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

Marco Bonomo, Carlos Carvalho, René Garcia, Vivian Malta, Rodolfo Rigato†

May 2021

Abstract

We propose a model that reconciles microeconomic evidence of frequent and large price changes with sizable monetary non-neutrality. Firms incur separate lump-sum costs to change prices and to gather and process some information about marginal costs. Additional relevant information is continuously available, and can be factored into pricing decisions at no cost. We estimate the model by Simulated Method of Moments, using price-setting statistics for the U.S. economy. The model with free idiosyncratic and costly aggregate information fits well both targeted and untargeted microeconomic moments and generates more than twice as much monetary non-neutrality as the Calvo model.

JEL classification codes: E00, E31

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

*Circulated previously as "Persistent Monetary Non-neutrality in an Estimated Model with Menu Costs and Partially Costly Information". For comments and suggestions, we thank Fabia Carvalho, Rui Castro, Virgiliu Midrigan, Oleksiy Kryvtsov, Julio Blanco, three anonymous referees, and seminar participants at the 2018 Inflation Targeting Conference of the Central Bank of Brazil, INSPER Workshop on Bayesian VAR and DSGE, 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. André Sztutman and Felipe Alves provided top-tier research assistance. Any errors are our own.

†Bonomo (Insper): bonomo@insper.edu.br, Carvalho (PUC-Rio): cvianac@econ.puc-rio.br, Garcia (Université de Montreal): rene.garcia@umontreal.ca, Malta (IMF): VMalta@imf.org, Rigato (Harvard): dinisrigato@g.harvard.edu.

1 Introduction

Aggregate inflation and output respond slowly to monetary shocks. In contrast, individual prices change somewhat frequently and by large amounts. 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), who provided comprehensive empirical evidence on price setting based on the data underlying the Consumer Price Index in the U.S., this field has experienced renewed interest and noteworthy developments. Many papers have expanded the frontier of so-called menu-cost models. Other work has analyzed the implications of explicit information frictions for price-setting behavior.[1] Menu-cost 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 reconciles microeconomic evidence of relatively frequent price adjustments with persistent real effects of nominal shocks. In our model, both price adjustments and the gathering and processing 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 at no cost.[5] Notice that the nature of information costs in our model goes beyond the costs of information acquisition. As in Reis (2006), information costs in our model encompass the process of deriving the implications of the acquired information for a firm's profit-maximizing price. Hence, to emphasize this important aspect of information costs, we will often refer to them as "information gathering and processing costs".

We employ a novel approach to solve our partial information model, which is not amenable

[1]For instance, Bonomo and Carvalho (2004), Reis (2006), and Alvarez, Lippi, and Paciello (2011) study models with lump-sum information costs, as first analyzed by Caballero (1989). Woodford (2009), Mackowiak and Wiederholt (2009), Matejka (2015a), (2015b), and Stevens (2019) study price setting in the presence of rational inattention, following Sims (2003).

[2]An early contribution is Almeida and Bonomo (2002), who show that disinflation entails only small output losses in an optimal state-dependent pricing model. There are exceptions, however, such as Gertler and Leahy (2008) and Nakamura and Steinsson (2010). These papers show that menu-cost 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 heterogeneity in price rigidity and real rigidities produced by input-output linkages.

[3]Most of these models tend to imply continuous price changes whenever the frictionless optimal price features a known drift. There are exceptions, however (see Section 2).

[4]The literature on price setting with information frictions usually assumes that all information is costly (e.g. Caballero 1989, Reis 2006, Moscarini 2004, Bonomo and Carvalho 2004, 2010, Woodford 2009, Alvarez, Lippi, and Paciello 2011, 2016, and Baley and Blanco 2019).

[5]Gorodnichenko (2008), Knotek (2010) and Klenow and Willis (2016) propose menu-cost models in which firms continuously incorporate partial information into pricing decisions.

to the solution approach employed by Alvarez, Lippi, and Paciello (2011). We then resort to the Simulated Method of Moments (SMM) to estimate the model, using price-setting statistics computed from micro price data from the U.S. Bureau of Labor Statistics. Finally, we study how the estimated economy reacts to monetary shocks.

The data favors a model with free idiosyncratic and costly aggregate information. It matches price-setting statistics that we target in the estimation, and also performs well when assessed against a set of untargeted statistics. At the same time, the estimated model produces a degree of monetary non-neutrality that exceeds that of the Calvo (1983) model by a factor of more than two. The reason is that most price changes reflect idiosyncratic information only, whereas aggregate information gathering and processing, which is key to the extent of monetary non-neutrality, happens only infrequently.

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.[6]

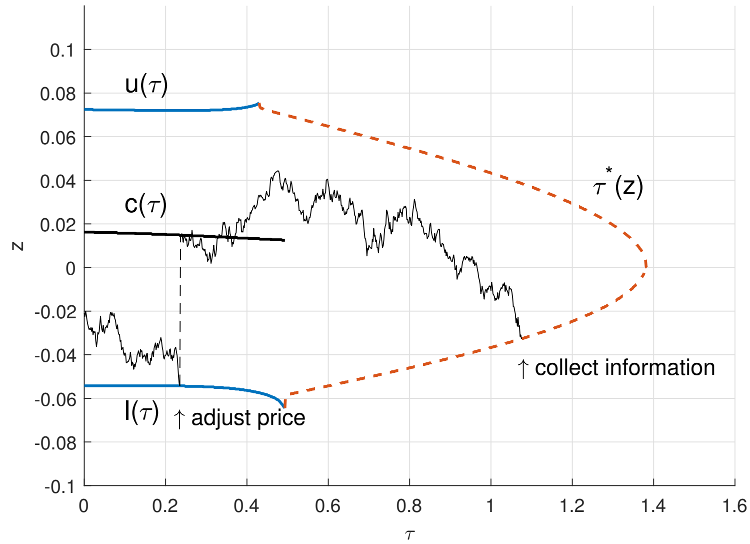
In our model, the difficulty comes from the fact that part of the relevant information about the frictionless optimal price is observed infrequently, owing to the lump-sum nature of the information cost. At the same time, since part of the relevant information is freely and continuously available, it may be optimal for firms to change their prices based on partial information. Hence, firms need to estimate foregone profits given available information.

To solve this problem, we develop 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 firms' foregone profits into two terms: a first component based on the firm's best estimate of its price discrepancy, and a second component due to 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.[7] Our framework can be used to study several

[6]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).

[7]See, e.g., Wilmott, Dewine, and Howison (1994).

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



models with adjustment costs and infrequent information, including most price-setting problems analyzed previously in the literature. Thus, our contribution is also methodological.[8]

In our partial information model, a firm always has the option to incur the menu cost (K) and make a price adjustment based on the current estimate of its price discrepancy (z). 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 boundary of the inaction region depicted as solid blue lines triggers price changes. When its estimated price discrepancy hits the lower (upper) boundary $l(\tau)$ ($u(\tau)$), the firm increases (decreases) its price and sets its estimated price discrepancy to $c(\tau)$. We refer to these as *partially informed* price changes. The subset of the boundary of the inaction region depicted as a dashed red line ($\tau^*(z)$) triggers information gathering and processing. Whenever that boundary is reached, the firm's estimated price discrepancy jumps, as the firm pays the cost to learn the cumulative 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 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.

[8]We first developed and employed this methodology in a retired working paper (Bonomo, Carvalho, and Garcia, 2011).

We estimate the partial information model under two alternative assumptions for the costly and free components of firms' frictionless optimal prices. In one model, price-setters can incorporate idiosyncratic information into their pricing decisions at no cost, while observing/processing information about aggregate shocks is costly. The other model reverses these assumptions. Finally, we estimate the Alvarez, Lippi, and Paciello (2011) model, in which all information is costly.

For each model, we estimate four parameters: information and menu costs, and the size of aggregate and idiosyncratic shocks. The models are estimated by SMM, minimizing the distance to five price-setting moments computed from micro price data from the U.S. Bureau of Labor Statistics.[9]

The model with costly aggregate and free idiosyncratic information yields an almost-perfect fit and is the only one that is not rejected by a test of overidentifying restrictions. In addition, only that model fits well three untargeted moments that the literature has focused on: the kurtosis of the distribution of price changes, the shape of the distribution of the duration of price spells, and the coefficient of a Coibion and Gorodnichenko (2015) regression of forecast errors on forecast revisions, which identifies the degree of information stickiness. In addition to fitting the data well, the story underlying the model with costly gathering and processing of aggregate information is plausible. While collecting aggregate information may seem to involve little cost, processing that information and deriving its implications to a firm's optimal price is arguably quite complex.

We then study the extent of monetary non-neutrality in the estimated economies.[10] The model favored by the data, with costly aggregate information, produces large and long-lasting real effects of monetary shocks, thus reconciling the aforementioned micro and macro evidence. The reason is that most price changes in that model reflect idiosyncratic information only. Such partially informed price changes do not contribute to offseting monetary shocks. Price changes that reflect aggregate information happen infrequently, leading to large and persistent monetary non-neutrality. In contrast, in the other two versions of the model, price changes that incorporate aggregate information happen frequently, and hence the real effects of monetary shocks are short lived.

Finally, we compare the degree of monetary non-neutrality in the partial information model with costly aggregate information with three benchmark models: Calvo (1983), Golosov and Lucas (2007), and a version of the latter model with fat-tailed idiosyncratic shocks. The

[9]The time series of the price-setting statistics used in estimation are frequency of price increases, frequency of price decreases, mean size of price changes, median size of price changes, and the frequency of price increases squared. In Section 4.3 we show how these moments identify the estimated parameters. We thank Oleksiy Kryvtsov for sharing the data with us.

[10]We follow the literature and measure non-neutrality through the area under the impulse response function of real output to a one percent increase in nominal aggregate demand.

three benchmark models are calibrated to match the frequency and mean size of price changes targeted in our estimation. In addition, the version of the Golosov and Lucas (2007) model with fat-tailed shocks is calibrated to generate the same kurtosis of the price change distribution as the partial information model. When compared to the Calvo (1983) model, the partial information model produces more than twice as much non-neutrality, whereas the Golosov and Lucas (2007) model with fat-tailed shocks generates only 70% as much.

The estimated model with costly aggregate information yields too large an estimate for the size of aggregate shocks when compared with the volatility of nominal US GDP. To assess whether our conclusions depend on such counterfactual feature, we also develop and estimate a multisector extension of the model that allows for sectoral productivity shocks, in which we discipline the scale of aggregate shocks using data on nominal US GDP. The estimated multisector model fits the data as well as the baseline model, and produces monetary non-neutrality that exceeds that of the Calvo model.

The rest of the paper is as follows. In Section 2 we discuss the related literature. Section 3 presents our general framework for solving price-setting problems with adjustment and/or information frictions. We then apply the framework to solve the partial infortmation model. Section 4 presents the data, estimation method, and estimation results. In that Section we also discuss parameter identification, and assess the performance of the models based on a set of untargeted moments. In Section 5, we analyze the aggregate effects of monetary shocks in the estimated economies, and compare the partial information model with costly aggregate information to benchmark models from the literature. The last section concludes.

2 Related literature

Several papers contribute to bridging the gap between microeconomic pricing facts and macroeconomic evidence on the real effects of monetary shocks (Gertler and Leahy 2008, Woodford 2009, Nakamura and Steinsson 2010, and Midrigan 2011 are a few prominent examples). Other papers also analyze the interaction between menu costs and information frictions, and their macroeconomic implications. 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 the previous 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. Knotek (2010) uses indirect inference methods to estimate a menu-cost model in which information about macroeconomic shocks is updated at random times, 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 needed to fit the data. Our model differs from his in a few dimensions. Firms face both adjustment and information costs, and design the optimal pricing policy taking both frictions into account. In addition, since our model is estimated with price-setting statistics only, its aggregate implications do not arise from the use of macroeconomic data in the estimation. An earlier paper by Bonomo and Garcia (2001) also studies a model with menu costs and exogenous information arrival at regular intervals. They solve for the optimal pricing policy and then study its macroeconomic implications.

Alvarez, Lippi, and Paciello (2011) and Alvarez, Lippi, and Paciello (2018) are closer to our paper in that they also study models of optimal price setting in the presence of menu costs and lump-sum information costs. In their model, there is no partial information and firms only entertain adjustment after incurring the information cost to observe their price gap.[11] This allows the authors to characterize the firm's problem and provide approximate analytical solutions, by focusing on the value function at times of observation. Under zero inflation drift, which they assume in their baseline model, the optimal policy entails two sequential decisions by the firm: whether or not to incur the menu cost today to change its price, and how long to wait until it is time to incur the observation cost again. The optimal duration of the spell until the next observation is decreasing in the (absolute value of the) chosen price gap — i.e., firms optimally choose to wait longer until the next observation when today's price gap is smaller. Alvarez, Lippi, and Paciello (2018) study the implications of these two pricing frictions for monetary non-neutrality. They solve for the equilibrium of an economy populated by many firms following that optimal pricing policy, and study how calibrated economies react to monetary shocks. For models calibrated to match the frequency and mean size of price changes observed in the data, they find that information costs lead to larger monetary non-neutrality.

