Kiel
Working Papers

Kiel Institute
for the World Economy

Optimal Inflation and Firms' Productivity Dynamics

by Henning Weber

No. 1685| February 2011

Web: www.ifw-kiel.de

Kiel Working Paper No. 1685 | February 2011

Optimal Inflation and Firms' Productivity Dynamics

Henning Weber

Abstract:
Empirical data indicate that firms tend to have below-average productivity upon entry and that they tend to experience post-entry productivity growth. I present a New Keynesian model with growth in firm-specific productivity and firm turnover that captures these two phenomena. The model predicts that the optimal rate of long-run inflation is positive and equal to growth in firm-specific productivity. When linearized at positive optimal inflation, the model is observationally equivalent to the basic New Keynesian model with homogenous productivity linearized at zero inflation. Optimal stabilization policies are the same in both models, and the Taylor principle ensures determinacy in either model.

Keywords: Optimal long-run inflation, trend inflation, heterogenous firms.

JEL classification: E01, E31, E32

Kiel Institute for the World Economy
24105 Kiel, Germany

Henning Weber
Telephone: +49 431 8814 572
E-mail: henning.weber@ifw-kiel.de

Optimal Inflation and Firms' Productivity Dynamics

Henning Weber*

February 20, 2011

PRELIMINARY, COMMENTS WELCOME.

Abstract

Empirical data indicate that firms tend to have below-average productivity upon entry and that they tend to experience post-entry productivity growth. I present a New Keynesian model with growth in firm-specific productivity and firm turnover that captures these two phenomena. The model predicts that the optimal rate of long-run inflation is positive and equal to growth in firm-specific productivity. When linearized at positive optimal inflation, the model is observationally equivalent to the basic New Keynesian model with homogenous productivity linearized at zero inflation. Optimal stabilization policies are the same in both models, and the Taylor principle ensures determinacy in either model.

*Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany, henning.weber@ifw-kiel.de, Tel: +49 431 8814-572. I thank particular Sebastian Braun, Andrea Colciago, Holger Görg, Paul Kramer, Wolfgang Lechthaler, Bartosz Mackowiak, Gianmarco Ottaviano, Chris Reicher, Dennis Snower, Carl Walsh, Lutz Weinke, and seminar participants in the Warsaw International Economic Meeting 2010 for discussions, comments, and suggestions. Any errors are mine.

1 Motivation

Many central banks around the globe maintain objectives for long-run inflation between 1 and 3 percent per year.[1] However, many academics in the field of monetary economics identify a very different range for optimal long-run inflation, namely from minus the real interest rate to zero.[2] These opposing views leave plenty of room for disagreement, and the recent nosedive in nominal short-term interest rates has moved the debate about the optimal rate of long-run inflation to center stage.

In this paper, I analyze optimal long-run inflation under plausible assumptions about firm-level productivity dynamics, and make a case for moderately positive long-run inflation. In firm-level data, firms move systematically through the productivity distribution over time. Firms tend to have below-average productivity upon market entry and they tend to experience post-entry growth in productivity through learning by doing, economies of scale, process innovation, or through changes in their product mix.

I consider a New Keynesian model with firm turnover and growth in firm-specific productivity (FIP) to capture the firm-level productivity dynamics observed in the data. In the FIP model, a firm enters the market with below-average productivity, and the firm's productivity grows at constant rate over time. The firm sets the optimal price for its product upon entry and resets the price infrequently in subsequent periods. The FIP model nests the basic New Keynesian model with homogenous productivity (HOP) as a special case.

My analysis delivers two main results. The first result is that optimal long-run inflation in the FIP model is positive and equal to growth in firm-specific productivity. This result emerges from the planner problem associated with the FIP model. In contrast, it is well known that, in the HOP model, optimal long-run inflation is zero. Zero long-run inflation allows firms in the HOP model to maintain the flexible-price markup at all times despite sticky nominal prices. Moreover, at zero long-run inflation, there is no dispersion in relative prices because firms do not benefit from changing their price at the flexible-price markup.

The FIP model overturns the zero-inflation (price-stability) result that is so prominent in the literature. In the FIP model, positive rather than zero long-run inflation is required for firms to maintain the flexible-price markup. Growth in firm-specific productivity induces a firm's real marginal costs to decline. Accordingly, for the firm to maintain the flexible-price markup, its real

[1] See Kuttner (2004), table 2, and Schmitt-Grohe and Uribe (2010), table 1.
[2] See Schmitt-Grohe and Uribe (2010).

price also must decline. It requires positive long-run inflation equal to growth in firm-specific productivity to convert sticky nominal prices into declining real prices.

The second result of my analysis concerns macro dynamics and the central bank's optimal stabilization of aggregate shocks. In order to analyze these dimensions, I linearize the FIP and the HOP models at their respective efficient flexible-price steady states. The second result is that both linearized models are observationally equivalent if the slope of the approximate Phillips curve is parameterized directly. Observational equivalence implies that both models generate the same impulse-response functions, and the same decomposition of macro data into aggregate shocks. It also follows from observational equivalence that the conditions for a determinate local equilibrium coincide across models. A prominent example is the Taylor principle, which requires that, for the equilibrium to be determinate, the central bank has to move the nominal interest rate more than one-for-one with inflation.

Observational equivalence extends to the central bank's optimal stabilization of shocks, and this result does not depend on whether the central bank acts with discretion or commitment. Optimal stabilization policies are the same across models because the welfare-based loss function of the central bank, derived as second-order expansion to the utility of the representative household, is the same in the FIP and in the HOP model. Accordingly, any linear-quadratic policy problem derived from either model, with the welfare-based loss function as the objective and equilibrium conditions as constraints, delivers the same optimal stabilization policy. The equivalence of optimal stabilization policies is consistent with the fact that optimal long-run inflation is distinct in the two models. Whereas the central bank in the FIP model stabilizes actual inflation around positive long-run inflation, the central bank in the HOP model stabilizes actual inflation around zero long-run inflation.

This paper is related to a comprehensive literature on optimal long-run inflation. In view of this literature, the paper's original contribution is to highlight the consequences of firm-level productivity dynamics for optimal long-run inflation. Thus, the main feature of my analysis is heterogenous productivity across firms, whereas the benchmark model in the literature features homogenous productivity across firms.

A prominent strand in the literature on long-run inflation finds that zero or negative long-run inflation is optimal. Goodfriend and King (2001) analyze the baseline model of the New Neoclassical Synthesis and make a case for long-run inflation near zero. Khan, King, and Wolman (2003), Amano, Moran, Murchison, and Rennison (2009), and Aruoba and Schorfheide (2009) find that long-run deflation between 0.5 and 2.5 percent per year is optimal. Khan, King, and Wolman (2003) use a model with four distortions that arise from staggered pricing, price-cost markups, shopping time, and money that is required for transactions. Amano, Moran, Murchison, and Rennison (2009) use a model with aggregate productivity growth. Aruoba and Schorfheide (2009) use a two-sector model with search-based microfounded money demand. Schmitt-Grohe and Uribe (2010) review main determinants of long-run inflation and conclude that optimal long-run inflation ranges from minus the real interest rate to zero. In contrast, this paper finds positive long-run inflation in a modified model of the New Neoclassical Synthesis.

A second strand in the literature emphasizes aggregate factors that imply positive long-run inflation while this paper emphasizes firm-specific factors that imply positive long-run inflation. Kim and Ruge-Murcia (2009) find that long-run inflation of 1.2 percent per year overcomes downwardly rigid nominal wages. Billi (forthcoming) and Coibion, Gorodnichenko, and Wieland (2010) account for the zero lower bound on nominal interest rates. Billi (forthcoming) finds that optimal long-run inflation is near 1 percent per year or close to 17 percent per year, depending on whether the central bank acts under commitment or discretion. Coibion, Gorodnichenko, and Wieland (2010) find that optimal long-run inflation is positive but less than 2 percent per year for various regimes of monetary policy.

In a related paper, Wolman (2009) studies long-run inflation in a two-sector model with sticky prices and sector-specific productivity growth. Productivity in each sector grows at its own rate, whereas real wages grow at the same rate in both sectors. In this model, the central bank cannot stabilize markups in both sectors at the same time, and mild deflation turns out to minimize social loss. My analysis highlights a different mechanism in that productivity growth remains identical across firms. Therefore, the central bank can stabilize all markups at the same time.

This paper is also related to the recent literature on the impact of long-run inflation on macro dynamics and optimal stabilization policies. A main assumption in this literature is that the positive inflation objectives observed in the real world are suboptimal. Ascari (2004), Hornstein and Wolman (2005), Ascari and Ropele (2007), Kiley (2007), Ascari and Ropele (2009), and Kobayashi and Muto (2010) adopt this assumption in New Keynesian models with homogenous productivity and show that the level of long-run inflation affects the plausibility of assumptions about how firms set prices, about the dynamics of macro variables, about determinacy properties of monetary-policy rules, and about optimal monetary-policy responses to shocks. A broad conclusion in this literature is that neither positive nor normative predictions of the New Keynesian model linearized at zero long-run inflation extrapolate to the realistic case of positive long-run inflation.

In contrast, a main result of my analysis is that the positive inflation objectives observed in the real world are (close to) optimal. This finding conflicts with the assumption that positive inflation objectives are suboptimal. Moreover, by observational equivalence with respect to the linearized HOP model, the linearized FIP model warrants extrapolating both positive and normative predictions of the basic New Keynesian model linearized at zero long-run inflation to the realistic case of positive long-run inflation.

The rest of this paper is organized as follows. Section 2 reviews existing evidence on firms' productivity dynamics and describes the model used in this paper. Section 3 analyzes the optimal rate of long-run inflation, and Section 4 presents the calibration. Output effects of suboptimal long-run inflation are quantified in Section 5. Section 6 describes macro dynamics and the central bank's optimal reaction to shocks. The observational equivalence of linearized models requires that the slope of the linearized Phillips curve is parameterized directly, and I argue below that this is a reasonable precondition. As robustness check, Section 7 considers the case when this slope is computed from structural parameters. Section 8 concludes.

2 Model

This section describes a model with growth in firm-specific productivity and with firm turnover that captures firms' productivity dynamics observed in the data. A special case of the model is the basic New Keynesian model without growth in firm-specific productivity and without firm turnover as in Clarida, Gali, and Gertler (1999), Woodford (2003), chapter 3, and Gali (2008), chapter 3. Before turning to the model, I briefly review the evidence on firms' productivity dynamics to motivate the model setup.

I split the evidence on firms' productivity dynamics into productivity differentials between new and incumbent firms and into post-entry growth of firms. The evidence on productivity differentials between new and incumbent firms suggests that new firms or new plants have 75 to 95 percent of the productivity of incumbents.[3] Productivity differentials of this magnitude reemerge across different time periods and across countries. Among others, Baily, Hulten, Campbell, Bres-

[3] Geroski (1995), Caves (1998), and Bartelsman and Doms (2000) survey evidence from longitudinal micro data.

nahan, and Caves (1992), Baldwin (1996), Foster, Haltiwanger, and Krizan (2001), Aw, Chen, and Roberts (2001), Farias and Ruano (2005), Baldwin and Gu (2006) and Wagner (2010a) examine productivity differentials in manufacturing. Foster, Haltiwanger, and Krizan (2001) also examine services.

The evidence on the post-entry productivity growth of firms or plants suggests that the productivity of surviving entrants grows between 2 and 3 percent per year. Thus, within a decade or so, the productivity of a new firm grows to a level similar to average productivity. Baily, Hulten, Campbell, Bresnahan, and Caves (1992) show that US manufacturing plants that were established between 1972 and 1977 and that were initially of below-average productivity attained above-average productivity by 1987. Foster, Haltiwanger, and Krizan (2001) estimate show that the productivity of US manufacturing firms grows at 2 percent per year. Huergo and Jaumandreu (2004) estimate that the productivity of Spanish manufacturing firms grows at 3 percent per year. Baldwin and Gu (2006) estimate that productivity of Canadian manufacturing firms grows at roughly 2 percent per year. The remainder of this section describes a model that is consistent with this evidence.

2.1 Firms

Firms are indexed by $j \in [0,1]$, and each firm produces a single product variety. As in Ghironi and Melitz (2005), I thus treat a firm as a production line for this particular variety. The technology of firm j needed to produce output y_{jt} is

$$y_{jt} = a_t g^{s_{jt}} \ell_{jt} \ .$$

Productivity growth $g \geq 1$ is common to all firms and independent of firm size. That firm growth is independent of firm size is Gibrat's law and a first approximation to the data. The integer variable $s_{jt} = 0, 1, 2, \ldots$ indicates the firm's age since market entry, and a new firm has $s_{jt} = 0$. When g equals unity, all firms are equally productive, as in the basic New Keynesian model. When g exceeds unity, established firms are more productive than young firms, as in the model used in Melitz (2003). The unique production input ℓ_{jt} is labor, which the firm hires in a perfectly-competitive market. Aggregate productivity a_t is a stationary stochastic process with mean $a > 0$. Growth in a_t is distinct from growth in firm-specific productivity and does not affect conclusions.

The assumption that firms achieve productivity gains mechanically over time reflects that the longer a firm uses its technology, the more efficient is its production. In reality, efficiency gains may occur through learning by doing, economies of scale, process innovation, or changes in the product mix of a background firm.

Firms enter and exit the economy in each period. At the beginning of a period, $\delta \in [0,1)$ new firms enter the economy. At the end of a period, δ firms exit the economy. Exiting firms are selected randomly.[4] Accordingly, firms with high and low levels of productivity are equally exposed to exit. In reality, exit of a firm with high productivity may occur because of a major shift in consumer taste, new regulation, or product liability; because the firm's market share is taken over by a new firm which produces a close substitute; or because the firm starts exporting and stops selling in the home market. Baily, Hulten, Campbell, Bresnahan, and Caves (1992) find that highly productive firms frequently exit at the industry level because they switch to different industries to optimize their product mix. Bernard, Redding, and Schott (2010) analyze the product-switching activities of US manufacturing firms and find that product switching enhances firms' efficiency. Along these lines, my assumption of nonselective exit is best interpreted as capturing both firms that switch industries and firms that die.

When a new firm enters the economy, it sets a price for its product. In subsequent periods, the firm resets its price with probability $(1-\alpha)$, $\alpha \in [0,1)$, each period until the firm exits the market. Firm j chooses its nominal price P_{jt} to solve

$$\max_{P_{jt}} \; E_t \sum_{i=0}^{\infty} (\kappa\beta)^i \; \tilde{\Omega}_{t,t+i} \left[P_{jt} y_{jt+i} - W_{t+i} \ell_{jt+i} \right]$$

$$y_{jt+i} = \left(P_{jt}/P_{t+i} \right)^{-\theta} y_{t+i}$$

$$y_{jt+i} = a_{t+i} g^{s_{jt+i}} \ell_{jt+i} \; .$$

Here, $\beta^i \tilde{\Omega}_{t,t+i}$ discounts nominal payoffs and $\kappa = \alpha(1-\delta)$ is the probability to produce tomorrow at old prices. The first constraint is demand for product j derived below, where P, W, and y denote aggregate price level, nominal wage and aggregate output, respectively.[5] The optimal

[4] In related work (Weber (unpublished)), I confirm that optimal long-run inflation is positive and equal to growth in firm-specific productivity in a version of this model with endogenous firm entry as in Bilbiie, Ghironi, and Melitz (2008) and Bilbiie, Ghironi, and Melitz (2010). However, endogenous firm entry requires one additional state variable to capture the dynamics of the number of firms. This state variable is absent in the basic New Keynesian model such that, by construction, the basic model cannot be observationally equivalent to models with endogenous entry. While endogenous firm entry is interesting it is not the main focus of this paper.

