Consumer Information Friction and Monetary Non-Neutrality: A Sequential Search Approach *

Xiaojie Liu†

Sep 2025

Click here to view the latest version

Abstract

This paper develops a model of monetary non-neutrality driven solely by consumer-side information frictions, in contrast to standard approaches that rely on firm-side nominal rigidities. Consumers face search frictions when shopping for goods and have imperfect information about the average nominal marginal cost. The key mechanism is that, following a monetary shock, consumers attribute some of the resulting price changes to firm-specific adverse shocks, inducing them to search for alternatives. To dissuade search, firms compress the markup and limit the passthrough of the shock to prices. Firms still pass through *idiosyncratic shocks* flexibly, while responses to *aggregate shocks* are dampened due to weakened strategic complementarities by consumers' information friction. Building on this mechanism, I develop a tractable medium-scale DSGE model that can be efficiently solved using sequence-space methods. Estimating the model via impulse response matching, I find the degree of information friction to be substantial, accounting for 68% of observed monetary non-neutrality. I further consider aggregate productivity shocks and apply it to interpret the inflation surge during the pandemic. Finally, I provide two pieces of empirical evidence that support the model's core mechanisms.

Keywords: Monetary non-neutrality, Price Stickiness, Phillips Curve, Consumer Search, Information Frictions, Incomplete Passthrough, Strategic Complementarity, Higher-Order Beliefs, Sequence-Space Jacobian, Inflation Surge

---

*I am deeply grateful for guidance of my supervisors: George-Marios Angeletos, Lawrence Christiano, Benjamin Jones, Sara Moreira, Alireza Tahbaz-Salehi. I also appreciate comments and discussions with Marty Eichenbaum, Gaetano Gaballo, Diego Huerta, Diego Känzig, Jennifer La'O, Jean-Paul L'Huillier, Chen Lian, Ryungha Oh, Alessandro Pavan, Harry Pei, Karthik Sastry, Michael Woodford, and Huo Zhen. I also thank seminar participants in SED meeting 2025.

†Kellogg School of Management, Northwestern University. Email: xiaojie.liu@kellogg.northwestern.edu

"This paper presents a theory that justifies price stickiness, namely, that firms, fearing to upset their customers, attribute a cost to price changes."

— Rotemberg (1982)

Monetary policy is widely understood to have substantial real effects in the short run. A large body of empirical literature has shown that both output and inflation decline following an unexpected increase in the interest rate.[1] Still, a fundamental question persists: Why don't changes in nominal policy rates simply lead to proportional changes in nominal wages and prices, leaving real quantities unaffected? A classic answer, dating back to Friedman (1968) and Phelps (1967), is that policy shocks are *not anticipated*.[2] Lucas (1972) shows that this unanticipated nature of shocks can be interpreted as a form of *imperfect information* under fully rational expectations. Sims (2003, 2006) propose that the unawareness of policy shocks can stem from agents' limited capacity to *pay attention to* all the relevant information to their decisions.

Firms are the agents that set prices. If they are not well informed about policy shocks, they may be unable to immediately and fully adjust prices in response. For this reason, it is natural to model price stickiness by assuming that firms, rather than consumers, cannot fully anticipate policy shocks due to imperfect information (Lucas, 1972; Woodford, 2003; Mankiw and Reis, 2002; Maćkowiak and Wiederholt, 2009). However, this approach comes with a less pleasant implication: to generate sufficiently large price stickiness to match the impulse responses of inflation to a monetary shock, firms must have extremely limited information about both nominal costs and nominal aggregate demand for a long period of time (Woodford, 2003).

This assumption is increasingly difficult to justify. In principle, firms post wages, negotiate input prices and thus should have accurate information about nominal costs. Furthermore, advances in computing and data storage have significantly improved firms' ability to forecast demand using big data (Begenau et al., 2018; Farboodi and Veldkamp, 2021). Recent survey evidence also challenges the premise of extreme informational frictions. Firms appear to be relatively well-informed about macroeconomic conditions and are capable of updating their inflation forecasts promptly in response to new information, such as the release of the monthly CPI (Candia et al., 2023; Yotzov et al., 2024).

In contrast, consumers are inherently less informed about nominal production costs and aggregate demand, simply since they neither produce goods nor bargain over input prices, and

[1] Christiano et al. (1999) identify this effect using timing restrictions in VAR. Recently, high-frequency identification approach helps resolve the endogeneity bias in the VAR approach and confirms this finding (Gertler and Karadi, 2015; Bauer and Swanson, 2023). Hazell et al. (2022) estimate the slope of Phillips curve to be very flat using cross-state variation in price indices. Ramey (2016) provides a great summary of this literature.

[2] Friedman famously noted, "there is always a temporary trade-off between inflation and unemployment; there is no permanent trade-off. The temporary trade-off comes not from inflation per se, but from unanticipated inflation, which generally means, from a rising rate of inflation" (1968, p. 11).

have far fewer resources to acquire the information. Empirical studies confirm this: consumers misperceive inflation (Binetti et al., 2024; Candia et al., 2023), overweight salient prices in forming expectations (Kumar et al., 2015), and anchor their beliefs on recent shopping experiences (D'Acunto et al., 2021). Furthermore, the consumer search literature documents substantial price dispersion even for identical goods, consistent with frictions in price discovery (Stigler, 1961; Kaplan and Menzio, 2015).

The limitation of firm-centric information-based theory and the new empirical findings raise an important question: is information friction-based explanation of monetary non-neutrality becoming untenable as policy shocks are increasingly *anticipated* or quickly learned by firms? Or alternatively, can we shift the foundation of price stickiness to consumer-side frictions, given that consumers appear to have limited tools for gathering and processing information about their economic environment? A prominent early advocate of this perspective is Rotemberg (1982), who argues that sluggish adjustment of prices may arise as firms *fear* to upset consumers if they fully adjust prices. Following this logic, Rotemberg (2005, 2010, 2011) and Eyster et al. (2015) develop theoretical models where consumers experience *disutility* when they pay prices that exceed their perceived markup. Firms thus set prices to avoid upsetting their customers. As a result, consumer expectations become part of the firms' pricing decisions.

This paper follows this route. The model builds on standard sequential search models (Wolinsky, 1986; Anderson and Renault, 1999) where consumers decide where to buy goods and they search sequentially over firms to learn prices and their fit. Each search is associated with a cost. The key difference of my model is that consumers do not feel regret or angry from perceived price unfairness. Instead, they respond by walking away and continuing to search for better prices and a better fit. The main mechanism operates as follows. Following a positive monetary shock, the nominal marginal costs increase. Firms tend to increase prices. However, consumers with *imperfect information* about the costs attribute much of this price increase to firm-specific adverse shocks. Marginal consumers who are initially indifferent between purchasing and searching are now incentivized to seek outside options. To dissuade search, firms compress markups, thereby limiting the passthrough of the monetary shock.

In this framework, the role of imperfect information parallels the original insights of Phelps (1969) and Lucas (1972): consumers rationally confuse an increase in nominal costs produced in all firms with an increase in the relative price of the firm they happen to visit. The key ingredient to such confusion is that firms are heterogeneous in productivity, and consumers lack information about each firm's idiosyncratic cost. As a result, any discrepancy between expected price and actual price will be attributed to higher or lower productivity of the firm. For example, a price above expectations may be seen as evidence of low productivity rather than a common nominal shock—prompting search rather than purchase.

I begin by characterizing the full-information rational expectations equilibrium. Under full information about policy shocks, I show that monetary policy is neutral. While this result is often taken for granted, it is not immediate in this context, as search frictions can be interpreted as a form of information friction on individual prices. Nonetheless, when consumers know the shock, they understand that, in equilibrium, all prices fully adjust proportionally to policy shocks. Relative prices remain unchanged, so do search decisions. The key insight is that search frictions alone are insufficient to generate monetary non-neutrality. Some form of imperfect information about aggregate shocks is still necessary to produce real effects.

One of the contributions of this paper is to show that the equilibrium in a setting with consumer sequential search, imperfect information about policy shocks, and heterogeneous firms can be summarized by a single equation, which resembles equation (2.7) in Woodford (2003) in its form,

$$\hat{p} = \Gamma \hat{mc} + (1 - \Gamma) E^c \hat{p} \tag{1}$$

where Γ is aggregate own-cost passthrough, namely passthrough of nominal marginal cost changes $\hat{mc}$, holding constant the response of other firms, and $E^c \hat{p}$ represents consumers' average expectation of price index. In Woodford (2003), firms are not well informed to fully pass policy shocks to prices. However, the mechanism here is more nuanced. First, the model preserves firms' ability to respond flexibly to *idiosyncratic shocks*. The reason is that consumers correctly interpret firm-specific price changes as reflecting relative price movements. In such cases, it is optimal for firms to pass through a fraction Γ of the shock and let marginal consumers choose to leave or stay. However, in the case of *aggregate shocks*, imperfect information drives a wedge between expected price index and actual price index. If firms were to fully pass through the aggregate shock, they would induce excessive consumer switching to other firms, as consumers misinterpreted an price increase as firm-specific, which reflects Rotemberg's core insight. Indeed, this holds true for all firms: consumers are less likely to stay and purchase during search, regardless of firms' productivity.

To further clarify the intuition, consider an extreme case in which policy shocks are completely unanticipated by consumers. In this scenario, an aggregate shock is perceived by consumers as indistinguishable from an idiosyncratic firm-level shock. As a result, the optimal passthrough is Γ. This outcome stands in stark contrast to firm-side imperfect information models in which firms fail to adjust prices altogether when shocks are fully unanticipated. The key difference lies in the informational asymmetry: in this model, firms know the policy shock. However, they understand that the shock is completely unanticipated to consumers, so they cannot respond to changes in other firms' prices, via the channel of *strategic complementarity*, i.e., setting prices close to competitors'.

The second contribution of this paper is to develop a medium-scale DSGE model that places the consumer theory of price stickiness at its core, and to estimate the magnitude of information friction using impulse response matching (Christiano et al., 2005). To better align the model with empirical evidence, I incorporate two additional frictions: habit formation and nominal wage rigidity. Habit

formation helps replicate the empirically observed hump-shaped response of output, while wage stickiness dampens the response of marginal costs to policy shocks (Christiano et al., 2005; Smets and Wouters, 2007). Consumers learn about policy shocks gradually, forming expectations based on a history of private noisy signals. Learning renders the deviation of output from potential arbitrarily persistent as signal noise increases (Woodford, 2003). I solve the model and compute impulse responses using the sequence-space method developed by Auclert et al. (2021), and follow Auclert et al. (2020) to accommodate learning dynamics in a flexible way. With this framework, I am able to address two central questions: (1) Can a model centered on consumer-side frictions—augmented with habit formation and wage rigidity—match the empirical impulse responses to monetary shocks identified via SVAR? and (2) How much do consumer-side frictions contribute to monetary non-neutrality?

I calibrate the model parameters to match the moments from the literature and my own empirical finding. In particular, the aggregate own-cost passthrough is a (roughly) sufficient statistic that summarizes all the parameters related to the search friction.[3] I take its value from Gopinath et al. (2011), Amiti et al. (2019) and Gagliardone et al. (2025). The degree of information friction among shoppers is then estimated using impulse response matching. Specifically, I estimate an external IV-SVAR following Gertler and Karadi (2015), using standard macroeconomic variables alongside household inflation expectations from the Michigan Survey of Consumers. The results reveal that household inflation expectations underreact significantly to monetary shocks with estimated average duration of information frictions to be 14-18 quarters, suggesting substantial information frictions. The key determinant of information friction in estimation is the relative impulse responses of CPI and cumulative inflation expectation. To analyze the monetary transmission mechanism, I switch off the information friction for shoppers in the model. As expected, the change in the price index becomes nearly five times larger, and expectations track the actual price level one-for-one—contrary to the empirical evidence. As a result, the degree of monetary non-neutrality is substantially diminished: the cumulative output response is reduced by about 65% and the peak output response shrinks by more than a half.

Finally, I consider aggregate productivity shocks in this framework. In contrast to standard New Keynesian models, the inflation response in this framework features a sharp initial increase followed by a rapid decline. This arises because, although nominal wage responses are dampened by wage rigidity and strategic complementarities are attenuated due to imperfect information, our model still allows firms to pass through a fraction Γ of the productivity shock to prices *immediately*, as implied by equation (1). I further examine the pandemic episode of elevated inflation. Beaudry et al. (2024) documents that during this period, household inflation expectations were closely aligned with actual inflation, suggesting minimal information frictions on the consumer side. I

[3]There is a demand allocation term relies on deep parameters (Baqaee et al., 2024). However, this effect is quantitatively small.

compute the model under the assumption of full information, the inflation response almost doubles. This application highlights the importance of consumer-side information friction in determining the slope of Phillips curve.

In the last part of the paper, I present two pieces of empirical evidence that support the key model mechanisms. First, I estimate how attention to inflation affects the effect of economic slack on core inflation. I find that higher attention, measured as in Pfäuti (2024), significantly increases the passthrough of headline inflation fluctuations and effect of labor market tightness. Second, using large-scale consumer panel data, I measure search activity by calculating the share of spending that occurs outside a household's routine store for each product category. I find that one standard deviation (60 bp) increase in unanticipated inflation for food and beverage leads to a 12.2 bp increase in the non-routine share of spending, consistent with model prediction.

Related Literature – This paper contributes to several strands of literature. First, my study is motivated by the importance of consumers' reactions to price changes in price-setting models (Rotemberg 2005, 2010, 2011; Kleshchelski and Vincent, 2009; Nakamura and Steinsson, 2011; Eyster et al., 2015; Matějka, 2015; Rebelo et al., 2024). These models usually assume that consumers experience disutility from perceived price unfairness or volatility. In contrast, this paper develops a consumer-based theory of price stickiness rooted in the tradition of Phelps (1969), Lucas (1972), Woodford (2003), Maćkowiak and Wiederholt (2009), Angeletos and La'o (2010), and Angeletos and La'O (2013). The model demonstrates that firms may not fully adjust prices to monetary shocks if shocks are not anticipated by their consumers, even firms themselves are fully informed. Therefore, my model offers new avenues for rational micro-foundations of price stickiness, which is consistent with recent empirical evidence on firm and household expectations.

Several papers are closely related. Gaballo and Paciello (2021) show that paid prices remain sticky because consumers, unaware of the monopoly firm's marginal costs, reallocate toward a competitive market after the monopoly increases its price after an expansionary shock. In contrast, I consider a sequential search model where all firms have monopoly power. I find that it is markup compression, not demand reallocation, that drives sluggish price adjustment. Quantitatively, reallocation has minimal effect on inflation. In addition, Bénabou and Gertner (1993) and L'Huillier (2020) focus on the strategic role of individual prices consumers observe as revealing the information about aggregate shocks. I specifically shut down this signaling channel. More broadly, my model remains tractable, accommodates standard DSGE features (Christiano et al., 2005; Smets and Wouters, 2007), and is solved efficiently using sequence-space methods (Auclert et al., 2021).