Our paper differs from theirs in several dimensions. First, our model features free partial information, and so firms can decide to adjust based on that information only — i.e., without incurring the information cost to learn about the costly component of marginal cost. Because of the presence of partial information, it turns out that our model is not amenable to the solution approach developed by Alvarez, Lippi, and Paciello (2011). The reason is that firms may adjust without incurring the information cost, and cannot plan the time of the next observation only as a function of the chosen price gap on information dates. Instead, firms need to decide at each point in time whether or not to incur the menu cost to change prices, and whether or not to incur the information cost to learn about the costly component of

[11] Alvarez, Lippi, and Paciello (2011) show that this is optimal as long as average inflation is small enough.

marginal costs. In order to solve our model, we develop a tractable unified framework for solving optimal time- and state-dependent price-setting problems, which is a methodological contribution in itself.

Second, in contrast to the common practice of calibrating models to a couple of price-setting statistics, we estimate different versions of our model — including one without partial information, as in Alvarez, Lippi, and Paciello (2011) and Alvarez, Lippi, and Paciello (2018) — using Simulated Method of Moments and a richer set of moments computed from micro price data. This allows us to conduct formal inference about model parameters, and to test alternative models using overidentifying restrictions. Such restrictions reject the model without partial information, while failing to reject the partial information model with costly aggregate information. The latter model also fits an important set of untargeted moments better.[12]

Finally, in terms of substantive macroeconomic lessons, we find that the estimated partial information model with costly aggregate information produces more than six times as much monetary non-neutrality as the estimated model without partial information. Although both models feature the same frequency of price changes, in the partial information model with costly aggregate information, only a small share of price changes incorporate aggregate information. Hence, firms react to monetary shocks slowly. In contrast, in the model without partial information, all price changes embody a response to monetary shocks.

Our paper is also related to Mackowiak and Wiederholt (2009). They study optimal price setting under rational inattention, in the presence of both idiosyncratic and monetary shocks. Because idiosyncratic shocks are much more volatile than monetary shocks, firms optimally choose to allocate attention to tracking the former, at the expense of aggregate shocks. As a consequence, firms face a filtering problem with relatively more precise signals about idiosyncratic shocks. In this context, monetary shocks, which are filtered based on noisier signals, have persistent aggregate effects.

Our estimated model that best fits the data has a similar flavor to Mackowiak and Wiederholt (2009), in that firms have more information about idiosyncratic shocks than about aggregate shocks.[13] Hence, most price changes reflect news about idiosyncratic cost components, and the aggregate price level displays a sluggish response to monetary shocks. But similarities end here. In their model, the information structure is endogenous, and their calibration implies that firms choose to have more information about idiosyncratic shocks. In contrast,

[12] Namely, the kurtosis of the distribution of price changes, the distribution of the duration of price spells, and Coibion and Gorodnichenko (2015) regressions of forecast errors on forecast revisions.

[13] The nature of information frictions in our model is different from that in Mackowiak and Wiederholt (2009), however. While their model features rational inattention as in Sims (2003), ours assumes lump-sum information costs, as in Caballero (1989) and Reis (2006).

we entertain nested models with partial information, with different assumptions about which piece of information is free and which piece is costly, and let the estimation select the model that best fits a rich set of price-setting statistics. This is done by formal estimation of the different models. Crucially, the moments we target in estimation include the frequency of price changes, which our model can fit because it features nominal price rigidities. This contrasts with the Mackowiak and Wiederholt (2009) model, in which prices change continuously.

Baley and Blanco (2019) study a model of price setting with menu costs in which firms have imperfect information about their idiosyncratic productivity. The latter is subject to both continuous shocks and infrequent jumps. Firms receive noisy signals about their productivity, and need to filter this information in order to decide whether or not to incur the menu cost to change prices. Infrequent jumps induce time variation in the degree of uncertainty that firms face, which the authors refer to as "uncertainty cycles". Baley and Blanco (2019) calibrate their model to a set of price-setting statistics and show that it leads to around 2.7 times as much monetary non-neutrality as the Golosov and Lucas (2007) model.

As in Baley and Blanco (2019), firms in our model also need to keep track of the uncertainty regarding their frictionless optimal prices in order to decide on price adjustments. In their model, in the absence of exogenous shocks to uncertainty, learning on the part of firms would lead to convergence over time to a constant amount of uncertainty. In contrast, firms in our model face increasing uncertainty over time, unless they incur the information cost to reduce such uncertainty. Hence, our model also features uncertainty cycles at the firm level. In addition, our model features aggregate shocks.

Our papers are otherwise quite different. Information costs in our model are lump-sum, as in Caballero (1989), Reis (2006), and Alvarez, Lippi, and Paciello (2011), whereas Baley and Blanco (2019) assume noisy information that leads to a signal-extraction problem. We formally estimate and test different models, and assess them against a set of important untargeted moments to which the literatures on price setting and on information frictions have devoted considerable attention. Finally, regarding aggregate implications, we find that our model generates more than thirteen times as much monetary non-neutrality as the Golosov and Lucas (2007) model.

3 A menu-cost model with partially costly information

3.1 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*. The loss in profits per unit of time from charging a given price is increasing in its distance to the (logarithm of the) frictionless optimal price (p_t^*). Price changes entail a lump-sum *menu cost* (K). In Appendix A, we present a simple general equilibrium model that yields an expression for a firm's p_t^* as the sum of two components — a common (nominal aggregate demand) and an idiosyncratic (productivity) component.[14] Information about one of these components is continuously and freely available, and can be factored into price-setting decisions at no cost. Gathering and processing information about the other component entails a lump-sum *information cost* (F).

At this point we remain agnostic about the nature of the two marginal cost components, and refer to them as the "free" and "costly" pieces of the information. Later, when we estimate the model, we entertain three alternatives — namely, costly aggregate information and free idiosyncratic information, a model which reverses these assumptions, and a model in which all information is costly, as in Alvarez, Lippi, and Paciello (2011). We then let the data select the model that yields the best fit.

Intuitively, because of the lump-sum nature of the information cost, a firm chooses to gather and process information about the costly component of p_t^* 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 have to be conditioned on the firm's best estimate of p_t^*, 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 its best estimate of p_t^*, given the firm's information.[15] 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 falls below 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 only on the time elapsed since the last information date, denoted τ (i.e., the last time the firm incurred the information cost F to gather full information about its frictionless optimal price). We thus write $l(\tau), c(\tau), u(\tau)$,

[14]To simplify notation, we omit firm indices from firm-level variables.

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

and refer to $l(\tau), u(\tau)$ as the adjustment boundaries 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 p_t^*. This amounts to a shock to the 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 benefit of learning 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 p_t^* 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 boundary 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.[16]

Figure 1 in the Introduction shows the optimal policy for illustrative parameter values.[17] In the sample path realization for the expected price discrepancy (z_t) that we depict, there is one partially informed adjustment before the firm decides to incur the cost to entertain information about the costly component of p_t^*. When the costly information is revealed, the time-elapsed variable τ is reset to zero, and the firm learns the cumulative innovation to p_t^* that occurred since the previous information date. 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 boundaries, whereas for large τ they are defined by the information boundary. When information about the costly component of p_t^* is not yet too outdated, some new partial information might lead to a large enough expected price discrepancy, inducing the firm to make a partially informed price adjustment. After some point (corresponding to $\tau \approx 0.5$ 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 p_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

[16]This is true as long as the underlying process for p_t^* is continuous (which we assume below). 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 to $c(0)$ immediately.

[17]Parameter values are reported in Section 3.3, where we present the model formally.

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 exercised or not depending on the new information.[18]

We now turn to the specifics of the model and the solution method.

3.2 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.[19] To avoid cluttering notation, we omit individual firms' subscripts from all variables. We assume the (log of the) frictionless optimal price p_t^* evolves according to:

$$dp_t^* = \mu dt - \sigma_{free} dW_{free,t} - \sigma_{cost} dW_{cost,t}, \tag{1}$$

where $W_{free,t}$ and $W_{cost,t}$ are independent standard Brownian motions. Information about $W_{free,t}$ is continuously and freely available, and costless to process. In contrast, gathering and processing information about $W_{cost,t}$ 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.[20] 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:

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

[18]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.

[19]In general, individual firms' problems and aggregation have to be handled jointly, because of general-equilibrium effects. We follow most of the literature and avoid this fixed-point problem by focusing on a parameterization of the model that implies strategic neutrality in price setting (specifically, log consumption utility and linear labor disutility). Under this parameterization, firms' profit losses due to a sub-optimal price are independent of aggregate conditions to a second-order approximation (see Gertler and Leahy 2008 and Alvarez and Lippi 2014, and also Appendix B).

[20]For a formalization of the firm's infinite horizon problem, see Bonomo, Carvalho, and Garcia (2011). Here we focus on its recursive formulation.

where E_t and Var_t denote, respectively, the conditional expectation and conditional variance given time t information. Most of the time, firms' information sets are (partially) outdated, reflecting information about the costly component of p_t^* obtained at the previous information date. Hence, $E_t p_t^* \neq p_t^*$, except on information dates. Notice also that, since firms have continuous access to partial information about p_t^*, the conditional variance $Var_t(p_t^*)$ refers to the component of marginal cost that is only observed at a cost — that is, to the $\sigma_{cost} dW_{cost,t}$ component in equation (1). The first term in the right-hand side of (2) represents the flow cost of deviating from the expected p_t^*. 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. With adjustment costs, 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 $t_0 < t$, $\tau \equiv t - t_0$ denotes the time elapsed since the last information date, and $z_t \equiv p_t - E_t p_t^*$ denotes the expected price discrepancy given the firm's information set. With these definitions, we can rewrite the firm's expected discrepancy cost (2) as a function of τ and z:

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

We can then write the value function at a time t — the optimized value of the firm's dynamic cost-minimization problem described above — in terms of the two state variables z_t and τ. In the inaction region, the value function, $V(z_t, \tau)$, obeys the following Bellman equation:

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

In the inaction region, z_t changes continuously because of both the drift μ and the free component of p_t^*, $W_{free,t}$. So, z_t evolves according the following stochastic differential equation:

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

Taking into account the process for the expected price discrepancy z_t (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 —- the Hamilton-Jacobi-Bellman equation:

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

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.

3.2.1 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 the value function must always satisfy

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

The boundaries 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:[21]

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

3.2.2 The information decision

Firms always have the option to incur the information cost F to gather and process information about the costly component of p_t^*, $W_{cost,t}$. Upon doing so, they learn the realization of $W_{cost,t}$ — 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 value function at the prevailing state with the expected value of a "lottery" that will yield the value at a new state after the realization of $W_{cost,t}$ is learned. Taking the lottery requires paying the information cost F. Optimality

[21] Readers familiar with impulse control problems may notice the absence of the usual smooth-pasting conditions. As shown in Appendix D, equation (8) implies smooth pasting at the optimal boundaries, for each τ.

requires that the current value function does not exceed the sum of the information cost and the expected value of the lottery:

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

where ε is a standard normal random variable. The information boundary, $\tau^*(z)$, is defined by points in the state space in which the firm is indifferent between getting informed and continuing with outdated information, at which (10) holds with equality. Thus, on information dates the expected price discrepancy receives a shock with distribution $N(0, {\sigma_{cost}}^2\tau^*(z))$, and the time elapsed variable τ is reset to zero, yielding the following "informational value-matching condition":

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

3.3 The optimal rule

We solve this pricing problem using a finite-difference method, which we describe in Appendix C. Figure 1 depicts the optimal policy for the following illustrative parameter values: $\mu = 0.05$, $\sigma_{cost} = \sigma_{free} = 0.05$, $K = 0.001$, $F = 0.001$, and $\rho = 0.025$. Under the optimal policy, the firm uses $W_{free,t}$-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 and process 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.