[5] It is a common assumption in models with heterogenous firms such as Melitz (2003) and Ghironi and Melitz (2005) that wages are identical across firms.

price equates the expected discounted sum of marginal revenues to the expected discounted sum of marginal costs. Once rearranged, this price satisfies

$$P^{\star}_{jt}\; g^{s_{jt}} = \frac{\theta}{\theta-1}\frac{E_t\sum_{i=0}^{\infty}(\kappa\beta/g)^i\;\tilde{\Omega}_{t,t+i}P^{\theta}_{t+i}\,(y_{t+i}/a_{t+i})\,W_{t+i}}{E_t\sum_{i=0}^{\infty}(\kappa\beta)^i\;\tilde{\Omega}_{t,t+i}P^{\theta}_{t+i}y_{t+i}}\;. \tag{1}$$

Here, I write s_{jt+i} as $i+s_{jt}$ because both a firm's age s_{jt} and index i are integers. The discounting effect of productivity growth g on the right-hand side indicates that, in future periods, higher productivity reduces marginal costs. Crucially, the right-hand side is independent of the firm index j.

Firms differ along two dimensions, the level of productivity and the date of the last price change (length of price spell). Both dimensions affect the optimal price (1). According to (1), firms maintain the same optimal price if they are identical along both dimensions. For aggregation, I replace firm index j by two new indices each representing one dimension of heterogeneity. One index refers to date of entry, whereas the other index refers to date of last price change. Denote the current price P_{jt} of firm j as

$$P_{jt} = P^{\star}_{t-(n+k),t-k}\;, \quad n=0,1,2,\ldots,\; k=0,1,2,\ldots\;.$$

Accounting for the date of entry, P_{jt} equals either the optimal price of the current period or the optimal price of some previous period. The first subscript $t-(n+k)$ indicates the date of market entry. The second subscript $t-k$ indicates the date of the last price change. Index n denotes the time between entry and price change.

It follows from pricing equation (1) that, for any two firms j and j', optimal time t prices are proportional:

$$P^{\star}_{jt} = g^{(s_{j't}-s_{jt})}\;P^{\star}_{j't}\;.$$

Let j denote the young firm, $s_{j't} > s_{jt}$, such that $s_{j't}-s_{jt}$ is a positive integer. The equation states that the price of the young firm exceeds the price of the old firm if $g>1$. Specifically, let firm j be of age k and have a price spell of k periods. Let firm j' be of age $n+k$ and have a price spell of k periods. Then, $P^{\star}_{t-k,t-k} = g^n\;P^{\star}_{t-(n+k),t-k}$. Optimal prices are proportional but not equal because the old firm is more productive and hence maintains a lower price.

2.2 Households

There are a large number of identical households. The representative household maximizes discounted expected lifetime utility:

$$\max_{\{Q_t, B_t, \ell_t, c_{jt}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left[u(c_t) - h(\ell_t) \right] \ , \quad 0 < \beta < 1 \ . \tag{2}$$

The flow budget constraint is

$$E_t[\Omega_{t,t+1} Q_{t+1}] + B_t + \int_0^1 P_{jt} c_{jt} \ dj \quad \leq Q_t + (1 + i_{t-1}) B_{t-1} + (1 - \tau_L) W_t \ell_t + T_t + D_t \ .$$

At time t, the representative household selects a financial portfolio of nominal claims that expire in period $t+1$, against other households. Financial markets are complete. Then, the price in period t of the portfolio with random payoff Q_{t+1} at time $t+1$ is $E_t[\Omega_{t,t+1} Q_{t+1}]$, where $\Omega_{t,t+1}$ is a unique stochastic discount factor to be determined. The household holds government bonds B_t which yield nominal interest rate i_t. The household spends on consumption and receives labor income $(1 - \tau_L) W_t \ell_t$, net of taxes. T_t denotes a lump-sum transfer and D_t denotes profits from ownership of firms. Household preferences over intermediate products j are $c_t = (\int_0^1 c_{jt}^{\frac{\theta-1}{\theta}} \ dj)^{\frac{\theta}{\theta-1}}$ with $\theta > 1$. Product demand is $c_{jt}/c_t = (P_{jt}/P_t)^{-\theta}$ and the cost-minimal price of c_t is $P_t = (\int_0^1 P_{jt}^{1-\theta} \ dj)^{\frac{1}{1-\theta}}$. Terminal conditions (not shown) require household solvency.

2.3 Government

Government consumption q_t is exogenous and is the same bundle of individual products j as c_t. The government budget constraint is $(1 + i_{t-1}) B_{t-1} + T_t + P_t q_t \leq B_t + \tau_L W_t \ell_t$. The government issues a riskless one-period bond B_t and collects labor taxes τ_L. Government expenses consist of debt service $(1 + i_{t-1}) B_{t-1}$, consumption expenditure $P_t q_t$ and transfers T_t to the representative household. Government income consists of new debt B_t and income from taxing labor. Terminal conditions (not shown) require government solvency. The government controls the nominal interest rate i_t.

2.4 Competitive Equilibrium and Aggregation

In equilibrium, firms set prices according to equation (1); the household maximizes lifetime utility (2) subject to the budget constraint (3) and the definition of composite consumption c_t; product markets clear such that $y_{jt} = c_{jt} + q_{jt}$; the labor market clears such that $\ell_t = \int_0^1 \ell_{jt} \ dj$; and bond and stock markets also clear. The government sets τ_L, complies with its budget constraint, and conducts monetary policy as specified below. The aggregate resource constraint $y_t = c_t + q_t$ holds and aggregate output is defined as $y_t = (\int_0^1 y_{jt}^{\frac{\theta-1}{\theta}} dj)^{\frac{\theta}{\theta-1}}$. The aggregate price level and aggregate technology are as follows.

Price Level. The unit mass of firms consists of many current and past entry cohorts,

$$1 = \sum_{s=0}^{\infty} (1-\delta)^s \delta \ .$$

Consider the cohort that entered $s \geq 0$ periods ago at time $t-s$, and normalize its mass to unity. At time t, the weighted average price $\Lambda_{1t}(s)$ of this cohort is

$$\Lambda_{1t}(s) = \begin{cases} (1-\alpha)\sum_{k=0}^{s-1} \alpha^k (P^\star_{t-s,t-k})^{1-\theta} + \alpha^s (P^\star_{t-s,t-s})^{1-\theta} & \text{if} \quad s \geq 1 \ , \\ (P^\star_{t,t})^{1-\theta} & \text{if} \quad s = 0 \ . \end{cases}$$

Rewrite the price level $P_t^{1-\theta} = \int_0^1 P_{jt}^{1-\theta} dj$ as sum over cohort prices weighted by cohort size,

$$P_t^{1-\theta} = \sum_{s=0}^{\infty} (1-\delta)^s \delta \Lambda_{1t}(s) \ .$$

Plug in for $\Lambda_{1t}(s)$ and simplify to obtain a recursive expression for the price level,

$$P_t^{1-\theta} = n_\gamma (1-\kappa\gamma)(P^\star_{t,t})^{1-\theta} + \kappa P_{t-1}^{1-\theta} \ . \tag{3}$$

Here, I define $\gamma = g^{\theta-1}$ and $n_\gamma = \delta/[1-(1-\delta)\gamma]$ and impose the condition $(1-\delta)\gamma < 1$ to obtain a finite price level. I return to this condition below. Transforming the price level into inflation delivers

$$1 = n_\gamma (1-\kappa\gamma)\,(p_t^\star)^{1-\theta} + \kappa \pi_t^{\theta-1} \ , \tag{4}$$

with $\pi_t = P_t/P_{t-1}$ and $p_t^\star = P^\star_{t,t}/P_t$.

Aggregate Technology. Combine labor market clearing with firm technology and product demand to link aggregate labor to aggregate output,

$$y_t = a_t \ell_t / \Delta_t \ .$$

This relationship is aggregate technology. Δ_t indicates the loss in the end-use value of output from dispersion in prices and productivity,

$$\Delta_t = \int_0^1 g^{-s_{jt}} p_{jt}^{-\theta} \, dj$$

defining p_{jt} as the relative price of firm j. Productivity dispersion remains constant over time because firm turnover is constant. Price dispersion varies over time because firms set prices depending on the time-varying state of the economy. Moreover, price dispersion is partly driven by productivity dispersion. In Yun (2005), Δ_t is driven by price dispersion only, and price dispersion arises exclusively from staggered pricing. Going through steps analogous to those used when deriving a recursive expression for the price level yields

$$\Delta_t = n_\gamma(1-\kappa\gamma)\,(p_t^\star)^{-\theta} + (\kappa/g)\pi_t^\theta \Delta_{t-1} \ . \tag{5}$$

Appendix A summarizes the equilibrium conditions of the competitive economy with sticky prices. In what follows, I frequently compare the FIP and the HOP models. This comparison involves changing two parameters, productivity growth g and the rate of firm turnover δ. However, to highlight the role of g, I keep δ unchanged throughout. Thus, I compare two models with firm turnover, one with homogenous productivity and one with firm-specific productivity. The firm-turnover model with homogenous productivity is equal to the basic model without firm turnover once I interpret $\kappa = \alpha(1-\delta)$ rather than α as the probability that a firm cannot adjusts its price. This interpretation constitutes a useful reference.

3 Analyzing Optimal Long-Run Inflation

The goal of this section is to determine the rate of long-run inflation that reconciles the steady state of the sticky-price model with the steady state of the efficient flexible-price model. This inflation rate is of interest in itself and because the sticky-price model is typically linearized at the efficient flexible-price steady state, and I shall treat the FIP model accordingly.[6] I start by solving

[6]Woodford (2003), chapter 6 and 7, discusses several reasons for pursuing this approach. First, a compact and intuitive way to understand the goals of welfare-maximizing monetary policy is to express the sticky-price model as a deviation from its corresponding flexible-price model. For this to be feasible, both models need to be linearized at the same steady state. Second, for the loss function accurate to second order to appropriately rank different policies based on a model accurate to first order, first-order terms must not appear in this loss function. One approach to ensure this is to expand utility near the efficient allocation. Third, in order to be able to assess the policies of interest accurately enough, these policies have to imply that model variables remain near

the planner problem of the model described so far for the efficient allocation. Then, I reconcile the flexible-price competitive equilibrium with the efficient allocation. Finally, I reconcile the sticky-price competitive equilibrium with the efficient flexible-price competitive equilibrium. All the derivations are valid for $g \geq 1$ and encompass the HOP model as a special case.

In the HOP model, two distortions push the sticky-price economy away from the efficient allocation: monopolistic competition among price setters and the fact that not all product prices adjust each period such that relative prices are potentially distorted. A labor tax is sufficient to offset the distortion from monopolistic competition, and zero long-run inflation copes with distortions in relative prices. As it turns out, in the FIP model, the same labor tax offsets the distortion from monopolistic competition, whereas positive rather than zero long-run inflation is required to account for distortions in relative prices.

3.1 Planner Equilibrium

I divide the solution of the planner equilibrium (PE) into two stages. The first stage determines the efficient allocation of labor across firms and aggregate technology that links aggregate output to aggregate labor. The second stage resolves the tradeoff between work and consumption. It is necessary to consider the first stage upfront because, with productivity growth $g > 1$, firms are heterogenous. The two-stage planning problem resembles the two-stage consumer problem in the competitive equilibrium (CE).

Aggregate Technology. The efficient allocation of aggregate labor ℓ_t^e across firms, with ℓ_t^e given, solves the problem:

$$\max_{\ell_{jt}^e} \quad \left(\int_0^1 (y_{jt}^e)^{\frac{\theta-1}{\theta}} \; dj \right)^{\frac{\theta}{\theta-1}} \qquad s.t. \quad y_{jt}^e = a \; g^{s_{jt}} \ell_{jt}^e \quad , \quad \ell_t^e = \int_0^1 \ell_{jt}^e \; dj$$

Rearranging optimality conditions yields

$$\ell_{jt}^e = \frac{g^{(\theta-1)s_{jt}}}{\int_0^1 g^{(\theta-1)s_{jt}} dj} \; \ell_t^e \; .$$

Firm j's efficient amount of labor ℓ_{jt}^e is proportional to total labor, ℓ_t^e. The factor of proportionality is related to firm j's productivity relative to average productivity. If productivity is homogenous across firms with $g = 1$, then labor is equally distributed across firms. If, how-

the steady state at which the model is linearized. One may guess (and later verify) that policies close to the efficient allocation fulfill this requirement.

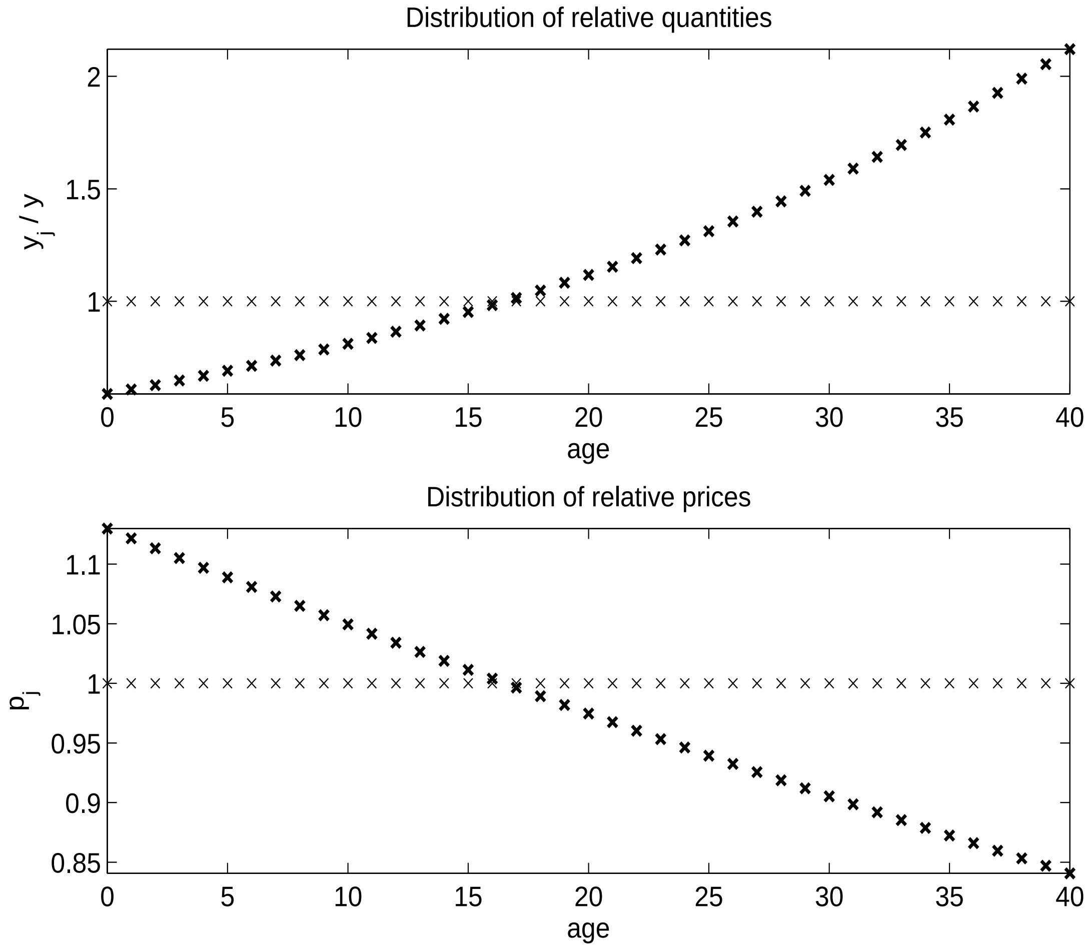

Figure 1: Distribution of relative quantities and relative prices consistent with the efficient allocation. Age is in quarters. Thick crosses corresponds to the FIP model with $g = 1.03^{1/4}$. Thin crosses correspond to the HOP model with $g = 1$.

ever, productivity is firm-specific with $g > 1$, then young firms employ less labor than average. Aggregation yields aggregate technology:

$$y_t^e = a_t \; \ell_t^e / \Delta^e \; .$$

The shift in output, which results from productivity dispersion and productivity growth, is constant over time and equals

$$\frac{1}{\Delta^e} = \left(\int_0^1 g^{(\theta-1)s_{jt}} dj \right)^{\frac{1}{\theta-1}} = \left(\frac{\delta}{1-(1-\delta)\gamma} \right)^{\frac{1}{\theta-1}} \; , \tag{6}$$

with $\gamma = g^{\theta-1}$. The output shifter $1/\Delta^e$ summarizes two effects. First, there is output loss because goods are produced in different quantities. Second, there is output gain because firms become more productive over time.

