Price Rigidities in U.S. Business Cycles*

Camilo Morales-Jiménez
Federal Reserve Board

Luminita Stevens
University of Maryland

October 29, 2024

We provide a structurally estimated time series for the degree of nominal price rigidities in the United States between 1978 and 2023. To model the price rigidity, we allow for stochastic state dependence in both the timing of price changes and the choice of what price to charge. We give a cost-based micro-foundation to this stochasticity, modeling firms that face information costs and menu costs when making decisions regarding their prices. Estimating the model on time series of moments from the distribution of price changes over time–in addition to time series of real economic activity and inflation–we find considerable monetary non-neutrality with medium-cycle volatility. Underlying our estimated series are a number of results that shed new light on the sources and dynamics of price rigidity. In particular, the model attributes most of the rigidities in price setting not to infrequent, but rather to inaccurate price adjustment. The timing of adjustments has been accurate, especially for price decreases. Menu costs have been small while information frictions have been much larger and more volatile. This volatility has implications for the effectiveness of monetary stabilization policy.

*The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System. We thank our discussant Anton Nakov and colleagues at numerous institutions and conferences for helpful conversations. Contact: camilo.moralesjimenez@frb.gov, stevens7@umd.edu.

1 Introduction

How severe are price rigidities in the U.S. economy? Have prices become more responsive to shocks over time? Does price rigidity vary over the business cycle, complicating the stabilization efforts of the Federal Reserve? These questions are at the core of monetary economics. Rigidities in prices change how the economy adjusts in response to any shock, be it a supply shock or a demand shock. They also determine to what extent monetary policy can stabilize fluctuations in inflation and real economic activity, and they determine the welfare costs of inflation. Consider the question of a soft or a hard landing for the U.S. economy following the inflation surge of 2021-2022. The answer to that question depends in part on how flexibly prices adjust – first to the inflationary shocks themselves, and second to the interest rate increases undertaken by the Federal Reserve in its efforts to lower inflation. Despite the large literature measuring and modeling price rigidity, uncertainty and disagreement regarding the severity of price rigidities persist, reflecting both the difficulty of extracting model-free empirical estimates of price rigidity from the data, and the lack of clarity regarding what frictions are most relevant when modeling this rigidity.[1]

This paper provides a structurally estimated time series for the degree of nominal price rigidity (NPR) in the United States between January 1978 and March 2023. The key feature of the estimation is the use of a *generalized* model of pricing frictions that nests major potential sources and allows data on price dynamics over time to identify the severity of different potential frictions. The measure of NPR in each period is given by what the structural model predicts the cumulative response of consumption would have been in reaction to a monetary policy shock, given the pricing frictions we estimate for that period. To our knowledge, this is the first such structurally estimated time series. The time series for nominal rigidities gives us an estimate for monetary non-neutrality and the effectiveness of U.S. monetary policy over time, which we then use to reassess different episodes such as the 2021-2022 period. Indeed, we establish that pricing frictions had been falling precipitously since the second half of 2016, well before the 2020 pandemic-related shocks, and had already exceeded historical lows by January 2021, when

[1] We discuss these challenges further in the literature review.

CPI inflation was a mere 1.5%. Hence, it is not surprising that the initial inflationary period and the subsequent monetary tightening saw rapid price adjustments: inflation was poised to respond rapidly and flexibly to any shock and subsequent policy response regardless of the source.

The underlying pricing frictions themselves are obtained using a model of the aggregate economy that is estimated on time series of the moments from the distribution of price changes, in addition to time series of real economic activity, interest rates, and inflation. Incorporating the dynamics of the distribution of price changes in the estimation is a novel use of the pricing micro data, much like the literature on heterogeneous agent models has increasingly used household income and wealth distributional data to inform models of the aggregate economy (for example, Bayer, Born & Luetticke, 2020; Bilbiie, Primiceri & Tambalotti, 2023; Auclert, Rognlie & Straub, 2020, in the growing HANK literature). Previous work uses *steady state averages* of distributional moments to estimate steady state pricing frictions. Here instead, we use time series for these moments. The distributional moments are constructed from the micro price data underlying the U.S. Consumer Price Index (CPI). These series were created by Nakamura, Steinsson, Sun & Villar (2018) and extended to 2023 by Montag & Villar (2023).[2] The use of time variation in the distribution of price changes enables us to identify the nature of the pricing frictions and to characterize if and how these frictions change over time.

We model price frictions flexibly, allowing for the possibility of rigidity and errors in both the timing of price changes and the repricing itself (the choice of what price to set when adjusting). We give a cost-based micro-foundation to these frictions, modeling firms that face both information costs and menu costs, and thus nesting the two main ways of endogenizing nominal price rigidities. In this way, we depart from conventional quantitative models with nominal rigidities in which the standard assumption is that once firms get to reoptimize their price, they choose the optimal, deterministic, full information price (or price path), regardless of whether the nominal friction is modeled using Calvo, menu costs, or an observation or inattention cost (for example, Reis, 2006; Woodford, 2009).

In the model, firms choose how much attention to pay to market conditions given the cost of obtaining more information and the cost of revising their current

[2] We thank Daniel Villar for sharing the time series of key pricing moments with us.

price. Information acquisition is modeled using rational inattention (Sims, 2003; Woodford, 2009). Since changing prices is costly, firms first decide if they want to change their price. They do so based on an imprecise awareness of the state of the economy, as in Woodford (2009). How accurate their timing decisions are is a choice that firms make state by state, weighing the marginal benefit versus the marginal cost in each state. This generates infrequent price adjustment and an endogenous degree of state-dependence in the timing of price changes. Second, when they decide to change their price, firms then need to decide what price to set. Unlike in prior models with sticky prices, this repricing decision is also based on an imprecise awareness of market conditions. A firm will choose state by state how much to spend on information in order to set its reset price accurately, resulting in reset prices that deviate stochastically from the full information reset price. This generates state-dependent price dispersion even conditional on price adjustment and arises as an additional source of nominal rigidity, beyond the usual infrequent adjustment margin. The model spans pricing behavior from fully flexible to fully random in terms of both timing and price levels, and we let the estimation on U.S. data pin down the degree of inaccuracy along each margin. The degree to which firms tolerate errors in pricing interacts with their tolerance for errors in the timing of price changes, and we provide a discussion of this interaction and show that under certain conditions it can rationalize Calvo-like price-setting as an optimal way to economize on repricing costs.

The estimation yields a sizable degree of NPR: on average, a shock to the federal funds rate of 25 basis points (bp) yields a cumulative change in consumption equal to 0.12 percent of annual steady-state consumption. This represents approximately 80% of the response the Calvo model would predict when calibrated to the same frequency and size of price changes. Between 1978 and 2023, the degree of non-neutrality exhibits no clear trend. This is surprising since one might imagine that technology has made both information gathering and repricing less costly. We interpret this as suggestive of the complexity that goes hand in hand with technological progress and data abundance: more is not always easier (Veldkamp & Chung, 2024).

Despite no clear trend, we find substantial medium-cycle volatility in the degree of NPR over time. We estimate above-average rigidity during the 1990s and the

2010s, and below-average rigidity at the beginning and end of the sample. There are no systematic patterns during recessions, with rigidity falling in the 1990 and 2001 recessions, but rising in the recessions of the early 1980s and in the Great Recession. This finding casts some doubt on the hypothesis of increased price flexibility during recessions. Moreover, movements in the NPR, in either direction, precede the start of recessions by one-to-two quarters and often persist beyond the end of recessions. For example, rigidity started rising in early 2007 and continued to rise until late 2011, when it plateaued. In mid-2016, rigidities began a steady decline that continued until early 2022. The increased price flexibility that we estimate starting around 2016 is particularly interesting. In hindsight, it suggests that we might have expected any inflationary shocks, should they occur, to be met with a sharper inflation response, rather than a sharper output response post-2016 versus pre-2016. This may explain why inflation surged so rapidly in 2021, and it points to the value of using distributional pricing data in real-time.

Underlying our estimated NPR series are several results that shed new light on the sources and dynamics of nominal rigidity. These results depart meaningfully from the conventional wisdom embedded in standard DSGE models with nominal frictions. At the heart of these implications is the interaction between firms' timing decisions and their repricing choices.

First, we estimate substantial inaccuracy in the repricing decision. Moreover, firm-level mistakes do not average out with aggregation. Pass-through of marginal costs to prices is incomplete and, in the parlance of literature, firms only partially close their price gaps when adjusting. This challenges standard models that assume perfect repricing conditional on adjustment. Our finding of a relatively high degree of errors in pricing breaks the connection between *adjustment* and *flexibility*: Even if prices are not very "sticky," in the sense that they are changing over time, they nevertheless only partially respond to economic conditions. Underscoring this dichotomy, we estimate only a modest correlation between the frequency of price changes and the estimated NPR. In fact, we show that under some conditions, a higher frequency of price changes can be associated with a higher degree of NPR, a *paradox of flexibility.*

Second, the timing of price changes has been fairly accurate: We estimate a small cost of determining when price changes are warranted, which yields a strongly state-dependent probability of adjustment, especially for price cuts. It seems that firms are able to determine quite accurately when their prices have become obsolete, but they have more difficulty determining the right reset price. This finding poses a challenge to the conventional models of sticky prices in which non-neutrality is a function of the frequency of adjustment and the strength of state-dependence in the timing of adjustment. In these models, in order to get meaningful non-neutrality, as we see in the data, one needs to weaken the state dependence in the timing of price changes. But with mistakes in repricing, that is no longer necessary. In the terminology of Caballero & Engel (2007), repricing errors mute the selection on the intensive margin, reducing the need to rely on weak selection on the extensive margin.

Third, even though prices change infrequently, most of the inaction reflects uncertainty about the right price to set, rather than an unwillingness to pay the adjustment cost. We estimate a small menu cost that accounts for only a small fraction of both adjustment costs and the total degree of price rigidity. In our model, firms understand that they risk picking the wrong price, so they often choose to forgo price changes altogether. This uncertainty provides an important micro-foundation for inaction that is quantitatively significant: introducing even a modest degree of errors in pricing can halve the frequency of adjustment for a given menu cost.

Fourth, in the time series, we estimate significant volatility in pricing frictions, primarily driven by volatility in information costs. We find that volatility has been strongly correlated with measures of exogenous uncertainty, such as those of Jurado, Ludvigson & Ng (2015) and Ludvigson, Ma & Ng (2021): When fundamental aggregate uncertainty is low, firms are less attentive to the environment, and monetary non-neutrality rises. This inverse relationship was particularly pronounced in the 1990s and in the 2010s, when measures of uncertainty were persistently below average.

Fifth, mistakes in repricing also break the tight relationship between monetary non-neutrality and the kurtosis of the distribution of price changes. We have known since the work of Alvarez, Le Bihan & Lippi (2016) that kurtosis relative to the frequency of price changes yields a sufficient statistic for the degree of NPR in a wide range of sticky price models, from the Calvo model to various types of menu cost models. In short, in models with perfect repricing, a lower frequency implies more rigidity, while a higher kurtosis implies a combination of many large price changes coexisting with many small price changes. This variability of price changes in turn increases non-neutrality for a given frequency of adjustment. Errors in pricing weaken this tight link first and foremost because they break the link between frequency and flexibility, as discussed above. Moreover, to the extent that firms do not full close their price gaps when adjusting their prices, the distribution of price gaps, and in particular its tails cannot be directly proxied for using the distribution of actual price changes.

Inaccuracy in pricing is consistent with a large body of evidence that economic choices are based on dispersed beliefs and are imprecisely related to optima in many contexts. Many studies have documented dispersion in actions and forecasts *conditional on adjustment*, both in survey data and in incentivized controlled laboratory experiments, including Mankiw, Reis & Wolfers (2003); Carroll (2003); Coibion & Gorodnichenko (2012); Magnani, Gorry & Oprea (2016); Cavallo, Cruces & Perez-Truglia (2017); Khaw, Stevens & Woodford (2017); Angeletos, Huo & Sastry (2021), among many others. Additional, more direct evidence of imprecision in price adjustment comes from studies of disaggregated price data that have compared actual price changes with firms' price gaps, using either measures of firms' marginal costs or proxies based on competitors' prices (Gagliardone, Gertler, Lenzu & Tielens, 2023; Karadi, Schoenle & Wursten, 2024)

Allowing for the possibility of mistakes in repricing also implies that Calvo is also no longer an upper bound on the degree of nominal rigidity. Varying the severity of information frictions regarding the timing of price adjustment spans the degree of state dependence in price setting, with the menu cost model at one end (when the information friction approaches zero) and Calvo (1983) at the other end (when the information friction is strong enough that the firm acquires no information to decide when to adjust its prices), as shown by Woodford (2009). But this result applies to models featuring perfect repricing. Adding errors in the repricing itself adds another layer of nominal rigidity. As a result, the model can feature larger non-neutrality than a Calvo model parameterized to have the same frequency of price adjustment. Intuitively, conditional on a price change, firms' reset prices respond less to aggregate conditions when firms price more inaccurately, making prices more rigid.

Finally, we also want to highlight our estimation method, which contributes to the literature that has sought to introduce heterogeneity in DSGE models. To our knowledge, this is the first Bayesian estimation of a model with rationally inattentive firms, and the first application of the sequence-space Jacobian (SSJ) method of Auclert, Bardóczy, Rognlie & Straub (2021) to a model with heterogeneous information. Moreover, since our estimation sample includes two periods in which the effective lower bound (ELB) was binding on the federal funds rate, we also show how to handle occasionally binding constraints with SSJ, by adapting the methods proposed by Guerrieri & Iacoviello (2015) and Kulish, Morley & Robinson (2017).

2 Additional Related Literature

Two important precursors to our work are the control cost pricing model of Costain & Nakov (2019) and the inattentive forecasting model of Khaw et al. (2017). Khaw et al. (2017) model rationally inattentive adjustment in both the timing of adjustment and the choice of a new forecast for individual decision-makers tracking the realizations of a slow-moving random variable. The model is then estimated on individual data from a controlled laboratory experiment. Costain & Nakov (2019) model price-setting firms that are subject to control costs in timing and repricing, and they use steady state moments of the distribution of price changes to estimate the severity of control costs on average. Since the control costs introduce errors in pricing that are uniform around the optimal price and hence average out with aggregation, they play little to no role in aggregate dynamics.

More broadly, our results build on several strands of the literature. First, we build on work that has sought to use moments from the micro pricing data to develop micro-founded models of nominal rigidities. For example, while the frequency of price adjustment is a sufficient statistic for the canonical Calvo (1983) model, Alvarez et al. (2016) prove that frequency relative to kurtosis pins down non-neutrality in a wide class of menu cost models, Berger & Vavra (2018) argue for the additional relevance of the standard deviation of price changes, and Luo & Villar (2021) suggest also taking into account the skewness of price changes.

Empowered by detailed empirical analyses of micro pricing patterns starting with the seminal work of Bils & Klenow (2004), Nakamura & Steinsson (2008), and Campbell & Eden (2014), a wave of menu cost models (*e.g.*, Golosov & Lucas Jr (2007); Nakamura & Steinsson (2010); Midrigan (2011); Alvarez & Lippi (2014); Vavra (2013)) have studied the contribution to NPR of different moments of the price change distribution. There is, nonetheless, an ongoing debate concerning the informativeness of various pricing moments for the degree of NPR as well as the degree to which the degree of NPR varies over time, and whether or not it is procyclical. Our results underscore that the extent to which particular steady-state pricing moments pin down NPR remains quite model-dependent.

Second, our framework nests models that generate non-neutrality via **infrequent** price adjustment with those that generate non-neutrality via the **incomplete** response of individual prices to shocks. In the first category, our model belongs to the class of generalized Ss models of infrequent adjustment such as Dotsey, King & Wolman (1999), Caballero & Engel (2007), and Woodford (2009), in which the probability of adjustment varies smoothly with the value of adjusting. In the second category, our model belongs to the class of models with imprecise price-setting (Woodford, 2003), in particular work that operationalizes the imprecision using tools from information theory (Maćkowiak & Wiederholt, 2009; Matějka, 2015; Turen, 2023; Afrouzi, 2020; Afrouzi & Yang, 2021). The first group of models, which make a firm's nominal price sticky over time, assume perfect repricing: Once a firm has the opportunity to reprice, the newly chosen price is a deterministic, full-information optimal choice. The models in the second group remove the impediments to changing prices every period, but instead relax the assumption of perfect repricing. We nest these two cases and allow the estimation to speak to their relative importance in generating monetary non-neutrality.

Third, by allowing both information frictions and nominal adjustment frictions to play potentially distinct roles in generating nominal rigidity in response to shocks, our paper relates to work that bridges these two approaches to endogenizing pricing frictions: Angeletos & La'O (2009) and Nimark (2008) study the interaction between Calvo (1983) price-setting and dispersed information a la Woodford (2003), while Klenow & Willis (2007) models a sticky-information version of menu cost pricing. Melosi (2014) estimates that imperfect common knowledge a la (Woodford, 2003) fits U.S. inflation and output time series better than a model with Calvo frictions alone. Alvarez, Lippi & Paciello (2011) present a theoretical analysis of price adjustment in the presence of menu costs and (fixed) information costs a la Reis (2006), and they also emphasize the interaction between the two sources of nominal rigidity.

Fourth, we contribute to the strand of literature that has studied how NPR varies with inflation. Empirical work has shown that once inflation exceeds high single digits, price rigidity starts to decline with inflation, rapidly reaching near-flexibility (Alvarez, Beraja, Gonzalez-Rozada & Neumeyer, 2019; Gagnon, 2009) Our model generates this state-dependence through an endogenous increase in information acquisition.

Fifth, our paper also relates to work that seeks to estimate the severity of information frictions over time more generally, such as Coibion & Gorodnichenko (2015). Our work is also complementary to Carvalho, Dam & Lee (2020), who study the degree of real rigidities and heterogeneity in price stickiness.

3 Model

To provide a credible structural estimation of the degree of NPR over time, we need a model that can accommodate different types of uniquely identifiable pricing frictions. This section presents such a model, in which we allow for errors in both the timing of price adjustment and the repricing decision. Information costs are the source of these errors and, together with menu costs, generate noisy, infrequently updated prices. We place the information and adjustment frictions on monopolistically competitive retailers while retaining the assumption of full information, flexible adjustment for other agents in the economy. We first describe the retailers problem, and then close the economy with a representative household, competitive intermediate goods producers, and fiscal and monetary authorities.

3.1 Monopolistically Competitive Retailers

A continuum of retailers j sell differentiated varieties and are monopolistically competitive price-setters in their product market and competitive price-takers in the market for their production input. They optimize subject to both information costs and menu costs.

Operating Profits Each retailer's demand is

$$y_{jt} = p_{jt}^{-\varepsilon_t} Y_t, \tag{1}$$

with $\varepsilon_t > 1$ denoting the elasticity of substitution, which is potentially time-varying, Y_t denoting final aggregate demand, and $p_{jt} = P_{jt}/P_t$ denoting the good's relative price, with

$$\left(\int p_{jt}^{1-\varepsilon_t} dj \right)^{\frac{1}{1-\varepsilon_t}} = 1. \tag{2}$$

The retailers' production function is

$$y_{jt} = e^{a_{jt}} x_{jt}, \tag{3}$$

where a_{jt} is an AR(1) process for idiosyncratic productivity and x_{jt} is the homogeneous input. Given its price, the retailer purchases whatever quantity of the intermediate good is needed to satisfy demand at that price.

Real operating profit per period is

$$\pi_{jt}^r = p_{jt} y_{jt} - p_t^x x_{jt} = \left[p_{jt}^{1-\varepsilon} - p_{jt}^{-\varepsilon} \left(\frac{p_t^x}{e^{a_{jt}}} \right) \right] Y_t, \tag{4}$$

where $p_t^x = P_t^x / P_t$ is the intermediate input's relative price.

Information Costs Firms are rationally inattentive to market conditions (Sims, 2003). They are rational, in that they optimize based on a complete understanding of the structure of their environment (payoff functions, shock processes, markets), but they must expend resources to learn the realizations of stochastic variables in real time.

Information acquisition is modeled as a choice that can be quantified and optimized using tools from information theory (Shannon, 1948, 1959). However, we assume that rather than being endowed with a fixed information capacity (Sims, 2003), firms can choose how much information to obtain, subject to a variable cost, as in Woodford (2009).

In our context, given the fixed menu cost of price adjustment, in each period firms must decide whether or not to update their price, and if so, what price to set. As is common in the RI literature, we assume that the costs of making the adjustment and pricing decisions contingent on the realized states are linear in the information acquired in order to make each decision,

$$\mathcal{C}^a_{jt} = \theta^a \, \mathcal{I}^a_{jt} \qquad \text{and} \qquad \mathcal{C}^p_{jt} = \theta^p \, \mathcal{I}^p_{jt}, \tag{5}$$

where θ^a is the unit cost of making a more informed decision about whether or not to change prices, θ^p is the unit cost of making a more informed price choice when adjusting prices, and $\mathcal{I}^a_{jt}$ and $\mathcal{I}^p_{jt}$ measure how much information is acquired for each decision (which we discuss further below). We allow for (but do not impose) potentially different unit costs, since they may reflect different managerial marginal costs of attention. If θ^a and θ^p are zero, the firm's problem collapses to a full information menu cost model.

Value of the Firm Firms acquire information and make pricing decisions to maximize

$$\mathcal{V}_0 = \left\{ E_{j0} \sum_{t=0}^{\infty} M_{0,t} \left[\pi^r_{jt} - \mathcal{C}^a_{jt} - \delta_{jt} \, \left(\kappa + \mathcal{C}^p_{jt} \right) \right] \right\}, \tag{6}$$

subject to (4) and (5), where $M_{0,t}$ is the stochastic discount factor used to discount real profit streams from date t to date 0, δ_{jt} is an indicator equal to 1 if the firm picks a new price in period t and 0 otherwise, and κ is the fixed cost of repricing. If the firm does not change its price in the period, it continues with its existing nominal price (there is no automatic indexation of prices to inflation).

Acquiring Information A firm's choice of how much information to obtain amounts to choosing how much its decisions condition on each realized state, relative to the best decisions the firm could make based on beliefs it has for free.

For each decision, the amount of information acquired is measured by Shannon's mutual information. For the adjustment decision, this is given by

$$\mathcal{I}^a_{jt} = E_t\left\{\mathcal{D}\Big(\Lambda_{jt} \parallel \bar{\Lambda}\Big)\right\}, \tag{7}$$

$$\mathcal{D}\big(\Lambda \parallel \bar{\Lambda}\big) = \Lambda \ln\left(\frac{\Lambda}{\bar{\Lambda}}\right) + \big(1-\Lambda\big)\ln\left(\frac{1-\Lambda}{1-\bar{\Lambda}}\right), \tag{8}$$

where Λ_{jt} denotes the probability that the firm adjusts its price in period t, after obtaining information about the realized state, $\bar{\Lambda}$ is the *reference probability* of adjustment, based on the firm's beliefs *before* obtaining current information, $\mathcal{D}$ is the Kullback-Leibler (KL) divergence of the choice distribution from the reference distribution, and expectations integrate over the joint distribution of idiosyncratic and aggregate states that the firm could face in period t.[3] Hence, the contribution to the firm's cost of conditioning the adjustment decision on a period's realized state is proportional to the probability of the firm finding itself in that state times the divergence of Λ from $\bar{\Lambda}$ in that state. The trade-off facing the firm captures the fact that the more Λ conditions on the realized state, the more it deviates from $\bar{\Lambda}$, and hence the higher is its cost. It also reflects the fact that all else equal, paying attention to more frequent states will cost more.