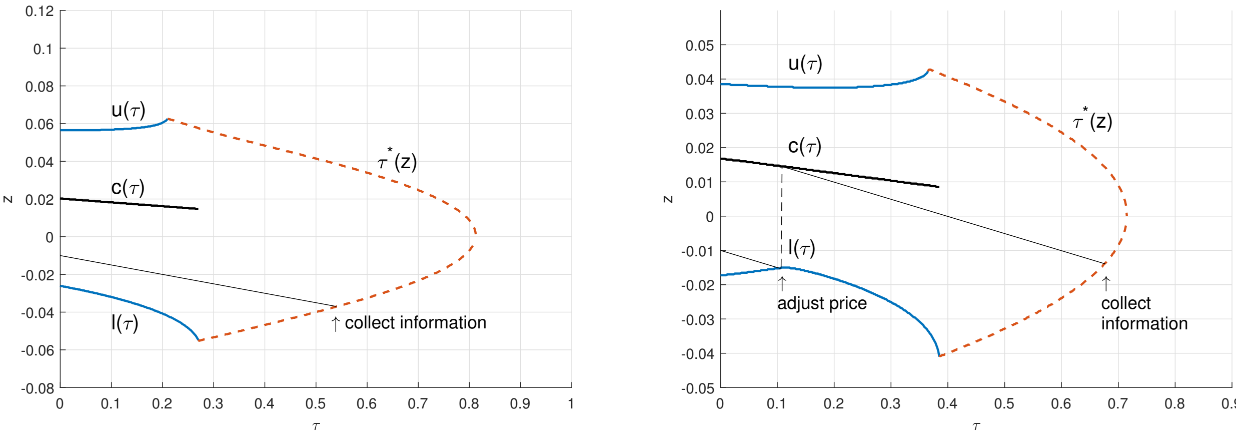
3.4 A particular case without partial information

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, Lippi, and Paciello (2011). We solve this problem by setting $\sigma_{free} = 0$ and relabeling $\sigma = \sigma_{cost}$ in our partial information model. Thus, the differential equation that characterizes the evolution of the value function inside the inaction region becomes:

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

Figure 2: Optimal pricing policy in the model with no partial information

(a) Illustrative parameter values (b) Smaller menu cost



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

Figure 2 illustrates the optimal pricing policy under adjustment and information costs when all information is costly. Panel a) is based on the following illustrative parameter values: $\mu = 0.05$, $\sigma_{cost} = 0.05\sqrt{2}$, $K = 0.001$, $F = 0.001$, and $\rho = 0.025$. The parameter σ_{cost} is set to make the volatility of innovations to the firm's frictionless optimal price equal to the total volatility in the illustrative parameterization of the partial information model underlying Figure 1. The solid (blue) lines $l(\tau)$, $u(\tau)$ are the adjustment boundaries of the inaction region that may trigger uninformed adjustments, while the dashed (red) line $\tau^*(z)$ is the information boundary that triggers information gathering and processing.

From the comparison of Figure 1 and panel a) in Figure 2, it is apparent that the information boundary is further to the right in the model with partial information (Figure 1). The reason for less frequent information gathering and processing in the partial information model is that in this case the firm has continuous information about one component of its frictionless optimal price. This diminishes its uncertainty and thus reduces the benefits of incurring the information cost to gather and process information as early as in the model without partial information. For this set of illustrative parameter values, this effect is quantitatively important: the maximum spell between two information dates is about twice as long in the model with partial information.

Panel b) illustrates the possibility of uninformed price adjustments. Alvarez, Lippi, and Paciello (2011) 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 trend inflation (deflation) is high enough or, for a given level of trend inflation ($\mu \neq 0$), if the menu cost is small enough relative to the information cost. In panel b), we illustrate the latter case by keeping the same parameter values as in panel a) and reducing the menu cost by about two thirds ($K = 0.0003$). This opens up the possibility of uninformed price changes, which do not occur under the optimal pricing policy depicted in panel a).

4 Estimating the model

We use time series of price-setting statistics computed from micro data from the U.S. Bureau of Labor Statistics to estimate the main parameters of the model by Simulated Method of Moments (SMM). Simulation-based methods are required, since the relationship between the model parameters and observed statistics is highly non-linear and complex. As detailed in the model in Appendix A, firms' frictionless optimal prices are subject to nominal aggregate demand shocks and idiosyncratic productivity shocks, both of which follow Brownian Motions with volatility parameters σ_{agg} and σ_{id}, respectively. We estimate our model under three different hypotheses regarding the nature of costly and free information.

In the first model we estimate, price-setters can incorporate all the information about idiosyncratic shocks into their pricing decisions at no cost, while observing and processing aggregate information is costly. This amounts to setting $\sigma_{cost} = \sigma_{id}$ and $\sigma_{free} = \sigma_{agg}$. We then estimate another specification that reverses these assumptions, i.e. sets $\sigma_{free} = \sigma_{id}$ and $\sigma_{cost} = \sigma_{agg}$. The last model is one in which all information is costly, as in Alvarez, Lippi, and Paciello (2011) — $\sigma_{free} = 0$ and $\sigma_{cost} = \sqrt{\sigma_{id}{}^2 + \sigma_{agg}{}^2}$. In Section 4.5, we estimate a multisector extension of our model in order to understand how missing sectoral shocks may affect our results.

4.1 Estimation method

Estimation by SMM involves minimizing the distance between data moments and moments obtained through simulation of the model. Moments are obtained as sample averages of the time series of price-setting statistics used in the estimation, which we describe in the next section.

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. $\Psi_{sim}(\Phi)$ is obtained by simulating time series of the relevant price-setting statistics, taking time averages to obtain the desired moments, and, finally, taking averages of moments across many simulations. The estimator $\hat{\Phi}$ is obtained by solving the following minimization problem:

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

where W is the optimal $q \times q$ weighting matrix, given by the inverse of the variance-covariance matrix of sample moments adjusted for simulation error (see, e.g., DeJong and Dave 2011).

4.2 Data

We estimate the model with monthly time series of price-setting statistics from February 1988 to January 2005 (204 months), constructed by Klenow and Kryvtsov (2008). Those statistics were computed from individual price changes (excluding sales) in the micro data underlying the Bureau of Labor Statistics' Consumer Price Index (CPI) for the top urban areas of the United States. In each month, the price-setting statistics are computed using the cross-section dimension.

The five time series we use are: i) frequency of price increases; ii) frequency of price decreases; iii) mean size of price changes; iv) median size of price changes; and v) frequency of price increases squared. Figure 3 shows the time series of these five price-setting statistics. As a first step for estimation, we remove possible seasonality in these series by regressing each against a set of seasonal monthly dummy variables, and proceed with the residuals added to the sample mean.

For the simulated data, we generate artificial 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 (13,500) 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 stationary distribution to which the economy converges in the absence of aggregate shocks.[22]

To reduce the number of parameters to be estimated, we set the inflation drift μ equal to the average CPI inflation over the period during which the micro data used to compute the empirical price-setting statistics were sampled (3.3% per year). In addition, we set the time-discount rate to $\rho = 2.5\%$ per year. Hence, for each of the three models described previously, this leaves us with four parameters to be estimated: $\sigma_{agg}, \sigma_{id}, K, F$.[23]

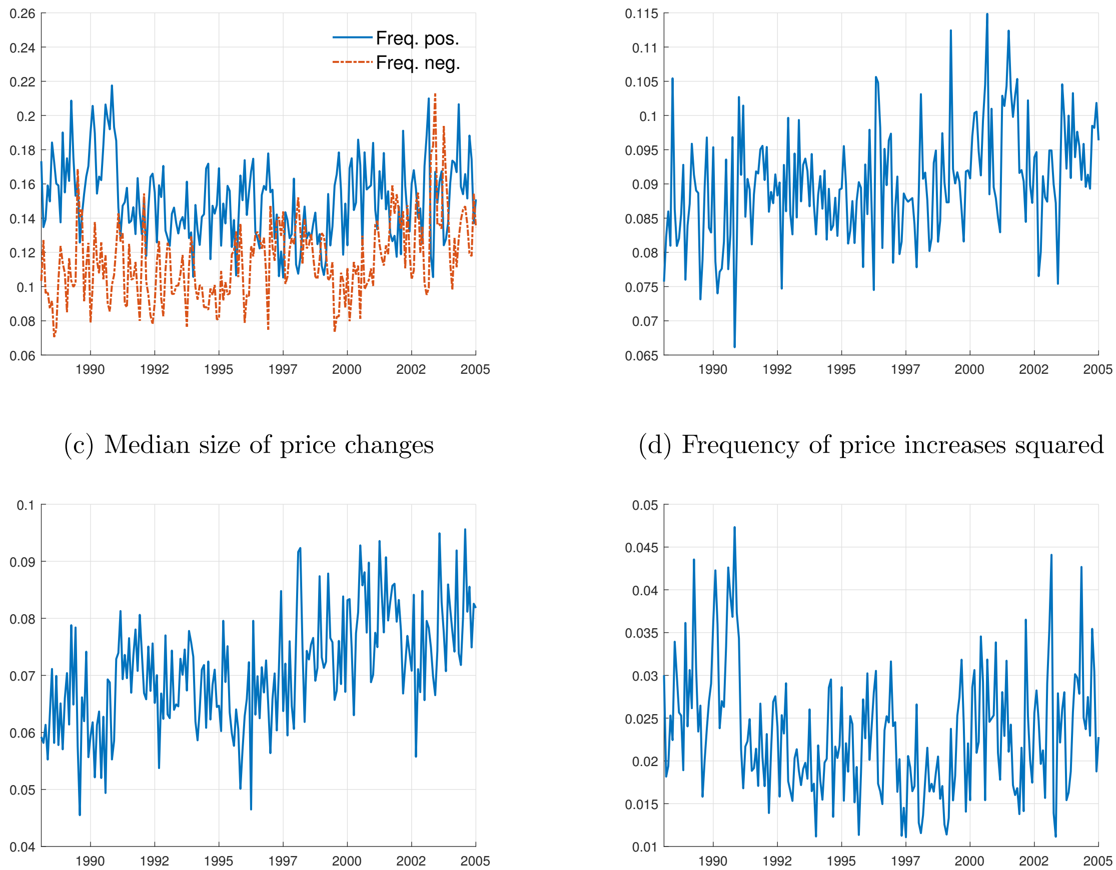
[22] To obtain the stationary distribution, we solve its associated Kolmogorov Forward Equation, following the numerical algorithm in Achdou et al. (2021).

[23] Monte Carlo simulations confirm that our SMM estimation with artificial time series of these price-setting statistics recovers the true parameter values used to generate the artificial samples.

Figure 3: Time series used in estimation

(a) Frequency of price increases and decreases (b) Mean size of price changes

(c) Median size of price changes (d) Frequency of price increases squared



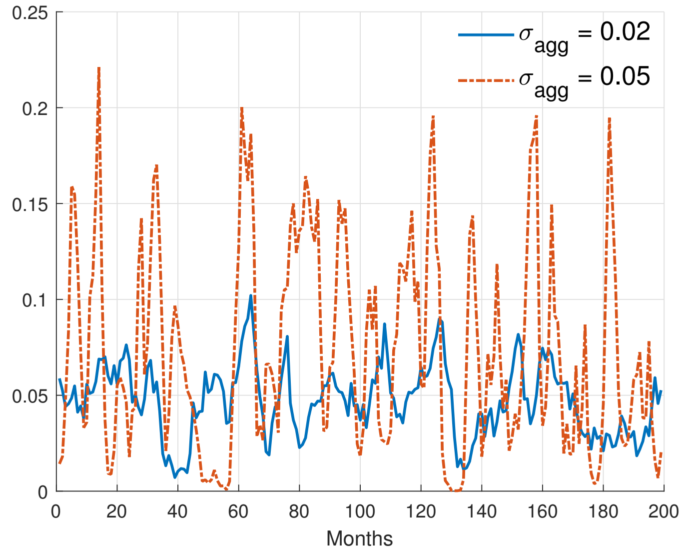
4.3 Parameter identification

In this section we provide some intuition for how the five moments used in the estimation of the model identify the parameters of interest (σ_{agg},σ_{id}, K, F). We do so by simulating, in the partial information model, the effect of changing one parameter at a time on the moments used in the estimation.[24] Although in principle changing one parameter should affect several moments, we focus on a subset of moments that are substantially affected by the parameter at hand. For this analysis, we vary parameter values around those used in the illustrative calibration of the partial information model that underlies Figure 1, given by (σ_{agg}, σ_{id}, K, F, μ, ρ)=(0.05, 0.05, 0.001, 0.001, 0.05, 0.025).

We start by illustrating how time series variation in price-setting statistics is informative of the scale of aggregate shocks. In the absence of aggregate shocks (i.e. $\sigma_{agg} = 0$) the

[24]The intuition for parameter identification in the model without partial information is similar, unless noted.

Figure 4: Simulated time series of the frequency of price increases for two values of σ_{agg}



distribution of firms in the space (τ, z) would settle at a stationary distribution. Once this distribution is reached, all price-setting statistics would remain constant over time.[25] Therefore, the volatility of aggregate shocks determines the time series variation of price-setting statistics. This effect is illustrated clearly in Figure 4, which depicts simulated time series of the frequency of price increases for two values of σ_{agg}. Higher aggregate volatility generates more variable frequency of price increases.[26] This provides intuition for how the uncentered second moment of the frequency of price increases (which is obtained as the time-series average of the frequency of price increases squared) identifies σ_{agg}.