Figure 1 plots the distribution of relative output y^e_{jt}/y^e_t as a function of a firm's age s_{jt}. For $g = 1$, distribution is uniform because all firms maintain the same technology and consumers derive the greatest value from aggregate consumption when products enter the composite in equal amounts. For $g > 1$, the planner faces a tradeoff because firms maintain different technologies. The planner has to trade off between a symmetric allocation of products favored by consumers and the fact that established products are produced more effectively than young products. The optimal resolution of this tradeoff is the distribution of relative quantities in Figure 1. Young firms produce below-average output and established firms produce above-average output.

Tradeoff between Labor and Consumption. To resolve the tradeoff between labor and consumption, the planner maximizes lifetime utility subject to aggregate technology and aggregate resource constraint:

$$\max_{\{c^e_t,\ell^e_t\}^\infty_{t=0}} \quad E_t \sum_{t=0}^{\infty} \beta^t [u(c^e_t) - h(\ell^e_t)] \quad s.t. \quad y^e_t = a_t \ell^e_t / \Delta^e \quad , \quad c^e_t = y^e_t - q_t \ .$$

Optimality requires that the marginal rate of transformation equals the marginal rate of substitution. To sum up, the PE consists of the two equations

$$\frac{h_\ell(\ell^e_t)}{u_c(c^e_t)} = \frac{a_t}{\Delta^e} \ , \qquad c^e_t + q_t = \frac{a_t \ell^e_t}{\Delta^e} \tag{7}$$

and the output shifter $1/\Delta^e$ in equation (6).

3.2 Competitive Equilibrium with Flexible Prices

Here, I derive the labor tax that reconciles the flexible-price CE with the PE. The flexible-price CE is the special case $\alpha = 0$ of the sticky-price CE summarized in Appendix A. It consists of the two equations

$$(1 - \tau_L)^{-1} \frac{\theta}{\theta - 1} \frac{h_\ell(\ell^f_t)}{u_c(c^f_t)} = \frac{a_t}{\Delta^f_t} \ , \qquad c^f_t + q_t = \frac{a_t \ell^f_t}{\Delta^f_t} \ . \tag{8}$$

Δ^f_t is the special case of equation (5) with flexible prices $\alpha = 0$. By this equation, the output shifter in the flexible-price CE is constant and equal to the output shifter in the PE:

$$\Delta^f_t = \Delta^e \ .$$

The finding suggests that the inefficiency from monopoly power does not distort the allocation of labor across firms in the flexible-price CE because monopoly power is the same across firms.

To show that this is indeed true, consider first the allocation of labor across firms in the flexible-price CE. Combine household demand and technology of product j to obtain $\ell^f_{jt} = a_t^{-1} g^{-s_{jt}} (P^{f\star}_{jt}/P^f_t)^{-\theta} y^f_t$. Divide by the analog condition for another firm and employ the cross-section recursion of prices derived from pricing equation (1) to obtain

$$\ell^f_{jt}/\ell^f_{j't} \quad = \quad g^{(\theta-1)[s_{jt}-s_{j't}]} \ .$$

In the PE, the ratio of labor of any two firms derives from the optimality condition of the labor-allocation problem and is equal to

$$\ell^e_{jt}/\ell^e_{j't} \quad = \quad g^{(\theta-1)[s_{jt}-s_{j't}]} \ .$$

Accordingly, the allocation of labor across firms in the flexible-price CE is efficient.

It is then straightforward to determine that a labor subsidy equal to $\tau^e_L = -1/(\theta-1)$ fully reconciles the flexible-price CE with the PE in equation (7) because this subsidy offsets the effect of firms' monopoly power. Moreover, this subsidy is the same in the FIP and the HOP model because decentralization does not distort the cross-section allocation of labor.

3.3 Competitive Equilibrium with Sticky Prices

I now determine the long-run inflation rate that reconciles the steady state of the sticky-price CE with the steady state of the flexible-price CE. The FIP model with sticky prices exhibits a well-defined aggregate steady state in which aggregate variables do not grow.[7] This is true despite the fact that firm-specific productivity grows at a positive rate because two polar forces, an expanding force and a contracting force, balance each other. The expanding force is growth in firm-specific productivity, which induces output of the average firm per cohort to grow. The contracting force is firm turnover. Exiting firms are randomly drawn such that the sample of exiting firms replicates the productivity distribution of the firm population. However, exiting

[7]By aggregate steady state I mean that aggregate shocks a_t and q_t are equal to their unconditional means, whereas both idiosyncratic shocks, firm exit and Calvo lottery, continue to operate. Steady-state output remains constant under two standard assumptions. First, aggregate labor does not grow $\ell_t = \ell$. Second, the central bank ensures constant long-run inflation $\pi_t = \pi$. Without loss of generality, I also assume that both exogenous processes a_t and q_t are stationary. Constant inflation implies a constant relative price of new firms $p^\star$ by equation (4). Constant inflation and constant relative price imply by equation (5) that the output shifter $1/\Delta$ remains constant. Using aggregate technology, it follows then that output is constant. See Appendix B for further details.

firms are replaced by new firms with low productivity.

The following three equations are derived from the CE with sticky prices in Appendix B and determine steady-state levels of Δ, c, ℓ conditional on π:

$$\frac{1/\Delta}{1/\Delta^e} = \left(\frac{1-\kappa\pi^\theta/g}{1-\kappa g^{\theta-1}}\right)\left(\frac{1-\kappa g^{\theta-1}}{1-\kappa\pi^{\theta-1}}\right)^{\frac{\theta}{\theta-1}} \tag{9}$$

$$(1-\tau_L)^{-1}\frac{\theta}{\theta-1}\frac{h_\ell(\ell)}{u_c(c)} = \frac{a}{\Delta^e}\left(\frac{1-\kappa\pi^{\theta-1}}{1-\kappa g^{\theta-1}}\right)^{\frac{1}{1-\theta}}\left(\frac{1-\kappa\beta\pi^\theta/g}{1-\kappa\beta\pi^{\theta-1}}\right) \tag{10}$$

$$c+q = a\ell/\Delta\ . \tag{11}$$

There are two distortions from inflation that do not arise in the flexible-price CE in equations (8). First, in equation (9), inflation alters the output shifter $1/\Delta$ in the sticky-price CE relative to $1/\Delta^e$. Second, in equation (10), inflation distorts the marginal rates of transformation and substitution. This distortion is unrelated to the wedge from monopolistic competition $\frac{\theta}{\theta-1}$ which also exists in the flexible-price CE and can be eliminated by imposing $\tau_L = \tau_L^e$. I summarize the first main result of the paper in the following proposition.

Proposition 1: Long-run inflation equal to growth in firm-specific productivity $\pi = g$ and the labor tax $\tau_L = \tau_L^e$ reconcile the sticky-price CE with the efficient flexible-price CE.

To prove this claim, evaluate equations (9)–(11) at $\pi = g$ and $\tau_L = \tau_L^e$ to obtain

$$\Delta = \Delta^e\ , \qquad \frac{h_\ell(\ell)}{u_c(c)} = \frac{a}{\Delta^e}\ , \qquad c+q = \frac{a\ell}{\Delta^e}\ .$$

Comparing these equations to the efficient flexible-price CE in steady state shows that $\pi = g$ eliminates both distortions from long-run inflation. This completes the proof.

It is instructive to consider each distortion from inflation in greater detail to understand why this particular inflation rate restores efficiency. I start with the distortion between marginal rates of transformation and substitution. A related analysis for the HOP model is contained in Goodfriend and King (1998) and King and Wolman (1999).

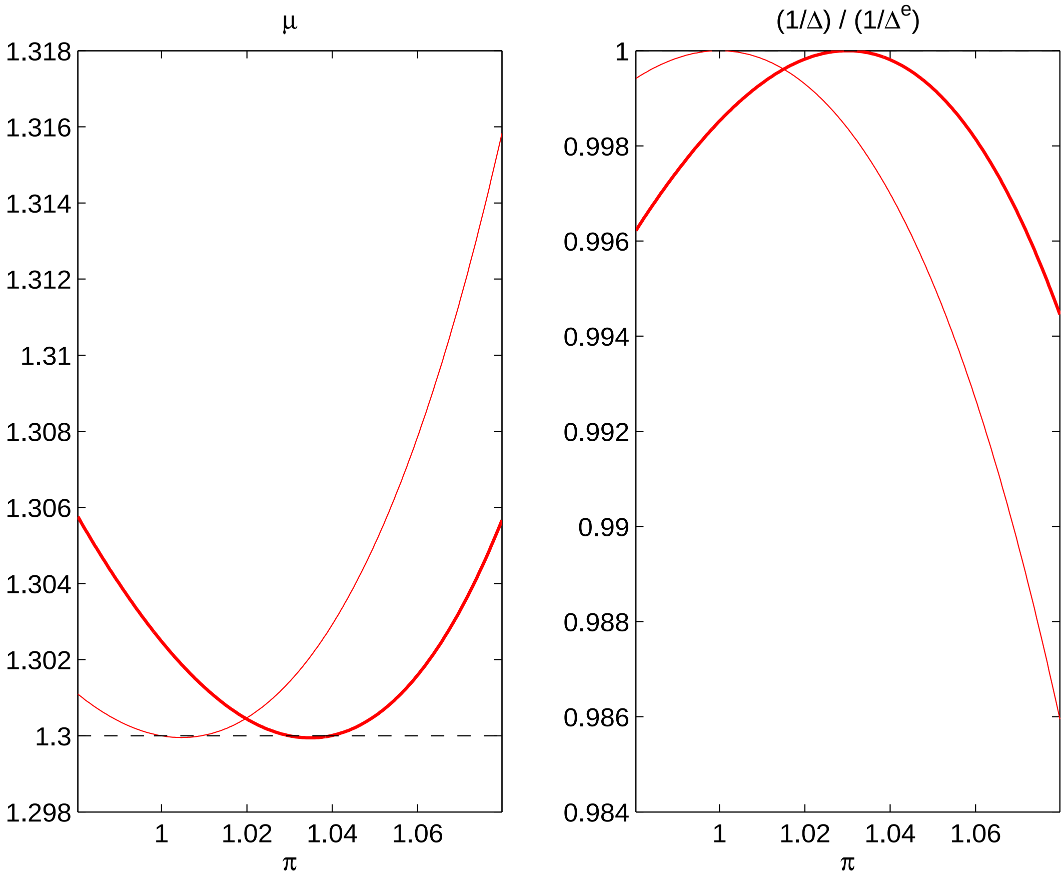

Figure 2: Average markup μ and output shifter relative to efficient output shifter $(1/\Delta)/(1/\Delta^e)$ plotted as a function of long-run inflation. Long-run inflation is annualized. Thick lines correspond to the FIP model with $g = 1.03^{1/4}$. Thin lines correspond to the HOP model with $g = 1$. Dashed line corresponds to static markup $\frac{\theta}{\theta-1}$.

Average Markup

It is straightforward to map the distortion between marginal rates of transformation and substitution into the average markup. Define the average markup as the inverse of real marginal costs of a firm with average efficient productivity, $\mu = a/(w\Delta^e)$. Equation (10) jointly with the optimality condition $w = (1-\tau_L)^{-1}h_\ell(\ell)/u_c(c)$ delivers the average markup as a function of inflation only,

$$\mu(\pi) \quad = \quad \frac{\theta}{\theta-1}\left(\frac{1-\kappa\beta\pi^{\theta-1}}{1-\kappa\beta\pi^{\theta}/g}\right)\left(\frac{1-\kappa g^{\theta-1}}{1-\kappa\pi^{\theta-1}}\right)^{\frac{1}{1-\theta}} \quad . \tag{12}$$

The left panel of Figure 2 plots $\mu(\pi)$ for the two cases $g > 1$ and $g = 1$. The average markup coincides with the static markup $\frac{\theta}{\theta-1}$ for long-run inflation equal to g. The average markup exceeds the static markup for long-run inflation below g, is minimal for some long-run inflation slightly above g, and exceeds the static markup again when long-run inflation exceeds g.

Two effects impinge on the average markup and generate this behavior. The anticipated-inflation effect works through firms that do adjust their price and dominates the average markup for long-run inflation above g. The pass-through effect works through firms that do not adjust their price and dominates the average markup for long-run inflation below g. The anticipated-inflation effect is best described by using the pricing equation (1). The pass-through effect is best described by using the price level (3). Jointly, both equations imply equation (12).

Consider the anticipated-inflation effect first. For a new firm, the steady-state pricing equation can be rearranged as

$$0 = \sum_{i=0}^{\infty} (\kappa\beta)^i \pi^{\theta i} \left[\frac{p^\star}{\pi^i} - \frac{\theta}{\theta - 1} \frac{w}{ag^i} \right] . \tag{13}$$

The square brackets contain the difference between the optimal real price $p^\star/\pi^i$ and the desired real price $\frac{\theta}{\theta-1} w/(ag^i)$, which is equal to the static markup over marginal costs. Over time, the difference between the optimal real price and the desired real price evolves depending on the value of π relative to g. If π equals g then $p^\star/\pi^i$ is equal to the desired price each and every period. Thus, inflation equal to g erodes $p^\star$ at exactly the right pace for the firm to maintain the static markup.[8] However, if π exceeds g then $p^\star$ is eroded faster than marginal costs decline, and future markups are compressed to below the static markup. Firms choose elevated markups in initial periods to counterbalance markup compression in future periods. Thus, with π above g, elevated markups of adjusting firms lift the current average markup in Figure 2.

In equation (13), anticipated inflation also implies that future differences between the optimal and the desired price are weighted more because $\pi^{\theta i}$ enters the discount factor. I refer to this as the market-share effect. With inflation, $p^\star/\pi^i$ declines over time and increases the firm's market share. When the market share is high, any difference between the optimal and the desired price reduces profits greatly. Accordingly, the higher trend inflation is, the more the firm is concerned with getting its price right in future periods. However, the market-share effect alone does not prevent $p^\star/\pi^i$ from being equal to the desired price in each period.

