Persistent Monetary Non-neutrality in an Estimated Model with Menu Costs and Partially Costly Information*

Marco Bonomo — Insper

Carlos Carvalho — Central Bank of Brazil, PUC-Rio

René Garcia — Université de Montréal

Vivian Malta — IMF

July 2018

Abstract

We propose a price-setting model which helps reconcile microeconomic evidence of relatively frequent price adjustments with persistent real effects of monetary shocks. In our model, both price adjustments and the gathering of some types of information are costly, requiring the payment of a lump-sum cost. Additional relevant information flows continuously, and can be factored into pricing decisions costlessly. We estimate three versions of the model by a Simulated Method of Moments, including a special case in which all information is costly. When idiosyncratic information is free and aggregate information is costly, our estimated model is able to match individual price-setting statistics for the U.S. and, at the same time, produce persistent monetary non-neutrality.

JEL classification codes: E00, E31

Keywords: menu costs, information costs, infrequent information, partial information, inattentiveness, optimal price setting, state-dependent pricing, time-dependent pricing.

*For comments and suggestions, we thank Fabia Carvalho, Rui Castro, Virgiliu Midrigan, Oleksiy Kryvtsov, and seminar participants at the INSPER Workshop on Bayesian VAR and DSGE 2017, Cleveland Fed's 2016 Conference on "Inflation: Drivers and Dynamics", ESWC 2015, SED 2015, REAP session at SBE 2014, LAMES-LACEA 2014, EEA-ESEM 2014, NASM 2014, LuBraMacro 2013, Central Bank of Chile, Universidade de Chile, EESP-FGV, EPGE-FGV, Bank of England, INSPER, Cambridge, Oxford, UCL, Columbia, Bank of Canada, FRB of Minneapolis, and FEA-USP. Rodolfo Rigato, André Sztutman, and Felipe Alves provided top-tier research assistance. Any errors are our own. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Central Bank of Brazil.

Emails: bonomo@insper.edu.br, cvianac@econ.puc-rio.br, rene.garcia@umontreal.ca, vmalta@imf.org.

1 Introduction

Aggregate inflation and output respond slowly to monetary shocks. In contrast, individual prices change somewhat frequently. Reconciling these two pieces of evidence is one of the key challenges for the literature that studies the microfoundations of monetary non-neutrality. Since the seminal work of Bils and Klenow (2004), this field has experienced renewed interest and noteworthy developments. Many papers have expanded the frontier of so-called menu-cost models In addition, recent work has analyzed the implications of explicit information frictions for price-setting behavior.[1] Menu costs models fit microdata reasonably well, but, as Golosov and Lucas (2007) point out, tend to generate little monetary non-neutrality.[2] On the other hand, information costs generate more realistic non-neutrality, but usually fail to produce nominal price rigidity.[3]

In this paper we propose a price-setting model that helps reconcile microeconomic evidence of relatively frequent price adjustments with persistent real effects of nominal shocks. In our price-setting model, both price adjustments and the gathering of some types of information about marginal costs are costly, requiring the payment of lump-sum costs.[4] Additionally, another relevant part of the information about firms' marginal costs flows continuously, and can be factored into pricing decisions somewhat costlessly.[5] We employ a novel approach to solve this model, and estimate alternative versions by a Simulated Method of Moments (SMM), using price-setting statistics computed from micro price data from the U.S. Bureau of Labor Statistics.[6] Finally, we study how the estimated economies react to monetary shocks.

The version of the model favored by the data features free idiosyncratic information and costly aggregate information. It is able to match individual price-setting statistics and, at the same time, produce persistent monetary non-neutrality. The reason is that most price changes reflect idiosincratic information only, whereas aggregate information, which is key to the extent of monetary non-neutrality, is collected infrequently.[7]

[1] For instance Bonomo and Carvalho (2004), Reis (2006), Woodford (2009), Maćkowiak and Wiederholt (2009), and Alvarez, Lippi and Paciello (2016). Caballero (1989) is a predecessor to some of this work.

[2] There are exceptions, however, such as Gertler and Leahy (2008) and Nakamura and Steinsson (2010). These papers show that these models can produce sizable non-neutrality, but this requires additional ingredients. Gertler and Leahy (2008) explore real rigidities generated by segmented input markets, and fat-tailed shocks. Nakamura and Steinsson (2010) explore real rigidities produced by input-output linkages.

[3] For example, models with lump-sum information costs tend to imply continuous price changes whenever the frictionless optimal price features a known drift. Models with information frictions based on rational inattention can produce nominal price rigidity (e.g. Matejka 2015a, 2015b).

[4] The literature on price setting with information frictions usually assumes only one type of information (e.g. Caballero 1989, Reis 2006, Moscarini 2004, Bonomo and Carvalho 2004, 2010, Woodford 2009, Gorodnichenko 2008, and Alvarez, Lippi, and Paciello 2011, 2016).

[5] Gorodnichenko (2008), Knotek (2010) and Klenow and Willis (2007) propose menu-cost models in which firms continuously incorporate partial information into pricing decisions.

[6] We thank Oleksiy Kryvtsov for providing us with those statistics.

[7] Readers familiar with Mackowiak and Wiederholt (2009) will recognize a flavor of their main result here. However, their model features only information frictions, and prices change continuously.

The optimal price-setting model we propose is not trivial to solve. It differs from pure state-dependent pricing problems, which can be cast in terms of controlling the discrepancy between a firm's current price and its frictionless optimal level. This price discrepancy determines foregone profits due to price rigidity, and is thus the relevant state variable based on which firms optimize. When the price discrepancy becomes large enough, firms incur the menu cost to adjust their prices.[8]

In our model, the difficulty comes from the fact that part of the relevant information about the frictionless optimal price is observed infrequently, owing to the lump-sum nature of the information cost. At the same time, since part of the relevant information is freely and continuously available, it may be optimal for firms to act upon it and change their prices. Hence, in our model firms generally do not know the exact value of foregone profits at a point in time, and must estimate it given available information.

To solve this problem, we employ a tractable unified framework for solving optimal time- and state-dependent price-setting problems. The key to making our approach tractable is to decompose the estimate of a firm's foregone profits into two terms: a first component based on the firm's best estimate of its price discrepancy, and a second component originated in the uncertainty about this estimate. Under commonly used assumptions, this uncertainty increases with the time elapsed since the last time the firm acquired full information (henceforth an *information date*). This decomposition allows us to define two state variables, based on which firms make decisions: the estimated price discrepancy given the firm's information set, and the time elapsed since its last information date. Hence, the optimal price-setting problem becomes a two-dimensional optimal stopping problem, which can be solved using methods commonly employed to price American options.[9] Our framework can be used to study several models with adjustment costs and infrequent information, including most price-setting problems analyzed previously in the literature. Thus, our contribution is also methodological.[10]

In our partial information model, a firm always has the option to incur the lump-sum adjustment cost (K) and make a price adjustment based on the current estimate of its price discrepancy. It can also incur the lump-sum information cost (F) to become fully informed. Figure 1 depicts the optimal policy for illustrative parameter values. It is characterized by an inaction region in the space defined by the two aforementioned state variables: the firm's estimated price discrepancy (z) and the time elapsed since the last information date (τ). Inside the inaction region, firms neither change prices nor gather information. The subset of the border of the inaction region depicted as solid blue lines dictates price changes. When its estimated price discrepancy hits $l(\tau)$ ($u(\tau)$), the firm increases

[8] Throughout, we follow most of the literature in working with a quadratic profit-loss function, which can be obtained as a second-order approximation around the frictionless optimal price. For textbook treatments, see Dixit (1993) and Stokey (2008).

[9] See, e.g., Wilmott et al. (1993).

[10] We first developed and employed this methodology in a retired working paper, which preceed this one (see Bonomo, Carvalho and Garcia 2011).

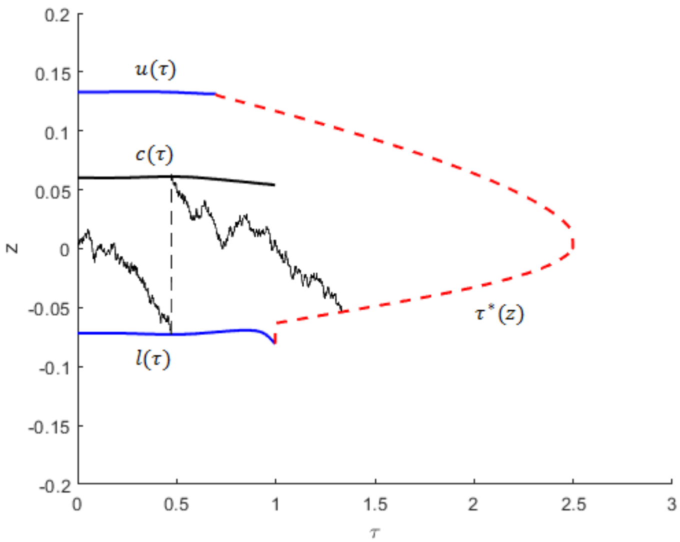

Figure 1: Optimal pricing policy under partially costly information (illustrative parameter values).

(decreases) its price to change its estimated price discrepancy to $c(\tau)$. We refer to these as *partially informed* price changes. The subset of the border of the inaction region depicted as a dashed red line ($\tau^*(z)$) triggers information gathering. At that point, the firm's estimated price discrepancy jumps, as the firm learns the cummulative unobserved shocks to the frictionless optimal price that took place since its last information date. It then decides whether or not to incur the menu cost to adjust its price. It does so whenever the learned price discrepancy falls outside the $(l(0), u(0))$ interval, in which case the firms sets its price discrepancy to $c(0)$. Hence, the optimal policy is both time- and state-dependent.

We estimate the partial information model under two alternative assumptions for the costly and free components of firms' frictionless optimal prices. In one version, price-setters can costlessly incorporate all the information about aggregate shocks into their pricing decisions, while observing/processing idiosyncratic shocks is costly. The other version of the model reverses these assumptions. Finally, we also estimate a version of the Alvarez et al. (2011) model, in which all information is costly and firms also face menu costs.

All three models fit the five targeted moments used in estimation reasonably well, but they perform distinctly in fitting the distribution of the duration of price spells – an "untargeted moment". The partial information model with costly aggregate information is the only one capable of generating a distribution that resembles the empirical distribution reported in Klenow and Kryvtsov (2008). Moreover, that version of the model is the only one that is not rejected by a test of overidentifying restrictions.