We now use comparative statics to illustrate how the moments used in estimation change as we vary parameter values one at a time, in Figure 5. Figure 5a shows that more volatile aggregate shocks σ_{agg} increase the uncentered second moment (and the variance) of the frequency of price increases. This is the comparative statics suggested by the simulated paths in Figure 4. To construct Figure 5a, for each value of σ_{agg}, we simulate 250 realizations of the economy with 13,500 firms over 204 months, and average the value of the moment of interest across simulations.

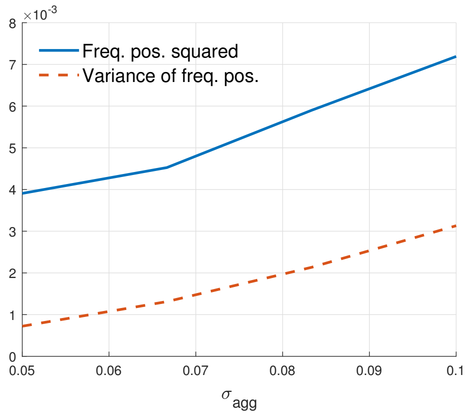
Figure 5b shows the effect of varying the scale of idiosyncratic shocks on the frequencies of price increases and decreases, and on their difference. Owing to the positive inflation drift μ, when idiosyncratic shocks are small (i.e. for small σ_{id}), price increases are much

[25] 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 based on which the target moments are computed.

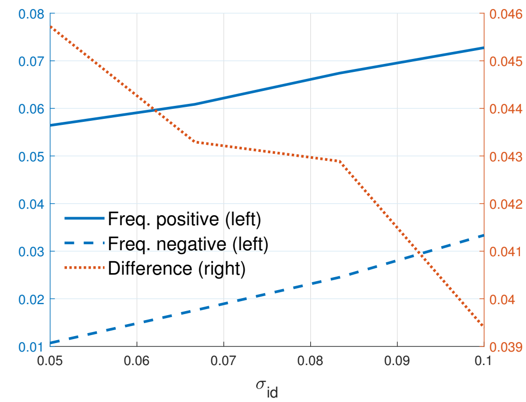
[26] This is also true of other price-setting statistics.

Figure 5: Parameter identification

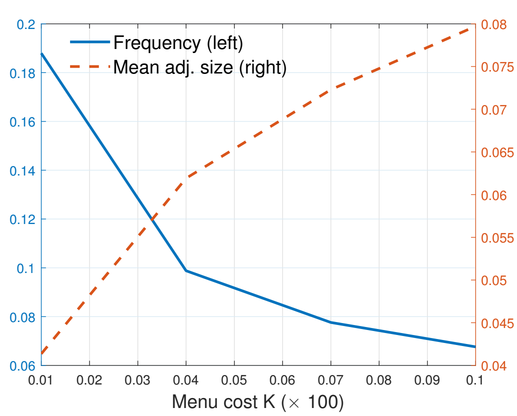
(a) Uncentered second moment and variance of frequency of price increases as functions of σ_{agg}



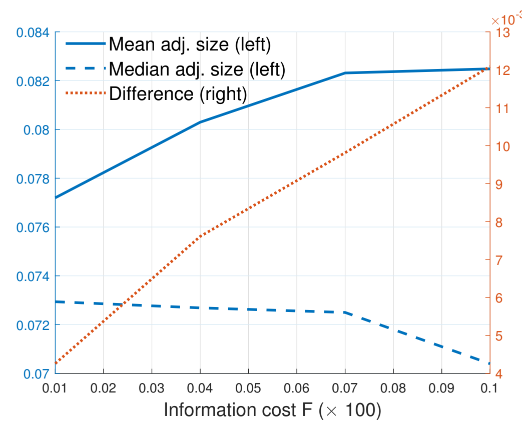
(b) Frequencies of price increases and decreases and their difference as functions of σ_{id}



(c) Frequency (left) and adjustment size (right) as functions of K



(d) Mean and median adj. size and their difference as functions of F

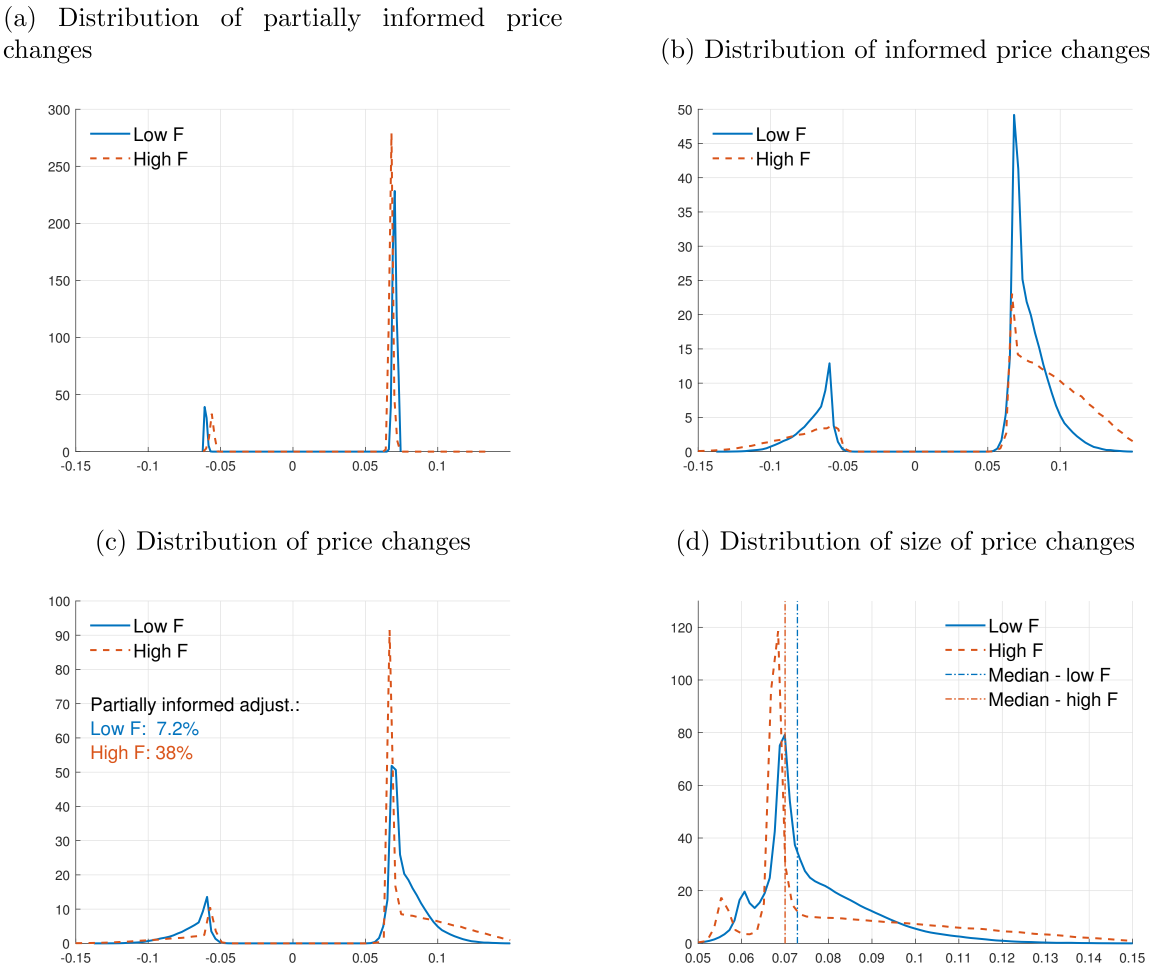


more common than price decreases. As we increase σ_{id}, the asymmetry generated by the inflation drift is attenuated by larger symmetric shocks, and hence the difference between the frequency of price increases and decreases falls. We use the same procedure to generate the comparative statics for the other parameters.

Identification of the menu cost K is straightforward, as the effects of varying the size of menu costs are well known. As K increases, the inaction region widens. As a consequence, price increases (and decreases) become less frequent and the mean adjustment size increases, as shown in Figure 5c. Thus, both the frequency of price increases (and decreases) and the mean adjustment size contribute to identification of K.

Finally, Figure 5d shows how the mean and median size of price adjustments change as we vary the information cost F. To understand this comparative statics, it is useful to recall that the distribution of price changes is a mixture of two underlying distributions: of fully

Figure 6: Distributions of fully and of partially informed price changes for different information costs



informed and of partially informed price changes. Fully informed price increases (decreases) have minimum size given by $c(0) - l(0)$ $(u(0) - c(0))$, and no upper bound. Hence, these price changes dominate the right tail of the distribution of absolute price changes. In turn, the distribution of partially informed price changes is less dispersed.[27] This composition of distributions is illustrated in Figure 6, for the illustrative calibration at hand.

Under this calibration, when the information cost F increases, the mean size of adjustments increases, whereas the median size decreases (Figure 5d).[28] The reason why the mean size increases is simple. When the information cost increases, the distribution of the size

[27]The size of partially informed adjustments is pinned down by the distances between the adjustment boundaries $l(\tau), u(\tau)$ and the target point $c(\tau)$. These distances do not vary much with τ.

[28]In the model without partial information, both the mean and median size of price changes increase with the information cost, but the difference between them still increases with F.

Table 1: Target and simulated statistics

Statistics	$f.pos.$	$f.neg.$	$avg.\,\lvert\Delta p\rvert$	$med.\,\lvert\Delta p\rvert$	$(f.pos.)^2$	$J-stat.$	$p-val.$
Data	0.150	0.115	0.090	0.071	0.023	-	-
No Partial info	0.148	0.120	0.087	0.077	0.022	101.7	0.000
Partial ($\sigma_{cost} = \sigma_{id}$)	0.147	0.121	0.088	0.076	0.023	88.84	0.000
Partial ($\sigma_{cost} = \sigma_{agg}$)	0.150	0.116	0.090	0.071	0.023	0.0284	0.866

of partially informed price adjustments does not change much (see Figure 6a).[29] In turn, informed price changes become less frequent, but larger on average (see Figure 6b), because unobserved innovations accumulate over longer spells. In the relevant range of parameters — which includes the ones used in the illustrative calibration, as well as the estimates of the model, presented below — the effect of larger informed price changes dominates, and thus the unconditional mean size of price changes goes up.

The median size of price changes decreases as a higher information cost is associated with a larger share of partially informed price changes (Figure 6c). The sizes of those partially informed adjustments fall below the median of the overall distribution (see panels in Figure 6). Since the share of partially informed price changes increases with the information cost, and their distribution is relatively invariant to F (see Figure 6a), the median of the overall distribution of the size of price changes shifts to the left (Figure 6d).

4.4 Estimation results

Table 1 compares targeted data moments with those produced by the different estimated models. Informal inspection of the moments produced by the models suggests that the partial information model with costly aggregate information fits the data better than the two other versions of the model — in fact, its fit is almost perfect (last row of Table 1). Indeed, a formal test of overidentifying restrictions shows this is the only model that is not rejected by the data (last two columns of Table 1). One may wonder whether the differences between moments produced by the three estimated models are economically meaningful. As we show below, in addition to fitting targeted moments better, the partial information model with costly aggregate information also performs better when assessed against a set of important untargeted moments, namely: the kurtosis of the distribution of price changes, the distribution of the duration of price spells, and the sensitivity of forecast errors to forecast revisions estimated through Coibion and Gorodnichenko (2015) regressions.

Table 2 reports parameter estimates for the three models. The model with partial in-

[29] This is so because the adjustment boundaries of the inaction region $(l(\tau), u(\tau))$, as well as the target expected discrepancy $(c(\tau))$, do not change much when the information cost varies.

