Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Demand Shocks, Hysteresis and Monetary Policy

Jae Sim

2022-080

Please cite this paper as:
Sim, Jae (2022). "Demand Shocks, Hysteresis and Monetary Policy," Finance and Economics Discussion Series 2022-080. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2022.080.

Demand Shocks, Hysteresis and Monetary Policy*

Jae Sim†

October 4, 2022

Abstract

This paper builds a micro-founded general equilibrium model of hysteresis in which changing composition of firms with heterogeneous qualities in response to demand shocks alter the total factor productivity of the economy through a process of "creative destruction". Hysteresis fundamentally challenges existing consensus on stabilization policies: the complete stabilization of demand shocks becomes suboptimal as demand creates its own supply; fiscal multiplier can be substantially larger than 1; an opportunistic monetary policymaker, who adopts a lenient policy reaction to positive demand shocks, but provides decisive monetary stimulus in response to negative demand shocks, can bring large welfare gains.

JEL Classification: E31, E32, E52, E58
Keywords: demand shocks, hysteresis, moentary policy

*The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors.

†The Board of Governors of the Federal Reserve System. Email: jae.w.sim@frb.gov

1 Introduction

On October 14, 2016, Janet L. Yellen, then the Chair of the Board of Governors of the Federal Reserve System at a research conference, called for a new research agenda to reconsider the traditional relationship between aggregate demand and supply, opening up a question that fundamentally challenges the underpinning of the philosophy behind the current consensus on the mechanism of how stabilization policies should work, a consensus that has not changed since Milton Friedman proposed the hypothesis of the natural rate of unemployment, about 50 years ago:

> "*Are there circumstances in which changes in aggregate demand can have an appreciable, persistent effect on aggregate supply?..., albeit more speculatively, strong demand could potentially yield significant productivity gains by, among other things, prompting higher levels of research and development spending and increasing the incentives to start new, innovative businesses.*"– Yellen (2016)

As Blanchard (2017) pointed out, the natural rate hypothesis of Milton Friedman had two subhypotheses, so called "independence hypothesis" and "accelerationist hypothesis". The first hypothesis has been much more influential than the second in the history of macroeconomics as the profession has drifted away from the second hypothesis over time while the first remains as a corner stone of any stabilization theory even today. It is the first hypothesis that Yellen (2016) asks the profession to call into question.

The first hypothesis posits that there exists a "natural rate of unemployment" or natural level of output that is determined by "real forces" and is independent of aggregate demand factors, including, among others, monetary policy. Ever since Friedman (1977), the independence hypothesis has been at the center of monetary economics and the majority of Keynesian economists has accepted the concept as unquestionable axiom, which can be seen in the expression such as "*exogenously* varying natural rate of output" (Woodford (2001)) or "*exogenous* variation in the "natural rate of output" as a result of any of several types of *real* disturbances" (Woodford (2003)).[1]

However, there was a small group of Keynesian economists who proposed and tested a counter-hypothesis known as hysteresis or history-dependence of the natural rate: Blanchard and Summers (1987), Mankiw (2001), Blanchard and Diamond (1994) and Ball (1999). For instance, Ball (1999) pointed out a possibility that "as demand pushes unemployment away from the current NAIRU, this causes the NAIRU itself to change over time". He then explained the difference in the natural rate of unemplyment rates between the EU and the U.S. by the difference in strength and duration in disinflation policies adopted by the two continents. He suggested that monetary policy, a primary demand factor, could affect the natural rates of unemployment. However, he did not develop structural models of hysteresis and the exact mechanisms of hysteresis are left as a black box.

There are at least two compelling reasons to reconsider hysteresis hypothesis at the current juncture: the "missing deflation puzzle" during the Global Financial Crisis in 2008 and the "missing

[1] In both quotes, the emphasis is by the author of this paper.

inflation puzzle" during the subsquent recovery. The wide resource gaps that opened up during the crisis generated fear of deflation. However, core inflation rates in both eurozone and U.S. economies were surprizing stable, hovering around 1 percent. The fear of deflation simply failed to materialize into reality. In contrast, the U.S. economy is now experiencing exactly opposite phenomenon. Even after the unemployment rate fell to 3.7 percent in 2018Q3, which is considered far below the influential estimates of the natural rate, for instance, 4.6 percent by the Congressional Budget Office, the core Personal Consumption Expenditure inflation rate in 2018Q3 remains at 2 percent year-over-year rate. Again, the fear that inflation was just around corner has failed to materialize into reality.

One popular explanation of these two puzzle is based on the notion of so-called "flat" Phillips curve. The notion can be understood in the context of the second hypothesis of Friedman (1977):

$$\pi_t - \pi_{t-1} = -a(u_t - u_t^N), \quad a > 0. \tag{1}$$

(1) is the accelerationist Phillips curve proposed by Friedman (1977). It says that inflation rate is accelerated when the unemployment rate u_t falls below the natural rate u_t^N. The proponents of flat Phillips curve argue that if a is closed to zero, the volatility of inflation rate must be much smaller than the volatility of unemployment gap.[2]

The goal of this paper is neither to prove nor to disprove the notion of flat Phillips curve. Rather, I show that there is an alternative explanation of the two puzzles: hysteresis. As pointed out by Yellen (2016), "persistent shortfalls in aggregate demand" during the Global Financial Crisis "could adversely affect the supply side of the economy". If so, the shocks that had elevated u_t also might have raised u_t^N during the crisis. Note that one of these shocks was monetary policy shock in that the binding zero lower bound constraint during this period was no different than monetary policy tightening shocks. In this situation, even though nothing happened to a of (1) and the Phillips curve was not "flat", the downward pressure on inflation rate might have been much more subdued if the natural rate of unemployment rate was closely tracking the actual rate of unemployment.

In reverse, if the demand shocks that had lowered u_t during the recovery, including extraordinary monetary policy stimulation through quantative easing and forward guidance, also had lowered u_t^N into an unprecedently low level, the upward pressure on inflation rate might have been tepid for this reason despite glaringly low level of unemployment rate. Thus hysteresis and reverse hysteresis can explain both "missing deflation" puzzle and "missing inflation" puzzle without relying on the notion of flat Phillips curve.

This paper builds a micro-founded general equilibrium model of hysteresis in which a changing composition of firms with heterogeneous qualities in response to demand shocks alters the total factor productivity (TFP henceforth) of the economy endogenously. To create the hysteresis channel, I introduce Baldwin and Krugman (1989)'s hysteresis model of firm dynamics in an otherwise standard New Keynesian model. Firms are created with sunk entry cost when new entrant firms realize good random draws of technology, good enough to justify the cost of entry. Incumbent firms face fixed

[2] An extreme form of this notion is recently delivered by the Director of the National Economic Council, Larry Kudlow, who says, "It's time to bury the Phillips curve". See https://twitter.com/larry_kudlow/status/913116542346375170.

operation costs such that when their productivity draws are bad enough, incumbent firms exit the market. Positive demand shocks raise the aggregate TFP by replacing relatively inefficient old firms with more efficient new firms. By a process of creative destruction, positive demand shocks change the composition of firms, elevating the effective level of productivity and keeping the inflation pressure from reaching a harmful level. Inverse Say's law holds: Demand creates its own supply. Negative demand shocks accomplish exactly the opposite by delaying the entries of firms with good qualities and helping firms with low qualities live on thanks to depressed factor prices, deteriorating overall TFP, reverse hysteresis. Lack of demand destroys supply capacity.

The size of the hysteresis effect in the model is dependent on the strength of the returns-to-specialization. With an increasing returns-to-specialization, a new entry creates a positive externality of "business enhansing" effect, making the hysteresis effects even stronger. With a decreasing returns-to-specialization, a new entry creates a negative externality of "business stealing" effect (see Spence (1976) and Mankiw and Whinston (1986)), making the hysteresis effects weaker.

Such hysteresis mechanism has important implications for monetary and fiscal policies. In a canonical New Keynesian model, aggregate demand disturbances are the objects of complete stabilization. A central bank tries to minimize the deviation of the economy from the flexible price economy. Often optimal monetary policy takes the form of increasing the reaction coefficients to inflation gap and/or unemployment gap coefficient indefinitely. Clarida, Galí, and Gertler (1999), for instance, formally show that increasing the coefficient for inflation gap indefinitely is optimal in the case of optimal monetary policy without commitment in the absence of cost-push shock, or equivalently, when the weight given to the output gap in the central bank's loss function converges to zero. In fact, in the absence of cost push shock, the optimal monetary policy is able to achieve zero inflation volatility by making the reaction coefficients of a given feedback rule arbitrarily large.[3]

However if the hysteresis mechanism described in the above is at work, a complete stabilization of positive demand shock becomes suboptimal as this is equivalent of "leaving money on the table". In this environment, demands shock have a mixed feature of demand shocks and supply shocks. I show that the welfare-maximizing monetary policy rule has finite coefficients for both inflation gap and the output gap. It is also shown that the welfare-maximizing inflation coefficient declines in the degree of returns-to-specialization, that is, the stronger the positive externality from firm entry, the more dovish monetary policy optimal.

More importantly, the presence of hysteresis implies that the welfare-maxizing monetary policy response to demand shocks should be asymmetric. If postive demand shocks lead to expansion in supply capacity, optimal monetary policy lets the demand shocks walk its way into supply by conducting lenient monetary policy against demand shocks. On the contrary, if negative demand shocks were to adversely affect supply capacity, optimal monetary policy should prevent such demand shocks from eating away supply capacity by providing the economy with massive monetary policy stimulus. To prove this intuition, I optimize the following asymmetric monetary policy rule to maxize the welfare

[3] Note that this property only holds perfectly in the simplest of all canonical models where the flexible price allocation is efficient. There are many mechanisms that can make this property to fail to hold, for instance, rigidity of real wage.

of the representative household,

$$r_t = r^* + \mathbf{1}(\pi_t \geq \pi^*)\rho_\pi^{(+)}(\pi_t - \pi^*) + \mathbf{1}(\pi_t < \pi^*)\rho_\pi^{(-)}(\pi_t - \pi^*) \tag{2}$$

and show that $\rho_\pi^{(+)} < \rho_\pi^{(-)}$ is optimal if the hysteresis is present. Furthermore, the degree of asymmetry in optimal monetary policy rule in terms of $\rho_\pi^{(-)} - \rho_\pi^{(+)}$ increases with the degree of returns-to-specialization.[4] Note that in the presence of positive externality due to returns-to-specialization, the competitive equilibrium delivered by free entry condition is inefficient because individual firms do not internalize the benefits of their own entries to incumbent firms. Hence optimal allocation should subsidize new firm entries. In the absence of active fiscal authority implementing subsidy policy, monetary policy could subsidize firm entries by implementing the asymmetric policy, i.e., $\rho_\pi^{(+)} < \rho_\pi^{(-)}$.

The welfare gains from such asymmetric monetary policy rules can be quite sizable. It is shown that depending on the degree of returns-to-specialization, asymmetric optimal monetary policy rule could bring large increase in the *mean* level of aggregate output on the order of 1.6 to 1.9 percent and 20 to 39 percent reduction in the standard deviation of aggregate output as compared with the economy with symmetric optimal monetary policy rule. Furthermore, asymmetric optimal monetary policy can bring 35 to 53 percent reduction in inflation volatility.

Monetary policy is not the only place where the presence of hysteresis mechanism fundamentally challenges our understadings of stabilization policies. I show that if the hysteresis mechanism is present and the monetary authority conducts its monetary policy using the optimized asymmetric rule, a fiscal multiplier can be as large as 3.7 to 4.8 depending on the strength of returns-to-specialization. DeLong and Summers (2012) has recently developed a macroeconomic model in which a fiscal multiplier can be substantially larger than 1 due to the hysteresis effect. However, their attempt faced a criticism that their model was not based on "first-order principles" (see Ramey (2012)). The results in this paper indicate that a model of fiscal multiplier due to hysteresis can be constructed from the first-order principles.

The rest of the paper is organized as follows: Section 2 develops the theoretical model; Section 3 discuss the main findings regarding business cycle properties under the baseline monetary policy rules, optimality of asymmetric monetary policy rule and fiscal multiplier under hysteresis and optimal asymmetric monetary policy; Section 4 concludes; Appendix A provides the proofs of the two Propositions in the main text. [5]

[4] The kink in the monetary policy rule (2) implies that the model cannot be solved by a conventional perturbation method and has to be solved and simulated by a fully nonlinear solution method. Specifically, the expanded path approach developed by Adjemian and Juillard (2011) is adopted for the nonlinear solution. See the online Appendix D to see more details regarding the nonlinear solution method.

[5] Online (not intended for publication) Appendix B provides the complete set of model equations; Online Appendix C shows how the non-stochastic steady state can be pinned down; Online Appendix D explains how the fully nonlinear solution method works.

2 The Model

We combine Ethier (1982)-Romer (1990)-type monopolistic competition model of time-varying variety with Baldwin and Krugman (1989)'s hysteresis model based on firm dynamics in an economy with nominal rigidity. Owing to endogenous exits, the number of firms do not grow in the long run. However, TFP becomes endogenous not because of changes in investment in research and development as in Romer (1990), but because the composition of firms with heterogeneous productivities changes in response to aggregate demand shocks.

2.1 Production Firms

2.1.1 Technology

There exists a continuum of product variety indexed by $[0, 1]$. The term, variety is interchangable with firm in this paper. Some firms in the continuum are active and some firms are inactive and exist only as "ideas". An active firm produces an intermediate good $y_t(i)$ using a constant-returns-to-scale Cobb-Douglas production funciton,

$$y_t(i) = m_t(i) x z_t(i) k_t(i)^a n_t(i)^{1-\alpha} \tag{3}$$

where $k_t(i)$ and $n_t(i)$ denote capital and labor. x is the aggregate productivity.[6] Throughout the analysis, x is assumed to be a constant. $z_t(i)$ is an i.i.d. idiosyncratic technology shock following a lognormal distribution,

$$\log z(i) \sim N(-0.5\sigma_z^2, \sigma_z^2)$$

with its cdf denoted by $F(\cdot)$. $m_t(i) \in \{0, 1\}$ denotes the activity status of firm i. If $m_t(i) = 1$, the firm is active in the current period. If $m_t(i) = 0$, the firm is inactive either because it has never entered the market or because it exited the market in the past.

Note that $\mathbb{E}_t[z_t(i)] = 1$. However $\mathbb{E}_t[z_t(i)|m_t(i) = 1]$ is not a constant and time-varies as the composition of active firms changes over time. The total measure of active firms today is given by

$$M_t \equiv \int_0^1 m_t(i) di.$$

The varieties of intermediate products are then sold to a representative firm, which produces the final good using a CES technology,

$$y_t = M_t^{1+v-\frac{\epsilon_p}{\epsilon_p - 1}} \left[\int_0^{M_t} y_t(i)^{\frac{\epsilon_p - 1}{\epsilon_p}} di \right]^{\frac{\epsilon_p}{\epsilon_p - 1}}, \quad \epsilon_p \in (1, \infty), \tag{4}$$

where ϵ_p is the elasticity of substitution among different varieties and v is the returns-to-specialization

[6] As will be made clear, $n_t(i)$ is itself a a composite input, combining a continuum of varieties of labor inputs $n_t(i, j)$ using a Dixit-Stiglitz aggregator, that is $n_t(i) = \left[\int_0^1 n_t(i, j)^{\frac{\epsilon_w - 1}{\epsilon_w}} dj \right]^{\frac{\epsilon_w}{\epsilon_w - 1}}$ where ϵ_w is the elasticity of substitution among different labor inputs.

(see Benassy (1996), Devereux, Head, and Lapham (1996) and Sim (2008)). If $v > 0$ ($v < 0$), there is an increasing (decreasing) returns-to-scale due to specialization. To see this, consider the fact that if all active firms produce an identical quantity, say, $\bar{y}$, the output of the final-good becomes $y = \bar{y}M^{1+v}$. $v > 0$ ($v < 0$) makes the output increasing(decreasing) in M more (less) than proportionately.

The price index implied by the zero profit condition of the final good sector is given by

$$P_t = M_t^{-\left(1+v-\frac{\epsilon_p}{\epsilon_p-1}\right)} \left[\int_0^{M_t} P_t(i)^{1-\epsilon_p} di\right]^{1/(1-\epsilon_p)}. \tag{5}$$

To streamline the notations, we define

$$\xi \equiv (1-\epsilon_p)\left(1+v-\frac{\epsilon_p}{\epsilon_p-1}\right). \tag{6}$$

Using (6), we can express the product demand for a variety i as

$$y_t(i) = M_t^{-\xi}\left[\frac{P_t(i)}{P_t}\right]^{-\epsilon_p} y_t. \tag{7}$$

We impose a parameter restriction such that $\xi > 0$ or equivalently $v < (\epsilon_p - 1)^{-1}$.

2.1.2 Firm Dynamics

When a new firm enters the market, the firm has to make an investment, which is sunk upon entry. This cost is denoted by a fraction $\gamma^E \in (0,1)$ of the non-stochastic steady state of aggregate output (y). Any active firm, whether new or incumbent, has to incur fixed operating cost in each period. This cost is denoted by a fration $\gamma^X (\leq \gamma^E)$ of the non-stochastic steady state of aggregate output.