Finally, we study the extent of monetary non-neutrality in the estimated economies. The model favored by the data, with costly aggregate information, produces long-lasting real effects of monetary shocks. In contrast, both the model with free aggregate information and the model in which all information is costly produce short-lived output responses.[11]

Several papers contribute to bridging the gap between fitting microeconomic pricing facts and macroeconomic evidence on the real effects of monetary shocks.[12] Some recent papers also explore mechanisms that rely on frequent adjustments based on idiossyncractic shocks, and less frequent macroeconomic information updating. Maćkowiak and Wiederholt (2009) argue that large idiosyncratic shocks lead firms to devote scarce attention to tracking them, at the expense of aggregate shocks. Klenow and Willis (2007) introduce infrequent information collection at exogenous intervals in a menu cost model. They use the model as a laboratory to analyze regressions of individual price changes on information that pre-dates their last price adjustment. Comparing regressions based on model-generated data with regressions based on BLS micro price data, they conclude that individual prices react to news prior to the last adjustment date, suggesting the presence of information frictions. Bonomo and Garcia (2001) solve for the optimal pricing policy in the presence of menu costs and exogenously infrequent information. They then aggregate such policies to study their macroeconomic implications.

The paper that is closest to ours is Knotek (2010), who uses indirect inference methods to estimate a menu cost model where information about macroeconomic shocks arrives infrequently, as in Mankiw and Reis (2002). He uses both macroeconomic data and price-setting statistics to estimate the model, and concludes that both menu costs and infrequent information are essential to fit the data. Our model differs from his because firms face both adjustment and information costs, and design the optimal pricing policy taking both frictions into account. In addition, our model is estimated with price-setting statistics only, and hence its aggregate implications do not arise from the need to fit macroeconomic data.

The rest of the paper is as follows. In Section 2.2 we present our general framework for solving price-setting problems with adjustment and/or information frictions. We then apply the framework to solve the model with those two frictions and partial information. Section 3 presents the data, estimation method and estimation results. In Section 4, we analyze the aggregate effects of monetary shocks in the estimated economies. The last section concludes.

[11] Alvarez, Lippi and Paciello (2018) study the effects of monetary shocks in calibrated versions of models with menu costs only, observation costs only, and both observation and menu costs, as in Alvarez, Lippi and Paciello (2011).

[12] Gertler and Leahy (2008), Woodford (2009), Nakamura and Steinsson (2010), and Midrigan (2011) are a few examples.

2 A model with menu costs and partially costly information

2.1 Informal description

We develop a model of price setting with costly price adjustments and partially costly information. In the absence of these frictions, each firm would set its price equal to its instantaneous profit-maximizing price – the so-called *frictionless optimal price.* In Appendix A, we present a simple general equilibrium model that yields an expression for the (logarithm of the) frictionless optimal price for a firm as the sum of two components – a common (nominal aggregate demand) and an idiosyncratic (productivity) component. Information about one of these components is continuously and freely available, and can be factored into price-setting decisions somewhat costlessly. Gathering and processing information about the other component entails a lump-sum *information cost* (F). At this point we remain agnostic about which component is costly. Hence, we refer to them as the "free" and "costly" components. Later, when we estimate the model, we entertain the two alternatives – namely, costly aggregate information and free idiosyncratic information, and vice-versa – and let the data inform us which component is better modeled as costly. The loss in profits per unit of time from charging a sub-optimal price is increasing in its distance from the frictionless optimal price. Price changes entail a lump-sum *menu cost* (K).

Intuitively, because of the lump-sum nature of the information cost, a firm will choose to gather and process information about the costly component of its frictionless optimal price only at certain dates, which we refer to as *information dates.* In between information dates, decisions about whether to incur the menu cost to change its price will have to be conditioned on the firm's best estimate of its frictionless optimal price, given its (partial) information. These two possible choices – changing prices or gathering information – imply an optimal inaction region, which we describe heuristically before spelling out the model in detail and explaining how we solve this pricing problem.

Let z_t denote the difference between the firm's price and the best estimate of its frictionless optimal price, given the firm's information.[13] We refer to z_t as the *expected price discrepancy.* Upon incurring the menu cost K, the firm can choose a new price, and will do so in order to set the expected price discrepancy z_t to an optimal level, which we denote by c_t. For a given information set, adjustment is only worthwhile if the expected price discrepancy is "large enough" to justify incurring the menu cost. This implies that at each point in time there are bounds l_t and u_t on the expected price discrepancy such that the firm will increase its price whenever z_t is less than l_t, and decrease its price whenever z_t exceeds u_t. The assumptions about the process for the frictionless optimal price that we specify subsequently imply that the policy functions $\{l_t, c_t, u_t\}$ do not depend on calendar time per se, but

[13] Under full information, this variable corresponds to the standard state variable used to solve menu-cost models (e.g., Dixit 1991).

only on the time elapsed since the last information date, denoted τ (i.e., the last time the firm paid the information cost F to gather full information about its frictionless optimal price). We thus write $\{l(\tau), c(\tau), u(\tau)\}$, and refer to $\{l(\tau), u(\tau)\}$ as the *adjustment bounds* of the inaction region.

Turning to the information decision, upon incurring the information cost F the firm learns the history of innovations to the costly component of its frictionless optimal price. This amounts to a shock to the best estimate of the price discrepancy that the firm held just prior to gathering information. At any given time t, gathering information is only worthwhile if the cost of not observing the innovations that have occurred since the last information date exceeds the information cost F. As long as the uncertainty associated with the best guess of the frictionless optimal price increases with the time elapsed since the last information date, the cost of not observing the underlying innovations will increase over time. Hence, for each expected price discrepancy z, the optimal policy specifies a time elapsed since the last information date ($\tau^*(z)$) that triggers information gathering. We refer to $\tau^*(z)$ as the *information border* of the inaction region. In practice, because firms have the options to adjust and to become informed, they will always find themselves within the inaction region. [14] Figure 1 in the Introduction shows the optimal policy for illustrative parameter values. We now turn to the specifics of the model and the solution method.

2.2 The model

We first solve the optimal price-setting problem of a single firm, and later estimate a model of an economy populated by a large number of such firms. Hence, to avoid cluttering notation, we omit individual firms' subscripts and denote a firm's frictionless optimal price by p_t^*. We assume p_t^* evolves according to:

$$dp_t^* = \mu dt - \sigma_f dW_{f,t} - \sigma_c dW_{c,t}, \tag{1}$$

where $W_{f,t}$ and $W_{c,t}$ are independent standard Wiener processes. Information about $W_{f,t}$ is continuously and freely available, and costless to process. In contrast, gathering and processing information about $W_{c,t}$ is costly. Any discrepancy between a firm's actual price p_t, and its frictionless optimal price p_t^* entails an instantaneous flow "cost" in the form of foregone profits. As we show in Appendix B, after a normalization, these *discrepancy costs* can be taken as being approximately equal to the square of the price discrepancy: $(p_t - p_t^*)^2$. The objective of firms is to minimize the present discounted value of expected total costs, which comprise flow discrepancy costs and lump-sum adjustment and information costs.[15]

[14] This is true as long as the underlying process for the frictionless optimal price is continuous (which we will assume). The only exceptions are information dates, on which the firm may learn that its price discrepancy is outside the adjustment bounds for $\tau = 0$. But in those cases the firm will choose to incur the menu cost and adjust immediately to $c(0)$.

[15] Appendix C formalizes the firm's dynamic optimization problem.

Under partial information about p_t^*, in order to evaluate the expected flow cost due to price discrepancies the firm must form a probabilistic assessment of p_t^* given its information. We can decompose the expected discrepancy cost at time t as:

$$E_t(p_t - p_t^*)^2 = (p_t - E_t p_t^*)^2 + Var_t\,(p_t^*)\,, \tag{2}$$

where Var_t denotes the conditional variance given time t information. Because the firm has continuous access to partial information about p_t^*, the conditional variance $Var_t\,(p_t^*)$ refers to the component of the frictionless optimal price that is only observed at a cost – that is, to the $\sigma_c dW_{c,t}$ component in equation (1).

The first term in the right-hand side of (2) represents the flow cost of deviating from the *expected* level of the frictionless optimal price. The second term represents the expected flow cost from not continuously entertaining full information about p_t^*. In the absence of adjustment costs, p_t would be set equal to $E_t p_t^*$, reducing the first part of the expected discrepancy cost to zero. Otherwise the firm must optimally solve the trade-off between letting p_t drift away from $E_t p_t^*$, and paying the cost to adjust based on partial information. As for the second term in (2), it is zero when information can be fully and continuously incorporated into the pricing decision at no cost, as in standard menu-cost models. If information gathering and processing is costly, the firm can reduce that second term at the expense of incurring the information cost.

For any given time t for which the last information date was at time $t_0 < t$, recall that $\tau \equiv t - t_0$ denotes the time elapsed since the last information date. Also, recall that $z_t \equiv p_t - E_t p_t^*$ denotes the expected price discrepancy at time t. With these definitions, we can rewrite the firm's expected discrepancy cost (2) as a function of τ and z:

$$E_t(p_t - p_t^*)^2 = f\,(z_t, \tau) \equiv z_t^2 + \sigma_c^2 \tau. \tag{3}$$

We can then write the value function at a time t – the optimized value of the firm's dynamic cost-minimization problem described above, and formalized in Appendix C – in terms of the two state variables z_t and τ. In the inaction region (i.e., absent any price change and/or costly information gathering), the value function, $V\,(z_t, \tau)$, obeys the following Bellman equation:

$$V\,(z_t, \tau) = \left(z_t^2 + \sigma_c^2 \tau\right) dt + e^{-\rho dt} E_t V(z_{t+dt}, \tau + dt). \tag{4}$$

Between information dates and in the absence of price adjustments (i.e., in the inaction region), z_t changes continuously because of both the drift μ and the free component of the frictionless optimal price, $W_{f,t}$:

$$dz_t = -\mu dt + \sigma_f dW_{f,t}. \tag{5}$$

Taking into account the process for the expected price discrepancy z (equation 5) and applying Ito's lemma, the evolution of the value function in the inaction region can be described by the following partial differential equation:

$$\frac{1}{2}\sigma_f^2 V_{zz}(z,\tau) - V_z(z,\tau)\mu + V_\tau(z,\tau) - \rho V(z,\tau) + z^2 + \sigma_c^2\tau = 0. \tag{6}$$

Solving for the value function requires the specification of boundary conditions, which are dictated by the price adjustment and information gathering decisions. We analyze each of these decisions in turn.