Table 2: Parameter Estimates

	σ_{agg}	σ_{id}	K	F
No Partial Info	0.028	0.172	0.00022	0.0006
t-statistic	7.73	51.75	4.16	2.59
Partial $\sigma_{agg} = \sigma_{free}$	0.026	0.178	0.00013	0.0009
t-statistic	11.18	40.71	2.64	4.03
Partial $\sigma_{agg} = \sigma_{cost}$	0.149	0.128	0.00030	0.0188
t-statistic	18.36	83.36	49.91	6.03

obs: t-statistics based on standard errors obtained with the delta method

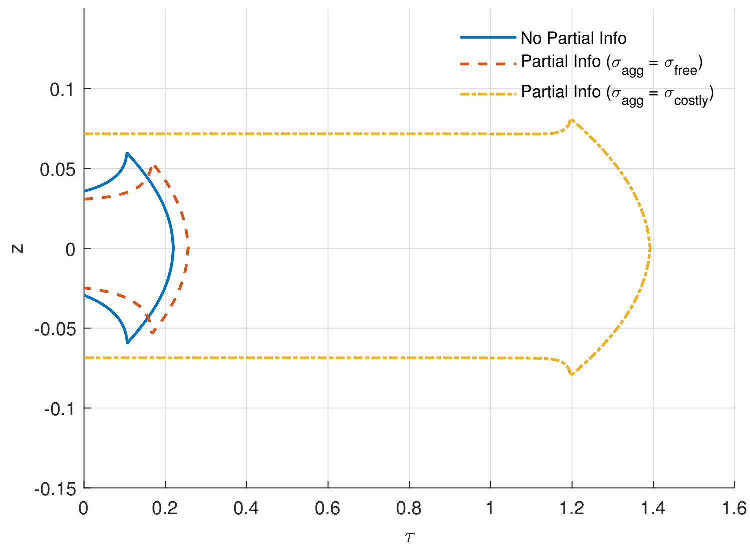
formation in which aggregate information is free and the model with no partial information produce relatively similar parameter estimates. Estimated menu costs are of the same order of magnitude in the three models. In contrast to the two other specifications, the model with costly aggregate information produces a larger estimate for the variance of aggregate shocks, coupled with a larger information cost. In our model, aggregate shocks are due to fluctuations in nominal GDP (see Appendix A). Hence, such a high estimate for the volatility of aggregate shocks is at odds with aggregate data, which we do not use when estimating the models. The high estimate for the volatility of aggregate shocks may proxy for omitted sectoral shocks. In Section 4.5, we present a multisector extension of the model with sectoral shocks in which we use aggregate data to discipline the volatility of aggregate shocks. The multisector model fits the micro moments equally well and is consistent with nominal GDP volatility. Most importantly, that model corroborates the conclusions that we draw below, based on the simpler one-sector model.

Figure 7 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 without partial information produce quite similar pricing policies. In contrast, the model with costly aggregate information leads to a very different optimal policy, with a bigger inaction region that encloses the other two. Its adjustment boundaries are wider, and costly information gathering and processing happens less frequently.

These differences in optimal policies imply different price-setting statistics (Table 3). Information gathering/processing occurs more often in the model with no partial information and in the model with free aggregate information: more than four times per year in both models, versus less than once per year in the model with costly aggregate information. In addition, those two models feature more frequent information gathering than price adjustments, in contrast with the model with costly aggregate information.[30] Another difference

[30] In interpreting results for the models with partial information, one should bear in mind that the frequency

Figure 7: Optimal policies implied by the estimated models



is that in those first two models almost all price 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, conditional on full information gathering/processing, price adjustments happen with a similar frequency — between 69-76% of the time.

The intuition underlying these results is straightforward. The estimated model with costly idiosyncratic information features small observed 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 the model with no partial information. Results 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 the volatile idiosyncratic marginal cost component reduces the incentives to gather costly information frequently. Hence, for a given information gathering cost F, the firm chooses to sample less frequently. In addition, the estimated information cost in the model with costly aggregate information is larger than in the other two models, further contributing to less frequent information gathering.

In terms of economic magnitudes, the estimated parameters for the model with costly aggregate information imply plausible adjustment and information costs when compared to

of information gathering and the frequency of observation do not coincide. Firms observe some information all the time, but only incur the cost to gather and process full information infrequently. Thus, the results of the estimated model with costly aggregate information do not contradict the survey evidence that suggests observation happens more often than price adjustments (for a summary of the available survey evidence, see Table 1 in Alvarez, Lippi, and Paciello 2018).

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

Baseline	No Partial	$\sigma_{cost} = \sigma_{id}$	$\sigma_{cost} = \sigma_{agg}$
Price adj. per year	3.19	3.20	3.27
fully informed	100%	95%	16%
partially informed	0%	5%	84%
Info. gatherings per year	4.62	4.01	0.76
resulting in price adj.	69%	76%	69%
Kurtosis of price change distribution	2.16	2.43	5.17

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%. In the model with costly aggregate information, assuming elasticity of substitution between goods of 3 (an upper bound obtained by Hobijn and Nechio 2018 with a low level of aggregation), the estimated parameters imply that firms spend, on average, 9% of annual steady-state profits on information and adjustment costs.[31] Menu costs account for 0.6% of annual profits, whereas information costs account for 8.4% 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.

A key feature of the model with costly aggregate information is that it endogenously generates a distribution of price adjustments with high kurtosis, as displayed is Table 3.[32] The model with costly idiosyncratic information, and the model with no partial information, as in Alvarez, Lippi, and Paciello (2011), generate less than half of that kurtosis. Alvarez, Le Bihan, and Lippi (2016) summarize available estimates of this moment for the U.S., which range from 4 to 5.1. Models of optimal price adjustment with Gaussian shocks to marginal costs tend to generate lower kurtosis, with the Golosov-Lucas model generating kurtosis of 1, while a multiproduct pricing model generates at most 3, which equals the kurtosis of the distribution of shocks to marginal costs. In order to produce realistic kurtosis

[31] As shown in Appendix B, in order to obtain those costs as a fraction of steady-state profits, cost parameters must be multiplied by $\theta(\theta-1)/2$, where θ is the elasticity of substitution between goods.

[32] The price-setting literature has devoted considerable attention to this statistic since Alvarez, Le Bihan, and Lippi (2016). The authors showed that in a large class of models calibrated to match the frequency of price-adjustments, the kurtosis of the price change distribution is a sufficient statistic for monetary non-neutrality, measured as the area below the impulse response function of output with respect to a monetary shock. Our model does not fall into the class models for which kurtosis is a sufficient statistic, as we further explore in Section 5.

of price changes, or to improve model fit with aggregate data, part of the literature resorts to leptokurtic shocks to marginal costs (Gertler and Leahy 2008, Midrigan 2011), multiproduct pricing (Midrigan 2011, Alvarez, Le Bihan, and Lippi 2016), or occasional free adjustments (Nakamura and Steinsson 2010, Alvarez, Le Bihan, and Lippi 2016).[33] In our model, there is no need for those features, since the mix of partially informed and fully informed adjustments engenders kurtosis in price changes — an untargeted moment — close to empirical estimates.

The three models have also sharply different implications for another untargeted statistic — namely, the distribution of the duration of price spells.[34] As shown in Figure 8, the model with no partial information and the model with free aggregate information produce counterfactual distributions, with almost 70-80% of the weight on 3-month price spells, and "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 smooth distribution that resembles the empirical distribution.[35]

It is easy to make sense of the differences in the distributions of price spells produced by the three estimated models. In the model with free aggregate information and in the model with no partial information, essentially all price changes are fully informed (see Table 3). This happens because, in these two estimated models, information gathering/processing happens roughly once every three months (see Table 3 and Figure 7), the inflation drift is small, and the rate of innovations to the freely observed component is low. In addition, conditional on information gathering/processing, price changes are quite likely (69-76%). Hence, a large proportion of price changes take place after spells of roughly three months. Conditional on choosing not to adjust after gathering information, firms have a high likelihood of adjusting on their subsequent information date — hence the second largest mass point in the distribution falls on spells of roughly twice the length of the modal price spell.[36] In contrast, the model with costly aggregate information produces a proportion of partially informed adjustments in excess of 80% (see Table 3). These price spells have quite variable lengths, as partially informed price changes are driven by volatile idiosyncratic shocks. Hence, this model produces a smooth histogram, which more closely resembles the data.

[33]Bonomo et al. (2020) analyze kurtosis of the price change distribution in a flexible class of multiproduct price-setting models that allows for partial synchronization of price changes within a firm.

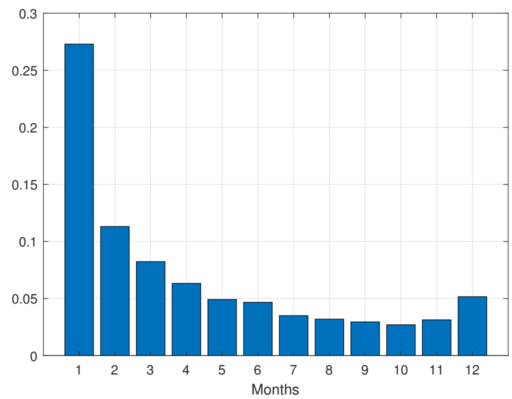
[34]Carvalho and Schwartzman (2015) show that the first two moments of the distribution of the duration of price spells are sufficient statistics for monetary non-neutrality, measured as the area below the impulse response function of output with respect to a monetary shock, in the class of time-dependent price-setting models.

[35]One may worry that the spread of this empirical distribution results from the pooling of heterogeneous distributions of price spells with little within-product variation. Klenow and Kryvtsov (2008), however, report substantial variation in the duration of price spells both within narrow product categories and even over time for a given quote line (see their Table V).

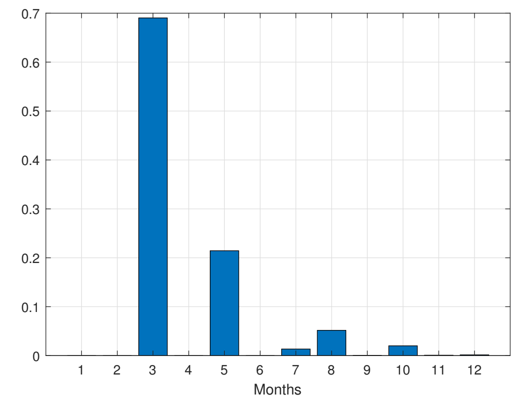
[36]The remaining features of the histograms arise due to the small differences between these two estimated models.

Figure 8: Distribution of price spells for models and data

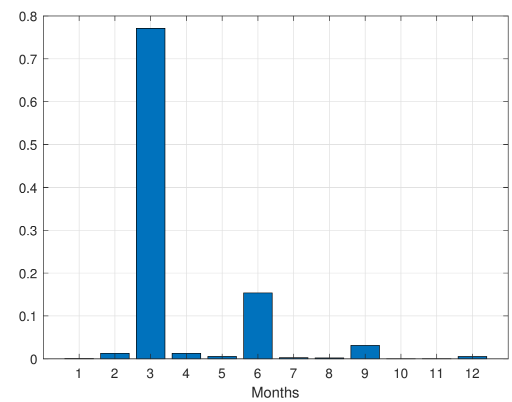
(a) Empirical distribution from Klenow and Kryvtsov (2008)



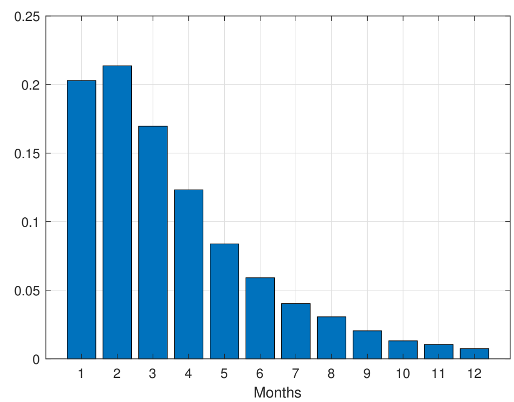
(b) Model without partial information



(c) Model with costly idiosyncratic information



(d) Model with costly aggregate information



Finally, in order to understand whether our model implies reasonable degrees of information rigidity, we follow Baley and Blanco (2019) and use simulated data to estimate the same regressions as in Coibion and Gorodnichenko (2015) (henceforth "CG"). Let x_t denote a given aggregate variable and $\mathbb{F}_t x_{t+h}$ be the average forecast at time t of this variable at $t+h$. In models in which firms do not have full information about aggregate conditions, the average forecast at a given point differs from the full-information expectation (denoted here E_t^*) — i.e. $\mathbb{F}_t x_{t+h} \neq E_t^* x_{t+h}$. Instead, it is a combination of lagged expectations of the form $E_{t-\tau}^* x_{t+h}$.

We work with firms' forecasts of nominal aggregate demand. We aggregate its path to produce a quarterly series, and estimate the following regression using simulated data:

$$x_{t+h} - \mathbb{F}_t x_{t+h} = \alpha + \beta(\mathbb{F}_t x_{t+h} - \mathbb{F}_{t-1} x_{t+h}) + \text{error term}. \tag{13}$$

Table 4: Coibion-Gorodnichenko regressions.

	CG	Free agg. info.	No partial info.	Costly agg. info.	Multisector
Constant	0.002	−0.000	0.000	−0.002	−0.000
	(0.144)	(0.004)	(0.005)	(0.022)	(0.025)
Slope	1.193	−0.051	0.165	1.756	0.994
	(0.497)	(0.144)	(0.198)	(0.589)	(0.359)
Obs.	173	173	173	173	173
R^2	0.195	0.007	0.011	0.118	0.071

We follow CG in using a one-year horizon ($h = 4$) and quarterly forecast revisions as the explanatory variable.