Let $V_t(z_t(i), m_{t-1}(i))$ denote the value of a firm with an idiosyncratic shock $z_t(i)$ and an activity status $m_{t-1}(i)$ in the previous period. To analyze the firm's entry/exit decision, 4 auxiliary value functions are defined as follows. For each of $(m_{t-1}(i), m_t(i)) \in \{0,1\} \times \{0,1\}$, an auxiliary value function is defined as

$$W_t(z_t(i), m_{t-1}(i), m_t(i)) = m_t(i)[\Pi_t(z_t(i)) - (1-m_{t-1}(i))\gamma^E y - \gamma^X y] + \mathbb{E}_t[Q_{t,t+1}V_{t+1}(z_{t+1}(i), m_t(i))] \tag{8}$$

where $\Pi_t(z_t(i))$ is the profit before entry and fixed costs, $Q_{t,t+1}$ is the stochastic discounting factor of the representative household. For instance, consider a new firm that has to decide to enter or not today. For this firm, $W_t(z_t(i), 0, 1)$ is the value of entry today while $W_t(z_t(i), 0, 0)$ is the value of delaying entry. Likewise, $W_t(z_t(i), 1, 1)$ and $W_t(z_t(i), 1, 0)$ are the value of delaying exit and the value of exiting for an incumbent firm. The value of a firm can then be expressed, regardless of its activity status in the previous period, as:

$$V_t(z_t(i), m_{t-1}(i)) = \max\{W_t(z_t(i), m_{t-1}(i), 0), W_t(z_t(i), m_{t-1}(i), 1)\}. \tag{9}$$

One can then define entry and exit threshold levels of productivity as a set of cutoff values of z that satisfy

$$W_t(z_t^E, 0, 0) = W_t(z_t^E, 0, 1) \tag{10}$$

in the case of entry problem ($m_{t-1}(i) = 0$) and

$$W_t(z_t^X, 1, 0) = W_t(z_t^X, 1, 1) \tag{11}$$

in the case of exit problem ($m_{t-1}(i) = 1$) where z_t^E and z_t^X are entry and exit threshold levels of idiosyncratic technology, respectively. In words, an inactive firm with the idiosyncratic technology level $z_t(i) = z_t^E$ is indifferent between staying inactive and entering the market. Likewise, an active firm with the idiosymcratic technology level $z_t(i) = z_t^X$ is indifferent between staying active and exiting the market. Using the definition of (8), one can express (10) and (11) as

$$(\gamma^E + \gamma^X)y_t = \Pi_t(z_t^E) + q_t^F \tag{12}$$

and

$$\gamma^X y_t = \Pi_t(z_t^X) + q_t^F \tag{13}$$

where

$$q_t^F \equiv \mathbb{E}_t\{Q_{t,t+1}[V_{t+1}(z_{t+1}(i), 1) - V_{t+1}(z_{t+1}(i), 0)]\}. \tag{14}$$

One can think of q_t^F as the expected surplus value tomorrow that is created by the decision to remain active today either by creating an additional firm or by postponing an exit. Note that since the profit is monotonically increasing in z, (12) and (13) imply that $z_t^E > z_t^X$ since the profit is monotonically increasing in z and $\gamma^E > 0$.

These conditions can be viewed as zero profit conditions (also "free entry" and "free exit"): the current profit and the expected surplus value of staying active today should cover fixed costs (inclusive of the entry cost in the case of a new entrant) at the entry and exit threshold technology levels.

Proposition 1 *The surplus value of firm given by (14) satisfies the following recursion:*

$$q_t^F = \mathbb{E}_t\left[Q_{t,t+1}\left(\int_{z\geq z_{t+1}^E} \gamma^E y dF(z) + \int_{z_{t+1}^X \leq z \leq z_{t+1}^E} \left[\Pi_{t+1}(z) - \gamma^X y + q_{t+1}^F\right] dF(z)\right)\right] \tag{15}$$

Proof See Appendix A.

Firm dynamics is then determined by the entry and exit conditions, (12) and (13), the dynamics of surplus firm value, (15) and the law of motion for the mass of active firms:

$$M_t = M_{t-1}[1 - F(z_t^X)] + (1 - M_{t-1})[1 - F(z_t^E)]. \tag{16}$$

The law of motion implies that only the incumbent firms whose idiosyncratic shocks are above the exit threshold z_t^X remain active ($M_{t-1}[1 - F(z_t^X)]$) and only the new firms whose idiosyncratic shocks are above the entry threshold z_t^E enter the market $((1-M_{t-1})[1-F(z_t^E)])$. For later purposes, standardized entry and exit thresholds are defined:

$$\nu_t^E \equiv \frac{\log(z_t^E)+0.5\sigma_z^2}{\sigma_z} \tag{17}$$

and

$$\nu_t^X \equiv \frac{\log(z_t^X)+0.5\sigma_z^2}{\sigma_z}. \tag{18}$$

Using (17) and (18), the law of motion (16) can be expressed as

$$M_t = M_{t-1}[1-\Phi(\nu_t^X)] + (1-M_{t-1})[1-\Phi(\nu_t^E)], \tag{19}$$

where $\Phi(\cdot)$ denotes the cdf of standard normal distribution.

2.1.3 Optimal Pricing

The firms face nominal rigidity in setting product prices. Nominal rigidity is introduced a là Rotemberg (1982). The symmetric equilibrium of Rotemberg (1982) under our timing assumption that price choice has to be made prior to the realization of idiosyncratic shock uniquely pins down the entry/exit thresholds. Firms are assumed to face quadratic adjustment cost when they change nominal prices. In particular, the generalized New Keynesian version by Ireland (2007) is assumed:

$$\frac{\varphi_p}{2}\left(\frac{P_t(i)}{\pi_{t-1}^{\chi}P_{t-1}(i)}-1\right)^2 y_t M_t^{-\xi} = \frac{\varphi_p}{2}\left(\frac{\pi_t}{\pi_{t-1}^{\chi}}\frac{p_t(i)}{p_{t-1}(i)}-1\right)^2 y_t M_t^{-\xi}, \quad \chi \in [0,1)$$

where $p_t(i) \equiv P_t(i)/P_t$, and π_t is gross inflation rate, i.e., $\pi_t \equiv P_t/P_{t-1}$. χ is the indexation parameter, which makes the Phillips curve both backward and forward looking. Note that it is assumed that the price adjustment cost is proportional to $y_t M_t^{-\xi}$ for the homogeneity of the Phillips curve.

Entrant firms were not active last period and the cost of changing its *own* relative price from last period is not defined. For this reason, entrant firms are assumed to face different price setting cost:

$$\frac{\varphi_p}{2}\left(\frac{P_t(i)}{\pi_{t-1}^{\chi}P_{t-1}}-1\right)^2 y_t M_t^{-\xi} = \frac{\varphi_p}{2}\left(\frac{\pi_t}{\pi_{t-1}^{\chi}}p_t(i)-1\right)^2 y_t M_t^{-\xi}.$$

This cost assumes that entrant firms face the cost of deviating from the average price last period. Taken together, these assumptions imply that the firm's real profit before fixed costs is

$$\Pi_t(i) = p_t(i)y_t(i) - r_t^K k_t(i) - w_t n_t(i) - \frac{\varphi_p}{2}\left[\frac{\pi_t}{\pi_{t-1}^{\chi}}p_t(i)\left(\frac{m_{t-1}(i)}{p_{t-1}(i)}+1-m_{t-1}(i)\right)-1\right]^2 y_t M_t^{-\xi} \tag{20}$$

where r_t^K and w_t are rental rate of capital and equilibrium real wage.

As mentioned, to formulate the optimal pricing problem in the context of symmetric equilibrium, it is assumed that the pricing decision has to be made before the realization of the idiosyncratic shock.

In this environment, the firms make decision on pricing based on the expected real marginal cost $\mathbb{E}_t^i[\mu_t(i)]$ where the expectation operator is defined with repect to the idiosyncratic uncertainty. The otimal pricing problem then solves

$$\max_{p_t(i)} \left\{ \left[p_t(i)^{1-\epsilon_p} - \mathbb{E}_t^i[\mu_t(i)]p_t(i)^{-\epsilon_p}\right] M_t^{-\xi} y_t - \frac{\varphi_p}{2}\left[\frac{\pi_t}{\pi_{t-1}^{\chi}} p_t(i)\left(\frac{m_{t-1}(i)}{p_{t-1}(i)} + 1 - m_{t-1}(i)\right) - 1\right]^2 y_t M_t^{-\xi} \right.$$
$$\left. + \mathbb{E}_t \sum_{s=t+1}^{\infty} \tilde{Q}_{t,s}\left[\left[p_s(i)^{1-\epsilon_p} - \mathbb{E}_s[\mu_s(i)]p_s(i)^{-\epsilon_p}\right] M_s^{-\xi_s} y_s - \frac{\varphi_p}{2}\left(\frac{\pi_s}{\pi_{s-1}^{\chi}}\frac{p_s(i)}{p_{s-1}(i)} - 1\right)^2 y_s M_s^{-\xi_s}\right]\right\}$$

where $\tilde{Q}_{t,s} \equiv Q_{t,s}\prod_{j=t+1}^{s}[1-\Phi(\nu_j^X)]$. The first order condition for this problem, after imposing the symmetric equilibrium condition $p_t(i) = p_{t+1}(i) = 1$, is

$$1 = \frac{\epsilon_p}{\epsilon_p - 1}\mathbb{E}_t^i[\mu_t(i)] - \frac{\varphi_p}{\epsilon_p - 1}\frac{\pi_t}{\pi_{t-1}^{\chi}}\left(\frac{\pi_t}{\pi_{t-1}^{\chi}} - 1\right)$$
$$+ \frac{\varphi_p}{\epsilon_p - 1}\mathbb{E}_t\left[Q_{t,t+1}[1-\Phi(\nu_{t+1}^X)]\frac{\pi_{t+1}}{\pi_t^{\chi}}\left(\frac{\pi_{t+1}}{\pi_t^{\chi}} - 1\right)\frac{y_{t+1}/M_{t+1}^{\xi}}{y_t/M_t^{\xi}}\right] \tag{21}$$

for both entrants and incumbents. This is the Phillips curve of the model. It is straightforward to show that the log-linear dynamics of inflation rate is captured by a both forward- and backword-looking Phillips curve:

$$\hat{\pi}_t = \frac{\chi}{1+\chi\beta[1-\Phi(\nu^X)]}\hat{\pi}_{t-1} + \frac{\beta[1-\Phi(\nu^X)]}{1+\chi\beta[1-\Phi(\nu^X)]}\mathbb{E}_t[\hat{\pi}_{t+1}] + \frac{\mu\epsilon/\varphi_p}{1+\chi\beta[1-\Phi(\nu^X)]}\mathbb{E}_t^i[\hat{\mu}_t], \tag{22}$$

where variables with hat are log-linearized values of the original variables and the variables without time subscripts are the values at the nonstochastic steady state.

2.1.4 Aggregate Production Function

Cost minimization, taking place after the realization of the idiosyncratic shock and hence after the entry/exit decision, requires

$$r_t^K = \alpha\mu_t(i)x_t z_t(i)\left[\frac{k_t(i)}{n_t(i)}\right]^{\alpha-1}$$

and

$$w_t = (1-\alpha)\mu_t(i)x_t z_t(i)\left[\frac{k_t(i)}{n_t(i)}\right]^{\alpha}$$

where $\mu_t(i)$ is the realized real marginal cost. These two conditions imply that the capital labor ratio is equalized across heterogeneous firms regardless of their levels of idiosyncratic technology, i.e.,

$$\frac{k_t(i)}{n_t(i)} = \left(\frac{\alpha}{1-\alpha}\right)\frac{w_t}{r_t^K} \equiv \kappa_t. \tag{23}$$

In this symmetric equilibrium, all active firms produce an identical level of output regardless of their productivity differentials. Since some firms are less productive than others, these firms need to use greater amounts of production inputs, but the ratio of production input must be equalized across all active firms. The equalization of product demand and supply implies

$$M_t^{-\xi}\left[\frac{P_t(i)}{P_t}\right]^{-\epsilon_p} y_t = x_t z_t(i) k_t(i)^\alpha n_t(i)^{1-\alpha}. \tag{24}$$

The following proposition shows the existence of aggregate production function and provides its analytical functional form.

Proposition 2 *Define* $k_t \equiv \int_0^1 m_t(i)k_t(i)di$, $n_t \equiv \int_0^1 m_t(i)n_t(i)di$. *The efficiency condition for factor utilization (23) and the market clearing condition (24) imply that the aggregate production function exists and is of the form:*

$$y_t = x\Xi_t k_t^\alpha n_t^{1-\alpha} \tag{25}$$

where

$$\Xi_t \equiv \frac{\{[(1-M_{t-1})[1-\Phi(\nu_t^E)] + M_{t-1}[1-\Phi(\nu_t^X)]\}^\xi}{(1-M_{t-1})[1-\Phi(\nu_t^E+\sigma_z)] + M_{t-1}[1-\Phi(\nu_t^X+\sigma_z)]}. \tag{26}$$

Proof See Appendix A.

Note that Ξ_t in the aggregate production function (25) can be thought of as the endogenous component of Solow residuals. Appendix A shows that the endogenous component of TFP is given by

$$\Xi_t = M_t^\xi \cdot \left(\int m_t(i)z_t(i)^{-1}di\right)^{-1}. \tag{27}$$

The more productive active firms are, the greater the endogenous TFP is. Therefore, if aggregate shocks that induce more entries of productive firms and exits of inefficient marginal firms, the TFP level of the economy is increased endogenously. Substituting (25) and (27) in the demand for an individual product and imposing the symmetric equilibrium condition yields

$$y_t(i) = \left(\int m_t(i)z_t(i)^{-1}di\right)^{-1} x k_t^\alpha n_t^{1-\alpha} \text{ for all } i \in [0,1]. \tag{28}$$

This shows that all active firms benefit from the increase in average quality of firms.

In our baseline calibration, we set $v = 0$ to exclude any effects of the externality stemming from the returns-to-specialization. In this case, competitive equilibrium is socially optimal. If $v > 0$, there exists "business enhansing" effects associated with new entry. In this case, the number of variety in the competitive equilibrium is too small relative to socially optimal level. If $v < 0$, there exists "business stealing" effects associated with new entry. In this case, the number of varieties in the competitive equilibrium is too large relative to socially optimal level (see Spence (1976) and Mankiw and Whinston

(1986)). The presence of such externality implies that the competive equilibrium analyzed in this paper is not socially optimal.[7] Later we explore the monetary policy implication of positive externality.

Corollary *The average marginal cost of the economy is given by*

$$\mathbb{E}_t^i[\mu_t(i)] = \frac{1}{x_t \Xi_t} \left(\frac{r_t^K}{\alpha} \right)^{\alpha} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha}. \tag{29}$$

Proof (25) is the sum of the products of all active firms. Hence, average level of output is given by

$$\frac{y_t}{M_t} = x_t \Xi_t \left(\frac{k_t}{M_t} \right)^{\alpha} \left(\frac{n_t}{M_t} \right)^{1-\alpha} \tag{30}$$

Using the average production function (30), the conditional factor demands can be derived as

$$\frac{n_t}{M_t} = \kappa_t^{-\alpha} \frac{1}{x_t \Xi_t} \frac{y_t}{M_t} \tag{31}$$

and

$$\frac{k_t}{M_t} = \kappa_t^{1-\alpha} \frac{1}{x_t \Xi_t} \frac{y_t}{M_t}. \tag{32}$$

Average total cost of producing average output y_t/M_t can be expressed as

$$\begin{aligned} TC\left(\frac{y_t}{M_t} \right) &= w_t \frac{n_t}{M_t} + r_t^K \frac{k_t}{M_t} \\ &= \frac{1}{x_t \Xi_t} \left(\frac{r_t^K}{\alpha} \right)^{\alpha} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha} \frac{y_t}{M_t} \end{aligned}$$

Hence the average marginal cost is given by $\mathbb{E}_t^i[\mu_t(i)] = TC'(y_t/M_t)$, which is equal to (29) in the Corollary. Q.E.D.

Not surprisingly, an increase in endogenous TFP lowers the average real marginal cost, which matters for pricing decisions as shown by log-linearized Phillips curve (22). This means that when endogenous TFP is increased during an expansion driven by positive demand shocks, inflation pressure is more subdued than in the economy without the endogenous TFP channel. This is precisely the meaning of the expression "demand creates its own supply", that is, the inverse Say's Law, a potential resolution to "missing inflation " puzzle. Conversely, (29) implies that when endogenous TFP is decreased during a contraction driven by negative demand shocks, deflation pressure is more subdued than in the economy without the endogenous TFP channel, a potential resolution to "missing deflation " puzzle.

[7] Sim (2008) shows how a constrained optimum can be constructed such that the total value of firms is maximized for the representative household subject to the mark-up pricing rule implied by the monopolistic competition in the presence of the externality in a real business cycle setting.