2.2.1 The adjustment decision

Because adjustment costs are lump-sum, price adjustments at any point in time minimize the value function at a given τ. Hence, the target point $c(\tau)$ when the time elapsed is τ satisfies:

$$c(\tau) = \arg\min_z V(z,\tau). \tag{7}$$

Since firms always have the option to incur the adjustment cost K and reset the expected discrepancy to $c(\tau)$, optimality implies that the value function must always satisfy:

$$V(z,\tau) \le K + V(c(\tau),\tau). \tag{8}$$

The bounds that define the adjustment inaction region, $l(\tau), u(\tau)$, are functions of τ that imply indifference between adjusting and not adjusting. Hence, they satisfy the value-matching conditions that obtain when (8) holds with equality:

$$\begin{aligned} V(l(\tau),\tau) &= K + V(c(\tau),\tau), \\ V(u(\tau),\tau) &= K + V(c(\tau),\tau). \end{aligned} \tag{9}$$

2.2.2 The information decision

Firms always have the option to incur the information cost F to gather and process information about the costly $W_{c,t}$ component. Upon doing so, they learn the realization of $W_{c,t}$ – or, equivalently, their frictionless optimal price p_t^* – and the time elapsed since the last information date, τ, is reset to zero. The decision of whether or not to get informed at any given point in time involves comparing the current value function with a "lottery" that will yield a new value function after the realization of $W_{c,t}$ is learned. Taking the lottery requires paying the information cost F. Optimality requires that the current value function does not exceed the sum of the information cost and the expected value of

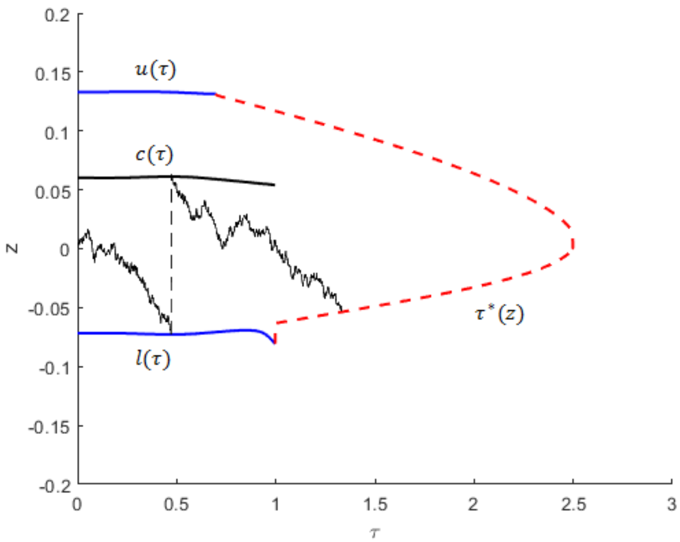

Figure 1: Optimal pricing policy under partially costly information (illustrative parameter values).

the lottery, that is:

$$V\left(z,\tau\right) \leq F + E\left[V\left(z + \sigma_c\sqrt{\tau}\varepsilon, 0\right)\right], \tag{10}$$

where ε is a standard normal random variable.

Points in the state space in which the firm is indifferent between getting informed and continuing with outdated information, for which (10) holds with equality, define the information boundary $(z, \tau^*(z))$. Thus, on information dates the expected price discrepancy receives a shock with distribution $N(0, \sigma_c^2\tau^*(z))$, and τ is reset to zero, yielding the following "informational value-matching condition":

$$V\left(z, \tau^*\left(z\right)\right) = F + E\left[V\left(z + \sigma_c\sqrt{\tau^*\left(z\right)}\varepsilon, 0\right)\right]. \tag{11}$$

2.3 The optimal pricing rule

We solve this pricing problem using a finite-difference method, which we describe in Appendix D. Figure 1, which we reproduce here for convenience, depicts the optimal policy for the following illustrative parameter values: $\mu = 0.2$, $\sigma_c = \sigma_f = 0.05$, $K = F = 0.002$, $\rho = 0.025$. Under optimal policy, the firm uses $W_{f,t}$-information between information dates and adjusts the expected price discrepancy to $c(\tau)$ whenever it hits the $l(\tau)$ or $u(\tau)$ boundaries of the inaction region. Whenever the expected price discrepancy hits the $\tau^*(z)$ boundary of the inaction region, the firm incurs the lump-sum cost F to gather information.

Partially informed price increases have size $c(\tau) - l(\tau)$, while partially informed price decreases have size $u(\tau) - c(\tau)$. In principle, those adjustment sizes depend on the time elapsed since the last information date. Fully informed adjustments are potentially much more variable in size, with lower bounds given by $c(0) - l(0)$ for price increases, and $u(0) - c(0)$ for price decreases.

In the sample path realization for the expected price discrepancy (z_t) that we depict in Figure 1, there is one partially-informed adjustments before the firm decides to incur the cost to entertain information about $W_{c,t}$. At that point, the time-elapsed variable τ is reset to zero, and the firm learns the cumulative innovation $W_{c,t} - W_{c,t-\tau^*(z)} \sim N\left(0, \sigma_c^2 \tau^*(z)\right)$. Then, the firm decides whether or not to incur the menu cost and change its price, depending on whether the price discrepancy is inside or outside the inaction region defined by $(l(0), u(0))$.

For small τ, the limits of the inaction region are dictated by the adjustment bounds, whereas for large τ they are defined by the information border. When information about $W_{c,t}$ is not yet too outdated, partial information about p_t^* might lead to a large enough expected price discrepancy, inducing the firm to make a partially-informed price adjustment. After some point (corresponding to $\tau \approx 1$ in Figure 1), making partially-informed adjustments is no longer optimal. The reason is that by that time the firm's information set has "depreciated" enough (due to the accumulation of unobserved innovations to $W_{c,t}$). Thus, a given expected discrepancy that might have triggered a partially-informed adjustment early on, will now trigger information gathering instead.

An interesting implication of optimal pricing behavior under adjustment and information costs, which can be glimpsed from the previous description, is that it is never optimal to make a partially-informed adjustment just prior to an information date. Rather than incurring the menu cost to make such an adjustment and then immediately incurring the information cost, it is always better to reverse the order of these actions and keep the option to adjust, to be exercized or not depending on the new information.[16]

2.4 A particular case with costly information only

In this subsection, we use our framework to tackle a particular case of the problem just described, in which all information is costly. This case is analyzed in Alvarez et al. (2011). They show that, when inflation is low enough, it is never optimal for firms to adjust prices without information. They fully characterize and solve the model under this assumption. Our solution method allows us to solve for the optimal pricing problem also when uninformed adjustments are optimal. This is the case if inflation is high enough or, for a given level of trend inflation $(\mu \neq 0)$, if adjustment costs are small enough relative to information costs.

We solve this problem by fixing $\sigma_f = 0$ and relabeling $\sigma = \sigma_c$ in the setting of Section 2. Thus,

[16] In Bonomo, Carvalho, and Garcia (2011) we show that the principle that it is never optimal to make an adjustment just prior to the arrival of relevant information is rather general. In particular, we illustrate this point in a setting in which information dates are exogenous and known to firms. In that case, the width of the inaction region increases to infinity just before an information date.

the differential equation which characterizes the evolution of the value function inside the inaction region becomes:

$$-\mu V_z(z,\tau) + V_\tau(z,\tau) - \rho V(z,\tau) + z^2 + \sigma^2\tau = 0.$$

The conditions related to the adjustment decision (7, 8 and 9) and information decision (10,11) remain the same.

Figure 2 illustrates the optimal pricing policy under adjustment and information gathering/processing costs when all information is costly, for the following illustrative parameter values: $\mu = 0.2$, $\sigma = 0.05\sqrt{2}$, $K = F = 0.002$, $\rho = 0.025$. The solid (blue) lines $l(\tau), u(\tau)$ are the adjustment bounds of the inaction region that trigger uninformed adjustments, while the dashed (red) line $\tau^*(z)$ is the information border that triggers information gathering. We illustrate the optimal pricing behavior with a sample path. Initially, z_t is close to zero. Due to the high enough drift μ, the expected price discrepancy hits the lower boundary $l(\tau)$, leading to an uninformed adjustment to $c(\tau)$. After that, the expected discrepancy drifts down until it touches the information boundary $\tau^*(z)$ at a point where $z \approx -0.05$ and $\tau \approx 0.9$. At that point the firm incorporates information into the pricing decision, as the expected discrepancy receives a shock with distribution $N\left(0, \sigma^2 \times 0.9^2\right)$. The time-elapsed variable τ is reset to zero, and the firm decides whether or not to pay the menu cost and change its price, depending on whether the just-learned price discrepancy is inside or outside the inaction region on information dates, defined by $(l(0), u(0))$. Once again, notice that it is never optimal to make an uninformed price adjustment just prior to an information date.

Note that in this parametrization the total variance of the firm's frictionless optimal price is the same as the one used to illustrate the optimal policy under partial information (Figure 1), and all other parameter values are equal. The difference is that here all information is costly. Hence, for a given expected price discrepancy, firms choose not to wait as long before paying the information cost to get informed.

3 Estimating the model

We use price-setting statistics computed from micro data from the U.S. Bureau of Labor Statistics to estimate the main parameters of the model by a Simulated Method of Moments (SMM). Simulation-based methods are required, since the relationship between the model parameters and observable statistics is highly non-linear and complex. We estimate the partial information model under two alternative hypotheses for the costly and free components. In one version, the price-setter can costlessly incorporate all the information about aggregate shocks into her pricing decision, while observing/processing idiosyncratic shocks is costly. The other version of the model reverts these assumptions. Finally, we also estimate the model in which all information is costly.

