Federal Reserve Bank of New York
Staff Reports

Costly Information, Planning Complementarities and the Phillips Curve

Sushant Acharya

Staff Report No. 698
November 2014

Costly Information, Planning Complementarities and the Phillips Curve
Sushant Acharya
Federal Reserve Bank of New York Staff Reports, no. 698
November 2014
JEL classification: D8, E3, E5

Abstract

Standard sticky information pricing models successfully capture the sluggish movement of aggregate prices in response to monetary policy shocks but fail at matching the magnitude and frequency of price changes at the micro level. This paper shows that in a setting where firms choose when to acquire costly information about different types of shocks, strategic complementarities in pricing generate planning complementarities. This results in firms optimally updating their information about monetary policy shocks less frequently than about idiosyncratic shocks. When calibrated to match frequent and large price changes observed in micro pricing data, the model is still capable of producing substantial non-neutralities. In addition, I use the model consistent Phillips curve and data from the Survey of Professional Forecasters to estimate the frequency at which firms update their information about monetary policy shocks. I find that the frequency of updating was higher in the 1970s compared to subsequent decades and hence conclude that monetary policy in the U.S. was relatively less effective prior to the 1980s.

Key words: sticky information, price rigidity, information choice, monetary non-neutrality, policy effectiveness

---

Acharya: Federal Reserve Bank of New York (e-mail: sushant.acharya@ny.frb.org). The author would like to thank Borağan Aruoba, Allan Drazen, John Haltiwanger, Ricardo Reis, John Shea, and Mirko Wiederholt for helpful comments. The author also thanks Julien Bengui, Javier Bianchi, Justin Burkett, Shu Lin Wee, and participants in the International Finance Division Seminar at the Federal Reserve Board and at the 2011 North American Winter Meeting of the Econometric Society for their suggestions. This paper is a modified version of Chapter 1 from my dissertation. The views expressed in this paper are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

Price stickiness is often assumed in macroeconomic models as a way of generating a Phillips curve relationship. Previous literature, dating as far back as Phelps (1970) and Lucas (1972), has stressed the importance of informational frictions to explain this relationship. More recently, *sticky information* (Mankiw and Reis, 2002; Reis, 2006) and *rational inattention* (Mackowiak and Wiederholt, 2009) have drawn on information frictions to explain the sluggish movement of prices in response to monetary policy shocks.

The importance of price stickiness in the propagation of monetary policy shocks has been questioned by recent micro-level pricing studies which suggest that prices may not be sticky. Bils and Klenow (2004) and Klenow and Kryvtsov (2008) show that the median non-housing consumer price changes once every 4.3 months. Conditional on a price change, the mean size of a price change is large in absolute terms, at about 11 percent. However, many macro studies (Christiano et al., 1999; Uhlig, 2005) find support for a role of sticky prices in the monetary transmission mechanism. Uhlig (2005) finds that only about 25 percent of the long-run response of the U.S GDP price deflator to a monetary policy shock occurs within the first year after the shock.

A possible reconciliation of the facts above is presented by Boivin et al. (2009). They find that disaggregated prices appear sticky in response to macroeconomic and monetary disturbances, but flexible in response to sector-specific shocks. Makowiak et al. (2009) also find that 100 percent of the long-run response of the sectoral price index to a sector-specific shock occurs in the month of the shock.

The sticky information approach squares well with the macroeconomic evidence on the effects of nominal shocks but struggles to explain frequently changing and volatile prices at the micro level as it assumes an exogenous arrival rate of new information. The sticky information literature, by assuming an **exogenous arrival rate**[1] for new information, cannot explain this differential rate of adjustment to idiosyncratic productivity and monetary shocks. In this paper, I present an endogenous mechanism within the sticky information setup which is able to rationalize frequent and large price changes at the micro level amid sluggish response in the aggregate price level to monetary shocks. The driving force behind this mechanism centers around the strategic motives of firms in choosing when to acquire new information. In the model, firms are allowed to change prices at zero cost but face a positive cost in updating their information, as in Reis (2006).

The basic mechanism can be summarized as follows. Strategic complementarity in pricing

[1] Reis (2006) models a choice of arrival rate of new information but features only one type of shock.

results in a strategic complementarity in the decision on when to acquire new information about the state of monetary policy, but *not* in the decision to acquire information about idiosyncratic productivity. This complementarity in planning results in a delay in the acquisition of information about monetary policy, but not about idiosyncratic productivity.[2]

When firm i gets new information about a positive nominal shock, its pricing response is dependent not only on the true state of monetary policy, but also on how other firms react to this nominal shock. If other firms do not adjust their prices, then it is not optimal for firm i to unilaterally increase its price, as relative prices matter for profit. A firm that chooses to observe new information about a recent monetary shock is forced to temper its price changes to compensate for the large fraction of firms that remain uninformed about this shock and whose prices therefore have not adjusted. This incentivizes the firm to delay acquisition of information about the monetary shock today, as it is optimal to only fully act on this information in the future (once other firms update their information) but the cost must be incurred today.

A delay in information acquisition about monetary policy by all firms implies that the aggregate price moves sluggishly and does not track money supply well. Together with pricing complementarity, this implies that the difference between the firm's current price and its target price is small. This implies that the loss from remaining uninformed and mis-pricing is small. Small losses from mis-pricing and delayed benefits from acquiring information give rise to planning complementarities which in turn lead to delayed information acquisition about the monetary state. Overall, prices increase less than proportionately on impact after a monetary shock but catch up over time as more firms update their information.

However, this complementarity in planning does not extend to a firm's decision to update its information about idiosyncratic productivity. A firm's benefit from being informed about their idiosyncratic state is not contingent on other firms' information about its individual productivity. Firms have no incentive to delay their decision to update their information about their idiosyncratic productivity and thus, do so frequently. Even if both the cost of acquiring information about monetary policy and the idiosyncratic productivity and the volatility of monetary shocks and idiosyncratic productivity are the same, firms will optimally prioritize acquiring new information about their idiosyncratic productivity. Consequently, prices change often and by large amounts in response to idiosyncratic shocks, while responding sluggishly to monetary policy shocks. The model calibrated to match the frequent and large price changes observed in micro pricing data, is

[2]This mechanism is distinct from the model of delay presented in Caballero (1989) which emphasizes information externalities in causing delay.

still capable of producing substantial non-neutralities.

The paper shows that with the appropriate modifications, Sticky Information models perform better than the existing time dependent and state dependent models at matching micro and macro pricing behavior. Time dependent models (Calvo, 1983; Yun, 1996) calibrated to match the macro sluggishness fail to match the micro facts while state dependent menu-cost models such as Golosov and Lucas (2007) fail to generate substantial non-neutrality once calibrated to match micro facts. Thus, this paper contributes to the growing literature which suggests the use of sticky information frictions over sticky-price frictions such as Calvo pricing.[3] This paper also shares similarities with the rational inattention literature. Models of rational inattention in the spirit of Mackowiak and Wiederholt (2009) assume that firms have finite information processing capacity and hence optimally choose how much information to acquire about monetary versus idiosyncratic shocks. However, a key difference in this paper is that firms choose *"when"* to update their information rather than *"how much"* information to acquire. This is consistent with survey evidence that firms conduct periodic meetings to update their pricing plans (See Zbaracki et al. (2004) and Blinder et al. (1998)). Furthermore, the pricing behavior of firms in the model can be described as setting price schedules over time rather than new prices at each point in time, which is consistent with empirical findings (Mankiw and Reis, 2010). Because the model has its roots in sticky information models, it has the added advantage of matching the term structure of disagreement for CPI inflation almost perfectly (Andrade et al., 2014).

Finally, I use data from the Survey of Professional Forecasters (Croushore, 1993) to estimate the frequency at which firms update their information about monetary policy shocks. This analysis can also be seen as an evaluation of effectiveness of US monetary policy over time in the spirit of Lucas (1973). Instead of looking at cross-country differences in the effectiveness of monetary policy, I examine how the effectiveness of monetary policy has changed since 1969. The 1970's were a period of many monetary "mistakes" (Romer and Romer, 2002). Monetary policy actively tried to exploit the Phillips curve relation and was more discretionary than in the subsequent decades. The model predicts that the central bank has a trade-off between the extent of discretionary policy changes and effectiveness of monetary policy: frequent shocks to monetary policy which are also large in magnitude causes firms to update their information about monetary shocks more frequently and hence the effect of monetary shocks on output is less persistent. Estimating the model-consistent

[3]See for example, Mankiw and Reis (2002), Mankiw and Reis (2006), Mankiw et al. (2004) etc. In addition, in recent work Kiley (2014) show that the standard New-Keynesian model exhibits policy paradoxes under the assumption of time dependent pricing but not under the assumptions of Sticky Information.

Phillips Curve separately over sub-samples covering the 1970's and the subsequent decades enables estimation of the duration for which firms choose to remain inattentive to changes in monetary policy. By looking at the estimates of this duration over the two sub-samples, I can infer that firms updated their information more frequently during the 1970's than in later decades. Hence monetary policy had a more persistent effect on output in later decades than during the 1970's. This result is complementary to the finding of Coibion and Gorodnichenko (2010) who in a different context find that when aggregate volatility fell, agents were paying less attention to the aggregate economy.

From a methodological point of view, the paper presents a model where aggregation is tractable and most results are in closed form. The solution to the model can be split into two parts, allowing the information choice to be solved as a deterministic control problem. This is similar methodologically to recent papers like Reis (2011), who decomposes a rational inattention problem into a stochastic part and a fully deterministic control problem.

The paper is arranged as follows. In the next section, I present the basic model. Section 3 contains a description of the equilibrium under costly and costless information. Section 4 contains the main analytical results. Section 5 presents the calibration strategy and discusses the quantitative performance of the calibrated model. I present the results of the empirical study in Section 6. I conclude in Section 7.

2 Model

The model combines features from Golosov and Lucas (2007), Hellwig and Veldkamp (2009) and Reis (2006). Time is continuous. The economy consists of a representative household and a unit mass of ex-ante identical monopolistically competitive firms indexed by $i \in [0, 1]$ which produce differentiated goods.

The economy is subject to two kinds of shocks: a monetary shock and firm-specific idiosyncratic shocks. The monetary authority is assumed to follow a money supply rule, which implies that the log of nominal money supply follows a Brownian motion with zero drift and variance σ_m^2

$$d \ln M(t) = \sigma_m dW(t) \tag{1}$$

where $W(t)$ is a standard Brownian motion. σ_m^2 can be thought as the degree of discretion of the monetary authority; a higher σ_m^2 corresponds to a higher degree of discretion as it means larger deviations from the deterministic component of the rule.

Firm specific productivity shocks $Z_i(t)$ are assumed to be i.i.d across firms and follow a mean reverting Ornstein-Uhlenbeck process with zero drift, rate of mean reversion η and variance σ_z^2 as in Golosov and Lucas (2007):

$$d \ln Z_i(t) = -\eta \ln Z_i(t) dt + \sigma_z dB_i(t) \tag{2}$$

where $B_i(t)$ is a standard Brownian motion such that for $j \neq j'$, B_j and $B_{j'}$ are independent. Each B_j is also independent of W.