Analogously, for the pricing decision, the amount of information obtained in order to decide *what price* to set is

$$\mathcal{I}^p_{jt} = E_t\left\{\mathcal{D}\Big(f_{jt}(p) \parallel \bar{f}(p)\Big)\right\}, \tag{9}$$

$$\mathcal{D}\left(f \parallel \bar{f}\right) = \int f(p)\ln\left(\frac{f(p)}{\bar{f}(p)}\right) dp, \tag{10}$$

where $f_{jt}(p)$ is the probability that the firm sets its price equal to p conditional on the information it acquires about the realized state, and $\bar{f}(p)$ is the reference probability of setting the price equal to p, based on the firm's beliefs about the right price to set prior to obtaining current information. As is the case for the adjustment decision, $\mathcal{D}$ is the KL divergence of the choice distribution from the

[3]The KL divergence gives a measure of how "far off" one would be, on average, if they assumed the first distribution when the true distribution were in fact the second distribution. Shannon's mutual information between two random variables x and y is the KL divergence of the joint distribution from the product of the marginal distributions.

reference distribution, and expectations integrate over the joint distribution of idiosyncratic prices, productivities, and aggregate states. Hence, the contribution to the total information flow of conditioning the pricing decision in a period on that period's state is equal to the probability of the firm finding itself in that state times the divergence of f_{jt} in that state from $\bar{f}$.

Reference Distributions Let $\Lambda_{ss}(p, a)$ denote the steady-state probability of adjustment of a firm with price p and idiosyncratic productivity a. We assume the firms' reference probability of adjustment $\bar{\Lambda}$ is the equilibrium frequency of adjustment in the steady state,

$$\bar{\Lambda} = \int \Lambda_{ss}(p, a)\, \tilde{\Omega}_{ss}(p, a)\, da\, dp, \tag{11}$$

which integrates the adjustment probability over the (endogenous) steady-state joint distribution of firm prices and productivities, $\tilde{\Omega}_{ss}$, before price review decisions have been made, but after the idiosyncratic shocks have been realized.

Similarly, the reference probability of charging each p in the set of possible prices is the steady state distribution of prices, after adjustments have been made. Letting $f_{ss}(p|a)$ denote the steady state probability with which a firm with idiosyncratic productivity a sets price p when adjusting, the reference distribution for prices integrates over idiosyncratic states:

$$\bar{f}(p) = \int f_{ss}(p \,|\, a)\, \Omega_{ss}(p, a)\, da, \tag{12}$$

where Ω_{ss} is the (endogenous) joint distribution of productivities and post-adjustment prices.

Discussion of Reference Distributions The assumption that firms use equilibrium distributions as their reference is motivated by the idea, plausible to us, that decision-makers with prior experience across a range of states may find it "easier" to have as references rules that they have observed work well on average, across many states.

By constraining the reference distributions to be the cross-sectional equilibrium distributions, we are using a slightly inefficient information structure relative to the pure RI solution. How does it compare to some possible alternatives?

One alternative is the control cost (CC) model of stochastic choice. This model would imply a reference distribution for each period that is uniform around the optimal action in that period. Choosing a more concentrated distribution entails a cost proportional to the divergence of the chosen distribution from the uniform. As the name suggests, these are models of costly control, rather than costly information. Such models allow for *unbiased* errors in the implementation of actions when the optimal action is known in each state, generating volatile, noisy actions. For example, Costain & Nakov (2019) apply control costs to price-setting in a general equilibrium monetary model.

On the other hand, RI models of costly information imply an endogenous reference distribution that is optimal given the decision-problem at hand. Note that rational decision-makers have strong incentives to develop sophisticated reference or default probabilities. A well-chosen default distribution can lower both the relative value of conditioning actions on the state in real time, as well as the cost of doing so. Hence, a rational decision-maker would want to use knowledge about the structure of the economy, the laws of motion of the shocks, and the shape of the objective function to choose well-adapted reference distributions that can serve as no-cost defaults.

What does a well-adapted no-cost default look like? In the RI model, it is one that gets as close as possible to conditioning on the state in real time, without actually doing so. Formally, the optimal reference distribution minimizes the choice distribution's average KL divergence from it, integrating over the distribution of possible states of the world that the decision-maker can expect to encounter.

Here we are using less efficient, though still endogenous reference distributions that take into account both the structure of the economy and the actions of others. It is in this sense that the model is behavioral RI model.[4]

Recursive Formulation We now define the firm's problem recursively and solve for each element of the optimal policy.

[4]See Woodford (2012) and Khaw, Stevens & Woodford (2019) for alternative deviations of the default distributions from the RI optima.

First, consider the choice of the adjustment probability $\Lambda_t\left(\tilde{p}, a\right)$ for a firm that begins period t with real price $\tilde{p}$ and idiosyncratic productivity a. This choice solves:

$$V_t^*(\tilde{p}, a) = \max_{\Lambda_t} \left\{ \Lambda_t \cdot \left[V_t^a(a) - \kappa\right] + (1 - \Lambda_t) \cdot V_t(\tilde{p}, a) - \theta^a \, \mathcal{D}\left(\Lambda_t \parallel \bar{\Lambda}\right) \right\}, \tag{13}$$

where the subscript t indicates dependence on the aggregate state and we have suppressed the arguments of the adjustment probability to ease notation. The retailer either adjusts to a new price, with probability $\Lambda_t\left(\tilde{p}, a\right)$, or continues with its current price, which occurs with probability $1 - \Lambda_t\left(\tilde{p}, a\right)$. In either case, it pays the cost of conditioning this period's adjustment probability on this period's state. This is the last term in the maximand.

If the firm continues with its existing price, it obtains $V_t(\tilde{p}, a)$, which consists of the flow operating profit at this price plus the expected discounted continuation value of entering the next period with this price,

$$V_t(p_{jt}, a_{jt}) = \pi_t(p_{jt}, a_{jt}) + E_t \left\{ M_{t,t+1} \, V_{t+1}^* \left(\tilde{p}_{j,t+1}, a_{j,t+1}\right) \right\}, \tag{14}$$

where expectations condition on the current state, $M_{t,t+1}$ is the real discount factor between the two periods, $\tilde{p}_{j,t+1} = p_{jt} P_t / P_{t+1}$ is the real price at the beginning of the next period given the current-period real price p_{jt} and the aggregate price levels, and V_{t+1}^* is the maximum attainable value the firm can expect in the next period (assuming optimal choices henceforth), which takes the form of equation (13).

If, instead, the firm adjusts its price, it pays the menu cost κ and can expect to obtain $V_t^a(a)$, the expected value under the optimal pricing policy, net of the information cost associated with deviating from the reference distribution in this state,

$$V_t^a(a) = \max_{f_t} \left\{ \int f_t(p \,|\, a) V_t(p, a) \, dp - \theta^p \mathcal{D}\left(f_t(p \,|\, a) \parallel \bar{f}(p)\right) \right\} \tag{15}$$

$$\text{s.t.} \quad \int f_t(p \,|\, a) \, dp = 1. \tag{16}$$

Optimal Choice Distributions The optimality condition for the choice of $\Lambda_t(\tilde{p}, a)$ equates the marginal value of a more accurate adjustment decision to its marginal cost, state by state. This yields an expression for the optimal log odds of adjustment given by

$$\ln\left(\frac{\Lambda_t(\tilde{p},a)}{1-\Lambda_t(\tilde{p},a)}\right) = \ln\left(\frac{\bar{\Lambda}}{1-\bar{\Lambda}}\right) + \frac{1}{\theta^a}\left[V_t^a(a) - V_t(\tilde{p},a) - \kappa\right]. \tag{17}$$

The model predicts a linear relationship between the conditional log odds, the unconditional log odds, and the net gain from adjusting the firm's price, with the unit cost θ^a governing the sensitivity of the adjustment decision to the value of adjusting. We emphasize that the firm can choose how much to deviate from the reference distribution state by state, which means that the firm can have state-dependent accuracy if the value of having more accurate timing of adjustments differs across states. The contribution to the total information flow of conditioning the adjustment decision on a period's state is equal to the probability of the firm finding itself in that state times the KL divergence of $\Lambda_t(\tilde{p}, a)$ from $\bar{\Lambda}$,

In the limit, $\theta^a \to \infty$ implies a constant probability of adjustment, as in the Calvo model, while as $\theta^a \to 0$ the adjustment decision converges to the deterministic Ss adjustment rule. For intermediate values of this cost, the adjustment probability is stochastically state-dependent, as in the information model of Woodford (2009) and as in the random menu costs model of Dotsey et al. (1999).

Now consider the optimal choice for the probability of charging a particular price in a particular state. This choice too can be made independently for each state and satisfies

$$f_t(p \mid a) = \frac{\bar{f}(p)\,\exp\left\{\frac{V_t(p,a)}{\theta^p}\right\}}{\int \bar{f}(\hat{p})\,\exp\left\{\frac{V_t(\hat{p},a)}{\theta^p}\right\}\,d\hat{p}} \tag{18}$$

for each p charged with positive probability in the steady state.

A price is charged with a higher probability in a particular state if it yields a higher value in that state compared with the average value across all possible prices under the reference distribution. The value of deviating from the reference must be high enough to compensate for the increase in information expenditure.

A lower attention cost θ^p enables a finer differentiation across states. As the cost approaches zero, the firm's repricing approaches a degenerate distribution for each state, centered on the optimal full-information reset price. Conversely, as θ^p increases, the firm differentiates pricing less and less across states and increasingly relies on the reference distribution. It does so first in states in which its continuation value is not too price sensitive, and eventually across all states.

Given the reference distributions, the optimal choice distributions are determined by equations (17) and (18). This gives us an optimization-based approach to generalizing the menu cost model to a stochastic version. Entropy reduction generates stochastic decisions: shrinking uncertainty to a degenerate distribution is often too costly, so the decision-maker is left with some residual uncertainty about the optimal course of action. In our context, this means that the firm acts probabilistically both in its decision about whether or not to change its price and in its decision about which price to charge. But the degree of randomness in choice is the result of a cost-benefit analysis and, as long as information is not infinitely costly, the firm will be more likely to adjust when the value of adjusting is higher and more likely to set a particular price when its continuation value at that price is higher compared with other possible prices. Moreover, the firm can specify the accuracy of its decisions in each state. For instance, in some states, it may not be worthwhile to expend resources on very precise information about market conditions, since perhaps in those states, the firm's payoffs are not very sensitive to having the correct price in place. On the other hand, other states of the world may make mispricing very costly, in which case the firm will want to condition its decisions more strongly on those states and be willing to pay the extra cost associated with that accuracy. We explore this state dependence in more detail in Section 4.4.

Price Distributions The law of motion for the joint distribution of prices and idiosyncratic states after all pricing decisions have been made is given by

$$\Omega_t(p,a) = [1-\Lambda_t(p,a)] \cdot \tilde{\Omega}_t(p,a) + \left[\int \Lambda_t(\hat{p},a)\,\tilde{\Omega}_t(\hat{p},a)\,d\hat{p}\right] \cdot f_t(p\,|\,a), \qquad (19)$$

where $\tilde{\Omega}_t(p,a)$ is the joint distribution at the beginning of the period, before any pricing decisions have been made, but after the realization of all shocks in the period. Hence, $\tilde{\Omega}_t(p,a)$ is given by last period's post-adjustment distribution ($\Omega_{t-1}(p,a)$) with all real prices eroded by inflation and idiosyncratic states transitioned to new values given the realization of period-t shocks.

In (19), the first term captures the mass of firms that start the period in state $p \times a$ and do not adjust, while the second term captures the mass of firms with idiosyncratic state a that adjust from any price $\hat{p}$ to end up with price p.

Price Dispersion and Resource Cost To complete the exposition of the pricing block of the model, it remains to define the total demand for the intermediate input, which is given by:

$$x_t^d \equiv \int x_{jt}\, dj = \int p_{jt}^{-\varepsilon_t}\, e^{-a_{jt}}\, Y_t\, dj = Y_t\, \Delta_t, \tag{20}$$

where $\Delta_t \equiv \int p_{jt}^{-\varepsilon_t} e^{-a_{jt}}\, dj$ is the equilibrium level of price dispersion in the economy and is pinned down by the distribution of prices:

$$\Delta_t = \int_{(p,a)} p^{-\varepsilon_t} e^{-a} d\Omega_t(p,a). \tag{21}$$

Given aggregate and idiosyncratic conditions, retailers make their pricing choices, generating a level of aggregate price dispersion Δ_t. In an economy with no pricing frictions, price dispersion would arise only due to differences in idiosyncratic productivity. With frictions, there is additional, inefficient price dispersion that causes misallocation across firms.

The resource cost of the pricing frictions is then given by

$$F_t = \int \theta^a \mathcal{D}\left(\Lambda_t(p,a) \parallel \bar{\Lambda}\right) d\tilde{\Omega}_t(p,a) + \int \Lambda_t(p,a) \Big[\kappa + \theta^p \mathcal{D}\left(f_t(\cdot|a) \parallel \bar{f}\right)\Big] d\tilde{\Omega}_t(p,a), \tag{22}$$

where the first term integrates over all review costs and the second term adds up the repricing cost of all firms that adjust in the period.

The measure of price dispersion and the resource cost associated with the price setting frictions are the only inputs into the non-pricing block of the model, allowing us to separate the information problem from the rest of the economy.

3.2 Closing the Model

Intermediate Goods Producers The supply of intermediate inputs is determined by a continuum of competitive producers who choose labor to maximize static real profits

$$\pi_t^x = p_t^x x_t - w_t L_t \tag{23}$$

subject to the production function

$$x_t = e^{a_t + z_t} L_t, \tag{24}$$

where L_t is labor input, w_t is the real wage, a_t is an AR(1) process for log aggregate productivity, and z_t is a random walk process that grows at the rate γ_{zt}, which is itself an AR(1) process.

Optimization by the intermediate goods producer yields

$$p_t^x = e^{-(a_t + z_t)}\, w_t \tag{25}$$

and market clearing for the intermediate input yields total labor demand

$$L_t = Y_t\, \Delta_t\, e^{-(a_t + z_t)}. \tag{26}$$

3.3 Households and Fiscal Authority

The representative household chooses streams of consumption C_t, labor supply L_t, and the real value of the risk-free bonds purchased in period t, B_t, to maximize lifetime utility,

$$\mathcal{U} = E_0 \sum_{t=0}^{\infty} \beta^t \zeta_t \left[\ln\left(C_t - hC_{t-1}\right) - \xi_t \cdot \left(\frac{L_t^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}} \right) + \chi_t \cdot B_t \right], \tag{27}$$

subject to the sequence of flow budget constraints

$$C_t + B_t \;=\; w_t L_t + D_t - T_t + B_{t-1} \frac{i_{t-1}}{\pi_t}, \tag{28}$$

and a no-Ponzi condition, where $\beta \in (0,1)$ is the household's discount factor, ζ_t

is a discount factor shock, $\nu \geq 0$ is the Frisch elasticity of labor supply, ξ_t is a shock to the relative disutility of working, and $h \in [0, 1)$ is the degree of habit in consumption, D_t are firm dividends, T_t are lump-sum taxes net of transfers, and $\pi_t \equiv \frac{P_t}{P_{t-1}}$ denotes the gross inflation rate between period $t-1$ and period t. The household supplies labor to the intermediate goods producers, earns a real wage w_t, and can invest in one-period bonds that earn a gross nominal rate i_t between period t and period $t+1$.

The exogenous discount factor shock affects the intertemporal Euler equation and has been shown by Justiniano, Primiceri & Tambalotti (2010) and others to be an important (reduced-form) driver of consumption fluctuations. It is also often used to drive the economy to the effective lower bound on the monetary authority's nominal policy rate (Eggertsson & Woodford, 2003). Shocks to the disutility of labor are introduced to affect the firm's marginal cost function. Lastly, including the real value of bond holdings in the household's utility function allows for "flight to quality" shocks χ_t. This is particularly relevant in an extension of the model that features investment in capital, since these shocks generate variation in the demand for risk-free bonds that allows the interest rate controlled by the monetary authority to deviate from the return on other assets. Originally, Krishnamurthy & Vissing-Jorgensen (2012) proposed introducing "convenience" assets, namely highly liquid, very safe assets, such as U.S. Treasuries, in the utility function, analogously to a money-in-the-utility specification.[5] Fisher (2015) shows how this specification endogenizes the risk preference shocks that Smets & Wouters (2007) introduce to the consumption Euler equation to help generate comovement between investment and consumption, and Campbell, Fisher, Justiniano & Melosi (2017) use it in a DSGE model that quantifies the effects of forward guidance on the U.S. economy since the Great Recession.

Let λ_t denote the Lagrange multiplier on the flow budget constraint. Household

[5] Krishnamurthy & Vissing-Jorgensen (2012) point to theoretical results demonstrating how assets with superior liquidity and perceived safety can command a premium. They also demonstrate the empirical relevance of this premium. Using U.S. data on corporate bond-Treasury spreads, Krishnamurthy & Vissing-Jorgensen (2012) estimate an average convenience yield of 73 basis point s from 1926 to 2008, nearly two thirds of which represents liquidity convenience, with the remainder representing safety convenience. This spread implies that the Treasury yield is lower than the actual risk-free rate and suggests care when parameterizing the risk-free rate in models that do not explicitly model the preference for convenience assets.

optimization yields

$$\left(\frac{1}{C_t - hC_{t-1}}\right) - E_t\left[\left(\frac{\beta\zeta_{t+1}}{\zeta_t}\right)\left(\frac{h}{C_{t+1} - hC_t}\right)\right] = \lambda_t, \tag{29}$$

$$\xi_t L_t^{\frac{1}{\nu}} = \lambda_t w_t, \tag{30}$$

$$\zeta_t\,\chi_t + E_t\left[\beta\,\zeta_{t+1}\,\lambda_{t+1}\left(\frac{i_t}{\pi_{t+1}}\right)\right] = \zeta_t\,\lambda_t, \tag{31}$$

along with the real discount factor between t and $t+1$:

$$M_{t,t+1} \equiv \frac{\beta\zeta_{t+1}\lambda_{t+1}}{\zeta_t\lambda_t}. \tag{32}$$

The final good is used for consumption C_t, government spending G_t, and to pay for the costs associated with pricing frictions F_t,

$$Y_t = C_t + G_t + F_t \tag{33}$$

with

$$Y_t = \left[\int y_{jt}^{\frac{\varepsilon_t - 1}{\varepsilon_t}}\,dj\right]^{\frac{\varepsilon_t}{\varepsilon_t - 1}}. \tag{34}$$

3.4 Wage Rigidity

For simplicity, we include reduced-form real wage rigidity given by

$$w_t = \delta^w \bar{w}^* + (1 - \delta^w) w_t^*, \tag{35}$$

where $\bar{w}^*$ is the steady state real wage and w_t^* is the competitive real wage.

3.5 Monetary and Fiscal Authorities

The monetary authority follows a Taylor rule that responds to deviations of inflation from its target and to deviations of output growth from long run growth, as in Sims (2013). The rule also features interest-rate smoothing and is subject to a zero lower bound. When not constrained by the lower bound, monetary policy implements

$$i_t^a = \rho_i \, i_{t-1}^a + (1-\rho_i) \left[i_{ss}^a + \phi_\pi \left(\pi_t^a - \pi_{ss}^a \right) + \phi_y \left(dy_t^a - \gamma_{ss}^a \right) \right] + e_t^\pi + e_t^i \,, \tag{36}$$

where i_t^a is the annual nominal rate in month t, π_t^a is the inflation rate over the most recent 12 months, dy_t^a is the real output growth over the most recent 12 months, γ_{ss}^a is steady-state annual output growth, and i_{ss}^a is the steady state nominal rate associated with steady state inflation π_{ss}^a. All rates are annualized rates. Policy is subject to a Gaussian AR(1) shock e_t^π with persistence ρ_π and innovations with standard deviation σ_π, and a Gaussian i.i.d. shock e_t^i with standard deviation σ_i.

We assume that government spending is funded by lump-sum consumer taxes and is a fixed share g of steady state output net of pricing frictions,

$$G_t = g(Y_{ss} - F_{ss}). \tag{37}$$

3.6 Shocks

We include a range of fundamental shocks, to avoid overstating the role the pricing frictions play in generating aggregate volatility. The aggregate exogenous shocks are to aggregate TFP (a_t), impatience (ζ_t), disutility of labor supply (ξ_t), bond demand (χ_t), markups (ε_t), the Taylor rule (e_t^π, e_t^i), and permanent productivity growth (γ_{zt}). In addition, we allow for unanticipated shocks to the menu cost, the marginal costs of acquiring information, and to the standard deviation of idiosyncratic shocks (σ_{ajt}). Variation in the pricing costs over time is interpreted as variation in the efficiency with which firms can process information and implement price changes. Allowing for this variation enables us to uncover to what extent nominal frictions move with the economy's cycle, and if there have been trends in pricing costs over time.

3.7 Balanced Growth

The source of long-run growth in the model is labor-augmenting technological progress, z_t, which grows at rate γ_{zt}. We detrend aggregate variables by z_t and we also require that the variables $b_t, \kappa_t, \theta_t^r, \theta_t^p$ grow at the same rate as z_t. Hence, we define $x_t \equiv x \cdot z_t$ for each element in this set.

4 Steady State Frictions

In this section we report what the pricing statistics over recent decades suggest about the nature and severity of pricing frictions in the United States, on average.

4.1 Steady State Data

We use statistics on price-setting patterns from the U.S. Consumer Price Index (CPI) to estimate the steady-state parameters. These statistics are based on the individual price quotes underlying the CPI, as constructed by Nakamura et al. (2018) for the sample starting in January 1978 and ending in December 2014, and extended by Montag & Villar (2023) to March 2023. We thank Daniel Villar for sharing the time series of these pricing moments with us.

The microdata underlying these statistics consist of approximately 80,000 monthly price quotes for products grouped into roughly 305 categories, or "entry-level items" (ELIs), which are then further aggregated into 13 major groups. Authors with access to the microdata can use the individual price quotes to construct empirical distributions of log price changes for each month, from which various pricing statistics are then calculated. For example, to calculate the frequency of price changes in each period, one computes the fraction of nonzero price changes across products within each ELI, and then the expenditure-weighted median across ELIs. Similarly, conditional on a price change, the absolute size of price changes is computed by taking the average log price change across products within each ELI and aggregating it to the expenditure-weighted median across ELIs. Higher moments are computed in a similar way, but by pooling data within each *major group* rather than within each ELI, and by taking an expenditure-weighted average across the 13 major groups.[6]

[6] Luo & Villar (2021) discuss the need to compute higher moments at the group level, due to sample size constraints at the ELI level: Since ELIs are narrowly defined categories, they sometimes have only a small number of observations. Higher moments are particularly sensitive to outliers, which is why a small number of observations is insufficient to compute them reliably within each ELI.

TABLE I: Data Moments

	Full Sample 1978-2023.Q1	Post-1984 1984-2023.Q1	Great Moderation 1984-2007.Q2	Baseline Model
Frequency	0.1131	0.1092	0.1005	0.0972
Size (abs. value)	0.0735	0.0744	0.0740	0.0750
Std. deviation	0.129	0.133	0.129	0.122
Skew	-0.131	-0.166	-0.142	-0.141
Kurtosis	11	10	11	11
Frequency up	0.076	0.071	0.069	0.067
Size up	0.069	0.069	0.070	0.073
Federal funds rate	0.0462	0.0354	0.0532	0.0532
Inflation rate	1.0354	1.0281	1.0308	1.0308
GDP growth rate	1.0151	1.0158	1.0204	1.0204

Notes: The shaded column shows the moments targeted in the steady state estimation. The pricing statistics report average values for moments constructed for the monthly distributions of log-price changes. *Frequency up* and *Size up* report the frequency and size of price increases. The average effective federal funds rate, CPI inflation, and per-capita real GDP growth for each sub-sample are annual rates. Data sources: Daniel Villar and FRED.

Table I reports average pricing moments for the full sample period and for two sub-samples: the post-1984 period, which may be of independent interest since it represents a period of established modern monetary policy, and the Great Moderation, which the Federal Reserve dates between January 1984 and June 2007, and which we target in our steady state calibration. The table also reports the average values for inflation, GDP growth, and federal funds rate, which we also target in our steady state estimation, and which we take from FRED.