2.2 Household

We model the household as a large family in which a continuum of family members exist and provide intermediate labor inputs to firms in monopolistically competitive labor markets. Firms use a composite labor inputs by aggregating intermediate labor inputs in a CES function:

$$n_t = \left[\int_0^1 n_t(j)^{\frac{\epsilon_w - 1}{\epsilon_w}} dj\right]^{\frac{\epsilon_w}{\epsilon_w - 1}}, \quad \epsilon_w \in (1, \infty) \tag{33}$$

The cost minimization problem of firms yields the labor demand for each intermediate labor input as

$$n_t(j) = \left(\frac{W_t(j)}{W_t}\right)^{-\epsilon_w} n_t \equiv w_t(j)^{-\epsilon_w} n_t, \tag{34}$$

where aggregate nominal wage index W_t satisfies

$$W_t = \left[\int_0^1 W_t(j)^{1-\epsilon_w} dj\right]^{1/(1-\epsilon_w)}.$$

The household faces a quadratic cost of adjusting real wage:

$$\frac{\varphi_w}{2}\left(\frac{W_t(j)/P_t}{W_{t-1}(j)/P_{t-1}} - 1\right)^2 n_s(j) = \frac{\varphi_w}{2}\left(\frac{w_t(j)}{w_{t-1}(j)}\frac{w_t}{w_{t-1}} - 1\right)^2 n_s(j)$$

where $w_t \equiv W_t/P_t$. Note that by supplying $n_t(j)$ units of labor input j, the households earn the following real income before tax:

$$\frac{W_t(j)}{P_t} n_t(j) = \frac{W_t(j)}{W_t}\frac{W_t}{P_t} w_t(j)^{-\epsilon_w} n_t = w_t w_t(j)^{1-\epsilon_w} n_t.$$

The representative household is endowed with GHH preferences (Greenwood, Hercowitz, and Huffman (1988)) with external habit. She maximizes her life-time utility:

$$U_t = \max_{c_s, b_{s+1}, w_s(j)} \mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \frac{1}{1-\sigma_u}\left[c_s - hc_{s-1} - \frac{1/\varsigma}{1+1/\theta}\left(\int_0^1 w_s(j)^{-\epsilon_w} n_s dj\right)^{1+1/\theta}\right]^{1-\sigma_u}$$

subject to a period-by-period budget constraint,

$$\begin{aligned} c_t = &\int_0^1 (1-\tau_t) w_t w_t(j)^{1-\epsilon_w} n_t dj + r_t^K k_t + \frac{1+r_{t-1}}{\pi_{t+1}} b_t - b_{t+1} \\ &- q_t^K [k_{t+1} - (1-\delta)k_t] - \int_0^1 \frac{\varphi_w}{2}\left(\frac{w_t(j)}{w_{t-1}(j)}\frac{w_t}{w_{t-1}} - 1\right)^2 n_t(j) dj + \Pi_t, \end{aligned} \tag{35}$$

where b_{t+1} is risk-free bonds, which are in zero net supply and Π_t is the transfer of aggregate profits. Note that the household chooses its labor supply by selecting the real wage of intermediate labor input on the demand curve $n_t(j) = w_t(j)^{-\epsilon_w} n_t$, acting as a monopolistically competitive labor supplier of

$n_t(j)$. The first order conditions for the household, after imposing the symmetric equilibrium condition, i.e., $w_t(j) = W_t(j)/W_t = 1$ and denoting the shadow value of the budget constraint by λ_t, are given by the following four conditions:

$$\lambda_t = \left(c_t - hc_{t-1} - \frac{1/\varsigma}{1+1/\theta} n_t^{1+1/\theta}\right)^{-\sigma_u} \tag{36}$$

$$\begin{aligned}(1-\tau_t)w_t = &\frac{\epsilon_w}{\epsilon_w - 1}\frac{n_t^{1/\theta}}{\varsigma} - \frac{\varphi_w}{\epsilon_w - 1}\frac{w_t}{w_{t-1}}\left(\frac{w_t}{w_{t-1}} - 1\right) \\ &+ \frac{\beta}{\epsilon_w - 1}\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\lambda_t}\frac{w_{t+1}}{w_t}\left(\frac{w_{t+1}}{w_t} - 1\right)\frac{n_{t+1}}{n_t}\right]\end{aligned} \tag{37}$$

$$1 = \beta\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\lambda_t}\frac{r_{t+1}^K + (1-\delta)q_{t+1}^K}{q_t^K}\right] \tag{38}$$

and

$$1 = \beta\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\lambda_t}\frac{1+r_{t+1}}{\pi_{t+1}}\psi_t\right], \tag{39}$$

where ψ_t is Smets and Wouters (2007)'s risk premium shock associated with nominal bond holdings, which follows an AR(1) process given by

$$\log\psi_t = \rho_\Xi \log\psi_{t-1} + \sigma_\psi\epsilon_{\psi,t}, \quad \epsilon_{\psi,t} \sim \text{i.i.d. } N(0,1). \tag{40}$$

The GHH preferences make the marginal cost of providing labor in the wage Phillips curve (37) independent of marginal utility of consumption. (37) shows that in the non-stochastic zero trend inflation steady state in which $\pi_t = \pi_{t-1} = 1$ and $w_t = w_{t-1}$, the real wage is determined by tax wage, $(1-\tau)^{-1}$, markup wedge $\epsilon_w/(\epsilon_w - 1)$ and the marginal diutility of labor controlled by the Frisch elasticity of labor supply:

$$w = (1-\tau)^{-1} \cdot \frac{\epsilon_w}{\epsilon_w - 1} \cdot \frac{n^{1/\theta}}{\varsigma}.$$

2.3 Investment-Goods Producers

To endogeneize the price of capital, it is assumed that there is a continuum of competitive investment-goods producers. These firms use the following CRS technology to produce investment goods. The combination of complete competition and CRS technology implies that the scale of the firms is indeterminate. Hence, the existence of a representative firm. This firm solves

$$\max_{x_s}\mathbb{E}_t\sum_{s=t}^{\infty} Q_{t,s}\left\{q_s^K i_s - \left[i_s + \frac{\kappa}{2}\left(\frac{i_s}{i_{s-1}} - 1\right)^2 i_{s-1}\right]\right\}.$$

The first order condition to the above problem is given by an investment Euler equation:

$$q_t^K = 1 + \kappa\left(\frac{i_t}{i_{t-1}} - 1\right) - \mathbb{E}_t\left\{Q_{t,t+1}\frac{\kappa}{2}\left[\left(\frac{i_{t+1}}{i_t}\right)^2 - 1\right]\right\}. \tag{41}$$

Since newly produced capital and the old capital are identical, q_t^K becomes the price of capital.

2.4 Closing the Model

As was mentioned in the introduction, for a monetary policy, an asymmetric Taylor rule is considered:

$$r_t = r_t^* + \mathbf{1}(\pi_t \geq \pi^*)\rho_\pi^{(+)}(\pi_t - \pi^*) + \mathbf{1}(\pi_t < \pi^*)\rho_\pi^{(-)}(\pi_t - \pi^*), \tag{42}$$

where π^* is the inflation target of the central bank. In the baseline analysis, we impose parameter restrictions $\rho_\pi^{(+)} = \rho_\pi^{(-)}$. Eventually, however, we allow the two coefficients to differ in the analysis of optimal monetary policy.

Regarding the natural rate of interest rate r_t^*, two cases are considered: First, in the baseline, $r_t^* = r^*$, that is, the natural rate is equal to the steady state nominal interest rate; Second, alternatively $r_t^* = r_t^F$, that is, the natural rate is set equal to the real interest rate of the flexible price economy without nominal rigidity. The two cases are considered because it is unclear whether government spending shock should be considered a pure demand disturbance that does not affect the flexible economy or should be considered a real disturbance that affects the flexable economy. In the latter case, Woodford (2003) recommends central banks to set the natural rate of interest rate according to the flexible price economy. While this distinction does not make any material differences for the main conclusions of this paper, it can make an important difference for the sign and size of fiscal multiplier.

Note that an inertial component of monetary policy rule is not considered. This is because the inertial component of the Taylor rule often works as an inefficient source of propagation. The feedback rule also excludes a term for output or unemployment gap. This is for two reasons. The focus of the analysis is on the effects of demand shocks, and hence the model economy features the divine coincidence. This choice also avoids the strong assumption that the central bank can observe the evolution of the endogenous TFP (or the natural rate of unemployment) in real time.

Final-goods market clearing condition is given by

$$\begin{aligned} y_t = {} & c_t + i_t + g_t + (1 - M_{t-1})[1 - \Phi(\nu_t^E)]\gamma^E y + M_t \gamma^X y \\ & + \frac{\kappa}{2}\left(\frac{i_t}{i_{t-1}} - 1\right)^2 i_{t-1} + \frac{\varphi_p}{2}\left(\frac{\pi_t}{\pi_{t-1}^X} - 1\right)^2 y_t M_t^{-1} + \frac{\varphi_w}{2}\left(\frac{w_t}{w_{t-1}} - 1\right)^2 n_t \end{aligned} \tag{43}$$

where g_t is government spending, which follows an AR(1) process

$$g_t = \rho_g g_{t-1} + \sigma_g \epsilon_{g,t} \tag{44}$$

Table 1: Calibration forBaseline Economy

Parameter	Value
Capital share	$\alpha = 0.3$
Elasticity of substitution in CES aggregator (goods)	$\epsilon_p = 2.5$
Elasticity of substitution in CES aggregator (labor)	$\epsilon_w = 10$
Returns-to-specialization	$v = 0.0$
Investment adjustment cost	$\kappa = 5$
Entry cost	$\gamma^E = 0.2$
Fixed cost	$\gamma^X = 0.2$
Volatility of idiosyncratic shock	$\sigma_z = 0.33$
Depreciation rate	$\delta = 0.025$
Time discount factor	$\beta = 0.99$
Consumption habit	$h = 0.5$
Constant relative risk aversion	$\sigma_u = 2$
Labor disutility	$\varsigma = 1$
Frisch elasticity of labor supply	$\theta = 2$
Price/wage adjustment cost	$\varphi_p = \varphi_w = 100$
Price indexation	$\chi = 0.5$
Inflation target	$\pi^* = 1$
Monetary policy reaction coefficients	$\rho_\pi^{(+)} = \rho_\pi^{(-)} = 2$
Persistence of risk premium shock and government spending shock	$\rho_\psi = \rho_g = 0.8$
Volatility of risk premium shock	$\sigma_\psi = 0.001$
Volatility of government spending shock	$\sigma_g = 0.0038$

For a balanced budget in each period, the tax rate on labor income should be adjusted according to

$$\tau_t = \frac{g_t}{w_t n_t} \tag{45}$$

3 Results

3.1 Calibration

Table 1 summarizes the calibration of the baseline economy. The capital share α is set equal to 0.3 as is standard in business cycle research. The elasticity of substitution (ϵ_p) in the CES aggregator is set equal to 2.5. This value is relatively small compared with other literature, implying a high price markup. However, given that there are sunk entry cost and fixed cost of operation, relatively large market power is required. In contrast, we assume a much less degree of market power of labor and set $\epsilon_w = 10$. Since we assume symmetric price and wage adjustment cost, $\varphi_p = \varphi_w = 100$, the choices of ϵ_w and ϵ_w implies that the wage Phillips curve is steeper than the price Phillips curve. Price indexation is calibrated as 0.5. This choice implies that the the coefficient of the lagged inflation term in the Phillips curve is approximatedly 0.33, which is in the middle of the range [0.1, 0.5] of the estimates reported by Schorfheide (2008).

Regarding the investment adjustment cost (κ), 5 is chosen to produce roughly 3 times more volatile investment series compared with output volatility. The depreciation rate of capital (δ) is 0.025 as is standard in the literature. Both entry cost (γ^E) and fixed cost of operation (γ^X) are set equal to 0.2 of average output level, which is consistent with mid-range of the estimates reported by Dunne, Roberts, Klimek, and Xu (2007). The volatility of idiosyncratic technology (σ_z) is chosen as 0.33. Given the choices of market power, sunk and fixed costs, and the volatility of idiosyncratic shock, the equilibrium number of active firms in the non-stochastic steady state is determined as $M = 0.912$.

Time discounting factor (β) is equal to 0.99, which can be considered standard and the constant relative risk aversion (σ_u) is 2 with the habit formation equal to 0.5. Both parameter choices are standard. Labor disutility parameter (ς) is set equal to 1 and the Firsch elasticity of labor supply (θ) is set as 2, which is consistent with the estimate of labor supply at the macroeconomic level. With this calibrations of nominal rigidity, the standard deviation of inflation rate is 4 percentage points in the baseline economy at the annual frequency. Gross inflation target (π^*) is equal to 1 and for the baseline monetary policy, a symmetric monetary policy is chosen: $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 2$.

Finally, for both risk-premium shock and government spending shock, a moderate degree of persistence is chosen: $\rho_\psi = \rho_g = 0.8$. This is because, as will be show, the firm dynamics and the endogenous fluctuations in TFP creates enough amount of endogenous persistence for equilbrium quantities and prices. One standard deviation of risk premium shock (σ_ψ) is chosen such that $\sigma_\psi/(1+r) = 40$ basis points in the steady state at an annual frequency. One standard deviation of government spending shock (σ_g) is chosen such that $\sigma_g/y = 40$ basis points in the steady state. With this calibration of shock volatilities given persistence levels, the two shocks have roughly equal variance decomposition share for output.

3.2 Model Dynamics

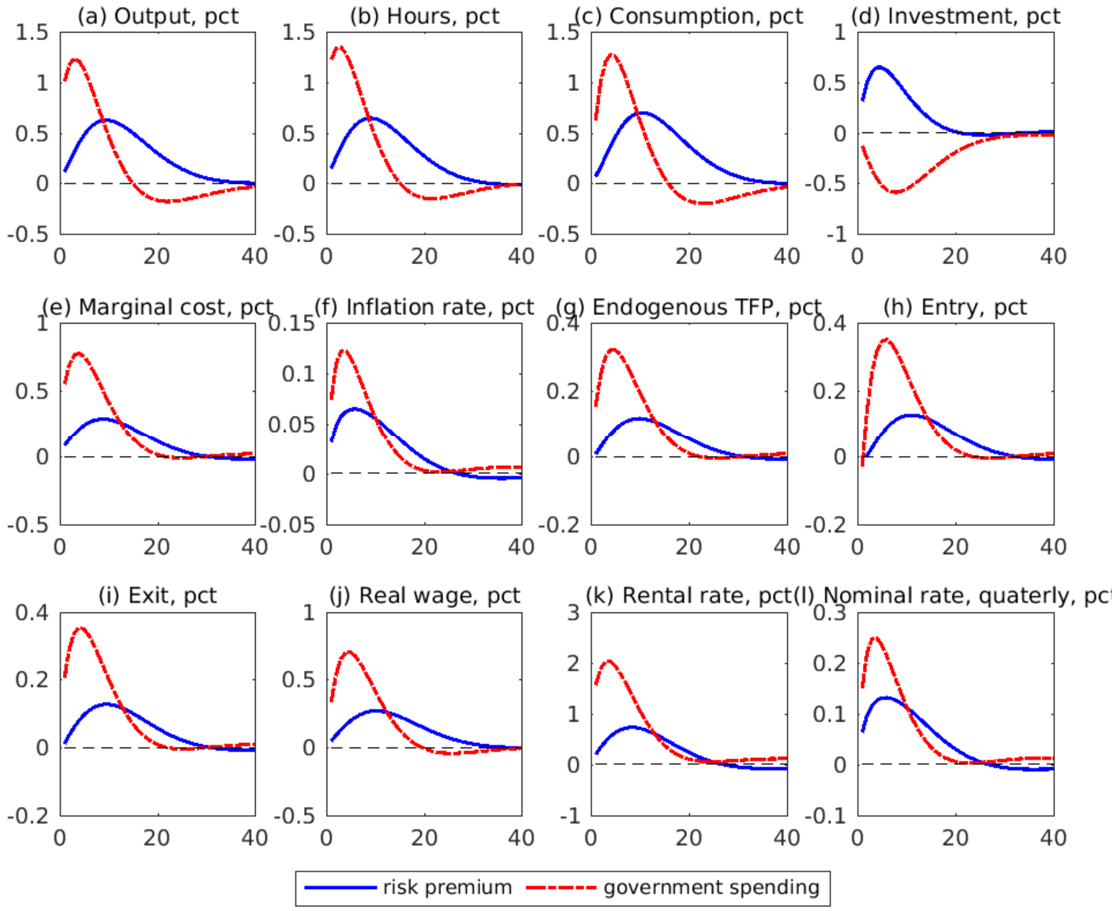
3.2.1 Impacts of Demand Shocks: Impulse Response Functions

In this subsection, I use impulse response functions to analyze the model dynamics under the demand shocks. Figure 1 shows the impacts of one standard deviation shocks to risk premium (blue solid lines) and to government spending (red dash-dotted lines). In response to the negative risk premium shock, output, hours, consumption and investment increase persistently as shown in panel (a)∼(d), which is standard as in the literature. What is unusual is the fact that the endogenous TFP increases persistently as shown in panel (g). The reason for this increase in TFP can be found in pabel (h) and (i): In response to the positive demand shock, firm entry increases persistently while firm exit is also increased persistently. Note that new entrant firms have higher quality than the average while exiting firms have lower quality than the average.