This paper takes a first step toward bridging recent empirical findings on consumer-side frictions with a structural model of price stickiness and monetary non-neutrality. Evidence shows households misperceive inflation (Kumar et al., 2015; D'Acunto et al., 2021; Candia et al., 2023), and that expectations vary systematically by IQ and gender (D'Acunto et al., 2022, 2023). Household

beliefs respond to Fed communication framing (Coibion et al., 2022) and are often attributed to political narratives like government spending (Binetti et al., 2024; Stantcheva, 2024). They tend to be more attentive to inflation in high-inflation environment (Cavallo et al., 2017; Weber et al., 2025). Incorporating household expectations into standard models helps explain inflation puzzles such as the missing disinflation in the Great Recession and pandemic inflation surge (Coibion and Gorodnichenko, 2015; Beaudry et al., 2024). This paper opens new venues for more research on linking empirical studies on people's inflation expectation to macroeconomic outcomes.

Outline – The rest of the paper is organized as follows. Section 1 presents the setup for static model. Section 2 characterizes the equilibrium. Section 3 presents the main results of this paper. Section 4 presents the dynamic general equilibrium model. Section 5 shows the calibration and estimation results, analyses the transmission mechanisms and consider aggregate productivity shocks. Section 6 shows empirical evidences. The last section concludes. Appendix A contains some of the proofs omitted from the text for the static model. Appendix B contains the proof of the dynamic model and details of calibration and estimation. Appendix C contains details of empirical setup and additional empirical evidence.

1 Static Model

In this section, I develop a macroeconomic model that has three types of agents: shoppers, firms, and a monetary authority. Shoppers face information frictions about monetary shocks and search frictions when deciding where to buy goods. In contrast, firms are fully informed. I start with the static model to explain the main mechanism.

Notation – I use lower case to denote $\log Y$ for any variable Y, i.e., $y = \log Y$ and lower case with hat to denote log-deviation from the steady-state value, i.e., $\hat{y} = \log Y - \log \bar{Y}$.

1.1 Setup

The timeline is as follows. Within a period, the monetary authority first sets the nominal wage. Shoppers are endowed with cash which also serves as the signal about the nominal wage. Firms post prices and shoppers search sequentially. A fraction of shoppers make the purchase in a given round and the remaining keep searching. The period ends until all shoppers make the purchase. I denote the round of search $r = 1, 2, 3, ...$ within a period. All rounds of search happen within one period.

Firm – The economy is populated with a unit mass of firms indexed $k \in [0, 1]$, each of which produces a differentiated product using the following production technology,

$$Y_k = A_k L_k \tag{2}$$

where L_k is the amount of labor employed. A_k is the firm's productivity, which is i.i.d. across firms. Specifically, I assume that $\log A_k$ is drawn from the normal distribution $\Phi_a(a) = \mathcal{N}(0, \sigma_a^2)$.

The marginal cost is $\frac{W}{A_k}$, where W is the nominal wage. The nominal wage represents the average nominal marginal cost. Labor is supplied outside this economy in the current setup.

Suppose the optimal pricing strategy $p^* : \mathcal{R}^2 \to \mathcal{R}$ exists, which I will explore in Section 2. It maps its states $\{a_k, w\}$ to the price of its good, i.e., $p_k = p^*(a_k, w)$. Heterogeneity in a_k across firms induces an endogenous price distribution $F(p|w)$, which has a density distribution $f(p|w)$. If p^* is monotone in both arguments, then the inverse mapping from p_k to a_k given w can be denoted by ${p^*}^{-1}(p, w)$. The price distribution conditional on w is given by,

$$F(p|w) = 1 - \Phi_a({p^*}^{-1}(p, w)) \tag{3}$$

Monetary Authority – The monetary authority draws the log nominal wage from the normal distribution $\mathcal{N}(\bar{w}, \sigma_w^2)$, which serves as the prior for shoppers. Denote the monetary shock by $\hat{w} = w - \bar{w}$.

Shopper – The economy is populated with a continuum of shoppers indexed by $i \in [0, 1]$. At the beginning of the period, each shopper is endowed with cash X_i. It follows,

$$X_i = W \exp\left(\sigma_x \varepsilon_{xi} - \frac{\sigma_x^2}{2}\right) \tag{4}$$

where ε_{xi} is i.i.d. across shoppers and it follows $\varepsilon_{xi} \sim \mathcal{N}(0, 1)$. It is also independent of any other shocks. The variance adjustment term $\sigma_x^2/2$ ensures that $\int X_i di = W$. The shopper treats X_i as signal about W. The expected log nominal wage is then given by $E(w|x_i) = \theta x_i + (1 - \theta)\bar{w}$, where $\theta = \frac{\sigma_x^{-2}}{\sigma_w^{-2} + \sigma_x^{-2}}$. It measures the information friction. Higher θ implies that signals are more informative. Based on the Bayes' rule, shopper i's posterior belief $H(w|x_i)$ follows $\mathcal{N}(E(w|x_i), (\sigma_w^{-2} + \sigma_z^{-2})^{-1})$.

In the rational expectations equilibrium, shoppers know the equilibrium pricing strategy p^*. Using the posterior belief about w, shoppers can construct their own perceived price distributions, $f(p|x)$, conditional on x:

$$f(p|x) = \int f(p|w)h(w|x)dw \tag{5}$$

The above equation yields two implications. First, when information is perfect, i.e., $x = w$, the perceived and objective price distributions coincide. Second, if $h(w|x_i)$ first-order stochastically dominates (FOSD) $h(w|x_j)$ and p^* increases in w, then $f(p|x_i)$ FOSD $f(p|x_j)$. Intuitively, if a shopper believes the nominal wage is on average lower, the perceived price distribution is shifted to the left.

Each shopper buys and consumes only one good.[4] The utility (or value) that shopper i gains from consuming good k is given by,

$$\log \frac{X_i}{P_k} + \frac{1}{\lambda}\epsilon_{ik} \tag{6}$$

[4]In Section 3.4, I extend the model by allowing shoppers to get access to multiple goods in one search.

where $\epsilon_{ik} \sim G$ is match utility between shopper i and good k, which is i.i.d across firms and shoppers.[5] G is triple continuously differentiable and its density function is g. It captures idiosyncratic consumer preferences for certain goods over others. Also, following the literature, I assume that G is *log-concave*.[6] The parameter λ controls the relative importance between two utilities. A larger λ implies more competition on prices. Since X_i only affects the level of utility and does not change the relative utilities across goods, I define the normalized utility as follows,

$$y_{ik} = -p_k + \frac{1}{\lambda}\epsilon_{ik} \tag{7}$$

According to (7), the value distribution of drawing a random good is a convolution of the perceived price distribution and the distribution of match utility,

$$\psi(y|x) = \int f(u-y|x)\lambda g(\lambda u)du \tag{8}$$

Shoppers search sequentially and randomly following Wolinsky (1986) and Anderson and Renault (1999). By incurring a search cost $\kappa > 0$, the shopper can visit a firm to learn both its price and the associated match utility. Each shopper is assigned with the first good for free.[7] Shoppers have free recall, meaning there are no additional costs for purchasing goods from firms they have previously visited. In addition, shoppers do not commit to any plan made before setting out to search.[8] I make the following simplifying assumption such that $f(p|x)$ remains fixed throughout all rounds of search.

Assumption 1. *Shoppers do not learn about monetary shocks from individual prices they observe during the search.*

This assumption is plausible if the variance of idiosyncratic productivity shocks is way larger than the variance of the monetary shocks. Even when a shopper consistently encounters firms charging high prices, she attributes this to bad luck rather than an increased nominal wage. I make

[5]The price distribution is not degenerated and the optimal prices are related to firm's idiosyncratic states, if the random utility term is match-specific. Otherwise, if it is only shopper-specific, it does not matter for search decisions. Firms compete only on price. The price distribution is degenerated to a single optimal price. If it is only good-specific, the optimal prices are not related to firm's productivity.

[6]Note that some commonly used distribution functions are log-concave, e.g., normal distribution, uniform distribution, Gumbel distribution. See Bagnoli and Bergstrom (2005) for a broad discussion of log-concavity that do and do not satisfy this condition. The assumption of log-concavity ensures that the hazard rate $\frac{g(x)}{1-G(x)}$ is monotonically increasing.

[7]Shoppers always participate the market with free first good.

[8]If the shopper formulated her search plan prior to search and she committed to that plan, then she would take into account the expected total search costs of sampling, and she would stop with a lower quality match if she were unlucky and happened to sample a sequence of firms for which she ill-suited. In the case of a shopper doing sequential search without commitment, she ignores past fixed costs of search as sunk. Burdett and Judd (1983) considers a search problem with commitment and homogeneous firms.

this assumption for tractability. First, optimal prices are non-linear in productivity and nominal wage. Bayesian updating about the aggregate state is analytically tractable. Second, to know which shoppers stop in each round of search, we need to keep track of each shopper's whole search history, which leads to exploding states. As I will show, our main mechanism remains valid as long as information friction exists. Therefore, this assumption does not affect my results.

Now I state the shopper's problem. Let v_{ir} denote the maximum value among the firms visited by shopper i up to the rth round. By definition, $v_{i1} = y_{i1}$. For $r > 1$, v_{ir} is recursively given by,

$$v_{ir} = \max\{v_{ir-1}, y_{ir}\} \tag{9}$$

In the rth round of the sequential search, the state of the shopper is v_{ir}. The shopper has the option to stop searching and accept v_{ir} or continue searching. The value function for the shopper, $U : \mathcal{R} \to \mathcal{R}$, in each state $v \in \mathcal{R}$, satisfies,

$$U(v|x) = \max\left\{v, -\kappa + U(v|x)\int_{-\infty}^{v} \psi(y|x)dy + \int_{v}^{\infty} U(y|x)\psi(y|x)dy\right\} \tag{10}$$

The maximum represents that the shopper can either receive the maximum value v up to this round and stop searching, or continue searching by incurring a search cost κ and drawing a random good. If the value of that good is lower than v, which occurs with probability $\int_{-\infty}^{v} \psi(y|x)\, dy$, the shopper will retain the value $U(v)$, since she has free recall. Otherwise, she will obtain a higher value from the newly drawn good and update v according to (9). The value function U is stationary only when Assumption 1 holds.[9] The shopper's problem is solved in two steps. First, she finds the U function that solves the functional equation (10). Second, she keeps sampling firms until v first exceeds the expression to the right of the comma in (10).

1.2 Equilibrium Definitions

The equilibrium concept is Perfect Bayesian Nash equilibrium (PBNE).[10] Formally, I define the equilibrium as follows:

Definition 1 (Equilibrium)**.** *A Perfect Bayesian Nash equilibrium is a triplet of allocation, prices, and beliefs such that*

1. *Firms choose the optimal pricing strategy p^* to maximize profits given the optimal search strategy.*

[9] Otherwise, with the information sets expanding over the rounds of search, the value function U should be indexed by the search round.

[10] In the standard search literature, prices are bounded and consumers know the range of prices. Therefore, consumers are able to detect the off-equilibrium prices. The literature imposes that consumers keep the same (passive) beliefs about the distribution of future prices on and off the equilibrium path. In our case, since productivity is assumed to have unbounded support, any positive price is an on-equilibrium price. The regulations on the off-equilibrium belief is not strictly needed in this model.

2. *Shoppers search without commitment. They do not update beliefs after observing prices during the search. Conditional on the information sets, they combine the optimal pricing strategy p^* and other exogenous distributions to compute $U(v|x)$ for each state v. Shoppers' optimal search strategy is then determined by the stopping rule as shown in (10).*

3. *Nominal wage and cash for each shopper are chosen exogenously.*

4. *Goods market clears.*

In addition, I define the full-information equilibrium in which shoppers know the nominal wage. It serves a natural benchmark for the incomplete-information equilibrium.

Definition 2 (Full-Information Equilibrium)**.** *A full-information equilibrium is the equilibrium defined above, except that shoppers know the nominal wage.*

I also refer to the full-information equilibrium where $w = \bar{w}$ as the full-information steady state.

2 Equilibrium Characterization

I now characterize the equilibrium. I proceed in three steps. First, I characterize the search and pricing strategies. Second, I show the existence and properties of full-information equilibrium. Third, I show the monetary neutrality under full information. Finally, I present search and pricing strategies under first-order approximation around the full-information steady state.

2.1 Characterization of Optimal Search Strategy

I first characterize the search strategy. Each shopper needs to first find the U function and then decide when to stop searching. The solution to the shopper's problem is presented in the following proposition,

Proposition 1. *Under Assumption 1, the optimal search strategy follows a threshold rule. The threshold is denoted $v^*(x)$. If $v < v^*(x)$, the shopper keeps searching; otherwise, she stops and makes the purchase. The threshold $v^*(x)$ decreases in x. It is determined by,*

$$\int \int_{\lambda(v^*(x)+p)}^{\infty} \left(\frac{1}{\lambda}\epsilon - p - v^*(x)\right) g(\epsilon) d\epsilon f(p|x) dp = \kappa \tag{11}$$

where $f(p|x)$ is the perceived price distribution and κ is the search cost per search.

Proof. See Appendix A. ■

The optimal search strategy is straightforward. The shopper keeps sampling firms until the target value $v^*(x)$ is reached. Indeed, in Appendix A, I show that the value function $U(v|x)$ is given by,

$$U(v|x) = \max\{v, v^*(x)\} \tag{12}$$

When $v < v^*(x)$, regardless of the current state v, the value of an additional search is constant, equal to the threshold $v^*(x)$, indicating that the shopper always opts to continue searching in this case. Conversely, when $v \geq v^*(x)$, the value function equals the state v, implying that the shopper accepts v.

The left-hand side of (11) represents the marginal utility of additional search at the threshold, i.e., when shopper's state is equal to $v^*(x_i)$. Suppose she samples another firm k, she will prefer the new good if $\frac{1}{\lambda}\epsilon_{ik} - p_k > v^*(x_i)$. Since the shopper can return without cost, the marginal utility obtained in this case is $\max\{\frac{1}{\lambda}\epsilon_{ik} - p_k - v^*(x_i), 0\}$. The threshold is achieved when the marginal utility is equal to the search cost. The threshold is decreasing in the shopper's signal x: A higher signal about the nominal wage leads to a lower threshold. Intuitively, when a shopper perceives the nominal wage to be lower, they believe the price distribution is shifted leftward. This increases the perceived probability of encountering a low-price firm, thereby raising the reservation value and increasing the likelihood of continued search.

This result generalizes the standard threshold rule found in the literature (see Weitzman, 1979; Wolinsky, 1986) by incorporating both an endogenous price distribution and incomplete information.[11] Proposition 1 demonstrates that the optimal search strategy retains the threshold structure.

2.2 Characterization of Optimal Pricing Strategy

I aggregate the optimal search decisions across shoppers and derive the demand allocation across firms and the resulting profits. Since shopper i only purchases the good k if $\frac{1}{\lambda}\epsilon_{ik} - p_k > v^*(x_i)$, the probability of purchasing from firm k is $1 - G\big(\lambda(v^*(x_i) + p_k)\big)$. This probability is the same for all rounds of search as learning from shopping is prohibited. The probability of any shopper purchasing from any firm in each search given w is given by,

$$\rho = \int \int \big(1 - G\big(\lambda(v^*(x) + p)\big)\big) d\Phi_x(x) f(p|w) dp \tag{13}$$

The probability ρ has following meaning. Suppose the mass of shoppers who visit any firm in the first round is one. A fraction ρ of these shoppers settle with the firms they visit in the first round. The remaining $1 - \rho$ shoppers search in the second round, a further $(1 - \rho)^2$ search in the third round, and so on. Then, the expected number of searches of each shopper is ρ^{-1}. Firms are atomistic and take ρ as given when setting prices. Furthermore, the total expenditure spent in firm k after all rounds of search is given by,

$$\omega_k = \frac{1}{\rho} \int \exp x\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x) \tag{14}$$

[11] In the canonical models, firms are homogeneous, implying a degenerate price distribution, and shoppers possess correct beliefs about this common price throughout their search.