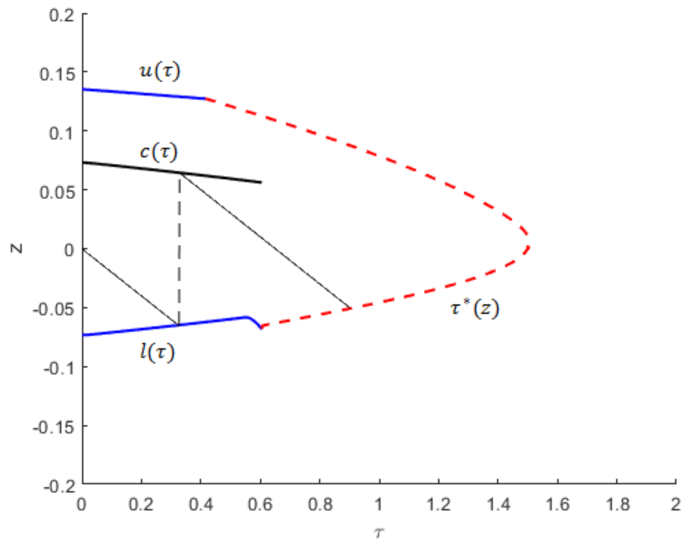

Figure 2: Optimal pricing policy in the model with no partial information (illustrative parameter values)

3.1 Estimation method

Estimation by SMM involves minimizing the distance between data moments and moments obtained through simulation of the model. This requires that, for each set of parameter values, we solve for the optimal pricing rule, and then simulate an economy with a large number of firms.

Let $\Phi \in \mathbb{R}^p$ be a vector of model parameters to be estimated, $\Psi_{data} \in \mathbb{R}^q$ $(q \geq p)$ a vector of data moments and $\Psi_{sim}(\Phi) \in \mathbb{R}^q$ a vector of the corresponding moments calculated from the model's simulated data. The estimator $\hat{\Phi}$ is obtained by solving the following minimization problem:

$$\min_{\Phi} \ \left(\Psi_{data} - \Psi_{sim}(\Phi)\right)' W \left(\Psi_{data} - \Psi_{sim}(\Phi)\right), \tag{12}$$

where W is the optimal $q \times q$ weighting matrix, given by the inverse of the variance-covariance matrix of the sample moments Ψ_{data} adjusted for simulation error.[17]

In accordance with the model detailed in Appendix A, all firms' frictionless optimal prices are subject to nominal aggregate shocks, and idiosyncratic shocks that are independent across firms. Estimating the version of the partial-information model in which idiosyncractic shocks are free and aggregate shocks are costly amounts to setting $\sigma_c = \sigma_a$ and $\sigma_f = \sigma_i$. The other version of the partial-information model obtains when $\sigma_c = \sigma_i$ and $\sigma_f = \sigma_a$. In the model in which all information is costly, $\sigma_f = 0$ and $\sigma_c = \sqrt{\sigma_i^2 + \sigma_a^2}$. In this model, aggregate and idiosyncratic shocks still have potentially

[17] We compute the artificial moments $\Psi_{sim}(\Phi)$ based on 50 artificial samples of the same time span as the actual data. Hence, the variance-covariance matrix underlying W equals 1.02 times the variance-covariance matrix of the sample moments in the data (see, e.g., DeJong and Dave 2011).

different effects on price-setting statistics, since only aggregate shocks generate time variation in aggregate statistics.[18]

We calibrate the average inflation rate to match the sample average over the period during which the micro data used to computed the empirical price-setting statistics were sampled. To reduce the number of parameters and save on computational time, we also set $\rho = 2.5\%$. Hence, for each of the three models, this leaves us with four parameters to be estimated: σ_a, σ_i, K, F.

3.2 Moments used in the estimation

We use monthly time series of price-setting statistics from February 1988 to January 2005 (204 months), constructed by Klenow and Kryvtsov (2008).[19] Those statistics were computed from individual price changes (excluding sales) in the micro data underlying the US Bureau of Labor Stastics CPI for the top urban areas of the United States. The sample period dictates the calibration of the average inflation rate ($\mu = 3.3\%$ per year).

The price-setting statistics used in the estimation are: i) frequency of positive price changes; ii) frequency of negative price changes; iii) mean size of price changes; iv) median size of price changes; and v) the frequency of positive price changes squared. We now provide some intuition for why matching these moments identifies our parameters of interest.[20]

The adjustment cost parameter K clearly affects the frequency of both positive and negative price changes. With a positive drift in the frictionless optimal price process ($\mu > 0$), positive adjustments are more frequent than negative adjustments. However, because the distributions of shocks are symmetric, the higher the variance of idiosyncratic shocks is, the smaller that asymmetry between positive and negative price adjustments becomes. Fully informed price adjustments tend to be larger and more variable than partially informed (or uninformed) adjustments.[21] An increase in information costs (F) tends to induce less frequent information collection, and a more dispersed distribution of informed price changes. This increases the mean size of price changes relative to its median. Finally, the incidence of aggregate shocks can produce meaningful time variation in the frequency of price changes – both positive and negative. Hence our use of the frequency of positive price changes squared as an estimation target.

[18] This is only strictly true in the limit with an infinite number of firms. In simulations with a finite number of firms, idiosyncratic shocks do produce (sampling) variation in aggregate price-setting statistics. Hence the importance of simulating the model with a number of firms that is comparable to the number of quote lines in the micro data from which the empirical moments are obtained.

[19] We thank Oleksiy Kryvtsov for sharing the time series with us.

[20] We used Monte Carlo simulations to confirm that our SMM estimation with these moments recovers the parameters of the model used to generate the artificial samples.

[21] The size of partially informed (or uninformed) adjustments is pinned down by the distance between the adjustment bounds $(l(\tau), u(\tau))$ and the target point $c(\tau)$. These distances do not vary that much with τ (see, e.g., Figure 1). In contrast, fully informed price increase (decreses) have minimum size given by $c(0) - l(0)$ $(u(0) - c(0))$, but no upper bound.

We simulate time series of 204 months for 50 economies with 13,500 firms each, and average simulated moments across the 50 artificial economies. The number of firms in each economy is similar to the number of quote lines (13,000-14,000) underlying the price-setting statistics that we use in estimation (see Klenow and Kryvtsov 2008). In each economy, the initial distribution of firms' state variables is drawn from the ergodic distribution.[22]

3.3 Results

Table 1 compares targeted data moments with those produced by the different estimated models. Informal inspection of the moments produced by the model with no partial information and by the model in which aggregate information is free suggests small differences compared to the data. The estimated partial information model with costly aggregate information, however, matches the data almost perfectly (last row of Table 1). A formal test of overidentification restrictions shows that the latter is the only model that is not rejected by the data (last two columns of Table 1).

Table 1: Target and simulated statistics

Statistic	$fr\Delta p^+$	$fr\Delta p^-$	$avg\,\lvert\Delta p\rvert$	$median\,\lvert\Delta p\rvert$	$(fr\Delta p^+)^2$	J-stat.	p-val.
Data	0.1503	0.1152	0.0900	0.0710	0.0233	-	-
No Partial	0.1481	0.1196	0.0873	0.0772	0.0224	101.7	0.0000
Partial $\sigma_a = \sigma_f$	0.1460	0.1207	0.0875	0.0761	0.0233	88.84	0.0000
Partial $\sigma_a = \sigma_c$	0.1500	0.1155	0.0901	0.0711	0.0232	0.0284	0.8661

Table 2 reports estimates of the model parameters, which show some important differences. The model with partial information in which aggregate information is free and the model with no partial information produce relatively similar estimates. In contrast, the model with costly aggregate information leads to a higher estimate for the variance of aggregate shocks coupled with significantly larger information costs.[23] Estimated menu costs are similar across the three models.

Figure 3 shows the optimal price-setting policies implied by the three estimated models. Notice that the model with partial information in which aggregate information is free and the model with costly information only produce quite similar pricing policies. In contrast, the model with costly aggregate information leads to a very different optimal policy, which encompasses the ones produced by the other two estimated models. Its adjustment bounds are wider than in the other two models, and information collection happens much less frequently.

[22] To obtain the ergodic distribution, we solve its associated Kolmogorov Forward Equation, following the numerical algorithm in Kaplan and Moll (2016). In earlier versions of the paper we obtained the ergodic distribution by simulation.

[23] This results might reflect the absence of sectoral shocks in our model. In the data, these shocks are a potential source of sizable variation in price-setting statistics, which in our one-sector model can only be captured by the aggregate shock. This possibility is worth investigating in future research.

Table 2: Estimated parameters

	$\boldsymbol{\sigma}_a$	σ_i	F	K
No Partial Info	0.0277	0.1723	0.00059	0.00022
standard deviation	0.0036	0.0033	0.00023	0.00005
t-statistic	7.73	51.75	2.59	4.16
Partial Info ($\sigma_a = \sigma_f$)	0.0258	0.1782	0.00091	0.00013
standard deviation	0.0023	0.0044	0.00023	0.00005
t-statistic	11.18	40.71	4.03	2.64
Partial Info ($\sigma_a = \sigma_c$)	0.1492	0.1280	0.0188	0.00030
standard deviation	0.0081	0.0015	0.0313	0.00001
t-statistic	18.36	83.36	6.03	49.91

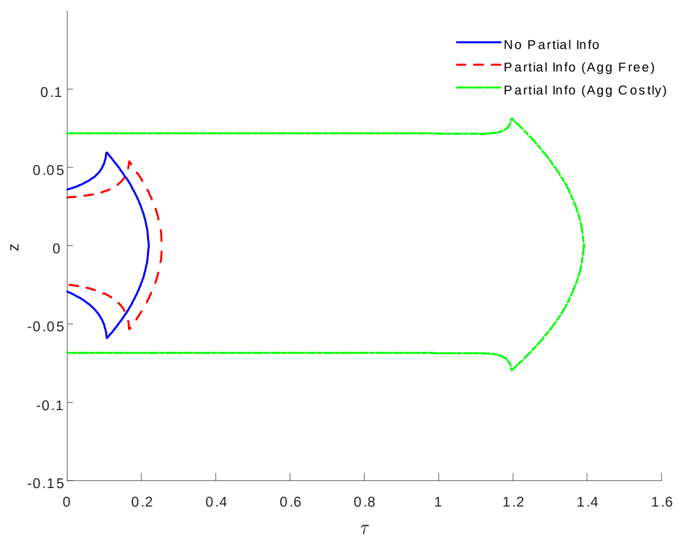

Figure 3: Optimal policies implied by the estimated models.

These differences are reflected in different price-setting statistics, reported in Table 3. Information collection occurs much more often in the model with no partial information and the model with free aggregate information (more than four times per year in both models, vs. less than once per year in the model with costly aggregate information). Another striking difference is that in those first two models almost all adjustments are fully informed. In sharp contrast, more than 80% of price adjustments in the model with costly aggregate information are based only on information about idiosyncratic shocks. In all models, when costly information is collected, price adjustments happen with a similar frequency – between $69-76\%$ of the time.