It is straightforward to see that, absent any information frictions — i.e. when $\mathbb{F}_t = E_t^*$ — it should be the case that $\beta = 0$. In words, when expectations are rational and agents are fully informed/attentive, forecast revisions should not explain forecast errors. CG show that models with staggered information or other frictions such as rational inattention generate $\beta > 0$. The authors then estimate (13) using data from the Survey of Professional Forecasters for different aggregate variables, and find a strong relationship between forecast errors and forecast revisions.

In order to make our results as comparable as possible to CG, we perform the following Monte Carlo simulation. For each model, we simulate $n = 100$ samples, each consisting of 173 quarters and 35 firms, thus mimicking the sample features in CG. We report the average and standard deviation (in parenthesis) of the estimated regression coefficients across samples.[37] Results are displayed in Table 4.

The model in which aggregate information is free yields essentially no relationship between forecast errors and forecast revisions. This is expected, as in that model firms have always have full-information about the variable they're forecasting. The model with no partial information also fails to generate statistically significant regression slopes. The reason is that, as Figure 7 and Table 3 show, information gathering/processing happens quite frequently in that model. Consequently, firms' information sets are not too outdated, and the resulting regression slope becomes small enough for the null hypothesis of $\beta = 0$ not to be rejected. In contrast, the model with costly aggregate information generates a statistically positive regression slope, which falls within the 90% confidence interval around CG estimates.

In Coibion and Gorodnichenko (2012), the authors assess the extent of forecasters' information rigidity by looking at how the average inflation forecast error responds to different

[37] We do not divide standard errors by $\sqrt{n}$. Reported standard errors are, therefore, estimates of what one would obtain in a typical sample of 173 quarters (with 35 forecasters).

types of shocks. Under the null hypothesis of full-information rational expectations, forecasts should adjust immediately. They find, however, that the average forecast adjusts sluggishly, consistent with economic agents updating their information sets every 6 or 7 quarters. Our estimated model with costly aggregate information implies firms collect information approximately every 5 quarters, falling slightly short of the aforementioned estimates. This is yet another untargeted moment for which the model generates plausible results.

As a summary of our main results, we find that the partial information model with costly aggregate information outperforms the other two — fits target moments better, and is the only model that is not rejected by the test of overidentifying restrictions. Moreover, that model fares better also when confronted with a set of three untargeted moments: the kurtosis of the price change distribution, the distribution of the duration of price spells, and the sensitivity of forecast errors to forecast revisions, obtained from Coibion-Gorodnichenko regressions.

One issue with the partial information model with costly aggregate information is that it produces a counterfactually large estimate for the volatility of aggregate shocks. As argued previously, this may proxy for large sectoral shocks that are also costly to observe, which are absent from the model. In the next section, we develop a multisector extension of the partial information model with costly aggregate and sectoral information, and use aggregate data to discipline the size of aggregate shocks. We find that the main results and conclusions just presented — as well as the aggregate implications of the partial information model with costly aggregate information — carry over to the multisector model.

4.5 A multisector model with partial information

The model is an extension of the partial information model with costly aggregate information. It features three sectors that are subject to sectoral productivity shocks ($W_{sec,t}$), in addition to the shocks featured in the baseline model. Consequently, firms' frictionless optimal prices now evolve as

$$dp_t^* = \mu dt - \sigma_{id} dW_{id,t} - \sigma_{sec} dW_{sec,t} - \sigma_{agg} dW_{agg,t},$$

where σ_{sec} gives the scale of sectoral shocks. Gathering and processing information about $W_{sec,t}$ and $W_{agg,t}$ is costly, whereas about $W_{id,t}$ is costless.[38]

We discipline the size of aggregate shocks by setting $\sigma_{agg} = 0.02$. This value is in line with the annualized volatility of U.S. nominal GDP growth. We then estimate remaining model parameters using the same set of moments used to estimate the one-sector models

[38] In the notation of the baseline model, this amounts to setting $\sigma_{free} = \sigma_{id}$ and $\sigma_{cost} = \sqrt{\sigma_{agg}^2 + \sigma_{sec}^2}$.

Table 5: Target and simulated statistics

Statistics	$f.pos.$	$f.neg.$	$avg\,\|\Delta p\|$	$med.\,\|\Delta p\|$	$(f.pos.)^2$	$J-stat$	$p-val$
Data	0.150	0.115	0.090	0.071	0.023	-	-
Partial ($\sigma_{agg} = \sigma_{cost}$)	0.150	0.116	0.090	0.071	0.023	0.0284	0.866
Multisector	0.149	0.116	0.090	0.071	0.023	0.6291	0.428

presented above.[39] In order to level the playing field and to keep the number of parameters to be estimated to a minimum, we assume K, F, σ_{id}, and σ_{sec} to be invariant across sectors. Hence, as before, we are left with four parameters to be estimated.

Table 5 compares targeted data moments with those produced by the partial information model with costly aggregate information and by its multisector extension. The multisector model also fits the data well and is not rejected by the test of overidentifying restrictions.

Estimated parameter values are displayed in Table 6. The volatility of sectoral shocks is estimated at 13% per year — 6.5 times the volatility of aggregate shocks. This ratio is similar in magnitude to the findings in McGrattan (2020), who estimates a multisector real business cycle model with aggregate and sectoral shocks to total factor productivity. Based on the estimates she provides, we obtain an average ratio of 5.3 across sectors.[40] When comparing parameter estimates with those of the single-sector model, the only noticeable difference is the reduction of the information cost. This leads to more frequent information gathering and processing, as can be inferred from panel a) in Figure 9, which shows inaction regions for both estimated models (see also Table 7).

When it comes to untargeted moments, the multisector model generates similar qualitative results as its one-sector counterpart. As seen in panel b) of Figure 9, the distribution of the duration of price spells generated by the multisector also resembles the empirical distribution reproduced in panel a) of Figure 8. The model also succeeds in producing realistic Coibion and Gorodnichenko (2015) regression coefficients (Table 4). In fact, it is the model that comes closest to their empirical estimates. Finally, the multisector model also generates higher kurtosis of the price change distribution than the partial information model with free aggregate information and the model without partial information (Table 7).

[39]The inflation drift and time-discount parameters are set to the same values used previously.

[40]To obtain this figure, we use estimates provided in Table 2 of McGrattan (2020). We compute the unconditional standard deviations of the sectoral and aggregate components of total factor productivity for each sector, take the ratio between the two, and average across sectors.

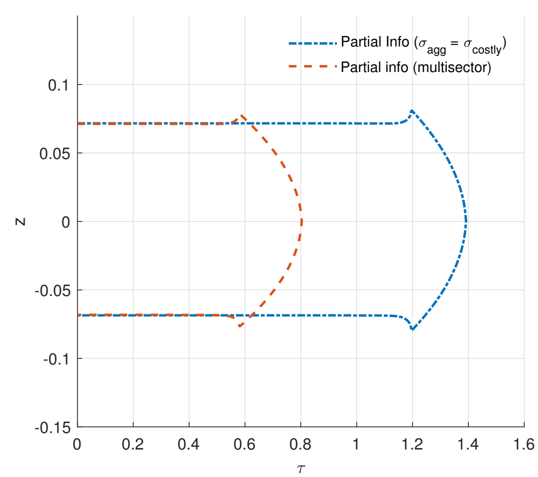
Table 6: Estimated parameters

	σ_{agg}	σ_{id}	K	F	σ_{sec}
Partial $\sigma_{agg} = \sigma_{cost}$	0.149	0.128	0.00030	0.0188	-
t-statistic	18.36	83.36	49.91	6.03	-
Multisector	0.02	0.127	0.00030	0.00423	0.1300
t-statistic	-	75.61	53.77	5.53	26.11

obs: t-statistics based on standard errors obtained with the delta method

Figure 9: Optimal policy and distribution of price spells for multisector model

(a) Optimal policy — comparison with $\sigma_{cost} = \sigma_{agg}$



(b) Distribution of price spells

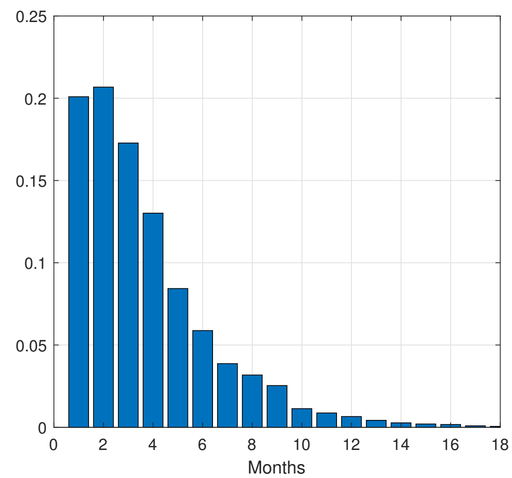


Table 7: Additional price-setting statistics implied by estimated models

Baseline	$\sigma_{cost} = \sigma_{agg}$	$\sigma_{cost} = \sqrt{\sigma_{sec}{}^2 + \sigma_{agg}{}^2}$
Price adj. per year	3.27	3.25
fully informed	16%	25%
partially informed	84%	75%
Info. gatherings per year	0.76	1.4
resulting in price adj.	69%	58%
Kurtosis of price change distribution	5.17	3.00

5 Real effects of nominal shocks

In this section, we study the real effects of nominal shocks in the estimated models. A firm's (log) frictionless optimal price equals the sum of (log) nominal aggregate demand and (log) idiosyncratic productivity. Log nominal aggregate demand m_t follows a Brownian motion with drift μ. For each estimated model, we simulate two million firms, starting from the ergodic steady state to which the economy converges after a sufficiently long spell without aggregate shocks. The (log) average price level p_t results from the aggregation of individual prices.

We follow the usual practice of analyzing the effects of a one-time monetary shock, which makes (log) nominal aggregate demand jump from m_0 to $m_0 + \zeta$ at time zero. After the shock, nominal aggregate demand resumes its trend, given by μ. Real (log) output, y_t, is given by:

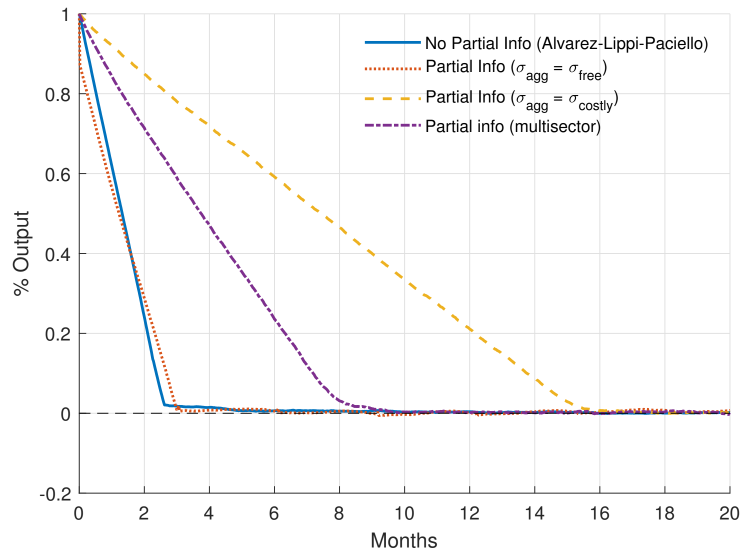
$$y_t = m_t - p_t.$$

5.1 Estimated models

Results for real output for the four models we estimate are shown in Figure 10. We simulate the effects of a 1% shock to nominal aggregate demand, and display impulse response functions (IRFs) for output, scaled by the size of the shock. Despite the fact that all models match the empirical frequency of price adjustments, they produce very different monetary non-neutrality. In the partial information model with free aggregate information and in the model with no partial information, output is back near its original level after approximately three months. In contrast, in the one-sector 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 much less frequently. The same intuition holds for the multisector model with costly sectoral and aggregate information, although real effects fade faster in this case, given the higher frequency of information gathering and processing.

To explore the roles of menu costs and information costs in producing the results reported in Figure 10, Figure 12 shows how changing K and F affects the extent of monetary non-neutrality in the model with costly aggregate information. Varying the menu cost — halving or doubling — around the estimated values has hardly any effect on aggregate dynamics (panel a)). In contrast, varying the information cost leads to meaningful differences in monetary non-neutrality (panel b)) — a result in line with the findings of Alvarez, Lippi,

Figure 10: Real effects of nominal shocks in estimated models



and Paciello (2018).