The dispersion in the cash in hand affects the expenditure allocation through the dispersion of thresholds and the covariance between the cash in hand and the threshold, as more cash in hand induces higher expectation of the nominal wage. The profit for firm k is then given by,

$$\pi_k = \frac{\omega_k}{P_k}(P_k - \frac{W}{A_k}) \tag{15}$$

The demand is derived by dividing the total expenditure spent in firm k by P_k. The profit is thus the total demand times the profit per sale. Monopolistic firms compete on prices. They maximize the profit in (15) with respect to prices.

Proposition 2. *Let μ_k denote the markup and e_k denote the elasticity of demand. Firm k charges a markup over its marginal cost,*

$$P_k = \mu_k \frac{W}{A_k}; \quad \mu_k = \frac{e_k}{e_k - 1} \tag{16}$$

The elasticity of demand e_k is determined by,

$$e_k = \lambda \frac{\int \exp(x) g\big(\lambda(v^*(x) + p_k)\big) d\Phi_x(x)}{\int \exp(x)\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x)} + 1 \tag{17}$$

Proof. See Appendix A. ■

The optimal price is a markup times the marginal cost. The elasticity of demand depends on the distribution of thresholds. This is a highly non-linear equilibrium. To understand the intuition, we first focus on the full-information benchmark.

2.3 Characterization of the Full-Information Equilibrium

In the full-information equilibrium, shoppers know the nominal wage. Let $v^*(w)$ denote the value of threshold in this equilibrium. From Proposition 1, it is given by,

$$\int \int_{\lambda(v^*(w)+p)}^{\infty} \big(\frac{1}{\lambda}\epsilon - p - v^*(w)\big) g(\epsilon) d\epsilon f(p|w) dp = \kappa \tag{18}$$

where $f(p|w)$ is the actual price distribution. The distribution of thresholds is reduced to a single value $v^*(w)$. It is easy to show that the average search per shopper $\rho = \int \big(1 - G(\lambda(v^*(w) + p))\big) f(p|w) dp$ and the optimal pricing strategy is given by,

$$P_k = \frac{e_k}{e_k - 1} \frac{W}{A_k} \tag{19}$$

$$e_k = \lambda \frac{g\big(\lambda(v^*(w) + p_k)\big)}{1 - G\big(\lambda(v^*(w) + p_k)\big)} + 1 \tag{20}$$

Firm's market power is determined by two factors. First, larger λ implies more competition on prices, inducing higher elasticity.[12] Second, the search friction naturally gives rise to monopoly

[12] Anderson et al. (1987) shows that without search frictions, if G is Gumbel distribution, the demand system is exactly CES and the elasticity of substitution is $\lambda + 1$.

power (Diamond, 1971). The effect of search friction on the elasticity is represented by a hazard function. The density g represents the number of marginal shoppers who are indifferent between making purchase and continuing searching that the firm will lose if it increases its price. The survival function $1 - G$ indicates that increasing prices will increase the profit obtained from all infra-marginal shoppers. The ratio between the two captures the trade-off when setting prices. Moreover, since G is log-concave, the hazard function increases in its argument. As a result, the elasticity of demand increases in the threshold and firm's own price. It has two implications. First, when firms face pickier shoppers implied by the higher threshold, demand becomes more elastic. Second, there is unique solution for price given productivity. Next, I establish the existence of the full-information equilibrium.

Theorem 1 (Existence of the Full-Information Equilibrium)**.** *There exists a unique full-information equilibrium in which consumers search actively.*

Proof. See Appendix A. ■

This theorem is proved in two steps. First, the elasticity of demand increases in price. Individual prices are uniquely determined by (19) and (20). Second, there is a unique solution to (18) for a given price distribution derived from optimal pricing strategy in (19) for a fixed $v^*(w)$. The threshold $v^*(w)$ monotonically decreases with the search cost. The equilibrium is then the fixed point of the reservation value and the price distribution. Note that there always exist a continuum of equilibria in which firms charge sufficiently high prices and shoppers do not search. However, there is only one equilibrium in which shoppers search actively. The following theorem shows our first main result regarding monetary non-neutrality.

Theorem 2. *In the full information equilibrium, monetary policy is neutral. Monetary shocks are fully passed through to firms' prices.*

Proof. See Appendix A. ■

This theorem establishes that when shoppers know the monetary shock, the passthrough from the monetary shock to prices is complete. This is true because in equilibrium, each firm understands that, if itself and all its competitors raise their prices one-to-one with a monetary shock, although consumers do not know individual prices of each firm, consumers understand the relative prices are unchanged and thus there is no gain from searching more or less. This belief is therefore consistent with firms' and shoppers' strategies. In sum, search friction alone is not sufficient to generate monetary non-neutrality. We need the interplay of both search and information frictions.

2.4 Approximate Optimal Strategies

Under incomplete information, both the threshold shown in Proposition 1 and the pricing strategy shown in Proposition (2) are highly non-linear. To proceed, I consider monetary shocks $\hat{w}$

drawn from a distribution with small standard deviation, i.e., $\sigma_w \to 0$. At the same time, I keep signal-to-noise ratio σ_w^2/σ_x^2 fixed, which correspond to a fixed level of information friction, θ.

Let $\hat{p}_k$ denote the log change in price in firm k after a nominal marginal cost shock. Unlike the standard demand system where an aggregate demand function is available and then the price index is naturally defined, in the model with search frictions, there is no obvious way to define a price index. However, the standard theory (Hulten, 1973, 1978) offers a simple non-parametric formula for the *log change* in price index. If preference is homothetic and stable,[13], which is the case here, the log change in the price index is the expenditure share-weighted log changes in all the prices:

$$\hat{p} = \int \hat{p}_k \bar{\omega}_k dk \tag{21}$$

where $\bar{\omega}_k = \frac{1}{\bar{p}}\big(1 - G\big(\lambda(v^*(\bar{w}) + p_k)\big)\big)$ is the expenditure share in the full-information steady state. The following proposition presents the first-order approximation to the optimal search strategy in (11) and the optimal pricing strategy in (17).

Proposition 3. *Fix the variance ratio $\frac{\sigma_z^2}{\sigma_w^2}, \frac{\sigma_s^2}{\sigma_w^2}$. To the first order as $\sigma_w \to 0$,*
***[Part 1]** The average threshold $\bar{v}^*(x) = \int v^*(x) d\Phi(x)$ is given by,*

$$\bar{v}^*(x) = v^*(\bar{w}) - \bar{E}(\hat{p}) \tag{22}$$

***[Part 2]** The elasticity of demand is given by,*

$$e_k = \lambda \frac{g\big(\lambda(v^*(w) + \bar{p}_k + \hat{p}_k - \bar{E}\hat{p})\big)}{1 - G\big(\lambda(v^*(w) + \bar{p}_k + \hat{p}_k - \bar{E}\hat{p})\big)} + 1 \tag{23}$$

where $\bar{E}\hat{p} = \theta\hat{p}$. In addition, e_k increases in $\hat{p}_k - \bar{E}\hat{p}$.

Proof. See Appendix A. ■

This proposition shows that the average threshold $\bar{v}^*(x)$ under incomplete information is equal to the threshold under full-information equilibrium when $w = \bar{w}$, adjusted downward the expected change in the price index. Because the expected change in prices is lower than the actual change due to information frictions, firms behave as if they are facing more selective consumers who are more prone to search.

The second part of the proposition characterizes the elasticity of demand. On the first order, the distribution of thresholds and covariance between cash in hand and thresholds vanish to zero

[13]Homotheticity requires that the income elasticity of demand must equal one for each good. Stability requires that consumers adjust their spending only in response to changes in income and relative prices. Baqaee and Burstein (2023) show how to generalize the standard theory to more general preferences. In our case, according to 15, the demand for firm k is proportional to the aggregate income $W = \int X_i di$, which implies homotheticity. In addition, there is no other factors, such as taste shocks, state of nature, other than relative prices and income, that can affect the demand. The match-specific shocks are integrated out.

and only $\bar{E}\hat{p}$ is retained. It is important that the elasticity depends on the perceived relative price, $\hat{p}_k - \bar{E}\hat{p}$, rather than the actual relative price typically seen in standard demand systems. When a consumer mistakes an individual firm's price increase for a rise in its relative price, they are more likely to initiate search. This is closely related to the intuition in Lucas (1972). Another perspective is that firms behave as though they are competing against competitors who charge lower prices. They reduce prices accordingly.

3 Results on Passthroughs

In this section, I first present define passthroughs and present the main result: the passthrough of a nominal marginal cost shock to the price index is generically incomplete. Then, I show monotone comparative statics and an extension to multiple goods in one firm.

3.1 Main Result

I first define passthroughs. Let φ_k denote the passthrough from the nominal wage shock $\hat{w}$ to price p_k. I refer to φ as correlated-shock passthrough. Following Amiti et al. (2019), I define φ as follows.

Lemma 1. *The price responds to both firm's own cost shocks and competitors' prices,*

$$\hat{p}_k = \gamma_k(\hat{w} - \hat{a}_k) + \xi_k \hat{p} \tag{24}$$

where γ_k is the own-cost passthrough and ξ_k is the cross-price passthrough.

$$\gamma_k = (1 - \frac{d\log\mu_k}{dp_k}\Big|_{\hat{w}=0})^{-1}; \xi_k = \frac{d\log\mu_k}{dp}\Big|_{\hat{w}=0}\gamma_k \tag{25}$$

where μ_k is the markup. The correlated-shock passthrough of each firm is given by,

$$\varphi_k = \gamma_k + \Phi\xi_k \tag{26}$$

The aggregate correlated-shock passthrough Φ is given by,

$$\Phi = \frac{\Gamma}{1-\Xi} \tag{27}$$

where $\Gamma = \int \gamma_k \bar{\omega}_k dk$, $\Xi = \int \xi_k \bar{\omega}_k dk$.

Proof. See Appendix A. ■

The lemma derives the own-cost and cross-price passthroughs using the markup elasticities, which, in equilibrium, can be obtained from 17. The correlated-shock passthrough is defined as the combination of these two components. By integrating both sides, the aggregate correlated-shock passthrough is derived. Importantly, the own-cost passthrough γ does not depend on the information friction θ as it is evaluated at the steady state. In contrast, the cross-price passthrough

ξ depends on how shoppers perceive the change in price index, which is shock-relevant. Note that shoppers know variables in the steady state and they know the information friction parameter θ, so they know the distribution of correlated-shock passthrough and Φ. Therefore, the average expectation of the change in price index is $\bar{E}\hat{p} = \Phi\bar{E}\hat{w} = \theta\hat{p}$. I now state our main results on passthroughs under incomplete information.

Theorem 3. *The model economy can be summarized by the following equation,*

$$\hat{p} = \Gamma\hat{w} + (1-\Gamma)\bar{E}^c\hat{p} \tag{28}$$

where $\bar{E}^c\hat{p}$ represents shoppers' average expectation of $\hat{p}$. The aggregate correlated-shock passthrough is given by,

$$\Phi = \frac{\Gamma}{1-\theta(1-\Gamma)} \tag{29}$$

$\Phi < 1$ *for any* $\theta < 1$ *given* $\Gamma < 1$.

Proof. See Appendix A. ■

This is the main theorem of this paper. It first frames the model economy as a beauty contest game, where firms exhibit strategic complementarity in response to *shoppers'* beliefs about the price index. This is distinguished from the classic beauty contest game between firms (Lucas, 1972; Woodford, 2003), where strategic complementarity is in terms of *firms'* beliefs.

The theorem also establishes that the aggregate correlated-shock passthrough is generically incomplete if there exists any information and search friction. It relies on two key parameters in this model: the aggregate own-cost passthrough, which is the sufficient statics for the effect of search friction on price index, and the degree of information friction, which affects the strategic complementarity in pricing. If $\kappa = 0$, this model reduces to the CES benchmark (Anderson et al., 1987), where consumers observe all prices and match utilities costlessly. Then, $\Gamma = 1$, we achieve monetary neutrality. Similarly, when $\theta = 1$, we have $\bar{E}^c\hat{p} = \hat{p}$, also yielding monetary neutrality. Any case in between gives monetary non-neutrality.

Figure 1 illustrates the three passthroughs. The right panel shows that under full information, the decrease in the own-shock passthrough over productivity is exactly offset by the increase in cross-price passthrough. The resulting correlated-shock passthrough is always one for firms of any productivity. This result aligns with Theorem 2, where I prove the exact result. The decomposition shows the contribution of each passthrough to the correlated-shock passthrough for firms with different productivities. High-productivity firms pass through less of their own cost shocks and more of other firms' prices. Notably, Amiti et al. (2019) establishes that the correlated-shock passthrough is one for broad preferences including nested CES and first-order Kimball demand family. The theorem complements their results and emphasizes that the complete information

about the price index is essential. Indeed, the complete knowledge of all prices is often assumed in these commonly used demand systems.

The left panel shows the case where shoppers have *zero* information about the shock. The cross-price passthrough is uniformly *zero* for any firm. As a result, the correlated-shock passthrough equals the own-cost passthrough. The intuition is that when shoppers do not perceive any change in the price index, firms are not able to pass any of the change in other firms' prices to their own prices. In this case, the monetary shock is as if an idiosyncratic cost shock to firms. In the intermediate cases, the increase in cross-price passthrough is still not sufficient to offset the decrease in own-shock passthrough, leading to incomplete correlated-shock passthrough.

Moreover, the individual correlated-shock passthrough is defined as $\varphi_k = \gamma_k + \Phi\xi_k$. The effect of strategic complementarity is further dampened by the aggregate correlated-shock passthrough as $\Phi < 1$. This induces a even smaller individual correlated-shock passthrough. Intuitively, the fact that shoppers understand $\Phi < 1$ reduces their expected change in price index for any given monetary shock. Firms then reduce response in pricing in response, resulting in even lower passthrough. Therefore, the incompleteness of the correlated-shock passthroughs is amplified by the firms' incentive of setting prices close to the price index, which is a form of real rigidity (Klenow and Willis, 2016).

3.2 Main Insight: Strategic Complementarity Attenuation

In Proposition 3, we have discussed that elasticity of demand for individual firms increases because shoppers perceive a dampened response of price index. Now, I "dig deeper" to show how such mechanism for *individual* firms relates to the attenuation on the strategic complementarity in pricing on the *aggregate* level. For this purpose, I will borrow heavily from game theory as in Morris and Shin (2002), Woodford (2003), and Angeletos and Lian (2023).