4.2 Steady State Parameters

We parameterize the model's stochastic steady state by targeting averages over the Great Moderation period since that was a period of relative macroeconomic stability. Nevertheless, prices displayed considerable volatility, as has been discussed

TABLE II: Steady State Parameters

Parameter		Baseline	$\kappa \doteq 0$	$\theta^a \doteq 0$	$\theta^p \doteq 0$	$K \doteq 5$
Mg. disutility of labor	ξ	1.7	1.8	2.1	1.1	1.3
Mg. utility of bonds	χ	0.058	0.059	0.059	0.059	0.057
Frisch elasticity	ν	2				
Discount factor	β	$0.6^{1/12}$				
Inflation (annual)	π^{ss}	1.0308				
Real growth (annual)	γ^{ss}	1.0204				
Gov't share G/C	g_c	0.25				
Fixed cost	κ	0.026	**0**	0.038	0.18	0.055
Timing accuracy cost	θ^a	0.097	0.29	**0**	2.58	0.0008
Repricing accuracy cost	θ^p	1.07	0.57	1.09	**0**	1.59
EOS among varieties	ε	11	11	11	5.8	6
SD(idios. shocks)	σ	0.077	0.05	0.09	0.19	0.067
Persistence(idios. shocks)	ρ	0.94	0.98	0.95	0.60	0.95

Notes: The model is estimated at the monthly frequency and parameters are set either at conventional values or to match averages for the Great Moderation period. The bottom six parameters are estimated jointly, targeting seven steady state pricing moments. The 'Baseline' column shows the best fitting parameter combination. The $K \doteq 5$ column shows values for the model that targets a steady state kurtosis of price changes of 5. The rest of the columns show parameter values when we re-estimate the model, shutting down one pricing friction at a time but continuing to target the same pricing moments.

in prior work (*e.g.*, Bils & Klenow, 2004; Golosov & Lucas Jr, 2007). In particular, the coexistence of large and frequent price cuts and price increases points to the importance of fairly large and frequent idiosyncratic shocks. Hence we will solve for a stochastic steady state with idiosyncratic shocks to firms' desired prices. The estimation of shocks away from the steady state will then make use of the entire data, including the volatile periods at the beginning and end of the sample.

Since the pricing moments aggregate monthly price changes, we set our model to a monthly frequency. In this way, we avoid having to map pricing moments based on monthly data into quarterly moments, which would require making assumptions about price rigidities in the transformation.

Table II presents the calibrated and estimated steady state parameter values. We set the steady state inflation rate to the realized average ($\pi^{ss} = 1.0307^{\frac{1}{12}}$), productivity growth to target the average real GDP growth per capita ($\gamma^{ss} = 1.0206^{\frac{1}{12}}$), and government spending to 25% of steady state consumption, per the realized sample average. The Frisch elasticity of labor supply is $\nu = 2$ and the relative disutility of working is $\xi = 1.68$, set to normalize employment in the steady state. Following Michaillat & Saez (2021), we parameterize wealth in the utility by setting the annual discount factor to 60% ($\beta = 0.6^{\frac{1}{12}}$) and then internally calibrating the parameter governing the marginal utility of bonds ($\chi = 0.0578$) such that the steady state nominal interest rate matches the average federal funds rate ($i^{ss} = 1.0532$). This specification helps the later estimation, when shocks push the economy to the effective lower bound on nominal interest rates.[7]

We estimate a vector of six parameters that jointly affect price-setting frictions, $\Theta = \{\varepsilon, \sigma, \rho, \kappa, \theta^a, \theta^p\}$, to target the averages of seven pricing moments: frequency, size, standard deviation, skew, and kurtosis of price changes, and frequency and size of price increases. We estimate Θ using

$$\Theta = \arg\min_{\mathcal{X}} \left(\mu(\mathcal{X}) - \mu^{data}\right) W \left(\mu(\mathcal{X}) - \mu^{data}\right)', \tag{38}$$

where $\mu(\mathcal{X})$ is the vector of model moments for a given pricing parameter $\mathcal{X}$, μ^{data} is the vector data moments, and the weighting matrix (W) is a diagonal matrix with the inverse of the data moments. We impose bounds on the parameters to avoid unreasonable values.[8]

We find a modest menu cost ($\kappa = 0.026$), a larger, though still relatively modest variable cost of information to determine if a price change is warranted ($\theta^a = 0.097$), and a large cost of accuracy in repricing ($\theta^p = 1.07$). Together, these costs imply a steady state level of expenditure on repricing that is 2.6% of steady state revenues.

Table III reports the breakdown of price-setting costs: Firms spend almost 0.3% of steady state sales on the fixed cost of price reviews. They spend approximately

[7]See also Cuba-Borda & Singh (2019).

[8]The bounds are $\theta^p, \theta^a \in [0, 10]$, $\kappa \in [0, 3]$, $\varepsilon \in [3, 11]$, $\sigma \in [0.03, 0.4]$, and $\rho \in [0, 0.996]$.

TABLE III: Steady State Outcomes

	Baseline	$\kappa \doteq 0$	$\theta^a \doteq 0$	$\theta^p \doteq 0$	$K \doteq 5$
Spending on repricing (share of revenues)					
Fixed cost ($\kappa \bar{\Lambda}$)	0.0028	0	0.0038	0.0140	0.0056
Review cost ($\theta^a I^a_{ss}$)	0.0076	0.0088	0	0.0021	0.0003
Repricing cost ($\theta^p I^p_{ss}$)	0.0158	0.0140	0.0199	0	0.0045
Total spending (F_t)	0.0262	0.0228	0.0237	0.0162	0.0104
Outcomes (relative to flex-price outcomes)					
Consumption	0.9258	0.9574	0.9352	0.8527	0.9656
Employment	1.0514	1.0359	1.0432	1.0211	1.0670
Wages	0.9493	0.9745	0.9552	0.8616	0.9975
Output	0.9507	0.9798	0.9579	0.8667	0.9758
Price Dispersion	1.1058	1.0573	1.0891	1.1781	1.0936

Note: Baseline steady state estimation and alternatives. Spending on pricing decisions is reported as a share of steady state revenues. Aggregate outcomes are reported as a share of the aggregate outcomes in an economy without pricing costs, but otherwise identically parameterized. The last column shows values for a model that targets a steady state kurtosis of price changes of 5 instead of the sample average of 11.

0.76% of revenues on acquiring information to decide whether or not to change their price, and roughly twice that amount (1.6% of revenues) to determine what price to charge, conditional on adjustment.

Compared with the flexible-price steady state with the same distribution of idiosyncratic shocks and the same elasticity of substitution among varieties, the economy with pricing frictions delivers significantly lower welfare. Steady state consumption is 7.4% lower, employment is 5.1% higher, and wages are 5.1% lower. Consumers work harder for less. Prices deviate from flexible price levels, and as a result, equilibrium price dispersion is 10.6% higher than the price dispersion that would be warranted given heterogeneity in productivities alone. Hence, although relatively modest at the level of each individual firm, these frictions aggregate to considerable losses for consumers.

Table IV presents the match between model and data for the pricing moments.

TABLE IV: Pricing Moments

	Baseline	$\kappa \doteq 0$	$\theta^a \doteq 0$	$\theta^p \doteq 0$	$K \doteq 5$
Freq. of price changes	0.0976	0.0880	0.0921	0.0751	0.0952
Absolute size	0.0751	0.0853	0.0794	0.0789	0.0792
Standard deviation	0.1223	0.1165	0.1447	0.0981	0.1067
Skew	-0.1419	-0.0983	-0.1271	-0.1394	-0.1406
Kurtosis	11	7	12.7	9	5
Freq. of price increases	0.0676	0.0577	0.0676	0.0483	0.0647
Size of price increases	0.0728	0.0870	0.0728	0.0874	0.0778

Pricing statistics for the baseline steady state estimation and alternatives. The last column shows values for a model that targets a steady state kurtosis of price changes of 5 instead of the sample average of 11.

Since we target more moments than we have parameters, the fit is not perfect, but it is quite close. In addition to the pricing costs, matching these moments also requires a high elasticity of substitution ($\varepsilon = 11$, which is higher than the value usually estimated in menu cost models, but closer to the values used in the trade literature), and highly volatile and persistent idiosyncratic shocks (which is typical in this literature, given the large volatility of price changes). The match is particularly notable since we use normally distributed shocks (rather than the leptokurtic shocks that are usually used in order to match the standard deviation and kurtosis of price changes).

Among all the targeted moments, which are quite standard, kurtosis deserves additional discussion, since the value for our data is much higher than the values typically reported in the literature, which tend to range between 3.5 and 5.5. This difference may be due to the broader range of goods in the CPI than in other data sets, and may reflect a combination of cross-sectional heterogeneity (for instance, differences across ELIs in repricing costs or in technologies) and measurement error (since higher order moments are much more difficult to estimate accurately). How much are results driven by this higher than usual level of kurtosis in the distribution of price changes? The last columns of Tables II, III, and IV report the estimation results of targeting a kurtosis of 5. Compared to our baseline estimation, estimating the model parameters targeting a kurtosis of 5 would yield a significantly higher value for the cost of repricing accuracy ($\theta^p = 1.6$ instead of 1.1), a lower elasticity of substitution ($\varepsilon = 6$ instead of 11), higher menu cost ($\kappa = 0.055$ vs. 0.026), lower values for θ^a, and smaller shocks. Since we unilaterally change the kurtosis target without changing any of the other moments, the fit with the data is weaker. Importantly, although it yields a different mix of fixed cost and variable reviewing costs, the main result that errors in pricing are the primary driver of pricing frictions is just as strong if not stronger.

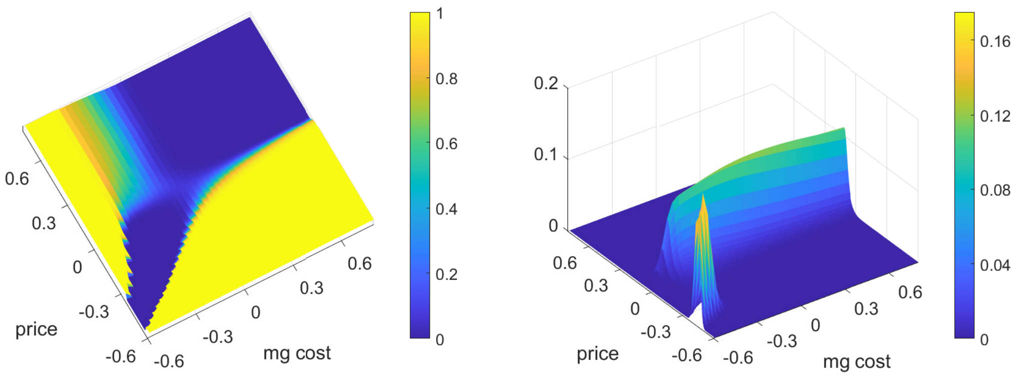
4.3 Restricting the Pricing Frictions

What does the model fit look like in the steady state if we restrict the type of pricing frictions? Columns in Tables II, III, and IV report the estimation results of targeting the same moments, but shutting down one pricing friction at a time. As expected, since we have fewer parameters, the match with the pricing moments is not as good. In fact, the version of the model that imposes no mistakes in pricing ($\theta^p \doteq 0$) yields the worst fit. But that is less important than how the parameters change as we vary the nature of the frictions.

First, note that when we impose $\kappa \doteq 0$, the model continues to deliver infrequent price adjustment, arising solely from the information friction: firms' uncertainty about the right price to charge and about when to change prices leads to a substantial amount of inaction (the frequency of price changes is 8.8% in the model with no menu costs, versus 9.8% in the baseline model with all three frictions.) Matching the pricing moments in this case requires a much larger cost of accurately timing price changes ($\theta^a = 0.29$ vs. 0.1 in the baseline model), smaller repricing errors, and smaller idiosyncratic shocks.

Second, imposing $\theta^a \doteq 0$ does not reduce the fit of the model too much, since that parameter was small to begin with, implying an accurate timing of price adjustment. Moreover, across all models with repricing errors, total spending on repricing is relatively stable, even though its composition changes.

Finally, to match the pricing moments when imposing no repricing errors ($\theta^p \doteq 0$) requires Calvo-like price adjustment (a very high cost of accurately timing price changes, $\theta^a = 2.6$) and also a much higher menu cost ($\kappa = 0.18$), and still the



(a) Hazard function

(b) Pricing policy

Figure 1: Pricing Policy

Note: This figure plots the optimal adjustment probability Λ and the optimal pricing policy f, in steady state, as functions of log price and log marginal cost.

fit is substantially weakened. Moreover, due to weak state-dependence in the timing of price changes, the model also generates larger relative price dispersion and hence significant consumption loses compared with the models with repricing errors (Table III).

4.4 Discussion of Steady-State Policies

What do the estimated parameters imply for the accuracy of pricing decisions? Inaccuracy in pricing arises along two dimensions: the decision of whether or not to review the existing price, as well as the decision of what price to set when deciding to change the existing price. The accuracy with which firms make these two decisions is captured in their adjustment probability and in their pricing policy, both of which are plotted in Figure 1.

We estimate fairly high accuracy in the timing of price adjustments, as indicated by the steep increase in the probability of adjustment away from the optimum, shown in the left panel of the figure. When a firm's existing price is close to the price it would expect to set upon repricing, the probability of adjustment is very close to zero. But for prices that are farther away, the probability of adjustment rises rapidly, and as a result, prices that are far from the optimum do not survive long. The adjustment occurs especially quickly when prices are relatively low, resulting in an asymmetric hazard. This asymmetry has been documented as a feature of rationally inattentive price adjustment (Woodford, 2009), and we find it is also a strong feature in historical U.S. data.

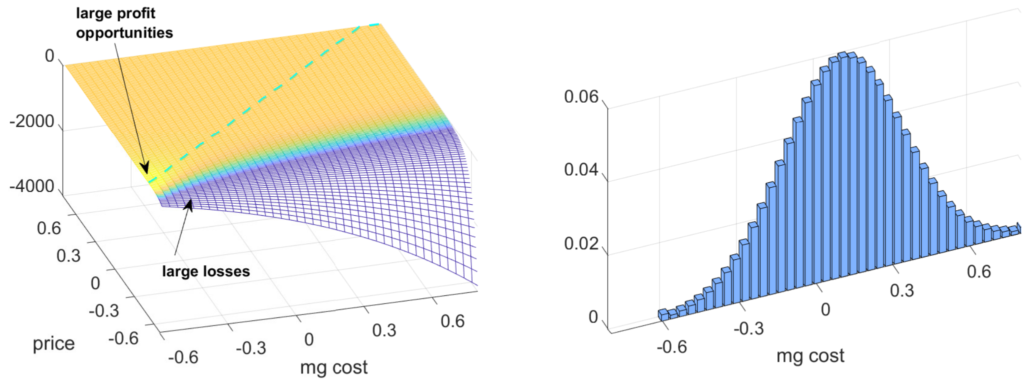
Conditional on adjustment, we estimate substantial errors in pricing. As shown in the right panel of Figure 1, prices are drawn from an imprecise, weakly state-dependent distribution. Two features of this policy stand out: First, there is significant price dispersion. For a given marginal cost, there is a wide range of prices that the firm could charge with a significant probability, reflecting a high degree of uncertainty about the right price to charge. Models that assume reset prices based on perfect information would yield a degenerate distribution, thus potentially significantly overstating the responsiveness of prices to shocks, conditional on adjustment.

Second, a pattern of occasional sales arises endogenously in the optimal pricing policy, purely as a way to capture highly profitable demand opportunities. The endogenous adjustment probability ensures that if the firm sets prices that are too low, it will change these prices quickly. Hence, the review policy effectively insures against pricing too low for too long. Given this insurance, firms choose to occasionally set low prices, to capture significant demand from competitors (the elasticity of substitution we estimate is high), knowing that if they are wrong in their assessment of demand or marginal cost, they will promptly get an accurate signal to revise prices, such that any potential profit loss will be short-lived.

Overall, we estimate a strongly state-dependent price adjustment probability but inaccurate prices. It seems firms can determine quite accurately when their prices have become obsolete, but they have more difficulty determining the right price level to set. This finding deviates from the conventional models of price rigidity, which assume perfect but infrequent repricing.

4.5 Incentives for Information Acquisition

How do incentives for information acquisition shape firms' pricing policies? Figure 2 plots firms' operating profit function and the distribution of marginal costs for the baseline parameterization. These two objects shape firms' incentives to acquire information and set prices accurately. The former governs the losses from



(a) Profit function

(b) Marginal cost distribution

Figure 2: Firm Incentives

Note: This figure plots the profit function in the left panel and the distribution of marginal costs in the right panel. The near-diagonal dashed line marks the full information flexible price optimum (constant markup over each marginal cost value.

mispricing, while the latter captures the incidence of these losses across states. In particular, the firm's operating profit per period – the flow profit before spending on information and adjustment costs – is asymmetric. The asymmetry shapes the attention of a firm trying to decide how to learn most efficiently about what price to set. Different considerations are at play, depending on where a firm finds itself in the distribution of marginal costs. Low current marginal costs provide an opportunity to generate significantly higher profits than average, especially given a high elasticity of substitution. Hence, firms have strong profit incentives to capitalize on these low-cost states. On the other hand, underpricing (relative to marginal cost) can generate large profit losses. As a result, firms have incentives to err on the side of over-pricing unless they are quite certain their costs are low. Hence, just based on the shape of the firm's profit function, we would expect firms to have a tendency to (i) price high when they do not have much information, and (ii) pay more attention when profit opportunities are high.

The incentive to set high prices is dampened by the frequency with which the firm expects to find itself in different states. With normally distributed shocks to marginal cost, the firm expects to spend more time closer to the center of the distribution, as shown in the second panel of Figure (2), which plots the distribution of marginal costs. This force induces firms to charge moderate prices more frequently, as these are likely to be close to optimal more often, thereby helping firms economize on information and adjustment costs.

4.6 Menu Cost Model Alternatives

To illustrate the incentives to price accurately in different states, we now compare our model to three menu cost alternatives: with perfect repricing, with repricing errors, and with timing errors. To show the differences across models more sharply, we begin with a higher menu cost than estimated in our baseline model, setting $\kappa = 0.07$.

Consider first the choices of firms that have perfect information for free, and are subject only to the menu cost when changing prices. Due to the fixed cost, there is a range of inaction in which firms keep their prices unchanged. Firms whose prices fall outside this range adjust to a dynamically optimal reset price. This price tracks the full information flex-price optimum very closely (albeit from above) for a wide range of marginal costs. The price only becomes less cost sensitive when marginal costs get unusually high. The firm understands that the very high cost will be fairly short lived and moreover, it is also one with relatively low demand due to the existance of a substantial mass of lower-price firms. As a result, the firm would rather set a lower price today than face the high probability of having to pay the menu cost to adjust prices down in the near future.

As has been extensively discussed in prior work, the resulting adjustment policy features asymmetric Ss bands of adjustment. As shown in Figure 3, for low marginal costs, the Ss bands are very narrow and the optimal reset price, which maximizes the firm's continuation value, closely tracks the flexible price (though it is slightly above it): Despite the low incidence of the low cost states, it is nonetheless worth it to the firm to price very accurately in these states, because the profit gains are so large. On the other hand, for higher marginal costs, the firm chooses instead a moderate price that applies to a wider range of states and enables it to save on the menu cost: Profit losses are not too large or sensitive to mispricing, so the Ss bands widen (especially for prices above the optimal reset price), and the reset price becomes less and less sensitive to the marginal cost.

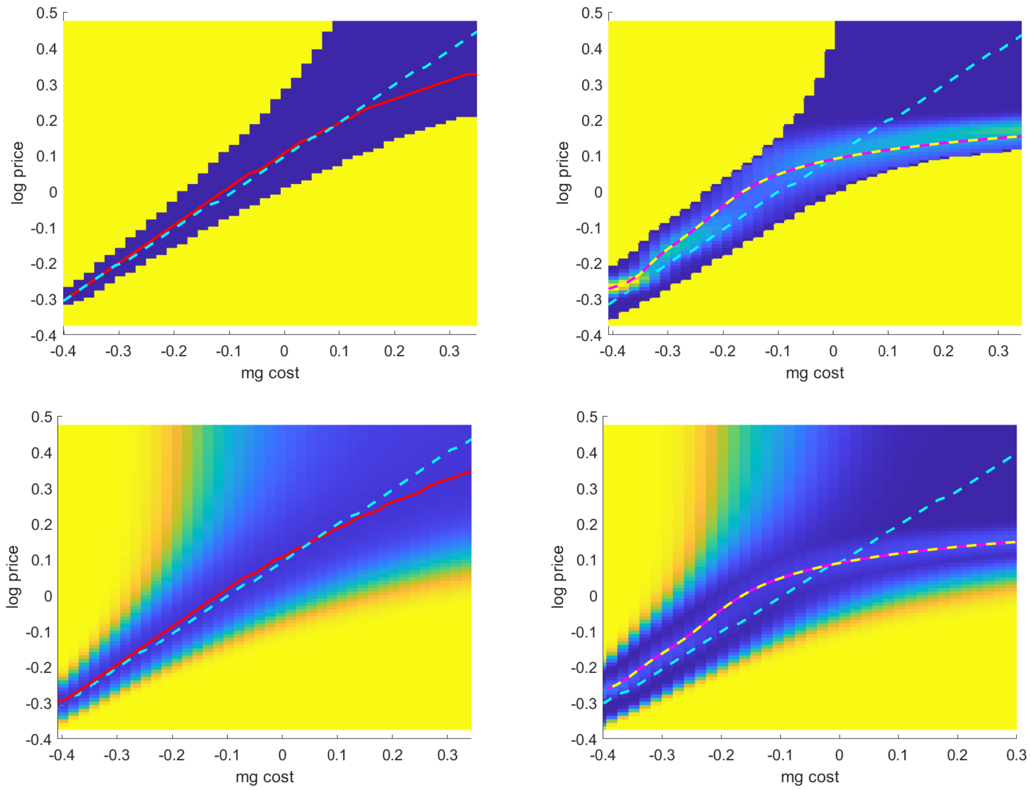


Figure 3: Policies Across Alternative Models

Note: This figure plots the adjustment policy and repricing policy for four models: (a) the menu cost model with perfect repricing ($\kappa = 0.07$), (b) the menu cost model with errors in pricing ($\kappa = 0.07$, $\theta^p = 0.1$), in the top right panel, (c) the menu cost model with errors in timing ($\kappa = 0.07$, $\theta^a = 0.1$), in the bottom left panel, and (d) the model with all three costs. In all panels, the cyan dashed line marks the full information flexible price optimum for each marginal cost value, and the red solid line marks the optimal reset price. Panels (b) and (d) (for which $\theta^p > 0$) plot the weighted average of the conditional reset price distribution, f, and also super-impose this reset distribution. mg cost stands for log-deviations of the marginal cost with respect to the average marginal cost in the economy.

The top right panel of Figure 3 shows the optimal policy for a new type of menu cost model that includes errors in pricing ($\theta^p > 0$). Instead of setting a deterministic reset price that maximizes its continuation value, the firm now sets a price probabilistically, drawn from an optimal distribution that maximizes its expected continuation value. The optimal distribution is given by an expression of the same form as in equation (18), with the appropriately adjusted value function.

Compared to the pure menu cost model, the errors-in-pricing model features additional endogenous price rigidity: The Ss bands widen and the frequency of price changes drops significantly. The possibility of mistakes in repricing arises as an important source of additional price rigidity, beyond the menu cost itself. A higher θ^p pushes toward a lower frequency of adjustment, biasing the firm toward inaction. This is a rational response, given the possibility of mistakes in pricing. As a result, models may over-estimate the size of adjustment costs needed to match the frequency of price changes, if they abstract from the possibility of mistakes in repricing. Furthermore, compared to the optimal reset price in the pure menu cost model, the weighted average price charged in different states now becomes even less sensitive to marginal costs. The repricing distribution features strong dampening of price responsiveness to marginal costs that are above average, while maintaining the downward flexibility of prices for low marginal costs. The pricing policy also features stronger over-pricing at low marginal cost values, as a way for the firm to protect itself against underestimating its marginal cost.