Such replacement of inefficient firms with efficient firms are reminiscent of "creative destruction" of Schumpeter (1994). The increase in firm exit during a demand driven boom is somewhat counter-intuitive provided that increase in product demand should improve the survival conditions for the marginal firms.[8] This is because the increase in product demand is dominated by the negative effects

[8] It is also counterfactual in that exit rate is asyclical in the data.

Figure 1: Impacts of Demand Shocks

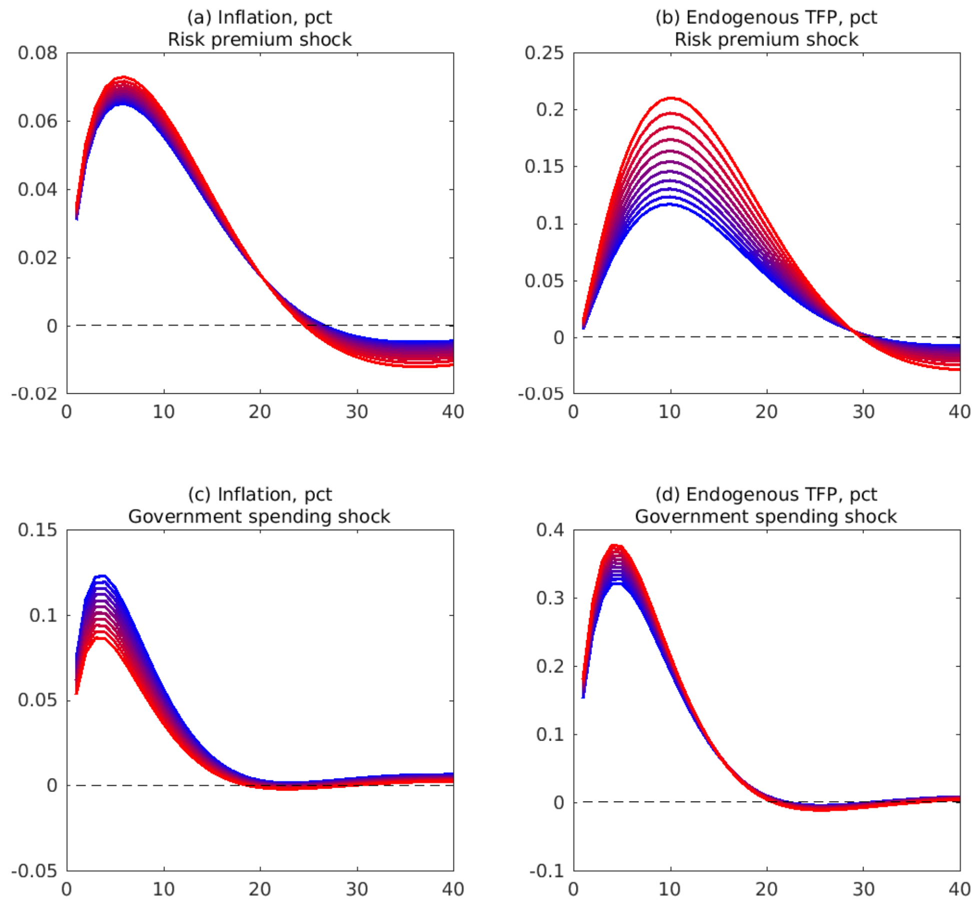


of increasing factor input prices shown panel (j) and (k). In the symmetric equilibrium, all firms produce the same quantity of goods, and the firms with relatively inefficient technology have to use disproportionately large amounts of production inputs, which makes the survival much more difficult for these firms. This is not the case for new entrant firms with relatively high productivity. Since these firms's technologies are highly efficient, their factor demands are not great in this symmetric equilibrium. As shown by (28), such creative destruction benefits all active firms.

The endogenous expansion in TFP makes the inflation pressure very much subdued as shown in panel (f). Despite the large increase in factor prices (especially the rental rate of capital), the increase in real marginal cost is mild in panel (e). This is because the endogenous increase in TFP offsets the large increases in factor prices as shown by (29), lowering the inflation pressure.

The dynamic pattern of the impact of government spending shock is different from the risk premium shock even though the two shock processes are modeled identically. The effects on aggregate output, hours and consumption (panel (a)∼(c)) are much more front-loaded compared with the effects of risk premium shock. This is because the spending shock works through goods market clearing condition (43) rather than through the variation in the dynamic profile of marginal utility of consumption. The spending shock is also different from the risk premium shock in that it crowds out private investment

Figure 2: Returns-to-Specialization and Endogenous TFP



Notes: The colors are arranged such that $v = 0.0$ is completely blue and as the v value is increased toward 0.1 the color becomes more redish and becomes complete red when $v = 0.1$.

as shown in panel (d). Except these two differences, a similar pattern appears to hold : the increase in endogenous TFP owing to the creative destruction; subdued inflation dynamics. These two mechanism allow for lenient monetary policy reactions, which then further strengthens economic expansion.[9]

3.2.2 Role of Returns-to-Specialization

The degree of endogenous expansion in supply capacity and the resulting relaxation in inflation pressure crucially depends on the degree of returns-to-specialization and the specific source of demand disturbances. To illustrate this point, in Figure 2, the degree of returns-to-specialization is varied from 0, which is the case in baseline calibration, to 0.1. In the figure, $v = 0$ is marked by complete blue color and as the degree of returns-to-specialization is increased toward 0.1, the color becomes more reddish, reaching complete red when $v = 0.1$. The figure focuses on two varables, inflation rate on the left panels and endogenous TFP on the right panels .

[9]The impulse response function analysis in this subsection is under the assumption that the central bank in the model conducts its monetary policy with a constant natural rate. As will be shown, the results for the spending shock are altered substantially under the alternative assumption about the time-varying natural rate.

First, consider the two upper panels, which show the impact of negative risk premium shock, a positive demand disturbance. As the degree of returns-to-specialization increases, the response of endogenous TFP becomes larger. However, despite the greater impact on TFP, the inflation response becomes more positive. This is because increase in factor prices dominates the increase in TFP, and as a result, greater degrees of returns-to-specialization lead to greater increases in real marginal cost. However, we note that additional increase in inflation pressure is pretty mild. Moving from $v = 0$ to $v = 0.1$ results in only about 1 basis poit difference in terms of the peak response of the inflation rate.

The lower panels show the case of government spending shock for the same range of returns-to-specialization. The same pattern emerges on the right panel: as the returns-to-specialization is increased, the endogenous TFP increases more and the difference between $v = 0$ and $v = 0.1$ at the peaks creates 0.1 percentage points difference in the TFP responses similar to the case of risk premium shock. However on the left panel, an opposite pattern appears: inflation response becomes smaller as the returns-to-specialization is strengthend and as a result, output expansion becomes stronger (not shown). This shows an interesting possibility that greater output expansions may be associated with milder inflation pressure if the increase in endogenous TFP more than offsets the increase in factor prices, a stronger case of inverse Say's Law: demand creates its own supply.

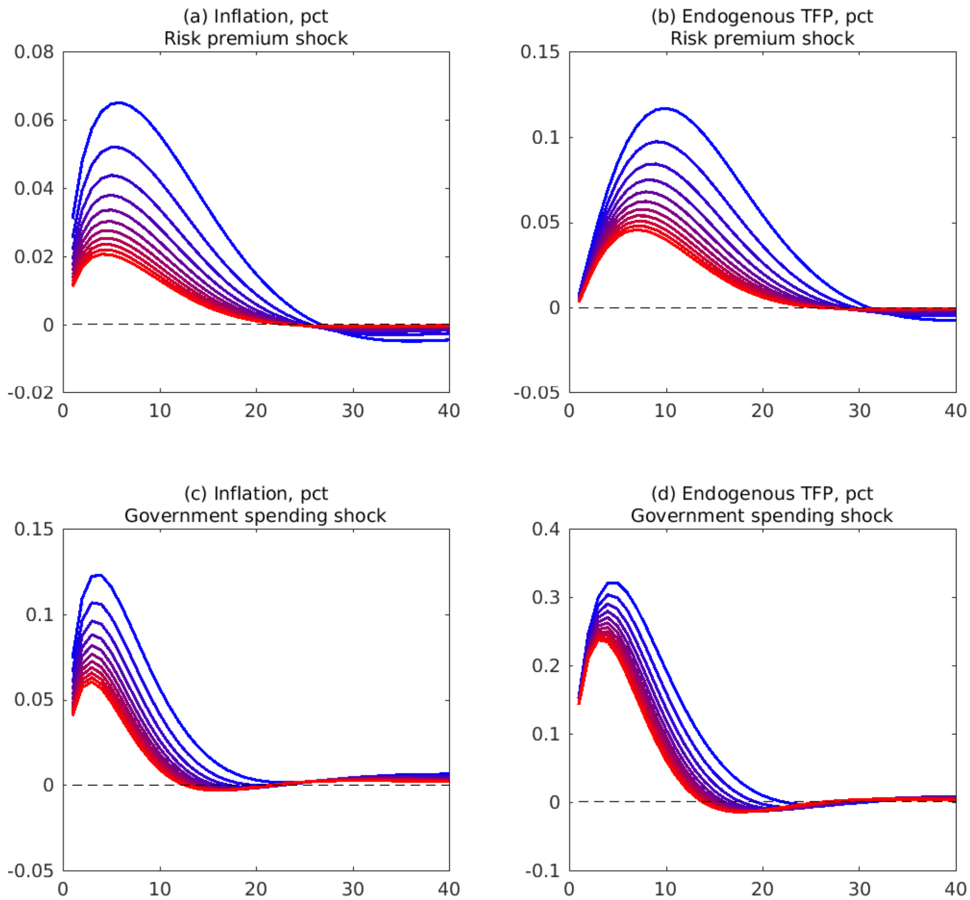
Before I move on to optimal monetary policy, I would like to emphasize that the hysteresis channel does not require the externality through positive returns-to-specialization. While it is true that positive returns-to-specialization may strengthen the degree of hysteresis effect, the existence of hysteresis channel does not require the presence of the externality.

3.3 Optimal Monetary Policy

The discussion of the model's impulse response functions indicates that the demand shocks considered in this paper has the aspects of both demand and supply shocks. What is optimal response of the central bank in this situation? In a canonical New Keynesian model, demand shocks are objects of complete stabilization especially. The central bank can achieve this by simply increasing inflation reaction coefficient indefinitely as shown by, for example, Clarida, Galí, and Gertler (1999), who show that as the the weight given to the unemployment gap in the loss function of the central bank approaches zero, which is equivalent with the importance of cost-push shock converging to zero, optimal inflation reaction coefficient should approach the infinity in a discretionary optimization without commitment.

Can one say the same thing in our environment in which "*changes in aggregate demand can have an appreciable, persistent effect on aggregate supply*" (Yellen (2016))? Below it is shown that to the extent that hysteresis is present, complete stabilization of demand shock is suboptimal when the cenral bank is constrained to adopt a symmetric policy rule. Furthermore, it is shown that when the central bank is allowed to use an asymmetric policy rule, the central bank should react to positive demand shock in a dovish manner while it has to react to negative demand shock much more vigorously, providing massive degree of monetary stimulus.

Figure 3: To Kill Inflation or to Save Hysteresis



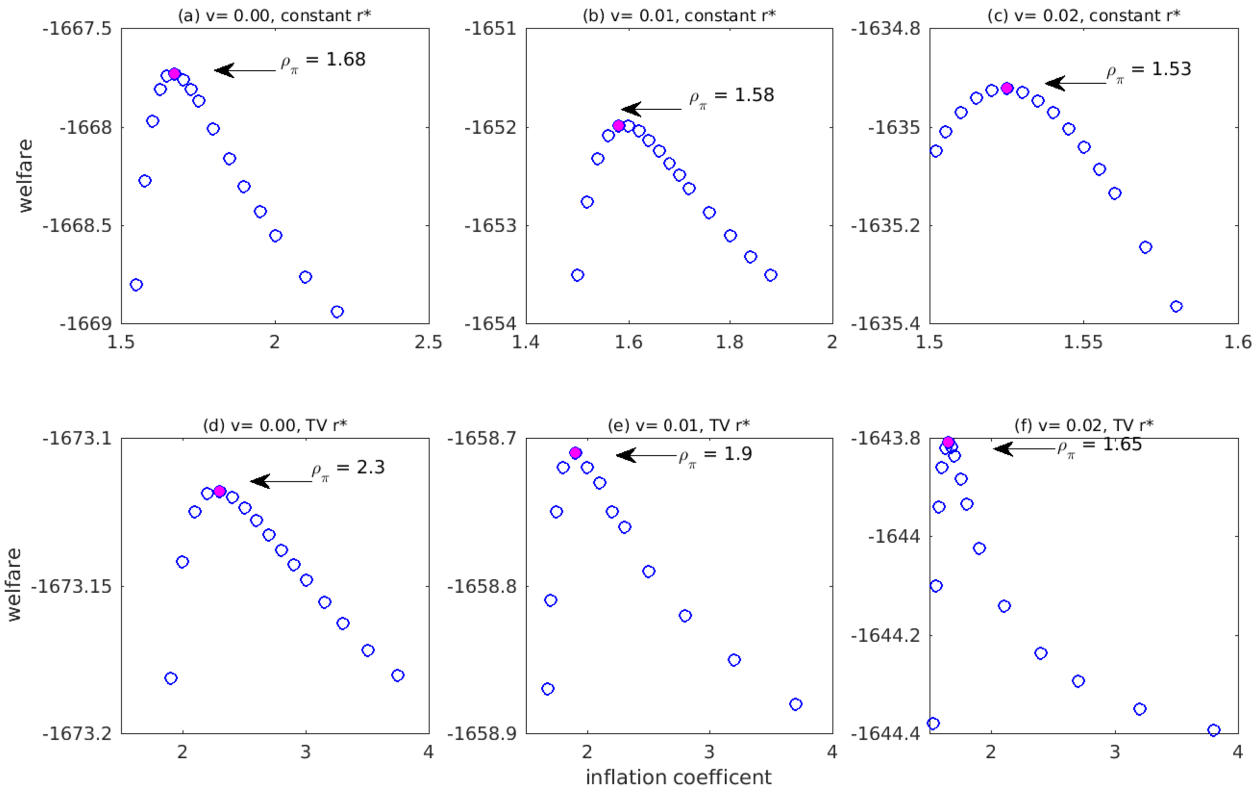
Notes: The colors are arranged such that $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 2.0$ is completely blue and as the $\rho_\pi^{(+)} = \rho_\pi^{(-)}$ value is increased toward 4.0 the color becomes more redish and becomes complete red when $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 4.0$.

3.3.1 Symmetric Monetary Policy

Figure 3 shows the effects of increasing inflation coefficient gradually from 2 to 4 in a setting of symmetric monetary policy rule with $\rho_\pi^{(+)} = \rho_\pi^{(-)}$. The figure is organized in a similar way that Figure 2 is organized: the left pannels show the responses of inflation rate and the right pannels the responses of endogenous TFP; the upper panels are the case of risk premium shock while the lower pannels are the case of government spending shock. When $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 2.0$, the color is complete blue. The color is gradually turning reddish as $\rho_\pi^{(+)} = \rho_\pi^{(-)}$ approaches 4.0, and the color becomes complete red when $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 4.0$.

The left pannels show that as the monetary policy becomes more hawkish, the inflation responses to the demand shocks become smaller. There is nothing special about this result. However, the right panels show that the same hawish monetary policy rules also kill the expansions of endogenous TFP in both cases of risk premium shock and fiscal spending shock. Stabilizing inflation is certainly welfare improving. However, shutting down the expansion of supply capacity may be detrimental for the welfare of the representative agent. In other words, in this economy, appointing Rogoff (1985)'s

Figure 4: Monetary Policy Conservatism and Welfare



Notes: The welfare of the representative agent is approximated by the second-order perturbation method. The volatilities of risk premium shocks and government spending shocks are calibrated such that each shock accounts for 50 percent of output variance decomposition in the long-run.