Corollary 1. *In the rational expectations equilibrium, the change in price index can be express as the infinite sum of higher-order beliefs of $\hat{w}$,*

$$\hat{p} = \Gamma \sum_{h=0}^{\infty} (1-\Gamma)^h \bar{E}^h \hat{w} \tag{30}$$

where shoppers' h-order average expectation $\bar{E}^h\hat{w} = \bar{E}(\bar{E}^{h-1}\hat{w})$, and $\bar{E}^0\hat{w} = \hat{w}$. Using the fact that $\bar{E}^h\hat{w} = \theta^h\hat{w}$, the aggregate correlated-shock passthrough is given by,

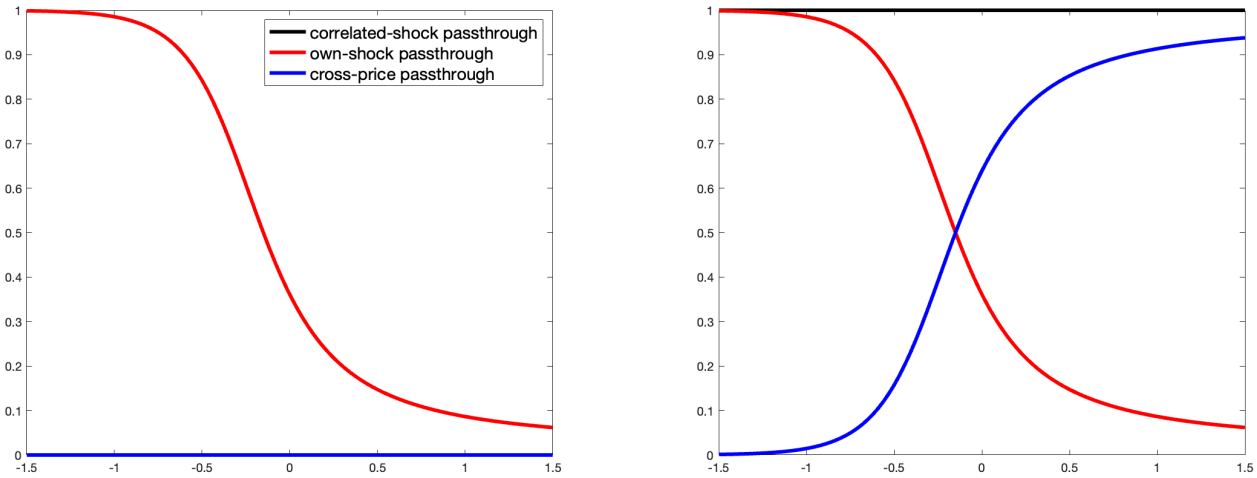
$$\Phi = \Gamma \sum_{h=0}^{\infty} (1-\Gamma)^h \theta^h \tag{31}$$

Proof. See Appendix A. ■

Equation (30) can be derived by iterating (28) in Theorem 3 forward[14]. We can interpret this infinite sum as the aggregate response of price index unfolding in the notional time within a period.

[14]To write the game into the infinite sum of higher-order beliefs as in (30), we need (i) firms are rational and (ii) shoppers know that firms are rational and (iii) the common knowledge of rationality among shoppers and (iv) firms

Figure 1: Distribution of three types of passthroughs under incomplete and complete information



Notes: The figure plots distributions of passthroughs based on one calibration of the model. The red line represents own-cost passthrough. The blue line represents cross-price passthrough. The black line represents correlated-shock passthrough . The left and right panels show passthroughs in the incomplete and complete information cases, respectively.

In the first round, when the monetary shock occurs, assume no firm has yet responded to the monetary shock and that firms expect other firms not to adjust their prices. The shock acts as an idiosyncratic shock to each firm, resulting in a price increase of $\Gamma\hat{w}$. In the second round, it becomes common knowledge among firms that the price index has increased by this amount. Due to the strategic complementarity of price-setting, firms pass an additional portion of the shoppers' perceived increase in the price index from the first round onto their prices, specifically $(1-\Gamma)\bar{E}\Gamma\hat{w}$, under the belief that other firms will not further adjust prices. This iterative process continues. In the $h+1$th round, the additional passthrough is $(1-\Gamma)^h\bar{E}^h\hat{w}$. The intuition connects to the literature on level-k thinking (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019).

Moreover, the information friction attenuates more the high-order strategic complementarity, since the higher-order beliefs of shoppers are more anchored to the prior. For θ close to zero (meaning a sufficiently large departure from common knowledge of the shock), the aggregate correlated-shock passthrough is arbitrarily close to the aggregate own-cost passthrough. But as θ increases (meaning a higher degree of common knowledge of the shock), the higher-order strategic complementarity rises rapidly and thus increases overall passthrough. Therefore, by varying θ between 0 and 1, we can thus span all the values between the aggregate own-cost passthrough and the full-information outcome.

know the common knowledge of rationality among shoppers. In Appendix A, I show which rationality is required in each step of derivation. The rational expectations equilibrium is a "super" concept that includes all these rationality.

3.3 Monotone Comparative Statics

In this section, I first show the cross-sectional properties of individual correlated-shock passthrough and limit results within an equilibrium. Then, I show comparative statics results on search and information frictions across equilibria. I define passthroughs on the primitive: $\varphi_k = \varphi(a_k)$, $\gamma_k = \gamma(a_k)$ and $\xi_k = \xi(a_k)$. The results also depends on the specific functional assumption on the distribution of match utility, G. Following the literature, I assume G is Gumbel distribution.

Proposition 4. *In a given equilibrium, markup $\mu(a)$ and cross-price passthrough $\xi(a)$ increases in productivity; own-cost passthrough $\gamma(a)$ and total passthrough $\varphi(a)$ decreases in productivity. Also, the following limit results hold:*

1. $\lim_{a\to\infty} \mu(a) = \infty$; $\lim_{a\to\infty} \varphi(a) = \Phi\theta$; $\lim_{a\to\infty} \gamma(a) = 0$; $\lim_{a\to\infty} \xi(a) = \theta$
2. $\lim_{a\to-\infty} \mu(a) = \frac{\lambda+1}{\lambda}$; $\lim_{a\to-\infty} \varphi(a) = 1$; $\lim_{a\to-\infty} \gamma(a) = 1$; $\lim_{a\to-\infty} \xi(a) = 0$

Proof. See Appendix A. ■

High-productivity firms exhibit lower own-cost passthrough because their lower prices attract marginal consumers with lower match utility. Given G follows Gumbel distribution, these consumers are more sensitive to relative prices, leading to a lower passthrough of own-cost shocks. For similar reasons, high-productivity firms have strong incentives to keep up with competitors' prices, resulting in stronger strategic complementarity. In addition, these firms set lower prices, and higher markups. Similar to oligopolistic CES models, the lower bound of markups is determined by the degree of substitutability λ between goods, while the upper bound approaches infinity. These results also align with with empirical evidence highlighted in the literature: (i) more productive firms charge higher markups (Amiti et al., 2014); (ii) more productive firms pass through less exchange rate shocks (Amiti et al., 2019).

Importantly, because information frictions dampen only the cross-price passthrough, high-productivity firms experience a more pronounced decline in cross-price passthrough and, therefore, correlated-shock passthrough. Coupled with their high expenditure share, these firms disproportionately contribute to the incompleteness of aggregate correlated-shock passthrough. In addition, according to the limit results in Proposition 4, the correlated-shock passthrough is bounded below by $\Phi\theta$. This implies that $\hat{p}_k - \bar{E}\hat{p}$ is always positive, resulting in uniformly higher perceived relative prices and more elastic demand for all firms. Consequently, following a positive monetary shock, shoppers are more likely to search and passthroughs are incomplete for all firms.

Next, I present results on the comparative statics of correlated-shock passthrough. I establish that, all else equal, (i) the correlated-shock passthrough is larger when the search cost is smaller for all firms, given G is Gumbel distribution, and (ii) the correlated-shock passthrough is larger when information friction is smaller for all firms. In Appendix A, I show the results for own-cost and cross-price passthroughs.

Proposition 5. *If $\varphi(a;\kappa)$ is correlated-shock passthrough for given κ, then for $\kappa_2 > \kappa_1$, $\varphi(a;\kappa_2) < \varphi(a;\kappa_1)$ for $\forall a$, and $\lim_{\kappa\to\infty} \varphi(a) = 0$, $\lim_{\kappa\to 0} \varphi(a) = 1$; If $\varphi(a;\theta)$ is correlated-shock passthrough for given θ, then for $\theta_2 > \theta_1$, $\varphi(a;\theta_2) > \varphi(a;\theta_1)$ for $\forall a$, and $\lim_{\theta\to 0} \varphi(a) = \gamma(a)$, $\lim_{\theta\to 1} \varphi(a) = 1$.*

Proof. See Appendix A. ■

This proposition establishes that two frictions are mutually reinforcing. Higher search cost increases the strategic complementarity in pricing as firms' demand are more sensitive to other firms' prices. Then, the information friction will have larger impact on the aggregate correlated-shock passthrough. Moreover, the limiting results show that passthrough can range from zero to one, allowing for a wide spectrum of monetary non-neutrality. In the dynamic model, shoppers will learn the shock over time. This proposition shows that if learning leads to the common knowledge of the shock, the monetary policy will be neutral in the long run.

Taken together, Theorem 3 establishes that the passthrough from monetary shocks to the price index is generically incomplete. Corollary 1 identifies the core mechanism: information frictions attenuate strategic complementarities in pricing. Proposition 4 and 5 further show that the incomplete passthrough is primarily driven by high-productivity firms. Also, aggregate passthrough declines as search and information frictions intensify.

3.4 Discussion

Search and Strategic Complementarity – The most common way to generate strategic complementarity in the literature relies on Kimball or Kimball-type preferences (Klenow and Willis, 2016). A sequential search model offers a natural microfoundation for such complementarity. Intuitively, relative prices affect consumers' decisions on whether to purchase from the firm they are currently visiting. In the model, the perceived price index of competing firms enters the shopper's reservation value for continued search.

Search costs, together with the relative importance of match-specific preference shocks and prices, jointly determine the elasticity and super-elasticity of demand, as shown in (20). Importantly, the existence of strategic complementarity hinges on positive search costs: as search costs approach zero, the model converges to the CES benchmark of Anderson et al. (1987), which does not generate strategic complementarity. Having established a micro-foundation for strategic complementarity, it is natural to ask what happens if consumers have incomplete information about the outside price index. The answer is given by equation (1), which turns out to be simple and elegant.

Interpretations – Theorem 3 relies on first-order approximation. The responses of price index is symmetric for both positive and negative monetary shocks by definition. Here, I give interpretations about the incomplete passthrough for both cases.

Consider a positive monetary shock, such as a 1% increase in the nominal wage. In response, firms tend to raise prices. If firms increase prices by 1%, they risk losing marginal shoppers who were previously indifferent between making a purchase and continuing their search. Conversely, if

firms do not raise prices at all, they forgo profits from infra-marginal shoppers by charging lower markups. To maximize profits, firms choose prices somewhere in between. In this case, Rotemberg (1982) provides the appropriate interpretation: firms, fearing to alienate marginal consumers, limit the passthrough of the shock.

Now consider a negative monetary shock, such as a 1% decrease in the nominal wage. If firms reduce prices by 1%, they attract more shoppers, who perceive themselves as fortunate to have found a high-productivity firm. On the other hand, if firms do not lower prices, they gain more profit from infra-marginal shoppers who are willing to pay higher markups. To maximize profits, firms opt not to fully decrease their prices. In this case, the appropriate interpretation is that firms limit the passthrough to extract greater profits from infra-marginal consumers.

Connections to New-Keynesian Literature – The New Keynesian Phillips Curve (NKPC), typically implemented through Calvo, departs from the tradition of Friedman (1968), Phelps (1967), and Lucas (1972) by imposing frictions directly on price-setting. Therefore, future change in marginal costs enter the current inflation. However, recent empirical work has cast doubt on this implications of the NKPC.

Models with Calvo frictions predict that changes in short-run expected future inflation have a large impact on current inflation (Werning, 2023). Menu cost models exhibit similar properties (Auclert et al., 2024). Yet, Coibion et al. (2020) and Rosolia (2024) exploit randomized information provision of recent inflation to firms in Italy and find small or even zero effect of changes in inflation expectations on current prices. Similarly, Benedek et al. (2015) and Buettner and Madzharova (2021) show that prices of non-durables do not adjust in anticipation of VAT changes, and they change at the implementation of tax changes. This issue is also reflected in the "forward guidance puzzle", where news about distant future policy shifts leads to implausibly large responses in current inflation, which is not consistent with data (Del Negro et al., 2024).

This paper returns to the tradition of explaining sticky prices through nowcast errors or past-period forecast errors. Expectations of future inflation remain important in the traditional sense emphasized by Friedman (1968): firms must take them into account when signing labor or supply contracts in the current period that will only be paid out in the next.

Extension to Multiple Goods – The baseline model only allows firms to produce one good and each shopper purchases and consume one good from one firm. In this extension, I allow multiple goods to be produced by one firm. The firm produces goods with different productivity for each good.

$$\log A_{kj} = \log A + a_k + a_{kj}$$

where a_{kj} is the productivity of producing good j by firm k. It is i.i.d following $a_{kj} \sim \mathcal{N}(0, \tilde{\sigma}^2_{ap})$. I assume that there is no search frictions when shopping within a firm. Shoppers decide which firm

to purchase form and then buy the CES aggregation of all the goods in the firm. Let P_k denote the CES price index of multiple goods in firm k. We have the following result.

Proposition 6. *Each firm charges same markup over all the products it sells.*

$$P_{kj} = \frac{e_k}{e_k - 1} \frac{W}{A_{kj}} \tag{32}$$

where e_k is the elasticity of demand uniform for all j. It is determined by

$$P_k = \frac{e_k}{e_k - 1} \frac{W}{A_k}; e_k = \lambda \frac{\int \exp(x)\big(g\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x)}{\int \exp(x)\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x))} + 1 \tag{33}$$

where P_k is the CES price index of P_{kj} and $p_k = \log(P_k)$. The passthrough of product-level productivity shocks a_{kj} increases toward one when the number of products increases.

Proof. See Appendix A. ■

This extension is consistent with the result in Hottman et al. (2016) and speaks to the empirical literature on passthrough of exchange rate shock to retail prices. Goldberg and Hellerstein (2013) and Nakamura and Zerom (2010) find complete passthrough of wholesale prices to retail prices for beer and coffee sales in retail stores.[15] The proposition shows that when there are many goods in one store, the passthrough of product-specific idiosyncratic shocks is closed to one. However, this does not mean that the passthrough of aggregate shocks and firm-specific shocks is complete.

4 Dynamic General Equilibrium Model

In this section, I present the full-fledged dynamic general equilibrium model, which completes and extends the static model.

First, I close the model—where the nominal wage and cash in hand are exogenous—in general equilibrium by introducing a new agent: the worker. In particular, there is a representative household which consists of a single worker and a continuum of shoppers. The worker makes the labor supply and aggregate consumption spending decisions, leaving only the decisions about where to buy particular goods to the shoppers. I assume the worker has full information about the model economy, and there is no communication between the worker and the shoppers. Under this setup, a monetary shock can affect both nominal marginal costs and nominal aggregate demand, yet remain unknown to the shoppers. Second, following Christiano et al. (2005), I introduce the nominal wage rigidity, which is crucial to dampen the response of nominal marginal cost in this model and generate plausible impulse responses that are consistent with the data. I will dissect mechanisms of how each of information friction and wage rigidity affects the impulse responses in detail in Section 5.3.