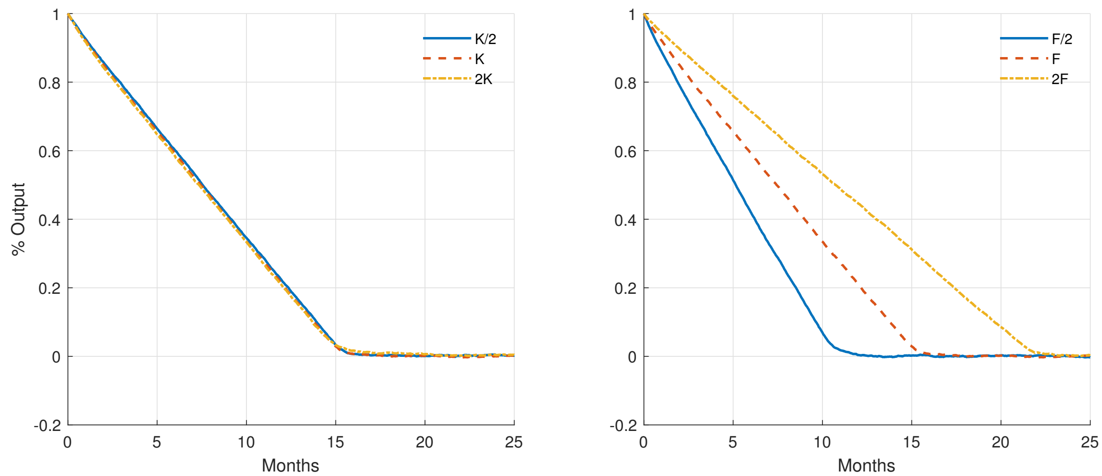
In principle, the model with no partial information and the model in which aggregate information is free are also capable of generating sizable monetary non-neutrality — this only requires larger price-setting frictions. When disciplined by the microdata, however, these two models imply a small degree of non-neutrality. The main difference between the estimated frictions in these two models and in the model 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 (e.g. Mankiw and Reis 2002).

5.2 Comparison to benchmark models

To further highlight the ability of our partial information model to generate persistent monetary non-neutrality, we compare it with three sticky price models commonly used in the literature: the Golosov-Lucas (GL) menu-cost model, the GL model with fat-tailed idiosyncratic shocks, and the Calvo model. The GL model is known to produce strong "selection effects" in price adjustment that lead to little monetary non-neutrality. Fat-tailed shocks were first introduced in the price-setting literature by Midrigan (2011), and are known to decrease selection and thus to generate stronger non-neutrality. The Calvo model has no price selection and is known to produce the maximum amount of monetary non-neutrality in a large class of sticky-price models (Carvalho and Schwartzman 2015, Alvarez, Le Bihan,

Figure 11: Monetary non-neutrality with costly aggregate information: varying K and F.

(a) Varying menu cost K (b) Varying information cost F



and Lippi 2016).

The three models are calibrated to generate the same frequency and average size of price adjustment as in our data. The GL model with fat-tailed shocks is calibrated to generate a kurtosis of price adjustments of 5.2,[41] which is the same as in our partial information model with costly aggregate information.

Results are displayed in Figure 12. As expected, the Golosov-Lucas model generates very little monetary non-neutrality, with output almost back at pre-shock levels after only two months. The GL model with fat-tailed shocks calibrated to generate the same kurtosis as the estimated partial information with costly aggregate information yields larger monetary non-neutrality, but not nearly as much as the partial information model. Even in the Calvo model, monetary non-neutrality is still less than half that of the partial information model.[42] In conclusion, the partial information model with costly aggregate information not only fits the microdata well, but also generates large and persistent monetary non-neutrality.

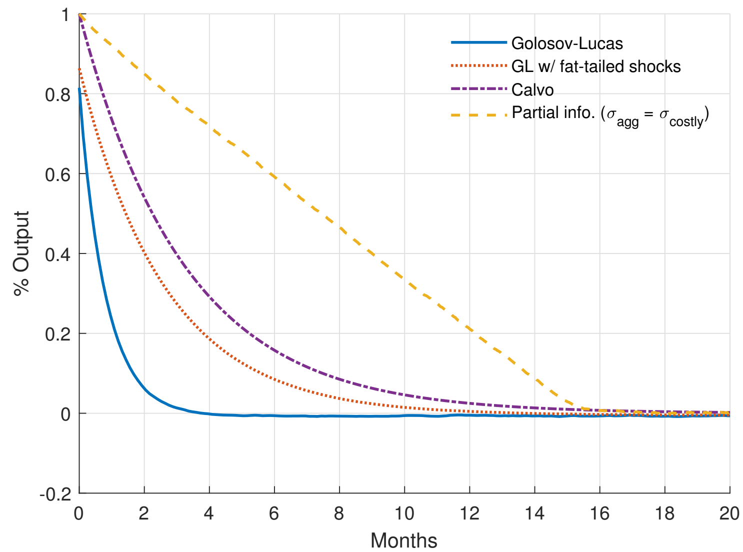
5.3 Kurtosis and monetary non-neutrality

Table 8 reports a measure of monetary non-neutrality — the area under the IRF of output in response to the monetary shock — and the kurtosis of the distribution of price changes

[41] To generate fat tails, we assume idiosyncratic shocks follow a Poisson process with a fixed arrival rate. Conditional on arriving, shocks follow a Laplace distribution.

[42] Another commonly used device to increase monetary non-neutrality in menu cost models is the introduction of random free adjustments, as in Nakamura and Steinsson (2010). The output response in such a model, however, would still be bounded above by the response of a Calvo model with the same frequency of price adjustments.

Figure 12: Real effects of monetary policy shocks in estimated models



for the four models analyzed in the previous section. The areas under the IRFs are reported relative to that of the GL model, which is normalized to unity.

Relative to the GL model, the Calvo model displays six times as much monetary non-neutrality and kurtosis. This follows from Alvarez, Le Bihan, and Lippi (2016), who show in a large class of models that, given the frequency of price adjustments, the kurtosis of the price change distribution is a sufficient statistic for monetary non-neutrality in the context of a small monetary shock.

While Calvo and GL are included in the class of models for which Alvarez, Le Bihan, and Lippi (2016) prove the sufficient statistic result, the GL model with fat-tailed shocks and the partial information model are not. Our results show that the mapping between kurtosis and monetary non-neutrality ceases to apply when we take these two additional models into account.[43,44]

Adding fat-tailed shocks to the GL model so as to generate the same kurtosis as in our estimated model (kurtosis of 5.2) increases monetary non-neutrality less than proportionately, by a factor of 4.2 (Table 8). More strikingly, the partial information model with costly aggregate information generates less kurtosis than the Calvo model, but more than twice as much monetary non-neutrality. In other words, when the partial information model is taken into account, the relationship between kurtosis and monetary non-neutrality ceases to be monotonic.[45]

[43] Notice that the frequency of price changes is the same in all four models that we analyze in this section.

[44] Dotsey and Wolman (2020) study different parametrizations of a menu-cost model and also find that Alvarez, Le Bihan, and Lippi (2016)'s sufficient statistic result does not hold in their model.

[45] We conjecture that the result of Alvarez, Le Bihan, and Lippi (2016) does not hold in our model be-

Table 8: Kurtosis and monetary non-neutrality in selected models

	Kurtosis of price changes	Monetary non-neutrality
Golosov Lucas	1	1
GL w/ fat-tailed shocks	5.2	4.2
Calvo	6	6
Partial information model	5.2	13.6

Note: Monetary non-neutrality measured as the area below the impulse response function of output with respect to a monetary shock, normalized so that non-neutrality in the GL model equals 1.

6 Conclusions

If each and every price change is effective in offsetting monetary shocks, nominal price rigidity cannot be the main source of monetary non-neutrality. This arises from the fact that, in the data, prices change too frequently to account for the sluggish response of aggregate prices and output to monetary shocks. For that reason, the sticky-price literature devotes substantial attention to mechanisms that contribute to mute the response of individual prices to monetary shocks.

Real rigidities, in the sense of Ball and Romer (1990), are one such mechanism. Strong real rigidities tend to induce strategic complementarities in pricing decisions, which, in the presence of staggered price setting, lead to partial adjustment of individual prices to monetary shocks.[46]

Information frictions can also generate persistent real effects of monetary shocks if they prevent individual price changes from fully reflecting monetary innovations. Matching the evidence on frequent and large price changes, however, requires that information be available about other shocks, to which firms can react.

In our model, firms have continuous information about idiosyncratic shocks that they can factor into pricing decisions at no cost. Information about monetary shocks, on the other hand, is costly to gather and process. This allows the model to generate frequent, large individual price changes, and, at the same time, large and persistent monetary non-neutrality.

cause most price changes reflect idiosyncratic shocks only, whereas price changes that matter for monetary non-neutrality are those that incorporate information about monetary shocks. Thus the mapping between unconditional price-setting statistics and monetary non-neutrality breaks down.

[46]Not all sources of real rigidity, however, are consistent with large individual price changes. For a thorough discussion of this issue, see Nakamura and Steinsson (2010).

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2021. “Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach.” Forthcoming in The Review of Economic Studies.

Almeida, Heitor and Marco Bonomo. 2002. “Optimal state-dependent rules, credibility and inflation inertia.” *Journal of Monetary Economics* 49:1317–1336.

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi. 2016. “The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach.” *American Economic Review* 106 (10):2817–2851.

Alvarez, Fernando and Francesco Lippi. 2014. “Price Setting with Menu Cost for Multiproduct Firms.” *Econometrica* 82 (1):89–135.

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

———. 2016. “Monetary shocks in models with inattentive producers.” *Review of Economic Studies* 83:421–459.

———. 2018. “Monetary shocks in models with observation and menu costs.” *Journal of the European Economic Association* 16 (2):353–382.

Baley, Isaac and Andres Blanco. 2019. “Firm Uncertainty Cycles and the Propagation of Nominal Shocks.” *American Economic Journal: Macroeconomics* 11 (1):276–337.

Ball, Laurence and David Romer. 1990. “Real rigidities and the non-neutrality of money.” *Review of Economic Studies* 57 (2):183–203.

Bils, Mark and Peter J. Klenow. 2004. “Some Evidence on the Importance of Sticky Prices.” *Journal of Political Economy* 112:947–985.

Bonomo, Marco and Carlos Carvalho. 2004. “Endogenous Time-Dependent Rules and Inflation Inertia.” *Journal of Money, Credit and Banking* 36 (6):1015–1041.

———. 2010. “Imperfectly-Credible Disinflation under Endogenous Time-Dependent Pricing.” *Journal of Money, Credit and Banking* 42 (5):799–831.

Bonomo, Marco, Carlos Carvalho, and Rene Garcia. 2011. “State-dependent Pricing under Infrequent Information: A Unified Framework.” Federal Reserve Bank of New York Staff Reports 455.

Bonomo, Marco, Carlos Carvalho, Oleksiy Krtytsov, Sigal Ribon, and Rodolfo Rigato. 2020. "Multi-Product Pricing: Theory and Evidence from Large Retailers." Mimeo.

Bonomo, Marco and Rene Garcia. 2001. "The Macroeconomic Effects of Infrequent Information with Adjustment Costs." *Canadian Journal of Economics* 34 (1):18–35.

Caballero, Ricardo. 1989. "Time Dependent Rules, Aggregate Stickiness and Information Externalities." Working Paper no. 428, Columbia University.

Calvo, Guillermo A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12 (3):383–398.

Carvalho, Carlos and Felipe Schwartzman. 2015. "Selection and Monetary Non-Neutrality in Time-Dependent Pricing Models." *Journal of Monetary Economics* 76:141–156.

Coibion, Olivier and Yuriy Gorodnichenko. 2012. "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy* 120 (1):116–159.

———. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8):2644–78.

DeJong, David N. and Chetan Dave. 2011. *Structural Macroeconometrics*. Princeton University Press.

Dixit, Avinash. 1991. "Analytical Approximations in Models of Hysteresis." *The Review of Economic Studies* 58 (1):141–151.

Dixit, Avinash J. 1993. *The Art of Smooth Pasting*. Routlege.

Dotsey, Michael and Alexander L. Wolman. 2020. "Investigating Nonneutrality In A State-Dependent Pricing Model With Firm-Level Productivity Shocks." *International Economic Review* 61 (1):159–188.

Gertler, Mark and John Leahy. 2008. "A Phillips Curve with an Ss Foundation." *Journal of Political* 116:533–572.

Golosov, Mikhail and Robert E. Lucas. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115:171–199.

Gorodnichenko, Yuriy. 2008. "Endogenous information, menu costs and inflation persistence." *Mimeo* .

Hobijn, Bart and Fernanda Nechio. 2018. "Sticker Shocks: Using VAT Changes to Estimate Upper-Level Elasticities of Substitution." *Journal of the European Economic Association* 17 (3):799–833.