2.1 Representative Household's Problem

The household derives utility from consumption of a final good (C) and from holding real balances $\left(\frac{M^D}{P}\right)$ while hours worked (n) reduce utility. The representative household's problem can be written as choosing the sequence of final good consumption, labor supply and real balances $\{C(t), n(t), M^D(t)/P(t)\}_{t=0}^{\infty}$ to maximize

$$E_0 \left\{ \int_0^{\infty} e^{-\rho t} \left[\frac{C(t)^{1-\gamma} - 1}{1-\gamma} - \alpha n(t) + \ln \left(\frac{M^D(t)}{P(t)} \right) \right] dt \right\}$$

subject to its budget constraint

$$M(0) \geq E_0 \left\{ \int_0^{\infty} Q(t) [P(t) c(t) + R(t) M^D(t) - \omega(t) n(t) - \Pi(t)] dt \right\} \tag{3}$$

where $Q(t)$ is the shadow price of nominal cash flows, and $\Pi(t)$ includes the nominal profits from firms and lump sum transfers. $R(t)$ is the nominal interest rate and satisfies the following:

$$Q(t) = e^{\int_t^{t+dt} R(s) ds} E_t\{Q(t + dt)\}$$

$C(t)$ is the *Dixit-Stiglitz* consumption aggregator

$$C(t) = \left[\int_0^1 c_i(t)^{\frac{\epsilon-1}{\epsilon}} di \right]^{\frac{\epsilon}{\epsilon-1}}$$

which aggregates consumption of a continuum of goods indexed $i \in [0, 1]$.

The first-order conditions with respect to $C(t)$, $n(t)$ and $M^D(t)$ can be written as

$$e^{-\rho t} C(t)^{-\gamma} = \lambda Q(t) P(t) \tag{4}$$

$$e^{-\rho t} \alpha = \lambda Q(t) \omega(t) \tag{5}$$

$$e^{-\rho t} \frac{1}{M^D(t)} = \lambda Q(t) R(t) \tag{6}$$

where λ is the multiplier on (3) and is independent of time. Utility maximization yields iso-elastic demand curves for each of the differentiated goods of the form

$$c_i(t) = C(t)\left(\frac{P_i(t)}{P(t)}\right)^{-\epsilon} \tag{7}$$

2.2 Firm's Problem

There is a continuum of ex-ante identical monopolistically competitive firms indexed by $i \in [0,1]$, each of whom produce a differentiated variety $y_{i,t}$. Firm i's production technology can be described by a non-increasing returns to scale production function

$$y_i(t) = AZ_i(t)L_i(t)^{\theta} \tag{8}$$

where $A > 0$ is a constant and $\theta \in (0,1]$. $Z_i(t)$ is the firm-specific idiosyncratic productivity shock and $L_i(t)$ is the amount of labor that the firm hires in an economy-wide labor market from households at a wage rate $\omega(t)$ at date t.[4] Firm i's nominal profit, excluding the cost of information acquisition, at date t can be written as

$$\pi_i(t) = P_i(t)c_i(t) - \omega(t)L_i(t) = P_i(t)c_i(t) - \omega(t)\left(\frac{c_i(t)}{AZ_i(t)}\right)^{\frac{1}{\theta}}$$

Each firm faces a fixed labor cost F_m , if they decide to acquire information about the monetary shock and F_z if they *plan* about the idiosyncratic state. The firm then chooses its process of prices $\{P_i(t)\}_{t=0}^{\infty}$, and a process of planning dates $\{D_i^m(t), D_i^z(t)\}_{t=0}^{\infty}$ to maximize its expected discounted profits:

$$E_0^i\left\{\int_0^{\infty} Q(t)\pi_i(t)dt - \int_0^{\infty} Q(t)\omega(t)[F_m dD_i^m(t) + F_z dD_i^z(t)]\right\} \tag{9}$$

taking as given $\{P(t), Q(t), \omega(t), C(t)\}_{t=0}^{\infty}$ and its information set at date 0. $dD_i^k(t) = 1$ refers to the firm's decision to plan about state $k \in \{m, z\}$ at date t, and $dD_i^k(t) = 0$ otherwise.

3 Equilibrium

Money Market Equilibrium. Money market equilibrium requires that $M^D(t) = M(t)$. The interest rate ensures that the demand for money equals the supply. Appendix A.1.1 shows that the nominal

[4] I abstract from capital in the production process. Including capital accumulation decisions complicates the analysis by potentially raising issues of *Forecasting the Forecasts of Others* as in Townsend (1983). Despite not modeling capital explicitly, I can evaluate the effect of fixed factors by varying the returns to scale to labor. $\theta < 1$ or decreasing returns can be interpreted as the presence of a fixed factor.

interest rate is constant in equilibrium.

$$R(t) = \rho - \frac{\sigma_m^2}{2} \ ,\forall t \tag{10}$$

Labor Market Equilibrium. Wages adjust to equate labor supply and labor demand, $n_t = \int_0^1 L_{i,t}di$. Equations (5) and (6) imply the following relationship between equilibrium the wage rate $\omega(t)$ and money supply $M(t)$:

$$\omega(t) = \alpha RM(t) \tag{11}$$

Thus, $\ln \omega(t)$ is also a Brownian motion with variance σ_m^2:

$$d \ln \omega(t) = d \ln M(t) = \sigma_m dW(t) \tag{12}$$

Goods Market Equilibrium. Equations (4) and (6), along with equation (10), imply that consumption depends on real money balances.

$$C(t) = \left[\frac{RM(t)}{P(t)}\right]^{\frac{1}{\gamma}} \tag{13}$$

Since there is no capital in the economy, the entire output is allocated either to consumption or to the resource cost of updating information. The resource constraint can be written as

$$Y(t) = C(t) + \chi(t) \tag{14}$$

where

$$\chi(t) = \frac{\omega(t)\left[F_m d\Gamma_t^m(t) + F_z d\Gamma_t^z(t)\right]}{P(t)}$$

represents the resource cost associated with the cost incurred by firms in updating their information. $d\Gamma_t^m(t)$ and $d\Gamma_t^z(t)$ represent the mass of firms updating their information about the aggregate and idiosyncratic states respectively and are formally defined in Section 3.2.

3.1 Costless Information

This is the case when $F_m = F_z = 0$ and will be referred to as the Costless Information case. Since information is free, firms update their information set at each instant and hence know the realizations of all shocks through the present.

Lemma 1. *In the Full Information case, firm i sets the optimal profit maximizing price given by*

$$\ln P_i^f(t) = \zeta \ln Z_i(t) + r \ln P^f(t) + (1-r)\ln M(t) \tag{15}$$

where $\zeta = \frac{-1}{\theta(1-\epsilon)+\epsilon}$ *and* $r = 1 - \frac{1+\gamma\theta-\theta}{\gamma(\theta(1-\epsilon)+\epsilon)}$

Proof. See Appendix A.1.2. □

In equation (15), r measures the strength of the strategic complementarity in pricing. Note that $\frac{\partial \ln P_i}{\partial \ln P} = r > 0$. If $r = 0$, then $\ln P_i(t) = \zeta \ln Z_i(t) + \ln M(t)$, i.e. there is no strategic complementarity. Strategic complementarity in pricing implies that a firm wants to set its price close to the average price.

Steady State: A steady state in this economy is defined by the case where there are no monetary or idiosyncratic productivity shocks, i.e., $\ln Z_i(t) = 0, \forall i \in [0,1]$ and $d \ln M(t) = 0$. All real variables are constant in the steady state and nominal variables grow at a constant rate. In the steady state, all firms would want to set the same price at each instant such that

$$d \ln P_i^{ss}(t) = d \ln P^{ss}(t) = 0$$

i.e. all prices increase at the rate of growth of money supply.

Proposition 1. *Classical Dichotomy: In the Full Information case, prices track nominal money balances. Monetary policy is neutral.*

Proof. Appendix A.1.3 shows that the aggregate price level tracks money supply perfectly. Equations (13) and (14) then demonstrate that real output is unaffected by a change in money supply. □

Prices reflect up-to-date information about the aggregate and idiosyncratic state at all points in time. Hence, prices adjust proportionally to changes in money supply such that nominal shocks have no effect on real output even in the short run.

3.2 Costly Information

A firm's problem can be thought of as one in which it is trying to track a target price which is a function of the not just of the monetary policy shocks but also how other firms respond to these shocks. In general, a firm does not observe the target price at each instant (since it is costly to do so) but knows the stochastic process of its target price.[5] The firm only observes a particular history of shocks when it incurs the associated cost, and hence can only observe its stochastically

[5] Since the firm observes all realizations of the exogenous shocks in the costless information case, it knows its target price exactly at all times.

evolving target price when it updates its information. The expected (net) lifetime loss to the firm from not updating its information set at each instant can be written as:

$$\begin{aligned} \mathcal{L} &= E_0^i \left\{ \int_0^\infty Q(t) \left[\pi(P_i^f(t); \mathcal{S}_i(t)) - \pi(P_i(t); \mathcal{S}_i(t)) \right] dt \right. \\ &\quad \left. + \int_0^\infty Q(t)\omega(t)[F_m dD_i^m(t) + F_z dD_i^z(t)] \right\} \end{aligned} \tag{16}$$

where $\mathcal{S}_i(t) = [P(t), M(t), Z_i(t)]$. Maximizing the objective in equation (9) is equivalent to minimizing the expression in equation (17). A second order Taylor expansion of the loss function in (17) around $P_i^f(t) - P_i(t) = 0$ yields

$$E_0^i \int_0^\infty e^{-\rho t} \left\{ [\ln P_i(t) - \ln P_i^f(t)]^2 + \mathcal{C}_m dD_i^m(t) + \mathcal{C}_z dD_i^z(t) \right\} dt \tag{17}$$

where $\mathcal{C}_m = \frac{\alpha(\epsilon-1)F_m}{\epsilon} \left[\frac{\theta-\theta\epsilon+\epsilon}{\theta}\right]$ and $\mathcal{C}_z = \frac{\alpha(\epsilon-1)F_z}{\epsilon} \left[\frac{\theta-\theta\epsilon+\epsilon}{\theta}\right]$. $\mathcal{C}_k$, $k = m, z$ can be interpreted as the cost in terms of labor of acquiring and processing information about the state k.

The first term in equation (17) is the *loss from mis-pricing* and can be interpreted as the loss from setting prices in an uninformed state. The second and third terms are the resource costs that a firm must bear when it *plans*/collects and processes new information about the aggregate and idiosyncratic state respectively. Incurring the costs to update information reduces the loss from mis-pricing.

For ease of exposition in solving the model, define the following

Define: $p_i(t) = \ln P_i(t), p(t) = \ln P(t), m(t) = \ln M(t), z_i(t) = \ln Z_i(t)$ and $p^*(t) = rp(t) + (1-r)m(t)$.

From the first order condition of the minimization problem specified in equation (17), firm i that last planned at $(\hat{\tau}_m, \hat{\tau}_z)$ will set price:

$$p_i(t) = E\{\overline{p}(t) \mid \mathcal{I}_{\hat{\tau}_m}, \mathcal{I}^i_{\hat{\tau}_z}\} = E\{p^*(t) \mid \mathcal{I}_{\hat{\tau}_m}\} + \zeta E\{z_i(t) \mid \mathcal{I}^i_{\hat{\tau}_z}\} \tag{18}$$

where $\mathcal{I}_{\hat{\tau}_m} = \{m(s)\}_{s \leq \hat{\tau}_m}$ and $\mathcal{I}^i_{\hat{\tau}_z} = \{z_i(s)\}_{s \leq \hat{\tau}_z}$ and $\overline{p}_i(t) \equiv p^*(t) + \zeta z_i(t)$ is the target (log) price that firm i wants to set to maximize profit. The aggregate price, and consequently the target price, depends on the decisions of all other firms and hence is determined endogenously. This distinguishes the current model from other research such as Bonomo et al. (2010) where this target price is exogenously specified. As mentioned earlier, the firm observes the true target price only when it simultaneously acquires information about both the monetary and idiosyncratic state. Acquiring information about the monetary state only allows the firm to calculate the first term of the target price in equation (18), while acquiring information about its idiosyncratic productivity reveals the second component. On dates when the firm does not update either type of information, the firm optimally chooses to set a price equal to the expected target price given its most recent information.