Consider now the pass-through effect that captures by how much the markup of a new firm moves the average markup. Define the markup of a new firm or the "marginal" markup $\mu^\star(\pi)$

[8] The FIP model suggests a simple rationale for firms not to index prices to inflation. Namely, inflation erodes prices in line with growth in firm-specific productivity.

as $\bar{P}^{\star}_{t,t}/(\bar{W}_t/a)$. Equation (13) can be converted into the marginal markup:

$$\mu^{\star}(\pi) = \frac{\theta}{\theta - 1}\left(\frac{1 - \kappa\beta\pi^{\theta-1}}{1 - \kappa\beta\pi^{\theta}/g}\right) . \tag{14}$$

To link the marginal markup to the average markup it is useful to link $\bar{P}_t$ and $\bar{P}^{\star}_{t,t}$. The aggregate price level does exactly this:

$$\bar{P}_t = \Delta^e \bar{P}^{\star}_{t,t}\left(\frac{1 - \kappa g^{\theta-1}}{1 - \kappa\pi^{\theta-1}}\right)^{\frac{1}{1-\theta}} \tag{15}$$

Divide both sides by marginal costs of a new firm and rearrange to obtain the link between marginal and average markup:[9]

$$\mu = \mu^{\star}\left(\frac{1 - \kappa g^{\theta-1}}{1 - \kappa\pi^{\theta-1}}\right)^{\frac{1}{1-\theta}} .$$

The term in round brackets represents inefficient pass-through when long-run inflation π and firms' productivity growth g are misaligned. Suppose $\pi = 1$ and $g > 1$ such that pass-through of $\mu^{\star}$ into μ is excessively high. Absent inflation, markups of many incumbent firms are above the static markup because their real marginal costs declined whereas their real prices remained constant. Accordingly, these firms produce insufficient amount of output and, thereby, receive too little weight in the average markup that weights a firm's markup by its relative output. In contrast, firms with the marginal markup overproduce and receive too much weight in the average markup. Vice versa, when $\pi > 1$ and $g = 1$, pass-through of of $\mu^{\star}$ into μ is insufficient because firms with the marginal markup receive too little weight in the average markup.

Productivity- and Price-Dispersion

The second distortion created by long-run inflation in the CE with sticky prices arises from productivity- and price-dispersion. The left panel of Figure 2 plots $1/\Delta$ over $1/\Delta^e$ in equation (9) as a function of inflation. The efficient shifter $1/\Delta^e$ indicates the amount of dispersion that is indispensable because firms' productivity differs. The plot shows that most levels of inflation create excess dispersion in relative prices and, by product demand $\bar{y}_{jt}/y = (\bar{P}_{jt}/\bar{P}_t)^{-\theta}$, in relative quantities. Excess dispersion in relative quantities reduces the end-use value of output because consumers prefer evenly distributed products. Excess dispersion arises whenever inflation is not equal to productivity growth $\pi \neq g$. In this case, firms do not manage to maintain the static

[9]Substituting equation (14) for the marginal markup again delivers equation (12).

markup while keeping their nominal price fixed. Accordingly, firms reset prices whenever feasible. Reset prices disperse relative prices because only a subset of prices is adjusted each period.

If inflation is equal to g, however, firms do maintain the static markup while keeping their nominal price fixed.[10] In this case, price stickiness imposes no constraint on firms because firms do not benefit from changing their price in the first place, and the distribution of relative prices resembles the one in the flexible-price CE. That is, if the central bank supports firms to maintain the static markup, at the same time, it minimizes dispersion in relative quantities. One policy instrument eliminates both distortions, and staggered pricing entails no social costs.

4 Calibration

A core parameter of this analysis is growth in firm-specific productivity g. I set firm-specific productivity growth to 3 percent per year or, taking one period as a quarter, to $g = 1.03^{1/4}$. This value is in line with the evidence on post-entry productivity growth discussed in Section 2. To see what this value implies for the productivity differential between new and incumbent firms, I derive the productivity distribution implied by the model. Rather than moving straight to the productivity distribution, it is useful to briefly consider the age distribution of firms. The age distribution is $J(s) = \delta(1-\delta)^s$, with s denoting age in quarters. The productivity distribution has discrete support because productivity $\phi(s) = g^s$ grows with age s, and age is discrete. The inverse of $\phi(s)$ is $s(\phi) = \ln(\phi)/\ln(g)$. Plug this expression into the age distribution to obtain the productivity distribution $J(s(\phi))$ or

$$J(\phi) = \delta(1-\delta)^{\ln(\phi)/\ln(g)} \ .$$

It holds that $\sum_\phi J(\phi) = 1$ for all $\phi(s) = g^s$ and $s = 0, 1, \ldots$.

I compute the productivity differential between new firms and the average incumbent firm from $J(\phi)$. At $g = 1.03^{1/4}$, the productivity of a new firm is 94.26 percent the productivity of the average incumbent firm. In light of the evidence reviewed in Section 2, a productivity differential of below 6 percent is at the very conservative end of the spectrum of estimates. Figure 3 plots the productivity distribution. For illustration, the figure contains two productivity distributions, one with $g = 1.03^{1/4}$ (crosses), and one with $g = 1.01^{1/4}$ (dots). Circles indicate the mean level

[10] Long-run inflation equal to g maximizes the ratio $\frac{1/\Delta}{1/\Delta^e}$, and the maximal value of this ratio is unity. Thus, there is no way to compress $1/\Delta$ to above its efficient level by means of inflation.

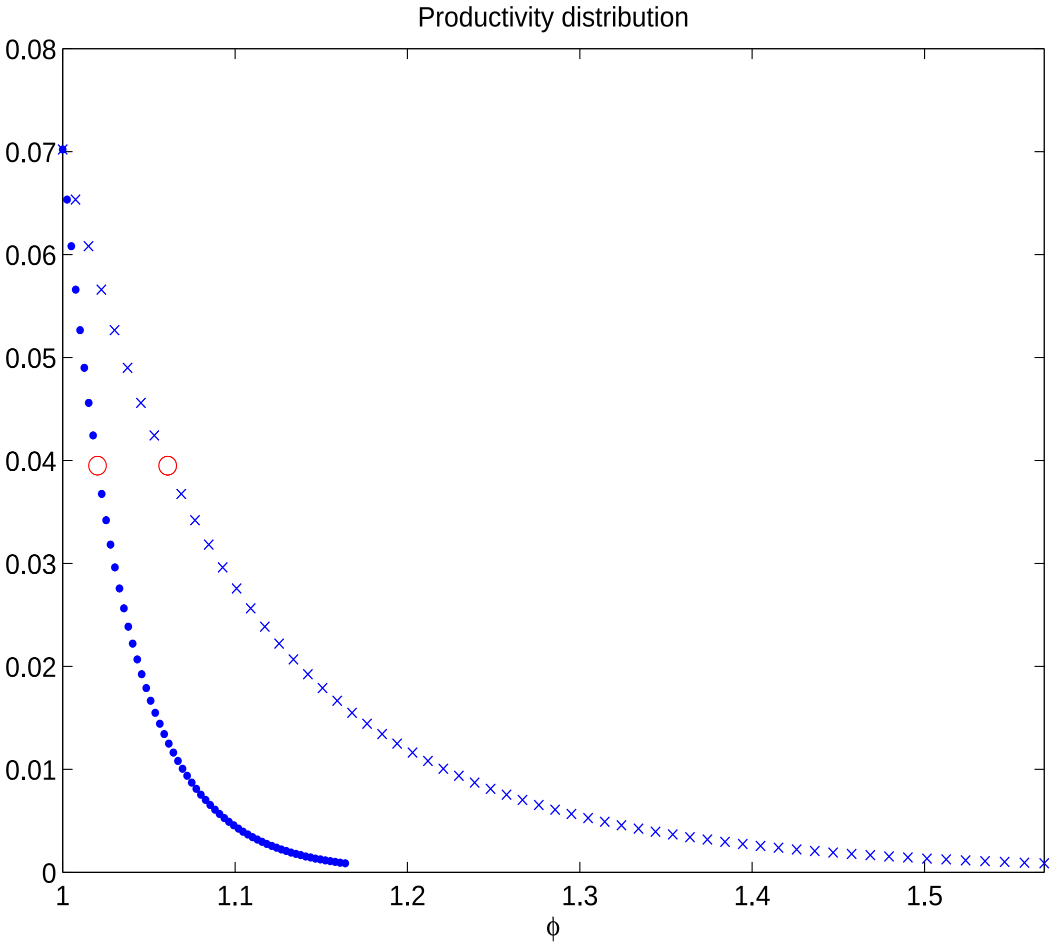

Figure 3: Productivity distribution $J(\phi)$ under the benchmark calibration $g = 1.03^{1/4}$ (crosses) and under the alternative calibration $g = 1.01^{1/4}$ (dots). Circles indicate the levels of productivity closest to the mean of the distribution.

of productivity. Evidently, the mean and dispersion increase in g, keeping the level of entry productivity constant. Young firms have below-average productivity and established firms have above-average productivity.

One alternative way to assess the calibration of g is to compute productivity differentials from estimated Pareto distributions that describe the productivity distribution in firm-level data well. To obtain the productivity differential in the Pareto case, suppose that the higher productivity ϕ is, the more established the firms are. Then, the productivity of a new firm is equal to the cutoff parameter ϕ_{min}, and the average productivity is equal to $E(\phi) = k\phi_{min}/(k-1)$, with

Pareto index $k \neq 1$.[11] The productivity differential is

$$\frac{\phi_{min}}{E(\phi)} = \frac{k-1}{k} \; .$$

Del Gatto, Mion, and Ottaviano (2006) estimate k close to 2 for the manufacturing sectors of 11 European countries. This value implies a productivity differential of 50 percent. Ghironi and Melitz (2005) calibrate k to 3.4 for a two-country model with endogenous entry into export markets.[12] In my setting, this value would imply a productivity differential of 29 percent. Overall, my calibration of g implies a very moderate productivity differential.

The remaining parameters are calibrated as follows. The flexible-price or static markup is 30 percent. This markup implies $\theta = 4\frac{1}{3}$, a value similar to the value used by King and Wolman (1996). Efficiency of the steady state requires subsidizing labor at rate $\tau_L^e = -0.3$. The probability α of not adjusting the price is 0.8059. The rate of product turnover δ is equal to 0.0694, which corresponds to the 25 percent product turnover per year reported in Broda and Weinstein (2010). Thus, $\kappa = \alpha(1-\delta)$ is equal to 0.75 such that the mean duration of a price spell in the censored distribution of prices is equal to four quarters. The discount factor β is set to 0.995. Period utility is parameterized as $u(c) = (c^{1-1/\sigma} - 1)/(1 - 1/\sigma)$ and $h(\ell) = \eta \ell^{1+\nu}/(1+\nu)$. I set σ equal to unity, which corresponds to the log-utility of consumption. The labor-supply elasticity ν is equal to 0.25. Then, the slope of the Phillips curve ζ is equal to 0.1061 in the FIP model and equal to 0.1269 in the HOP model.[13] The ratio of government consumption over output $s_q = q/y$ is 0.2, and $\eta = 3$ implies steady-state labor close to 0.5.

5 Output Effects of Long-Run Inflation

In order to illustrate the output effects of long-run inflation in the FIP model, I solve equations (9) to (11) for the function $y = y(\pi)$. Figure 4 plots the function for the benchmark calibration. In the FIP model (thick line), output increases with long-run inflation as long as inflation is zero

[11] Discrete $J(\phi)$ can be connected to the continuous Pareto distribution $f(\phi)$ by deriving parameters k and ϕ_{min}, such that $J(\phi) = f(\phi)$ for all $\phi = g^s$. However, it is misleading to interpret parameters obtained in such a way in light of the empirical data because $f(\phi)$ differ from $J(\phi)$. The difference is that the mass of J is not evenly spread over the continuous support of f because the values of ϕ grow exponentially and thus are not equidistant.

[12] Firms that enter export markets usually have higher productivity than firms that do not. This does not contradict the observation that new exporters tend to have lower productivity than incumbent exporters. Indeed, for West Germany, Wagner (2010b) shows that new exporters are less productive than incumbent exporters but are more productive than nonexporters (Table 3).

[13] For robustness, I also consider σ equal to 6.25, which is the value used in Woodford (2003), Table 6.1. This value implies that ζ is equal to 0.0318 in the FIP model and ζ is equal to 0.0381 in the HOP model, which is in line with estimates in Cecioni (2010) or Altig, Christiano, Eichenbaum, and Linde (undated).

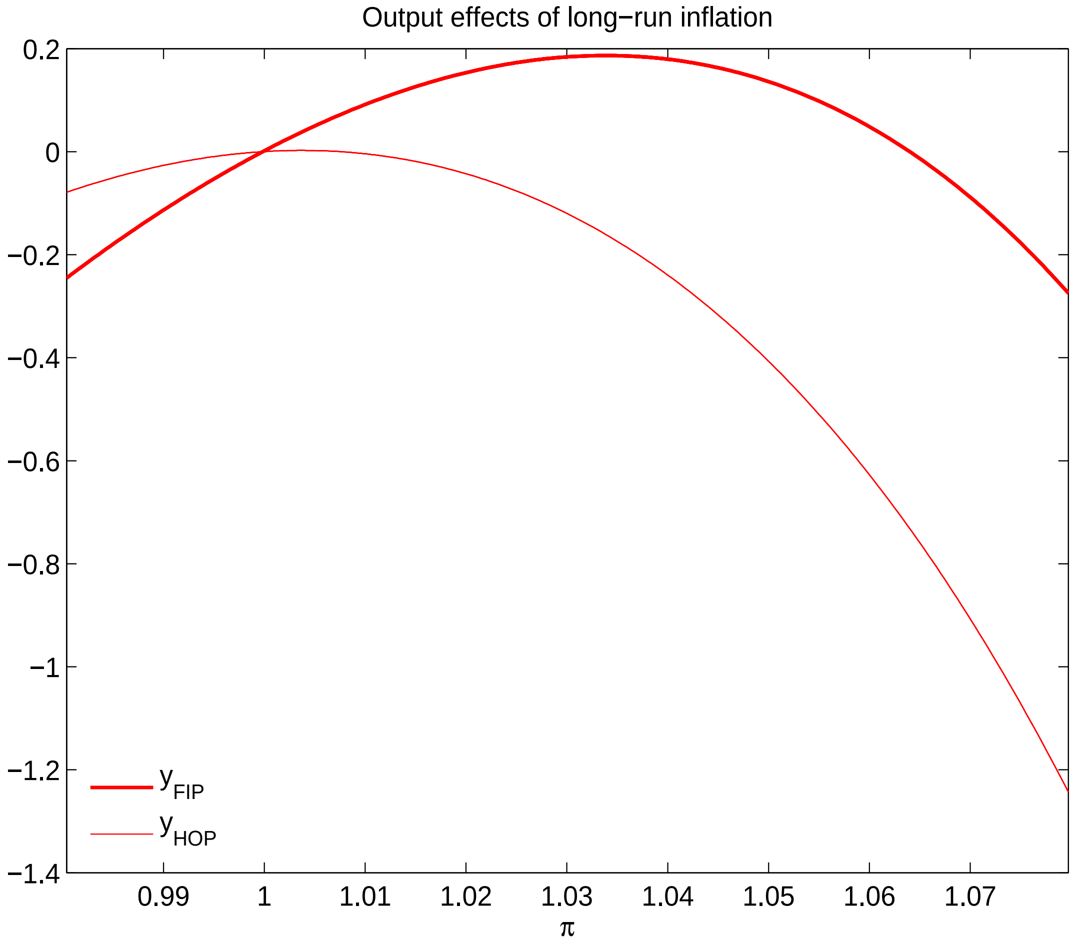

Figure 4: Output as a function of long-run inflation in the FIP model and the HOP model. Output is the percentage deviation from the level of output when long-run inflation is zero. Long-run inflation is annualized. Thick line corresponds to FIP model with $g = 1.03^{1/4}$. Thin line corresponds HOP model with $g = 1$.

or moderately positive; the relationship inverts for high levels of long-run inflation. In the HOP model (thin line), output falls for inflation to slightly above zero and higher.