King, Robert G. and Alexander L. Wolman. 1999. "What Should the Monetary Authority Do When Prices are Sticky?" *in Taylor, John B., ed. Monetary Policy Rules* :349–398Chicago: University of Chicago Press.

Klenow, Peter J. and Oleksiy Kryvtsov. 2008. "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *Quarterly Journal of Econoimcs* 123 (3):863–904.

Klenow, Peter J. and Jonathan L. Willis. 2007. "Sticky information and sticky prices." *Journal of Monetary Economics* 54 (S):79 – 99.

———. 2016. "Real Rigidities and Nominal Price Changes." *Economica* 83:443–472.

Knotek, Edward. 2010. "A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment." *Journal of Money, Credit and Banking* 42:1543–1564.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie. 2001. "Beyond Balanced Growth." *The Review of Economic Studies* 68 (4):869–882.

Mackowiak, Bartosz and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99 (3):769–803.

Mankiw, Gregory N. and Ricardo Reis. 2002. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (4):1295–1328.

Matejka, Filip. 2015a. "Rigid Pricing and Rationally Inattentive Consumer." *Journal of Economic Theory* 158:656–678.

———. 2015b. "Rationally inattentive seller: Sales and discrete pricing." *Review of Economic Studies* 83:1125–1155.

McGrattan, Ellen R. 2020. "Intangible capital and measured productivity." *Review of Economic Dynamics* 37:S147–S166.

Midrigan, Virgiliu. 2011. "Menu Costs, Multi-Product Firms and Aggregate Fluctuations." *Econometrica* 79:1139–1180.

Moscarini, Giuseppe. 2004. "Limited Information Capacity As a Source of Inertia." *Journal of Economic Dynamics and Control* 28:2003–2035.

Nakamura, Emi and Joń Steinsson. 2010. "Monetary Non-neutrality in a Multi-Sector Menu Cost Model." *Quarterly Journal of Economics* 125 (3):961–1013.

Oksendal, Bernt. 2000. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag.

Reis, Ricardo. 2006. "Inattentive Producers." *Review of Economic Studies* 73 (3):793–821.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3).

Stevens, Luminita. 2019. "Coarse pricing policies." *Review of Economic Studies* 87 (1):420–453.

Stokey, Nancy. 2008. *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. Princeton University Press.

Wilmott, Paul, Jeff Dewine, and Sam Howison. 1994. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press.

Woodford, Michael. 2009. "Information-constrained state-dependent pricing." *Journal of Monetary Economics* 56 (S):100–124.

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

Online Appendix

A General equilibrium model

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

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

subject to the budget constraints:

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

Utility is defined over the composite consumption good $C_t \equiv [\int_0{}^1 (C_{it}/A_{it})^{((\theta-1)/\theta)} di]^{(\theta/(\theta-1))}$ 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:

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

The demand for an individual variety is:

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

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.[47]

The static profit-maximizing price for firm i, P^*_{it} (also referred to as its frictionless optimal price), is given by the usual markup rule:

$$P^*_{it} = \frac{\theta}{\theta - 1}\frac{W_t}{A_{it}}.$$

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:

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

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.

B Profit function approximation

Here we derive the quadratic approximation to the static profit-maximization problem used in the main text. We omit time subscripts for conciseness. Write real flow profits as:

$$\begin{aligned}\Pi &= \frac{P_i}{P}C_i - \frac{W}{PA_i}C_i \\ &= A_i^{1-\theta}\left(\frac{P_i}{P}\right)^{1-\theta} C - \frac{W}{P}A_i^{-\theta}\left(\frac{P_i}{P}\right)^{-\theta} C,\end{aligned}$$

[47] More generally, assumptions relating preference and technology processes have been used previously in the literature on "balanced growth" in multisector models (e.g. Kongsamut, Rebelo, and Xie 2001).

where the second line uses the demand function derived in the previous section. Write P_i/P as $(P_i/P_i^*) \times (P_i^*/P)$ and use the definition of P_i^* to obtain

$$\begin{aligned} \Pi &= A_i^{1-\theta}\left(\frac{P_i^*}{P}\right)^{1-\theta}\left(\frac{P_i}{P_i^*}\right)^{1-\theta} C - \frac{\theta-1}{\theta} A_i^{-\theta}\left(\frac{P_i^*}{P}\right)^{-\theta}\left(\frac{P_i}{P_i^*}\right)^{-\theta} C \\ &= \left(\frac{\theta}{\theta-1}\right)^{1-\theta}\left(\frac{W}{P}\right)^{1-\theta} C\left[\left(\frac{P_i}{P_i^*}\right)^{1-\theta} - \frac{\theta-1}{\theta}\left(\frac{P_i}{P_i^*}\right)^{-\theta}\right]. \end{aligned}$$

Denote the real wage as $w = W/P$ and price gap as $x_i = \log P_i - \log P_i^*$. The profit function then becomes

$$\Pi\left(C, w, x_i\right) = \left(\frac{\theta}{\theta-1}\right)^{1-\theta} w^{1-\theta} C\left[e^{-(\theta-1)x_i} - \frac{\theta-1}{\theta} e^{-\theta x_i}\right].$$

We want to take a second-order approximation of Π around a frictionless steady state $(\bar{C}, \bar{w}, 0)$. Note that aggregate variables enter multiplicatively. Since, by definition, profits are maximized at $x_i = 0$, it is easy to check that

$$0 = \left.\frac{\partial \Pi}{\partial x_i}\right|_{x_i=0} = \left.\frac{\partial^2 \Pi}{\partial x_i \partial w}\right|_{x_i=0} = \left.\frac{\partial \Pi}{\partial x_i \partial C}\right|_{x_i=0}.$$

Therefore the second-order expansion, after some simplification, becomes

$$\Pi \approx \bar{\Pi} - \frac{1}{2}\bar{\Pi}\theta(\theta-1)x_i^2 + \text{terms independent of } x_i.$$

The terms independent of x_i depend only on the aggregate variables w and C, which the firm takes as given. Hence they do not affect the optimal pricing policy. After discarding them, we can express the loss function in terms of profit losses relative to the frictionless steady state as

$$L = \frac{\bar{\Pi} - \Pi}{\bar{\Pi}} = \frac{\theta(\theta-1)}{2} x_i^2 \propto x_i^2.$$

C Numerical solution

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:

$$z \approx n \triangle z,$$

$$\tau \approx m \triangle t,$$

$$V_\tau \approx \frac{v_{n,m+1} - v_{n,m}}{\triangle t},$$

$$V_z \approx \frac{v_{n,m} - v_{n-1,m}}{\triangle z},$$

$$V_{zz} \approx \frac{v_{n+1,m} - 2v_{n,m} + v_{n-1,m}}{(\Delta z)^2}.$$

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

$$\rho v_{n,m} = (n \triangle z)^2 + \sigma_c^2 m \triangle t + p^0 v_{n,m} + p^- v_{n-1,m} + p^+ v_{n+1,m} + \frac{1}{\triangle t} v_{n,m+1}, \tag{14}$$

where

$$p^0 = -\frac{\mu}{\triangle z} - \left(\frac{\sigma_f}{\triangle z}\right)^2 - \frac{1}{\triangle t},$$

$$p^- = \frac{\mu}{\triangle z} + \left(\frac{\sigma_f}{2\triangle z}\right)^2,$$

$$p^+ = \left(\frac{\sigma_f}{2\triangle z}\right)^2.$$

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 (14) 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.

D Optimal adjustment and smooth-pasting

brazilSome readers may wonder why smooth pasting conditions did not appear in our optimal price-setting problem. In fact, we implement optimal conditions as an inequality due to the option of adjusting or not, which implies that the firm will exercise the option of adjusting when it is optimal to do so. This generates a HJB variational inequality which can be proven to imply smooth pasting conditions (see Oksendal 2000). We also provide an heuristic argument adapted from Dixit (1993).

We implement optimality conditions in adjusment through the following condition:

english

$$V(z,\tau) \leq K + V(c(\tau),\tau), \tag{15}$$

where

$$c(\tau) = \arg\min_z V(z,\tau). \tag{16}$$

Recall that in the inaction region, the value function satisfies the HJB

$$\rho V(z,\tau) = z^2 + {\sigma_c}^2\tau - V_z(z,\tau)\mu + \frac{1}{2}{\sigma_f}^2 V_{zz}(z,\tau) + V_\tau(z,\tau). \tag{17}$$

Thus, equations (15) and (17) imply in the following condition:

$$\rho V(z,\tau) = \min\left\{ z^2 + {\sigma_c}^2\tau - V_z(z,\tau)\mu + \frac{1}{2}{\sigma_f}^2 V_{zz}(z,\tau) + V_\tau(z,\tau), \rho K + \rho V(c(\tau),\tau) \right\},$$

which can be rewritten as:

$$\max\left\{ \rho V(z,\tau) - z^2 - {\sigma_c}^2\tau + V_z(z,\tau)\mu - \frac{1}{2}{\sigma_f}^2 V_{zz}(z,\tau) - V_\tau(z,\tau), V(z,\tau) - \min_z V(z,\tau) - K \right\} = 0. \tag{18}$$

Equation (18) is called an HJB variational inequality. One can prove that (18) implies the following smooth-pasting conditions (see Oksendal 2000):

$$V_z(l(\tau),\tau) = V_z(u(\tau),\tau) = V_z(c(\tau),\tau) = 0.$$

Here we proceed with an heuristic approach adapted from Dixit (1993). The fact that the target point $c(\tau)$ minimizes the value function for a given τ implies:

$$V_z(c(\tau),\tau) = 0.$$

Now we show that optimal adjustment at the upper trigger point $u(\tau)$ implies

$$V_z(u(\tau)) = V_z(c(\tau), \tau).$$

Start from the fact that the value matching condition applies:

$$V(u(\tau), \tau) = V(c(\tau), \tau) + K. \tag{19}$$

Suppose $V_z(c(\tau), \tau) > V_z(u(\tau), \tau)$. Then, together with condition (19), this would make $V(z, \tau) < V(c(\tau), \tau) + K$ for all $z \in (u(\tau), u(\tau) + \varepsilon)$ for some small $\varepsilon > 0$, which would make inaction optimal at $u(\tau)$, contradicting optimal adjustment at $u(\tau)$.

Now suppose that $V_z(c(\tau), \tau) < V_z(u(\tau), \tau)$. Consider the discrete binomial process that approximates the Brownian motion process for z: z moves up by Δh with probability $p = \frac{1}{2}\left[1 - \frac{\mu}{\sigma^2}\Delta h\right]$, where $\Delta h = \sigma\sqrt{\Delta t}$, and down by $-\Delta h$ with probability $1-p$. Now, instead of the firm adjusting at $u(\tau)$, let it consider waiting for the next small time step Δt, and then revisiting the decision. If the increment is $+\Delta h$, adjust, and if it is $-\Delta h$, continue. We show next that the strategy of waiting is better if $V_z(c(\tau), \tau) < V_z(u(\tau), \tau)$, contradicting optimal adjustment at $u(\tau)$. By waiting, the firm gets:

$$\begin{aligned}
&\left(z^2 + \sigma^2\tau\right)\Delta t + (1 - \rho\Delta t)\left(p\left(V\left(c(\tau) + \Delta h, \tau + \Delta t\right) + K\right) + (1-p)\,V\left(u(\tau) - \Delta h, \tau + \Delta t\right)\right) \\
&= p\left(V(c(\tau), \tau) + V_z(c(\tau), \tau)\Delta h + V_\tau(c(\tau), \tau)\Delta t + K\right) + (1-p)\,V\left(u(\tau) - \Delta h, \tau + \Delta t\right) \\
&= \left(pV(u(\tau), \tau) + V_z(c(\tau), \tau)\Delta h\right) + (1-p)\,V\left(u(\tau) - \Delta h, \tau + \Delta t\right) \\
&= V(u(\tau), \tau) + pV_z(c(\tau), \tau)\Delta h - (1-p)\,V_z(u(\tau), \tau)\Delta h \\
&= V(c(\tau), \tau) + K + \frac{1}{2}\left[V_z(c(\tau), \tau) - V_z(u(\tau), \tau)\right]\Delta h \\
&< V(c(\tau), \tau) + K,
\end{aligned}$$

where we used value matching conditions, definition of p, and we have retained the leading terms of the expansion of order Δh (notice that Δt is of order Δh^2). Thus, the policy of waiting and adjusting in case a positive increment realizes, and continuing otherwise reduces the cost with respect to the policy of adjusting at $u(\tau)$. This rules out $V_z(c(\tau), \tau) < V_z(u(\tau), \tau)$. Thus the smooth pasting condition at the upper trigger point must be valid. A similar argument applies for showing that smooth pasting must be valid at the lower trigger point $l(\tau)$.