Since information acquisition is costly, firms choose not to update their information at each date and at any date t, the economy is characterized by a cross sectional distribution of firms, $\Gamma_t(\tau_m, \tau_z)$, with different vintages of information. $\Gamma_t(\tau_m, \tau_z)$ denotes the fraction of firms that updated their information about monetary shocks prior to date τ_m and about their own idiosyncratic productivity prior to date τ_z. Accordingly, the marginal distribution, $\Gamma_t^m(\tau_m) \in [0, 1]$, refers to the mass of firms at time t that acquired information about monetary policy prior to date τ_m. $\Gamma_t^z(\tau_z)$ is defined in similarly. $d\Gamma_t^m(\tau)$ and $d\Gamma_t^z(\tau)$ denote the measure of firms that acquire information exactly at date τ about monetary policy and their idiosyncratic productivity respectively.[6] The evolution of $\Gamma_t^k(\tau)$, $k = m, z$ can be written recursively as

$$\Gamma_{t+dt}^k(\tau) = \Gamma_t^k(\tau) - \int_{-\infty}^{\tau} \mathcal{D}_t^k(s) d\Gamma_t^k(s) \ , \forall t \leq \tau \text{ and } k = m, z \tag{19}$$

where $\mathcal{D}_t^k(s)$ is the probability that a firm that most recently acquired information about the aggregate (idiosyncratic)state at date s will acquire information about the aggregate (idiosyncratic) state again at date t.

A firm's pricing decision potentially depends on all the past realizations of the aggregate and idiosyncratic state, making the dimensionality of the firm's problem infinite. I use the method of undetermined coefficients to find an analytical solution to the firm's problem in Appendix A.1.4.[7]

Lemma 2. *In equilibrium:*

1. *The aggregate (log) price $p(t)$ follows the following process:*

$$p(t) = \sigma_m \int_{-\infty}^{t} \frac{[1 - \Gamma_t^m(\tau)](1 - r)}{1 - r + r\Gamma_t^m(\tau)} dW(\tau) \tag{20}$$

[6] In other words, the fraction $1 - \Gamma_t^m(\tau_m)$ is the fraction of firms that know all the realizations of the aggregate state up to date $\tau_m : \{m_s\}_{s \leq \tau_m}$ and $1 - \Gamma_t^z(\tau_z)$ has the analogous interpretation for the idiosyncratic state.

[7] The basic approach to finding the solution is a two step procedure. First, I solve for the optimal price a firm sets conditional on any information set as shown in equation (18). I then solve for the firm's optimal choice of when to update information to minimize its loss from being uninformed. The solution to this step takes the form of a deterministic control problem.

2. *The (log) price set by firm i at date t when it last updated its information at dates $(\hat{\tau}_m, \hat{\tau}_z)$ is given by:*

$$p_i(t) = \sigma_m \int_{-\infty}^{\hat{\tau}_m} \frac{1-r}{1-r+r\Gamma_t^m(\tau)} dW(\tau) + \zeta z_i(\hat{\tau}_z) e^{-\eta(t-\hat{\tau}_z)} \tag{21}$$

Proof. The proof is similar to the one in Hellwig and Veldkamp (2009) and can be found in Appendix A.1.4. □

Having determined the price set by a firm given its information set, next, we need to solve for the optimal planning horizon for each firm. In the present setup, firms choose when to update their information, so they optimally decide their information set at each instant. The difference between the forecasted and actual target price of the firm with information set $\{\mathcal{I}_{\hat{\tau}_m}, \mathcal{I}_{\hat{\tau}_z}\}$, is given by:

$$p_i(t) - \bar{p}_i(t) = \sigma_m \int_{\hat{\tau}_m}^{t} \frac{1-r}{1-r+r\Gamma_t^m(\tau)} dW(\tau) + \zeta \left[z_i(\hat{\tau}_z) e^{-\eta(t-\hat{\tau}_z)} - z_i(t) \right] \tag{22}$$

The first term is the error in forecasting the aggregate component of the target price given that forecasts are formed with respect to the information set $\mathcal{I}_{\hat{\tau}_m}$. The second term is the forecast error in the idiosyncratic component of the target price where the forecasts are made with respect to $\mathcal{I}^i_{\hat{\tau}_z}$. Since $W(t)$ and $B_i(t)$ are standard Brownian motions with unit variance, equations (22) and (17) imply that the loss of this firm from mis-pricing can be expressed as

$$L(t, \hat{\tau}_m, \hat{\tau}_z) = \underbrace{\sigma_m^2 \int_{\hat{\tau}_m}^{t} \left[\frac{1-r}{1-r+r\Gamma_t^m(\tau)} \right]^2 d\tau}_{L_1(t,\hat{\tau}_m)} + \underbrace{\zeta^2 \frac{\sigma_z^2}{2\eta} \left[1 - e^{-2\eta(t-\hat{\tau}_z)} \right]}_{L_2(t,\hat{\tau}_z)}$$

The first and second terms are the variances of the forecast errors associated with forecasting the aggregate component and the idiosyncratic component of the target price respectively. Since the loss function can be separated into a purely aggregate part and a purely idiosyncratic part, the problem of when to update information about each state can be solved as two separate problems.

The equilibrium concept in this incomplete information economy is a stationary Bayesian-Nash Equilibrium. Two structures of equilibria arise naturally: synchronized and staggered. In the synchronized equilibrium, all firms choose to update their information about state k, $k = m, z$ at the same date. In the staggered stationary equilibrium, only a fixed fraction of firms *plans* about the aggregate and idiosyncratic state at each date. Existing research suggests that price changes are not synchronized. For example, Lach and Tsiddon (1992) find a lack of synchronization of price changes in Israel. Thus, I focus on the **stationary staggered equilibrium**, which is empirically more relevant.[8]

I concentrate on the pricing problem of firm i. In equilibrium, all other firms acquire information about the aggregate state every T_m periods and about their idiosyncratic state every T_z periods. Thus, the proportion of firms acquiring information about the aggregate state over any interval is given $\frac{1}{T_m}dt$ and about the idiosyncratic state is $\frac{1}{T_z}dt$. As a result, in the staggered equilibrium

$$\Gamma_t^k(\tau) = \begin{cases} 0 & \text{if } \tau < t - T_k \\ 1 - \frac{t-\tau}{T_k} & \text{if } t - T_k \leq \tau < t \end{cases}$$

for $k = m, z$.

Consider a firm which updates its information about the monetary policy shock today. The decision of when to update its information about the monetary state next minimizes the loss from being uninformed. The problem of a firm that updates its information today can be written as:

$$\mathcal{L}_1(\hat{\tau}_m) = \min_{\tau'_m \geq \hat{\tau}_m} \int_{\hat{\tau}_m}^{\tau'_m} e^{-\rho(s-\hat{\tau}_m)} L_1(s, \hat{\tau}_m) ds + e^{-\rho(\tau'_m - \hat{\tau}_m)} [\mathcal{C}_m + \mathcal{L}_1(\tau'_m)] \tag{23}$$

where

$$L_1(t, \hat{\tau}) = \begin{cases} \sigma_m^2 \int_{\hat{\tau}}^{t} \frac{(1-r)^2}{(1-r\frac{(t-\tau)}{T_m})^2} d\tau & \text{if } \hat{\tau} \geq t - T_m \\ \sigma_m^2 \int_{t-T_m}^{t} \frac{(1-r)^2}{(1-r\frac{(t-\tau)}{t})^2} ds + \sigma_m^2 (t - T_m - \hat{\tau}) & \text{if } \hat{\tau} < t - T_m \end{cases}$$

Similarly the problem on when to next acquire information on the idiosyncratic state conditional on obtaining information today can be written as:

$$\mathcal{L}_2(\hat{\tau}_z) = \min_{\tau'_z \geq \hat{\tau}_z} \int_{\hat{\tau}_z}^{\tau'_z} e^{-\rho(s-\hat{\tau}_z)} L_2(s, \hat{\tau}_z) ds + e^{-\rho(\tau'_z - \hat{\tau}_z)} [\mathcal{C}_z + \mathcal{L}_2(\tau'_z)] \tag{24}$$

where

$$L_2(t, \hat{\tau}) = \frac{\zeta^2 \sigma_z^2}{2\eta} \left[1 - e^{-2\eta(t-\hat{\tau})} \right]$$

The solution to the firm's problem can be seen as a threshold rule. The loss from being inattentive depends on the variance of the forecast error which is strictly increasing in the duration since the firm last updated its information.[9] Once the expected loss is large enough, the firm chooses to incur the fixed cost of acquiring information about that state and resets the forecast error variance associated with that component of the target price to zero. It then conditions its forecasts of the

[8] Acharya (2013) discusses some aspects of the synchronized equilibrium.

[9] Since the loss depends deterministically on time since the firm last updated, its information, the decision of the firm can also be seen as a threshold on the forecast error variance or the time since last update.

target price on this newly expanded information set. The threshold is chosen such that if the firm did not incur the fixed cost to update its information set, the loss from mis-pricing would result in larger overall losses than the cost of obtaining information and reducing the forecast error.

Proposition 2. ***Optimal Planning Horizon**: Each firm chooses to update its information set about monetary policy shocks every T_m^* periods and about its idiosyncratic productivity every T_z^* periods.*

1. *The unique optimal horizon for planning about the monetary shock T_m^* is implicitly defined by*

$$\mathcal{C}_m = \sigma_m^2 T_m^* \int_0^{T_m^*} (1-r)e^{-\rho s} \frac{T_m^* - s}{T_m^* - rs} ds \tag{25}$$

2. *The unique optimal horizon for planning about the idiosyncratic productivity shock T_z^* is implicitly defined by*

$$\mathcal{C}_z = \frac{\zeta^2 \sigma_z^2}{2\eta} \int_0^{T_z^*} e^{-\rho\delta}(e^{-2\eta\delta} - e^{-2\eta T_z^*})d\delta \tag{26}$$

Proof. Appendix A.1.5 derives the expression for the T_m^*. The expression for T_z^* can be derived using the same procedure. □

Proposition 3. *If $\mathcal{C}_k > 0$, then $T_k^* > 0$ for $k = m, z$.*

Proof. Plugging $T_k^* = 0$ into the RHS of equation (25) or equation (26) yields zero on the RHS, which is a contradiction; i.e. it is never optimal for a firm to update its information about the aggregate state or the idiosyncratic state at each instant unless doing so is costless. □

4 Analytical Results

4.1 Differential adjustment to nominal and idiosyncratic productivity shocks

For this section, I set $\eta = 0$. As a result both, the money supply and the idiosyncratic shock processes follow a drift-less Brownian motion

$$\begin{aligned} \ln M(t) &= \sigma_m dW(t) \\ \ln Z_i(t) &= \sigma_z dB_i(t) \end{aligned}$$

Equation (26) in the limit as $\eta \to 0$ can be written as

$$\mathcal{C}_z = \zeta^2 \sigma_z^2 \int_0^{T_z^*} e^{-\rho\delta}(T_z^* - \delta)d\delta \tag{27}$$

Lemma 3. *T_m^* is increasing in the strength of the strategic complementarity, i.e.* $\frac{\partial T_m^*}{\partial r} > 0, \forall r \in [0,1]$

Proof. Lemma 3 can be verified by applying the Implicit Function Theorem on equation (25). □

Proposition 4. *For $r \in (0,1)$ and normalizing $\zeta = 1$, if $\sigma_m = \sigma_z = \sigma$ and $\mathcal{C}_m = \mathcal{C}_z = \mathcal{C}$, then* $T_m^* > T_z^*$

Note that equation (26) is of the same form as equation (27) with r set to 0 (with $|\zeta|$ normalized to 1). Lemma 3 then implies that $T_z^* < T_m^*$. Proposition 4 implies that firms choose to incorporate new information about idiosyncratic shocks into prices more often than information about aggregate shocks, even when both shocks are equally volatile and when the costs associated with updating information about each shock are equivalent. Empirical evidence suggests that idiosyncratic productivity is highly volatile relative to aggregate shocks. Thus, in the calibrated model firms update their information about idiosyncratic productivity at a substantially higher frequency. Hence, the model is able to explain differential adjustment of prices to idiosyncratic and aggregate shocks.