Quantitatively, the output effects of long-run inflation are moderate in the FIP model but large in the HOP model. In Figure 4, output in the FIP model falls by 0.27 percent when long-run inflation increases from zero to 8 percent per year, whereas output in the HOP model falls by 1.24 percent. With θ equal to 8 rather than $4\frac{1}{3}$, output in the FIP model falls by 1.29 percent, whereas output in the HOP model falls by 3.92 percent when annual inflation increases from zero to 8 percent. With σ equal to 6.25 rather than unity and θ equal to $4\frac{1}{3}$, output in the FIP model falls by 0.84 percent, whereas output in the HOP model falls by 3.74 percent. These large output effects of long-run inflation in the HOP model are in line with results in Ascari (2004) and Bakhshi, Khan, Burriel-Llombart, and Rudolf (2007).

It has been argued that the large output effects of long-run inflation in the HOP model are implausible. They arise because positive inflation erodes sticky nominal prices. However, price erosion provides a very strong incentive for firms to change their prices when long-run inflation exceeds certain levels.[14] In the model of Levin and Yun (2007), firms select the frequency of price adjustment endogenously. Whereas moderate inflation implies a nontrivial reduction of output in this model, this effect disappears at high inflation rates because firms decide to adjust their prices more often. In the model of Dotsey, King, and Wolman (1999) with state-dependent pricing, positive long-run inflation makes it more likely that each firm will adjust its price in a given period. Moreover, it is more likely that firms with outdated prices will adjust, and the maximum duration of price spells will be shortened endogenously. The HOP model with time-dependent Calvo pricing fails to capture these effects of long-run inflation.

The output effects of positive long-run inflation are much more moderate and thus more plausible in the FIP model than in the HOP model. In the FIP model, firms actually benefit from positive inflation being equal to growth in firm-specific productivity because this amount of price erosion enables firms to maintain the static markup at all times despite sticky nominal prices. One consequence of this is that, in the FIP model, Calvo pricing remains a parsimonious benchmark assumption for levels of long-run inflation consistent with the inflation objectives of central banks in many industrialized economies.

Clearly, most levels of inflation in Figure 4 lead to an inefficient steady state because g is kept constant while π varies. For the inefficient steady state to exist, inflation cannot exceed $\pi < (g/\kappa)^{1/\theta}$. This upper bound ensures a finite optimal price $p^\star$ in equation (1). For the benchmark calibration, annual inflation above 31.31 percent conflicts with the upper bound.[15] The limit falls to 15.90 percent with θ equal to 8. Inflation rates above these values have been extremely rare in industrialized economies in recent decades.

In the FIP model, there is also an upper bound on long-run inflation in *efficient* steady state which is distinct from the bound in *inefficient* steady state. The additional bound arises because efficiency requires $\pi = g$, and g has to obey $g < (1-\delta)^{\frac{1}{1-\theta}}$. This bound derives from equation (4) and guarantees new firms a positive market share. Appendix B provides further details. For

[14]Graham and Snower (2004) argue along these lines for the case of sticky wages. Kiley (2002) considers a variation of the theme, namely, that welfare costs of long-run inflation are much larger in a model with Calvo pricing than in a model with Taylor pricing. One approach to mitigate the effects of inflation in steady state is price indexation (Yun (1996)). Cogley and Sbordone (2008) argue that price indexation is not consistent with micro data on prices.

[15]The upper bound on inflation, which to some extent is an artefact of pure Calvo pricing, can be weakened or entirely overcome by assuming endogenous price stickiness as in Bakhshi, Khan, Burriel-Llombart, and Rudolf (2007), by assuming pricing contracts as in Taylor (1980), or by assuming truncated Calvo pricing as in Amano, Ambler, and Rebei (2007).

my benchmark calibration, the upper bound on efficient long-run inflation is close to 9 percent per year, and encompasses the 3 percent efficient long-run inflation implied by this calibration. Whereas the upper bound on efficient long-run inflation is loose for intermediate values of θ, it tightens for high values of θ.[16] High values of θ reduce the market share of new products because consumers are more willing to substitute away from new products with a high relative price. Accordingly, when θ equals 8 rather than $4\frac{1}{3}$, the upper bound falls to 4.2 percent per year. However, this value still encompasses the estimate of efficient long-run inflation implied by the FIP model.

Another way to assess the upper bound is to estimate efficient long-run inflation using average inflation in the data. Fuhrer (2009), Table 2, reports averages for various measures of annual US inflation that all remain well below the upper bound of 9 percent for data from 1959 to 2008 and for various subsamples. Moreover, the upper bound encompasses inflation rates, averaged by decade, in major industrialized economies from the 1960s to present when exempting the high inflation episodes in the 1970s (Bakhshi, Khan, Burriel-Llombart, and Rudolf (2007), Table 2).

6 Macro Dynamics and Monetary Stabilization Policy

The second set of results concerns the dynamics of aggregate variables and the central bank's optimal reaction to aggregate shocks. In order to analyze these aspects, I consider the model with firm-specific productivity (FIP) and the model with homogenous productivity (HOP) when each model is linearized at its efficient flexible-price steady state.

6.1 Macro Dynamics

A local equilibrium of the FIP model calculated to the first order is determined by

$$\begin{aligned} \hat{\pi}_t &= \beta E_t \hat{\pi}_{t+1} + \zeta(\pi) x_t \\ x_t &= E_t x_{t+1} - s_c \sigma (\hat{i}_t - E_t \hat{\pi}_{t+1} - \hat{r}_t^f) \\ \left[\nu + (\sigma s_c)^{-1}\right] \hat{y}_t^f &= (\sigma s_c)^{-1} s_q \hat{q}_t + (1+\nu)\hat{a}_t \\ s_c \sigma \hat{r}_t^f &= -E_t(1 - L^{-1})[\hat{y}_t^f - s_q \hat{q}_t] \end{aligned} \tag{16}$$

[16]The evidence from models with firm turnover favors intermediate rather than high values of θ. Lewis and Poilly (2010) estimate θ equal to 3.31 in a monetary model with endogenous entry, a value similar to the one estimated in Bernard, Eaton, Jensen, and Kortum (2003). Furthermore, entry barriers exist in the real world even though these barriers do not matter in the FIP model with exogenous firm turnover. Entry barriers are likely to increase the markups of incumbent firms and thus support intermediate values of θ. Moreover, assuming decreasing returns to scale in firms' technology would push the upper bound outwards.

plus a specification of monetary policy. Variable x denotes the output gap $\hat{y} - \hat{y}^f$, and $\hat{y}^f$ and $\hat{r}^f$ denote the natural levels of output and the natural real rate, respectively. A hat on top of a variables denotes percentage deviation from steady state, and L denotes the lag operator. The first equation is the New Keynesian Phillips Curve (NKPC). The slope of the NKPC $\zeta(\pi)$ depends on π and the structural parameters,

$$\zeta(\pi) = \frac{(1 - \kappa\pi^{\theta-1})(1 - \kappa\pi^{\theta-1}\beta)}{\kappa\pi^{\theta-1}}[\nu + (\sigma s_c)^{-1}] \ . \tag{17}$$

I suppose here (and reconsider below) that the slope of the NKPC ζ is parameterized directly and therefore remains the same across models. Thus, I obtain the following proposition.

Proposition 2: Macro dynamics in the linearized FIP model and in the linearized HOP model are observationally equivalent if the slope of the approximate New Keynesian Phillips Curve is parameterized directly.

Recall that the FIP model nests the HOP model as the special case g, and thus π, which is equal to unity. Accordingly, it is easy to verify from equations (16) and (17) that both linearized models are observationally equivalent if the slope of the NKPC ζ is parameterized directly because all other parameters remain independent of either g or π.

Observational equivalence of linearized models emerges despite the fact that the FIP model is consistent with the positive inflation objectives pursued by many central banks around the world, whereas the HOP model is not consistent with these objectives. It follows from the equivalence result that all the first and second moments of macro variables coincide across models. For instance, both models imply the same impulse-response functions, and both models decompose macro time-series into exactly the same sequences of aggregate shocks. Accordingly, both models suggest the same positive interpretation of historic episodes in light of these shocks.

Equivalence further implies that conditions for the equilibrium to be unique coincide across models. One such condition for the case when monetary policy follows the rule $\hat{\imath}_t = f_\pi \hat{\pi}_t + f_x x_t$ is the Taylor principle. Roughly speaking, the Taylor principle states that f_π should exceed unity. More precisely, as Bullard and Mitra (2002) show in their proposition 1, the condition $\zeta(f_\pi - 1) + (1 - \beta)f_x > 0$ ensures uniqueness in the model, as summarized by equations (16). Accordingly, the Taylor principle ensures determinacy in both models.

The equivalence result is in contrast with previous findings in the literature. First, Ascari

(2004) shows that macro dynamics of the basic New Keynesian model with Calvo pricing change dramatically if the model is linearized at positive rather than zero inflation.[17] He concludes that it is not appropriate to compare data simulated from the model linearized at zero inflation with actual data characterized by positive average inflation. According to my analysis, such a comparison is appropriate as long as it can be convincingly argued that average inflation approximates efficient long-run inflation in the FIP model reasonably well. Second, Kiley (2007) and Ascari and Ropele (2009) show that determinacy conditions in the basic New Keynesian model with Taylor pricing (Kiley) or Calvo pricing (Ascari and Ropele) are sensitive to long-run inflation. Most prominently, the Taylor principle no longer guarantees determinacy at moderately positive long-run inflation. My analysis does not support this claim.

A main assumption underlying my proposition 2 is that the slope of the NKPC ζ remains the same in the FIP and the HOP model. This assumption is reasonable because both models deliver exactly the same slope estimate if the slope is estimated from either model by means of macro data, as it is done frequently in the literature. A similar situation arises in Altig, Christiano, Eichenbaum, and Linde (undated) who study a model with firm-specific capital and a nested model with homogenous capital and exploit observational equivalence across models to obtain a better match between macro and micro evidence on the duration of price spells.[18] One alternative to assuming that ζ is the same across models is to compute ζ from calibrated structural parameters. In this case, the two models do make different predictions because π differs across models and ζ depends on π. In Section 7, I trace out these differences, which turn out to be small.

6.2 Monetary Stabilization Policy

In order to assess the central bank's optimal reaction to shocks, I derive the welfare-based loss function of the central bank as a second-order expansion to the utility of the representative household in the FIP and the HOP model.

Proposition 3: The welfare-based loss function $\mathcal{L}$ calculated to the second order is proportional

[17] A similar finding emerges in Hornstein and Wolman (2005) for a model with firm-specific capital and price setting according to Taylor (1980). Hornstein and Wolman also find a large impact of long-run inflation on the determinacy region of monetary-policy rules.

[18] Along similar lines, Appendix C derives micro implications of macro estimates of the slope of the Phillips curve. There, I show that the FIP model mitigates the conflict that exists in the HOP model between micro data on price durations and estimates of ζ obtained from macro data.

to

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \hat{\pi}_t^2 + [\zeta(\pi)/\theta] x_t^2 \right\} \ . \tag{18}$$

It follows straight from the proposition that the welfare-based loss function is the same in the FIP and the HOP model if the slope of the NKPC ζ is parameterized directly and therefore remains the same across models. The reason for this is that long-run inflation π matters for the welfare loss (18) only through this slope, which enters the relative weight attached to stabilizing the output gap. The weight is proportional to ζ, and the factor of proportionality θ^{-1} is independent of long-run inflation.

Equivalence of the approximate loss functions in the FIP and the HOP model implies that the central bank's optimal reaction to shocks is identical in the two approximate models. The optimal reaction to shocks derives from a linear-quadratic policy problem with objective (18), constraints (16), shock processes, and given initial conditions. Under the stated assumptions, all the components of this problem are the same in the approximate FIP and the approximate HOP model. Equivalence of optimal stabilization policies does not depend on assumptions about central-bank commitment and extends to the optimal coefficients of simple policy rules.

Ascari and Ropele (2007) analyze optimal stabilization policy in the basic New Keynesian model with Calvo pricing linearized at positive long-run inflation. They posit a loss function of the same form as equation (18) but treat the weight on stabilizing the output gap as a primitive parameter. Here, I can relate this weight to structural parameters. Furthermore, a main result in Ascari and Ropele (2007) is that optimal discretionary policy renders the equilibrium with positive long-run inflation indeterminate.[19] This result does not carry over to my analysis.

Coibion, Gorodnichenko, and Wieland (2010) derive the welfare-based loss function calculated to the second order at a steady state with positive long-run inflation. They perform these calculations in a New Keynesian model with optimal long-run inflation equal to zero absent the zero lower bound. They find that the weight attached to stabilizing inflation increases in

[19] Ascari and Ropele minimize period loss $\hat{\pi}_t^2 + \chi x_t^2$ with $\chi > 0$ under discretion subject to their NKPC, which contains a cost-push shock $u_t = \rho_u u_{t-1} + \epsilon_t$ with $\epsilon_t \sim (0, \sigma^2)$ and $\rho_u \in [0, 1)$. For positive long-run inflation, their NKPC also contains the expectation of an auxiliary variable. The law of motion of the auxiliary variable is entirely forward-looking for their benchmark calibration. No such auxiliary variable enters the NKPC derived here. Accordingly, minimizing objective (18) subject to $\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \zeta(\pi) x_t + u_t$ yields a determinate equilibrium as long as $\xi(\pi) > 0$. Solution details are in Woodford (2003)'s proposition 7.5, or in Clarida, Gali, and Gertler (1999), Section 3.

long-run inflation because the marginal disutility from price dispersion is increasing in long-run inflation. My weight attached to stabilizing inflation does not display this effect because positive long-run inflation equal to g in my model implies the efficient amount of price dispersion.

6.3 Optimal Long-Run Inflation Target

The welfare-based loss function implies that optimal inflation targets are distinct in the FIP and the HOP model even though the optimal reaction to shocks is identical in these models. This property of the loss function is similar to the fact that optimal long-run inflation is distinct in both models.

In order to show that optimal inflation targets differ across models, rearrange the definition $\hat{\pi}_t \equiv \log(\pi_t/\pi)$ as $\hat{\pi}_t = (\pi_t - \pi)/\pi$, which is calculated to the first order, and plug it into the period loss to obtain

$$((\pi_t - \pi)/\pi)^2 + [\zeta(\pi)/\theta]x_t^2 \ . \tag{19}$$

Inspecting this equation reveals that the central bank in the FIP model stabilizes actual inflation around positive long-run inflation $\pi > 1$, whereas the central bank in the HOP model stabilizes actual inflation around zero long-run inflation $\pi = 1$.

Ascari and Ropele (2007) assign a positive inflation target to the central bank in the basic New Keynesian model with homogenous productivity to analyze optimal stabilization policy. Accordingly, the central bank in their model pursues suboptimal long-run policy because the optimal inflation target is zero rather than positive in this model. They motivate this assumption with the observation that inflation targets in the real world are often positive. In the FIP model examined here, it is actually optimal for the central bank to pursue a positive inflation target.