The errors-in-pricing model features a distribution of reset prices that is most dispersed in the middle, reflecting the fact that the firm fine-tunes its pricing accuracy depending on the state. The firm's optimal conditional distribution of reset prices is tight at low marginal costs because low costs are highly profitable opportunities that are worth capturing accurately. On the other hand, it is also tight at high marginal costs, but for an entirely different reason: high marginal costs present opportunities to save on information and adjustment costs by not differentiating prices across states too much. Most of the time, however, firms are in the middle range with the widest price dispersion.

The bottom left panel of Figure 3 shows the optimal policy for a menu cost model with errors in timing, similar to the model of Woodford (2009). In this case, $\theta^p = 0$, and the firm sets the optimal reset price deterministically, but $\theta^a > 0$, such that the firm adjusts its price *probabilistically*, with a continuous adjustment probability $\Lambda \in (0, 1)$, that takes a form similar to that in equation (17).

Compared to the pure menu cost model, the probability of adjustment is significantly above zero everywhere, even at the optimal reset price, since the firm is never certain of the state, and hence sometimes adjusts even when it should not. Away from the optimal reset price, the probability of adjustment increases gradually and asymmetrically: low costs are worth identifying accurately since, as before, they offer highly profitable opportunities, while the rest of the time, adjustment is Calvo-like even for modest information frictions. As θ^a increases, the probability of adjustment Λ becomes more and more Calvo-like, approaching a constant probability of adjustment $\bar{\Lambda}$.

Finally, the bottom right panel of Figure 3 reproduces the optimal policy for the menu cost models with errors in both timing and pricing. Timing and repricing accuracy are now chosen to be jointly optimal, and hence they interact to optimize the use of information. Mistakes in price-setting make firms more careful when changing prices so that the timing of adjustments becomes more state-dependent. Relative to the model of Woodford (2009), the introduction of mistakes in pricing makes the probability of adjustment Λ more well-shaped and with a wider region of near-inaction. On the other hand, mistakes in timing make the firm pay more attention to the price it sets, reducing over-pricing in low-cost states, slightly increasing sensitivity to marginal cost in high-cost states, and overall reducing the dispersion in the distribution of reset prices.

These interactions provide a new way to rationalize Calvo-like behavior. Firms that can learn what the right price is very easily do not need to worry about timing their price changes. They can change prices with some constant probability, as is assumed in the Calvo model. But if figuring out the right price is very difficult, firms should pay close attention to when their prices are wrong, thus making their timing decision more state-dependent. In practice, our estimates suggest that firms may find it easier to learn they are setting the wrong price than to know how to fix it, as indicated by the relative sizes of the two information costs.

4.7 Implied Non-neutrality

What do these pricing frictions imply for the degree of monetary non-neutrality? Table V reports the cumulative impulse response (CIR) of consumption to a 25 bp shock to the federal funds rate for our model and for alternatives that turn off different pricing frictions.

The model deviates from the conventional wisdom, which posits that in order to get meaningful non-neutrality, one needs weak state dependence in the timing

TABLE V: Non-Neutrality

Model	CIR
Baseline model	0.112
Model with kurtosis $\doteq 5$	0.180
Model with $\kappa \doteq 0$	0.097
Model with $\theta^a \doteq 0$	0.117
Woodford (2009) model ($\theta^p \doteq 0$)	0.145

Notes: The table reports the cumulative response of consumption, as a percent of annual steady state consumption, in response to a 25-bp impulse to the federal funds rate.

of price changes. For example, with perfect repricing, menu cost models need some auxiliary features to weaken the selection in terms of who adjusts in response to a shock. But with mistakes in repricing, firms can end up with a suboptimal price even if they correctly decide when to adjust. Hence mistakes in the timing of price changes are no longer strictly necessary for non-neutrality.

Mistakes in repricing mean that in the aggregate there is a lower effective frequency of adjustment. We can connect this discussion to the discussions of selection in time-dependent and state-dependent models (Caballero & Engel, 2007; Golosov & Lucas Jr, 2007; Auclert, Rigato, Rognlie & Straub, 2022; Gagliardone et al., 2023). In effect, our model features negative selection on the *intensive* margin that either amplifies the lack of selection along the extensive margin in the Calvo model or works to offset the positive selection along the extensive margin in the menu cost model. As a result, the menu cost model may overstate flexibility both because of over-estimating selection in terms of who adjusts and because of not considering the trade-offs between paying attention to *both* the pricing and the timing decisions.

4.8 Calvo and the Bounds on Non-neutrality

Firms with an accurate adjustment policy will not devote as much effort to implementing a very precise pricing decision: they will tolerate mistakes in price levels because the adjustment probability is steep so that these mistakes won't survive long. On the other hand, firms that have a more imprecise adjustment policy that conditions on the state more weakly, will choose to pay more attention to the prices they charge, so that even if they mistakenly decide to change prices, there is a relatively high probability that they will choose a suitable price.

This interaction illustrates how both the Calvo model and the menu cost model can overstate price flexibility. Information frictions suggest that if the adjustment is Calvo-like, the firm will try to make the pricing more accurate, while if the adjustment is accurate like in the menu cost model, the pricing will be imprecise.

With perfect repricing, varying the severity of information frictions regarding the *timing* of price adjustment spans the degree of state dependence in price setting, with the menu cost model at one end (when the information friction approaches zero) to Calvo (1983) at the other end (when the information friction is strong enough that the firm acquires no information to decide when to adjust its prices), as shown by Woodford (2009). However allowing for errors in the repricing decision itself adds a new source of mispricing, and hence non-neutrality. As a result, Calvo need no longer be the upper bound on price rigidity: As the cost of setting accurate prices increases, mistakes in repricing increase, and as a result, non-neutrality also increases.

5 Statistics and Simulations

In this section we discuss the identification of our pricing costs and the mapping between these costs and the degree of non-neutrality. We are particularly interested in the following questions: First, what is it about the pricing data that points to the importance of errors in pricing rather than in the timing of price changes? Second, is there a sufficient statistic that captures the frictions in price setting in this information-constrained model of price setting, as is the case with a wide class of full-information menu cost models (Alvarez et al., 2016)?

Identification rarely requires an extensive discussion in typical models of nominal rigidities since it is often a straightforward exercise. In both Calvo and menu cost models, the frequency of price changes is used to pin down either the Calvo parameter or the size of the menu cost, and the properties of the profit function and the distribution of idiosyncratic shocks are used to match the size, standard deviation, and kurtosis of price changes (for example, large shocks Golosov & Lucas Jr (2007) and fat tails as in Midrigan (2011)).

As for models with information frictions, they either do not target pricing moments, focusing only on the aggregate response of output to a monetary policy shock, or they take the same approach as the menu cost literature, setting the information friction parameter jointly with the idiosyncratic shock process.[9]

Since our model nests different sources of pricing frictions, it becomes important to pinpoint which moments pins down which friction, especially since we are estimating steady state frictions that deviate from the conventional wisdom, which has focused on imperfect timing of price adjustment rather than imperfect repricing.

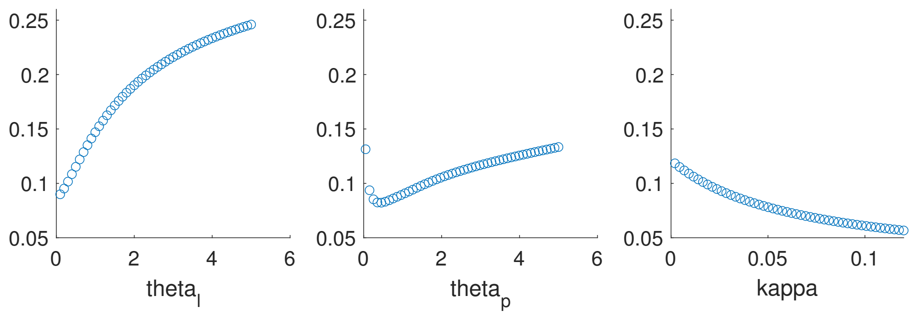
To address such questions we conduct simulations of the model around the estimated steady state, where in each simulation we vary one or more of the parameters of interest.

The panels in Figure 4 plot how the pricing moments vary with the three key pricing parameters (θ_p, θ_a, and κ), fixing the distribution of idiosyncratic shocks and the elasticity of substitution at their estimated steady state levels.

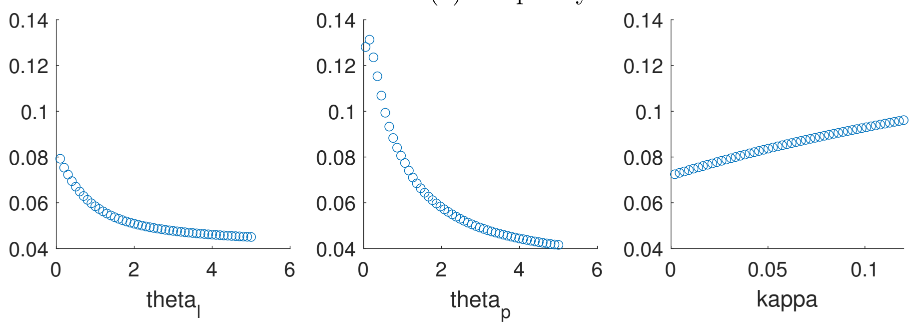
The simulations reveal some non-linearity in the effects that θ^p has on pricing moments. Most notably, increasing θ^p from 0 to 0.2 initially lowers the frequency of adjustment and raises the standard deviation of price changes. But once $\theta^p > 0.2$, a higher θ^p monotonically raises the frequency of adjustment and reduces the standard deviation of price changes.

Kurtosis also has a nonlinear relationship with θ^p: it increases as θ^p increases, but after a certain point, it starts decreasing as large deviations from the reference distributions are not profitable even for very productive firms. Even though kurtosis starts decreasing, the CIR keep increasing a prices become more anchored to the reference distribution, and hence more rigid.

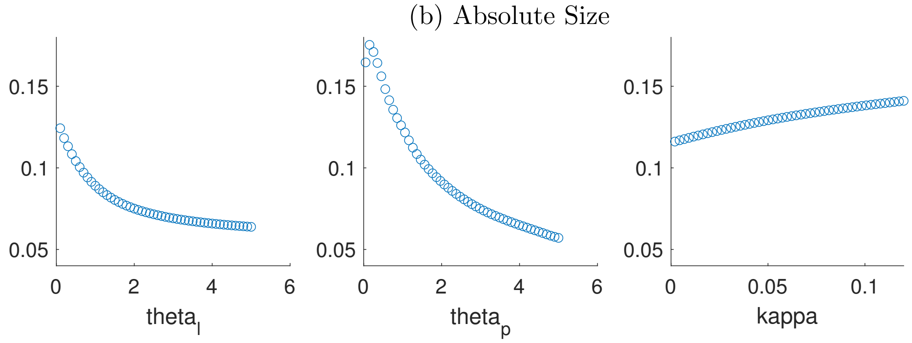
[9] An exception is the paper by Aruoba, Oue, Saffie & Willis (2024), who parameterize their menu cost model using not pricing moments, but rather moments from the cross-sectional distribution of demand and productivity estimated by Foster, Haltiwanger & Syverson (2008) on manufacturing data.



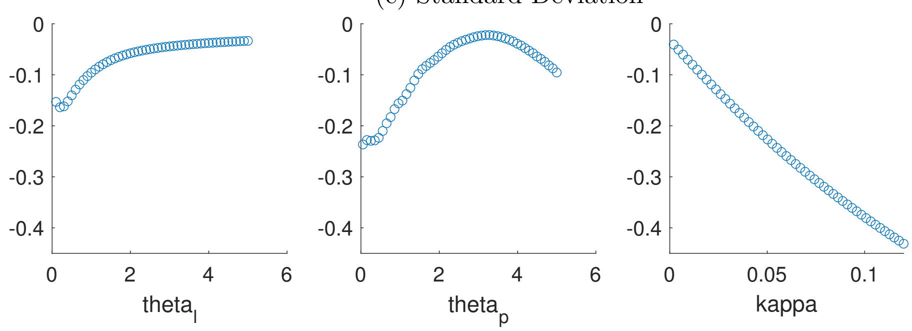
(a) Frequency



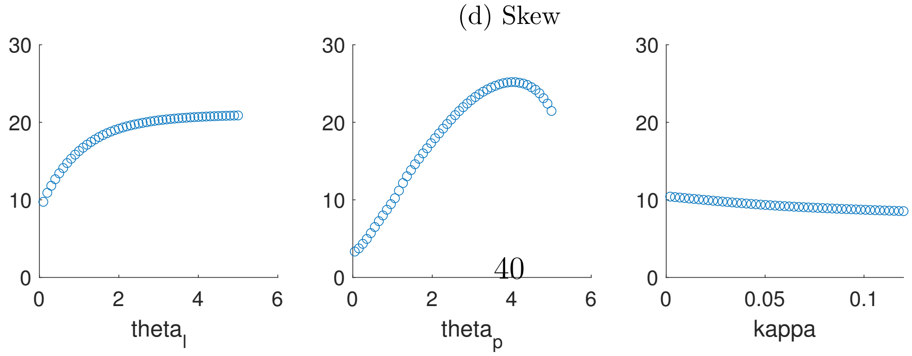
(b) Absolute Size



(c) Standard Deviation



(d) Skew



(e) Kurtosis

Figure 4: Identification of Pricing Costs in Simulations

6 Estimation

In this section, we report results from our Bayesian estimation of some of the model's parameters. To our knowledge, this is the first use of Bayesian techniques to estimate a model featuring rationally-inattentive and heterogeneous agents with an occasionally binding constraint on the interest rate. In this section, our aim is to exploit the time variation in the distribution of price changes to estimate variations in the pricing parameters and, as a consequence, price rigidity over time.

6.1 Estimation Approach

We estimate the model using Bayesian techniques. In the likelihood function evaluation, to gain both speed and accuracy, we compute the equilibrium dynamics using the Sequence-Space Jacobian (SSJ) method of Auclert et al. (2021). So far, the SSJ approach has been successfully used in a variety of models with heterogeneity across households; here, we show how it can also be effectively used in models with heterogeneous information.

We extend the SSJ method to handle occasionally binding constraints, since our sample period includes two episodes during which the lower bound was binding on the federal funds rate. The B describes how we adapt the methods proposed by Guerrieri & Iacoviello (2015) and Kulish et al. (2017) to handle the occasionally binding constraint when solving the model dynamics and evaluating the likelihood function using the SSJ method.

To reduce computational time, we only estimate the shock processes, we fix the habit formation and wage rigidity parameters based on the literature, and we run a separate Bayesian estimation for the parameters of the Taylor rule for the monetary authority. Even though we gain speed and accuracy with SSJ, the presence of the two ELB periods makes the likelihood evaluation time-consuming, in part, because we have a monthly model with a large number of observations (541 months).

The estimation includes fundamental shocks to preferences, technologies, and policies (whose realizations are not observable for free to firms) as well as shocks to the pricing frictions themselves, interpreted as shocks to attention or efficiency of

TABLE VI: Calibrated Parameters

Parameter	Value	
Habit in consumption	$h = 0.67$	Havranek et al. (2017)
Wage rigidity	$\delta_w = 0.08$	Monthly rigidity
Taylor Rule		Separately estimated
Persistence	$\rho_i = 0.925$	
Inflation coefficient	$\phi_\pi = 1.7$	
GDP growth coefficient	$\phi_y = 1.0$	

Notes: The table reports the values of the parameters that are calibrated rather than estimated in the dynamic estimation. The procedure used to estimate the Taylor rule parameters is detailed in Appendix D.

information processing and implementation of decisions. We set habit formation in consumption to $h = 0.67$, based on the meta-study of DSGE models of Havranek, Rusnak & Sokolova (2017) and, based on typical values, real wage rigidity implies that wages change once a year on average ($\delta_w = 0.083$). Additionally, we run a separate Bayesian estimation for the parameters of the Taylor rule for the monetary authority. Table VI reports the parameter values we use, and Appendix D describes the estimation of the Taylor rule coefficients in more detail.

6.2 Data

For the model estimation, we use standard macroeconomic data in the estimation of DSGE models, and we complement that data with time series for moments of the distribution of price changes computed based on the individual price quotes underlying the CPI. Appendix A provides a detailed description of data sources and transformations.

We use quarterly real GDP growth per capita, quarterly CPI inflation, the quarterly average for the federal funds rate, and the 2 and 5 years treasury yields. Following Kulish et al. (2017), we include long-term interest rates in our estimation to aid with identification at the Effective Lower Bound. For simplicity, we map the long-term interest rates to our model based on the expectations hypothesis by relating yields on long-term bonds with agent's beliefs about the future path of the federal funds rate:

$$2 \text{ year treasury yield} = c^2 + \frac{1}{24} E_t \left[\sum_{j=0}^{23} i^{12}_{t+j} \right] + \eta^2_t + \eta^y_t$$
$$5 \text{ year treasury yield} = c^5 + \frac{1}{60} E_t \left[\sum_{j=0}^{59} i^{12}_{t+j} \right] + \eta^5_t + \eta^y_t$$

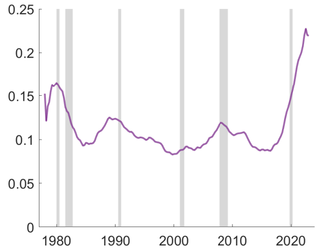
where c^2 and c^5 are yield specific and time invariant components set to match the average difference between the federal funds rate and these yields. η^2, and η^5 are i.i.d. yield specific shocks, and η^y is an exogenous component, common to all yields, that follows an AR(1) process.

In addition to these standard macroeconomic series, we use the time series for five moments of the distribution of price changes in the US. These are the frequency of price changes, the mean absolute size, the standard deviation, the skewness, and the kurtosis of the log-price changes in the Consumer Price Index. As mentioned in the steady state estimation, these statistics are based on the individual price quotes underlying the CPI, as constructed by Nakamura et al. (2018) for the sample starting in January 1978 and ending in December 2014, and extended by Montag & Villar (2023) to March 2023.[10]

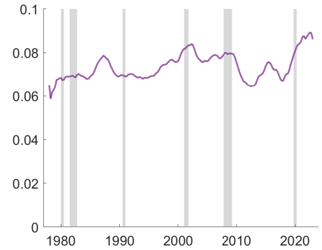
Figure 5 plots the time series for the pricing moments used in the estimation. Of particular interest, note that the frequency, size, and standard deviation of log-price changes all started increasing and kurtosis started decreasing well before the 2020 recession. While neither of these moments is sufficient on its own to pin down non-neutrality, these synchronized movements already suggest that we might expect a change in the degree of price flexibility starting around 2016.

Other notable patterns in these pricing series are (*i*) a relatively high and volatile frequency of price adjustment, ranging between 10% and over 20% monthly, (*ii*) a more stable size of price changes, ranging between 6% and 8% per month, (*iii*) a rising dispersion in price changes, measured by the standard deviation of price changes, which increases from less than 10% at the beginning of the sample to more than 15% by the end of the sample, (*iv*) volatile skew and kurtosis, both trending down over time.

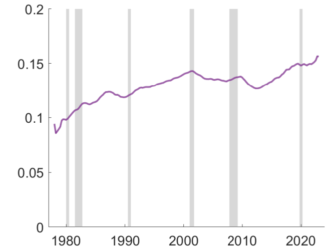
[10]We again thank Daniel Villar for sharing the time series of these pricing moments with us.



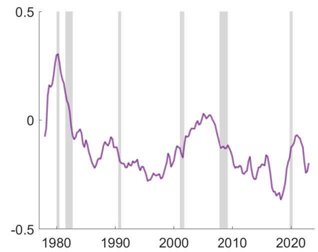
(a) Frequency



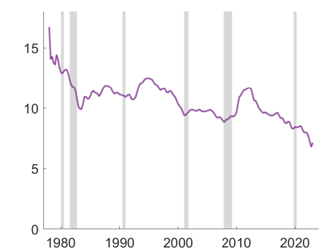
(b) Mean Absolute Size



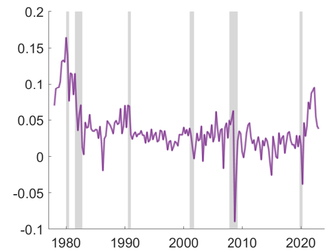
(c) Standard Deviation



(d) Skew



(e) Kurtosis



(f) Inflation

Figure 5: Pricing Moments Over Time

Note: This figure plots the smoothed pricing series used for the model estimation. Shaded areas mark NBER recession dates.

6.3 Prior and Posterior Distributions

We include a range of fundamental shocks, to avoid overstating the role the pricing frictions play in generating aggregate volatility. In addition to shocks to pricing frictions, the aggregate exogenous shocks are to idiosyncratic risk ($\sigma_{a,t}$), aggregate TFP (a_t), impatience (ζ_t), labor supply (ξ_t), bond demand (χ_t), markups (ε), monetary policy (ϵ_t), trend inflation (π_t^*), permanent productivity growth (γ_t), and shocks to yield curve ($\eta_t^2, \eta_t^5, \eta_t^y$). Table VII reports the prior distributions for these parameters along with the posterior mode and standard deviation.

Regarding the standard deviation of the shock innovations, we selected a normal distribution for shocks affecting the pricing parameters, long-term interest rates, and labor supply. Following Ferroni, Grassi & León-Ledesma (2019), we selected this prior, whose support includes zero, to avoid estimating "nonexistent" shocks, as this prior distribution doesn't force the exogenous processes to be different from zero. For all other standard deviations, we selected a standard inverse gamma distribution. In addition to these priors, following Del Negro & Schorfheide (2008), we impose an implicit prior over the model-implied variance of the observable variables and the covariance between the federal funds rate, GDP growth, and inflation. In particular, our prior states that those specific model-implied-covariances follow a normal distribution with parameters determined by the data moments.[11] Finally, to avoid exploring unreasonable areas of the parameter space, we impose that the the filtered value of the menu cost is not below zero and that the filtered log-deviations for θ_p and θ_a are not greater than three in absolute value. More details can be found in Appendix E.

Of particular interest, a one standard innovation to θ_p, θ_a, and κ would result in increases of approximately 3.6, 1.8, and 3.4 percent, respectively.[12]

6.4 Pricing Frictions and NPR Over Time

Figure 6 plots the estimated series for θ^p, θ^a, and κ when the model parameters are evaluated at their posterior mode. We find meaningful variation in these parameters over time, despite allowing for movements in markups and the volatility of idiosyncratic shocks and not forcing these pricing parameters to be non-zero.[13]

We estimate a large and variable repricing cost that has remained above 1 for most of the sample period, indicating substantial pricing frictions. The repricing cost was not particularly elevated at the beginning of the sample, despite the high inflation of the time. Similarly, while the repricing cost appears to rise in most recessions, this cost reached historically high values during booming or recovery periods. In particular, the mid-1990s and the early-2010s. It is noticeable that the repricing cost has been following since 2016 and reached historically low levels. The cost of timing price changes is much smaller and only had one modest increase in the mid-1990s and one large spike in the early-2010s. These spikes coincide with

[11]For example, we imposed that the model-implied variance for the federal funds rate distributes normal with mean (μ_i) and standard deviation (σ_i) equal to the mean and ω times standard deviation of the quadratic deviation of the federal funds rate with respect to the model's steady state, where ω is a scalar that we set to 0.25. Hence, $\mu_i = \sum_t x_t^i/T$ and $\sigma_i^2 = \omega^2 \sum_t \left(x_t^i - \mu_i\right)^2/T$, where $x_t^i = \left(i_t^{data} - i_{ss}\right)^2$. Appendix E presents formally this implicit prior.

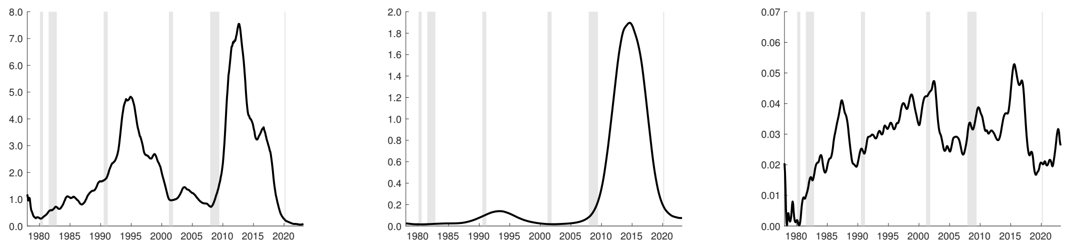
[12]For κ this is given by $0.091/\kappa_{ss} = 3.43$.