Strategic complementarity in pricing spills over into information acquisition decisions about monetary policy and causes a *delay* in information acquisition about monetary shocks. This can be seen as a combination of two forces.

Strategic complementarity in pricing implies that a firm seeks to limit the gap between their price and the average price as profits depend on relative prices. The firm's pricing response to a monetary shock is contingent not just on the monetary policy but also on how other firms respond to the nominal shock. When a firm gets new information about monetary policy, it realizes that a large fraction of firms remain uninformed and price according to their old information. As a result it its pricing response to current monetary shocks. This diminishes the value of obtaining information about the shock today, as the firm faces an upfront cost of acquiring information today while the benefit is accrued only in the future. Strategic complementarity in price setting, thus feeds into strategic complementarity in information acquisition and incentivizes a firm to delay its acquisition of information about monetary shocks till others acquire new information. Thus, individual prices and hence the aggregate price move sluggishly in response to monetary shocks.

Furthermore, since the aggregate price moves sluggishly, the gap between a firm's current price and its target price is small in terms of the loss from being uninformed about monetary shocks. This small loss from mis-pricing further reinforces a firms' motive to delay acquiring new information about monetary shocks and strengthen the planning complementarity. The combination of the planning and pricing complementarity implies that aggregate price moves sluggishly in response to monetary shocks.

However, a firm's price response to its own idiosyncratic productivity shock is not contingent on the actions of others, as each firm is too small and cannot affect the average price. The absence of strategic complementarities in pricing with respect to idiosyncratic productivity implies that prices respond fully to idiosyncratic shocks. Moreover, this absence of strategic complementarity in pricing extends to a lack of complementarities associated with planning over the idiosyncratic state. Hence, there is no delay in information acquisition about idiosyncratic productivity unlike the case of information acquisition regarding monetary policy. As a result, prices will reflect new information about idiosyncratic productivity more quickly than information about monetary shocks. Thus, this model is able to explain differential adjustment of prices to different shocks. Prices respond frequently to idiosyncratic productivity shocks but not monetary shocks.

4.2 Mean Reverting Idiosyncratic Productivity Shocks

Empirical studies have found that relative price changes are frequent, large but transitory. Furthermore, Mankiw and Reis (2010) and other research has documented that many price changes follow what seem like predetermined patterns that follow simple algorithms, and actual resetting of price plans based on new information seems less frequent. The model

Suppose firm i updates its information about idiosyncratic productivity today at $t = t_0$ and observes that $z(t_0) = z_0$ is above the mean. Since its marginal cost is lower this period, the firm can afford to set a lower relative price and attract more demand. However, since productivity is mean reverting, the firm expects productivity to fall back to the average at a rate η. The firm chooses not to update its information between t_0 and $t_0 + T_z^*$. During this period, it sets prices according to a simple pricing plan that it determined at t_0. The price plan stipulates that prices be raised over time towards the average price to compensate for the increasing marginal costs, and is set to track the target price as closely as possible. The price plan can be written as

$$p_i(t) = \zeta z_0 e^{-\eta(t-t_0)} \text{for } t_0 \leq t \leq T_z^* \tag{28}$$

At $t_0 + T_z^*$, the firm updates information about its idiosyncratic productivity and resets the price plan. Suppose it observes that $z(t_0 + T_z^*) = z_1$. It then resets the price plan to

$$p_i(t_0 + T_z^* + dt) = \zeta z_1 e^{-\eta dt} \text{ for } 0 \leq dt < T_z^*$$

As can be seen from the price plan in equation (28), the model predicts that relative price changes are transitory which is consistent with empirical findings.

Empirical evidence suggests that idiosyncratic shocks are very volatile and hence firms would tend to update their information about idiosyncratic shocks frequently. Prices incorporate new information more often, which results in large jumps in prices from firms resetting their price plans (See Figure 1).

In the present setup, firms update their information about idiosyncratic productivity every T_z^* [defined in equation (26)] periods and reset price plans every time they do so. They set prices according to this new plan at every subsequent date until they update their information again. Prices change at each date but only incorporate new information at discrete intervals which correspond to resetting of price plans at the arrival of new information. This is consistent with Blinder et al. (1998), who find evidence of managers' adjusting their price plans.

The model struggles at generating flat price paths. In terms of the model, flat price paths would occur when at the planning date t_0, the firm observes $z(t_0) = 0$ at the planning date. Under a diffusion process, the probability of drawing $z = 0$ is a zero measure event. Thus, without introducing any other friction, observing a flat price path is a zero measure event. A possible solution to this problem might be found in a rational inattention setup as in Woodford (2009). In such a setting, rationally inattentive firms might choose to set prices out of a discrete set even if the shock process is continuous. This would complicate the aggregation problem and is left for future research.

4.3 Volatility of Monetary Policy Shocks and Effectiveness of Monetary Policy

Proposition 5. *For $r \in (0,1)$, duration for which firms remain inattentive about state $k \in \{m, z\}$ is decreasing in the volatility of that particular state. In other words*

$$\frac{\partial T_k^*}{\partial \sigma_k^2} < 0 \ , \ \forall k \in \{m, z\}$$

Proof. The proof is provided for the relation between T_m^* and σ_m^2. The proof for T_z^* is analogous.

Define

$$F_m(T_m, \sigma_m^2) = \sigma_m^2 T_m \int_0^{T_m} (1-r) e^{-\rho s} \frac{T_m - s}{T_m - rs} ds - \mathcal{C}_m$$

Equation (25) can be seen as the case $F_m(T_m^*, \sigma_m^2) = 0$. Appendix A.1.5 shows that $\partial F_m(T_m^*, \sigma_m^2)/\partial T_m^* > 0$. A simple application of the Implicit Function Theorem then implies:

$$sign\left[\frac{dT_m^*}{d\sigma_m^2}\right] = -sign\left[\frac{\partial F_m(T_m^*, \sigma_m^2)}{\partial \sigma_m^2}\right] = - \text{ , since } 0 < r < 1$$

□

Proposition 5 has direct implications for the effectiveness of monetary policy. The longer firms choose to stay inattentive about changes to monetary policy, the more persistent is the effect of monetary policy on output. This is because the effect of monetary policy on output dies out as soon as all firms learn about the monetary policy change. If each firm updates its information frequently then all firms will become informed about a change in policy more quickly, which then reduces the effect of that change in monetary policy.

Thus, the model predicts that a monetary authority has a trade-off between effectiveness of monetary policy and the level of discretion it can use in the conduct of monetary policy. A more discretionary monetary policy, in terms of the model, corresponds to a higher variance of monetary policy σ_m^2. The model predicts that monetary policy will have more persistent effects on output in regimes with lower σ_m^2. Another way to look at this prediction is that periods with high inflation volatility (high σ_m^2) correspond to periods where the persistence of the effect of monetary policy on output is lower. I use this prediction of the model to analyze the effectiveness of monetary policy in the US since the 1970's in Section 6.

5 Calibration and Numerical Results

As is standard in the New Keynesian literature (e.g. Gertler and Leahy (2008)), I simulate the model around a zero long run inflation steady state. I draw on existing literature for the values of the preference parameters ρ, γ, α, and ϵ. The discount rate ρ is set to 0.04, following an annual calibration. The risk aversion parameter γ is set to to 2. The elasticity of substitution parameter ϵ is set to 10. The dis-utility of labor α is set to 9, which implies that roughly 33 percent of the unit time endowment is allocated to labor in steady state. The exponent on labor in the production function is calibrated to the standard value of $\theta = 2/3$. This calibration yields $r = 0.7917$ which measures the strength of the strategic complementarity in price setting. Woodford (2003) suggests values of r between 0.75 and 0.9. I set $\sigma_m = 0.0248$, which corresponds to the standard deviation of annual inflation in the Klenow and Kryvtsov (2008) data set. To calibrate the cost of acquiring and processing information, I draw on existing literature. Studies such as Chevalier et al. (2003) estimate menu costs to be of the order of about 0.75 per cent of a firm's revenue. Zbaracki et al. (2004) report that information processing costs are 6 times as large as menu costs. Thus, I set the cost of acquiring information as $0.75 \times 6 = 4.5$ percent of steady state revenue. Without any strong reason to set the cost of acquiring information about monetary shocks differently from that about the idiosyncratic state, I set $\mathcal{C}_m = \mathcal{C}_z$.

This leaves two parameters to be calibrated: the variance σ_z^2 and the rate of mean reversion η of idiosyncratic productivity shocks. I follow a calibration strategy similar to Golosov and Lucas (2007). First, I shut down the monetary policy shock. Thus, all price changes are in response to idiosyncratic shocks. I calibrate $\hat{\sigma}_z^2 \equiv \zeta^2 \sigma_z^2$ and η by targeting the average number of price reviews per year and the average size of a price change conditional on a price increase.[10] I choose the parameters to minimize the sum of the squared differences between the values of the two targets in the data and the model generated counterparts.

Alvarez et al. (2011) report that the median number of price reviews per year in the US is between 2 and 3. I target the median number of price reviews a year to be 3. Golosov and Lucas (2007) find that the average price change conditional on a price increase is 0.095 for regular price changes in the Klenow and Kryvtsov (2008) dataset. Minimizing the squared difference yields $\hat{\sigma_z} = 0.58$, which corresponds to a quarterly standard deviation of 0.14, and $\eta = 0.6652$, which corresponds to a persistence of 0.85 in a quarterly AR(1). I simulate the model with the shortest time period being a month to calculate the average size of price changes. The calibration exercise yields 0.086 as the average price change conditional on an increase and 3.9 price reviews a year. Details of the calibration are summarized in Table 1.

I calculate the optimal planning horizon about the monetary state from equation (25). The half-life of the response of output is approximately 5.2 years.[11] Thus, unlike Golosov and Lucas

[10] Equations (15) and (18) show that ζ determines the sensitivity of the profit-maximizing price to the idiosyncratic state variable and has the same effect as the variance of the idiosyncratic state variable. Therefore, I normalize $|\zeta|$ to one and only choose the variance of the idiosyncratic state variable. Thus, I calibrate $\hat{\sigma}_z^2 \equiv \zeta^2 \sigma_z^2$ instead of just σ_z^2.

[11] If I let $\mathcal{C}_m \neq \mathcal{C}_z$, then the model is able to match any characteristics of both the micro level data and sluggish responses of prices. On setting $\mathcal{C}_m$ to be 0.1 percent of steady state revenue, a positive shock to monetary policy results in a positive effect on output for approximately 6 quarters which is the length of the non-neutrality estimated in Christiano et al. (1999). It is not unreasonable to believe that the cost of acquiring information about monetary policy might be substantially smaller than the cost associated with the idiosyncratic state. Information has characteristics of a public good and becomes cheaper as more people choose to acquire it. All firms are looking for the same information about monetary policy while they only want information about their own idiosyncratic state. Thus, the higher demand for information about monetary policy makes the cost of acquiring information about it much lower.