Positive optimal inflation targets are rarely encountered in models of the New Neoclassical Synthesis that do not impose the zero lower bound on nominal interest rates. Woodford (2003), chapter 6, shows that zero target inflation remains a robust prediction, at least approximately, in many extensions of the basic model. Erceg and Levin (2006), Huang and Liu (2005), Benigno (2004), and Aoki (2001) analyze cyclical variation in relative prices across sectors and show that such variation justifies temporary, but not permanent, deviation from price stability.

7 Flattening the Slope of the Phillips Curve

The previous discussion assumes that the slope of the NKPC is parameterized directly. This assumption is plausible because a slope estimate obtained from macro data is the same regardless of whether the linearized FIP or HOP model is employed for estimation. One alternative to estimation is to calibrate structural parameters underlying the slope ζ using micro evidence. In this case, higher long-run inflation implies a flatter slope, and the two models do make different predictions. In this section, I trace out these differences. I start with the impact of long-run inflation on the slope of the NKPC. Then, I compare macro dynamics under a simple monetary-policy rule and assess welfare consequences of optimal stabilization policies.

7.1 How Long-Run Inflation Flattens the Phillips Curve

In order to link long-run inflation π and the slope of the NKPC ζ, I look more closely into the pricing equation (1) and into the inflation equation (4). Jointly, these equations imply the NKPC.

Linearizing the pricing equation (1) at the efficient steady state delivers

$$\hat{p}_t^\star = E_t \sum_{i=0}^{\infty} (\kappa\gamma\beta)^i \left[(1-\kappa\gamma\beta)\hat{s}_{t+i}^\star + \kappa\gamma\beta\hat{\pi}_{t+i+1}\right] .$$

The discount rate $\kappa\gamma\beta$ increases in long-run inflation because γ increases in long-run inflation. This captures the market-share effect in steady state. With positive long-run inflation, the firm's relative price falls over time and future market shares increase. The firm anticipates high future market shares and weights future price determinants more. The market-share effect also implies that real marginal costs $\hat{s}_{t+i}^\star$ are weighted less and expected inflation $\hat{\pi}_{t+i+1}$ is weighted more. Expected inflation is weighted more because it indicates more erosion of the firm's price when inflation is expected to be above trend.

Inflation equation (4) pins down the relative market share of a new firm in steady state,

$$(p^\star)^{1-\theta} = [1-(1-\delta)g^{\theta-1}]/\delta .$$

Increasing g reduces the market share of a new firm because increasing g boosts average productivity while the productivity of a new firm remains fixed. A small market share means that the optimal relative price of new firms $p_t^\star$ has less impact on inflation because the aggregate price level $P_t = \int_0^1 P_{jt}(c_{jt}/c_t)dj$ weighs prices by relative quantities. The linearized inflation equation captures this effect:

$$\hat{\pi}_t = \left(\frac{1-\kappa\gamma}{\kappa\gamma}\right)\hat{p}^{\star}_t \ .$$

The optimal price of new firms affects inflation less the higher is γ.

Combining the pricing equation and the inflation equation yields the NKPC (16) with its slope ζ determined by equation (17). The higher long-run inflation is the smaller this slope is and the flatter the Phillips curve is. This happens for two reasons. First, firms set their prices with greater emphasis on future price determinants because positive long-run inflation increases future market shares. Second, with positive long-run inflation, the current market share of new firms falls relative to the market share of the average firm.[20] Therefore, inflation responds less sensitively to the optimal price of new firms.

7.2 Macro Dynamics

How do changes in the slope of the NKPC that are triggered by long-run inflation affect macro dynamics? Figure 5 plots the dynamics of the output gap, inflation, and interest rates in the FIP model (crosses) and in the HOP model (circles) after a cost-push shock of one percent. Monetary policy is the same across models, $\hat{i}_t = f_\pi \hat{\pi}_t + f_x x_t$ with $f_\pi = 1.5$ and $f_x = 0.125$.

Qualitatively, shock transmission is identical in both models. The shock increases inflation, and monetary policy responds by raising the nominal interest rate more than inflation. Elevated real rates encourage the household to postpone consumption such that sticky-price output falls. The output gap falls, too, because flexible-price output remains unaffected by the shock. By the NKPC, the central bank partially stabilizes inflation through the drop in the output gap.

Quantitatively, inflation and the output gap vary slightly more in the FIP model than in the HOP model. The equilibrium beyond the dynamics in Figure 5 is (imposing $\sigma = \nu = 1$)

$$\begin{pmatrix} x_t \\ \hat{\pi}_t \end{pmatrix} = [\ \zeta(f_\pi - \rho_u) + (1-\rho_u\beta)(1-\rho_u+f_x)\ ]^{-1} \begin{pmatrix} -(f_\pi - \rho_u) \\ 1-\rho_u+f_x \end{pmatrix} u_t \ .$$

The slope of the NKPC ζ appears in the square brackets. For the moment, let $\rho_u = 0$. A flatter slope increases the (absolute) responses of both inflation and the output gap to cost-push shocks because the term in square brackets falls. It follows from the NKPC that a given movement in the output gap achieves less inflation stabilization when ζ is small. By the policy rule, the nominal

[20] Quantitatively, both channels are about equally important.

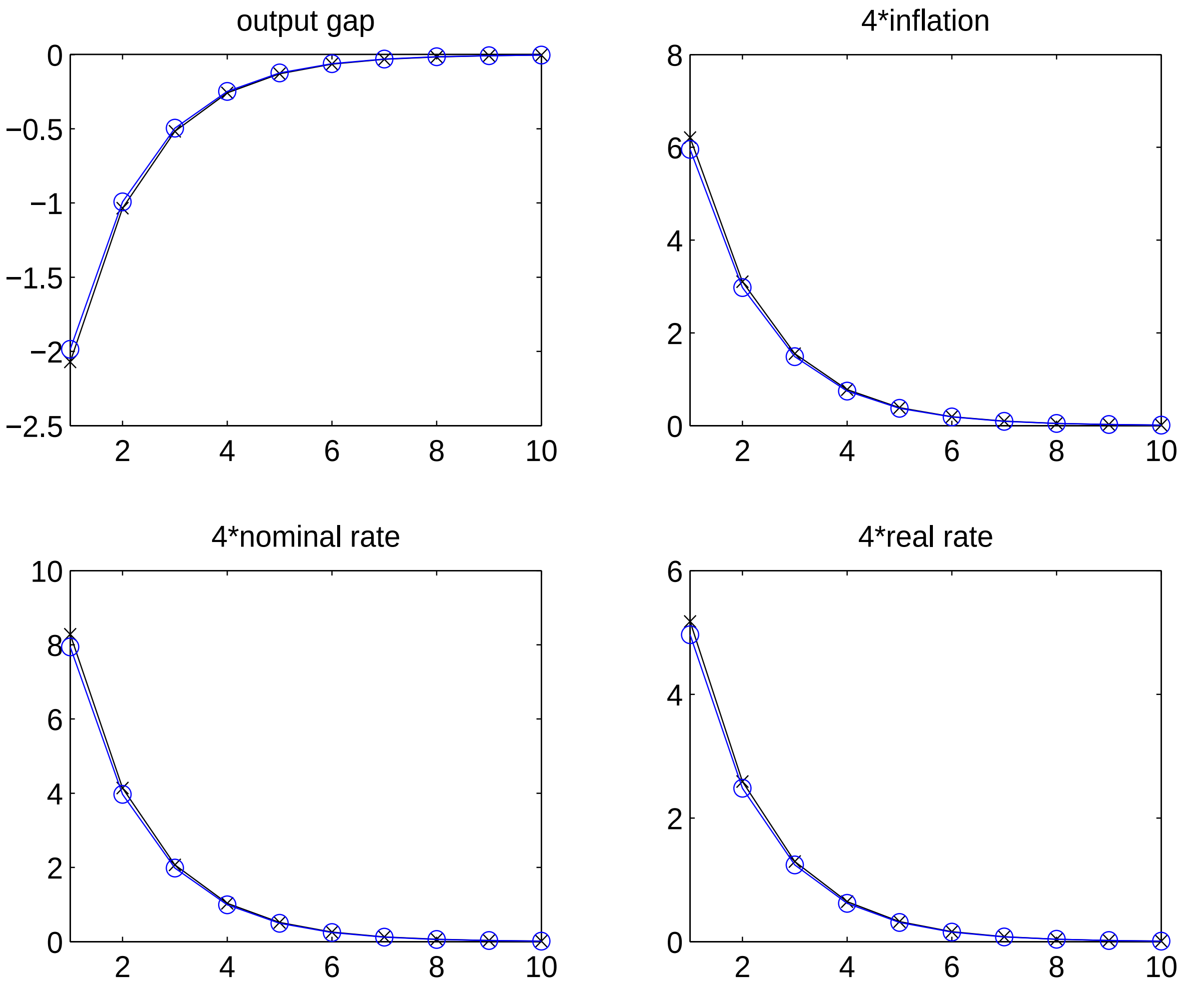

Figure 5: Impulse responses to the same positive u_t shock with $\rho_u = 0.5$ and unit size in FIP model (crosses) and HOP model (circles). Monetary policy follows the same rule $\hat{i}_t = f_\pi \hat{\pi}_t + f_x x_t$ in both models.

interest rate thus varies more. By the Euler equation, the output gap varies more. With $\rho_u > 0$, reducing ζ still decreases the term in square brackets even though expectations interfere with the direct effect. Overall, when shocks are autocorrelated, a flatter slope increases the equilibrium variation in inflation and the output gap even more.

A compact way to summarize the quantitative differences between both models is through the standard deviation σ_u of the residual of the cost-push shock. Dynamics coincide across models if σ_u in the FIP model is set to 95.88 percent of its value in the HOP model.[21] Changing the AR coefficient of the shock alters this factor; $\rho_u = 0.8$ implies a factor of 91.21 percent, and $\rho_u = 0$ implies a factor of 98.01 percent. Overall, quantitative differences are small.

[21] Proportionality of impulse-response functions (IRFs) fails when the model contains endogenous states or when there are shocks to the natural real rate. However, differences between IRFs remain moderate in these cases, too.

7.3 Monetary Stabilization Policy

I turn now to the effect of a flatter slope of the Phillips curve on optimal stabilization policies under central-bank discretion and commitment. Optimal stabilization of cost-push shocks in either model derives from minimizing the loss function (18) subject to the NKPC $\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \zeta(\pi) x_t + u_t$, conditional on the shock process and initial conditions. A discretionary central bank cannot credibly commit to future action such that the optimal targeting rule under discretion is contemporaneous,

$$\hat{\pi}_t + \frac{\lambda}{\zeta(\pi)} x_t = 0 \ .$$

Parameter $\lambda > 0$ denotes the relative weight on stabilizing the output gap rather than inflation in the loss function. The central bank affects private-sector expectations in case it credibly commits to future action. Accordingly, the optimal targeting rule under commitment is history-dependent,

$$\hat{\pi}_t + \frac{\lambda}{\zeta(\pi)} (x_t - x_{t-1}) = 0 \ .$$

In general, optimal targeting rules differ across the FIP and the HOP model because $\zeta(\pi)$ differs across models.

There is one important exception for λ equal to the welfare-based weight $\zeta(\pi)/\theta$ that is arguably the most relevant case. In this case, optimal targeting rules are

$$\hat{\pi}_t + \theta^{-1} x_t = 0 \quad , \qquad \hat{\pi}_t + \theta^{-1}(x_t - x_{t-1}) = 0 \ .$$

Differences across models that arise under these policies originate exclusively from differences in the slope of the NKPC and therefore from shock transmission.

Figure 6 plots responses in the FIP model (crosses) and the HOP model (circles) for optimal policy under commitment to a cost-push shock with $\rho_u = 0$. The weight on stabilizing the output gap is the welfare-based weight. In both models, deflation is optimal in the periods after the shock. Deflation prevents large increases in optimal prices $P^{\star}_{t-n,t}$ because firms anticipate the deflation. Through product demand, large increases in the optimal price would induce the output gap to fall by more than what the figure shows.

Quantitatively, stabilization is more successful in the HOP model than in the FIP model

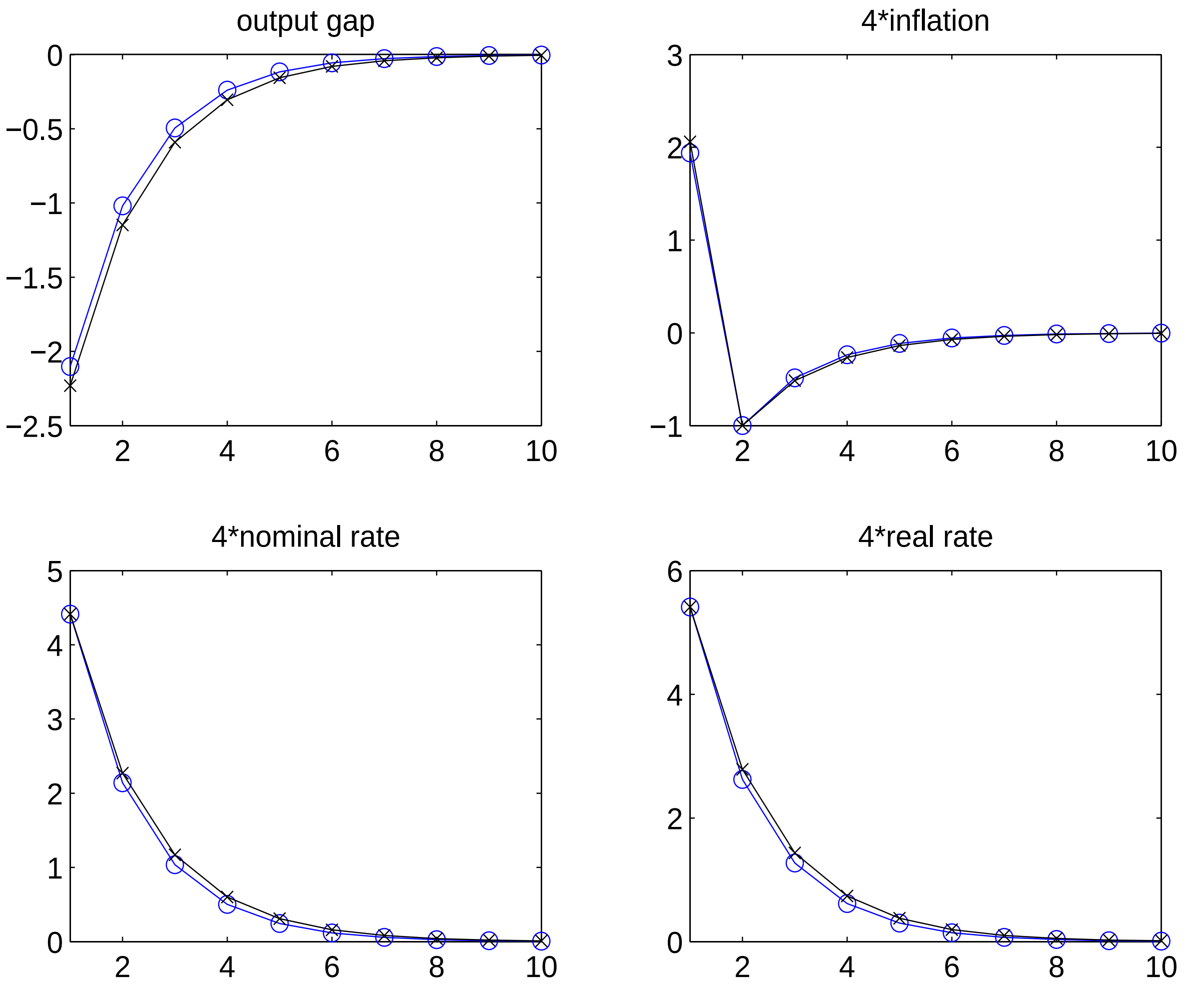

Figure 6: Impulse responses in the FIP model (crosses) and the HOP model (circles) for optimal policy under commitment to a cost-push shock u_t with $\rho_u = 0$.