[15] See Gopinath et al. (2011) for a summary of this literature.

In addition, the dynamic model features persistent monetary shocks to the interest rate and slow learning about the underlying shock on the shopper side. For simplicity, I assume away security markets where shoppers can trade claims that are contingent on the search process and final choices.[16]

Timeline – Time is discrete denoted by $t = 0, 1, 2, ...$ The timeline is as follows. The period is divided into morning and afternoon. In the morning, the monetary authority sets the interest rate, and the worker makes decisions on the labor supply, the bond position, and the aggregate consumption spending. In the afternoon, firms post prices, and shoppers search sequentially and decide where to buy. They are constrained by the amount of cash allocated to them. The worker has full information and shoppers have incomplete information about monetary shocks.

Monetary Authority – The monetary authority sets the nominal interest rate. It follows the Taylor rule,

$$i_t = \phi \hat{\pi}_t + v_{mt} \tag{34}$$

where v_{mt} follows,

$$v_{mt} = \rho_m v_{mt-1} + \sigma_m \varepsilon_{mt} \tag{35}$$

where ε_{mt} is the monetary shock and it follows $\varepsilon_{mt} \sim \mathcal{N}(0, 1)$.

Firm – Let A_{kt} denote the firm's productivity. It follows,

$$\log A_{kt} = \rho_a \log A_{kt} + \sigma_a \varepsilon_{akt} \tag{36}$$

where the idiosyncratic productivity shock follows $\varepsilon_{akt} \sim \mathcal{N}(0, 1)$; ρ_a is the persistence.

Worker – In each household, the worker maximizes expected discounted utility with a discount factor $\beta \in (0, 1)$. Period utility is defined over the integration of values which shoppers will obtain from search in the afternoon, given by

$$\int \left(\log \left(\frac{X_{it}}{P_{k(i)t}} \right) + \frac{1}{\lambda} \epsilon_{ik(i)t} \right) di$$

where $k(i)$ denotes the firm selected by shopper i. Due to log utility, the level of consumption spending X_{it} does not influence the shopper's search decision, allowing the worker's and shopper's problems to be separated. Also, I assume that the allocation of the aggregate consumption spending X_t across shoppers is i.i.d across shoppers and time, according to the rule specified in (4). Therefore, the worker only needs to determine the aggregate consumption C_t, which

[16]For example, a security might provide positive returns if a shopper experiences a long sequence of unfavorable draws or if the final choice only slightly exceeds the threshold. Mongey and Waugh (2024) show that the demand allocations in a standard discrete-choice model without search frictions can be different when the market is complete.

implies the aggregate consumption spending $X_t = P_t C_t$, and then transfer randomly the individual consumption spending to each shopper for them to decide where to buy goods.

To match the hump-shaped response of output, following Smets and Wouters (2007), I assume that the worker has external habits. The worker can also save in risk-free bonds B_t (in zero net supply) that pay an interest rate of R_t. The worker's problem writes,

$$\max_{B_t, X_t, L_t} E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(C_t - \lambda_c C_{t-1}) - \frac{L_t^{1+\eta}}{1+\eta} \right)$$
$$\text{s.t.} \quad P_t C_t + B_t = W_t L_t + R_{t-1} B_{t-1} + \Pi_t$$

where L_t is the labor effort, W_t is nominal wage, and Π_t is total nominal profits of firms. The parameter λ_c represents how consumption sticks to its last-period level. Using goods market clearing, the optimization problem gives the first-order Euler equation as in Smets and Wouters (2007),

$$\hat{y}_t = c_1 \hat{y}_{t-1} + (1 - c_1) \hat{y}_{t+1} - c_2 (\hat{i}_t - \hat{\pi}_{t+1}) \tag{37}$$

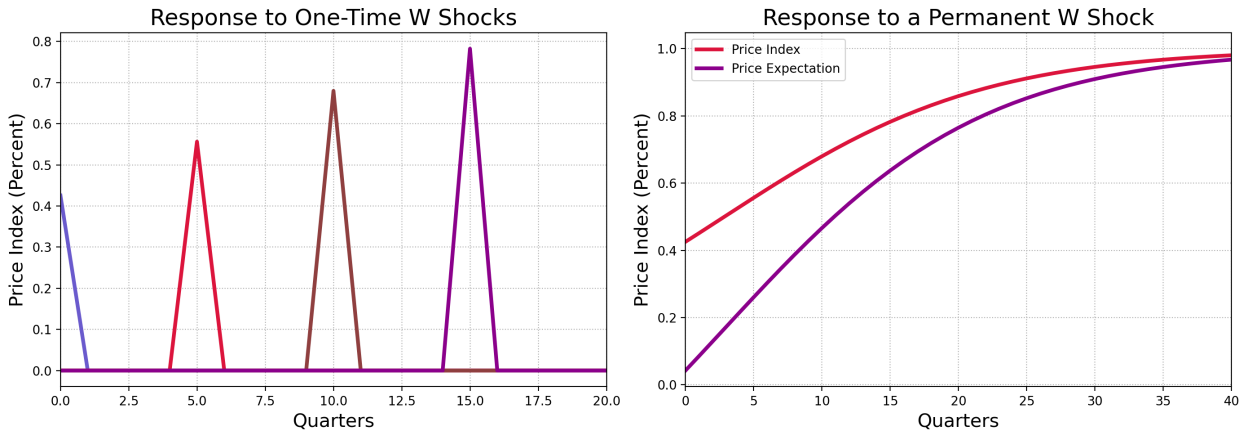
where $c_1 = \frac{\lambda_c}{1+\lambda_c}$ and $c_2 = \frac{1-\lambda_c}{1+\lambda_c}$. Larger λ_c not only dampens the response of output to the change in the real interest rate but anchor its level to the last-period output response. In Section 5.4, I consider an alternative setup following Auclert et al. (2020) to generate hump-shaped output response.

Shopper – In the afternoon, each shopper receives the consumption spending X_{it} and solve the static problem in (10) period by period. In principle, they can infer the information that workers received in the morning from X_{it}. However, since X_{it} is a signal about endogenous variables, incorporating it would significantly increase computational complexity, which is beyond the scope of this paper (See Auclert et al. (2020) for a discussion.). For tractability, I set $\sigma_x \to \infty$ and only consider exogenous signals as follows.

Following Auclert et al. (2020) and Angeletos and Huo (2021), I assume that each shopper receives noisy information about the monetary shock. In particular, Auclert et al. (2020) model consumer learning about persistent monetary shocks by decomposing the shock into a sequence of independent news shocks $\{v_{mt}\}_{t=0}^{\infty}$. I adopt a similar framework, assuming that shoppers begin receiving signals about each news shock starting from date 0. Specifically, for a news shock that will hit at data t, shoppers receive independent private signals in each period $j \leq t$. Then, applying standard Bayesian updating, shoppers' average belief about this shock at date $j \leq t$ is $\bar{E}_j v = \frac{\sum_{\tau=0}^{j} \sigma_{s\tau}^{-2}}{\sigma_m^{-2} + \sum_{\tau=0}^{j} \sigma_{s\tau}^{-2}} v$. Following Auclert et al. (2020), I choose the sequence $\sigma_{s\tau}$ such that $\bar{E}_j v = 1 - (1-\theta)^j$, for $j = 1, 2, \ldots$, where θ is a constant representing both the degree of information friction and the speed of learning for shoppers.[17] Stacking equation (28) across t into a

[17]The exponentially decaying information friction induced by the time-varying sequence of variances of signals ensures that average beliefs about the monetary shock converge fast. In contrast, using a constant $\sigma_{s\tau}$ across all τ results

Figure 2: Properties of the Passthrough Matrix $\mathbf{\Psi}$



Notes: The figure plots the responses of price index to one-time nominal wage shocks (left panel) and a permanent nominal wage shock (right panel) when $\theta = 0.1$. In addition, the average expectation of price index is plotted in the right panel.

vector-valued equation and rearranging, we obtain,

$$\hat{\boldsymbol{p}} = \underbrace{\Gamma(\boldsymbol{I} - (1-\Gamma)\boldsymbol{\Theta})^{-1}}_{\mathbf{\Psi}} \hat{\boldsymbol{w}} \tag{38}$$

where $\boldsymbol{\Theta}$ is a diagonal matrix with entries $1-(1-\theta)^j$ for $j = 0, 1, 2, \ldots$. Following Auclert et al. (2024), I denote the passthrough matrix as $\mathbf{\Psi} = \Gamma(\boldsymbol{I} - (1-\Gamma)\boldsymbol{\Theta})^{-1}$. It characterizes the response of the price level to nominal wage changes.

The left panel of Figure 2 presents the responses of price index to one-time nominal wage shocks. Two properties are worth noting. First, this information-based Phillips curve exhibits no forward-looking behavior: firms pass through the nominal wage shock to prices contemporaneously, at the time the shock occurs. Second, for shocks that materialize later, shoppers have more time to acquire information, resulting in a larger price index response. Both properties resemble the firm-side information-based Phillips curve discussed in Mankiw and Reis (2002). The idea that this or last period's today's inflation expectation matters for today's inflation is closely related to Friedman (1968) and Lucas (1972). The right panel of Figure 2 shows the response of price index and average expectation of price index to a permanent nominal wage shock. When learning is slow, i.e., $\theta = 0.1$, the price index converges to 1 in about 10 years. The expectation initially undershoots significantly the price index and gradually converge to it over time. Following Auclert et al. (2024), I then derive the *generalized Phillips curve* by substituting $\hat{\boldsymbol{w}} = \hat{\boldsymbol{w}}^r + \hat{\boldsymbol{p}}$, where $\hat{w}^r$ is the log-deviation of real wage,

in unrealistically slow learning. For example, if the signal-to-noise ratio is $1/9$, then even at date 100, the average belief remains 9% below the true shock.

and solving the fixed point,

$$\hat{\boldsymbol{\pi}} = \underbrace{(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{\Psi}\,(\boldsymbol{I} - \boldsymbol{\Psi})^{-1}}_{\boldsymbol{K}}\,\hat{\boldsymbol{w}}^{r} \tag{39}$$

where $\boldsymbol{L}$ is the lag matrix with entries of 1 one line below the diagonal. Equation (39) maps any change in real wage $\hat{\boldsymbol{w}}$ to inflation. Matrix $\boldsymbol{K}$ represents the slope of the *generalized Phillips curve*.

The response of the price index to a 1% nominal wage shock is bounded below by Γ, which is typically 0.4. This lower bound is attained in the limiting case where shoppers receive *zero* information about the shock. In contrast, nominal rigidity à la Calvo—where only 25% of firms can adjust their prices each period, which is typically assumed in the literature—implies a more muted price index response of just 0.12, given the strategic complementarity is 0.6. To generate a more realistic inflation response in the model, I introduce nominal wage rigidity. As shown in equation (28), dampening the response of nominal wages directly reduces the magnitude of the price index response.

Labor Union – Following Erceg et al. (2000) and Auclert et al. (2021), the worker supplies labor to unions, which differentiate the labor into types of labor services denoted by L_{jt}. A competitive labor packer aggregates these differentiated services using a Kimball aggregator (Kimball, 1995), and sells the composite labor input to firms at a nominal wage W_t. Firms use the aggregated labor service L_t in production, where aggregation satisfies:

$$\int_j G_w\left(\frac{L_{jt}}{L_t}\right) dj = 1, \tag{40}$$

with G_w denoting the Kimball aggregator. The aggregator satisfies $G_w'(x) = \exp\left\{\frac{1-x^{\frac{\zeta_w}{\varepsilon_w}}}{\zeta_w}\right\}$, where ε_w is the elasticity of substitution across labor types and ζ_w is the super-elasticity. Wages are set by unions subject to nominal rigidities à la Calvo: with probability $1-\lambda_w$, unions can re-optimize wages in each period, while otherwise wages are indexed to past inflation.[18] The union's optimization problem implies, to the first-order, the following wage Phillips curve:

$$\hat{w}_t^r = w_1(\hat{w}_{t-1}^r + \pi_{t-1}) + w_2(\mathbb{E}_t\hat{w}_{t+1}^r + \pi_{t+1}) - \pi_t - \kappa_w(\hat{w}_t^r - \eta\hat{l}_t - \hat{y}_t), \tag{41}$$

where $w_1 = \frac{1}{\beta}$, $w_2 = \frac{\beta}{1+\beta}$, and the slope parameter $\kappa_w = \frac{(1-\beta\lambda_w)(1-\lambda_w)}{\lambda_w}\frac{\varepsilon_w}{\varepsilon_w - 1 + \zeta_w}$. The slope κ_w governs the extent to which the gap between the real wage and the marginal rate of substitution between leisure and consumption is passed to the real wage.

Reallocation and Aggregate Productivity – A positive monetary shock induces more search under incomplete information, with larger information frictions leading to even greater search intensity and vice versa. Shoppers are more likely to be drawn to high-productivity, low-price firms,

[18] Christiano et al. (2016) derive wage inertia from their specification of how firms and workers negotiate wages. They argue that this micro-foundation outperforms a variant of the standard New Keynesian Calvo sticky wage model.

leading to an expansion in demand for these firms and a contraction for low-productivity firms. This reallocation of demand improves allocative efficiency and enhances aggregate productivity (Baqaee and Farhi, 2020; Baqaee et al., 2024). To see this formally, the aggregate production function is given by:

$$\hat{y}_t = \hat{l}_t + (\hat{v}_t - \hat{p}_t) \tag{42}$$

where $\hat{v}_t = \hat{w}_t \int \bar{\alpha}_k \varphi_{kt}\, dk$ and $\bar{\alpha}_k = \frac{\bar{\omega}_k \bar{\mu}_k^{-1}}{\int \bar{\omega}_k \bar{\mu}_k^{-1}\, dk}$ represents the steady-state labor demand share, where $\bar{\mu}_k$ represents steady-state markup. Importantly, markups are not uniform across firms, so the change in aggregate productivity is nonzero.

To determine its sign, since high-productivity firms have higher markups, the weight $\bar{\alpha}_k$ effectively down-weights the incomplete passthrough φ_{kt} for high-productivity firms and up-weights it for low-productivity firms. Given that φ_{kt} is lower for more productive firms, it follows that:

$$\hat{v}_t / \hat{w}_t > \hat{p}_t / \hat{w}_t$$

Thus, a positive monetary shock induces an increase in aggregate productivity through endogenous reallocation of demand. Higher productivity dampens the response of price index and amplifies the monetary non-neutrality. Following the same method as computing $\hat{p}_t$, $\hat{v}_t$ is given by,

$$\hat{\boldsymbol{v}} = \underbrace{\tilde{\Gamma}(\boldsymbol{I} - (1-\tilde{\Gamma})\boldsymbol{\Theta})^{-1}}_{\tilde{\boldsymbol{\Psi}}} \hat{\boldsymbol{w}} \tag{43}$$

where $\tilde{\Gamma} = \int \bar{\alpha}_k \gamma_k dk$ and $\tilde{\Gamma} > \Gamma$. Then, the aggregate production function is given by,

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{l}} + (\tilde{\boldsymbol{\Psi}} - \boldsymbol{\Psi})\hat{\boldsymbol{w}} \tag{44}$$

Equilibrium – In the limit of small monetary shock, the equilibrium dynamics of variables $\{\hat{\pi}_t, \hat{p}_t, \hat{w}_t^r, \hat{w}_t, \hat{v}_t, \hat{c}_t, \hat{i}_t\}$ is governed by equations (37), (41), (44), (34), (43) and two equations that connect nominal and real variables: $\hat{w}_t = \hat{w}_t^r + \hat{p}_t$ and $\hat{\pi}_t = \hat{p}_t - \hat{p}_{t-1}$.