	Parameter	Source/Target
Dis-utility of Labor	$\alpha = 9$	Steady state labor = 0.34
Coefficient RRA	$\gamma = 2$	Std. in literature
Share of Labor in output	$\theta = \frac{2}{3}$	Std. in literature
Elasticity of substitution across goods	$\epsilon = 10$	Std. in literature
Discount Rate	$\rho = 0.04$	Std. Annual Calibration
S.D. of monetary shock	$\sigma_m = 0.0248$	Annual standard deviation of inflation from Klenow and Kryvtsov (2008)
Cost of updating information	$\mathcal{C} = 0.0088$	Chevalier et al. (2003), Zbaracki et al. (2004)
S.D. of idiosyncratic shock	$\hat{\sigma}_z = 0.58$	S.D. of price changes conditional on increase = 0.91, Golosov and Lucas(2007)
Persistence of idiosyncratic shock	$\eta = 0.6652$	# Price Reviews/year= 3, Alvarez et al. (2011)

Table 1: Baseline Parameter Values

(2007), this model is able to generate substantial non-neutralities and a strongly persistent effect of a shock to monetary policy on output.

Golosov and Lucas (2007) note that the strength of the non-neutrality is increasing in the share of the fixed factor which can be measured as $1 - \theta$. This is because the degree of strategic complementarity (r) in price setting is decreasing in $1 - \theta$. Even with $\theta = 0.99$, i.e, even if the importance of the fixed factor and the degree of the strategic complementarity ($r = 0.0872$) are much lower than suggested by Woodford (2003), the model still generates significant non-neutrality. The half-life of the effect on output for $\theta = 0.99$ is 2.8 years.

Figure 1 shows the paths of prices set by a firm under costly information compared to under costless information in response to a sample path of idiosyncratic shocks. Monetary Shocks are shut down in the figure, so prices are only reacting to changes in idiosyncratic productivity. The

This can be also seen from the fact that one might learn about the state of monetary policy from reading a newspaper which is much cheaper than a firm hiring consultants to find out about its demand and cost advantages. Veldkamp and Wolfers (2007) formalize such an argument.

dashed line corresponds to the path of a firm's price under costless information. The solid line shows the reaction of price in an environment with costly information. The vertical lines indicate price reviews, i.e periods when firms update their information. The vertical dotted lines indicate a price review. The price jumps at the arrival of new information. Prices under costly information are certainly smoother but not very different from those under full information. The realized loss in profit from mis-pricing is not very large, since firms choose to update their information about idiosyncratic productivity frequently.

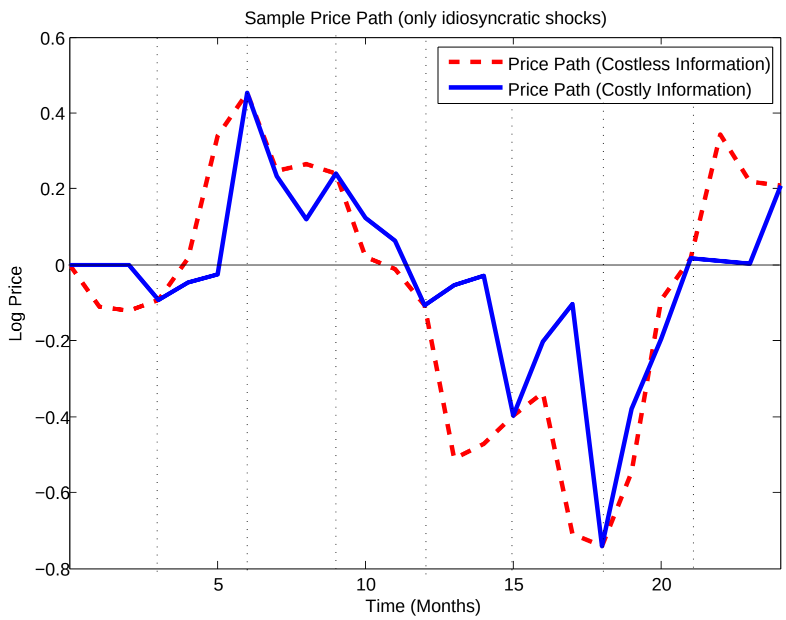

Figure 1: Price Paths of Representative firm in response to idiosyncratic shocks

The impulse response of output and the aggregate price to a positive monetary shock of one standard deviation magnitude can be seen seen in Figure 2. : The top panel depicts the response of aggregate price and the bottom panel depicts the deviation of output from the *natural level*. T^*, in the figure, is the optimal planning horizon for the monetary state. The staggered nature of the equilibrium results in a large fraction of firms being uninformed about the current true state. This allows for the model to generate sufficient heterogeneity in firm behavior. Midrigan (2011) points out that heterogeneity between firms in Golosov and Lucas (2007) is very limited and hence prices adjust very quickly to a shock in their model. Pricing and planning complementarities result in a sluggish response of the aggregate price to a monetary shock which results in output being different

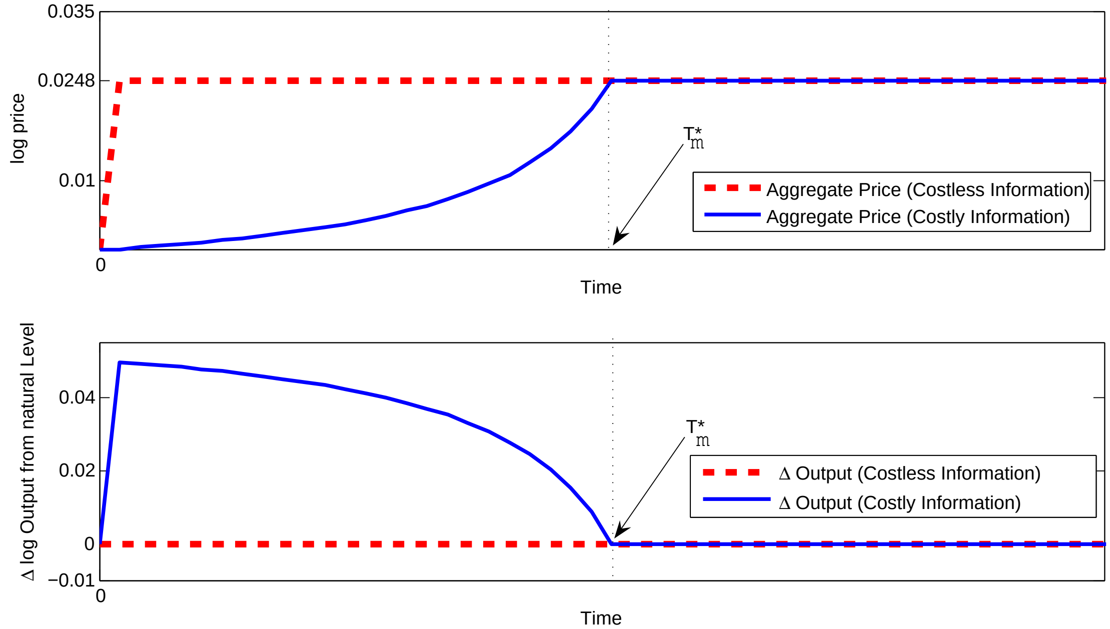

Figure 2: Impulse Response to a monetary expansion

from the natural level. Once all firms update their information sets such that each of them has incorporated the shock to monetary policy into their prices, the aggregate price adjusts *fully* to the shock and output reverts back to the natural level.

Figure 3 shows how the aggregate price evolves under costly information in reaction to a sample path for monetary policy shocks as compared to the case with costless information. The dashed line is the path of (log) money supply. The dashed line also corresponds to the path of the aggregate (log) price in reaction to monetary policy shocks under costless information. The solid line corresponds to the response of the aggregate price under the costly information case. Changes in the aggregate price are much smaller in the case with costly information compared to the case with costless information. Furthermore, the aggregate price responds slowly to shocks when information is costly. From Proposition 1, in a setting with costless information, the aggregate price tracks money supply perfectly. Thus, a fall(rise) in the dashed line in the figure corresponds to a contractionary (expansionary) shock to money supply. The solid line, representing the path of the aggregate price with costly information lags behind the dashed line. With costly information, the aggregate price starts to fall well after a contractionary shock to money supply and displays inertia as it continues to fall even after there has been an expansionary shock to monetary policy.

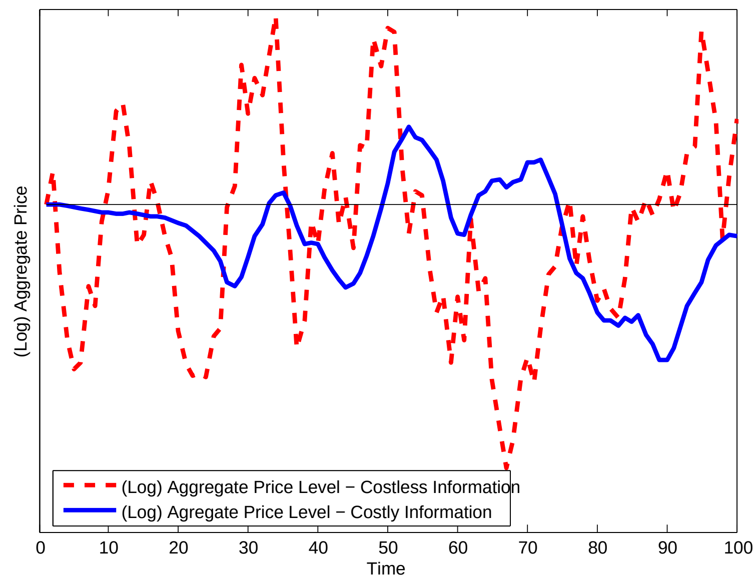

Figure 3: Sample Price Path of the aggregate price in response to a sequence of monetary policy shocks.

Figure 4 plots the distribution of price changes along a sample price path. As Midrigan (2011) and Klenow and Kryvtsov (2008) point out, standard SDP models generate what has been termed the *missing middle*, i.e they fail to explain small price changes. Midrigan (2011) also points out that the distribution of price changes exhibits excess kurtosis. The current model is capable of generating this excess kurtosis and is able to explain a large number of both small price changes and at the same time, also large price changes. Since there is no cost associated with changing prices per se, the firm changes prices even when only small price changes are warranted. This is not the case in menu cost models.

6 Effectiveness of Monetary Policy over time

In this section, I look at changes in the effectiveness of US monetary policy since 1969. This empirical study is in the spirit of Lucas (1973). Instead of looking at cross-country differences in the effectiveness of monetary policy as in Lucas (1973), I examine changes in the effectiveness of US monetary policy over time. Economists generally agree that the 1970s, a period of high volatility

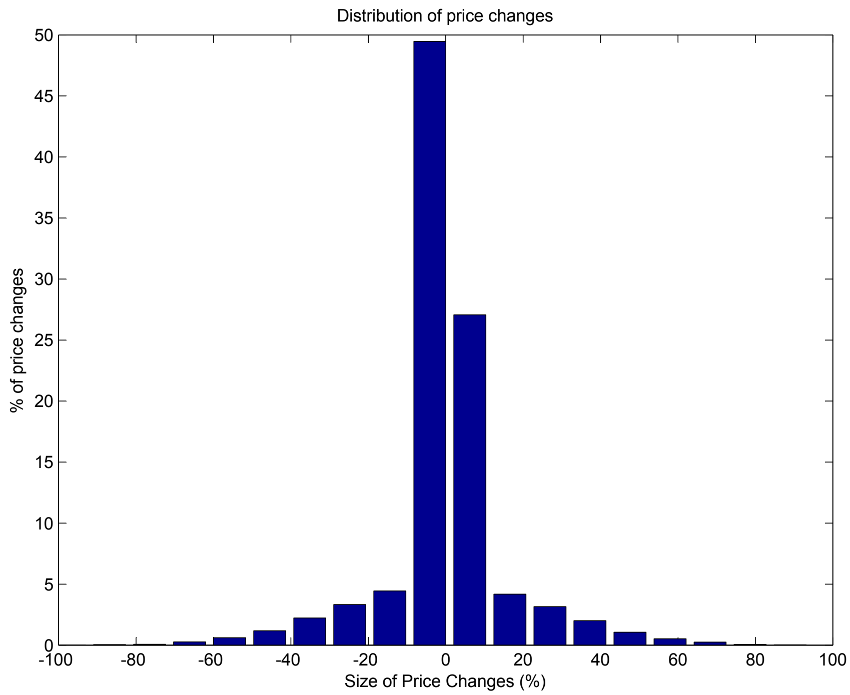

Figure 4: Distribution of size of price changes

in both output and inflation, was also a period in which monetary policy was highly discretionary and performed quite poorly, relative to both earlier and later periods (Romer and Romer, 2002).