[13]Figure E.1 in Appendix E shows the filtered times series for ϵ_t and $\sigma_{a,t}$ along with their data decomposition.

TABLE VII: Model Priors and Posteriors

Parameter Name		Prior Dist	Prior Mode	Prior SD	Posterior Mode	Posterior SD	90% HPD	
Standard deviation of shock innovations (x100)								
Price setting cost	θ^p	norm	1.000	0.500	3.538	0.077	3.445	3.649
Price review cost	θ^a	norm	1.000	0.500	1.795	0.100	1.710	1.987
Menu cost	κ	norm	0.026	0.013	0.091	0.001	0.090	0.093
Idiosyncratic risk	σ	norm	0.008	0.002	0.034	0.001	0.033	0.034
Markup	ε	invg	0.055	0.014	5.694	0.025	5.659	5.725
Permanent TFP	γ	invg	0.190	0.048	1.434	0.049	1.362	1.479
Impatience	ζ	invg	1.118	0.280	2.524	0.028	2.491	2.562
Transitory TFP	a	invg	0.165	0.041	19.147	0.120	19.022	19.330
Bond demand	ξ	invg	2.333	0.583	99.960	0.315	99.248	99.966
Labor supply	χ	norm	0.000	0.041	0.004	0.024	0.005	0.065
Trend inflation	π^*	invg	0.116	0.029	0.117	0.041	0.084	0.194
Monetary policy	i	invg	0.083	0.021	0.574	0.038	0.520	0.615
Term premia	η^y	norm	0.000	1.000	0.667	0.036	0.626	0.721
2yrs yield	η^2	norm	1.211	0.605	1.480	0.089	1.404	1.632
5yrs yield	η^5	norm	1.138	0.569	0.011	0.047	0.011	0.131
Autocorrelation								
Price setting cost	θ^p	beta	0.500	0.150	0.959	0.001	0.957	0.961
Price review cost	θ^a	beta	0.500	0.150	0.980	0.002	0.977	0.981
Menu cost	κ	beta	0.500	0.150	0.854	0.003	0.850	0.857
Idiosyncratic risk	σ	beta	0.500	0.150	0.958	0.001	0.956	0.959
Markup	ε	beta	0.500	0.150	0.002	0.001	0.001	0.005
Permanent TFP	γ	beta	0.025	0.150	0.003	0.024	0.003	0.054
Impatience	ζ	beta	0.500	0.150	0.974	0.001	0.973	0.975
Transitory TFP	a	beta	0.900	0.150	0.001	0.000	0.000	0.002
Bond demand	ξ	beta	0.900	0.150	0.502	0.003	0.498	0.505
Labor supply	χ	beta	0.025	0.150	0.032	0.120	0.028	0.347
Trend inflation	π^*	beta	0.500	0.150	0.401	0.148	0.362	0.798
Term premia	η^y	beta	0.500	0.150	0.976	0.003	0.971	0.980

Notes: norm and invg refer to the normal and inverse gamma distributions, respectively. HPD: High Probability Density.



(a) θ^p (b) θ^a (c) Standard Deviation

Figure 6: Estimated Pricing Frictions

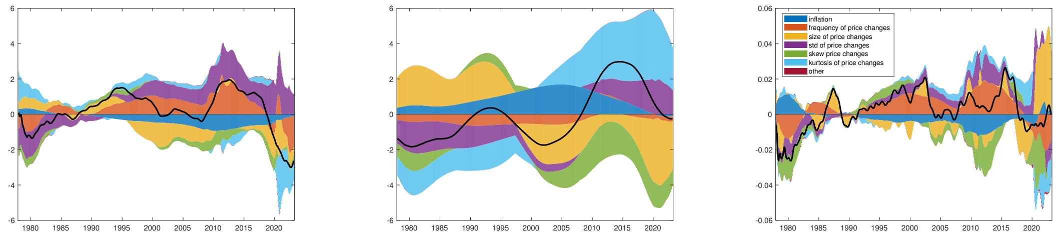
Note: This figure plots the estimated series for θ^p, θ^a, and κ when the model parameters are evaluated at the optimization mode. Shaded areas represent NBER recession dates.

the large increases in the repricing cost. Finally, we estimate a small menu cost over our sample period that was almost zero early in our sample and that has exhibit modest fluctuations since the mid-1980s.

Figure 7 plots the data decomposition for our pricing parameters, which allows us to gauge what features of our data inform the dynamics of our pricing parameters. Some notable features of these figures are that (*i*) most of the variation in these pricing parameters is informed by the moments of the distribution of price changes. The rest of the variation is almost entirely explained by inflation. (*ii*) While some moments seem to be, in general, more informative about the pricing parameters, all moments used in the estimation contain meaningful information about the dynamics of the pricing parameters. And (*iii*) the relative importance of different moments of the price change distribution varies over time. For example, while the frequency and standard deviation of log-price changes informed the spikes in θ^p during the mid-2010s, the decline in θ^p since 2016 was mainly informed by the size and kurtosis of price changes.

What do these dynamics implied for pricing rigidities? Figure 8 reports the estimated degree of NPR in each period by computing the cumulative impulse response (CIR) of consumption to a 25 basis points change in the federal funds rate, given the estimated pricing parameters for that period. The series reports the cumulative response as a percentage of annual steady state consumption.[14]

[14]We keep all other parameters at their baseline value.



(a) θ^p (b) θ^a (c) κ

Figure 7: Estimated Pricing Frictions: Data Decomposition

Note: This figure plots the data decomposition for the estimated series for θ^p, θ^a, and κ when the model parameters are evaluated at the optimization mode. θ^p and θ^a are log-deviation from steady state values, and κ is deviation from steady-state value.

A larger value of the CIR implies stronger monetary non-neutrality. For each point in time, we solve the model using the pricing parameters at that time and recalculating the reference probability of price adjustment $\bar{\Lambda}$ and the reference distribution of prices $\bar{f}$ while keeping all other parameters at their baseline values.

To gauge the importance of each type of friction in this estimate, Figure 9 shows what our estimated series would look like if we had only seen movements in one of the margins, while the other two margins remain at their steady state values. This figure shows that variations in the repricing cost have contributed the most to variations in the degree of pricing rigidities, that variations in the cost of timing price changes are only noticeable in the mid-2010s, and that variations in the menu cost have contributed little to the degree of pricing rigidities.

Surprisingly, we find no trend over time in the CIR, despite the fact that technology has arguably made repricing easier. However, we do find substantial volatility in the degree of NPR over time. A growing literature has argued that monetary policy may be less effective during downturns.[15] Our results suggest that this chain of reasoning does not always apply during US recessions. For example, we estimate that price rigidities rose modestly during the Great Recession.

Our estimation shows that price rigidities reached a maximum value in the mid-2010s, a period in which the pricing cost and the cost of timing price changes were

[15] This argument is based on evidence that price dispersion increases during downturns, which implies more price flexibility in existing models.

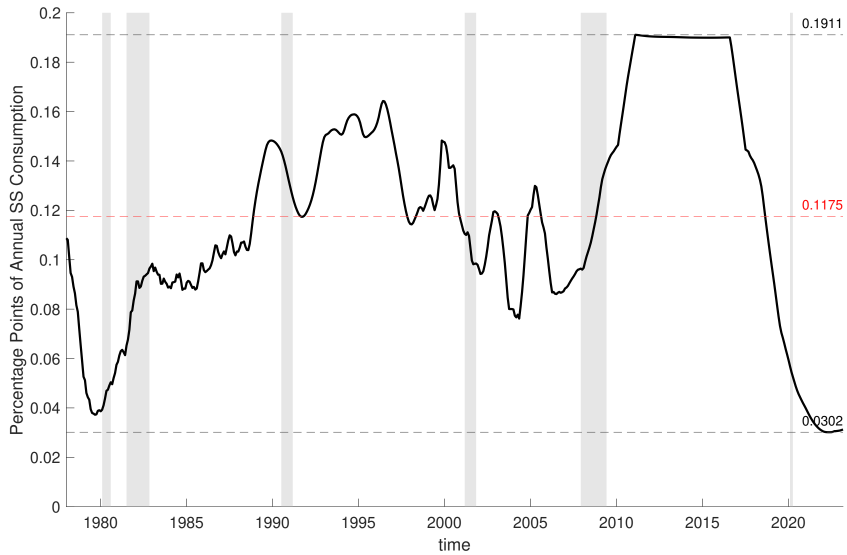


Figure 8: Implied Nominal Rigidity Over Time

Note: This figure plots the model-implied degree of nominal price rigidities over time (NPR). The horizontal dashed lines mark the minimum, mean, and maximum values realized over the sample period. Shaded areas represent NBER recession dates. The NPR series was smoothed with a centered moving average of 13 months.

so high that an additional increase in those costs could not make price rigidities increase. Our estimation also suggests that price rigidities started to decline well before the COVID-19 pandemic, pointing to the risk of higher inflation even before the subsequent shocks.

How to reconcile our findings with the growing literature that has documented diminished monetary policy effectiveness in downturns? We believe the answers lies in the distinction between the transmission of a central bank's policy to the short term real interest rate and the effect that rate has on consumption, investment, employment, and other macroeconomic variables. In this article, we concentrate on the first part, while the cyclicality and state-dependence of the second part has been the focus of papers such as those of Berger, Milbradt, Tourre & Vavra (2021); Eichenbaum, Rebelo & Wong (2022); McKay & Wieland (2021).

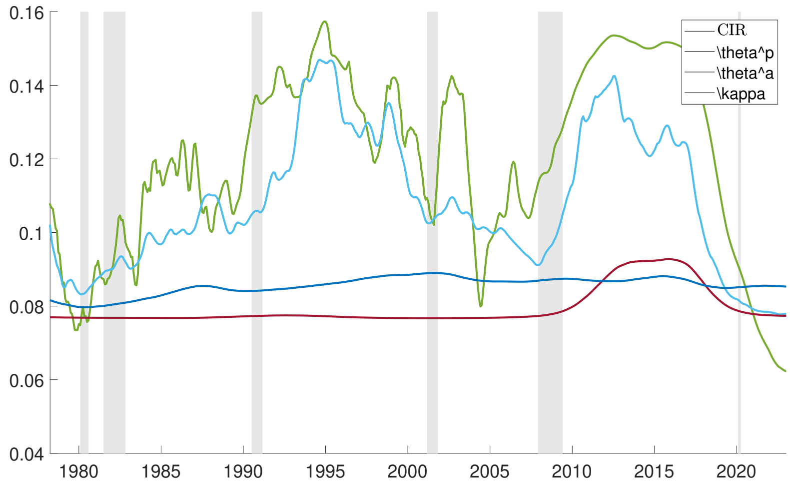


Figure 9: Role of Frictions in Nominal Rigidity Variation Over Time

6.5 Economic Interpretation of Pricing Parameters and NPR Variability

We showed that our pricing parameters exhibit meaningful fluctuations over time that are informed by different moments of the pricing distributions. In some recessions, like in the Great Recession, these parameters started to increase suggesting either higher uncertainty about the right price to set, or fewer resources devoted to price-setting (perhaps as managers substitute away from managing prices to managing liquidity and cutting costs). But these parameters reached high values during booming or recovering periods: mid-1990s and mid-2010s. How can we interpret these dynamics? Figure evidence suggesting that pricing parameters are negatively correlated with measures of aggregate uncertainty. This evidence suggests that period of low uncertainty are periods in which the benefit of accurate pricing and timing of price changes is relatively low and fewer resources are devoted to price-setting.

7 Conclusion

This article estimates the degree of nominal price rigidity in the U.S. economy between 1978 and 2023. We identify costly information as the main friction that

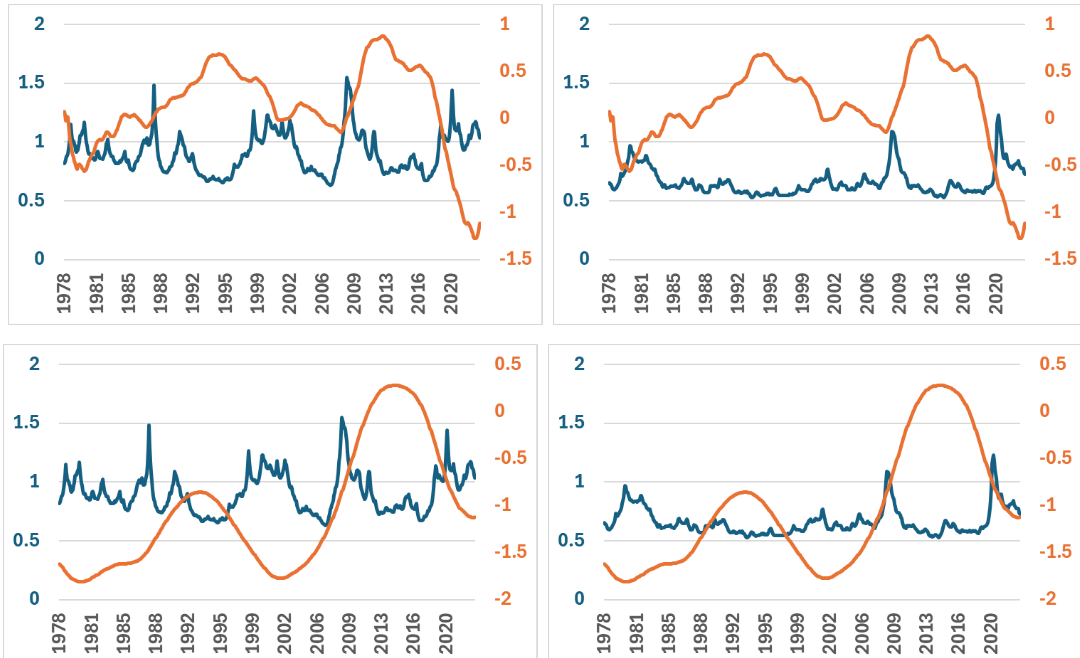


Figure 10: Information frictions and aggregate uncertainty

prevents firms from adjusting prices more flexibly in response to shocks, with information about the right price to charge, conditional on adjustment, being the most significant driver of price rigidity. These results contribute to our understanding of how efficiently the U.S. economy has adjusted to shocks in recent decades, and how effectively policymakers have stabilized aggregate demand.

On net, what do our results suggest for inflation and monetary control going forward? We emphasize the endogeneity and variability in the degree of state-dependence in price setting: First, our estimation results give great weight to firms' choices of how much attention to devote to choosing prices accurately. They suggest that while firms generally know with fairly high accuracy when their prices are outdated, they are much less certain about what the right price to charge is. Second, we find that firms' attention to market conditions is variable over time. This variability implies state-dependence in the cost of disinflation over time.

More work is needed to measure the attention firms devote to price setting versus other operational decisions. But our finding that mispricing is a major driver of monetary non-neutrality connects models of nominal rigidities to the much broader literature that has documented stochasticity in choice in a wide range of contexts. While stochastic choice may appear at odds with classic principles of optimization of well-specified objective functions, in this paper, we microfound it with rational acquisition of costly information. But it is worth separating the stochasticity result from the model through which we endogenize this stochasticity. We leave to future research other possible sources of randomness in decision-making (*e.g.*, deliberate, or exploratory randomization, model uncertainty, or responding to consumer constraints). The important message is that whatever its source, the consequence of stochastic choice is often a systematic bias in the response of the aggregate price level to shocks. Stochasticity need not be divorced from but can rather be understood as a cause of bias (Woodford, 2020).

References

Afrouzi, Hassan (2020), "Strategic inattention, inflation dynamics, and the non-neutrality of money," CESifo working paper.

Afrouzi, Hassan & Choongryul Yang (2021), "Dynamic rational inattention and the Phillips curve," *SSRN 3770462* .

Alvarez, Fernando, Martin Beraja, Martin Gonzalez-Rozada & Pablo Andrés Neumeyer (2019), "From hyperinflation to stable prices: Argentina's evidence on menu cost models," 134(1): 451–505.

Alvarez, Fernando, Hervé Le Bihan & Francesco Lippi (2016), "The real effects of monetary shocks in sticky price models: a sufficient statistic approach," *American Economic Review* 106(10): 2817–51.

Alvarez, Fernando & Francesco Lippi (2014), "Price setting with menu cost for multi-product firms," *Econometrica* 82(1): 89–135.

Alvarez, Fernando, Francesco Lippi & Luigi Paciello (2011), "Optimal price setting with observation and menu costs," *The Quarterly Journal of Economics* 126(4): 1909–1960.

Angeletos, George-Marios, Zhen Huo & Karthik A Sastry (2021), "Imperfect macroeconomic expectations: Evidence and theory," 35(1): 1–86.

Angeletos, George-Marios & Jennifer La'O (2009), "Incomplete information, higher-order beliefs and price inertia," *Journal of Monetary Economics* 56: S19–S37.

Aruoba, S Borağan, Eugene Oue, Felipe Saffie & Jonathan Willis (2024), "Non-Constant Demand Elasticities, Firm Dynamics and Monetary Non-Neutrality: Role of Demand Shocks," Tech. rep., National Bureau of Economic Research.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie & Ludwig Straub (2021), "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models," *Econometrica* 89(5): 2375–2408.

Auclert, Adrien, Rodolfo D Rigato, Matthew Rognlie & Ludwig Straub (2022), "New Pricing Models, Same Old Phillips Curves?" Tech. rep., National Bureau of Economic Research.

Auclert, Adrien, Matthew Rognlie & Ludwig Straub (2020), "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model," Tech. rep., National Bureau of Economic Research.

Bayer, Christian, Benjamin Born & Ralph Luetticke (2020), "Shocks, frictions, and inequality in U.S. business cycles," Forthcoming, The American Economic Review.

Berger, David, Konstantin Milbradt, Fabrice Tourre & Joseph Vavra (2021), "Mortgage prepayment and path-dependent effects of monetary policy," *American Economic Review* 111(9): 2829–78.

Berger, David & Joseph Vavra (2018), "Dynamics of the US price distribution," *European Economic Review* 103: 60–82.

Bilbiie, Florin O, Giorgio Primiceri & Andrea Tambalotti (2023), "Inequality and business cycles," Tech. rep., National Bureau of Economic Research.

Bils, Mark & Peter J Klenow (2004), "Some evidence on the importance of sticky prices," *Journal of Political Economy* 112(5): 947–985.

Caballero, Ricardo J & Eduardo MRA Engel (2007), "Price stickiness in Ss models: New interpretations of old results," *Journal of Monetary Economics* 54: 100–121.

Calvo, Guillermo A (1983), "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics* 12(3): 383–398.

Campbell, Jeffrey R & Benjamin Eden (2014), "Rigid prices: Evidence from us scanner data," *International Economic Review* 55(2): 423–442.

Campbell, Jeffrey R, Jonas DM Fisher, Alejandro Justiniano & Leonardo Melosi (2017), "Forward guidance and macroeconomic outcomes since the financial crisis," *NBER Macroeconomics Annual* 31(1): 283–357.

Carroll, Christopher D (2003), "Macroeconomic expectations of households and professional forecasters," *The Quarterly Journal of Economics* 118(1): 269–298.

Carvalho, Carlos, Niels Arne Dam & Jae Won Lee (2020), "The cross-sectional distribution of price stickiness implied by aggregate data," *Review of Economics and Statistics* 102(1): 162–179.

Cavallo, Alberto, Guillermo Cruces & Ricardo Perez-Truglia (2017), "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments," 9(3): 1–35.

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

Coibion, Olivier & Yuriy Gorodnichenko (2015), "Information rigidity and the expectations formation process: A simple framework and new facts," 105(8): 2644–78.

Costain, James & Anton Nakov (2015), "Precautionary price stickiness," *Journal of Economic Dynamics and Control* 58: 218–234.

Costain, James & Anton Nakov (2019), "Logit price dynamics," *Journal of Money, Credit and Banking* 51(1): 43–78.

Cuba-Borda, Pablo & Sanjay R Singh (2019), "Understanding persistent stagnation," *Available at SSRN 3247350* .

Del Negro, Marco & Frank Schorfheide (2008), "Forming priors for DSGE models (and how it affects the assessment of nominal rigidities)," *Journal of Monetary Economics* 55(7): 1191–1208, ISSN 0304-3932.

Dotsey, Michael, Robert G. King & Alexander L. Wolman (1999), "State-dependent pricing and the general equilibrium dynamics of money and output," *Quarterly Journal of Economics* 114: 655–690.

Eggertsson, Gauti B & Michael Woodford (2003), "The zero bound on interest rates and optimal monetary policy," *Brookings Papers on Economic Activity* 2003(1): 139–233.

Eichenbaum, Martin, Sergio Rebelo & Arlene Wong (2022), "State-dependent effects of monetary policy: The refinancing channel," *American Economic Review* 112(3): 721–61.

Ferroni, Filippo, Stefano Grassi & Miguel A. León-Ledesma (2019), "Selecting structural innovations in DSGE models," *Journal of Applied Econometrics* 34(2): 205–220.

Fisher, Jonas DM (2015), "On the structural interpretation of the Smets–Wouters risk premium shock," *Journal of Money, Credit and Banking* 47(2-3): 511–516.

Foster, Lucia, John Haltiwanger & Chad Syverson (2008), "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?" *American Economic Review* 98(1): 394–425.

Gagliardone, Luca, Mark Gertler, Simone Lenzu & Joris Tielens (2023), "Anatomy of the Phillips Curve: Micro Evidence and Macro Implications," Tech. rep., National Bureau of Economic Research.

Gagnon, Etienne (2009), "Price setting during low and high inflation: Evidence from Mexico," *The Quarterly Journal of Economics* 124(3): 1221–1263.

Golosov, Mikhail & Robert E Lucas Jr (2007), "Menu costs and Phillips curves," *Journal of Political Economy* 115(2): 171–199.

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

Havranek, Tomas, Marek Rusnak & Anna Sokolova (2017), "Habit formation in consumption: A meta-analysis," *European Economic Review* 95: 142–167.

Jurado, Kyle, Sydney C Ludvigson & Serena Ng (2015), "Measuring uncertainty," *American Economic Review* 105(3): 1177–1216.

Justiniano, Alejandro, Giorgio E Primiceri & Andrea Tambalotti (2010), "Investment shocks and business cycles," *Journal of Monetary Economics* 57(2): 132–145.

Karadi, Peter, Raphael Schoenle & Jesse Wursten (2024), "Price selection in the microdata," *Journal of Political Economy Macroeconomics* 2(2).

Khaw, Mel Win, Luminita Stevens & Michael Woodford (2017), "Discrete adjustment to a changing environment: Experimental evidence," *Journal of Monetary Economics* 91: 88–103.

Khaw, Mel Win, Luminita Stevens & Michael Woodford (2019), "Empirical models of inattentive choice," Working Paper.

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

Krishnamurthy, Arvind & Annette Vissing-Jorgensen (2012), "The aggregate demand for treasury debt," *Journal of Political Economy* 120(2): 233–267.

Kulish, Mariano, James Morley & Tim Robinson (2017), "Estimating DSGE models with zero interest rate policy," *Journal of Monetary Economics* 88: 35–49.

Ludvigson, Sydney C, Sai Ma & Serena Ng (2021), "Uncertainty and business cycles: exogenous impulse or endogenous response?" *American Economic Journal: Macroeconomics* 13(4): 369–410.

Luo, Shaowen & Daniel Villar (2021), "The skewness of the price change distribution: A new touchstone for sticky price models," *Journal of Money, Credit and Banking* 53(1): 41–72.

Maćkowiak, Bartosz A. & Mirko Wiederholt (2009), "Optimal sticky prices under rational inattention," *The American Economic Review* 99(3): 769–803.