5 Calibration

In this section, I first calibrate the aggregate own-cost passthrough and estimate the parameters governing search frictions. I also quantify the effect of demand reallocation on aggregate productivity. Next, I estimate the information friction by matching the empirical impulse responses from a structural VAR (SVAR) with those generated by the model.

5.1 Calibration of Search Friction

The search cost in the model is described as a utility cost, which is hard to measure in the data. Fortunately, the aggregate own-cost passthrough is a sufficient statistic for "deep" parameters

Table 1: Baseline Calibration of the Model

Parameter	Description	Value
κ	Search cost	0.25
λ	Relative importance of match utility	4.19

Notes: The table reports the calibrated values for parameters that are related to the aggregate own-cost passthrough.

related to the search friction, i.e., search cost κ, relative importance of match utility λ, and standard deviation and persistence of idiosyncratic productivity σ_a and ρ. Another property of aggregate own-cost passthrough is that it is evaluated at the steady state and it does not depend on the information friction. Therefore, we can separate the calibration of these two frictions.

Amiti et al. (2019) and Gagliardone et al. (2025) estimate the aggregate own-cost passthrough to be around 0.4-0.5 using comprehensive data on Belgian manufacturing firms. [19] Recognizing that buyer-seller relationships in tradable sectors often involve well-informed firms, I also considers estimates using data from non-tradable sectors. In particular, Gopinath et al. (2011) use data from a retail chain containing wholesale cost information and report substantial heterogeneity in own-cost passthrough, with a median of approximately 0.5 for U.S. stores and 0.25 for Canadian stores. Since they cannot control for the pricing behavior of *all* the other firms, as Amiti et al. (2019) do, their estimates may be upward biased due to unobserved strategic complementarities from firms not in this dataset. In sum, the estimates based on data from manufacturing and retail sectors are similar. Both exhibit the strong strategic complementarity in pricing. Based on these findings, I calibrate $\Gamma = 0.4$.

Next, I quantify the effect of demand reallocation and also provide a sanity check of the implied values of the parameters governing the search friction. To proceed, I use the estimate of elasticity of substitution from DellaVigna and Gentzkow (2019). They use NielsenIQ Retailer Scanner database and find that the average elasticity of substitution across stores and products is 0.25, implying a markup of 1.67. In the data, one retailer sell a wide range of products. In Section 3.4, I show that if one firm sell multiple products and there is no search friction within the store, it charges the same markup for all products. In addition, the idiosyncratic productivity dispersion σ_a and persistence are set to be consistent with Decker et al. (2020), i.e., $\sigma_a = 0.3$, $\rho_a = 0.6$.

Table 1 summarizes the calibrated parameters, and Table 2 presents the fit of the model to data. Despite its parsimonious structure, the model is successful in matching key moments in the data. They have three implications. First, although the targeted elasticity of demand is 2.5, the implied λ is just 4.2. Remember $\lambda + 1$ represents the elasticity of demand when $\kappa \to 0$ as shown in Proposition

[19] Whether the sum of own-cost and cross-price passthroughs is not statistically different from one depends on empirical specifications. For example, Amiti et al. (2019) show that, this sum is not statistically different from one. In contrast, Gagliardone et al. (2025) use a specification derived from standard New Keynesian models and estimate the price stickiness is about 0.7, indicating the sum of passthroughs less than one.

Table 2: Model Fit

	Moment	Model	Data	Source
M1	Average markup	1.67	1.67	DellaVigna and Gentzkow (2019)
M2	Aggregate own-cost passthrough	0.4	0.4	Amiti et al. (2019); Gopinath et al. (2011); Gagliardone et al. (2025)

Notes: The table summarizes the moments, model and data values of these moments, and the sources of the empirical values of these moments.

4. This suggests that the search friction accounts for a substantial bulk of the market power. Second, based on the calibrated values of deep parameters, we can infer the average search per shopper ρ^{-1} equal to 1.37. This suggests a relatively large search cost and about 73% shoppers make the purchase in the first round of search. Therefore, Assumption 1 is plausible, as shoppers only visit a limited number of firms on average. Finally, the implied $\tilde{\Gamma}$ based on calibrated parameters is 0.426, which is larger than Γ as expected. However, the magnitude is very small. The reallocation effect is therefore second-order.

5.2 Calibration of Information Friction

I first introduce data and present the empirical findings. Then, I use impulse response matching to estimate the information friction.

Data and Method – The dataset used in the main specification consists of six monthly variables: the industrial production index; the average hourly earnings of private production employees deflated by the GDP deflator; the consumer price index (CPI) including food and energy; consumer inflation expectations from the Michigan Survey of Consumers; the excess bond premium; and the 1-year government bond rate. The sample spans from January 1978—the earliest date for which inflation expectation data are available—through December 2019.

I include the CPI measure that incorporates food and energy prices, motivated by a large body of evidence showing that households form inflation expectations based on salient and frequently encountered prices. For example, Kumar et al. (2015) show that households are more attentive to highly visible prices, such as those for gasoline. Similarly, D'Acunto et al. (2021) find that consumers learn about inflation through their retail shopping experiences. To ensure consistency between the actual inflation series and what households observe and internalize in forming expectations, I therefore use the CPI measure that includes food and energy components.

The Michigan Survey of Consumers does not provide a nowcast measure. I therefore use the lagged one-month–ahead inflation forecast as a proxy for the nowcast: $E_{t-1}\pi_t = E_{t-1}\pi_{t+12}/12$. Then, I take the cumulative sum of this monthly series as a proxy of the expectation of log price index. Note that this constructed series is not identical to the expected log price index, as it relies on the average monthly inflation expectation rather than the true expected monthly inflation rates.

To make the expected log price index directly comparable to the actual log price index, I follow the same procedure to transform the raw CPI series into the same format.[20] The results are similar when using the raw CPI series (See Appendix B). The sample period is 1978M1–2019M12, as the Michigan Survey provides inflation expectation data only from 1978 onward.

I begin by estimating the reduced form of a monthly six-variable VAR, using twelve lags for each variable. The five reduced-form residuals are modeled as linear combinations of five underlying structural shocks. The goal is to identify how the structural shocks to monetary policy affect the contemporaneous reduced-form residuals. Once these shocks are identified, the VAR can be used to trace out their dynamic effects on the macroeconomic variables. To identify exogenous variation for the monetary shocks, I employ external instruments for the one-year government bond rate based on high-frequency surprises in futures market prices around FOMC announcements following the approach in Gertler and Karadi (2015) and Bauer and Swanson (2023). I use the constructed monetary policy shock series provided by Bauer and Swanson (2023).

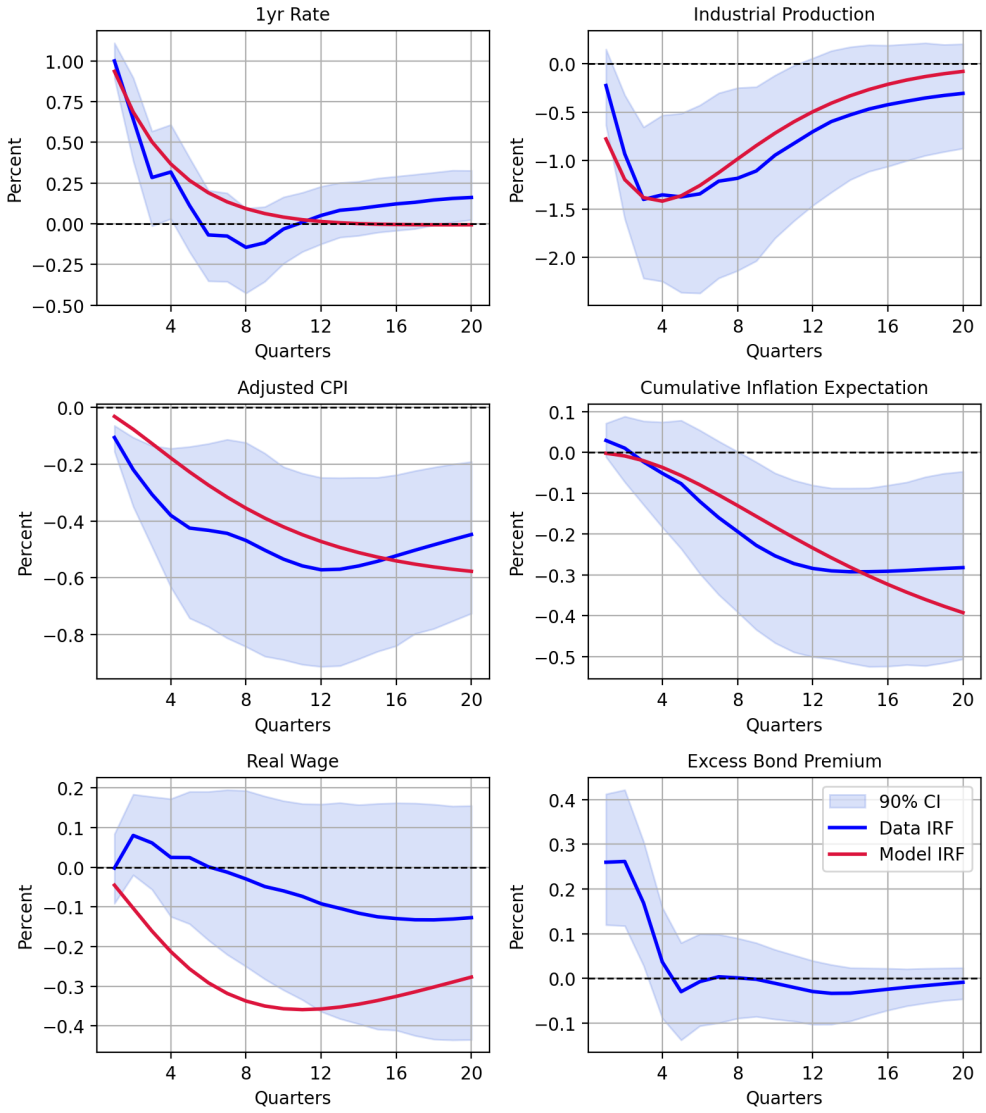
Empirical Findings – Figure 3 presents the main results. The blue lines depict the empirical impulse responses of four macroeconomic variables following a 100 basis point increase in the one-year government bond rate. Industrial production exhibits a hump-shaped response, consistent with existing literature (Gertler and Karadi, 2015), declining by approximately 150 bps in one to two years. adjusted CPI falls on impact by about 10 bps and continues to decline, reaching a total reduction of 50 bps over three years. In contrast, cumulative inflation expectation does not respond on impact, then gradually declines, falling by about 30 bps over the same horizon. These empirical patterns have two key implications. First, inflation expectations systematically undershoot realized inflation, supporting the core assumption of the model—namely, that consumers possess incomplete information about monetary shocks. Second, consumer learning appears remarkably slow: even after three years, consumers have internalized only about two-third of the total decline in cumulative inflation. If we just focus on the relative responses of inflation and inflation expectation, a simple back-of-the-envelope calculation shows: $1-(1-\theta)^{20}=2/3$, which yields $\theta=0.053$, implying an average duration about 18 quarters. I will show in what follows that this provides a useful reference point for the structural estimate of θ obtained through impulse response matching.

Impulse Response Matching – Following Christiano et al. (2005), I use impulse response matching to estimate 4 parameters: nominal wage rigidity λ_w, shoppers' average information friction θ, worker's habit λ_c, and persistence of monetary shock ρ_m. Collecting them in the vector $\Lambda=\{\lambda_w,\lambda_c,\theta_s,\rho_m\}$. Let $\boldsymbol{J},\widehat{\boldsymbol{J}}$ denote model and empirical impulse responses. The estimator $\hat{\Lambda}$ solves,

$$\min_{\Lambda}\left(\boldsymbol{J}(\Lambda)-\widehat{\boldsymbol{J}}\right)'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{J}(\Lambda)-\widehat{\boldsymbol{J}}\right)$$

[20] In particular, I calculate yearly inflation as the percentage change from a year ago of CPI. Then, I divide this inflation by 12 and take the cumulative sum of this monthly series as my new series of CPI.

Figure 3: Empirical and Baseline Model Impulse Responses



Notes: The figure plots the impulse responses for 6 variables following a 100 bp shock on the 1-year government bond rate. The sample period is 1978M1–2019M12. The blue lines represent the empirical impulse responses and the red lines represent the model counterparts. The shaded areas are 90% confidence intervals derived from wild bootstrap. The red line represents the impulse responses implied in Section 4. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

$\boldsymbol{\Sigma}$ is a diagonal matrix that has confidence interval of empirical impulse responses in its entries. I also give more weights to inflation expectation as it is crucial in estimating θ.

Table 3 summarizes the calibrated parameters. The aggregate own-cost passthrough as the associated passthrough matrix that describes the effect of demand reallocation are taken from Section 5.1. In addition, I follow the standard procedure in the literature (Smets and Wouters, 2007)

Table 3: Calibrated and Estimated Parameters.

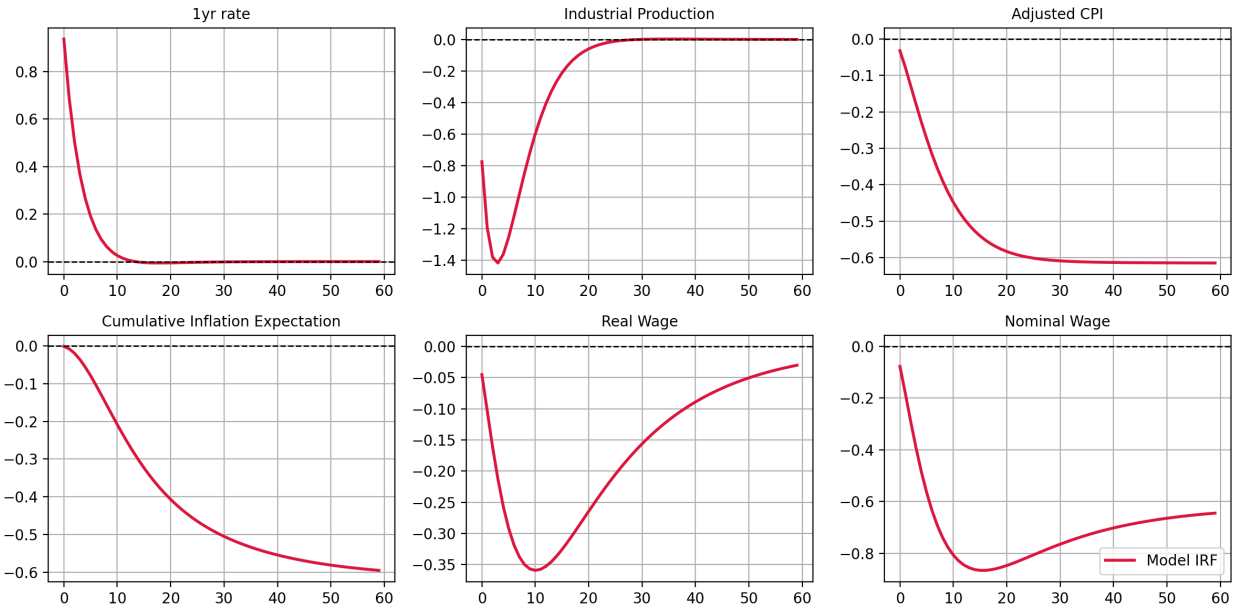
Panel A: Calibrated parameters			Panel B: Estimated parameters		
Param.	**Description**	**Value**	**Param.**	**Description**	**Value**
σ	EIS	1	θ	Information friction	0.055
η	Frisch elasticity	0.5	λ_c	Habit formation	0.798
β	Discount factor	0.998	λ_w	Calvo wage stickiness	0.931
ϕ	Taylor rule coefficient	2	ρ_m	Persistence of monetary shock	0.777
ε_w	Elasticity across labor types	3			
λ_w	Superelasticity across labor types	10			
Γ	Aggregate own-cost passthrough	0.4			
$\tilde{\Gamma}$	Demand reallocation	0.426			

of assuming a wage markup of 1.5 which implies the elasticity across labor types equal to 3, and a Kimball superelasticity for wages of $\zeta_w = 10$.