In terms of the model, I interpret a larger variance of the shocks to monetary policy as characterizing a more discretionary regime. The larger the volatility of the monetary policy shocks, σ_m^2, the larger the variation in monetary policy which is explained by shocks and less by the systematic component of the monetary policy rule. Consider the extreme example where $\sigma_m^2 = 0$. In that case, any change in monetary policy is determined entirely by the systematic component of the monetary policy rule and corresponds to the zero discretion case viewed through the lens of the model. In terms of the current model, the 1970's would correspond to a higher σ_m^2 than in subsequent periods. Following Proposition 5, this would imply that $T_{1970's} > T_{post1970's}$, i.e. monetary policy had less persistent effects on output in the 1970's than in subsequent periods.

To test the above claim, I estimate T from equation (30) over separate sub-samples. I use the 1970's as the first sub-sample and the subsequent years as the second sub-sample. Following

Bernanke and Mihov (1998), I use the following two breakpoints for a change in policy regime between the 1970's and the subsequent period. The first breakpoint is 1979(Q3), which corresponds to the announcement of chairman Volcker's new operating regime. The second breakpoint is 1982(Q1), which Bernanke and Mihov (1998) claim as roughly corresponding to the abandonment of targeting of non-borrowed reserves in favor of interest rate targeting.

The model-consistent Phillips Curve (in discrete time) can be written as:

$$\begin{aligned}\pi_t &= \frac{p_t - E_{t-T}[p_t]}{T} + (1-r)\frac{y_t - E_{t-T}[y_t]}{T} \\ &\quad + \sum_{j=0}^{T-1} E_{t-1-j}[\pi_t + (1-r)\Delta y_t] + \epsilon_t\end{aligned} \tag{29}$$

where $T \in \{1, 2, 3, 4, ...\}$ is the number of quarters each firm chooses to wait before it updates its information about monetary policy, p_t is the log price at time t, y_t is the output gap, $r \in \mathbb{R}$ is the strength of the strategic complementarity in pricing, and ϵ_t is a measurement error which is i.i.d across time. The full derivation is available in Appendix A.1.6. Compared to the standard version of the Sticky Information Phillips Curve (SIPC) (Mankiw and Reis, 2002), this version of the model demonstrates an optimal relation between T and the number of lags that need to be included to estimate the Phillips Curve relationship consistently.[12]

To proxy for the forecasts of firms, I use data from the Survey of Professional Forecasters (Croushore, 1993). The survey provides individual responses of a panel of forecasters with up to four quarter ahead forecasts for the level of the price index and nominal GDP. Consistent forecasts of real GDP, inflation and the output gap are based on this information.[13] To calculate the values of the consolidated forecast in every quarter period, I calculate the median and mean forecasts within the cross-section of professional forecasters. Using the median or the mean does not alter results qualitatively.[14] The main limitation of the dataset is that forecasts are available only up to 4 quarters ahead which limits the parameter space for T to $\{1, 2, 3, 4\}$.[15]

[12] See Appendix A.2.1 for a detailed discussion.

[13] As in Coibion (2010), I assume that the forecasters know the CBO's measure of potential output at the time of forecasting. This enables me to calculate the forecasted output gap as the log difference between the forecasted real GDP and the measure of potential output for a particular quarter.

[14] Since the data in the period prior to the 1980's has some missing values for forecasts 4 quarters ahead of some of the variables of interest, I impute these values by averaging the data for the neighboring quarters which are available. Alternatively, dropping these entries does not alter the results qualitatively.

[15] $T = 1$ implies that each firm updates information every quarter and so there is no sticky information. In this case the SIPC relationship breaks down and equation (30) is not valid. Thus, I focus on $T \in \{2, 3, 4\}$

As a preliminary check, I estimate the parameters of equation (30) by the method of Maximum Likelihood Estimation.[16] Estimating the model over the whole sample yields an estimate of $\hat{T} = 4$. This corresponds to one price review every year,[17] which is consistent with other empirical findings such as Blinder et al. (1998). $T = 4$ is also consistent with the Mankiw and Reis (2002) estimate of the exogenous arrival rate of new information $(1 - \lambda = 0.25)$.

Estimating the model over the sub-samples separated by the first break point gives us the point estimates of $\widehat{T}_{t<1979(Q3)} = 3$ and $\widehat{T}_{t\geq 1979(Q3)} = 4$. Using the second break point also yields the estimates $\widehat{T}_{t<1982(Q1)} = 3$ and $\widehat{T}_{t\geq 1982(Q1)} = 4$. Thus, firms were updating information more frequently in the 1970's, a period which in terms of the model corresponds to a higher σ_m^2 . Conversely, firms update information more slowly in periods where monetary policy is less discretionary and σ_m^2 is low. Finally, I test the null hypothesis $H_0 : T_{t<1979(Q3)} \leq T_{t\geq 1979(Q3)}$ against the complement, and also test the null that $H_0 : T_{t<1982(Q1)} \leq T_{t\geq 1982(Q1)}$ against its complement.

Through a likelihood ratio test,[18] I am able to reject the hypothesis that $T_{t<1979(Q3)} \leq T_{t\geq 1979(Q3)}$ in favor of the alternative $T_{t<1979(Q3)} > T_{t\geq 1979(Q3)}$ at the 1% level of significance, i.e. firms updated their information more frequently prior to 1979(Q3). Moreover, using the second breakpoint, I am able to reject the hypothesis that $T_{t<1982(Q1)} \leq T_{t\geq 1982(Q1)}$ in favor of the alternative $T_{t<1982(Q1)} > T_{t\geq 1982(Q1)}$ at the 5% level of significance. These results are supportive of the model's predictions. Through the lens of the model, the aggregate price responded more quickly to monetary policy shocks in the 1970's than in subsequent decades, and hence monetary policy had a relatively less persistent effect on output in the 1970's.

7 Conclusion

Unlike the sticky information literature, the model presented in this paper is capable of explaining the differential adjustment of prices in response to monetary and idiosyncratic shocks. By relying on costly information rather than physical menu costs, the model is also consistent with the findings of Zbaracki et al. (2004), who find that information processing costs associated with pricing decisions are extremely important in determining pricing behavior. The model relies on sluggishness in prices to explain monetary non-neutrality. This sluggishness in price adjustment is caused by the decision of all firms to delay the acquisition of new information. This channel is shown to be

[16] See Appendix A.2.2 for details.

[17] The null hypothesis of $H_0 : T = 4$ v/s $H_a : T < 4$ cannot be rejected even with 90% confidence in a likelihood ratio test.

[18] see Appendix A.2.1 for details.

particularly important by Klenow and Willis (2007) who find that price changes in the CPI reflect older information than would be predicted by a costless-flexible information environment.

One of the key implications of the model is that the persistence of output in response to a monetary shock should be decreasing in the degree of monetary policy discretion, or alternatively the level of inflation volatility. Using data from the 1970's and subsequent decades to estimate the model-consistent Phillips curve, I confirm the hypothesis that the persistence was indeed lower in the 1970's, which has been identified as a period where monetary policy was actively trying to exploit the Phillips curve relation and was highly discretionary.

The model is also able to better account for features seen in micro pricing data as documented in Mankiw and Reis (2010). When calibrated to match frequently changing prices at the micro level, the model is still able to generate substantial non-neutralities. This is an improvement over standard menu cost models, as it is able to explain a sluggish response of the aggregate price to monetary shocks despite large and frequent price changes at the micro level to idiosyncratic shocks. In addition, the model does not suffer from the problem of the 'missing middle' and is able to generate a highly kurtotic distribution of price changes as has been documented by some researchers.

A shortcoming of the model is that it struggles to explain the existence of spells during which prices do not change at all. This is because price changes are costless, and thus firms are not deterred from changing prices by small amounts. Prices do not adjust continuously at low levels of inflation in the data. Adding menu costs to change prices does would resolve this problem but reintroduces the dilemma of the 'missing middle'. A possible resolution of this discrepancy can potentially be found in a rational inattention setup like in Woodford (2009), as mentioned earlier. This is left for future research.

References

Acharya, Sushant, "Informational Frictions in Macroeconomics." Dissertation, University of Maryland 2013.

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello, "Optimal Price Setting With Observation and Menu Costs," *The Quarterly Journal of Economics*, 2011, *126* (4), 1909–1960.

Andrade, Philippe, Richard K. Crump, Stefano Eusepi, and Moench Emanuel, "Fundamental Disagreement," Staff Reports No. 655, Federal Reserve Bank of New York 2014.

Bernanke, Ben S. and Ilian Mihov, "The liquidity effect and long-run neutrality," *Carnegie-Rochester Conference Series on Public Policy*, December 1998, *49* (1), 149–194.

Bils, Mark and Peter J. Klenow, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 2004, *112* (5), 947–985.

Blinder, Alan S., Elie R. D. Canetti, David E. Lebow, and Jeremy B. Rudd, *Asking about prices: a new approach to understanding price stickiness*, Russell Sage Foundation, 1998.

Boivin, Jean, Marc P. Giannoni, and Ilian Mihov, "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data," *American Economic Review*, March 2009, *99* (1), 350–84.

Bonomo, Marco, Carlos Carvalho, and Ren Garcia, "State-dependent pricing under infrequent information: a unified framework," Staff Reports 455, Federal Reserve Bank of New York 2010.

Caballero, Ricardo, "Time Dependent Rules, Aggregate Stickiness And Information," *Discussion Paper Series, MIT*, 1989.

Calvo, Guillermo A., "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 1983, *12*, 383–398.

Chevalier, Judith A., Anil K Kashyap, and Peter E. Rossi, "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, March 2003, *93* (1), 15–37.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Monetary policy shocks: What have we learned and to what end?," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1B, Elsevier, 1999.

Coibion, Olivier, "Testing the Sticky Information Phillips Curve," *The Review of Economics and Statistics*, 04 2010, *92* (1), 87–101.

_ **and Yuriy Gorodnichenko**, "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," NBER Working Papers 16537, National Bureau of Economic Research, Inc November 2010.

Croushore, Dean, “Introducing: the survey of professional forecasters,” *Business Review*, Nov 1993, pp. 3–15.

Gertler, Mark and John Leahy, “A Phillips Curve with an Ss Foundation,” *Journal of Political Economy*, 06 2008, *116* (3), 533–572.

Golosov, Mikhail and Robert E. Lucas, “Menu costs and Phillips Curves,” *Journal of Political Economy*, 2007, *115* (2), 171–199.

Hellwig, Christian and Laura Veldkamp, “Knowing what others Know: Coordination motives in information acquisition,” *Review of Economic Studies*, 2009, *76*, 223–251.

Kiley, Michael T., “Policy Paradoxes in the New Keynesian Model,” Finance and Economics Discussion Series 2014-29, Federal Reserve Board of Governors 2014.