The intuion for these results is clear. The estimated model with costly idiosyncratic information features small aggregate shocks. As a result, they rarely lead to partially informed adjustments. Hence, this partial information model produces results that are similar to those in which the (small)

Table 3: Additional price-setting statistics implied by estimated models

Baseline	No Partial Info	Partial Info ($\sigma_a = \sigma_f$)	Partial Info ($\sigma_a = \sigma_c$)
K	0.00022	0.00013	0.0003
F	0.00059	0.00091	0.0188
Price adjustments per year	3.19	3.20	3.27
fully informed adjustments	100%	95%	16%
adjustments w/o (full) info	0%	5%	84%
Information gatherings per year	4.62	4.01	0.76
resulting in immediate adj.	69%	76%	69%

aggregate shocks are also costly to observe, as in the model with costly information only. Things are different in the model with costly aggregate information. Since the flow profit loss is a convex function of the price discrepancy, the fact that the firm can react to news about its volatile idiosyncratic component reduces the incentives to gather costly information frequently. Hence, for a given information gathering cost F, the firm would choose to sample less frequently. In addition, the estimated information cost in the model with costly aggregate information is substantially larger than in the other two models, further contributing to less frequent information gathering.

In terms of economic magnitudes, the estimated parameters imply plausible adjustment and information costs when compared to available evidence. Zbaracki et al. (2004) in their case study of a large industrial manufacturer identify three components associated with the costs of pricing decisions and price changes: "physical costs" (menu costs), managerial costs (information gathering, decision-making, and communication costs), and customer costs (communication and negotiation costs). These costs amount to, respectively, 0.7%, 4.6%, and 14.7% of the firm's net margin, adding up to 20% of net margin. For the model in which aggregate information is costly, the estimated parameters imply that firms spend, on average, 3% of annual steady state profits on information and adjustment costs. Menu costs account for 0.2% of annual profits, whereas information costs account for 2.8% of profits. The ratio of information expenses to expenses incurred for posting new prices is around 14:1. This compares with a ratio of managerial costs to physical costs of 6.5:1 in Zbaracki et al. (2004). If one uses the ratio of managerial plus customer costs to physical costs, the result is 27:1.

The estimated models have sharply different implications for an untargeted statistic – namely, the distribution of the duration of price spells. In particular, the models with no partial information and with free aggregate information produce counterfactual distributions, with "gaps" that do not appear in the empirical distribution reported by Klenow and Kryvtsov (2008). In contrast, the model with costly aggregate information leads to a distribution that resembles that empirical distribution that Klenow and Kryvtsov report (see Figures 4-7, below).

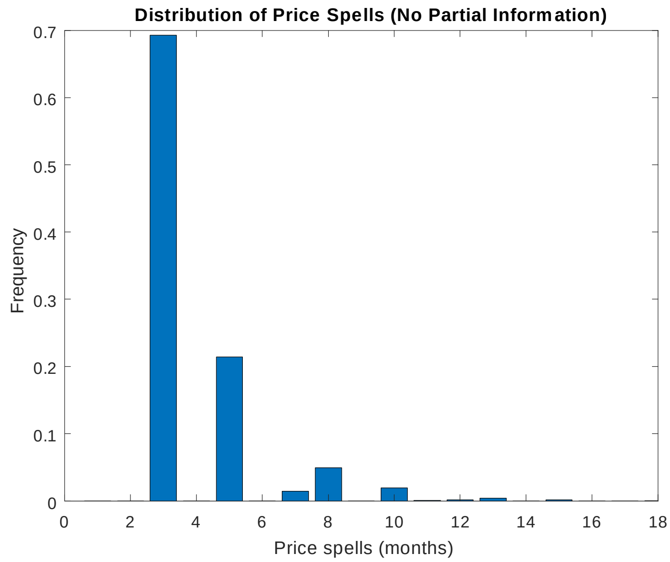

Figure 4: No partial information.

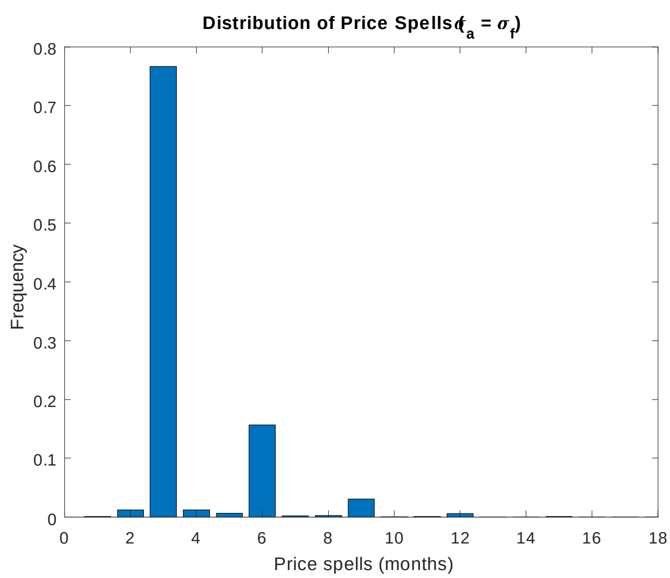

Figure 5: Partial information model with free aggregate information.

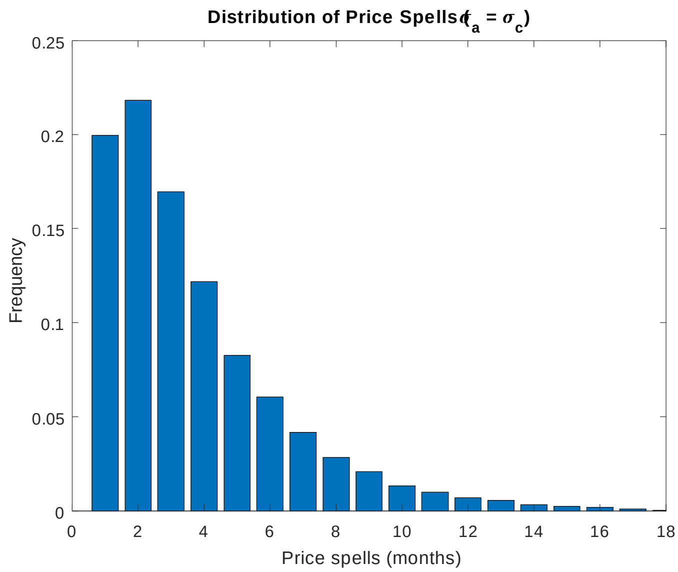

Figure 6: Partial information model with costly aggregate information.

880 *QUARTERLY JOURNAL OF ECONOMICS*

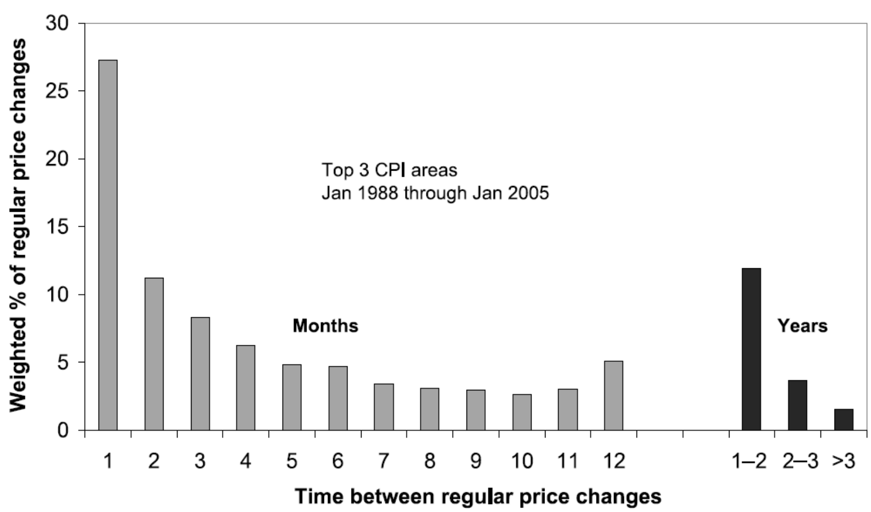

FIGURE IV
Weighted Distribution of Times between Regular Price Changes

Figure 7: Empirical distribution of durations of (regular) price spells from Klenow and Kryvtsov (2008).

4 Real effects of nominal shocks

In our model, the (log of the) frictionless optimal price equals the sum of the (log of) nominal aggregate demand and the (log of) idiosyncractic productivity. We assume that nominal aggregate demand $\mathcal{M}_t$ follows a geometric Brownian motion with trend growth μ. To study real effects of monetary shocks, however, we follow the usual practice of analyzing the effects of a one-time monetary shock, starting from the ergodic steady state to which the economy converges after a sufficiently long spell without aggregate shocks. We shock the economy with a one-time innovation that makes nominal aggregate demand jump from $\mathcal{M}_0$ to $\mathcal{M}_0 (1 + \zeta)$ at time zero. After the jump, nominal aggregate demand resumes its trend, given by μ.

For each estimated model, we simulate two million firms, starting from the ergodic distribution. We then shock nominal aggregate demand as described above and aggregate individual prices to obtain the average price level. Real output (in logs) is given by:

$$y_t = \log \mathcal{M}_t - p_t.$$

Results for real output are shown in Figure 8 for a shock of size $\zeta = 1\%$, as a proportion of the shock. Despite the fact that all three models display similar frequency of price adjustments, they display very different aggregate dynamics. In all models, output increases immediately, and reverts back to the original level, but at different rates. In the models with no partial information and with free aggregate information, output is back near its original level after approximately three months. In contrast, in the model with costly aggregate information it takes more than a year for real effects to dissipate. The intuition behind this result is clear: while in the first two models firms react to aggregate information very frequently, in the model with costly aggregate information, most of the time firms react only to idiosyncratic information, and take account of aggregate shocks only infrequently. Notice that these results obtain in a model without any strategic complementarities in price setting.

To explore the roles of menu costs and information costs in producing the results reported in Figure 8, Figure 9 shows how changing K and F affects the persistence of monetary non-neutrality in the three models. Varying the menu cost – halving or doubling – around the estimated values has hardly any effect on aggregate dynamics. In contrast, varying information costs leads to meaningful differences in monetary non-neutrality – a result that is in line with the findings of Alvarez et al. (2018). This difference is much more pronouced in the model with costly aggregate information.