The model-implied impulse responses at the estimated parameters are shown in red in Figure 3. Overall, the model provides a close fit to the empirical counterparts. In Appendix B, I present results from estimating a Calvo model, where the wage rigidity and the persistence of the monetary policy shock are fixed at the values reported in Table 3. The Calvo parameter is estimated to be around 0.86, i.e., 86% firms cannot adjust prices in each quarter. Figure B.1 shows that the model yields a good fit for the inflation and output responses. However, it fails to capture the response of inflation expectations. This is expected: under full information, inflation expectations move one-for-one with actual inflation. As such, the Calvo model lacks the mechanism necessary to generate the observed wedge between expected and actual inflation. In contrast, the ability of our model to jointly match both series is clearly a success. The estimated information friction parameter for shoppers show a large degree of information friction, $\theta = 0.055$, implying an average duration of incomplete information of 18 quarters. This high level of friction is necessary for the model to replicate the muted response of inflation expectations, which empirically decline by only about two-third of the change in actual inflation in five year. This is consistent with the simple back-of-envelop calculation derived before, indicating that the information friction is mainly identified based on the relative responses of inflation and inflation expectation in the estimation. This estimate is in the ballpark of the estimate of consumer-side information friction in the literature. For example, Auclert et al. (2020) estimates the average duration of household inattention to be 15 quarters to match the hump-shaped response of output in a HANK model.

The persistence of monetary shock is estimated to be 0.78 which is lower than typically estimated in the literature. The reason is that the response of industrial production is less pronounced given the variables and sample period and the response of 1-year government bond rate is not informative about the persistence. I show that the persistence may vary according to different

Figure 4: Longer-Horizon Model Impulse Responses



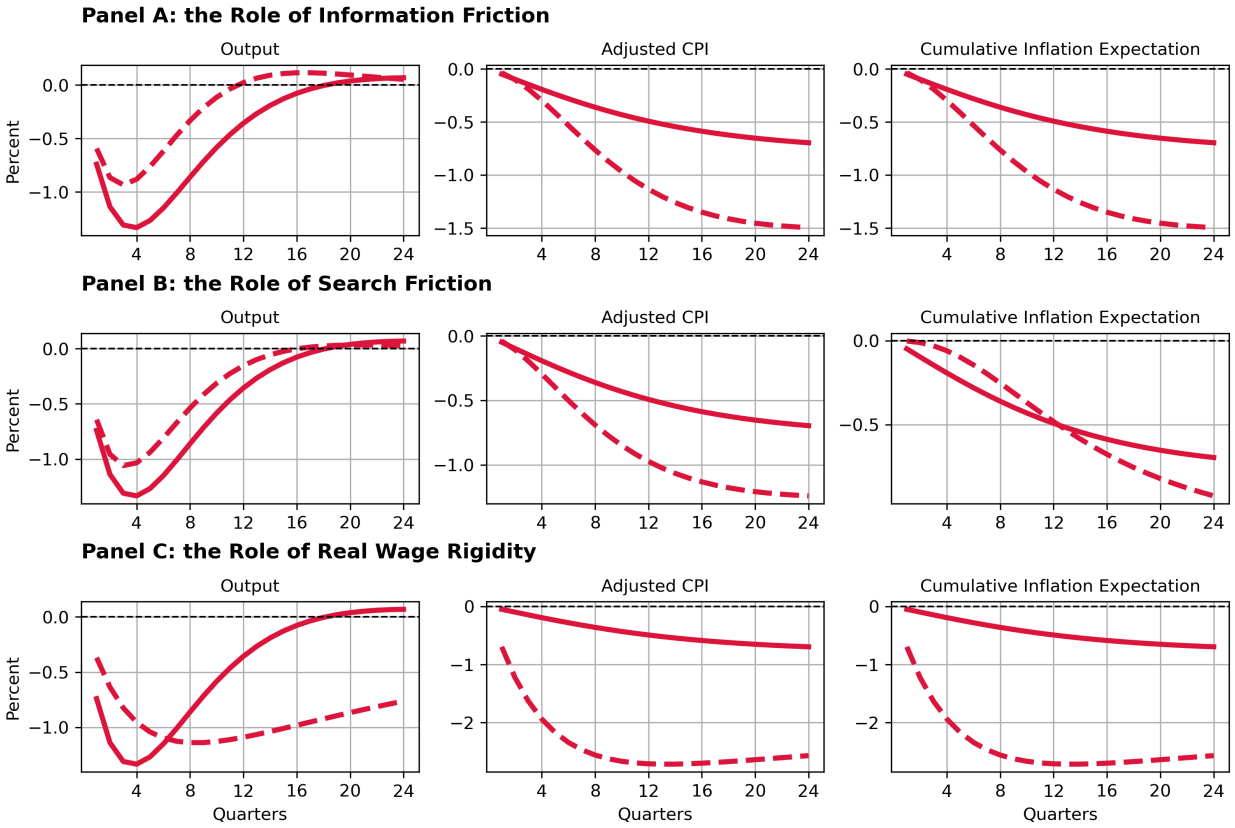
Notes: The figure plots the impulse responses for the longer horizons. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

choices of variables and different specifications of Euler equation. However, the estimates of the key parameter of interest θ remain robust. I discuss them in Section 5.4 and Appendix B.

The estimated wage rigidity in Table 3 is $\lambda_w = 0.931$, indicating a higher degree of stickiness than in Christiano et al. (2005) and Smets and Wouters (2007). This parameter is identified from two sources. First, it must match the impulse response of real wages. Empirically, real wages display little movement over the horizon, consistent with substantial rigidity. This finding is in line with recent evidence using alternative samples and variables (Gagliardone and Gertler, 2023). Second, wage stickiness also affects the dynamics of CPI inflation: lower nominal wage adjustment implies a smaller price response. Because the estimation places more weight on CPI and inflation expectations, the fit for the real wage response is imperfect. A higher λ_w than reported in Table 3 would improve the fit for real wages but further dampen the CPI and inflation-expectation responses. A richer treatment of labor market frictions may be needed to fully capture the real wage dynamics. Nonetheless, the estimate is comparable to Auclert et al. (2020), who calibrate wage and price rigidities to match impulse responses to monetary shocks identified with external instruments from Romer and Romer (2004). Since externally identified inflation responses are typically more muted than those obtained from standard Cholesky decompositions, greater nominal rigidity is required in the model to replicate the empirical evidence.

Figure 4 displays longer-horizon model impulse responses up to 60 quarters. The adjusted CPI reaches its minimum already at 20 quarters, as shown in Figure 3, and then remains flat

Figure 5: Decomposing the Transmission Channels of Monetary Policy



Notes: For all three panels, the solid line plots the baseline impulse responses at the estimated parameters in Table 3 as the benchmark. In Panel A, dashed lines represent impulse responses when $\theta = 1$, holding all other parameters at their estimated values. In Panel B, dashed lines represent the case when $\Gamma = 0.8$ and $\tilde{\Gamma} = 0.805$. In Panel C, dashed linse represent impulse responses when the wage rigidity is zero, i.e., $\lambda_w = 0$.

thereafter. Inflation expectations continue to adjust and fully converge only after 60 quarters, reflecting substantial information frictions. Industrial production, by contrast, returns to steady state relatively quickly, within 20 quarters. The response of real wages is more persistent, consistent with a high degree of wage stickiness. Nominal wages initially decline by more than 80 bps, reflecting the combined drop in real wages and the price index. Over time, as the effect of the shock on real wages dissipates, nominal wages gradually recover and stabilize around 60 bps below the baseline.

In sum, information friction is disciplined by the relative responses of cumulative inflation and inflation expectations; wage rigidity contributes the remaining part of sluggishness of the CPI response and matches the real wage response; habit formation is essential for generating the hump-shaped output response; and the persistence of the monetary shock governs both the persistence of impulse responses of other variables and magnitude of the output response.

5.3 Monetary Policy Transmission Channels

In this section, I discuss monetary policy transmission channels in the dynamic general equilibrium model. In particular, I focus on the role of information friction, nominal wage rigidity and habit formation in contributing to the model-implied impulse responses.

The Role of Information Friction – To understand the importance of information friction, I switch off the information friction for shoppers in the model by setting $\theta_s = 1$, holding all other parameters at their estimated values. Panel A in Figure 5 presents the results. The solid line represents the impulse responses generated from the baseline model. The dashed line shows the responses when shoppers have full information. As expected, the decline in the price index becomes substantially larger—nearly three times greater than in the baseline model. As a result, the degree of monetary non-neutrality is significantly reduced. The cumulative output response is reduced by about 68% and the peak output response shrinks by more than 50 bps. Not surprisingly, the expectation of price index now tracks the actual price index exactly.

The Role of Search Friction – The effect of search friction is governed by the parameter $1 - \Gamma$, which captures the degree of strategic complementarity. According to Proposition 5, lower search friction implies a higher passthrough of idiosyncratic shocks and reduced strategic complementarity. To evaluate the role of search friction, I increase Γ to 0.8 and replot the model's impulse responses. The corresponding demand reallocation parameter adjusts to $\tilde{\Gamma} = 0.805$. Panel B displays the results: the responses of the price index and inflation expectations more than double, while the cumulative output response declines by approximately 46%.

Intuitively, increasing Γ alters the relative contributions of the two sources of price stickiness. A higher Γ places greater weight on nominal wage rigidity and reduces the influence of the information friction channel. Since the latter accounts for a larger share of observed price stickiness (68%), this shift leads to stronger inflation response and a weaker output response.

The Role of Wage Rigidity – To assess the role of wage rigidity, I turn off wage rigidity in the model by setting $\lambda_w = 0$. As shown in Panel C, the absence of wage rigidity results in an immediate and pronounced decline in the price index—approximately 90 bps on impact. The price level continues to fall, reaching a decline of about 300 bps within two years. As a result, the initial output response is attenuated. However, after three years, inflation becomes positive, prompting a rebound in the price index and a rise in the real interest rate due to the Taylor rule. This leads to a sharper contraction in output relative to the baseline model.

The underlying mechanism is as follows. When wages are flexible, nominal marginal costs fall sharply in response to the shock, driving large movement in inflation. However, shoppers' expectations of the price index adjust slowly and remain largely unchanged. As shown in equation 28, firms have an incentive to set prices close to the expected price index. Therefore, as the initial deflationary force dissipates, firms begin to raise prices in response to the inertia in

price expectations. In contrast, with wage rigidity, nominal marginal costs and inflation decline more gradually, giving shoppers' expectations time to adjust. This reduces the divergence between expected and actual prices, thereby smoothing out the inflation dynamics. These results confirm the role of wage rigidity in preventing a sharp change in marginal costs (Christiano et al., 2005) and generating reasonable inflation responses.

Taken together, information friction, search friction or strategic complementarity, and wage rigidity generate substantial monetary non-neutrality, and jointly match the impulse responses of macroeconomic variables and inflation expectation.

5.4 Robustness: Joint Restrictions on Information Friction

I have shown that the relative responses of CPI and inflation expectation are the key determinant of information friction on the consumer side. However, in reality, such frictions may affect more than just consumers' beliefs about inflation. For example, Auclert et al. (2020) argue that standard models with consumption habits are align with the large marginal propensities to consume (MPCs) observed in the data. Instead, they propose that an Euler equation modified by information frictions offers a more plausible explanation. In their framework, the hump-shaped output response arises because, under incomplete information, workers underestimate the true increase in the real interest rate. In a representative-agent setting, where output responds only through intertemporal substitution, such misperception leads to muted adjustments in consumption spending and thus output. As workers gradually learn about the shock, output begins to decline more; however, by that point, the shock dissipates, giving rise to a hump-shaped response.

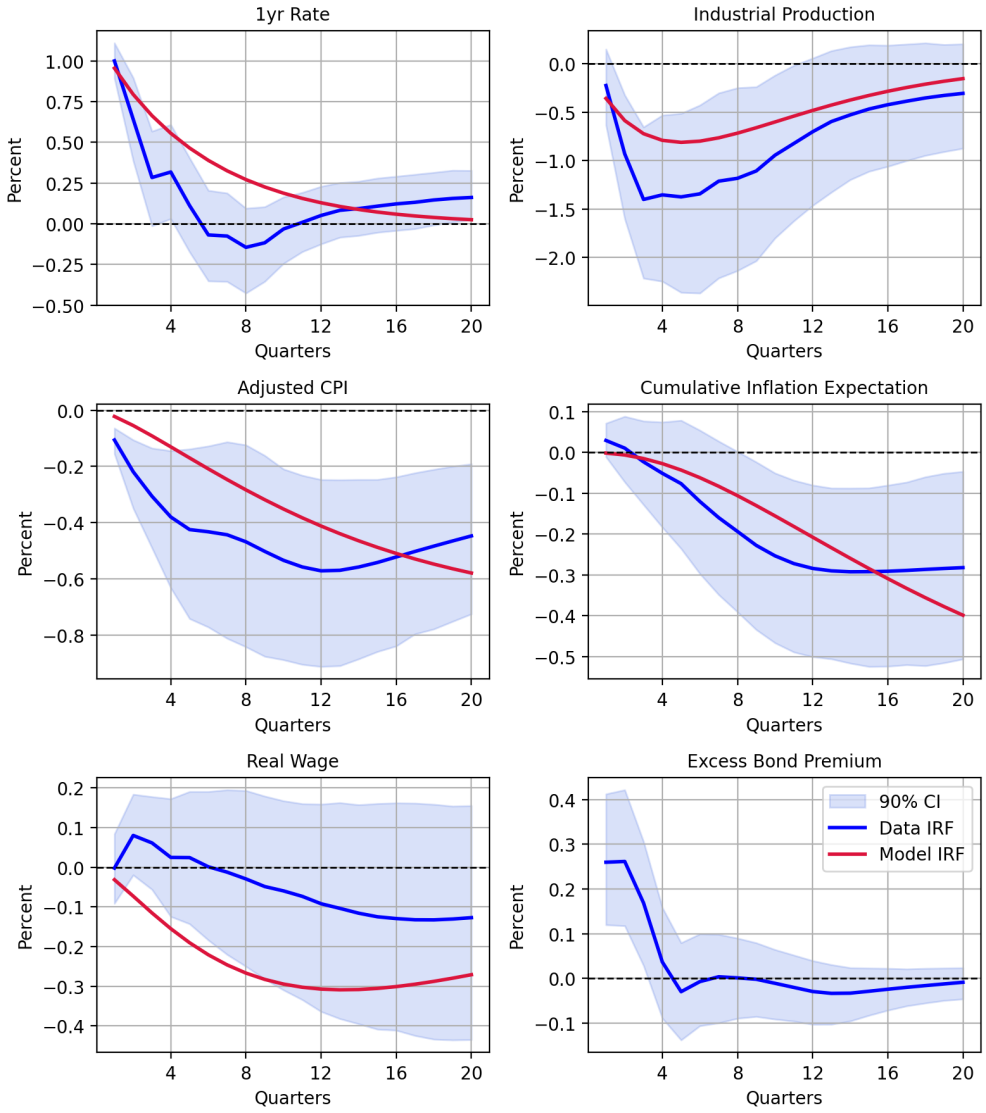
I conduct the following exercise: suppose that the true driver of the hump-shaped consumption response is indeed information friction, as argued by Auclert et al. (2020). This imposes an additional restriction on consumer-side information frictions. In particular, I consider an alternative setup in which there is a continuum of households, each consisting of one worker and a continuum of shoppers. Each household receives a noisy signal about policy shocks in every period. A key constraint is that the worker and shoppers within the same household make joint decisions on consumption, saving, and shopping choices based on the same information set. The bond supply is zero, and the bond market clears.