Klenow, Peter J. and Jonathan L. Willis, “Sticky information and sticky prices,” *Journal of Monetary Economics*, 2007, *54* (Supplement 1), 79 – 99.

_ **and Oleksiy Kryvtsov**, “State-Dependent or Time-Dependent Pricing:Does It Matter for Recent U.S. Infation?,” *Quarterly Journal of Economics*, 2008.

Lach, Saul and Daniel Tsiddon, “The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data,” *Journal of Political Economy*, 1992, *100*, 349–89.

Lucas, Robert E, “Expectations and the Neutrality of Money,” *Journal of Economic Theory*, 1972, *4*, 103–124.

_ , “Some International Evidence on Output-Inflation Trade-Offs,” *American Economic Review*, 1973, *63*, 326–334.

Mackowiak, Bartosz and Mirko Wiederholt, “Optimal Sticky Prices under Rational Inattention,” *American Economic Review*, 2009, *99* (3).

Mankiw, Gregory N. and Ricardo Reis, “Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve,” *Quarterly Journal of Economics*, 2002, *117*, 1295–6.

_ **and** _ , “Pervasive Stickiness,” *American Economic Review*, 2006, *96* (2), 164–169.

Mankiw, N. Gregory and Ricardo Reis, "Imperfect Information and Aggregate Supply," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3 of *Handbook of Monetary Economics*, Elsevier, 2010, chapter 5, pp. 183–229.

_ , _ **, and Justin Wolfers**, "Disagreement about Inflation Expectations," in "NBER Macroeconomics Annual 2003, Volume 18" NBER Chapters, National Bureau of Economic Research, Inc, October 2004, pp. 209–270.

Makowiak, Bartosz, Emanuel Moench, and Mirko Wiederholt, "Sectoral price data and models of price setting," *Journal of Monetary Economics*, 2009, *56, Supplement* (0), S78 – S99. Supplement issue: December 12-13, 2008 Research Conference on 'Monetary Policy under Imperfect Information' Sponsored by the Swiss National Bank (http://www.snb.ch) and Study Center Gerzensee (www.szgerzensee.ch).

Midrigan, Virgiliu, "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *Econometrica*, 2011, *79* (4), 1139–1180.

Phelps, Edmund S., "Introduction: The New Microeconomics in Employment and Inflation Theory.," in Edmund S. Phelps et al., ed., *Microeconomic Foundations of Employment and Inflation Theory*, Vol. 1B, New York: Norton, 1970.

Reis, Ricardo, "Inattentive Producers," *Review of Economic Studies*, 2006.

_ , "When Should Policymakers Make Announcements?," *Working Paoer, Columbia University*, 2011.

Romer, Christina D. and David H. Romer, "The evolution of economic understanding and postwar stabilization policy," *Proceedings*, 2002, pp. 11–78.

Townsend, Robert M., "Forecasting the Forecast of Others," *Journal of Political Economy*, 1983, *91*, 546–588.

Uhlig, Harald, "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics*, 2005, *52*, 381–419.

Veldkamp, Laura and Justin Wolfers, "Aggregate shocks or aggregate information? Costly information and business cycle comovement," *Journal of Monetary Economics*, September 2007, *54* (Supplemen), 37–55.

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

_ , "Information-constrained state-dependent pricing," *Journal of Monetary Economics*, 2009, *56, Supplement*, S100 – S124.

Yun, Tach, "Nominal price rigidity, money supply endogeneity and business cycles," *Journal of Monetary Economics*, 1996, *37*, 345–370.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen, "Managerial and Customer Dimensions of the Costs of Price Adjustment: Direct Evidence From Industrial Markets," *Review of Economics and Statistics*, 2004, *86* (2), 514–533.

A Mathematical Appendix

A.1 Proofs

A.1.1 Existence of a constant equilibrium nominal interest rate

Let $R(t) = R$, then, by definition

$$Q(t) = e^{Rdt} E_t\{Q(t + dt)\}$$

From equation (6):

$$E_t\left(\frac{Q(t + dt)}{Q(t)}\right) = e^{-\rho dt} E_t\left(\frac{R(t)M^D(t)}{R(t + dt)M^D(t + dt)}\right)$$

Impose $R(t) = R(t + dt) = R$:

$$E_t\left(\frac{Q(t + dt)}{Q(t)}\right) = e^{-\rho dt} E_t\left(\frac{M^D(t)}{M^D(t + dt)}\right) = e^{\left(-\rho + \frac{\sigma_m^2}{2}\right)dt}$$

Thus,

$$R = \rho - \frac{\sigma_m^2}{2}$$

A.1.2 Proof of Lemma 1

The Full-Information case is when $F_m = F_z = 0$. All firms adjust prices in response to all shocks at every instant. Plugging in equations (4) - (9) and imposing market clearing into the firm's objective can be rewritten as

$$\begin{aligned} E_0^i \quad & \left\{\frac{1}{\lambda}\int_0^\infty e^{-\rho t}\left[\left(\frac{RM(t)}{P(t)}\right)^{\frac{1}{\gamma}-1}\left(\frac{P_i(t)}{P(t)}\right)^{1-\epsilon} - \frac{\alpha}{[AZ_i(t)]^{\frac{1}{\theta}}}\left(\frac{RM(t)}{P(t)}\right)^{\frac{1}{\gamma\theta}}\left(\frac{P_i(t)}{P(t)}\right)^{-\frac{\epsilon}{\theta}}\right]dt \right. \\ & \left. - \alpha\left[F_m\int_0^\infty e^{-\rho t} dD_i^m(t)dt + F_z\int_0^\infty e^{-\rho t} dD_i^z(t)dt\right]\right\} \end{aligned}$$

Each firm sets $P_i(t)$ so as to maximize the objective function above:

$$P_i^f(t) = \left[\frac{\epsilon\alpha}{\theta(\epsilon - 1)}\frac{R^{\frac{1+\theta\gamma-\theta}{\gamma\theta}}}{A^{\frac{1}{\theta}}}\right]^{\frac{\theta}{\theta(1-\epsilon)+\epsilon}} Z_i(t)^{-\frac{1}{\theta(1-\epsilon)+\epsilon}} M(t)^{\frac{1+\gamma\theta-\theta}{\gamma(\theta(1-\epsilon)+\epsilon)}} P(t)^{1-\frac{1+\gamma\theta-\theta}{\gamma(\theta(1-\epsilon)+\epsilon)}}$$

Define $A = \left[\frac{\epsilon\alpha}{\theta(\epsilon-1)}R^{\frac{1+\gamma\theta-\theta}{\gamma\theta}}\right]^\theta$ so that the initial constant term goes to 1. Thus,

$$P_i^f(t) = Z_i(t)^{-\frac{1}{\theta(1-\epsilon)+\epsilon}} M(t)^{\frac{1+\gamma\theta-\theta}{\gamma(\theta(1-\epsilon)+\epsilon)}} P(t)^{1-\frac{1+\gamma\theta-\theta}{\gamma(\theta(1-\epsilon)+\epsilon)}}$$

Taking logs on both sides

$$\ln P_i^f(t) = \zeta \ln Z_i(t) + r \ln P^f(t) + (1-r) \ln M(t) \tag{30}$$

where $\zeta = \frac{-1}{\theta(1-\epsilon)+\epsilon}$ and $r = 1 - \frac{1+\gamma\theta-\theta}{\gamma(\theta(1-\epsilon)+\epsilon)}$.

A.1.3 Proof of Proposition 1

The price index $P = \left[\int_0^1 P_i(t)^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}$ can be approximated by

$$p(t) = \int_0^1 p_i(t)di$$

around any symmetric equilibrium. Thus, integrating equation (30) over $i \in [0,1]$ yields:

$$\begin{aligned} p^f(t) &= \int_0^1 p_i^f(t)di = \zeta \int_0^1 z_i(t)di + rp^f(t) + (1-r)m(t) \\ &= rp^f(t) + (1-r)m(t) \end{aligned}$$

which implies

$$p^f(t) = m(t)$$

A.1.4 Proof of Lemma 2

Guess that $p(t)$ follows the path

$$p(t) = \sigma_m \int_{-\infty}^{t} g_t(\tau)dW(\tau)$$

Plugging this guess into the expression for $p^*(t)$ yields

$$p^*(t) = \sigma_m \int_{-\infty}^{t} [1 - r + rg_t(\tau)]dW(\tau)$$

Note that

$$E\{p^*(t) \mid \mathcal{I}_{\hat{\tau}_m}\} = \sigma_m \int_{-\infty}^{\hat{\tau}_m} [1 - r + rg_t(\tau)]dW(\tau)$$

and

$$E\{z_i(t) \mid \mathcal{I}^i_{\hat{\tau}_z}\} = z_i(\hat{\tau}_z)e^{-\eta(t-\hat{\tau}_z)}$$

Thus, from equation (18), firm i with the information set $\mathcal{I}^i_t = \mathcal{I}_{\hat{\tau}_m} \times \mathcal{I}^i_{\hat{\tau}_z}$ is

$$p_i(t) = \sigma_m \int_{-\infty}^{\hat{\tau}_m} [1 - r + rg_t(\tau)]dW(\tau) + \zeta z_i(\hat{\tau}_z)e^{-\eta(t-\hat{\tau}_z)}$$

The aggregate (log) price can be derived by integrating over the two distributions Γ_t^m and Γ_t^z.

$$p(t) = \sigma_m \int_{-\infty}^{t} [1 - \Gamma_t^m(\tau)][1 - r + r g_t(\tau)] dW(\tau)$$

Using the method of undetermined coefficients yields

$$g_t(\tau) = \frac{[1 - \Gamma_t^a(\tau)](1 - r)}{1 - r + r\Gamma_t^a(\tau)}$$

Thus,

$$p(t) = \int_{-\infty}^{t} \frac{(1 - r)(1 - \Gamma_t^m(\tau))}{1 - r + r\Gamma_t^m(\tau)} dW(\tau)$$

Plugging this into the expression for

$$\begin{aligned} p^*(t) &= r\sigma_m \int_{-\infty}^{t} \frac{([1 - \Gamma_t^a(\tau)]1 - r)}{1 - r + r\Gamma_t^a(\tau)} dW(\tau) + (1 - r)\sigma_m \int_{-\infty}^{t} dW(\tau) \\ &= \sigma_m \int_{-\infty}^{t} \frac{1 - r}{1 - r + r\Gamma_t^a(\tau)} dW(\tau) \end{aligned}$$

A firm's expectation of $p^*(t)$ follows:

$$E\{p^*(t) \mid \mathcal{I}_{\hat{\tau}_m}\} = \sigma_m \int_{0}^{\hat{\tau}_m} \frac{1 - r}{1 - r + r\Gamma_t^m(\tau)} dW(\tau) \tag{31}$$

This follows from the fact that $E_s W(t) = 0$, $\forall t > s$ as $W(t)$ is a standard Brownian motion. Since firm i at date t sets price $p_i(t)$ given by:

$$\begin{aligned} p_i(t) &= E\{p^*(t) \mid \mathcal{I}_{\hat{\tau}_m}\} + \zeta E\{z_i(t) \mid \mathcal{I}_{\hat{\tau}_z}^i\} \\ &= \sigma_m \int_{0}^{\hat{\tau}_m} \frac{1 - r}{1 - r + r\Gamma_t^m(\tau)} dW(\tau) + \zeta z_i(\hat{\tau}_z) e^{-\eta(t - \hat{\tau}_z)} \end{aligned}$$

where $E\{z_i(t) \mid \mathcal{I}_{\hat{\tau}_z}^i\} = z_i(\hat{\tau}_z) e^{-\eta(t - \hat{\tau}_z)}$ is the expression for the conditional expectation of $z_i(t)$ which follows an Ornstein-Uhlenbeck process.