because the slope of the NKPC is steeper in the HOP model. The mechanics are the same as those underlying Figure 5. Inspecting Figure 6 more carefully reveals that responses die out less quickly in the FIP model. To illustrate the reason, I iterate the policy rule backward and assume that inflation and output gap are in steady state before the shock to obtain

$$-\theta^{-1} x_t = \hat{\pi}_t + \hat{\pi}_{t-1} + \cdots + \hat{\pi}_1 \ , \quad t \geq 2 \ .$$

In equilibrium, current and past inflation rates are related to the output gap simply because policy is history-dependent. History dependence projects accumulated past responses into current responses and, thereby, transforms less stabilization into more persistent dynamics. Figure 6 reveals that the magnitude of this effect is fairly small. Ascari and Ropele (2007) find large effects of positive long-run inflation on the persistence of responses as proxied by the magnitude

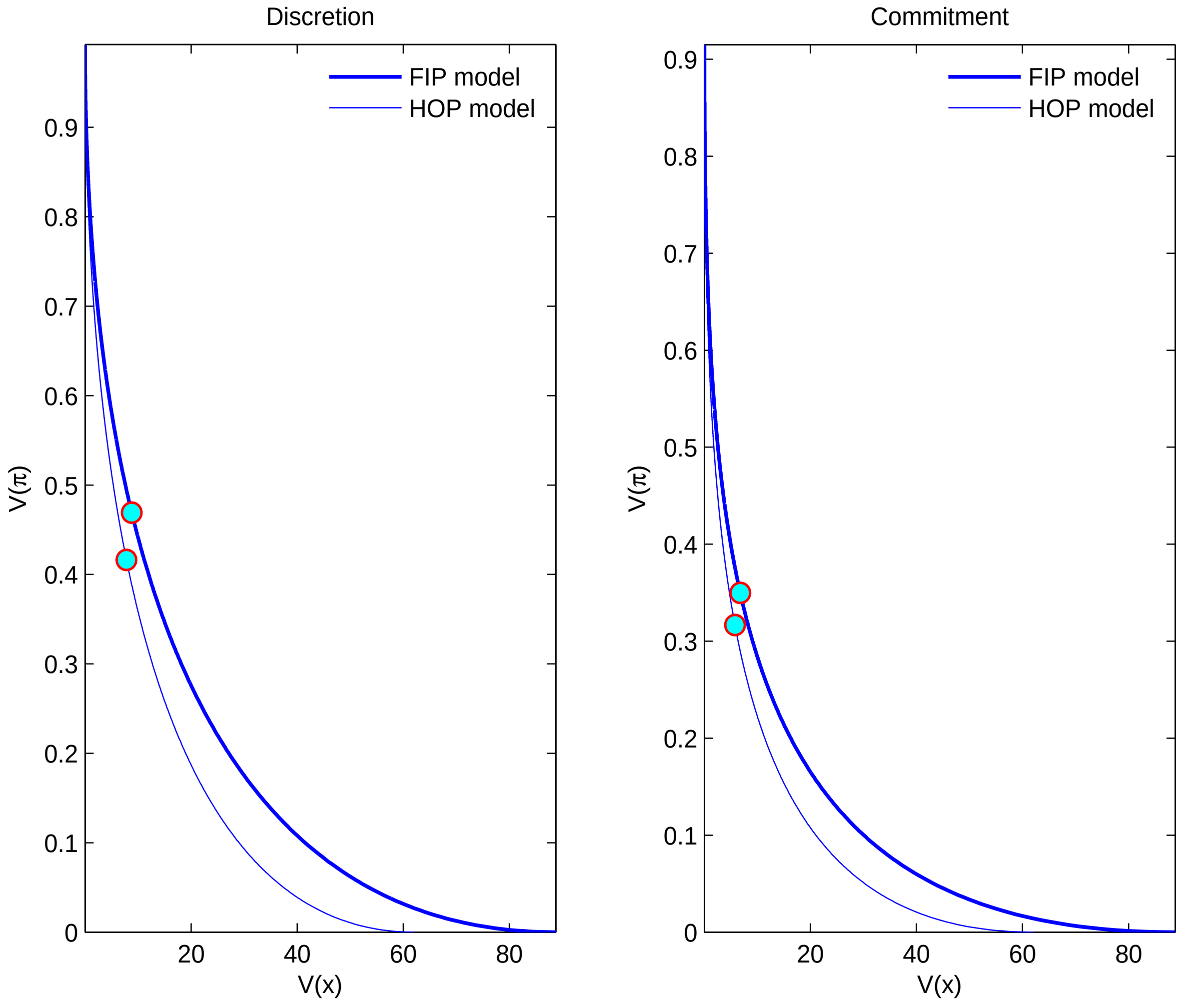

Figure 7: Efficiency frontiers in FIP model (thick lines) and HOP model (thin lines) under discretion and commitment for u_t shock with $\rho_u = 0$ and $\sigma_u = 1$. $V(\hat{\pi})$ and $V(x)$ denote unconditional variance in inflation and output gap. A circle indicates the combination of unconditional variances in inflation and output gap under the welfare-based weight on the output gap.

of stable eigenvalues. Presumably, one reason for this is that the functional form of the optimal policy rule under commitment in their model changes when long-run inflation is positive rather than zero.

The finding that stabilization is more successful in the HOP model than in the FIP model is very general. Figure 7 plots efficiency frontiers for optimal policy under commitment and discretion in the FIP model (thick lines) and the HOP model (thin lines). A circle indicates the variances in inflation and the output gap associated with the optimal policy under the welfare-based loss with $\lambda = \zeta(\pi)/\theta$. Varying λ traces out the frontier. Large values of λ imply a small variance in the output gap and a large variance in inflation because the central bank cares most about stabilizing the output gap.

Stabilization in the HOP model is more successful than in the FIP model because the efficiency frontier of the HOP model is closer to the origin for all values of λ. The frontiers in the FIP and the HOP model diverge for small values of λ. The reason for this is that, whenever the central bank aims to stabilize inflation, it can do so only by moving the output gap. Accordingly, differences in the slope of the NKPC matter most when the central bank puts a large weight on stable inflation. Frontiers converge to each other for large values of λ because the central bank does not even attempt to stabilize inflation anymore. Accordingly, different slopes of the NKPC no longer matter.

To quantify the losses associated with the flatter NKPC in the FIP model, I compute the percentage difference between losses as $(\mathcal{L}_{FIP} - \mathcal{L}_{HOP})/\mathcal{L}_{HOP}$ for λ equal to the welfare-based weight. For discretionary policy, the loss in the FIP model is 6.159 percent larger than the loss in the HOP model. For committed policy, the same quantity is equal to 6.157 percent. Relative welfare losses associated with a flatter NKPC are nonnegligible. However, there is not much a monetary policymaker can do to reduce losses in the FIP model because policies are already optimal.

8 Conclusion

In micro data, it appears that firms move systematically within the productivity distribution. In particular, firms tend to have below-average productivity upon entry, and firms tend to experience post-entry growth in productivity. The purpose of this paper is to highlight the consequences of firm-level productivity dynamics for optimal long-run inflation and for monetary stabilization policy. I work with a New Keynesian model and add two assumptions to capture firms' productivity dynamics. The first assumption is that a firm's productivity grows mechanically with the age of the firm. The second assumption is that firm turnover is positive. Jointly, these assumptions make firms heterogenous in terms of productivity. The model nests the basic New Keynesian model with homogenous productivity as a special case.

The model delivers two main results. First, optimal long-run inflation is positive and equal to growth in firm-specific productivity. This result differs markedly from the price-stability result typically obtained with models of the New Neoclassical Synthesis. Positive long-run inflation is optimal because, by growth in firm-specific productivity, the desired real price of a firm declines over time. Thus, positive optimal inflation erodes sticky nominal prices at exactly the right pace to reduce elevated markups and excess dispersion in relative prices. Apparently, many central banks around the globe pursue moderately positive rather than zero long-run inflation. The model used here suggests this is (close to) optimal.

The second result is that, when linearized at positive long-run inflation, the model with firm-specific productivity is observationally equivalent to the nested New Keynesian model with homogenous productivity linearized at zero long-run inflation. This result has several consequences that differ from results reported in the literature. For one, the Taylor principle still guarantees determinacy in the model with positive long-run inflation. Moreover, the dynamics of real-world inflation, which is positive on average, are frequently interpreted through the lenses of the basic New Keynesian model linearized at zero long-run inflation, and this interpretation has been criticized. However, the equivalence result obtained here supports this interpretation.

A tightly related result is that both linearized models prescribe the same optimal stabilization policies because the welfare-based loss function of the central bank, derived as a second-order expansion to the utility of the representative household, is also the same in both models. Accordingly, my model suggests that optimal stabilization policies, which hold in the basic New Keynesian model near zero long-run inflation, are robust to moderately positive long-run inflation. The related literature argues for the contrary.

Overall, these results suggest that firms' productivity dynamics observed in micro data matter substantially for conclusions about monetary policy. In the wake of the recent financial turmoil, Williams (2009), Blanchard, Dell'Ariccia, and Mauro (2010), and McCallum (2010), among others, debate consequences of raising inflation targets to above their current levels in order to provide central banks with more leeway to cope with large adverse shocks. My results contribute to this debate that firm-level productivity dynamics call for moderately positive long-run inflation even without taking into account the zero lower bound on nominal interest rates.

It is interesting to extend the current analysis in future work. First, while this paper finds positive optimal long-run inflation in a cashless economy, the literature emphasizes that the opportunity costs arising from holding money imply negative optimal long-run inflation. Future work could assess the relative importance of firm-level productivity dynamics versus the opportunity costs from holding money for optimal long-run inflation. Second, in line with the evidence, my model emphasizes firm-level productivity differentials and firm-level productivity growth. Recently, Foster, Haltiwanger, and Syverson (2008) have suggested that demand factors are another important distinction across firms. Future work thus might examine models that also account for demand factors. Finally, it would be interesting to extend the current analysis to models with endogenous firm growth and endogenous firm turnover, and this might require resorting to numerical methods. These extensions would likely make estimates of optimal long-run inflation more reliable quantitatively.

References

Altig, D., L. Christiano, M. Eichenbaum, and J. Linde (undated): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," *Review of Economic Dynamics*.

Amano, R., S. Ambler, and N. Rebei (2007): "The Macroeconomic Effects of Nonzero Trend Inflation," *Journal of Money, Credit and Banking*, 39(7), 1821–1838.

Amano, R., K. Moran, S. Murchison, and A. Rennison (2009): "Trend inflation, wage and price rigidities, and productivity growth," *Journal of Monetary Economics*, 56(3), 353–364.

Aoki, K. (2001): "Optimal monetary policy responses to relative-price changes," *Journal of Monetary Economics*, 48(1), 55–80.

Aruoba, S. B., and F. Schorfheide (2009): "Sticky Prices Versus Monetary Frictions: An Estimation of Policy Trade-offs," NBER Working Papers 14870, National Bureau of Economic Research, Inc.

Ascari, G. (2004): "Staggered Prices and Trend Inflation: Some Nuisances," *Review of Economic Dynamics*, 7(3), 642–667.

Ascari, G., and T. Ropele (2007): "Optimal monetary policy under low trend inflation," *Journal of Monetary Economics*, 54(8), 2568–2583.

——— (2009): "Trend Inflation, Taylor Principle, and Indeterminacy," *Journal of Money, Credit and Banking*, 41(8), 1557–1584.

Aw, B. Y., X. Chen, and M. J. Roberts (2001): "Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing," *Journal of Development Economics*, 66(1), 51–86.

Baily, M. N., C. Hulten, D. Campbell, T. Bresnahan, and R. E. Caves (1992): "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity. Microeconomics*, 1992, pp. 187–267.

Bakhshi, H., H. Khan, P. Burriel-Llombart, and B. Rudolf (2007): "The New Keynesian Phillips curve under trend inflation and strategic complementarity," *Journal of Macroeconomics*, 29(1), 37–59.

Baldwin, J. R. (1996): "Productivity Growth, Plant Turnover and Restructuring in the Canadian Manufacturing Sector," Analytical Studies Branch Research Paper Series 1995087e, Statistics Canada, Analytical Studies Branch.

Baldwin, J. R., and W. Gu (2006): "Plant turnover and productivity growth in Canadian manufacturing," *Industrial and Corporate Change*, 15(3), 417–465.

Bartelsman, E. J., and M. Doms (2000): "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, 38(3), 569–594.

Benigno, P. (2004): "Optimal monetary policy in a currency area," *Journal of International Economics*, 63(2), 293–320.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–1290.

Bernard, A. B., S. J. Redding, and P. K. Schott (2010): "Multiple-Product Firms and Product Switching," *American Economic Review*, 100(1), 70–97.

Bilbiie, F., F. Ghironi, and M. Melitz (2008): "Monetary Policy and Business Cycles with Endogenous Entry and Product Variety," in *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, pp. 299–353. National Bureau of Economic Research, Inc.

Bilbiie, F., F. Ghironi, and M. J. Melitz (2010): "Endogenous Entry, Product Variety, and Business Cycles," Discussion paper.

Billi, R. M. (forthcoming): "Optimal Inflation for the U.S. Economy," *American Economic Journal: Macroeconomics*.

Blanchard, O., G. Dell'Ariccia, and P. Mauro (2010): "Rethinking Macroeconomic Policy," *Journal of Money, Credit and Banking*, 42(s1), 199–215.

Broda, C., and D. E. Weinstein (2010): "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 100(3), 691–723.

Bullard, J., and K. Mitra (2002): "Learning about monetary policy rules," *Journal of Monetary Economics*, 49(6), 1105–1129.

Caves, R. E. (1998): "Industrial Organization and New Findings on the Turnover and Mobility of Firms," *Journal of Economic Literature*, 36(4), 1947–1982.

Cecioni, M. (2010): "Firm entry, competitive pressures and the US inflation dynamics," Temi di discussione (Economic working papers) 773, Bank of Italy, Economic Research Department.

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

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010): "Inflation-Gap Persistence in the US," *American Economic Journal: Macroeconomics*, 2(1), 43–69.

Cogley, T., and A. M. Sbordone (2008): "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review*, 98(5), 2101–26.

Coibion, O., Y. Gorodnichenko, and J. F. Wieland (2010): "The Optimal Inflation Rate in New Keynesian Models," NBER Working Papers 16093, National Bureau of Economic Research, Inc.

Del Gatto, M., G. Mion, and G. I. P. Ottaviano (2006): "Trade Integration, Firm Selection and the Costs of Non-Europe," CEPR Discussion Papers 5730, C.E.P.R. Discussion Papers.

Dotsey, M., R. G. King, and A. L. Wolman (1999): "State-Dependent Pricing And The General Equilibrium Dynamics Of Money And Output," *The Quarterly Journal of Economics*, 114(2), 655–690.

Erceg, C., and A. Levin (2006): "Optimal monetary policy with durable consumption goods," *Journal of Monetary Economics*, 53(7), 1341–1359.

