percent deviation from its pre-shock trend level.

Figure 3: Exogenous Productivity Comparison



Source: Authors' calculations.
Note: The figure reports one realization of output, wage inflation, the productivity growth rate, and the nominal interest rate from a two-state Markov chain for the natural interest rate under alternate policy equilibria. In period 1, the natural interest rate becomes negative, stays there for 28 quarters, and then returns to the full employment steady state. Exogenous growth denotes optimal policy in the exogenous growth benchmark setting (shutting down changes in R&D and TFP growth) from same steady state as the endogenous growth calibration. The optimal rule (dashed) denotes the optimal commitment equilibrium allocation with endogenous growth. TFP growth rate and wage inflation are plotted as (annualized) percentage deviation from their respective steady states. Output is shown as percent deviation from its pre-shock trend level.

Figure 4: Alternate Rules at the Zero Lower Bound



Source: Authors' calculations.
Note: The figure reports one realization of output, wage inflation, the productivity growth rate, and the nominal interest rate from a two-state Markov chain for the natural interest rate under alternate policy equilibria. In period 1, the natural interest rate becomes negative and stays there for 28 quarters, and then returns to the full employment steady state. The realizations under a Taylor rule, Markov perfect equilibrium (or discretionary) optimal policy, optimal commitment policy, hysteresis targeting, and nominal wage-level targeting rule are shown. TFP growth rate and wage inflation are plotted as (annualized) percentage deviation from their respective steady states. Output is shown as percent deviation from its pre-shock trend level.

Tables

Table 1: Baseline Calibration of Parameters

	Value	Source/Target
Discount Rate	$\beta = 0.99$	Standard Value
Steady-State Wage Markup	$\lambda_w = 0.10$	Standard Value
Calvo Probability of Wage Adjustment	$(1 - \theta_w) = 1 - 0.75$	Standard Value
Inverse Frisch Elasticity	$\nu = 2$	Standard Value
Step-Size of Innovation	$\gamma = 1.55$	$4z = 3.6\%$
Innovation Cost Parameter	$\delta = 38.01$	$g = 2\%$
Inverse Innovation Elasticity	$\varrho = 1.90$	R&D/GDP $= 2\%$
Probability of Patent Expiration	$\eta = 0.0285$	$4(z + \eta) = 15\%$
Labor Share	$1 - \alpha = 0.5$	

Table 2: Policy Rules at the ZLB: Welfare Comparison

Policy Rule	*Relative Welfare Loss*	*Permanent Output Gap*
Optimal Rules		
Discretion (MPE)	100	−4.05
Commitment	4.84	−0.74
Commitment with higher wt on $\hat{g}_t$	18.58	0
Simple Rules		
Strict Inflation Target	100	−4.05
Hysteresis Targeting	18.58	0
Wage Level Targeting	21.61	−2.26
Nominal GDP Level Targeting	26.04	−2.51

Notes: These values report the relative welfare loss (in percent) starting from an efficient steady state. Losses are expressed in consumption equivalent units relative to the optimal discretionary rule (MPE). Under discretion, the welfare loss is 0.0048% of steady state consumption, and is normalized to 100%. The computation details are given in Appendix E.2. The true relative weight on the productivity growth rate gap is 1.52. Under a weight value of 16.78, the permanent output gap is zero. See text.

Table 3: Policy Rules at the ZLB: Welfare Comparison for Range of ϱ

Innovation Intensity ϱ	1.20	1.50	Benchmark 1.90	2.40	2.80
Permanent Output Gap (Percent)					
Discretion (MPE)	−6.74	−5.27	−4.05	−3.13	−2.64
Commitment	−0.24	−0.64	−0.74	−0.70	−0.64
Hysteresis Targeting	0	0	0	0	0
Wage Level Targeting	−2.26	−2.60	−2.26	−1.85	−1.60
Nominal GDP Level Targeting	−2.40	−2.91	−2.51	−2.03	−1.75
Relative Welfare Loss (Percent of Discretionary Equilibrium)					
Discretion (MPE)	100	100	100	100	100
Commitment	0.02	2.86	4.84	5.53	6.23
Hysteresis Targeting	0.04	8.64	18.58	24.29	29.12
Wage Level Targeting	0.21	16.55	21.51	21.46	22.86
Nominal GDP Level Targeting	0.27	21.18	26.04	24.63	25.59

Notes: These values report the relative welfare loss (in percent) starting from an efficient steady state. Losses are expressed in consumption equivalent units relative to those under optimal discretionary rule. The consumption equivalent welfare loss under discretion ranges between 0.0046% and 0.1489% of steady state consumption. Two baseline parameters are adjusted: innovation intensity elasticity, $(1/\varrho)$, and research cost, δ, to maintain the efficient-BGP growth rate and innovation rate. Across various calibrations, we also recalibrate the probability of escape from ZLB $(1-\mu)$ and the natural rate of interest $(r_S^n < 0)$ so as to keep (normalized) output drop and inflation drop under optimal discretion fixed at −7.5% and −1%, respectively. See text.