A.1.5 Proof of Proposition 2

The problem of finding the optimal planning horizon about the aggregate state can be reformulated with the time since when last information was acquired. Define $\delta_m = t - \hat{\tau}_m$ and $\delta_z = t - \hat{\tau}_z$ as the time since firm i last acquired information about the aggregate state and about the idiosyncratic state respectively. Thus, the two Bellman equations above can be rewritten as

$$\mathbb{L}_1(\delta_m) = \min_{\delta_m' \geq \delta_m} \int_{0}^{\delta_m' - \delta_m} e^{-\rho s} L_1(s) ds + e^{-\rho(\delta_m' - \delta_m)} [\mathcal{C}_m + \mathbb{L}_1(0)] \tag{32}$$

$$\mathbb{L}_2(\delta_z) = \min_{\delta_z' \geq \delta_z} \int_{0}^{\delta_z' - \delta_z} e^{-\rho s} L_2(s) ds + e^{-\rho(\delta_z' - \delta_z)} [\mathcal{C}_z + \mathbb{L}_2(0)] \tag{33}$$

where

$$L_1(\delta_m) = \begin{cases} \sigma_m^2 \int_0^{\delta_m} \frac{(1-r)^2}{(1-r\frac{s}{T_m})^2} ds & \text{if } \delta_m \leq T_m \\ \sigma_m^2 \int_0^{T} \frac{(1-r)^2}{(1-r\frac{s}{T_m})^2} ds + \sigma_m^2(\delta_m - T_m) & \text{if } \delta_m > T_m \end{cases}$$

and

$$L_2(\delta_z) = \zeta^2 \sigma_z^2 \int_0^{\delta_z} ds = \zeta^2 \sigma_z^2 \delta_z$$

The solution to the first Bellman equation is characterized by the optimal planning horizon (for aggregate money shocks) T_m^*, iff $t > s + T_m^*$. The first order conditions with respect to $\hat{\delta}_m$ for the problem described in equation (32), using the fact that the optimal horizon is T_m^*, we can write

$$L_1(T_m^*) = \rho[\mathcal{C}_q + \mathbb{L}_1(0)] \tag{34}$$

where

$$\mathbb{L}_1(0) = \frac{\int_0^{T_m^*} e^{-\rho s} L_1(s) ds + e^{-\rho T_m^*} \mathcal{C}_m}{1 - e^{-\rho T_m^*}} \tag{35}$$

Following equations (34) and (35), in equilibrium since $T_m = T_m^*$, it must be the case that

$$L_1(T_m^*) = \frac{\rho}{1 - e^{-\rho T_m^*}} \left(\int_0^{T_m^*} e^{-\rho s} L_1(s) ds + \mathcal{C}_m \right)$$

or

$$\mathcal{C}_m = \int_0^{T_m^*} e^{-\rho \delta} [L_1(T_m^*) - L_1(\delta)] d\delta$$

Recall that

$$L_1(\delta) = \sigma_m^2 \int_0^{\delta} \frac{(1-r)^2}{\left(1 - r\frac{s}{T_m}\right)^2} ds$$

if $\delta \leq T_m$. Define $\Theta = \frac{s}{T_m^*}$, $L_1(\delta)$ can be written as

$$L_1(\delta) = \sigma_m^2 T_m^* \int_0^{\delta/T_m^*} \left(\frac{1-r}{1-r\Theta} \right)^2 d\Theta = \sigma_m^2 (1-r)^2 T_m^* \frac{\delta}{T_m^* - r\delta}$$

which can be used to write

$$L_1(T_m^*) - L_1(\delta) = (1-r) T_m^* \frac{T_m^* - \delta}{T_m^* - r\delta}$$

Therefore, T_m^* is implicitly defined by

$$F_m(\sigma_m, r, \mathcal{C}_m, T_m^*) = 0$$

where

$$F_m(\sigma_m, r, \mathcal{C}_m, T_m) = \sigma_m^2 T_m \int_0^{T_m} (1-r)e^{-\rho s}\frac{T_m - s}{T_m - rs}ds - \mathcal{C}_m$$

Note that

$$\frac{\partial F_m}{\partial T_m^*} = \sigma_m^2(1-r)\int_0^{T_m^*} \frac{T_m^* - s}{T_m^* - rs}ds + \sigma_m^2(1-r)T_m^* \int_0^{T_m^*} \frac{s(1-r)}{(T_m^* - rs)^2}ds > 0 \text{ for } 0 < r < 1$$

Since, $F_m(\sigma_m, r, \mathcal{C}_m, 0) = -\mathcal{C}_m < 0$ and $\frac{\partial F_m}{\partial T_m^*} > 0$, F_m crosses zero only once, $T_m^* > 0$ is unique.

A.1.6 Derivation of the Phillips Curve

To estimate the Phillips curve, I use a discrete time approximation to the problem. In the staggered equilibrium, a particular firm i which last updated information at date τ, at time t wants to set the (log) price

$$p_i(t;\tau) = E_\tau\left[p(t) + (1-r)y(t)\right]$$

where F_τ denotes the forecast of the (log) target price at time t based on information as of time τ. $E_\tau p(t)$ is the expected log price at t as of τ and $E_\tau y(t)$ is the expected output gap (in percentages) as at t as of τ.

Since in the staggered symmetric equilibrium, a constant fraction $\frac{1}{T}$ of firms update their information every period, at time t, no firm has information older than $t - T - 1$, where T is the optimal planning horizon(that was determined as an equilibrium object). Thus, a mass $\frac{1}{T}$ of firms have information which is current as of date $s \in [t - T + 1, t]$. Thus, the aggregate (log) price at time t is the average of the price set by each set of firms with different vintages:

$$\begin{aligned} p(t) &= \sum_{s=0}^{T-1} \frac{E_{t-s}\left[p(t) + (1-r)y(t)\right]}{T} \\ &= -E_{t-T}\left[\frac{p(t) + (1-r)y(t)}{T-1}\right] + \sum_{s=0}^{T-1} \frac{E_{t-1-s}\{p(t) + (1-r)y(t)\}}{T-1} \\ &\quad + \frac{(1-r)y(t)}{T-1} \end{aligned} \tag{36}$$

Similarly, the aggregate (log) price at $t-1$ can be written as

$$p(t-1) = \sum_{s=0}^{T-1} \frac{E_{t-1-s}\{p(t-1) + (1-r)y(t-1)\}}{T} \tag{37}$$

Subtract equation (36) from (37) to get equation (30) minus the measurement error:

$$\pi(t) \;=\; \frac{p(t)+(1-r)y(t)}{T} + \sum_{s=0}^{T-1} \frac{E_{t-s-1}\{\pi(t)+(1-r)\Delta y(t)\}}{T}$$
$$-\frac{E_{t-T}\{p(t)+(1-r)y(t)\}}{T} \qquad (38)$$

A.2 Data Appendix

A.2.1 Estimation of the SIPC

The standard SIPC (Mankiw and Reis, 2002) can be written as

$$\pi_t = \frac{(1-\lambda)(1-r)}{\lambda} y_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j}[\pi_t + (1-r)\Delta y_t]$$

where $1-\lambda$ is the exogenously specified rate of arrival of new information and also corresponds to the fraction of firms receiving new information every period. It is necessary to truncate the number of lagged variables as estimating an equation with infinite regressors is not feasible. However, the determination of the number of lags to be included is not supported by any economic theory or restriction from within the model. Additionally, truncation introduces another error term which is usually correlated with the regressors, causing estimates to be generally inconsistent. [See Coibion (2010) for a more detailed discussion on this inconsistency]

The setup in the present paper presents a slightly modified approach which lets one circumvent these problems. Given that each firm updates its information every T quarters, in a staggered setting, a fraction $\frac{1}{T}$ of firms update their information each period. This implies that, at any date t, there is no firm that has information older than vintage $t-T$. This makes the optimal choice of lags equal to T. This forms a long run identifying restriction ensuring that monetary policy is neutral in the long run. Since all firms know about any changes to monetary policy prior to $t-T+1$, the aggregate price has responded proportionally to those changes in policy and hence there is no effect on output anymore. The standard SIPC imposes a different long run identifying restriction to ensure long run neutrality: $\lim_{t\to\infty} \lambda(1-\lambda)^t = 0$. This restriction in the standard SIPC implies that all firms receive new information at least asymptotically. However, this long run restriction does not provide a criterion for picking the number of lags to be used in the estimation.

A.2.2 Maximum Likelihood Estimation

The equation to estimate is

$$\begin{aligned}\pi(t) &= \frac{p(t)+ry(t)}{T} + \sum_{s=0}^{T-1}\frac{E_{t-s-1}\{\pi(t)+r\Delta y(t)\}}{T} \\ &\quad -\frac{E_{t-T}\{p(t)+ry(t)\}}{T} + \epsilon_t \end{aligned} \tag{39}$$

where $\epsilon_t \sim N(0,\sigma^2)$ is a measurement error i.i.d across time. The parameter space is defined as follows

$$\Theta = \{(T, r, \sigma) \mid T \in \{2,3,4\}, r \in \mathbb{R}, \sigma \in \mathbb{R}_+\}$$

I estimate the parameters using Maximum Likelihood estimation. The Likelihood function given data X can be written as

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{t_N-t_1} exp\left\{-\frac{1}{2\sigma T}\sum_{t=t_1}^{t_N}\left(T\pi(t)-[p(t)+ry(t)]+E_{t-T}[p(t)+ry(t)]\cdots\right.\right.$$
$$\left.\left.-\sum_{s=0}^{T-1}E_{t-s-1}[\pi(t)+r\Delta y(t)]\right)^2\right\}$$

A.2.3 Hypothesis Testing

To test Proposition 5, I estimate the model over two subsamples. Thus, the model to be estimated is

$$\begin{aligned}\pi(t) &= \mathcal{I}(t<break)\left[\frac{p(t)+r_{pre}y(t)}{T_{pre}} - \frac{E_{t-T_{pre}}\{p(t)+r_{pre}y(t)\}}{T_{pre}}\cdots\right. \\ &\quad \left.+\sum_{s=0}^{T_{pre}-1}\frac{E_{t-s-1}\{\pi(t)+r_{pre}\Delta y(t)\}}{T_{pre}}\right]\cdots \\ &\quad +\mathcal{I}(t\geq break)\left[\frac{p(t)+r_{post}y(t)}{T_{post}} - \frac{E_{t-T_{post}}\{p(t)+r_{post}y(t)\}}{T_{post}}\cdots\right. \\ &\quad \left.+\sum_{s=0}^{T_{post}-1}\frac{E_{t-s-1}\{\pi(t)+r_{post}\Delta y(t)\}}{T_{post}}\right] + \varepsilon_t\end{aligned}$$

To implement the Likelihood ratio test I calculate the maximized log likelihood of the restricted model where I restrict $T_{pre} \leq T_{post}$ and the unrestricted model where T_{pre} and T_{post} are left unrestricted. The test statistic takes the following form

$$LR = -2\left[\ln\mathcal{L}(restricted) - \ln\mathcal{L}(unrestricted)\right]$$

which is distributed as a χ^2 with 1 degree of freedom. The critical values for 1% is 6.6349, 5% is 3.8415 and for 10% is 2.7055.

For the first breakpoint, with the median forecasts, the test statistic is 10.386 and the null can be rejected with 99% confidence.; with mean forecasts, the test statistic is 4.6132 and the null can be rejected with 95% confidence.

For the second breakpoint, with the median forecasts, the test statistic is 10.386 and the null can be rejected with 99% confidence.; with mean forecasts, the test statistic is 4.6132 and the null can be rejected with 95% confidence.