In principle, the model with no partial information and the model in which aggregate information is free are also capable of generating sizable monetary non-neutrality – this only requires larger price-setting frictions. However, when disciplined by the microdata, these two models imply a small degree

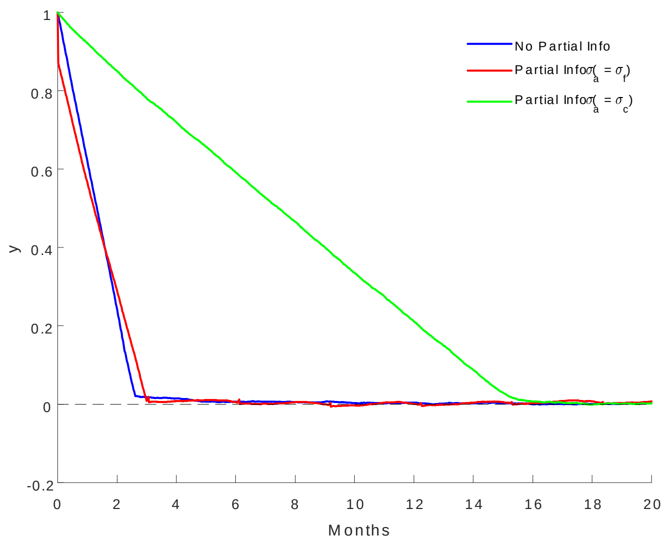

Figure 8: Real effects of monetary shocks in estimated models.

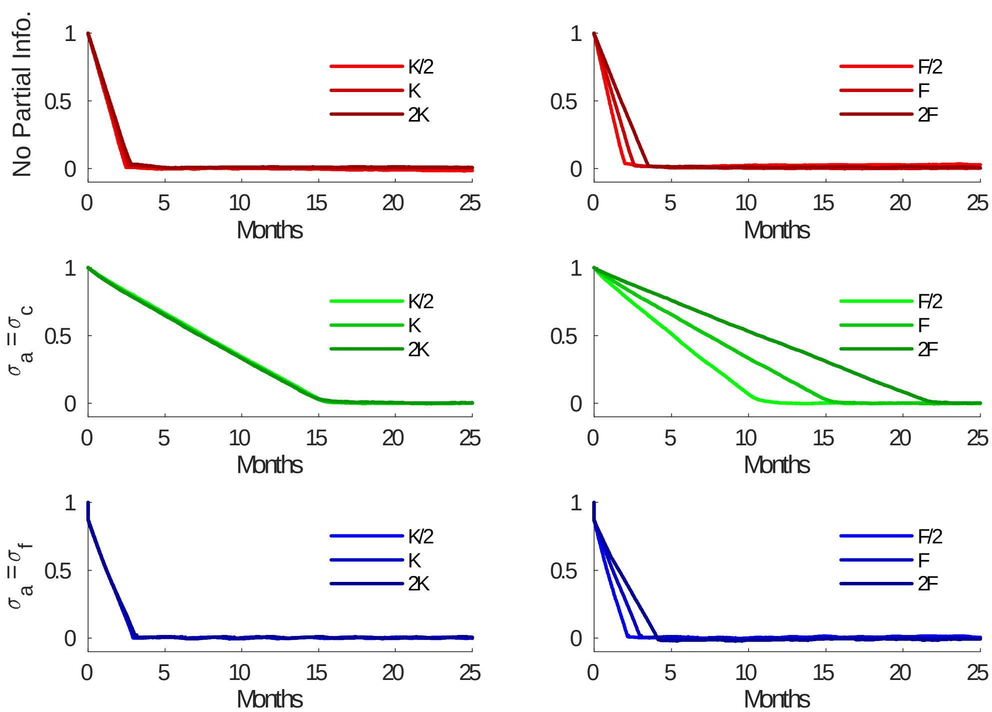

Figure 9: Real effects of monetary shocks: varying K and F.

of non-neutrality. The main difference between the estimated frictions in these two estimated models and in the one in which aggregate information is costly is the size of the information cost, which is significantly larger in the latter model. Hence, our findings are in accordance with the price-setting literature, which shows that information frictions are a powerful source of monetary non-neutrality.

5 Conclusion

We propose a price-setting model which helps to reconcile micro evidence of relatively frequent price adjustments with persistent real effects of monetary shocks. In our model, both price adjustment and the gathering of some types of information are costly, requiring the payment of lump-sum costs. Additionally, another relevant part of information on firms' marginal costs is continuously available, and can be factored into pricing decisions costlessly. When we assume that aggregate information is costly, and idiosyncratic marginal cost information is free, our model is able to match individual price-setting statistics for the U.S. and, at the same time, produce persistent monetary non-neutrality. Continuously available information about idiosyncratic shocks leads to low frequency of aggregate information collection, which turns out to be crucial for the persistence of real effects of monetary shocks.

References

[1] Almeida, H. and M. Bonomo (2002), "Optimal State-Dependent Rules, Credibility and Inflation Inertia," *Journal of Monetary Economics* 49: 1317-1336.

[2] Alvarez, F., L. Guiso, and F. Lippi (2012), "Durable Consumption and Asset Management with Transaction and Observation Costs," *American Economic Review.*102 (5): 2272-2300.

[3] Alvarez, F., F. Lippi, and L. Paciello (2011), "Optimal Price Setting with Observation and Menu Costs," *Quarterly Journal of Economics* 126: 1909-1960.

[4] ____________ (2016), "Monetary shocks in models with inattentive producers," *Review of Economic Studies* 83: 421-459.

[5] ____________ (2018), "Monetary Shocks with Observation and Menu Costs," *Journal of the European Economic Association* (https://doi.org/10.1093/jeea/jvx013).

[6] Bils, M. and P. Klenow (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy* 112: 947-985.

[7] Bonomo, M. and C. Carvalho (2004), "Endogenous Time-Dependent Rules and Inflation Inertia," *Journal of Money, Credit and Banking* 36: 1015-1041.

[8] ____________ (2010), "Imperfectly-Credible Disinflation under Endogenous Time-Dependent Pricing," *Journal of Money, Credit and Banking* 42: 799-831.

[9] Bonomo, M., C. Carvalho, and R. Garcia (2011), "State-dependent Pricing under Infrequent Information: A Unified Framework," Federal Reserve Bank of New York Staff Reports #455.

[10] ____________ (2001), "The Macroeconomic Effects of Infrequent Information with Adjustment Costs," *Canadian Journal of Economics* 34: 18-35.

[11] Caballero, R. (1989), "Time Dependent Rules, Aggregate Stickiness and Information Externalities," Working Paper no. 428, Columbia University.

[12] Gertler, M. and J. Leahy (2008), "A Phillips Curve with an Ss Foundation," *Journal of Political Economy* 116: 533-572.

[13] Golosov, M. and R. Lucas (2007),"Menu Costs and Phillips Curves," *Journal of Political Economy* 115: 171-199.

[14] Gorodnichenko, Y. (2008), "Endogenous Information, Menu Costs and Inflation Persistence," NBER Working Paper No. 14184.

[15] DeJong, D. and C. Dave (2011), *Structural Macroeconometrics*, Princeton University Press, second edition.

[16] Dixit, A. (1993), *The Art of Smooth Pasting*, Routledge.

[17] Kaplan, G. and B. Moll (2016), "Liquid and Illiquid Assets with Fixed Adjustment Costs," mimeo.

[18] King, R. and A. Wolman (1999), "What Should the Monetary Authority Do When Prices Are Sticky?," in J. Taylor (org): *Monetary Policy Rules*, University of Chicago Press.

[19] Klenow, P., and O. Kryvtsov (2008), "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *Quarterly Journal of Economics* 123: 863-904

[20] Klenow, P. and J. Willis (2007), "Sticky Information and Sticky Prices," *Journal of Monetary Economics* 54: 79-99.

[21] Knotek, E. (2010), "A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment," *Journal of Money, Credit, and Banking* 42: 1543–64.

[22] Kongsamut, P., S. Rebelo and D. Xie (2001), "Beyond Balanced Growth," *Review of Economic Studies* 68: 869-882.

[23] Maćkowiak, B., and M. Wiederholt (2009), "Optimal Sticky Prices under Rational Inattention," *American Economic Review* 99: 769-803.

[24] Matějka, F. (2015a), "Rigid pricing and rationally inattentive consumer," *Journal of Economic Theory* 158: 656-678.

[25] ______________ (2015b), "Rationally inattentive seller: Sales and discrete pricing," *Review of Economic Studies* 83: 1125–1155.

[26] Mankiw, G., and R. Reis (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics* 117: 1295-1328.

[27] Midrigan, V. (2011), "Menu Costs, Multi-Product Firms and Aggregate Fluctuations," *Econometrica* 79: 1139-1180.

[28] Moscarini, G. (2004), "Limited Information Capacity as a Source of Inertia," *Journal of Economic Dynamics and Control* 28: 2003-2035.

[29] Nakamura, E. and J. Steinsson (2010), "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model," *Quarterly Journal of Economics* 125: 961-1013.

[30] Reis, R. (2006), "Inattentive Producers," *Review of Economic Studies* 73: 793-821.

[31] Stokey, N. (2008), *The Economics of Inaction: Stochastic Control Models with Fixed Costs*, Princeton University Press.

[32] Wilmott, P., J. Dewynne, and S. Howison (1993), *Option pricing: mathematical models and computation*, Oxford Financial Press.

[33] Woodford, M. (2009), "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics* 56, Supplement: S100-S124.

[34] Zbaracki, M., M. Ritson, D. Levy, S. Dutta and M.J., M. Bergen (2004), "Managerial and Customer Dimensions of the Costs of Price Adjustment: Direct Evidence from Industrial Markets," *Review of Economics and Statistics* 86: 514-533.