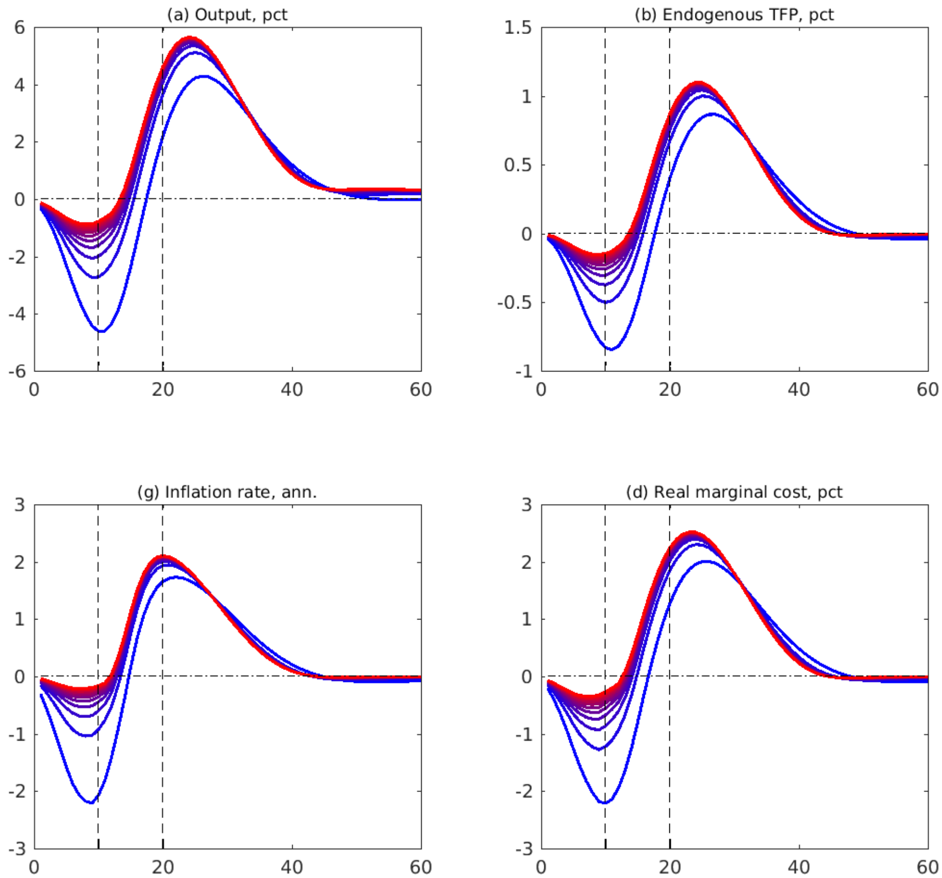
"conservative central banker" may harm the welfare of the society. Which of these opposing effects dominates is the important question at hand.

Figure 4 illustrates how the welfare changes as the monetary policy becomes more conservative. If the gains from stabilization of inflation dominates the loss from shutting down the hysteresis channel, the welfare of the representative agent should be maximized at large numbers for $\rho_\pi^{(+)} = \rho_\pi^{(-)}$.[10] The upper three panel shows the cases when the monetary authority adopts a constant natural rate of interest rate because the government spending shock does not affect the flexible economy while the lower three panel shows the cases when the monetary authority adopts a time-varying natural rate of interest because the spending shock does affect the flexible economy (Woodford (2003)). The left, middle and right panels show the cases with the returns-to-specialization equal to 0.00, 0.01 and 0.02.

First, consider the upper panels (a)∼(c). These three panels show that the welfare maximizing (symmteric) monetary policy coefficients are finite and in fact very small. In other words, the central bank that aims to maximize the welfare of the representative agents finds complete stabilization of demand shocks suboptimal and let the demand shocks follow through such that the hysteresis effects can be conserved. Furthermore, the comparisons of panels (a)∼(c) reveals that as the degree of

[10]The welfare is approximated by the second order perturbation method. Later when the asymmetric policy rule is used, the expanded path approach is used to handle the kink in the monetary policy rule.

Figure 5: Asymmetric Monetary Policy



Notes: $\rho_\pi^{(+)} = 2$ for all lines. $\rho_\pi^{(-)}$ is raised from 2 to 12. The color of the lines is completely blue when $\rho_\pi^{(-)} = 2$, becomes more reddish as the value of $\rho_\pi^{(-)}$ becomes closer to 12, and becomes completely red when $\rho_\pi^{(-)} = 12$. In the first 10 periods, the economy is subject to a sequence of positive risk premium shock, ρ_ψ[1.0 1.0 1.5 1.5 1.5 1.5 1.0 0.5 0.25 0.125]. In the next 10 periods, the economy is subject to a sequence of negative risk premium shocks with exactly the same absolute values but with negative signs.

returns-to-specialization increases, optimized inflation coefficients become even smaller. These results indicate that the benefits from letting the hysteresis effects follow through can be far greater than the costs of having fluctuations in inflation and output.

Next, consider the bottom three panels (d)$\sim$(f). The experiment is identical with the upper panels except that the government spending shock is modeled as a real shock that affects the flexible economy and the central bank is assumed to adjust the natural rate of interest rate according to the real rate in the flexible economy such that when inflation gap is completely stabilized, the nominal interest rate is set according to the real rate of the flexible economy. Our two findings are robust against this alternative assumption: welfare maximizing inflation coefficient is finite and becomes smaller as the degree of returns-to-specialization gets stronger.

3.3.2 Asymmetric Monetary Policy

So far we have assumed that the policy response of the central bank is symmetric against positive and negative demand shocks. However, the results of welfare maximizing symmetric monetary policy raises an entirely different question: if there are benefits from letting the hysteresis channel operate in the economy, should not the central bank selectively choose to stabilize the impacts of negative demand shocks only? While positive demand shocks may expand supply capacity through hysteresis, negative demand shocks do the opposite: negative demand shocks may lead to contraction in supply capacity. If so, can the central bank allow the former while preventing the latter selectively?

To illustrate the effects of such opportunistic monetary policy, in Figure 5, 11 monetary policy rules are used against a sequence of positive risk premium shocks (negative demand shocks) for the first 10 periods. These shocks are then alternated with a sequence of negative risk premium shock (positive demand shocks) for the second 10 periods.[11] The baseline monetary policy is specified as $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 2$. Alternative monetary policy rules are specified as $\rho_\pi^{(+)} = 2$ and $\rho_\pi^{(-)} \in \{3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, that is, the coefficient for positive inflation gap is held constant at 2 while the coefficient for negative inflation gap is gradually increased toward 12. The color of the lines is complete blue when $\rho_\pi^{(+)} = \rho_\pi^{(-)} = 2$, becomes increasingly more reddish as $\rho_\pi^{(-)}$ is raised toward 12, and becomes complete red $\rho_\pi^{(-)} = 12$.[12]

The blue lines in the figure (located at the bottom of each panel) show the response of symmetric monetary policy. In this case, the economy responds symmetrically to negative and positive demand shocks. As the monetary policy becomes more asymmetric, the downside risk of the economy is more limited. For example, under the symmetric monetary policy, out declines about 4.5 percent in panel (a). As the monetary policy becomes more asymmetric, such large output decline is prevented and maximum decline is limited below 1 percent. As a result, negative hysteresis effect is also limited as shown by the contraction of the endogenous TFP in panel (b). Consequently downward pressure on inflation and real marginal cost shown in panels (c) and (d) are much more subdued.

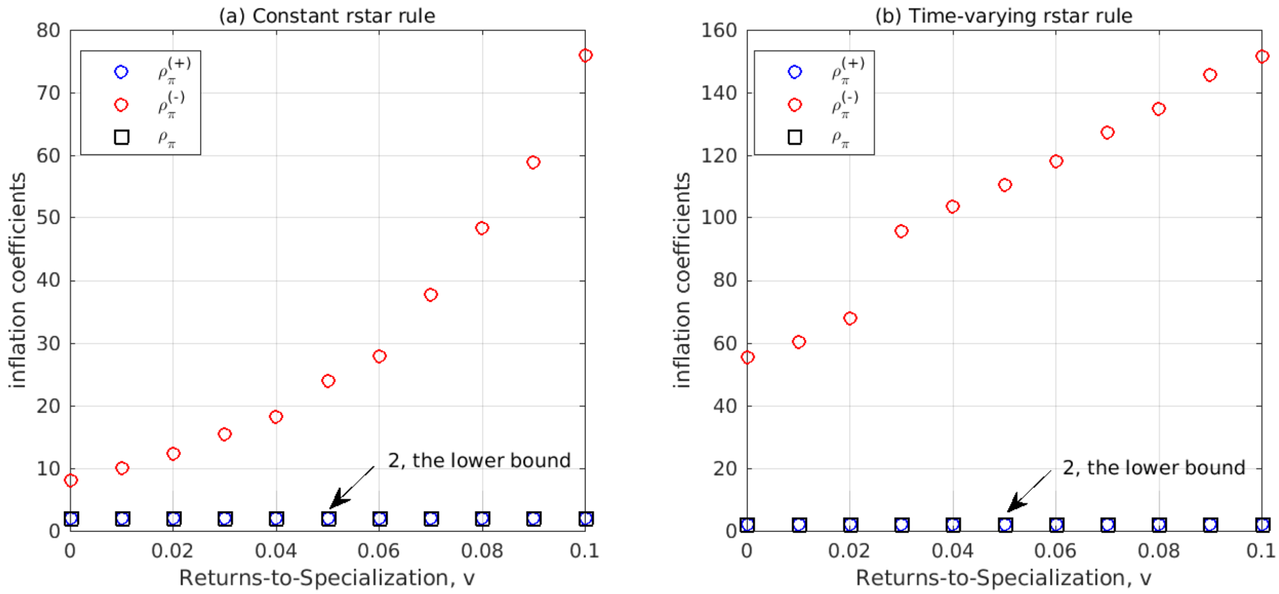
Note that the policy switch operates through observing positive and negative inflation gaps rather than observing the shocks or the estimates of resource slack.[13] Even after negative demand shocks turn into positive demand shocks at the eleventh period, monetary policy switch is not executed since the negative inflation gap persists. So exact time in which monetary policy switch happens is different for each alternative monetary policies shown in the figure. In general, the switch (from $\rho_\pi^{(-)}$ to $\rho_\pi^{(+)}$) is sooner when $\rho_\pi^{(-)}$ is larger (more asymmetric) as the negative inflation gap is less persistent

[11] Exact sequence is given by $\sigma_\psi \times$ [1.0 1.0 1.5 1.5 1.5 1.5 1.0 0.5 0.25 0.125 -1.0 -1.0 -1.5 -1.5 -1.5 -1.5 -1.0 -0.5 -0.25 -0.125].

[12] Note that such an economy with a kinked monetary policy rule cannot be simulated by conventional perturbation method. I use so called extended path approach of Adjemian and Juillard (2011), which provides a fully nonlinear solution based on solving two-point boundary value problems. The solution satisfies nonlinear model equations exactly even when the model is kinked or has occasionally binding constraint. The method improves upon popular piecewise linear solution such as Guerrieri and Iacoviello (2015) in that the model is not perturbed in any order in any place (for instance arround the steady state). See Appendix D for more detailed information.

[13] This is the reason why the experiment is realistic in the sense that monetary policymakers rarely observe directly the shocks that hit the economy in real time.

Figure 6: Optimized Asymmetric Monetary Policy



Notes: $\rho_\pi^{(+)}$ (blue circles) and $\rho_\pi^{(-)}$ (red circles) are chosen from an interval [2,1000] to maxize the welfare of the representative agent. ρ_π (black squares) is the case when the optimization is subject to a constraint, $\rho_\pi^{(+)} = \rho_\pi^{(-)}$.

endogenously.[14]

Figure 5 makes it clear that the more asymmetric the monetary policy becomes, the more right skewed the distributions of the economy becomes, changing the means of equilibrium quantities and prices. For instance, while the theoretical model has zero trend inflation rate in the non-stochastic steady state, the asymmetric monetary policy rules introduce inflation bias: $\mathbb{E}[\pi - 1] > 0$, that is, the unconditional mean inflation rate is strictly positive. Can such a policy rule be welfare improving? This may be the case if such policy improves the mean of production capacity sufficiently enough by inducing highly right-skewed distribution for endogenous TFP.

To answer this question, the asymmetric monetary policy rule is optimized to maximize the welfare: $\mathbb{E}[U] \approx 1/T \sum_{t=1}^{T} U_t$.[15] In practice, $\rho_\pi^{(+)}$ and $\rho_\pi^{(-)}$ are chosen from an interval, $[2, 1000]$. Note that to the extent that such asymmetric policy rules are detrimental to the welfare of the household, the optimization program can always avoid such asymmetry by choosing $\rho_\pi^{(-)} = \rho_\pi^{(+)}$. Also note that I do not force a constraint $\rho_\pi^{(-)} \geq \rho_\pi^{(+)}$ in the optimization. If the optimization program finds $\rho_\pi^{(-)} < \rho_\pi^{(+)}$ optimal, it will do so.

Figure 6 shows the optimization results. There are three types of markers shown in the figure: blue circles for $\rho_\pi^{(+)}$; red circles for $\rho_\pi^{(-)}$; black squares for symmetric monetary policy ρ_π. The left panel shows the cases with a constant natural rate of interest while the right panel shows the cases

[14] Once all switches happen in the 11 cases, the monetary policy rules are identical across the 11 cases. However, the economy behaves slightly differently despite the identical monetary policy rule. This is mainly because of the backward looking component of price Phillips curve.

[15] In this experiment, I cannot compute the analytical moments due to the kink in monetary policy rule. Instead, I simulate the economy with identical set of demand shocks to assess a particular policy rule. T is set equal to 5,000. Note that $\mathbb{E}[U]$

with a time-varying natural rate of interest. Each panel shows 11 cases of different degrees of returns-to-specialization with $v = 0$ shown at the beginning of the horizontal axes.

First, consider the left panel. In the baseline economy with constant returns-to-specialization, the optimal degree of asymmetry is sizable: $\rho_\pi^{(+)}$ is optimized at the lower bound of the optimization; a large number close to 10 is chosen for $\rho_\pi^{(-)}$. In words, the welfare optimizing monetary policy chooses the most dovish reaction coefficient against positive demand shocks (more accurately, against positive inflation pressure), but chooses the most hawkish reaction coefficient againt negative demand shock as if the cenral banker were a "conservative central banker" only to the downside risk.

As the degree of increasing returns-to-specialization goes up, the size of positive externality increases. Consequently, the benefits from the asymmetry in monetary policy grows and the optimized negative inflation gap coefficient is monotonically increasing in the returns-to-specialization. At $v = 0.1$, the optimized $\rho_\pi^{(-)}$ is close to 80. In contrast, positive inflation gap coefficient is always optimized at the lower bound.

Note that it is not just the blue circles that are located at the bottom of the figure. The black circles are also located at the bottom. This means that *symmetric* monetary policy coefficient is always optimized at the lower bound. This suggests that when such a *Janus-faced* central banker is not available, then it is optimal to appoint a central banker that is dovish all the time in this economy.

Next, the right panel shows that the main results shown in the left panel are completely robust when we adopt an alternative assumption for the natural rate. Even when government spending shock is considered a real shock that affects the flexible economy and the monetary authority adjusts its natural rate according to the real rate of the flexible economy, the same results are obtained: positive inflation gap coefficient is always chosen at the lower bound; negative inflation gap coefficient is much larger than the positive one and the degree of asymmetry increases monotonically in the degree of returns-to-specialization; when symmetry is imposed, the welfare maximizing inflation coefficient is always found at the lower bound. An added feature in this case is that the optimized values for $\rho_\pi^{(-)}$ are much larger. Even in the baseline economy without positive returns-to-specialization, the optimized value is close to 60.

Table 2 compares key moments of the model economy under optimized symmetric and asymmetric monetary policy rules. Line 1 (symmetric) and 2 (asymmetric) compare the mean aggregate outputs when the degree of returns-to-specialization is varied. Line 3 shows that mean aggregate output can be increased substantially by switching from symmetric policy rule to asymmetric rule. The differences from the symmetric monetary policy rules can reach as large as 2 percent. Furthermore, line 4∼6 show that by switching to the asymmetric monetary policy, the central bank reduce output volatility 20∼40 percent. As suggested by Figure 5, such reduction in output volatility is due to the elimination of downside risks. In summary, the asymmetric monetary policy can change the overall properties of output distribution, increasing mean, but reducing large amount of volatility.[16]

[16]Note that even when the monetary policy is symmetric, the distributions of endogenous variables are not completely symmetric although underlying shocks are drawn from Gaussian process. This is because the model economy is nonlinear and the solution method is fully nonlinear. Hence, the positive mean inflation rates under the symmetric monetary policy rules can be thought of as the results of nonlinearity of the model economy.

Table 2: Mean and Standard Deviation of Output and Inflation Rate: Constant Natural Rate

	Returns-to-Specialization, v				
	0.00	0.02	0.04	0.06	0.08
1. E(y) under optimized symmetric policy, A	0.80	0.82	0.84	0.85	0.87
2. E(y) under optimized asymmetric policy, B	0.81	0.83	0.84	0.86	0.88
3. (B/A-1) x 100	1.13	1.20	1.25	1.28	1.23
4. SD(y) under optimized symmteric policy, C	3.46	3.58	3.73	3.95	4.08
5. SD(y) under optimized asymmetric policy, D	2.78	2.70	2.63	2.56	2.50
6. (D/C-1) x 100	-19.71	-24.51	-29.55	-35.09	-38.71
7. E($\pi - 1$) under optimized symmetric policy, pct, E	0.11	0.09	0.07	0.06	0.05
8. E($\pi - 1$) under optimized asymmetric policy, pct, F	0.66	0.65	0.61	0.58	0.54
9. SD(π) under optimized symmetric policy, pct, G	1.59	1.53	1.48	1.44	1.38
10. SD(π) under optimized asymmetric policy, pct, H	1.04	0.92	0.82	0.73	0.65
11. (H/G-1) x 100	-34.59	-39.64	-44.38	-49.12	-52.77

Notes: Asymmetric monetary policy coefficients, $\rho_\pi^{(+)}$ and $\rho_\pi^{(-)}$, and symmetric monetary policy coefficient, ρ_π are chosen from an interval [2,1000] to maximize the welfare of the representative agent. Inflation statistics are at the quarterly frequency.