Magnani, Jacopo, Aspen Gorry & Ryan Oprea (2016), "Time and state dependence in an Ss decision experiment," *American Economic Journal: Macroeconomics* 8(1): 285–310.

Mankiw, N Gregory, Ricardo Reis & Justin Wolfers (2003), "Disagreement about inflation expectations," 18: 209–248.

Matějka, Filip (2015), "Rationally inattentive seller: Sales and discrete pricing," *The Review of Economic Studies* 83(3): 1125–1155.

McKay, Alisdair & Johannes F Wieland (2021), "Lumpy durable consumption demand and the limited ammunition of monetary policy," *Econometrica* 89(6): 2717–2749.

Melosi, Leonardo (2014), "Estimating models with dispersed information," *American Economic Journal: Macroeconomics* 6(1): 1–31.

Michaillat, Pascal & Emmanuel Saez (2021), "Resolving New Keynesian anomalies with wealth in the utility function," *Review of Economics and Statistics* 103(2): 197–215.

Midrigan, Virgiliu (2011), "Menu costs, multiproduct firms, and aggregate fluctuations," *Econometrica* 79(4): 1139–1180.

Montag, Hugh & Daniel Villar (2023), *Price-Setting During the Covid Era*, FEDS Notes.

Nakamura, Emi & Jón Steinsson (2008), "Five facts about prices: A reevaluation of menu cost models," *The Quarterly Journal of Economics* 123(4): 1415–1464.

Nakamura, Emi & Jón Steinsson (2010), "Monetary non-neutrality in a multisector menu cost model," *The Quarterly Journal of Economics* 125(3): 961–1013.

Nakamura, Emi, Jón Steinsson, Patrick Sun & Daniel Villar (2018), "The elusive costs of inflation: Price dispersion during the US Great Inflation," *The Quarterly Journal of Economics* 133(4): 1933–1980.

Nimark, Kristoffer (2008), "Dynamic pricing and imperfect common knowledge," *Journal of Monetary Economics* 55(2): 365–382.

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

Reiter, Michael (2009), "Solving heterogeneous-agent models by projection and perturbation," *Journal of Economic Dynamics and Control* 33(3): 649–665, ISSN 0165-1889.

Shannon, Claude E (1948), "A mathematical theory of communication," *The Bell System Technical Journal* 27(3): 379–423.

Shannon, Claude E. (1959), "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec* 4(142-163): 1.

Sims, Christopher A. (2003), "Implications of rational inattention," *Journal of Monetary Economics* 50(3): 665–690.

Sims, Eric (2013), "Growth or the Gap? Which Measure of Economic Activity Should be Targeted in Interest Rate Rules?" *Mimeograph, University of Notre Dame* .

Smets, Frank & Rafael Wouters (2007), "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American economic review* 97(3): 586–606.

Turen, Javier (2023), "State-dependent attention and pricing decisions," *American Economic Journal: Macroeconomics* 15(2): 161–189.

Vavra, Joseph (2013), "Inflation dynamics and time-varying volatility: New evidence and an Ss interpretation," *The Quarterly Journal of Economics* 129(1): 215–258.

Veldkamp, Laura & Cindy Chung (2024), "Data and the aggregate economy," *Journal of Economic Literature* 62(2): 458–484.

Woodford, Michael (2003), "Imperfect common knowledge and the effects of monetary policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, J. Stiglitz P. Aghion, R. Frydman & M. Woodford, eds., pp. 25–58, Princeton University Press, Princeton, NJ.

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

Woodford, Michael (2012), "Inattentive valuation and reference-dependent choice," Unpublished Manuscript.

Woodford, Michael (2020), "Modeling imprecision in perception, valuation, and choice," *Annual Review of Economics* 12: 579–601.

A Data Description

We set our model to a monthly frequency and aggregate the model-simulated series to a quarterly frequency to align the model-simulated data with the U.S. data. In these notes, the time subscript t refers to a month, and $\tilde{x}$ is variale x detrended by trend-productivity.

Growth rate of real per-capita GDP Data for quarterly real GDP of chained 2012 dollars and seasonally adjusted was retrieved from FRED (series GDPC1). This series is divided by the quarterly average of the monthly civilian noninstitutional population, 16 yr+ available at FRED (series CNP16OV). Growth rates are computed by log differences (quarterly growth rate). Hence, the measurement equation for quarterly GDP growth is given by:

$$\begin{aligned} dy_t^q =& log\left(G\tilde{D}P_t + \frac{G\tilde{D}P_{t-1}}{\gamma_{zt}} + \frac{G\tilde{D}P_{t-2}}{\gamma_{zt-1}}\right) \\ &- log\left(G\tilde{D}P_{t-3} + \frac{G\tilde{D}P_{t-4}}{\gamma_{zt-3}} + \frac{G\tilde{D}P_{t-5}}{\gamma_{zt-4}}\right) \\ &+ log\left(\gamma_{zt}\gamma_{zt-1}\gamma_{zt-2}\right) \end{aligned} \tag{A.1}$$

where

$$G\tilde{P}D_t = \tilde{c}_t - \tilde{G}_t \tag{A.2}$$

Quarterly CPI inflation rate Data for the Consumer Price Index for All Urban Consumer was retrieved from FRED (series CPIAUCSL). We take the quarterly average of this monthly series and compute quarterly inflation as the ratio of the quarterly CPI index between two months minus one. Hence, our measurement equation is links the quarterly inflation in the data (π_t^q) to the model variables as follows:

$$\pi_t^q = \pi_t \cdot \pi_{t-1} \cdot \pi_{t-2} - 1 \tag{A.3}$$

Quarterly Federal Funds Rate We retrieve the Federal Funds rate from FRED (series DFF) and take the quarterly average. Hence, the quarterly federal funds rate in the data (i_t^q) is linked to our model variables as follows:

$$i_t^q = \frac{\left(i_t^{12} + i_{t-1}^{12} + i_{t-2}^{12}\right)}{3} \tag{A.4}$$

Quarterly 2 year Treasury Yields We retrieve data for the Market Yield on U.S. Treasury Securities at 2-Year Constant Maturity from FRED (series DGS2). We take the quarterly average of this series. The measurement equation for this series (i_t^{y2}) is linked to the model variables as follows:

$$i_t^{y2} = c_2 + E_t\left[\sum_{j=0}^{23} i_{t+j}^{12}\right] + \eta_t^2 + \eta_t^y \tag{A.5}$$

where c^2 is a constant, η_t^2 is a yield specific i.i.d. shock, and η_t^y is an exogenous shock common to all shocks that follows an AR(1) process.

Quarterly 5 year Treasury Yields We retrieve data for the Market Yield on U.S. Treasury Securities at 5-Year Constant Maturity from FRED (series DGS5). We take the quarterly average of this series. The measurement equation for this series (i_t^{y5}) is linked to the model variables as follows:

$$i_t^{y5} = c_5 + E_t\left[\sum_{j=0}^{59} i_{t+j}^{12}\right] + \eta_t^5 + \eta_t^y \tag{A.6}$$

where, as before, c^5 is a constant, and η_t^5 is a yield specific i.i.d. shock.

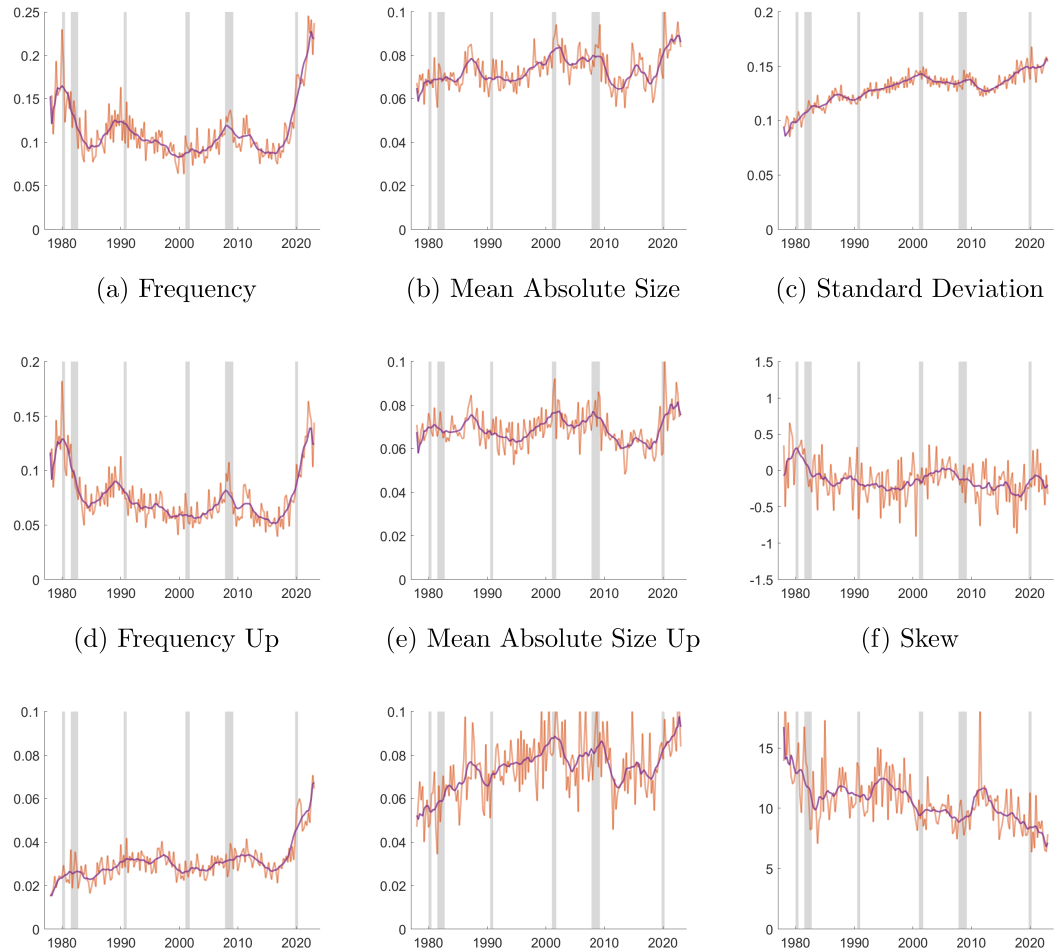
Pricing Moments We use 7 moments of the distribution of log-price changes based on the individual price quotes underlying the CPI, as constructed by Nakamura et al. (2018) for the sample starting in January 1978 and ending in December 2014, and extended by Montag & Villar (2023) to March 2023. Those moments are: frequency, size, standard deviation, skewness, and kurtosis of all log-price changes, and the frequency and size of log-price increases. Daniel Villar kindly shared these aggregate series with us. We take the quarterly average of these series, and we smooth them using a **XXX**. In the steady state estimation, we make use of all these moments. In the Bayesian estimation we make use of the 5 moments describing log-price changes. We link these moments to our model variables by computing these moments for the non-zero log-price changes in our model. Figure (A.1) plots the time series of the different pricing moments we use, both raw and smoothed. The estimation uses the smoothed series. The distribution of log-price changes is based on approximately 80,000 observations per month. Goods are classified in approximately 305 categories, or "entry-level items" (ELIs), which are then further aggregated into 13 major groups. Authors with access to the microdata can use the individual price quotes to construct empirical distributions of log-price changes for each month, from which various pricing statistics are then calculated. The frequency, size, and standard deviation of all log-price changes along with the frequency and size of log-price increases are computed across products within each entry-level item (ELI), and then the expenditure-weighted median across ELIs to compute an aggregate measure. The skewness and kurtosis of all log-price changes are computed by pooling data within each *major group*, rather than within each ELI, and then taking an expenditure-weighted average across the 13 major groups. Luo & Villar (2021) discuss the need to compute higher moments at the group level, due to sample size constraints at the ELI level: Since ELIs are narrowly defined categories, they sometimes have only a small number of observations. Higher moments are particularly sensitive to outliers, which is why a small number of observations is insufficient to compute them reliably within each ELI.

The Effective Lower Bound The federal funds rate was at the effective zero lower bound (ELB) twice during our sample period: between January 2009 and December 2015, following the Great Recession, and between March 2020 and February 2022, following the Covid Recession.

As explained in Appendix B, for the likelihood evaluation, we need a series with the expected ELB duration at each month in which the federal funds rate was at the ELB. We calibrate that series using the Blue Chip data between 2008 and 2010, and the survey of primary dealers since 2011. For each survey, we compute the expected ELB duration, in months, for each month as described below. Then, our monthly series correspond to the BlueChip series between 2008 and December 2010, and to the survey of professional forecasters since January, 2011.

Using the Blue Chip microdata, we compute the expected ELB duration (in quarters) for each forecaster and month. Blue Chip is a monthly survey, but participants are asked for the expected federal funds rate in quarter intervals. For example, the expected federal funds rate value in 2010Q1. Then, we compute the median expected ELB duration for each month across forecasters. By construction, this expected duration is top-coded, as the Blue Chip survey only asks for the expected federal funds rate value for the next 5 quarters. However, before 2011, the median expected ELB duration is less than 5 quarters. We compute the expected duration in months by assuming a lag of one week between data collection and publication and by using the FOMC meetings calendar. If a quarter has two FOMC meetings, we take the simple average of the expected duration associated with those two meetings.

Since January 2011 and until 2015, the survey of primary dealers asks for the *"Most Likely Quarter and Year of First Target Rate Increase"*, later in our sample, they asked for the specific FOMC meeting. Based on the median answer for those questions, the date in which the survey was received, and the FOMC meetings calendar, we compute the expected ELB duration in months. As with the Blue Chip survey, we take the simple average among the expected durations associated



(a) Frequency (b) Mean Absolute Size (c) Standard Deviation

(d) Frequency Up (e) Mean Absolute Size Up (f) Skew

(g) Frequency Down (h) Mean Absolute Size Down (i) Kurtosis

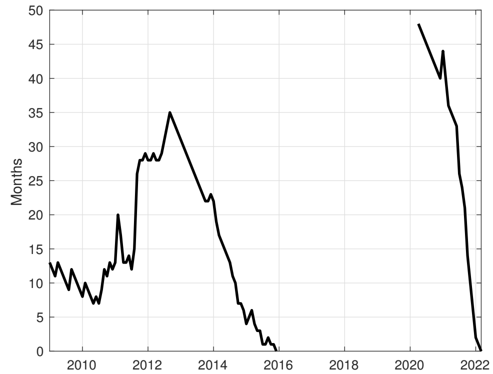
Figure A.1: Data Pricing Moments

Note: These panels plot the raw and MA-smoothed series for the pricing moments based on the U.S. CPI data from 1978 to 2023.Q1. Shaded areas represent NBER recession dates.

with the meetings in each quarter, and we get a monthly series by interpolating and rounding. Figure A.2 plots the expected ELB duration series used for the estimation.

Recession Dates The NBER dating committee lists dates for peaks and troughs in economic activity. In the NBER's convention, the first month of a recession is the month following the peak, and the last month of a recession is the month of a trough. Therefore, we define the start month of a recession as peak plus one

Figure A.2: Expected ELB duration



Note: This Figure plots the expected ELB duration for each month used in the Bayesian estimation of the model. We use the BlueChip survey between 2008 and 2010, and the survey of primary dealers since 2011.

month and the end month of a recession as the trough. For example, in 2020, the peak economic activity was reached in February 2020 and the trough was reached in April 2020, yielding a two-month recession: March and April 2020.

B SSJ, Occasionally Binding ELB Constraint, and Model Estimation

For linear models written in a recursive formulation, Kulish et al. (2017) and Guerrieri & Iacoviello (2015) show that, for an expected ELB duration, the law of motion of the economy can be written as a time-varying linear policy function. As a result, the likelihood evaluation can be computed based the Kalman filter with time-varying coefficients. For example, Kulish et al. (2017) show that DSGE models can be estimated for sample periods including the ELB period by replacing the federal funds rate with a time series of expected ELB durations as an observable. The expected ELB duration does not have to be model consistent. In other words, in absence of any other shocks, the federal funds rate could be expected to be above the ELB at a time period different than the implied by the expected ELB duration, adding another form of monetary policy shocks. Kulish et al. (2017) even propose to estimate the ELB duration.

In the context of this paper, where we use SSJ to solve and estimate the model, how can we estimate the model during the ELB period? We show that one possibility is to recover the time-varying and recursive formulation of the model based on SSJ IRFs for different ELB durations. Then, we can find the (time-varying) MA representation of the model to compute the log-likelihood of the model.[16] This approach is time consuming as the matrix operations can be computationally demanding for large number of state variables and a large number of draws. However, this approach may be preferred for running stochastic simulations.

Instead, we propose a new and efficient way of computing the log-likelihood during the ELB period for a given expected ELB duration using SSJ. This procedure is equivalent to the aforementioned possibility, but our approach is faster. Our approach is to model the expected ELB duration as a sequence of anticipated monetary policy shocks. Hence, for each month that the Federal Funds rate is at the ELB, given a sequence of shocks up to that month, and given the expected duration of the ELB at that month, agents in the economy expected a sequence of anticipated monetary policy shocks for as long as they expected the ELB to bind. First, in section B.1, we show how to compute impulse responses with SSJ when the ELB is expected to bind for m periods, and we show that those responses are identical to what we would get based on the recursive formulation of the model based on Reiter (2009) and Kulish et al. (2017). Second, in section B.2 we show how to get the time-varying recursive formulation of the policy functions based on the IRFs for different ELB durations, and we show that the responses from this method also match the previous methods. In section B.3, we show that these re-

[16]Another possibility is to compute the Kalman filter recursively based on the time-varying policy functions.

sponses are also equivalent to assuming a sequence of anticipated monetary policy shocks for m periods. In section B.4, we combined these results to show how to evaluate the likelihood of the data when the ELB is binding during some periods.

B.1 IRFs when the ELB is binding using SSJ

Assuming that the ELB is not binding, the linearized system of equations describing the equilibrium in SSJ can be written as:

$$F_x dX + F_z dZ = 0 \tag{B.1}$$

where X and Z represent the paths of the endogenous and exogenous variables, respectively. The Jacobians F_x and F_z are stacked matrices that represent different parts of the model. Hence, at the core of our model, the Jacobian F_x is given by:

$$F_x = \left[F_x^P; F_x^{px}; F_x^{ARC}; F_x^{Euler}; F_x^{Taylor}\right] \tag{B.2}$$
$$F_z = \left[F_z^P; F_z^{px}; F_z^{ARC}; F_z^{Euler}; F_z^{Taylor}\right] \tag{B.3}$$

where subscripts refer to the equations related to the aggregate price index (P), the aggregate marginal cost (px), the aggregate resource constraint (ARC), the euler equation (Euler), and the Taylor rule (Taylor). For example, F_x^{Taylor} and F_z^{Taylor} are the Jacobians describing the path of the linearized Taylor rule in response to a path of the endogenous variables (dX) and exogenous variables (dZ). F_x^{Taylor} is a matrix of size T by nT, and F_z^{Taylor} is a matrix of size T by eT, where T is a large horizon for which the Jacobian is computed, n is the number of endogenous variables, and e is the number of exogenous variables. In our case, the endogenous variables are the inflation rate (π), output (Y), consumption (C), interest rates (i), and the marginal cost (px). Hence, in SSJ, the responses of the endogenous variables to unanticipated shocks is given by:

$$dX = -F_x^{-1} F_x dZ \tag{B.4}$$

How to compute the responses when the interest rate is expected to bind for m periods? In this case, we would have to modify the Jacobian F_x^{Taylor} to reflect that the interest rate is at the ELB for the first m periods. Hence, in this case, the linearized system of equations describing the equilibrium is given by:

$$F_x^* dX + F_z^* dZ + C^* = 0 \tag{B.5}$$

where C^* is a column vector of size nT, and F_x^* and F_z^* are identical to F_x and F_z except for those rows describing the Taylor rule. Those Jacobians describing the path of the Taylor rule for the first m periods should reflect that in deviations with respect to the steady state:

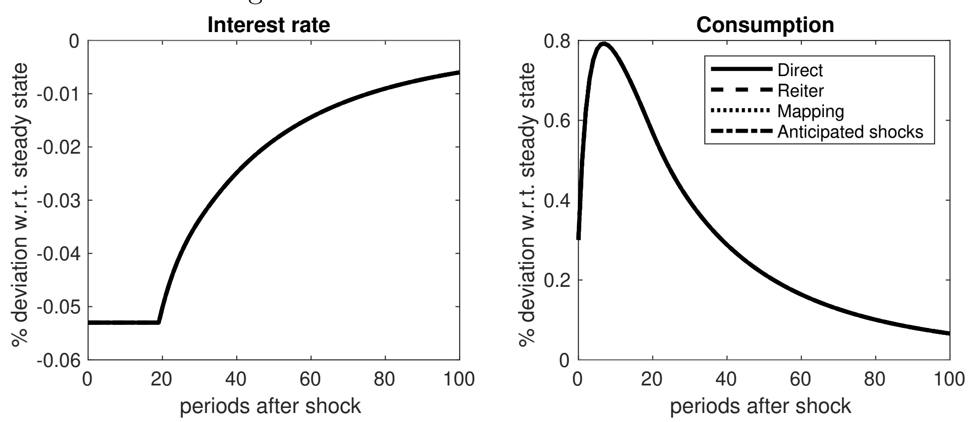
$$\hat{i}_t = -i^{ss} \tag{B.6}$$

Hence, $F_x^{*Taylor}$ will be equal to the identity matrix for the first m periods, and then equal to F_x^{Taylor} for periods between $m+1$ and T. $F_z^{*Taylor}$ is a matrix of zeros, and C^{Taylor} is equal to $-i^{ss}$ for the first m periods and zero everywhere else.[17] In this case, the responses to unanticipated shocks are given by:

$$dX = -F_x^{-1}\left(C^* + F_z^* dZ\right) \tag{B.7}$$

Figure B.1 plots the interest rate and consumption responses to a 1% productivity shock when the interest rate is expected to be at 0 for 20 periods. The solid lines plot the responses based on the method described in this section (which we denote by "direct"), and the dash lines represent the same responses but computed based on the recursive formulation of the model (using the Reiter (2009) method) and then employing the method proposed by Kulish et al. (2017) for the ELB. These lines are almost identical and are on top of each other.[18][19]

Figure B.1: IRFs at the Zero Lower Bound.



Note: This Figure plots the impulses responses to a 1% increase in productivity when the interest rate is expected to be at 0 for 20 periods. "Direct" refers to the approach described in section B.1. "Reiter" refers to the responses computed using the Reiter (2009) Method and Kulish et al. (2017). "Mapping" refer to the computation of the responses using the time-varying and recursive formulation of the model recovered from the SSJ solution as explained in section B.1.

[17]Here we assume that the ELB for the (net) interest rate is equal to 0. If the ELB is different than zero, $\hat{i}_t = -i^{ss} + \underline{i}$, where $\underline{i}$ is the lower bound in the interest rate.

[18]The responses are also identical for all other endogenous variables in response to all exogenous shocks. We also checked that the responses are identical in both cases when the interest rate is expected to be at 0 between periods j and $j+m$, where $j > 1$. Additional graphs are available up to request.

[19]The results presented in this Appendix are based on a smaller grid for the marginal cost, as the computations with the Reiter method were taking hours with a grid size of 41 points for the marginal cost. In this Appendix, we used a grid with 5 points for the marginal cost.

B.2 Recovering time-varying recursive policy functions

Now, given that the goal is to evaluate the likelihood function, how can we use the Kalman filter and SSJ for this purpose? Auclert et al. (2021) show how to recover the unconditional recursive formulation of the policy functions given the IRFs computed based on SSJ. However, when the ELB is binding, the recursive formulation of the policy functions is time-varying. When the ELB is binding, we show that we can use the IRFs for different ELB durations computed in the previous section to recover the time-varying policy functions. In particular, denote the recursive formulation of the policy functions as:

$$X_t = C_t + P_t X_{t-1} + D_t E_t, \tag{B.8}$$

where matrices C_t, P_t, and D_t describe the policy rules at time t when the expected ELB duration is equal to $duration_t$. To recover these matrices from the SSJ solution when the ELB is expected to bind for m periods, one needs to (1) compute the IRFs for an expected duration equal to 1, (2) recover C_1, P_1, and D_1 from that solution following Auclert et al. (2021), and (3) repeat for each duration equal to 2, 3... m.