6 Appendix A

We derive the frictionless optimal price in a simple general equilibrium framework. A representative consumer maximizes expected discounted utility:

$$E_{t_0} \int_{t_0}^{\infty} e^{-\rho(t-t_0)} \left[\log(C_t) - H_t\right] dt,$$

subject to the budget constraints:

$$B_t = B_0 + \int_0^t W_r H_r dr - \int_0^t \left(\int_0^1 P_{ir} C_{ir} di \right) dr + \int_0^t T_r dr + \int_0^t \Lambda_r dQ_r + \int_0^t \Lambda_r dD_r, \text{ for } t \geq 0.$$

Utility is defined over the composite consumption good $C_t \equiv \left[\int_0^1 (C_{it}/A_{it})^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}}$ with $\theta > 1$, where C_{it} is the consumption of variety i, and A_{it} is a relative-preference shock. P_{it} is the price of variety i, H_t is the supply of labor, which commands a wage W_t, B_t is total financial wealth, T_t are total net transfers, including any lump-sum flow transfer from the government, and profits received from the firms owned by the representative consumer. Q_r is the vector of prices of traded assets, D_r is the corresponding vector of cumulative dividend processes, and Λ_r is the trading strategy, which satisfies conditions that preclude Ponzi schemes. The associated consumption price index, P_t, is given by:

$$P_t = \left[\int_0^1 P_{it}^{1-\theta} di \right]^{\frac{1}{1-\theta}}. \tag{13}$$

The demand for an individual variety is:

$$C_{it} = A_{it}^{1-\theta} \left(\frac{P_{it}}{P_t} \right)^{-\theta} C_t. \tag{14}$$

Firms hire labor to produce according to the following production function:

$$Y_{it} = A_{it} H_{it}.$$

Note that we assume that the productivity shock is perfectly correlated with the relative-preference shock in the consumption aggregator. This has precedence in the sticky-price literature (for instance, King and Wolman 1999 and Woodford 2009). Our specific assumption follows Woodford (2009), and aims to produce a tractable profit-maximization problem that can be written as a price-setting "tracking problem" in which the firm only cares about the ratio of the two stochastic processes driving profits, which will be specified below.[24]

The static profit-maximizing price for firm i, P_{it}^* (also referred to as its *frictionless optimal price*),

[24] More generally, assumptions relating preference and technology processes have been used previously in the literature on "balanced growth" in multi-sector models (e.g. Kongsamut et al. 2001).

is given by the usual markup rule:

$$P_{it}^{*} = \frac{\theta}{\theta - 1} \frac{W_t}{A_{it}}. \tag{15}$$

From the representative household's labor supply:

$$\frac{W_t}{P_t} = C_t,$$

which leads to:

$$P_{it}^{*} = \frac{\theta}{\theta - 1} \frac{P_t C_t}{A_{it}}.$$

In logarithms (lowercase variables denote logarithms throughout), this reads:

$$p_{it}^{*} = \log\left(\frac{\theta}{\theta - 1}\right) + \log\left(P_t C_t\right) - \log\left(A_{it}\right).$$

Ignoring the unimportant constant and assuming appropriate exogenous stochastic processes for nominal aggregate demand and for idiosyncratic productivity yields the specifications used throughout the main text.

7 Appendix B

Here we derive the quadratic approximation to the static profit-maximization problem used in the main text. Write real flow profits as:

$$\Pi\left(\frac{P_i}{P}, C, A_i\right) = A_i^{1-\theta} \frac{P_i}{P} \left(\frac{P_i}{P}\right)^{-\theta} C - \frac{W}{P A_i} A_i^{1-\theta} \left(\frac{P_i}{P}\right)^{-\theta} C,$$

where P_i is the price *charged* by firm i. We can use the labor supply equation to express the real wage as a function of aggregate consumption ($\frac{W}{P} = C$), and rewrite the expression for real flow profits as:

$$\Pi\left(\frac{P_i}{P}, C, A_i\right) = A_i^{1-\theta} \left(\frac{P_i}{P}\right)^{1-\theta} C - C^2 A_i^{-\theta} \left(\frac{P_i}{P}\right)^{-\theta}.$$

Let $\overline{\Pi}$ be the steady-state level of real profits in a frictionless economy (upper bars denote steady-state values):[25]

$$\overline{\Pi} \equiv \Pi\left(\frac{P_i^{*}}{P}, \overline{C}, A_i\right).$$

[25] A constant level of aggregate consumption requires the restriction $\left[\int_0^1 A_{it}^{\theta-1} di\right]^{\frac{1}{1-\theta}} = 1$, which we assume holds throughout the paper.

We want to approximate the loss function $\overline{\overline{L}}$ defined as:

$$\overline{\overline{L}}\left(\frac{P_i^*}{P}, \frac{P_i}{P}, C, A_i\right) = \frac{\Pi\left(\frac{P_i^*}{P}, C, A_i\right) - \Pi\left(\frac{P_i}{P}, C, A_i\right)}{\overline{\overline{\Pi}}}$$
$$= \frac{\Pi\left(\frac{P_i^*}{P}, C, A_i\right) - \Pi\left(\frac{P_i}{P}, C, A_i\right)}{\Pi\left(\frac{P_i^*}{P}, \overline{C}, A_i\right)} \frac{\Pi\left(\frac{P_i^*}{P}, C, A_i\right)}{\overline{\overline{\Pi}}}, \tag{16}$$

The second ratio in (16) can be written as:

$$\frac{\Pi\left(\frac{P_i^*}{P}, C, A_i\right)}{\overline{\overline{\Pi}}} = \frac{A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{1-\theta} C - C^2 A_i^{-\theta}\left(\frac{P_i^*}{P}\right)^{-\theta}}{\overline{C} - \overline{C}^2} \tag{17}$$
$$= \frac{A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{1-\theta} C - \frac{\theta-1}{\theta} C A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{1-\theta}}{\overline{C} - \frac{\theta-1}{\theta}\overline{C}}$$
$$= A_i^{1-\theta}\frac{C}{\overline{\overline{C}}}\left(\frac{P_i^*}{P}\right)^{1-\theta}$$
$$= \left(\frac{C}{\overline{\overline{C}}}\right)^{2-\theta},$$

where we use the facts that $\frac{P_i^*}{P} = \frac{\theta}{\theta-1}\frac{C}{A_i}$ and $\overline{C} = \frac{\theta-1}{\theta}$. Note how the link between preferences and technology makes the idiosyncratic shock drop from the expression for maximized profits.

The first ratio in (16) is the proportional profit loss due to the "suboptimal" price P_i. It is convenient to rewrite it as:

$$\frac{\Pi\left(\frac{P_i^*}{P}, C, A_i\right) - \Pi\left(\frac{P_i}{P}, C, A_i\right)}{\Pi\left(\frac{P_i^*}{P}, C, A_i\right)} = 1 - \frac{\Pi\left(\frac{P_i}{P}, C, A_i\right)}{\Pi\left(\frac{P_i^*}{P}, C, A_i\right)}.$$

The profit ratio in the above expression can be written as:

$$\frac{\Pi\left(\frac{P_i}{P}, C, A_i\right)}{\Pi\left(\frac{P_i^*}{P}, C, A_i\right)} = \frac{A_i^{1-\theta}\left(\frac{P_i}{P}\right)^{1-\theta} C - C^2 A_i^{-\theta}\left(\frac{P_i}{P}\right)^{-\theta}}{A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{1-\theta} C - C^2 A_i^{-\theta}\left(\frac{P_i^*}{P}\right)^{-\theta}}$$
$$= \frac{A_i^{1-\theta}\left(\frac{P_i}{P}\right)^{1-\theta} - \frac{\theta-1}{\theta}\frac{P_i^*}{P} A_i^{1-\theta}\left(\frac{P_i}{P}\right)^{-\theta}}{A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{1-\theta} - \frac{\theta-1}{\theta}\frac{P_i^*}{P} A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{-\theta}}$$
$$= \theta\frac{\left(\frac{P_i}{P}\right)^{1-\theta} - \frac{\theta-1}{\theta}\frac{P_i^*}{P}\left(\frac{P_i}{P}\right)^{-\theta}}{\left(\frac{P_i^*}{P}\right)^{1-\theta}}$$
$$= \theta\left(\frac{P_i^*}{P_i}\right)^{\theta-1} - (\theta-1)\left(\frac{P_i^*}{P_i}\right)^{\theta},$$

so that:

$$\frac{\Pi\left(\frac{P_i^*}{P},C,A_i\right)-\Pi\left(\frac{P_i}{P},C,A_i\right)}{\Pi\left(\frac{P_i^*}{P},C,A_i\right)}=1-\theta\left(\frac{P_i^*}{P_i}\right)^{\theta-1}+(\theta-1)\left(\frac{P_i^*}{P_i}\right)^{\theta}. \tag{18}$$

As before, note how the link between preference and technology makes the idiosyncratic shock drop from the expression above.

Combining (17) and (18), and keeping the relevant arguments of the loss function, we obtain:

$$\overline{\overline{L}}\left(\frac{P_i^*}{P},\frac{P_i}{P},C,A_i\right)=\overline{L}\left(\frac{P_i^*}{P_i},C\right)=\left(\frac{C}{\overline{\overline{C}}}\right)^{2-\theta}\left[1-\theta\left(\frac{P_i^*}{P_i}\right)^{\theta-1}+(\theta-1)\left(\frac{P_i^*}{P_i}\right)^{\theta}\right]. \tag{19}$$

We can rewrite the loss function $\overline{L}$ in terms of logarithms:

$$G\left(p_i^*-p_i,c\right)=e^{(2-\theta)c}\left[\left(1-\theta e^{(\theta-1)\left(p_i^*-p_i\right)}\right)+(\theta-1)\,e^{\theta\left(p_i^*-p_i\right)}\right].$$