Table 3: Mean and Standard Deviation of Output and Inflation Rate: Time-Varying Natural Rate

	Returns-to-Specialization, v				
	0.00	0.02	0.04	0.06	0.08
1. E(y) under optimized symmetric policy, A	0.81	0.82	0.84	0.85	0.87
2. E(y) under optimized asymmetric policy, B	0.82	0.84	0.85	0.87	0.88
3. (B/A-1) x 100	1.63	1.78	1.86	1.88	1.93
4. SD(y) under optimized symmteric policy, C	2.16	2.35	2.58	2.88	3.30
5. SD(y) under optimized asymmetric policy, D	1.51	1.58	1.67	1.78	1.89
6. (D/C-1) x 100	-30.32	-32.46	-35.13	-38.41	-42.91
7. E($\pi - 1$) under optimized symmetric policy, pct, E	0.06	0.07	0.07	0.07	0.08
8. E($\pi - 1$) under optimized asymmetric policy, pct, F	1.01	0.44	0.43	0.42	0.40
9. SD(π) under optimized symmetric policy, pct, G	1.59	1.53	1.48	1.44	1.38
10. SD(π) under optimized asymmetric policy, pct, H	1.04	0.92	0.82	0.73	0.65
11. (H/G-1) x 100	-45.37	-46.71	-48.60	-50.98	-54.19

Notes: Asymmetric monetary policy coefficients, $\rho_\pi^{(+)}$ and $\rho_\pi^{(-)}$, and symmetric monetary policy coefficient, ρ_π are chosen from an interval [2,1000] to maximize the welfare of the representative agent. Inflation statistics are at the quarterly frequency.

Line 7 and 8 show the mean inflation rates of the economy under the optimized symmetric and asymmetric monetary policy rules. Two important features emerge from line 7 and 8. First, the asymmetric monetary policy rules introduce sizable inflation bias. For instance, in the case of baseline returns-to-sepecialization, the asymmetric rule increases inflation bias as much as 55 basis points at the quarterly frequency, an equivalent of 220 basis points at the annual frequency. It is a remarkable result that despite such a large increase in inflation bias, the asymmetric policy improves the welfare. This result confirms the conjecture that Blanchard, Cerutti, and Summers (2015) made in discussing the monetary policy implication of hysteresis phenomenon :

> "The implication is straightforward, namely that monetary policy should react more strongly to output movements, relative to inflation. For example, by being more aggressive early on, this would reduce the increase in unemployment, and by implication, reduce the increase in the number of long term unemployed. It also implies that stabilizing inflation is definitely not the optimal policy" – (Blanchard, Cerutti, and Summers (2015), p. 25)

Second, both inflation biases under the symmetric and asymmetric policies tend to decline as the degree of returns-to-specialization increases. This is one sense in which the inverse Say's Law, "demand creates its own supply", can be interpreted. Furthermore, line 9∼11 show that adopting the asymmetric policy rule can result in large reduction in inflation volatility on the order of 35∼53 percent. The fact that a stronger returns-to-specialization allows for a larger reduction in inflation volatility is another sense in which the inverse Say's Law can be interpreted. Again, as suggested by Figure 5, the reduction in inflation volatility is entirely due to the elimination of downside volatility.

Table 3 provides the results of the same experiment in an environment where government spending shock is modeled as a real shock that affects the flexible economy and the monetary authority adjusts the natural rate according to the real rate in the flexible economy. This table shows that the results obtained in Table 2 are completely robust under this environment as well. There are minor differences. The increases in the mean of aggregate output tend to be slightly smaller while the inflation bias introduced by the asymmetric policy tend to be larger. The reduction in both inflation and output volatilities due to the asymmetric policy rule much larger in this alternative environment.

3.4 Fiscal Multiplier under Optimal Monetary Policy

We finish our analysis by discussing the implications of the hysteresis for fiscal multiplier. DeLong and Summers (2012) has made an argument that in a depressed economy with nominal interest rate stuck at the liquidity trap, a small degree of hysteresis will make fiscal multiplier much larger than 1, in particular, if such spending shock is not met with monetary policy tightening. Their argument was met by a criticism that their model is based on the first principles of optimization (see Ramey (2012)).

I show that the presence of hysteresis in a model that is based on the first principles can make fiscal multiplier much larger than 1 even when the spending is financed by distortionary taxation provided that the government spending shock is modeled as a pure nominal shock that does not affect the flexible economy and the monetary authority does not adjust its natural rate accordingly. Furthermore, the argument does not even rely on the assumption of the economy being in a depression or nominal interest rate being at the zero lower bound.

The fiscal multiplier is defined as

$$FM \equiv \frac{\sum_{t=1}^{100} \Delta y_t}{\sum_{t=1}^{100} \Delta g_t}.$$

Table 4: Fiscal Multipliers

	Returns-to-Specialization, v				
	0.00	0.02	0.04	0.06	0.08
FM under optimized symmetric rule, $r_t^* = r$	3.72	4.10	4.44	4.74	4.84
FM under optimized asymmetric rule, $r_t^* = r$	3.72	4.10	4.44	4.74	4.99
FM under optimized symmetric rule, $r_t^* = r_t^F$	-2.16	-1.96	-1.75	-1.52	-1.30
FM under optimized asymmetric rule, $r_t^* = r_t^F$	-2.16	-1.96	-1.74	-1.52	-1.29

where Δy_t is the impulses response of output and government spending to the fiscal shock $\epsilon_{g,t}$. Table 4 reports the fiscal multipliers under two circumstances: First, the government spending shock is considered a pure nominal shock that does not affect the flexible economy and accordingly, the central bank does not adjust its natural rate ($r_t^* = r$); Second, the government shock is considered a real shock that affects the flexible economy and the central bank sets its natural rate according to the real interest rate in the flexible economy ($r_t^* = r_t^F$). In both cases, we assume that asymmetric monetary policy rule is optimized in each case both under symmetric rule and under asymmetric rule.

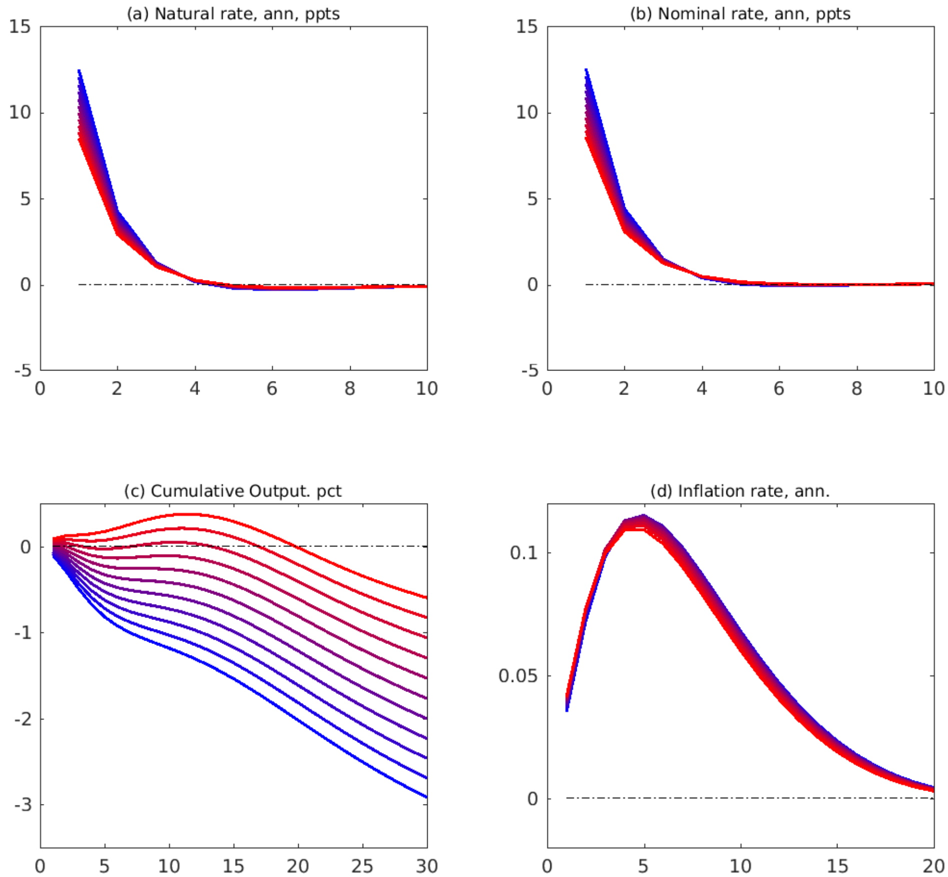
The first line of Table 4 shows the fiscal multipliers in the case of constant natural rate under symmetric monetary policy rule for various degrees of returns-to-specialization. The first column shows that the fiscal multiplier can be as large as 3.7 in the baseline economy with constant returns-to-specialization. Furthermore, the fiscal multiplier is shown to increase monotonically as the degree of returns-to-specialization strengthens. The second line indicates that there are no material differences in the magnitudes of fiscal multiplier regardless of monetary policy taking symmetric forms or not. This is because inflation gap in this experiment never turns negative and the switch to asymmetric rule is almost never called in. These results suggest that if hysteresis mechanism is at work, the fiscal multiplier can be indeed substantially larger than 1 as the increase in demand due to fiscal expansion leads to improvement in the average quality of active firms and all active firms benefits from endogenously improved TFP. Conversely, the same results also suggest that the fiscal austerity dominating the political dialogues in advanced economies can result in harmful effects on the growth potentials of these economies through hysteresis channel.

The third and fourth lines of Table 4 tell a different story: when the central bank sets the natural rate according to the flexible rate at the time of fiscal expansion, fiscal multiplier can turn negative and the magnitudes of negative effects can be quite substantial. How can the two opposite results be reconciled with each other? Since the only difference between the two cases is the monetary policy setting, the answer must be found there.

To answer this question, I show, in Figure 7, the responses of the natural rate (the real interest rate in the flexible economy, panel (a)) and the nominal rate (panel (b)). In both cases, the responses die out rather quickly and the panels are drawn only for the first 10 periods for this reason. Each panel shows 11 cases of different degrees of returns-to-specialization from $v = 0$ to $v = 0.1$ with the color of the lines gradually transitioning from complete blue ($v = 0$) to complete red ($v = 0.1$).

In response to the fiscal expansion, the natural rate jumps substantially as much as more than 12

Figure 7: Natural Rate and Fiscal Multiplier



Notes: $\rho_\pi^{(+)}$ (blue circles) and $\rho_\pi^{(-)}$ (red circles) are chosen from an interval [2,1000] to maxize the welfare of the representative agent. ρ_π (black squares) is the case when the optimization is subject to a constraint, $\rho_\pi^{(+)} = \rho_\pi^{(-)}$.

percentage points in panel (a). Why does the real interest rate jump in response to fiscal spending? This is the effect of general equilibrium restriction: given that the aggregate supply is fixed, in order to allow for fiscal expansion, private investment and consumption must decline to create resource slack, and the only way to create such a slack especially when there is no increase in labor supply through wealth effect due to the GHH preferences, the equilibrium real interest rate must jump.

When the central bank in the economy with nominal rigidity sets its natural rate according to the real interest rate of the flexible price economy, the nominal interest rate set by the Taylor rule also follows a similar time path because any movement in the nomial rate due to small inflation gap as shown in panel (d) is dominated by the natural rate component. Overall initial impact on aggregate output can be negative or positive depending on whether the decline of private spending or the increase in public spending dominates. The cummulative output paths shown in panel (c) shows that the former dominates when the degree of returns-to-specialization is particularly low. As the degree of returns to specialization becomes relatively high, the initial impact on output turns positive. Even in these cases, the cumulative effect on output eventually turns negative.[17]

[17]The negativity of fiscal multiplier in this case is not a general phenomenon for two reasons. First, If the preferences are

The results shown in Table 4 raises a question: Should the government spending shock be modeled as demand shock or as real shock? Clearly there are two strands of traditions in New Keynesian literature. Smets and Wouters (2007) classify the government spending shock as a demand shock, in fact, the most important source of output variation in one year horizon. Woodford (2003) classifies it as a real shock that affects the flexible price economy. The results shown in Table 4 suggest that if government spending shocks can be charaterized as nominal spending shocks, active hysteresis effects can make fiscal expansion much stronger than in conventional models.[18]

4 Conclusion

The traditional notion of stabilization implicitly assumes that there exists a given level of maximum production capacity that is determined by non-demand, real factors and the role of the stabilization policy is to minimize the deviation of the actual economy from the capacity. This notion does not fit very well the most recent experience on inflation front: where there seemed to be large resource gaps, we failed to see large deflation pressure, and where there seemed to be little room for more expansion, we failed to observe rising inflation pressure. It is in this context that Yellen (2016) has called for a new research agenda in macroeconomics: reconsidering the traditional dichotomy between aggregate demand and supply.

This paper is an attempt to respond to such a call. A structural general equilibrium model of hysteresis based on a simple model of firm dynamics is constructed and monetary policy implications of hysteresis phenomenon is studied. The main conclusion is that if hysteresis effects exist, complete stabilization of demand shocks is suboptimal. Another main conclusion is that if hysteresis provides a channel through which the history of demand may affect today's supply capacity, the attitudes of policymakers to positive and negative demand shocks should be fundamentally asymmetric: positive demand shocks should be let run through since such demand shocks build their own supply; negative demand shocks should be met by strong stimulus policy before the demand shocks destroy their own supply. The welfare gains from adopting such asymmetric policy tools are shown to be substantial.

separable for consumption utility and labor disutility, labor supply has wealth effect such that the decline of consumption increases the marginal utility of wage earning, which can offset the negative effect of increasing tax on labor supply. Second, if the Frisch elasticity of labor supply is sufficiently elastic, the fiscal multiplier can turn positive even without separable utility. However, this requires the elasticity to be greater than 5, which can be considered unrealistically large.

[18]The differences in real effects of fiscal spendings as nominal shocks and as real shocks can be interpreted in the following way. When the fiscal spending shock is modeled as a nominal shock, an implicit assumption is that the central bank targets its nominal interest rate according to the Taylor rule without changing the natural rate in response to the fiscal shock. While the fiscal spending is financed by a distortionary taxation in the model, one alternative way of modeling it is money-financed spending. In this case government spending is financed by increased reserve in the system. Since the central bank targets the nominal rate, it wants to drain the increased reserve by an open market operation, holding the interest rate close to target. This is one way to understand the constant natural rate. The case of government spending shock as a real shock can be interpreted as bond financed fiscal spending. If government issues more bonds, the public relinquishes its reserves to obtain newly issued government bonds. If the central bank does not targets its nominal rate, reduced supply of reserve leads to an increase in interest rate provided that money demand is declining in nominal rate, which can be modeled through money-in-utility specification. The increase in the nomial rate due to the increase in natural rate can be interpreted as tantamount to the bond financing case. I leavc the task of describing two alternative funding methods for government spending as future works.

References

Adjemian, S., and M. Juillard (2011): "Accuracy of the Extended Path Simulation Method in a New Keynesian Model with Zero Lower Bound on the Nominal Interest Rate," Unpublished manuscript, Université du Maine.

Baldwin, R., and P. Krugman (1989): "Persistent Trade Effects of Large Exchange Rate Shocks," *The Quarterly Journal of Economics*, 104(4), 635–654.

Ball, L. (1999): "Aggregate Demand and Long-Run Unemployment," *Brookings Papers on Economic Activity*, 1999(2), 189–251.

Benassy, J.-P. (1996): "Taste for variety and optimum production patterns in monopolistic competition," *Economics Letters*, 52(1), 41 – 47.

Blanchard, O., E. Cerutti, and L. Summers (2015): "Inflation and Activity Two Explorations and their Monetary Policy Implications," Working Paper 21726, National Bureau of Economic Research.

Blanchard, O. J. (2017): "Should We Reject the Natural Rate Hypothesis?," Working Paper Series WP17-14, Peterson Institute for International Economics.

Blanchard, O. J., and P. Diamond (1994): "Ranking, Unemployment Duration, and Wages," *The Review of Economic Studies*, 61(3), 417–434.

Blanchard, O. J., and L. H. Summers (1987): "Hysteresis in unemployment," *European Economic Review*, 31(1-2), 288–295.

Clarida, R., J. Galí, and M. Gertler (1999): "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37(4), 1661–1707.

DeLong, J. B., and L. H. Summers (2012): "Fiscal Policy in a Depressed Economy," *Brookings Papers on Economic Activity*, 44(1 (Spring), 233–297.