Dotted lines in Figure B.1 represent the responses computed based on the time-varying and recursive formulation of the model recovered from the SSJ solution, which we denote by "mapping". These three lines are almost identical and are on top of each other.

B.3 ELB and anticipated monetary policy shocks

In this section, we show that responses presented above are equivalent to assuming anticipated monetary policy shocks that guarantee (in expectation) that the interest rate will be at 0 for m periods. While this method is less efficient to compute specific responses, this method is very efficient for the likelihood computation.

The basic idea of this method is the following: suppose that the monetary authority follows a Taylor-type policy rule. But, during ELB periods, the monetary authority activates anticipated shocks: today it announces future changes the the interest rate. Those anticipated shocks are such that the interest rate will be expected to be 0 for m periods.

Formally this procedure works as follows. The linearized system of equations describing the equilibrium in SSJ is given by:

$$F_x dX + F_z dZ + F_n dN = 0 \tag{B.9}$$

where, compared to (B.1), N is the path of the anticipated monetary policy shocks. dN is a column vector of size $T \cdot a$, where a is the number of anticipated monetary policy shocks.[20] F_n is the Jacobian of the endogenous equations with respect to the anticipated monetary policy shocks.[21] Then, the responses of the endogenous variables to shocks is given by:

$$dX = -F_x^{-1}\left(F_z dZ + F_n dN\right) \tag{B.10}$$

From B.10, we can extract the responses of the interest rate, which we can denote by:

$$di = AdZ + BdN \tag{B.11}$$

Now, we want to find the sequence of anticipated monetary policy shocks (dN) that make the expected interest rate be 0 for m periods. Hence, based on (B.11), we can solve for dN such that:

$$di = C^i = AdZ + BdN \tag{B.12}$$

$$dN = -B^{-1}\left(C^i + AdZ\right) \tag{B.13}$$

where C^i is a column vector of size T, with the first m entries equal to $-i^{ss}$.[22] Now, substituting (B.13) into (B.10)

$$dX = -F_x^{-1}\left[F_z dZ - F_n B^{-1}\left(C^i + AdZ\right)\right] \tag{B.14}$$

$$dX = -F_x^{-1}\left[-F_n B^{-1} C^i + \left(F_z - F_n B^{-1} A\right) dZ\right] \tag{B.15}$$

Note the similarities between (B.15) and (B.7). Both expression would be identical if:

$$C^* = -F_n B^{-1} C^i \tag{B.16}$$

$$F_z^* = F_z - F_n B^{-1} A \tag{B.17}$$

In fact, Figure B.1 also includes the responses to a 1% increase in productivity when the interest rate is expected to be at 0 for 20 periods based on the "anticipated shocks" method. This method deliver identical responses to the previous methods.

While the anticipated shocks method involves more operations than the "direct" method presented in the previous section, we will shock in the next section that the anticipated shocks method facilitates the computation of the likelihood.

[20] For example, if the monetary authority announces shocks for the next 4 periods, $a = 4$.

[21] In our case, this Jacobian is zero for all endogenous equations except for the rows associated with the Taylor rule.

[22] As before, we assume that the ELB for the interest rate is equal to 0. If the ELB is different than zero, the first m entries of C^i will be $-i^{ss} + \underline{i}$, where $\underline{i}$ is the lower bound in the interest rate.

Note that recovering the time-varying and recursive formulation of the policy functions is time consuming to evaluate the likelihood function because it implies multiple large matrix operations for each set of parameters. However, given a set of parameters, recovering this approach is very efficient to compute stochastic simulations with ELB periods, because the matrix operations have to the executed only once. In contrast, for stochastic simulations, the anticipated shocks method can be time consuming.

B.4 Our algorithm. Steps to compute likelihood when the ELB is binding

The challenging part of computing the likelihood consists of combining the fact that, during ELB period, there is sequence of past shocks and there is an expected duration of the ELB. In the previous two sections, where we presented the "direct" and "anticipated shocks" methods, we implicitly assumed that we were departing from the steady state of the economy. In other words, there was not a sequence of past shocks. In this section, we borrow the intuition from the "anticipated shocks" method to solve this problem. In particular, we can now make the anticipated sequence of shocks not only a function of the expected ELB duration and current shocks but also a function of the sequence of *past* aggregate shocks. To achieve this result, we: (1) get the MA representation of the economy based on (B.10). This give us the response of the endogenous variables to *past* shocks. (2) Based on the MA representation of the economy we create a vector describing the expected path of the interest rate. (3) We solve for the current set of anticipated shocks such that the interest rate is expected to be at zero for m periods.

Note, again, that the difference with respect the IRFs presented in the previous two sections is that we are also conditioning on the *past* realizations of the shocks.

Computing the likelihood

1. Solving the model using SSJ, include L anticipated monetary policy shocks, where L should be as large as the maximum expected ELB duration. This results in an MA representation of the endogenous variables (dX) of the form:

$$dX = MA\epsilon^{t-1} + \tilde{\alpha}_0\tilde{\epsilon}_t + \beta_0\mu_t \tag{B.18}$$

where:

$$\epsilon_t = [\tilde{\epsilon}_t \quad \mu_t]' \tag{B.19}$$

$$\epsilon^t = [\epsilon_t \quad \epsilon_{t-1} \quad ... \quad \epsilon_{t-T}]' \tag{B.20}$$

$$\mu_t = [\varepsilon_{0,t} \quad \varepsilon_{1,t} \quad ... \quad \varepsilon_{L,t}]' \tag{B.21}$$

$$MA_j = \left[\alpha_j \quad \alpha_{j+1} \quad ... \quad \alpha_T \quad \underbrace{0 \;\; 0 \;\; \cdots \;\; 0}_{\text{j times}} \right] \quad \forall j = 0 \cdots T \tag{B.22}$$

$$\alpha_j = [\tilde{\alpha}_j \quad \beta_j]' \quad \forall j = 0 \cdots 1...T \tag{B.23}$$

μ_t is the vector of unanticipated and anticipated monetary policy shocks at time t, $\tilde{\epsilon}_t$ is the vector of all other shocks in the economy at time t, ϵ_t is the combination of the those two (i.e. all aggregate shocks at time t), and ϵ^t is the history of ϵ_t until time t. β_j is the response of the endogenous variables at time t to anticipated and unanticipated monetary policy shocks at time $t-j$. Similarly, $\tilde{\alpha}_j$ is the response of endogenous variables at time t to all other aggregate shocks at time $t-j$. MA_j is the endogenous variables response to all shocks up to period t.[23]

2. Extract the MA representation for the Federal Funds rate distinguishing between monetary policy shocks (anticipated and non-anticipated) and all other shocks in the economy. Hence, the federal funds rate at time t is given by:

$$\hat{i}_t = MA_1^i \epsilon^{t-1} + \tilde{\alpha}_0^i \tilde{\epsilon}_t + \beta_0^i \mu_t \tag{B.24}$$

where:

$$MA_j^i = \left[\alpha_j^i \quad \alpha_{j+1}^i \quad ... \quad \alpha_T^i \quad \underbrace{0 \;\; 0 \;\; \cdots \;\; 0}_{\text{j times}} \right] \quad \forall j = 0 \cdots T \tag{B.25}$$

$$\alpha_j^i = \left[\tilde{\alpha}_j^i \quad \beta_j^i\right]' \quad \forall j = 0 \cdots 1...T \tag{B.26}$$

β_j^i is the response of the interest (i) at time t to anticipated and unanticipated monetary policy shocks at time $t-j$. Similarly, $\tilde{\alpha}_j^i$ is the response of the interest (i) at time t to all other aggregate shocks at time $t-j$. MA_j^i is the interest rate response to all shocks up to period t.

[23]Because the solution is truncated at horizon T in SSJ, the last j elements of MA_j are equal to zero.

3. Get the implied vector of monetary policy shocks for time t, as a function of past monetary policy shocks and as a function of the other aggregate shocks (current and past). At time t, the expected interest rate for $t+1$ is:

$$E_t\left[\hat{i}_{t+1}\right] = \begin{bmatrix}\alpha_2^i & \alpha_3^i & \dots & \alpha_T^i & 0\end{bmatrix} \epsilon^{t-2} + \tilde{\alpha}_1^i \tilde{\epsilon}_t + \beta_1^i \mu_t \tag{B.27}$$

We can then group, the expected interest rate between t and $t+L$ as:

$$\begin{bmatrix} \hat{i}_t \\ E_t\left[\hat{i}_{t+1}\right] \\ \vdots \\ E_t\left[\hat{i}_{t+L}\right] \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_1^i & \alpha_2^i & \alpha_3^i & \cdots & \alpha_{T-2}^i & \alpha_{T-1}^i & \alpha_T^i \\ \alpha_2^i & \alpha_3^i & \alpha_4^i & \cdots & \alpha_{T-1}^i & \alpha_T^i & 0 \\ \alpha_3^i & \alpha_4^i & \alpha_5^i & \cdots & \alpha_T^i & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \alpha_{L+1}^i & \alpha_{L+2}^i & \alpha_{L+3}^i & \cdots & 0 & 0 & 0 \end{bmatrix}}_{\Omega_t} \epsilon^{t-1} + \underbrace{\begin{bmatrix} \tilde{\alpha}_0^i \\ \tilde{\alpha}_1^i \\ \tilde{\alpha}_2^i \\ \vdots \\ \tilde{\alpha}_L^i \end{bmatrix}}_{\omega_t} \tilde{\epsilon}_t + \underbrace{\begin{bmatrix} \beta_0^i \\ \beta_1^i \\ \beta_2^i \\ \vdots \\ \beta_L^i \end{bmatrix}}_{\lambda_t} \mu_t \tag{B.28}$$

Now, if the ELB is expected to bind for L periods:

$$\begin{bmatrix} \hat{i}_t \\ E_t\left[\hat{i}_{t+1}\right] \\ \vdots \\ E_t\left[\hat{i}_{t+L}\right] \end{bmatrix} = \underbrace{\begin{bmatrix} -i^{ss} \\ -i^{ss} \\ \vdots \\ -i^{ss} \end{bmatrix}}_{\zeta_t} \tag{B.29}$$

Hence, we can get the implied monetary policy shocks for time t as:

$$\mu_t = \lambda_t^{-1}\left[\zeta_t - \Omega_t \epsilon^{t-1} - \omega_t \tilde{\epsilon}_t\right] \tag{B.30}$$

Matrices, Z, ω, Ω, and λ are denoted with a subscript t as they are a funciton of the expected ELB duration at time t. We can re-group the elements in equation (B.30) such that:

$$\mu_t = \lambda_t^{-1}\zeta_t + \phi_t \tilde{\epsilon}^t + \psi_t \mu^{t-1} \tag{B.31}$$

where the first columns in ϕ_t correspond to the column in $-\lambda_t^{-1}\Omega_t$ associated with $\tilde{\epsilon}^{t-1}$, and the last columns correspond to $-\lambda_t^{-1}\omega_t$, and the columns of $-\lambda_t^{-1}\Omega_t$ associated with μ^{t-1} correspond to ψ_t.

4. Get the implied sequence of monetary policy shocks as a function of current and past aggregate shocks other than anticipated monetary policy shocks. Assuming that the ELB binds between periods t and $t+E$, stack

$$\begin{bmatrix}\mu_t\\ \mu_{t+1}\\ \vdots\\ \mu_{t+E}\end{bmatrix} = \underbrace{\begin{bmatrix}\lambda_t^{-1}\zeta_t\\ \lambda_{t+1}^{-1}\zeta_{t+1}\\ \vdots\\ \lambda_{t+E}^{-1}\zeta_{t+E}\end{bmatrix}}_{\mathcal{C}} + \underbrace{\begin{bmatrix}0 & 0 & \cdots & 0 & \phi_t^0 & \phi_t^1 & \cdots & \phi_t^{T-1} & \phi_t^T\\ 0 & 0 & \cdots & \phi_{t+1}^0 & \phi_{t+1}^1 & \phi_{t+1}^2 & \cdots & \phi_{t+1}^T & 0\\ \vdots & & & & & & & & \\ \phi_{t+E}^0 & \phi_{t+E}^1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0\end{bmatrix}}_{\Phi} \begin{bmatrix}\tilde{\epsilon}_{t+E}\\ \tilde{\epsilon}_{t+E-1}\\ \vdots\\ \tilde{\epsilon}_{t+1}\\ \tilde{\epsilon}_t\\ \tilde{\epsilon}_{t-1}\\ \vdots\\ \tilde{\epsilon}_{t-T}\end{bmatrix}$$

$$+ \underbrace{\begin{bmatrix}0 & 0 & \cdots & 0 & 0 & 0\\ \psi_{t+1}^1 & 0 & \cdots & 0 & 0 & 0\\ \psi_{t+2}^2 & \psi_{t+2}^1 & \cdots & 0 & 0 & 0\\ \vdots & & & & & \\ 0 & 0 & \cdots & \psi_{E+1}^2 & \psi_{E+1}^1 & 0\end{bmatrix}}_{\Psi} \begin{bmatrix}\mu_t\\ \mu_{t+1}\\ \vdots\\ \mu_{t+E-1}\\ \mu_{t+E}\end{bmatrix} \tag{B.32}$$

where

$$\phi_t = \begin{bmatrix}\phi_t^0 & \phi_t^1 & \phi_t^2 & \cdots & \phi_t^T\end{bmatrix} \tag{B.33}$$

$$\psi_t = \begin{bmatrix}\psi_t^1 & \psi_t^2 & \cdots & \psi_t^T\end{bmatrix} \tag{B.34}$$

Hence, the implied sequence of monetary policy shocks is given by:

$$\mu^{t+E} = (I-\Psi)^{-1}\left[\mathcal{C} + \Phi\tilde{\epsilon}^{t+E}\right] \tag{B.35}$$

which implies that for each period t:

$$\mu_t = \mathcal{A}_t^\mu + \mathcal{B}_t^\mu \tilde{\epsilon}^t \tag{B.36}$$

Note that in our model $\mu_t = 0$ for all time periods before the first ELB episode. But this analysis is easy to extend to a model with active anticipated monetary policy shocks in non-ELB periods.

5. Plug the vector of monetary policy shocks (B.36) into the (B.18) to get the time t MA representation for the endogenous variables of the form:

$$dX_t = \mathcal{A}_t + \mathcal{B}_t\tilde{\epsilon}^t \tag{B.37}$$

6. Using (B.37) compute the likelihood of as in Auclert et al. (2021) (section 5.3). In particular, given the vector dX_t^{obs} of n_{obs} observables and a sample size of T_{obs}:

$$dX_t^{obs} = B \cdot dX_t + u_t \tag{B.38}$$

where u_t is a vector of measurement errors, and B is a selector matrix. We then stack the covariances of these observables in a large symmetric matrix V of size $n_{obs} \cdot T_{obs}$. Then, the likelihood function is given by:

$$\mathcal{L}\left(d\mathbf{X}^{obs} \mid \Theta\right) = (2\pi)^{-\frac{T_{obs}}{2}} \mid V \mid^{-\frac{1}{2}} exp\left\{\left[d\mathbf{X}^{obs} - \mathbf{A}\right]' V^{-1}\left[d\mathbf{X}^{obs} - \mathbf{A}\right]\right\} \tag{B.39}$$

where $d\mathbf{X}^{obs}$ is the stacked vector of observables, $\mathbf{A}$ is the staced vector of constants $(B \cdot \mathcal{A}_t)$, and Θ is the vector of model parameters.[24]

[24] Auclert et al. (2021) discuss how to quickly evaluate the determinant of V as well as the quadratic form.

C Proof of proposition ??

In steady state, for a given real input price p^x and aggregate demand Y, if $\Lambda(\tilde{p},a)$ and $f(p\,|\,a)$ solve the firm's problem for pricing parameters θ^p, θ^a, and κ, then $\Lambda(\tilde{p},a)$ and $f(p\,|\,a)$ solve the firm's problem for real input price p^x, aggregate demand $\tilde{Y}$, and for any set of pricing parameters $\tilde{\theta}^p$, $\tilde{\theta}^a$, and $\tilde{\kappa}$ such that:

$$\frac{\tilde{\theta}^p}{\tilde{Y}} = \frac{\theta^p}{Y}$$
$$\frac{\tilde{\theta}^a}{\tilde{Y}} = \frac{\theta^a}{Y}$$
$$\frac{\tilde{\kappa}}{\tilde{Y}} = \frac{\kappa}{Y}$$

Proof. First, we proof that the value of a firm is homogeneous of degree 1 in the aggregate output as long as the pricing parameters (θ^p, θ^a, and κ) are constant, relative to aggregate output. Then, using this result, we proof that the optimal choice distributions are homogeneous of degree 0 in the aggregate output as long as the pricing parameters are constant, relative to aggregate output.

Value of a firm Denoting $\bar{x} = \frac{x}{Y}$ for $x = \theta^p, \theta^a, \kappa$, we can denote the value of a firm in steady state as:

$$V(p_j, a_j; Y) = \pi(p_j, a_j; Y) + E\Big\{ M\, V^*\left(\tilde{p}_j, a_j; Y\right) \Big\} \tag{C.1}$$

where

$$\begin{aligned}
\tilde{p} &= \frac{p}{\pi} \\
\pi(p_j, a_j; Y) &= Y p_j^{-\varepsilon}\left(p_j - p^x\right) \\
V^*(p, a; Y) &= \max_{\Lambda} \Big\{ \Lambda(a) \cdot \left[V^a(a;Y) - \bar{\kappa} Y\right] + (1 - \Lambda(a)) \cdot V(p, a; Y) - \bar{\theta}^a Y\, \mathcal{D}\left(\Lambda(a) \parallel \bar{\Lambda}\right) \Big\}, \\
V^a(a; Y) &= \max_{f} \left\{ \int f(p\,|\,a) V(p, a; Y)\, dp - \bar{\theta}^p Y \mathcal{D}\left(f(p\,|\,a) \parallel \bar{f}(p)\right) \right\} \\
& \quad \text{s.t.} \int f(p\,|\,a)\, dp = 1.
\end{aligned}$$

Substituting, and denoting Λ^* and $f^*(p)$ the optimal distributions:

$$\begin{aligned}
V(p_j, a_j; Y) &= \pi(p_j, a_j; Y) \\
&\quad + E\Big\{ M \Lambda^* \cdot \left[\int f(p\,|\,a) V(p, a; Y)\, dp - \bar{\theta}^p Y \mathcal{D}\left(f(p\,|\,a) \parallel \bar{f}(p)\right) - \bar{\kappa} Y \right] \Big\} \\
&\quad + E\Big\{ M\left(1 - \Lambda^*\right) \cdot V(\tilde{p}, a; Y) - M \bar{\theta}^a Y\, \mathcal{D}\left(\Lambda^* \parallel \bar{\Lambda}\right) \Big\}
\end{aligned} \tag{C.2}$$

Hence, note that the value of the firm is homogenous of degree 1 in aggregate output

$$
\begin{aligned}
V(p_j, a_j; Y) &= Y \cdot V(p_j, a_j; 1) \\
V(p_j, a_j; 1) &= \pi(p_j, a_j; 1) \\
&\quad + E\Big\{M\Lambda^* \cdot \Big[\int f(p \,|\, a) V(p, a; 1)\, dp - \bar{\theta}^p \mathcal{D}\left(f(p \,|\, a) \parallel \bar{f}(p)\right) - \bar{\kappa}\Big]\Big\} \\
&\quad + E\Big\{M\left(1 - \Lambda^*\right) \cdot V(\tilde{p}, a; 1) - M\bar{\theta}^a \mathcal{D}\left(\Lambda^* \parallel \bar{\Lambda}\right)\Big\}
\end{aligned}
$$

Substituting this result in $V^a(a;Y)$, we can also verify that $V^a(a;Y) = Y \cdot V^a(a;1)$

Optimal choice distribution The review distribution in steady state is given by:

$$
\ln\left(\frac{\Lambda(\tilde{p}, a)}{1 - \Lambda(\tilde{p}, a)}\right) = \ln\left(\frac{\bar{\Lambda}}{1 - \bar{\Lambda}}\right) + \frac{1}{\theta^a}\left[V^{a;Y}(a) - V(\tilde{p}, a; Y) - \kappa\right]
$$

Given that the value of the firm is homogenous of degree 1 in aggregate output, and substituting $\bar{x} = \frac{x}{Y}$ for $x = \theta^p, \theta^a, \kappa$

$$
\begin{aligned}
\ln\left(\frac{\Lambda(\tilde{p}, a)}{1 - \Lambda(\tilde{p}, a)}\right) &= \ln\left(\frac{\bar{\Lambda}}{1 - \bar{\Lambda}}\right) + \frac{1}{\bar{\theta}^a Y}\left[V^a(a;1)Y - V(\tilde{p}, a; 1)Y - \bar{\kappa}Y\right] \\
\ln\left(\frac{\Lambda(\tilde{p}, a)}{1 - \Lambda(\tilde{p}, a)}\right) &= \ln\left(\frac{\bar{\Lambda}}{1 - \bar{\Lambda}}\right) + \frac{1}{\bar{\theta}^a}\left[V^a(a;1) - V(\tilde{p}, a; 1) - \bar{\kappa}\right]
\end{aligned}
$$

Namely, $\Lambda(\tilde{p}, a)$ is homogenous of degree 0 in aggregate output as long as pricing parameters are constant relative to aggregate output. Regarding the pricing distribution, in steady state:

$$
\begin{aligned}
f(p \,|\, a) &= \frac{\bar{f}(p)\, \exp\left\{\frac{V(p,a;Y)}{\theta^p}\right\}}{\int \bar{f}(\hat{p})\, \exp\left\{\frac{V(\hat{p},a;Y)}{\theta^p}\right\} d\hat{p}} \\
f(p \,|\, a) &= \frac{\bar{f}(p)\, \exp\left\{\frac{V(p,a;1)Y}{\bar{\theta}^p Y}\right\}}{\int \bar{f}(\hat{p})\, \exp\left\{\frac{V(\hat{p},a;1)Y}{\bar{\theta}^p Y}\right\} d\hat{p}} \\
f(p \,|\, a) &= \frac{\bar{f}(p)\, \exp\left\{\frac{V(p,a;1)}{\bar{\theta}^p}\right\}}{\int \bar{f}(\hat{p})\, \exp\left\{\frac{V(\hat{p},a;1)}{\bar{\theta}^p}\right\} d\hat{p}}
\end{aligned}
$$

Namely, the pricing disribution is also homogenous of degree 0 in aggregate output as long as pricing parameters are constant relative to aggregate output. □

D Taylor Rule Estimation

To estimate the parameters governing the Taylor rule, we run a Bayesian estimation on the federal funds rate between 1984Q1 and 2007Q2. The Taylor rule is given by:

$$\begin{aligned} i_t &= \rho_i i_{t-1} + (1-\rho_i)\left[i_{ss} + \phi_\pi\left(\pi^a_t - \bar{\pi}^*\right) + \phi_y(dy^a_t - dy_{ss})\right] + \epsilon^i_t + e^\pi_t \\ e^\pi_t &= \rho_\pi e^\pi_{t-1} + \epsilon^\pi_t \\ \epsilon^i_t &\sim \mathcal{N}(0, \sigma_i) \\ \epsilon^\pi_t &\sim \mathcal{N}(0, \sigma_\pi) \end{aligned}$$

where i_t is the federal funds rate at quarter t, π^a_t is the annual CPI inflation rate at quarter t, and dy^a_t is the annual quarterly GDP growth. ϵ^i is an i.i.d. monetary policy shock, and e^π_t is a persistent shock to the inflation target (π^*). We set i_{ss}, dy_{ss}, and $\bar{\pi}$ to the sample averages of the Federal Funds rate, annual quarterly GDP growth, and annual CPI inflation.