Farias, J. C., and S. Ruano (2005): "Firm productivity, heterogeneity, sunk costs and market selection," *International Journal of Industrial Organization*, 23(7-8), 505–534.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001): "Aggregate Productivity Growth. Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, NBER Chapters, pp. 303–372. National Bureau of Economic Research, Inc.

Fuhrer, J. C. (2009): "Inflation persistence," Working Papers 09-14, Federal Reserve Bank of Boston.

Gali, J. (2008): *Monetary Policy, Inflation, and the Business Cycle*. Princeton University Press.

Geroski, P. A. (1995): "What do we know about entry?," *International Journal of Industrial Organization*, 13(4), 421–440.

Ghironi, F., and M. J. Melitz (2005): "International Trade and Macroeconomic Dynamics with Heterogeneous Firms," *The Quarterly Journal of Economics*, 120(3), 865–915.

Goodfriend, M., and R. G. King (1998): "The new neoclassical synthesis and the role of monetary policy," Working Paper 98-05, Federal Reserve Bank of Richmond.

——— (2001): "The case for price stability," Working Paper 01-02, Federal Reserve Bank of Richmond.

Graham, L., and D. J. Snower (2004): "The real effects of money growth in dynamic general equilibrium," Working Paper Series 412, European Central Bank.

Hornstein, A., and A. L. Wolman (2005): "Trend inflation, firm-specific capital, and sticky prices," *Economic Quarterly*, (Fall), 57–83.

Huang, K. X., and Z. Liu (2005): “Inflation targeting: What inflation rate to target?,” *Journal of Monetary Economics*, 52(8), 1435–1462.

Huergo, E., and J. Jaumandreu (2004): “Firms’ age, process innovation and productivity growth,” *International Journal of Industrial Organization*, 22(4), 541–559.

Khan, A., R. G. King, and A. L. Wolman (2003): “Optimal Monetary Policy,” *Review of Economic Studies*, 70, 825–860.

Kiley, M. T. (2002): “Partial Adjustment and Staggered Price Setting,” *Journal of Money, Credit and Banking*, 34(2), 283–98.

Kiley, M. T. (2007): “Is Moderate-to-High Inflation Inherently Unstable?,” *International Journal of Central Banking*, 3(2), 173–201.

Kim, J., and F. J. Ruge-Murcia (2009): “How much inflation is necessary to grease the wheels?,” *Journal of Monetary Economics*, 56(3), 365–377.

King, R., and A. L. Wolman (1999): “What Should the Monetary Authority Do When Prices Are Sticky?,” in *Monetary Policy Rules*, NBER Chapters, pp. 349–404. National Bureau of Economic Research, Inc.

King, R. G., and A. L. Wolman (1996): “Inflation Targeting in a St. Louis Model of the 21st Century,” NBER Working Papers 5507, National Bureau of Economic Research, Inc.

Klenow, P. J., and B. A. Malin (2010): “Microeconomic Evidence on Price-Setting,” NBER Working Papers 15826, National Bureau of Economic Research, Inc.

Kobayashi, T., and I. Muto (2010): “A note on expectational stability under non-zero trend inflation,” MPRA Paper 22952, University Library of Munich, Germany.

Kuttner, K. N. (2004): “A Snapshot of Inflation Targeting in its Adolescence,” in *The Future of Inflation Targeting*, ed. by C. Kent, and S. Guttmann, RBA Annual Conference Volume. Reserve Bank of Australia.

Levin, A., and T. Yun (2007): “Reconsidering the natural rate hypothesis in a New Keynesian framework,” *Journal of Monetary Economics*, 54, 1344–1365.

Lewis, V., and C. Poilly (2010): “Firm Entry and the Monetary Transmission Mechanism,” *manuscript*.

Linde, J. (2005): “Estimating New-Keynesian Phillips curves: A full information maximum likelihood approach,” *Journal of Monetary Economics*, 52(6), 1135–1149.

McCallum, B. T. (2010): “Should Central Banks Raise their Inflation Targets,” .

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Nakamura, E., and J. Steinsson (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 123(4), 1415–1464.

Schmitt-Grohe, S., and M. Uribe (2010): "The Optimal Rate of Inflation," NBER Working Papers 16054, National Bureau of Economic Research, Inc.

Taylor, J. B. (1980): "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 88(1), 1–23.

Wagner, J. (2010a): "Entry, Exit and Productivity: Empirical Results for German Manufacturing Industries," *German Economic Review*, 11, 78–85.

——— (2010b): "The Post-Entry Performance of Cohorts of Export Starters in German Manufacturing Industries," IZA Discussion Papers 5183, Institute for the Study of Labor (IZA).

Weber, H. (unpublished): "Firm entry, firm heterogeneity and inflation dynamics," Discussion paper.

Williams, J. C. (2009): "Heeding Daedalus: Optimal inflation and the zero lower bound," Working Paper Series 2009-23, Federal Reserve Bank of San Francisco.

Wolman, A. L. (2009): "The optimal rate of inflation with trending relative prices," Working Paper 09-02, Federal Reserve Bank of Richmond.

Woodford, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Yun, T. (1996): "Nominal price rigidity, money supply endogeneity, and business cycles," *Journal of Monetary Economics*, 37(2), 345 – 370.

——— (2005): "Optimal Monetary Policy with Relative Price Distortions," *American Economic Review*, 95(1), 89–109.

A Competitive Equilibrium

The following equations summarize the competitive equilibrium with sticky prices. The flexible-price model is the special case $\alpha = 0$. The parameters are defined in the text. Real profits are denoted as $d_t = D_t/P_t$.

$$
\begin{aligned}
1 &= \beta E_t \frac{u_c(c_{t+1})}{u_c(c_t)} \; \frac{1+i_t}{\pi_{t+1}} \\
w_t &= (1-\tau_L)^{-1} \; \frac{h_\ell(\ell_t)}{u_c(c_t)} \\
1 &= \frac{\delta(1-\kappa\gamma)}{1-(1-\delta)\gamma}(p_t^\star)^{1-\theta} + \kappa\pi_t^{\theta-1} \\
\Delta_t &= \frac{\delta(1-\kappa\gamma)}{1-(1-\delta)\gamma}(p_t^\star)^{-\theta} + (\kappa/g)\pi_t^{\theta}\Delta_{t-1} \\
p_t^\star &= \frac{\theta}{\theta-1}\frac{\gamma_{2t}}{\gamma_{1t}} \\
\gamma_{1t} &= 1 + \kappa\beta E_t \frac{u_c(c_{t+1})}{u_c(c_t)}\,(y_{t+1}/y_t)\,\pi_{t+1}^{\theta-1}\gamma_{1t+1} \\
\gamma_{2t} &= s_t^\star + (\kappa\beta/g) E_t \frac{u_c(c_{t+1})}{u_c(c_t)}\,(y_{t+1}/y_t)\,\pi_{t+1}^{\theta}\gamma_{2t+1} \\
s_t^\star &= w_t/a_t \\
y_t &= a_t\ell_t/\Delta_t \\
d_t &= y_t\,(1 - s_t^\star\Delta_t) \\
y_t &= c_t + q_t \; .
\end{aligned}
$$

B Steady State

Conditional on π, the steady-state variables $w, c, \ell, p^\star, \gamma_1, \gamma_2$, and Δ are determined by

$$
\begin{aligned}
w &= (1-\tau_L)^{-1}\,\frac{h_\ell(\ell)}{u_c(c)} \\
p^\star &= \left(\frac{\delta}{1-(1-\delta)\gamma}\right)^{\frac{1}{\theta-1}} \left(\frac{1-\kappa\pi^{\theta-1}}{1-\kappa\gamma}\right)^{\frac{1}{1-\theta}} \\
\Delta &= \frac{\delta(1-\kappa\gamma)}{1-(1-\delta)\gamma}(p^\star)^{-\theta} + (\kappa/g)\pi^\theta\Delta \\
p^\star &= \frac{\theta}{\theta-1}\left(\frac{1-\kappa\beta\pi^{\theta-1}}{1-(\kappa\beta/g)\pi^\theta}\right)\frac{w}{a} \\
\gamma_1 &= \frac{1}{1-\kappa\beta\pi^{\theta-1}} \\
\gamma_2 &= \frac{w/a}{1-(\kappa\beta/g)\pi^\theta} \\
c+q &= a\ell/\Delta\ .
\end{aligned}
$$

Conditions for existence of the steady state are

$$
\pi < \left(\frac{1}{\kappa}\right)^{\frac{1}{\theta-1}} \quad , \quad \pi < \left(\frac{g}{\kappa}\right)^{\frac{1}{\theta}} \quad , \quad g < \left(\frac{1}{1-\delta}\right)^{\frac{1}{\theta-1}} \ .
$$

The two restrictions on inflation ensure that γ_1, γ_2, Δ, and π itself are well defined in steady state. The restriction on g ensures $(1-\delta)\gamma < 1$ such that $\delta/(1-(1-\delta)\gamma)$ and thereby $p^\star$ remains finite. In efficient steady state with $\pi = g$, the second condition collapses to the first condition. Moreover, the first condition holds whenever the third condition holds because $\kappa = \alpha(1-\delta)$ and $\alpha \leq 1$. Accordingly, with $\pi = g$, existence of the steady state requires only that

$$
\pi < \left(\frac{1}{1-\delta}\right)^{\frac{1}{\theta-1}} \ .
$$

C Micro Implications of Macro Estimates of the Phillips Curve

The slope of the NKPC is frequently estimated from macro data.[22] Using a structural model, such an estimate translates into an estimate of the mean duration of a price spell. A question that is debated much in the literature is how this macro estimate of the mean duration of a price spell compares to micro evidence on the same quantity. Generally, macro data imply a low estimate that implies a long mean duration of price spells. A long mean duration of price spells is often considered to conflict with micro evidence. Micro evidence indicates a short mean duration of price spells (Klenow and Malin (2010), Nakamura and Steinsson (2008)).

To connect micro and macro evidence, denote the estimate of the slope of the NKPC obtained from macro data as $\hat{\zeta}$. The mean duration of a price spell, which is derived from the censored distribution of price spells, is equal to $1/(1-\kappa)$.[23] I relate both by solving the slope

$$\zeta = \tfrac{(1-\kappa\gamma)(1-\kappa\gamma\beta)}{\kappa\gamma}[\nu + (\sigma s_c)^{-1}]$$

for $\kappa\gamma$ conditional on $\hat{\zeta}$. Denote the solution as $\widehat{\kappa\gamma}$ with $\widehat{\kappa\gamma}_{1/2} = -a_0/2 \pm \sqrt{a_0^2/4 - 1/\beta}$ and $a_0 = (1 + \beta + \hat{\zeta}/[\nu + (\sigma s_c)^{-1}])/\beta$. Of the two solutions $\widehat{\kappa\gamma}_{1/2}$, I chose the one within the unit interval. The value of κ consistent with a particular estimate $\hat{\zeta}$ is then $\hat{\kappa} = \widehat{\kappa\gamma}/\pi^{\theta-1}$. The higher π is the lower the value of $\hat{\kappa}$ that is required to match $\hat{\zeta}$. The macro estimate of the mean duration of a price spell consistent with $\hat{\zeta}$ is then $1/(1-\hat{\kappa})$.

Figure 8 shows that values of g above unity reduce the mean duration of price spells implied by a particular slope estimate $\hat{\zeta}$. Here, $\hat{\zeta}$ is computed based on structural parameters underlying the FIP model. I discuss in the main text that the slope ζ (not its estimate) falls when $g = \pi$ increases. To offset the tendency of ζ to fall with π, a lower probability κ for the average price to remain constant is required. As a result, the mean duration of prices falls. With annualized g equal to 1.03, the mean duration of price spells implied by macro estimates is below 4 quarters in the FIP model rather than right above 4.3 quarters as in the HOP model. Accordingly, the FIP model mitigates the conflict between micro and macro evidence on price stickiness.

[22] Linde (2005) estimates the NKPC in a small-scale model similar to the one considered here. More recently, Cogley and Sbordone (2008) and Cogley, Primiceri, and Sargent (2010) have derived respective estimates.

[23] I consider the *censored* distribution of prices to isolate the effect of g from the effect of δ. It is a matter of calibration whether one takes into account forced item substitutions or not in order to infer the mean duration of a price spell.

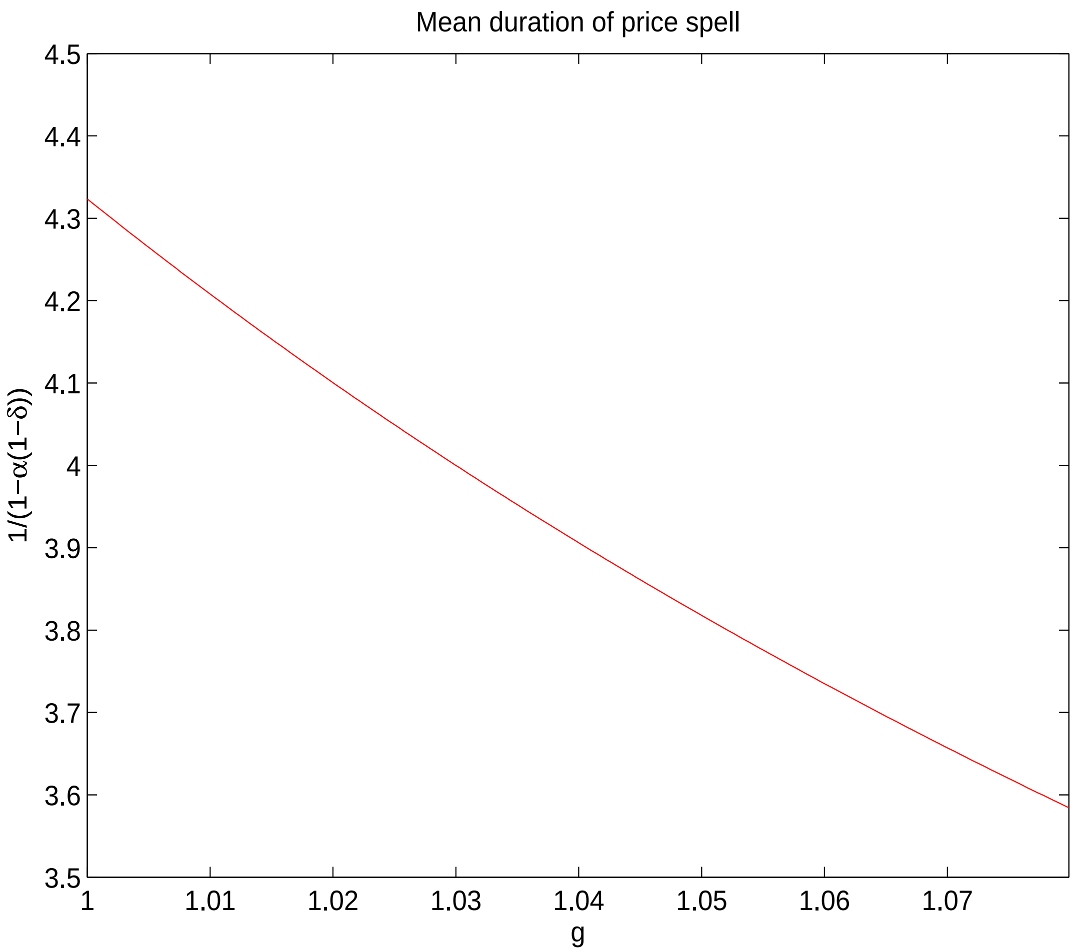

Figure 8: Mean duration of a price spell in quarters in the truncated price distribution as a function of g. The plot shows annualized productivity growth g.