Devereux, M. B., A. C. Head, and B. J. Lapham (1996): "Aggregate fluctuations with increasing returns to specialization and scale," *Journal of Economic Dynamics and Control*, 20(4), 627 – 656.

Dunne, T., M. Roberts, S. Klimek, and Y. Xu (2007): "The Dynamics of Market Structure and Market Size in Two Health Services Industries," Working Papers 07-26, Center for Economic Studies, U.S. Census Bureau.

Ethier, W. J. (1982): "National and International Returns to Scale in the Modern Theory of International Trade," *The American Economic Review*, 72(3), 389–405.

Fair, R. C., and J. B. Taylor (1983): "Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models," *Econometrica*, 51(4), 1169–1185.

Friedman, M. (1977): "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, 85(3), 451–472.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.

Guerrieri, L., and M. Iacoviello (2015): "OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily," *Journal of Monetary Economics*, 70, 22 – 38.

Ireland, P. N. (2007): "Changes in the Federal Reserve's Inflation Target: Causes and Consequences," *Journal of Money, Credit and Banking*, 39(8), 1851–1882.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1994): *Continuous Univariate Distributions*, vol. 1. Wiley, 2nd edn.

Mankiw, N. G. (2001): "The Inexorable and Mysterious Tradeoff between Inflation and Unemployment," *Economic Journal*, 111(471), 45–61.

Mankiw, N. G., and M. D. Whinston (1986): "Free Entry and Social Inefficiency," *The RAND Journal of Economics*, 17(1), 48–58.

Ramey, V. A. (2012): "Discussant comments on "Fiscal Policy in a Depressed Economy" by J. Bradford DeLong and Lawrence H. Summers," Brookings Institution, June 26, 2012.

Rogoff, K. (1985): "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, 100(4), 1169–1189.

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98(5), 71–102.

Rotemberg, J. J. (1982): "Monopolistic Price Adjustment and Aggregate Output," *The Review of Economic Studies*, 49(4), pp. 517–531.

Schorfheide, F. (2008): "DSGE model-based estimation of the New Keynesian Phillips curve," *Economic Quarterly*, (Fall), 397–433.

Schumpeter, J. (1994): *Capitalism, socialism and democracy*. Routledge New York.

Sim, J. (2008): "Uncertainty-Driven Business Cycle," Unpublished working paper, The Board of Governors of the Federal Reserve System.

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

Spence, M. (1976): "Product Selection, Fixed Costs, and Monopolistic Competition," *The Review of Economic Studies*, 43(2), 217–235.

WOODFORD, M. (2001): “The Taylor Rule and Optimal Monetary Policy,” *The American Economic Review*, 91(2), 232–237.

——— (2003): *Interest and prices.* Princeton Univ. Pr., Princeton, NJ.

YELLEN, J. L. (2016): “Macroeconomic Research After the Crisis,” Speech by Janet L. Yellen, the Chair of the Boart of Governors of the Federal Reserve System at ”The Elusive ’Great’ Recovery: Causes and Implications for Future Business Cycle Dynamics” 60th annual economic conference sponsored by the Federal Reserve Bank of Boston, Boston, Massachusetts.

Appendices

A Proofs of Propositions

A.1 Proof of Proposition 1

$W_t(z_t(i), m_{t-1}(i), m_t(i))$ can be written as one of the following four:

$$W_t(z_t(i), 0, 0) = \mathbb{E}_t[Q_{t,t+1} V_{t+1}(z_{t+1}(i), 0)] \tag{A.1}$$

$$W_t(z_t(i), 0, 1) = \Pi_t(i) - (\gamma^E + \gamma^X)y + \mathbb{E}_t[Q_{t,t+1} V_{t+1}(z_{t+1}(i), 1)] \tag{A.2}$$

$$W_t(z_t(i), 1, 0) = \mathbb{E}_t[Q_{t,t+1} V_{t+1}(z_{t+1}(i), 0)] \tag{A.3}$$

and

$$W_t(z_t(i), 1, 1) = \Pi_t(i) - \gamma^X y + \mathbb{E}_t[Q_{t,t+1} V_{t+1}(z_{t+1}(i), 1)] \tag{A.4}$$

By equating (A.1) and (A.2) at $z_t(i) = z_t^E$ on the one hand, and (A.3) and (A.4) at $z_t(i) = z_t^X$ on the other hand, one can derive the entry and exit conditions:

$$(\gamma^E + \gamma^X)y - \Pi_t(z_t^E) = q_t^F \tag{A.5}$$

and

$$\gamma^X y - \Pi_t(z_t^X) = q_t^F \tag{A.6}$$

where q_t^F, by using (9) and (14), can be expressed as

$$q_t^F = \mathbb{E}_t\left\{\int_0^\infty Q_{t,t+1}\left[\begin{array}{c} \max\{W_{t+1}(z_{t+1}(i), 1, 1), W_{t+1}(z_{t+1}(i), 1, 0)\} \\ -\max\{W_{t+1}(z_{t+1}(i), 0, 1), W_{t+1}(z_{t+1}(i), 0, 0)\}\end{array}\right] dF(z)\right\}. \tag{A.7}$$

To evaluate the integration, the support of z is divided into three regions: $[0, z_{t+1}^X)$; $[z_{t+1}^X, z_{t+1}^E)$; $[z_{t+1}^E, \infty)$. This partition of $\mathbb{R}^+$ assumes that $z_{t+1}^E > z_{t+1}^X$. This assumption is verified at the end of this appendix. In the first region, the integrand is equal to

$$\begin{aligned}
&W_{t+1}(z_{t+1}(i), 1, 0) - W_{t+1}(z_{t+1}(i), 0, 0) \\
&= \mathbb{E}_t[Q_{t,t+1} V_{t+1}(z_{t+1}(i), 0)] - \mathbb{E}_t[Q_{t,t+1} V_{t+1}(z_{t+1}(i), 0)] \\
&= 0
\end{aligned}$$

Incumbent firms find it optimal to exit ($W_{t+1}(z_{t+1}(i), 1, 1) < W_{t+1}(z_{t+1}(i), 1, 0)$) and dormant firms find it optimal to delay the entry ($W_{t+1}(z_{t+1}(i), 0, 1) < W_{t+1}(z_{t+1}(i), 0, 0)$).

In the second region, the integrand is equal to

$$\begin{aligned}
&W_{t+1}(z_{t+1}(i), 1, 1) - W_{t+1}(z_{t+1}(i), 0, 0) \\
&= \Pi_{t+1}(i) - \gamma^X y + \underbrace{\mathbb{E}_{t+1}\left\{Q_{t+1,t+2}\left[V_{t+1}(z_{t+2}(i), 1) - V_{t+1}(z_{t+2}(i), 0)\right]\right\}}_{= q_{t+1}^F}
\end{aligned} \tag{A.8}$$

In this region, the idiosyncratic shock is not low enough to make an incumbent firm want to exit the market ($W_{t+1}(z_{t+1}(i), 1, 1) > W_{t+1}(z_{t+1}(i), 1, 0)$). However, the idiosyncratic shock is not high enough to make a new entrant firm want to enter the market ($W_{t+1}(z_{t+1}(i), 0, 1) < W_{t+1}(z_{t+1}(i), 0, 0)$). No firms change its status as incumbent firms find it optimal to delay exit and dormant firms still find it optimal to delay entry. This is the inaction region.

In the last region, the integrand is equal to

$$\begin{aligned} & W_{t+1}(z_{t+1}(i),1,1) - W_{t+1}(z_{t+1}(i),0,1) \\ = & \Pi_{t+1}(i) - \gamma^X y + \mathbb{E}_t[Q_{t,t+1}V_{t+1}(z_{t+1}(i),1)] \\ & -\{\Pi_{t+1}(i) - (\gamma^E + \gamma^X)y + \mathbb{E}_t[Q_{t,t+1}V_{t+1}(z_{t+1}(i),1)]\} \\ = & \gamma^E y \end{aligned} \tag{A.9}$$

In this region, the cost shock is so high that no firm wants to in the market. Dormant firms remain dormant and incumbent firms exit the market. Hence using (A.8) and (A.9), one can evaluate the value of marginal firm as

$$q_t^F = \mathbb{E}_t\left[Q_{t,t+1}\int_{z_{t+1}^E}^{\infty}\gamma^E y dF(z) + \int_{z_{t+1}^X}^{z_{t+1}^E}\left[\Pi_{t+1}(z) - \gamma^X y + q_{t+1}^F\right]dF(z|\sigma)\right].$$

Q.E.D.

A.2 Proof of Proposition 2

The market clearing condition at the individual good-level is given by

$$M_t^{-\xi}\left[\frac{P_t(i)}{P_t}\right]^{-\epsilon_p} y_t = x_t z_t(i) k_t(i)^{\alpha} n_t(i)^{1-\alpha} \tag{A.10}$$

Imposing the symmetric equilibrium condition, dividing both sides by $z_t(i)M_t^{-\xi}$ and integrating the both sides of (A.10) with respect to the firm index yields

$$y_t \int_0^{M_t} z_t(i)^{-1} di = x_t M_t^{\xi}\kappa_t^{\alpha}\int_0^{M_t} n_t(i)di = x_t M_t^{\xi} k_t^a n_t^{1-\alpha}. \tag{A.11}$$

To evaluate the left hand sides, note that $1/z_t$ follows a lognormal distribution, $\log(1/z_t) = -\log(z_t) \sim N(0.5\sigma_z^2, \sigma_z^2)$. Also note that of the total number of firms M_t, $(1-M_{t-1})[1-\Phi(\nu_t^E)]$ of new entrant firms and $M_{t-1}[1-\Phi(\nu_t^X)]$ of incumbent firms have average inverse productivities,

$$\mathbb{E}(z^{-1}|z \geq z_t^E) = \frac{1}{1-F(z_t^E)}\int_{z\geq z_t^E} z^{-1}dF(z) = \frac{\Phi(\tilde{\nu}_t^E - \sigma_z)}{\Phi(\tilde{\nu}_t^E)}$$

and

$$\mathbb{E}(z^{-1}|z \geq z_t^X) = \frac{1}{1-F(z_t^X)}\int_{z\geq z_t^X} z^{-1}dF(z) = \frac{\Phi(\tilde{\nu}_t^X - \sigma_z)}{\Phi(\tilde{\nu}_t^X)},$$

respectively where

$$\tilde{\nu}_t^E \equiv \frac{\log(1/z_t^E) - 0.5\sigma_z^2}{\sigma_z} = -\nu_t^E \tag{A.12}$$

and

$$\tilde{\nu}_t^X \equiv \frac{\log(1/z_t^X) - 0.5\sigma_z^2}{\sigma_z} = -\nu_t^X \tag{A.13}$$

(See Johnson, Kotz, and Balakrishnan (1994) regarding the properties of lognormal distribution and truncated integral). Hence, the integration of inverse idiosyncratic shocks can be expressed as

$$\begin{aligned}\int_0^{M_t} z_t(i)^{-1}di &= (1-M_{t-1})[1-\Phi(\nu_t^E)]\mathbb{E}(z^{-1}|z \geq z_t^E) + M_{t-1}[1-\Phi(\nu_t^X)]\mathbb{E}(z^{-1}|z \geq z_t^X) \\ &= (1-M_{t-1})[1-\Phi(\nu_t^E)]\frac{\Phi(\tilde{\nu}_t^E - \sigma_z)}{\Phi(\tilde{\nu}_t^E)} + M_{t-1}[1-\Phi(\nu_t^X)]\frac{\Phi(\tilde{\nu}_t^X - \sigma_z)}{\Phi(\tilde{\nu}^X)}\end{aligned}$$

Using (A.12), (A.13), and the fact that $\Phi(-x) = 1 - \Phi(x)$ in general, we can rewrite the above as

$$\int_0^{M_t} z_t(i)^{-1}di = (1-M_{t-1})[1-\Phi(\nu_t^E + \sigma_z)] + M_{t-1}[1-\Phi(\nu_t^X + \sigma_z)] \tag{A.14}$$

Substituting (A.14) in (A.11) and defining

$$\Xi_t \equiv \frac{\{[(1-M_{t-1})[1-\Phi(\nu_t^E)] + M_{t-1}[1-\Phi(\nu_t^X)]\}^\xi}{(1-M_{t-1})[1-\Phi(\nu_t^E + \sigma_z)] + M_{t-1}[1-\Phi(\nu_t^X + \sigma_z)]},$$

one can express the aggregate production function as

$$y_t = x_t \Xi_t k_t^a n_t^{1-\alpha}.$$

Q.E.D.

A.3 Analytical Expression of Value of Marginal Firm

To derive an analytical expression for this, note that the profit in symmetric equilibrium is given by

$$\begin{aligned}\Pi_{t+1}(z_{t+1}(i)) &= y_{t+1}(i) - w_{t+1}n_{t+1} - r_{t+1}^K k_{t+1} \\ &= y_{t+1}(i)\left\{1 - \frac{1}{x_{t+1}z_{t+1}(i)}\left[w_{t+1}\left(\frac{\alpha}{1-\alpha}\frac{w_{t+1}}{r_{t+1}^K}\right)^{-\alpha} + r_{t+1}^K\left(\frac{\alpha}{1-\alpha}\frac{w_{t+1}}{r_{t+1}^K}\right)^{1-\alpha}\right]\right\} \\ &= y_{t+1}M_{t+1}^{-\xi}\left[1 - \frac{1}{x_{t+1}z_{t+1}(i)}\left(\frac{w_{t+1}}{1-\alpha}\right)^{1-\alpha}\left(\frac{r_{t+1}^K}{\alpha}\right)^{\alpha}\right].\end{aligned} \tag{A.15}$$

Note that if we define $u_t = 1/z_t$ such that $dz_t = -1/u_t^2 du_t$,

$$\begin{aligned}\int_{z_{t+1}^X}^{z_{t+1}^E} \frac{1}{z_{t+1}} f(z)dz &= -\int_{1/z_{t+1}^X}^{1/z_{t+1}^E} \frac{1}{u_t} f\left(\frac{1}{u_t}\right) du_t \\ &= \int_{1/z_{t+1}^E}^{1/z_{t+1}^X} \frac{1}{u_t} f\left(\frac{1}{u_t}\right) du_t = \int_{u_{t+1}^E}^{u_{t+1}^X} u_t g(u_t) du_t\end{aligned}$$

where g is the pdf of lognormal distribution of $u_t = 1/z_t$. We then define normalized thresholds

$$\tilde{\nu}_{t+1}^X \equiv \frac{\log u_{t+1}^X - 0.5\sigma_z^2}{\sigma_z} = \frac{\log(1/z_{t+1}^X) - 0.5\sigma_z^2}{\sigma_z} = -\nu_{t+1}^X$$

and

$$\tilde{\nu}_{t+1}^E \equiv \frac{\log u_{t+1}^E - 0.5\sigma_z^2}{\sigma_z} = \frac{\log(1/z_{t+1}^X) - 0.5\sigma_z^2}{\sigma_z} = -\nu_{t+1}^E$$

Hence

$$\int_{z_{t+1}^X}^{z_{t+1}^E} \Pi_{t+1}(z)dF(z) = y_{t+1}M_{t+1}^{-\xi}\{\Phi(\nu_{t+1}^E) - \Phi(\nu_{t+1}^X) \\ - \left(\Phi(\tilde{\nu}_{t+1}^X - \sigma_{z,t+1}) - \Phi(\tilde{\nu}_{t+1}^E - \sigma_{z,t+1})\right)\Xi_{t+1}\mathbb{E}_{t+1}[\mu_{t+1}(i)]\}$$

where we use

$$\frac{1}{x_t}\left(\frac{r_t^K}{\alpha}\right)^\alpha \left(\frac{w_t}{1-\alpha}\right)^{1-\alpha} = \Xi_t\mathbb{E}_t^i[\mu_t(i)].$$

Since in general $\Phi(-x) = 1 - \Phi(x)$,

$$\begin{aligned}\Phi(\tilde{\nu}_{t+1}^X - \sigma_z) - \Phi(\tilde{\nu}_{t+1}^E - \sigma_z) &= 1 - \Phi(\nu_{t+1}^X + \sigma_z) - [1 - \Phi(\nu_{t+1}^E + \sigma_z)] \\ &= \Phi(\nu_{t+1}^E + \sigma_z) - \Phi(\nu_{t+1}^X + \sigma_z)\end{aligned}$$

One can then express the value of marginal firm as

$$\begin{aligned} q_t^F &= \mathbb{E}_t\left\{Q_{t,t+1}\left[(1-\Phi(\nu_{t+1}^E))\gamma^E y + \left(\Phi(\nu_{t+1}^E) - \Phi(\nu_{t+1}^X)\right)(q_{t+1}^F - \gamma^X y)\right]\right\} \\ &+ \mathbb{E}_t\left\{Q_{t,t+1}\frac{y_{t+1}}{M_{t+1}^\xi}\left[\Phi(\nu_{t+1}^E) - \Phi(\nu_{t+1}^X) - (\Phi(\nu_{t+1}^E + \sigma_z) - \Phi(\nu_{t+1}^X + \sigma_z))\Xi_{t+1}\mathbb{E}_{t+1}[\mu_{t+1}(i)]\right]\right\}\end{aligned} \tag{A.16}$$