Hence, for a set of parameters $\Theta = \{\rho_i, \phi_\pi, \phi_y, \rho^\pi, \sigma^i, \sigma^\pi\}$, we implement the Kalman filter on the Taylor rule and evaluate the Likelihood. We impose upper and lower bounds on our parameters, and we impose a normal prior distribution on all these parameters, as shown in Table D.1.

TABLE D.1: TAYLOR RULE PARAMTERS: PRIORS AND POSTERIOR

		Prior		Bounds	
Parameter	Mode	Mean	Std	Lower	Upper
ϕ_π	1.6977	2.0000	0.2500	1.00	3.00
ϕ_y	1.004	0.5000	0.2500	0.00	3.00
ρ_i	0.7914	0.5000	0.0500	0.00	0.95
ρ_π	0.2288	0.1200	0.0500	0.03	0.99
σ_i	0.3500	0.0625	0.0625	0.00	1.00
σ_π	0.0934	0.0625	0.0625	0.00	1.00

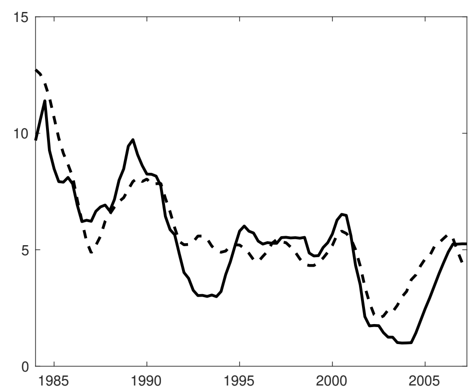
Notes: The table reports the posterior mode of the estimated Taylor rule parameter along with their prior mean, prior standard deviation, and lower and upper bounds. The prior distribution on all parameters is normal.

We restrict the Taylor rule parameter on inflation (ϕ_π) to be between 1 and 3, and our prior is that π_π distributes normal with mean equal to 2 and standard deviation equal to 0.25. Similarly, we impose the Taylor rule parameter on GDP growth (ϕ_y) to be between 0 and 3, and our prior is that ϕ_y distributes normal with mean equal to 0.5 and standard deviation equal to 0.25. We constraint the interest smoothing parameters (ρ_i) to be between 0 and 0.95, and we set the mean and stand deviations of its prior to 0.5 and 0.05. Figure D.1 plots the Federal Funds Rate (solid line) along with the predicted rate with the parameters evaluated at their mode (dash line). In particular, the predicted rate equals:

$$i_t^p = \rho_i i_{t-1}^p + (1 - \rho_i) \left[i_{ss} + \phi_\pi \left(\pi_t^a - \bar{\pi}^* \right) + \phi_y (dy_t^a - dy_{ss}) \right]$$

with $i_0^p = i_0$. In general, the predicted rate tracks well the movements in the actual Federal Funds Rate.

Figure D.1: Federal Funds Rate and Fit



Note: This Figure plots the Federal Funds rate (solid line) along with the predict federal funds rate when the parameters are evaluated at their posterior mode (second column of Table D.1). The predicted rate equals $i_t = \rho_i i_{t-1}^p + (1 - \rho_i) \left[i_{ss} + \phi_\pi \left(\pi_t^a - \bar{\pi}^* \right) + \phi_y (dy_t^a - dy_{ss}) \right]$ with $i_0^p = i_0$.

E Estimation details

E.1 Overall Prior Distribution

In our Bayesian estimation, in addition to the standard informative priors over the model parameters, we specified an implicit prior over selected business cycle moments generated by our model, following Del Negro & Schorfheide (2008). Denote μ^{obs} as the vector of selected business cycle moments from the data, and $\mu(\Theta)$ as the model generated moments given the set of parameters Θ. $\mu(\Theta)$ relates to the data business cycle moments as follows:

$$\mu^{obs} = \mu(\theta) + \eta^{obs} \tag{E.1}$$

where η^{obs} is a vector of measurement errors that distributes normal with matrix of variance covariance equal to $\Sigma^{\eta^{obs}}$. We express (E.1) in terms of a conditional density (likelihood function) and use Bayes theorem in combination with a marginal density (denoted by $\pi(\Theta)$) to generate a conditional distribution (our overall prior distribution) that reflects our beliefs about the selected business cycle moments:

$$p\left(\Theta \mid \mu^{obs}\right) = \mathcal{L}\left(\mu(\Theta) \mid \mu^{obs}\right) \cdot \pi(\Theta) \tag{E.2}$$

where $\pi(\Theta)$ refers to the informative priors listed in Table ??. Hence, our posterior distribution is given by:

$$p(\Theta \mid d\mathbf{X}^{obs}, \mu^{obs}) \propto \quad \mathcal{L}(d\mathbf{X}^{obs} \mid \Theta)p\left(\Theta \mid \mu^{obs}\right) \tag{E.3}$$

$$\propto \quad \mathcal{L}(d\mathbf{X}^{obs} \mid \Theta)\mathcal{L}\left(\mu(\Theta) \mid \mu^{obs}\right) \cdot \pi(\Theta) \tag{E.4}$$

where $\mathcal{L}(d\mathbf{X}^{obs} \mid \Theta)$ is the likelihood function of the data defined in (B.39). The selected business cycle moments are: the variance of the observables, the covariance between the federal funds rate and GDP growth and CPI inflation, and the covariance between GDP growth and inflation. We assume that $\left(\Sigma^{\eta^{obs}}\right)^{-\frac{1}{2}}$ is a diagonal matrix with entries equal to ω times the standard deviation of the data moment, where ω is a scalar that we set to 0.25. Hence, for each moment m, we assume that the standard deviation of the measurement error associated with that moment equals σ^m:

$$\sigma^m = \omega \left[\sum_{t=1}^{T^{obs}} \frac{\left(x_t^m - \bar{x}^m\right)^2}{T^{obs}}\right]^{\frac{1}{2}} \tag{E.5}$$

where x_t^m refers to the data moment at time t, and $\bar{x}$ is the sample average of x. For example, $x_t^m = (dy_t^q - \bar{dy}^q)^2$ for the variance of GDP growth, and $x_t^m = (dy_t^q - \bar{dy}^q)(\pi_t^q - \bar{\pi}^q)$ for the covariance between GDP growth and quarterly inflation.

To compute the model generated moments $\mu(\Theta)$, we make use of the time-varying MA representation presented in Appendix B (equation (B.37)). Given the time-varying MA representation of the observables:

$$dX_t^{obs} = \mathcal{A}_t^{obs} + \mathcal{B}_t^{obs}\epsilon^t \tag{E.6}$$

we stack them to get:

$$d\mathbf{X}^{obs} = \mathbf{A} + \mathbf{B}\epsilon \tag{E.7}$$

Given that the steady state of our model matches the sample average of our observables, we are interested in computing:

$$E\left\{\left[d\mathbf{X}^{obs}\right]'\left[d\mathbf{X}^{obs}\right]\right\} = \Sigma(\Theta) \tag{E.8}$$

$$= \mathbf{A}'\mathbf{A} + \mathbf{B}'\Sigma^{\epsilon}\mathbf{B} \tag{E.9}$$

Where Σ^{ϵ} equals to the matrix of variance covariance of the model shocks.[25] Note that $\Sigma(\Theta)$ is a big square and symmetric matrix of size $n^{obs} \cdot T^{obs}$.

Define matrix $\sigma(\Theta)_t$ as the square and symmetric matrix of size $nobs$ form by the rows and columns of $\Sigma(\Theta)$ starting in row and column $(t-1)nobs+1$ and ending in row and column $t \cdot nobs$. Matrix $\sigma(\Theta)_t$ is the expected sample covariance of the observables in period t or, in other words, the expected quadratic deviation of the observables from their sample average in period t. Then, the expected matrix of variance covariance generated by the model is given by:

$$\bar{\sigma}(\Theta) = \sum_{t=1}^{T^{obs}} \frac{\sigma(\Theta)_t}{T^{obs}} \tag{E.10}$$

Hence, our selected model generated moments used in our prior ($\mu(\Theta)$) correspond to the associated elements of matrix $\bar{\sigma}(\Theta)$.

Note that our model generated moments condition on the ELB episodes. In other words, our model generated moments are not unconditional, the moments are conditional on the ELB periods and expected duration for those.

Other parameter restrictions: In addition to these priors, to avoid exploring unreasonable areas of the parameter space, we get the filtered series for each draw of parameters and discard those draws that imply a value of the menu cost below zero or log-deviations for θ_p and θ_a greater than three in absolute value.

E.2 Posterior Sampler

We use a standard Metropolis Hasting Algorithm with similar specification as in Kulish et al. (2017). The algorithm is the following:

[25] Note that the last term in (E.9) equals to matrix V of Appendix B.

Algorithm: At the beginning of each iteration j, and given an initial set of parameters Θ_{j-1}

1. Randomly select how many parameters to update from a uniform distribution between $\lceil 0.2n_\Theta \rceil$ and n_Θ, where n_Θ is the number of parameters in Θ.

2. Randomly select which parameters to update.

3. Construct a proposed set of parameters Θ_j^p. To do this, we use a multivariate Student t distribution with 12+p degrees of freedom centered at Θ_{j-1} and with a matrix of variance covariance given by Σ^{Θ^p}, which will be specified below. p is te number of parameters being updated.

4. Compute the acceptance ratio (AR):

$$AR_j^\Theta = \frac{p(\Theta_j^p \mid d\mathbf{X}^{obs}, \mu^{obs})}{p(\Theta_{j-1} \mid d\mathbf{X}^{obs}, \mu^{obs})} \tag{E.11}$$

   set $AR_j = 0$ if the proposal includes inadmissible values.

5. Accept the proposal with probability $min\{AR_j, 1\}$ and set $\Theta_j = \Theta_j^p$. Otherwise, set $\Theta_j = \Theta_{j-1}$.

Matrix of variance covariance Σ^Θ: To compute the matrix of variance covariance Σ^Θ, we run a chain of 150,000 draws using a matrix of variance covariance equal to the diagonal matrix of the Hessian at the optimization mode scaled by κ^{mcmc}, which we set to 0.0575. Then, we drop the first 50,000 draws and compute the matrix of variance covariance resulting from this chain, which we denote by Σ^Θ. Then, for the proposal density, we can decomposte matrix Σ^Θ as follows:

$$\Sigma^\Theta = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{E.12}$$

where the first block refers to the fixed parameters at iteration j. Hence, the specific matrix of variance covariance used in iteration j equals to:

$$\Sigma^{\Theta^p} = \omega^{mcmc}\Sigma_{22|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \tag{E.13}$$

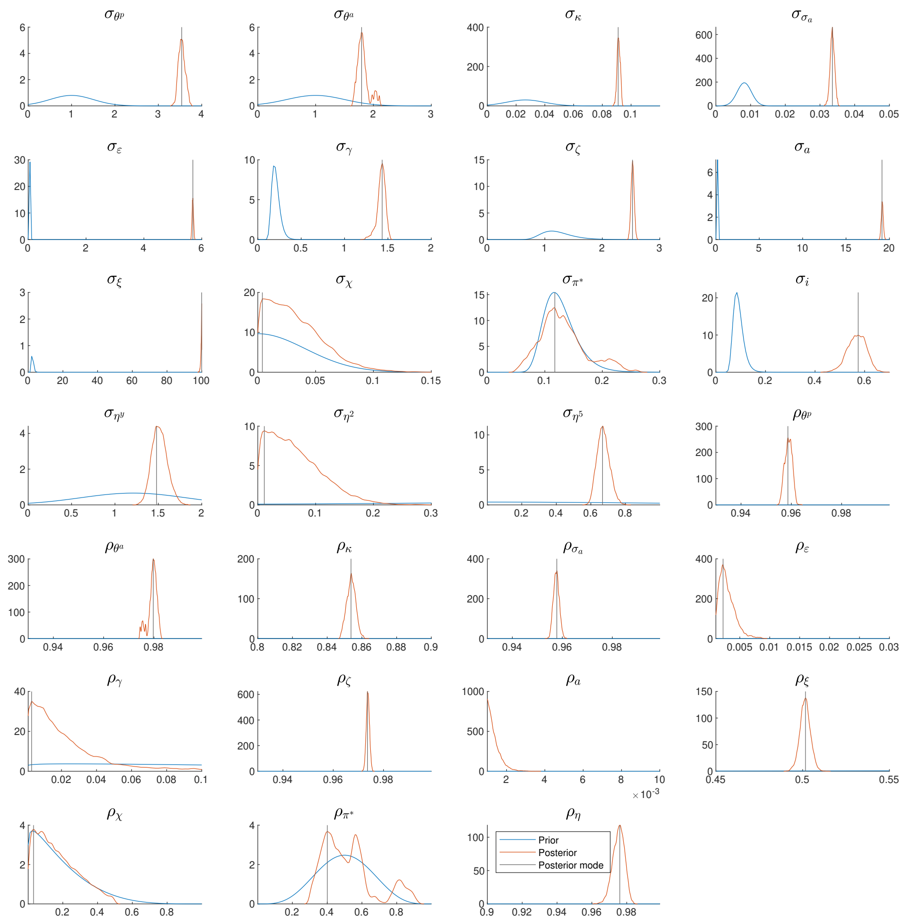
where ω^{mcmc} is a scaling parameter that we set to 0.2 to target an acceptance rate of 30%.

Sample chain: We used a single chain of 500,000 draws with an average acceptance rate of 30%. We drop the first 100,000 draws to compute statistic on the the posterior distribution.

E.3 Prior and Posterior distributions

Figure E.1 plots the prior and posterior distributions and Figure E.2 plots the trace plots. Solid black lines represent the posterior mode.

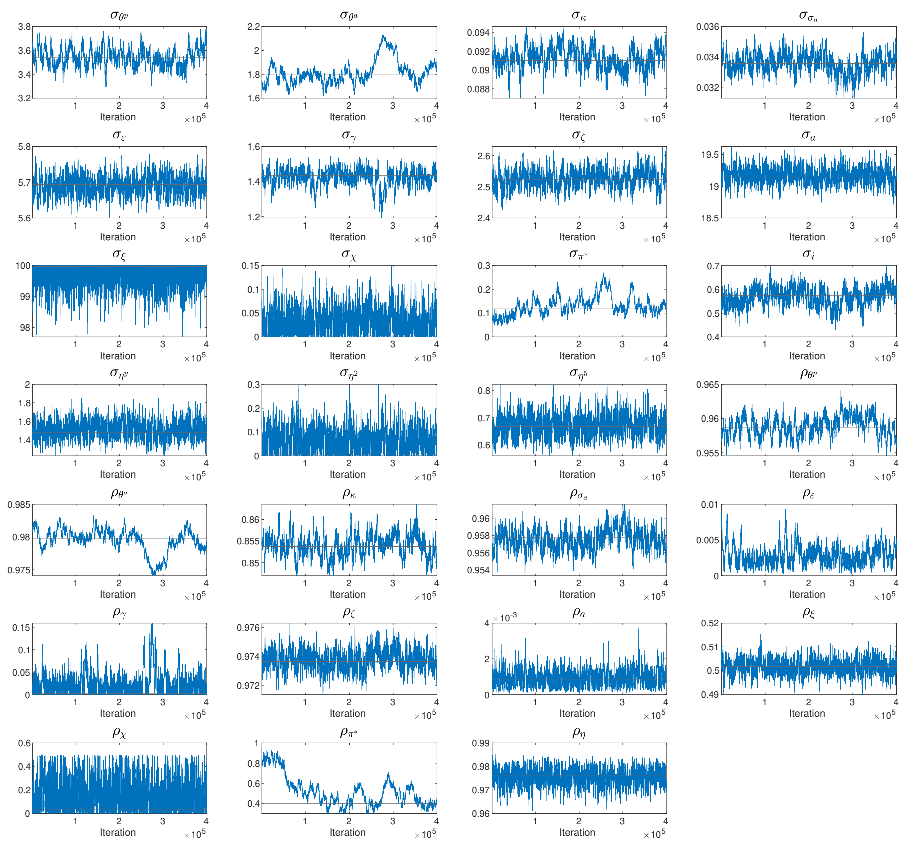
Figure E.1: Prior and posterior distribution



E.4 Elasticity of demand and volatility o idiosyncratic cost

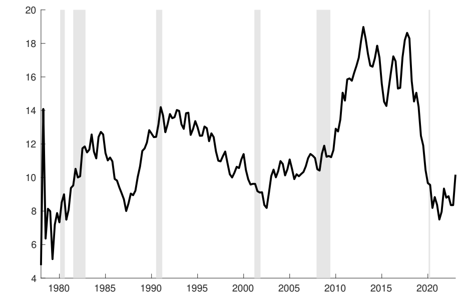
Figure E.3 plots the filtered series for the elasticity of demand (ε) and volatility of idiosyncratic shocks σ_a when the model is evaluated at the posterior mode.

Figure E.2: Trace plots

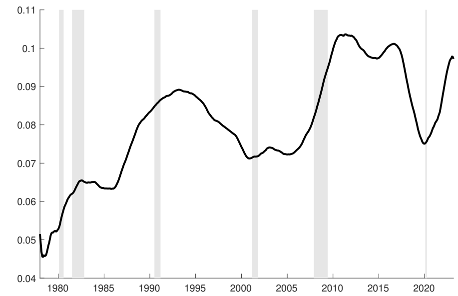


Note: Horizontal line represents posterior mode.

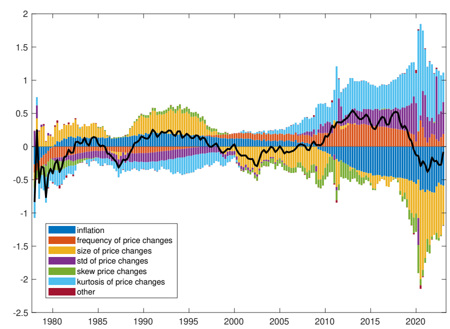
Figure E.3: Elasticity of Demand and Volatility of Idiosincratic Cost



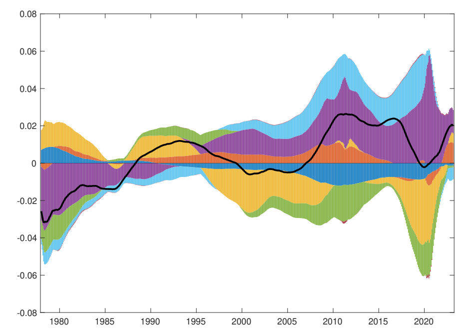
(a) ε: Filtered series



(b) σ_a: Filtered series



(c) ε: Data decomposition



(d) σ_a: Data decomposition

Top panel presents the estimated series for the elasticity of demand (ε) and volatility of idiosyncratic shocks σ_a. The bottom panel presents the data decomposition of the deviation of σ_a from its steady state value and the data decomposition for the log-deviations of ε for its steady state value.

F Detrended Model Equations

[MAKE SURE THAT NOTATION IS CONSISTENT WITH FINAL VERSION OF THE MODEL]

Aggregate price index:

$$P_t = 1 = \left[\int p_{jt}^{1-\varepsilon}\Omega_{jt}dj\right]^{\frac{1}{1-\varepsilon}} \tag{F.1}$$

Aggregate resource constraint:

$$\Delta_t \tilde{Y}_t = \tilde{C}_t + \tilde{G}_t + \tilde{F}_t \tag{F.2}$$

where

$$\tilde{F}_t = \int \left[\tilde{\kappa} + \tilde{\mathcal{C}}_{jt}^a + \lambda_{jt}\tilde{\mathcal{C}}_{jt}^p\right]\tilde{\Omega}_{jt}dj \tag{F.3}$$

Euler equation:

$$\zeta_t\,\chi_t + E_t\left[\frac{\beta}{\gamma_{t+1}}\,\zeta_{t+1}\,\tilde{\lambda}_{t+1}\left(\frac{i_t}{\pi_{t+1}}\right)\right] = \zeta_t\,\tilde{\lambda}_t \tag{F.4}$$

where

$$\tilde{\lambda}_t = z_t\lambda_t = \left(\frac{1}{\tilde{C}_t - h\frac{\tilde{C}_{t-1}}{\gamma_t}}\right) - E_t\left[\left(\frac{\beta\zeta_{t+1}}{\zeta_t}\right)\left(\frac{h}{\gamma_{t+1}\tilde{C}_{t+1} - h\tilde{C}_t}\right)\right] \tag{F.5}$$

Marginal cost:

$$e^{a_t}\tilde{p}_t^x = \delta^w e^{a_t}\frac{\xi_t\left(\frac{\tilde{Y}_t}{e^{a_t}}\Delta_t\right)^{\frac{1}{\nu}}}{\tilde{\lambda}_t} + (1-\delta^w)\tilde{p}^{x*} \tag{F.6}$$

where we made use of:

$$p_t^x = \frac{w_t}{e^{a_t}} \tag{F.7}$$

$$w_t^* = \frac{\xi_t L_t^{\frac{1}{\nu}}}{\lambda_t} \tag{F.8}$$

$$L_t = \frac{\tilde{Y}_t}{e^{a_t}}\Delta_t \tag{F.9}$$

and where

$$\Delta_t = \int_j e^{-a_{jt}}p_{jt}^{-\varepsilon}\Omega_{jt}dj \tag{F.10}$$

Taylor rule:

$$i_t = i_t^{target} + \epsilon_{rt}, \tag{F.11}$$

$$(i_t^{target})^p = \rho_i (i_{t-1})^p + (1-\rho_i)\left[(i_{ss})^p + \phi_\pi \left(\Pi_{j=0}^{p-1}\pi_{t-j} - (\pi_t^*)^p\right) + \phi_y (dy_t - dy_{ss})\right], \tag{F.12}$$

G Alternative Reference Distributions

One novelty of the model we consider is the fact that firms use the equilibrium distribution of prices as their reference distribution. This feature introduces an interesting layer of strategic complementarities (SC): whereas the typical SC story is that each firm has incentives to keep its price close to that of its competitors, here each firm benefits from keeping the *distribution* from which it draws its prices close to the distribution of its competitors' prices. How does this fixed point relationship between individual policies and the aggregate distribution of prices affect pricing frictions?

Consider two alternatives that break the strategic complementarities induced by this fixed point. First, suppose that each firm is purely rationally inattentive such that it uses as its reference the distribution that minimizes the distance to the choice distribution that the firm uses over the life of its policy. Formally, this is given by

$$\bar{f} = ... \tag{G.1}$$

Similarly, the RI-optimal reference frequency of adjustment is given by

$$\bar{\Lambda} = ... \tag{G.2}$$

This formulation does not introduce any new parameters, but it places a different restriction on what the reference distribution can be. As a result, it not only implies weaker strategic complementarities across firms, but also a lower sensitivity of the cross-sectional distribution of prices to the value of θ^p. how to show this?

We find that this formulation does not fit the data as well: the reference distribution is too constrained in the RI model. Instead, it seems that firms mimic each other more than would be RI-optimal. This finding is not entirely surprising: Khaw et al. (2017) found that relaxing the RI formulation provided a better fit to how people forecasted a random variable in a laboratory setting.

A second alternative is that firms use as reference a uniform distribution, as in Costain & Nakov (2015, 2019). This assumption can be rationalized in a control-cost (CC) framework, in which firms know the optimal action to take but face an implementation friction: they need to exert effort to take the optimal action, orelse they end up with an action that is drawn from a uniform distribution around that optimum. Exerting effort is costly, with the cost proportional to the KL divergence from this uniform default.

One big difference between the RI-based and the CC frameworks is that in the latter the distribution of prices from which the firm draws is always centered on the optimal price. As a result, any repricing mistakes wash out on average, and any nonneutrality is once again determined only by the frequency and timing accuracy of price changes. Conversely, in the RI-based models, whether with mimicking firms or with purely RI firms, the optimal distribution f is skewed: As discussed earlier and shown in Figure (3), the distribution is centered on values slightly *above* the optimum at low marginal cost levels and becomes increasingly less sensitive to the optimal price as marginal cost increases. This asymmetry is the result of the firm economizing on its informationa acquisition costs. In the CC framework, since the firm already knows the optimal price, the skew in the resulting distribution is much weaker.