The exact loss function $G\left(p_i^*-p_i,c\right)$ can be used in the optimal price-setting problems. However, the presence of aggregate consumption in the expression implies that solving for the optimal pricing rule in the presence of pricing frictions involves a fixed-point problem, even in the absence of strategic complementarity or substitutability in price setting. To make the optimal pricing problem more tractable, we eliminate the effect of aggregate output by assuming $\theta=2$ (as in Danziger 1999, and Bonomo and Carvalho 2010). In addition, we take a second-order Taylor expansion of flow profit losses around the frictionless optimal price, based on which we analyze the price-setting problems discussed in the paper:

$$flow\ profit\ losses\,(p_{it})\propto\left(p_{it}-p_{it}^*\right)^2.$$

8 Appendix C

In this appendix we show how we formalize firms' intertemporal optimization problems in the presence of frictions. For brevity we focus on the case with dissociated information gathering/processing and adjustment costs, which is more involved. The other cases are simpler and can be formalized analogously. Let $\widehat{F}$ and $\widehat{K}$ denote the levels of, respectively, the information and adjustment costs. Formally, the pricing problem of a firm may be written as:[26]

$$\widetilde{V}\left(s_{t_0}\right)=\max_{\left\{\left\{t_j,t_{j_n},X_{t_{j_n}}\right\}_{n=0}^{N_j}\right\}_{j=1}^{\infty}}E_{t_0}\sum_{j=0}^{\infty}e^{-\rho(t_j-t_0)}E_{t_j}\left[\begin{array}{c}-e^{-\rho(t_{j+1}-t_j)}\widehat{F}+\int_{t_j}^{t_{j_0}}e^{-\rho r}\Pi\left(\frac{X_r}{P_r},C_r,A_r\right)dr\\+e^{-\rho\left(t_{j_0}-t_j\right)}\sum_{n=0}^{N_j}\left[\begin{array}{c}\int_{t_{j_n}}^{t_{j_{n+1}}}e^{-\rho r}\Pi\left(\frac{X_r}{P_r},C_r,A_r\right)dr\\-e^{-\rho\left(t_{j_n}-t_{j_0}\right)}\widehat{K}\end{array}\right]\end{array}\right],$$

[26] We drop the i subscripts in order to simplify the notation.

where for all $r \notin \left\{t_{0_0}, ...t_{0_1}, t_{0_{N_0}}, ...t_{1_0}, ...\right\}$ $X_r = X_{r^-}$. $\widetilde{V}(s_{t_0})$ denotes the expected present value of real profits Π, net of adjustment and information costs, when the state of the economy is s_{t_0}. The sequence t_j denotes the information dates - dates in which the firm incurs the information gathering and processing cost. The sequence t_{j_n} denotes the dates in which the firm incurs the menu cost and changes its price, and the sequence $X_{t_{j_n}}$ denotes the prices chosen on these dates.

Let $V^*(s_{t_0})$ denote the expected present value of profits of a hypothetical identical firm in the same economy that does not face any pricing friction. Then,

$$V^*(s_{t_0}) = E_{t_0}\left[\int_{t_0}^{\infty} e^{-\rho r}\Pi\left(\frac{P_r^*}{P_r}, C_r, A_r\right) dr\right],$$

where P_r^* is the individual price that maximizes real profits at time r, i.e. the frictionless optimal price of the firm. With this auxiliary value function, $\widehat{V}(s_{t_0}) \equiv V^*(s_{t_0}) - \widetilde{V}(s_{t_0})$ is the minimized expected present value of the real profit losses due to the existence of information and adjustment costs, and our problem can be equivalently rewritten in terms of $\widehat{V}(s_{t_0})$.

Defining $\widehat{L}\left(\frac{P^*}{P}, \frac{X}{P}, C, A\right) \equiv \Pi\left(\frac{P^*}{P}, C, A\right) - \Pi\left(\frac{X}{P}, C, A\right)$ to be the instantaneous real profit loss due to a "suboptimal" price X, and normalizing the pricing problem by the steady-state level of real profits in a frictionless economy, $\overline{\Pi}$, we can rewrite the firm's program as:

$$\overline{V}(s_{t_0}) = \min_{\left\{\{t_j, t_{j_n}, X_{t_{j_n}}\}_{n=0}^{N_j}\right\}_{j=1}^{\infty}} E_{t_0}\sum_{j=0}^{\infty} e^{-\rho(t_j - t_0)} E_{t_j}\left[\begin{array}{c} e^{-\rho(t_{j+1}-t_j)}\overline{F} + \int_{t_j}^{t_{j_0}} e^{-\rho r}\overline{L}\left(\frac{P_r^*}{X_r}, C_r\right) dr \\ +e^{-\rho\left(t_{j_0}-t_j\right)} \sum_{n=0}^{N_j}\left[\begin{array}{c} \int_{t_{j_n}}^{t_{j_{n+1}}} e^{-\rho r}\overline{L}\left(\frac{P_r^*}{X_r}, C_r\right) dr \\ +e^{-\rho\left(t_{j_n}-t_{j_0}\right)}\overline{K} \end{array}\right] \end{array}\right],$$

where $\overline{V}(s_{t_0}) \equiv \frac{\widehat{V}(s_{t_0})}{\overline{\Pi}}$, $\overline{L}\left(\frac{P_r^*}{X_r}, C_r\right) \equiv \frac{\widehat{L}\left(\frac{P_r^*}{P_r}, \frac{X_r}{P_r}, C_r, A_r\right)}{\overline{\Pi}}$, $\overline{F} \equiv \frac{\widehat{F}}{\overline{\Pi}}$, $\overline{K} \equiv \frac{\widehat{K}}{\overline{\Pi}}$. Note that from (19), $\overline{L}(.,.)$ only depends on the ratio of the frictionless optimal price to the charged price, and on aggregate consumption.

Finally, assuming $\theta = 2$, and relying on the same second-order Taylor approximation of flow profit losses used in Appendix B, the firm's (approximate) pricing problem can be written as:

$$V(s_{t_0}) = \min_{\left\{\{t_j, t_{j_n}, x_{t_{j_n}}\}_{n=0}^{N_j}\right\}_{j=1}^{\infty}} E_{t_0}\sum_{j=0}^{\infty} e^{-\rho(t_j - t_0)} E_{t_j}\left[\begin{array}{c} e^{-\rho(t_{j+1}-t_j)}F + \int_{t_j}^{t_{j_0}} e^{-\rho r}\left(x_r - p_r^*\right)^2 dr \\ +e^{-\rho\left(t_{j_0}-t_j\right)} \sum_{n=0}^{N_j}\left[\begin{array}{c} \int_{t_{j_n}}^{t_{j_{n+1}}} e^{-\rho r}\left(x_r - p_r^*\right)^2 dr \\ +e^{-\rho\left(t_{j_n}-t_{j_0}\right)}K \end{array}\right] \end{array}\right],$$

where $V(s_{t_0}) \equiv \frac{\overline{V}(s_{t_0})}{2}$, $F \equiv \frac{\overline{F}}{2}$ and $K \equiv \frac{\overline{K}}{2}$.

The firm's problem at any information date labeled t_0^0 is given by the following intertemporal optimization program:[27]

$$V\left(s_{t_0^0}, p_{t_{-1}^{N_{-1}}}\right) = \tag{20}$$

$$\min_{\left\{t_j^0, \left\{t_j^n, p_{t_j^n}\right\}_{n=1}^{N_j}\right\}_{j=1}^{\infty}} E_{t_0^0} \sum_{j=0}^{\infty} E_{t_j^0} \left\{ \begin{array}{c} e^{-\rho\left(t_{j+1}^0 - t_0^0\right)} F + \int\limits_{t_j^0}^{t_j^1} e^{-\rho\left(r - t_0^0\right)} \left(p_{t_{j-1}^{N_{j-1}}} - p_r^*\right)^2 dr \\ + \sum\limits_{n=1}^{N_j - 1} \left[\int\limits_{t_j^n}^{t_j^{n+1}} e^{-\rho\left(r - t_0^0\right)} \left(p_{t_j^n} - p_r^*\right)^2 dr + e^{-\rho\left(t_j^n - t_0^0\right)} K \right] \\ + \int\limits_{t_j^{N_j}}^{t_{j+1}^0} e^{-\rho\left(r - t_0^0\right)} \left(p_{t_j^{N_j}} - p_r^*\right)^2 dr + e^{-\rho\left(t_j^{N_j} - t_0^0\right)} K \end{array} \right\},$$

where $s_{t_0^0}$ is the initial state, $p_{t_{-1}^{N_{-1}}}$ the inherited price, ρ is the time discount rate, and E_t denotes the expectation operator conditional on time t information.

9 Appendix D

In order to find the optimal rule $\{l(\tau), c(\tau), u(\tau)\}$, we need to find the value function. We start by discretizing the partial differential equation (6) over a grid with time-increments Δt and discrepancy-increments Δz, using an implicit finite-difference method. We make the following approximations:

$$\begin{aligned} z &\approx n \triangle z, \\ \tau &\approx m \triangle t, \\ V_\tau &\approx \frac{v_{n,m+1} - v_{n,m}}{\triangle t}, \end{aligned} \tag{21}$$

$$\begin{aligned} V_z &\approx \frac{v_{n,m} - v_{n-1,m}}{\triangle z}, \\ V_{zz} &\approx \frac{v_{n+1,m} - 2v_{n,m} + v_{n-1,m}}{(\Delta z)^2}. \end{aligned}$$

We can then obtain the following discretization for (6):

$$\rho v_{n,m} = (n \triangle z)^2 + \sigma_c^2 m \triangle t + p^0 v_{n,m} + p^- v_{n-1,m} + p^+ v_{n+1,m} + \frac{1}{\triangle t} v_{n,m+1}, \tag{22}$$

[27] The formalization of the problem starting at an arbitrary date is similar, but heavier on notation.

where

$$
\begin{aligned}
p^0 &= -\frac{\mu}{\triangle z} - \left(\frac{\sigma_f}{\triangle z}\right)^2 - \frac{1}{\triangle t}, \\
p^- &= \frac{\mu}{\triangle z} + \left(\frac{\sigma_f}{2 \triangle z}\right)^2, \\
p^+ &= \left(\frac{\sigma_f}{2 \triangle z}\right)^2.
\end{aligned}
$$

We apply the following solution algorithm. We guess values for the function $v_{n,m}$ for a large grid of times and expected discrepancies. It is important to impose conditions (7)-(11), which state that at any time and discrepancy the price setter will incur the information and/or the adjustment cost, if it is advantageous for her to do so. We then choose a time T large enough to exceed the optimal time interval between information dates for any initial discrepancy z_{t_0}. For such time T, find the z that minimizes $V(z, T)$, denoted $c(T)$, and impose conditions (8) and (10) to determine the new value at T. Then, use the difference equation (22) to find the value function at time elapsed $\tau = T - \triangle t$. Next, impose conditions (8) and (10) to determine the new value at $T - \triangle t$ and so on, until time $\tau = 0$. At that point, test if the value function at each time and discrepancy is close enough (according to some convergence criterion set a priori) to the value function at the previous iteration. Otherwise, begin another iteration. After convergence, use conditions (7), and (9) for each τ to find $c(\tau)$, $u(\tau)$, and $l(\tau)$ and condition (11) to determine $\tau^*(z)$ for any given discrepancy.