The system of equations then replace (15) in the main text with (A.16). (A.5), (A.6) and (A.15) imply

$$\begin{aligned}(\gamma^E + \gamma^X)y - y_tM_t^{-\xi}\left[1 - \frac{1}{x_tz_t^E}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}\left(\frac{r_t^K}{\alpha}\right)^\alpha\right] &= q_t^F; \\ \gamma^X y - y_tM_t^{-\xi}\left[1 - \frac{1}{x_tz_t^X}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}\left(\frac{r_t^K}{\alpha}\right)^\alpha\right] &= q_t^F.\end{aligned}$$

Solving these two expression for $1/z_t^E$ and $1/z_t^X$ verifies the guess $z_t^E > z_t^X$:

$$\begin{aligned}1/z_t^E &= x_t\left(\frac{1-\alpha}{w_t}\right)^{1-\alpha}\left(\frac{r_t^K}{\alpha}\right)^\alpha\left[1 - \frac{(\gamma^E + \gamma^X)y - q_t^F}{y_tM_t^{-\xi}}\right] \\ &< x_t\left(\frac{1-\alpha}{w_t}\right)^{1-\alpha}\left(\frac{r_t^K}{\alpha}\right)^\alpha\left[1 - \frac{\gamma^X y - q_t^F}{y_tM_t^{-\xi}}\right] = 1/z_t^X.\end{aligned}$$

B System of Equations

There are 25 endogenous variables in the economy with nominal rigidity, including the Wicksellian natural output, which should be determined in the flexible block of the economy:

$$Y_t = \begin{bmatrix} z_t^E & z_t^X & q_t^F & \nu_t^E & \nu_t^X & M_t & r_t^K & w_t & \Xi_t & \pi_t \\ y_t & k_t & n_t & q_t^K & Q_{t,t+1} & r_t & \lambda_t & c_t & i_t & \mathbb{E}_t^i[\mu_t(i)] \\ \tau_t & x_t & \psi_t & g_t & y_t^* \end{bmatrix}.$$

Of these, the output under price flexibility, y_t^*, has to be determined outside the economy with nominal rigidity. The 27 equations are given as follows.

$$(\gamma^E + \gamma^X)y - y_tM_t^{-\xi}\left[1 - \frac{1}{x_tz_t^E}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}\left(\frac{r_t^K}{\alpha}\right)^\alpha\right] = q_t^F \tag{B.1}$$

$$\gamma^X y - y_t M_t^{-\xi}\left[1 - \frac{1}{x_t z_t^X}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}\left(\frac{r_t^K}{\alpha}\right)^{\alpha}\right] = q_t^F \tag{B.2}$$

$$\begin{aligned} q_t^F &= \mathbb{E}_t\left\{Q_{t,t+1}\left[(1-\Phi(\nu_{t+1}^E))\gamma^E y + \left(\Phi(\nu_{t+1}^E) - \Phi(\nu_{t+1}^X)\right)(q_{t+1}^F - \gamma^X y)\right]\right\} \\ &+ \mathbb{E}_t\left\{Q_{t,t+1}\frac{y_{t+1}}{M_{t+1}^{\xi}}\left[\Phi(\nu_{t+1}^E) - \Phi(\nu_{t+1}^X) - (\Phi(\nu_{t+1}^E + \sigma_z) - \Phi(\nu_{t+1}^X + \sigma_z))\Xi_{t+1}\mathbb{E}_{t+1}[\mu_{t+1}(i)]\right]\right\} \end{aligned} \tag{B.3}$$

$$\nu_t^E = \sigma_z^{-1}[\log(z_t^E) + 0.5\sigma_z^2] \tag{B.4}$$

$$\nu_t^X = \sigma_z^{-1}[\log(z_t^X) + 0.5\sigma_z^2] \tag{B.5}$$

$$M_t = M_{t-1}[1 - \Phi(\nu_t^X)] + (1 - M_{t-1})[1 - \Phi(\nu_t^E)] \tag{B.6}$$

$$\left(\frac{\alpha}{1-\alpha}\right)\frac{w_t}{r_t^K} = \frac{k_t}{n_t} \tag{B.7}$$

$$\mathbb{E}_t^i[\mu_t(i)] = \frac{1}{x_t \Xi_t}\left(\frac{r_t^K}{\alpha}\right)^{\alpha}\left(\frac{w_t}{1-\alpha}\right)^{1-\alpha} \tag{B.8}$$

$$\Xi_t \equiv \frac{\{[(1-M_{t-1})[1-\Phi(\nu_t^E)] + M_{t-1}[1-\Phi(\nu_t^X)]\}^{\xi}}{(1-M_{t-1})[1-\Phi(\nu_t^E+\sigma_z)] + M_{t-1}[1-\Phi(\nu_t^X+\sigma_z)]} \tag{B.9}$$

$$\begin{aligned} 1 &= \frac{\epsilon_p}{\epsilon_p - 1}\mathbb{E}_t^i[\mu_t(i)] - \frac{\varphi_p}{\epsilon_p - 1}\frac{\pi_t}{\pi_{t-1}^{\chi}}\left(\frac{\pi_t}{\pi_{t-1}^{\chi}} - 1\right) \\ &+ \frac{\varphi_p}{\epsilon_p - 1}\mathbb{E}_t\left[Q_{t,t+1}[1-\Phi(\nu_{t+1}^X)]\frac{\pi_{t+1}}{\pi_t^{\chi}}\left(\frac{\pi_{t+1}}{\pi_t^{\chi}} - 1\right)\frac{y_{t+1}/M_{t+1}^{\xi}}{y_t/M_t^{\xi}}\right] \end{aligned} \tag{B.10}$$

$$y_t = x_t \Xi_t k_t^a n_t^{1-\alpha} \tag{B.11}$$

$$\lambda_t = \left(c_t - hc_{t-1} - \frac{1/\varsigma}{1+1/\theta}n_t^{1+1/\theta}\right)^{-\sigma_u} \tag{B.12}$$

$$\begin{aligned} (1-\tau_t)w_t &= \frac{\epsilon_w}{\epsilon_w - 1}\frac{n_t^{1/\theta}}{\varsigma} - \frac{\varphi_w}{\epsilon_w - 1}\frac{w_t}{w_{t-1}}\left(\frac{w_t}{w_{t-1}} - 1\right) \\ &+ \frac{\beta}{\epsilon_w - 1}\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\lambda_t}\frac{w_{t+1}}{w_t}\left(\frac{w_{t+1}}{w_t} - 1\right)\frac{n_{t+1}}{n_t}\right] \end{aligned} \tag{B.13}$$

$$1 = \mathbb{E}_t\left[Q_{t,t+1}R_{t+1}\psi_t\right] \tag{B.14}$$

$$1 = \mathbb{E}_t\left[Q_{t,t+1}\frac{r_{t+1}^K + (1-\delta)q_{t+1}^K}{\psi_t q_t^K}\right] \tag{B.15}$$

$$q_t^K = 1 + \kappa\left(\frac{i_t}{i_{t-1}} - 1\right) - \mathbb{E}_t\left\{Q_{t,t+1}\frac{\kappa}{2}\left[\left(\frac{i_{t+1}}{i_t}\right)^2 - 1\right]\right\} \tag{B.16}$$

$$Q_{t,t+1} = \beta\frac{\lambda_{t+1}}{\lambda_t} \tag{B.17}$$

$$r_t = r^* + \mathbf{1}(\pi_t \geq \pi^*)\rho_\pi^{(+)}(\pi_t - \pi^*) + \mathbf{1}(\pi_t < \pi^*)\rho_\pi^{(-)}(\pi_t - \pi^*) \tag{B.18}$$

$$\begin{aligned} y_t = c_t + i_t + g_t + \frac{\kappa}{2}\left(\frac{i_t}{i_{t-1}} - 1\right)^2 i_{t-1} + \frac{\varphi_p}{2}\left(\frac{\pi_t}{\pi_{t-1}^X} - 1\right)^2 y_t M_t^{-\xi} \\ +(1 - M_{t-1})[1 - \Phi(\nu_t^E)]\gamma^E y + M_t\gamma^X y + \frac{\varphi_w}{2}\left(\frac{w_t}{w_{t-1}} - 1\right)^2 n_t \end{aligned} \tag{B.19}$$

$$k_{t+1} = (1 - \delta)k_t + i_t \tag{B.20}$$

$$\tau_t = \frac{g_t}{w_t n_t} \tag{B.21}$$

$$\log x_t = \rho_x \log x_{t-1} + \sigma_x \epsilon_{x,t} \tag{B.22}$$

$$\log \psi_t = \rho_\Xi \log \psi_{t-1} + \sigma_\psi \epsilon_{\psi,t} \tag{B.23}$$

$$g_t = \rho_g g_{t-1} + \sigma_g \epsilon_{g,t} \tag{B.24}$$

C Non-Stochastic Steady State

In the steady state government spending is assumed to be zero, i.e., $g = 0$. This also implies that the steady state labor income tax rate is equal to zero, i.e., $\tau = 0$. To solve for the non-stochastic steady state, guess z^E and z^X. This guess determines

$$\nu^E \equiv \sigma_z^{-1}[\log(z^E) + 0.5\sigma_z^2], \tag{C.1}$$

$$\nu_t^X \equiv \sigma_z^{-1}[\log(z^X) + 0.5\sigma_z^2], \tag{C.2}$$

$$M = \frac{1 - \Phi(\nu^E)}{1 - [\Phi(\nu^E) - \Phi(\nu^X)]}, \tag{C.3}$$

and

$$\Xi_t \equiv \frac{\{[(1 - M)[1 - \Phi(\nu^E)] + M[1 - \Phi(\nu^X)]\}^\xi}{(1 - M)[1 - \Phi(\nu^E + \sigma_z)] + M[1 - \Phi(\nu^X + \sigma_z)]}. \tag{C.4}$$

Steady state inflation rate is assumed to be zero. (B.10) then pins down the steady state marginal cost as

$$\mathbb{E}[\mu(i)] = \frac{\epsilon_p - 1}{\epsilon_p} \tag{C.5}$$

Since the real marginal cost must also satisfy

$$\mathbb{E}[\mu(i)] = \frac{1}{x\Xi}\left(\frac{r^K}{\alpha}\right)^{\alpha}\left(\frac{w}{1-\alpha}\right)^{1-\alpha}, \tag{C.6}$$

substituting (C.5) in (C.6) and solving this for w yields

$$w = (1-\alpha)\left[x\Xi\mathbb{E}[\mu(i)]\left(\frac{\alpha}{r^K}\right)^{\alpha}\right]^{1/(1-\alpha)}. \tag{C.7}$$

The GHH preference then pins down the level of labor as

$$n = (\varsigma w)^{\theta} \tag{C.8}$$

Aggregate output can be written in terms of capital-labor ratio:

$$y = x\Xi\left(\frac{k}{n}\right)^{\alpha} n \tag{C.9}$$

From (B.7) and (B.8), one can derive the capital-labor ratio as

$$\frac{k}{n} = \frac{w}{r^k}\frac{\alpha}{1-\alpha} \tag{C.10}$$

where r^k is given by $r^k = \beta^{-1} - (1-\delta)$. Substituting (C.8) and (C.10) in (C.9) yields the aggregate output. Aggregate capital can be determined as

$$k = \left(\frac{w}{r^k}\frac{\alpha}{1-\alpha}\right)n.$$

Substituting this in the resource constraint and solving the expression in c yields

$$c = y - \delta k - (1-M)[1-\Phi(\nu^E)]\gamma^E y + M\gamma^X y \tag{C.11}$$

In the steady state, the marginal value of firm satisfies

$$q^F = \gamma^X y - yM^{-\xi}\left[1 - \frac{1}{xz^X}\left(\frac{w}{1-\alpha}\right)^{1-\alpha}\left(\frac{r^K}{\alpha}\right)^{\alpha}\right] \tag{C.12}$$

To update on the initial guess, three zero functions are defined as follows:

$$h^1(z^E, z^X) \equiv (\gamma^E + \gamma^X)y - yM^{-\xi}\left[1 - \frac{1}{xz^E}\left(\frac{w}{1-\alpha}\right)^{1-\alpha}\left(\frac{r^K}{\alpha}\right)^{\alpha}\right] - q^F \tag{C.13}$$

and

$$\begin{aligned} h^2(z^E, z^X) \equiv & -q^F + \beta\left[(1-\Phi(\nu^E))\gamma^E y + \left(\Phi(\nu^E) - \Phi(\nu^X)\right)(q^F - \gamma^X y)\right] \\ & + \beta\frac{y}{M^{\xi}}\left[\Phi(\nu^E) - \Phi(\nu^X) - (\Phi(\nu^E + \sigma_z) - \Phi(\nu^X + \sigma_z))\Xi\mathbb{E}^i[\mu(i)]\right] \end{aligned} \tag{C.14}$$

A numerical root finder based on Newton's method is used to find a solution $\{z^E, z^X\}$ that satisfies

$$0 = h^1(z^E, z^X) = h^2(z^E, z^X).$$

D Numerical Method

The extended path method by Adjemian and Juillard (2011) can be considered as a variant of Fair and Taylor (1983). For any given states and realized shocks, the method finds the equilibrium values of endogenous variables by solving two-point boundary value problems. The method can be best illustrated by introducing a matrix notation. Let y_t^s denotes $N \times 1$ vector of endogenous variables where $s = 1, ..., S$ and $t = 0, ..., T$. S is the size of the simulation and T is the size of the two-point boundary value problems. At $s = 1$, the states are given by the non-stochastic steady state, denoted by y_0^0. At $s = 1$, shocks, denoted by $M \times 1$ vector ϵ_1, realize.

If the functional forms of policy functions are known, today's solution is determined by $y_1^1 = f(y_0^0, \epsilon_1)$. However, the functional forms are not known. Instead we obtain y_1^1 by solving the two boundary value problems, also known as "shooting algorithm". The optimality of today's solution y_1^1 is ensured by checking whether or not the entire time path from y_1^1 to y_1^T of the two-point boundary value problem is consistent with the model's nonlinear equations, which may contain kinks. This process is done under the assumption of perfect foresight. In other words, all shock variables are evaluated at their means of the distribution, i.e., zeros. Once this condition is satisfied, the solutions of the boundary value problems except today's equilibrium values, namely $y_2^1, ..., y_T^1$, are discarded

$$Y = \begin{bmatrix} & \epsilon_1 & 0 & \cdots & 0 & & & & \\ y_0^0 & y_1^1 & y_2^1 & \cdots & y_T^1 & & & & \\ & \downarrow & \epsilon_2 & 0 & \cdots & 0 & & & \\ & y_0^1 & y_1^2 & y_2^2 & \cdots & y_T^2 & & & \\ & & \downarrow & \epsilon_3 & 0 & \cdots & 0 & & \\ & & y_0^2 & y_1^3 & y_2^3 & \cdots & y_T^3 & & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & & \downarrow & \epsilon_S & 0 & \cdots & 0 \\ & & & & y_0^{S-1} & y_1^S & y_2^S & \cdots & y_T^S \end{bmatrix} \tag{D.1}$$

At $s = 2$, the initial conditions y_0^1 are set equal to y_1^1, the solution value at $s = 1$. At $s = 2$, shocks, ϵ_2 realize. With the initial condition y_0^1 and the realized shocks ϵ_2, another two-point boundary value problem is solved to obtain $y_1^2, ..., y_T^2$. Again the solutions of the boundary value problems except today's equilibrium values, namely $y_2^2, ..., y_T^2$ are discarded. At $s = 3$, the initial conditions y_0^2 are set equal to y_1^2 and shocks ϵ_3 realize. We then solve another two-point boundary value problem is solved to obtain $y_1^3, ..., y_T^3$. Again y_1^3 is kept as tomorrow's initial condition, and $y_2^3, ..., y_T^3$ is discarded. This process should be iterated.

The matrix (D.1) illustrates this process. For instance, row 3 describes the boundary value problem of $s = 1$ and row 5 describes the boundary value problem of $s = 2$. One can then see $y_0^0, ..., y_0^{S-1}$ as time-varying initial conditions, $y_1^1, ..., y_1^S$ as model's nonlinear solution and $y_T^1 = \cdots = y_T^S = y_0^0$ as time-invariant terminal conditions. In each date s, finding solutions of size T for the two point boundary value problem involves initial guessing and updating, which relies on Newton-based method.

As mentioned in the main text, this method greatly improves upon popular linear or piecewise linear perturbation solutions as the solution is fully consistent with model's nonlinear equations, and is quite versatile in handling arbitrary numbers of kinks. There is no approximation error associated with perturbation degree or with perturbation point. However, there are two drawbacks in this method. First, this solution method is expensive in that size $N \times S$ solution requires finding $N \times S \times T$ solutions to the boundary value problems. Second, like any other linear or piecewise linear solutions, it assumes perfect foresight and ignores the Jensen's inequality effect of uncertainty.