The information structure of workers are similar to shoppers described in Section 4, which is characterized by a single parameter θ. The optimization problem gives the usual first-order Euler equation with information friction-enhanced MPC (marginal propensity to consume) matrix. Following the sequence-space Jacobian approach developed by Auclert et al. (2021), we have,

$$\hat{\boldsymbol{y}} = \boldsymbol{M}^{r}\hat{\boldsymbol{r}} + \boldsymbol{M}^{y}\hat{\boldsymbol{y}} \tag{45}$$

where $\hat{\boldsymbol{y}}$ is vector of output over time. $\hat{\boldsymbol{r}}$ is vector of real interest rate over time, where $\hat{r}_t = \hat{i}_t - \hat{\pi}_{t+1}$. Matrix $\boldsymbol{M}^{r}$ and $\boldsymbol{M}^{y}$ are information-friction modified Jacobians, which measures partial and general

Figure 6: Empirical and Model Impulse Responses with Joint Restriction



Notes: The figure plots the impulse responses for 4 variables following a 100 bp shock on the 1-year government bond rate. The sample period is 1978M1–2019M12. The blue lines represent the empirical impulse responses and the red lines represent the counterparts generated by the model that uses information friction to generate hump-shaped output response. The shaded areas are 90% confidence intervals derived from wild bootstrap. The red line represents the impulse responses implied in Section 4. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

equilibrium responses of output to real interest rate shocks. In Appendix B, I discuss these Jacobians in detail.

Figure 6 shows that the model continues to fit the empirical impulse responses well under the alternative setup. Table 4 reports the corresponding calibrated parameters. Notably, the estimate of persistence increases from 0.78 to 0.86, which is consistent with the estimates in the literature. The

Table 4: Robust Calibration of the Model

Parameter	Description	Value
θ	Shoppers' information friction	0.057
λ_w	Calvo wage stickiness	0.931
ρ_m	Persistence of monetary shock	0.859

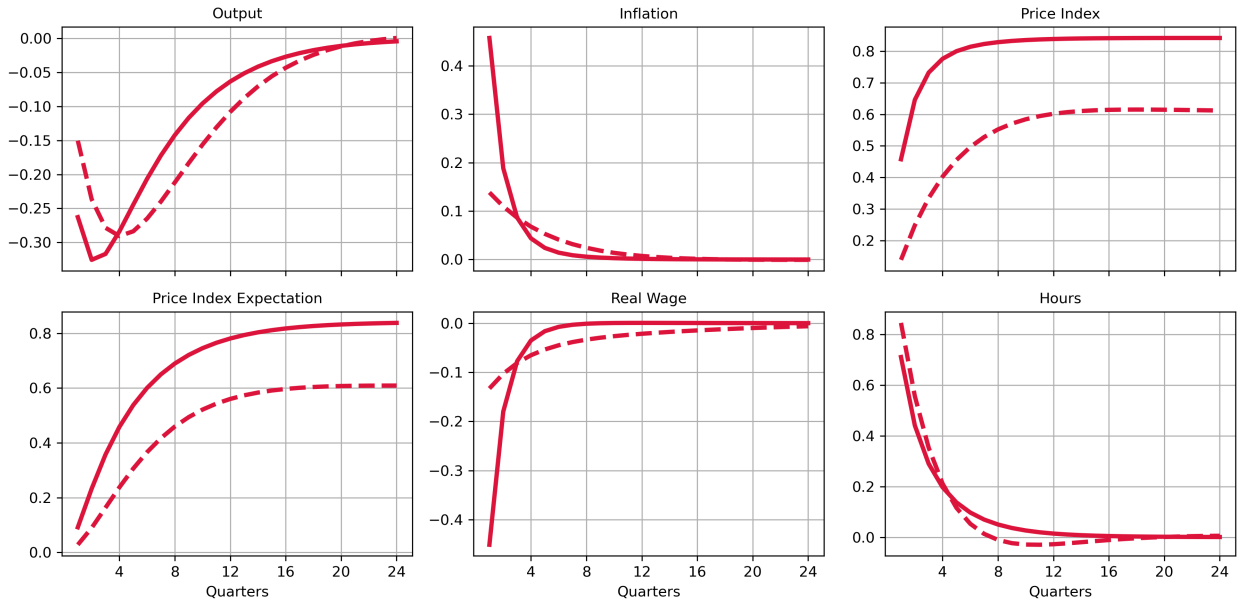
reason is that with information frictions that affect the Euler equation, households respond less to a given shock. Therefore, to match the same impulse response of output, larger persistence of the shock is needed. On the contrary, the habit formation in the baseline amplifies the persistence of output responses. In Figure B.3, I show that due to higher persistence, the CPI drops 20 bps more than the baseline in the long run.

Importantly, the estimated degree of information friction remains nearly unchanged, underscoring the robustness of this key parameter. This stability reflects the fact that the estimate is largely pinned down by the relative responses of CPI and cumulative inflation expectations. The implied duration is broadly consistent with the findings of Auclert et al. (2020). While their study relies on a HANK framework to match impulse responses identified from Romer and Romer (2004) monetary shocks using data from 1969–1996, the present paper employs a different model, sample period, and identification strategy. Despite these differences, both studies yield similar estimates of consumer-side inattention, strengthening confidence in the robustness of this parameter.

To further evaluate the role of information frictions, I conduct two counterfactual exercises. First, I eliminate the information friction related to the shopping decisions while maintaining frictions in consumption-saving decisions. The results, shown in Figure B.2, indicate that the cumulative output response declines by 70%, slightly above the baseline magnitude of 68%. Second, I reverse the experiment—removing the information friction in the consumption-saving decision while preserving it in the CPI. In this case, output drops sharply by 550 bps on impact and gradually returns toward baseline in an exponential manner. Moreover, it nearly doubles the responses of both cumulative inflation and inflation expectations. This highlights the additional role of information friction in the Euler equation in dampening inflation dynamics. However, the underlying mechanisms differ. The information friction in the Euler equation attenuates inflation primarily by weakening the intertemporal substitution channel. In contrast, the friction related to the CPI directly reduce strategic complementarity in firms' pricing decisions, which exerts a quantitatively larger effect on inflation dynamics. This distinction is evident in both Panel A and Panel B of Figure B.2.

Finally, I conduct additional robustness checks reported in Appendix B: (i) replacing adjusted CPI with the raw CPI series; and (ii) replacing the industrial production index with the Brave–Butters–Kelley real GDP index. While the estimated persistence of monetary shocks and

Figure 7: Model Impulse Responses After a Productivity Shock



Notes: The figure plots the impulse responses for 6 variables following a 100 bp shock on the aggregate productivity. The solid lines represent the impulse responses calibrated to values in Table 1 and $\theta = 0.2$, $\rho_a = 0.8$. The dashed lines represent the Calvo model with wage rigidity and habit formation in preferences. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

the degree of wage stickiness vary across these specifications, the estimate of information frictions remains robust. Across exercises, the implied average duration of inattention ranges from 14 to 18 quarters.

5.5 Aggregate Productivity Shocks

In this section, I extend the model to incorporate aggregate productivity shocks. I show the superior properties of inflation and output responses generated from this model compared to the standard Calvo model with wage rigidity and habit formation.

Setup and Impulse Responses – The framework can also incorporate the aggregate supply shocks. In particular, in the static model, suppose firms' productivity has two components,

$$\log A_k = \log A + \sigma_a \varepsilon_{ak} \tag{46}$$

where A is the aggregate productivity shock. Let $a = \log A$. It draws from $\mathcal{N}(0, \sigma_A^2)$. Shoppers receive noisy signals that can inform them about the aggregate productivity shock.

Since both the nominal wage and the aggregate productivity affect prices only through the marginal costs, the optimal pricing strategy is homogeneous of degree zero in $\{W, A\}$, which implies,

$$p^*(a_k, w, a) = p^*(a_k, w - a) \tag{47}$$

Any positive change in aggregate productivity acts equivalently to a proportional decrease in the nominal wage. Therefore, I can similarly define $\hat{p}_k = -\varphi_k \hat{a}$ and aggregate correlated-shock passthrough Φ.

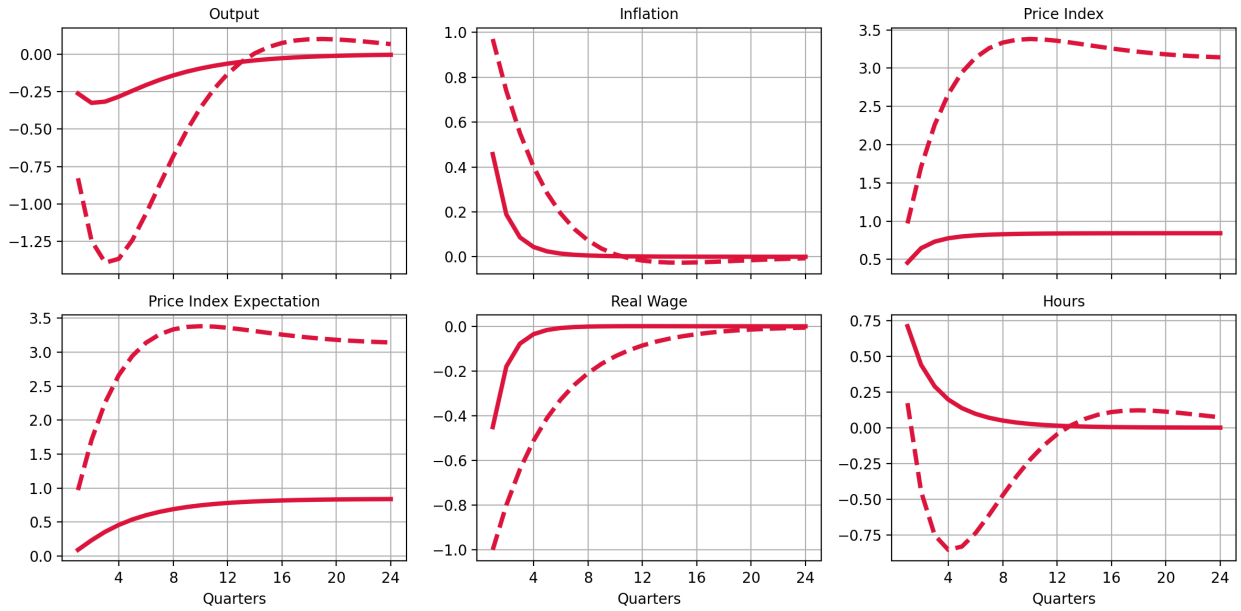
The intuition mirrors that of monetary shocks. Following a negative productivity shock—such as an increase in input prices due to supply chain disruptions—shoppers with incomplete information may misattribute the resulting price increases to idiosyncratic productivity shocks. This misperception induces them to search more intensively for alternative firms. In response, firms face a trade-off: raising prices aggressively risks losing marginal consumers, while holding prices steady sacrifices profits on infra-marginal consumers. The resulting optimal pricing strategy, therefore, features incomplete passthrough of the productivity shock.

The setup for the dynamic general equilibrium model under aggregate productivity shocks is detailed in Appendix B. For calibration, I hold wage rigidity and habit formation fixed at the values reported in Table 1. Motivated by evidence from Kumar et al. (2015), who document that consumers are highly responsive to fluctuations in salient prices such as gasoline and food, I calibrate the shoppers' information friction parameter to $\theta = 0.2$, implying 5-quarter average duration of information friction. For illustration, I also set the persistence of productivity shock to $\rho_a = 0.8$. The solid lines in Figure 7 displays the impulse responses following a 100 basis point negative productivity shock. For comparison, I also consider an enhanced Calvo pricing model incorporating wage rigidity and habit formation, with a Calvo parameter of $\lambda_p = 0.25$, implying 4-quarter average duration of price cell. Under this specification, the output and hours responses are broadly similar to those two models.

However, inflation dynamics differ notably between the two models. In the baseline case, inflation rises sharply by 50 bps on impact but quickly returns to zero. In contrast, the Calvo model generates a more muted inflation response—about 15 bps—followed by a slower decline. This reflects differences in the transmission of nominal marginal cost shocks to prices. Although information frictions reduce strategic complementarities in pricing, and wage rigidity dampens part of the nominal marginal cost, the overall nominal marginal cost still rises by 100 bps due to the exogenous productivity shock, leading to a price increase of roughly 40 bps—consistent with what is observed in Figure 7. By contrast, the Calvo model directly attenuates the passthrough of the entire marginal cost, including both wage and productivity shock, to prices, resulting in a much muted inflation response. Furthermore, since firms anchor their prices around expected price index, the muted response of consumers' expectation of price index results in the faster decline in inflation as the shock dissipates in the baseline model.

Taken together, these results highlight the flexibility of the baseline model in jointly matching two empirically observed regimes: (i) dampened responses of inflation and inflation expectations, alongside significant monetary non-neutrality following monetary shocks; and (ii) pronounced but less persistent inflation response following productivity shocks.

Figure 8: Illustration of Pandemic Inflation Surge



Notes: The figure plots the impulse responses for 6 variables following a 100 bp shock on the aggregate productivity. The solid lines represent the impulse responses calibrated to values in Table 1 and $\theta = 0.2$, $\rho_a = 0.8$. The dashed lines represent the Covid case by setting $\theta = 1$. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

Application to the Pandemic Inflation Surge – A puzzle has emerged in the dynamics of inflation and unemployment during the pandemic. Inflation, likely triggered by supply chain disruptions, surged immediately after the pandemic and remained persistently elevated—contrary to the predictions of a flat Phillips curve estimated using pre-COVID data Hazell et al., 2022. Several theories have been proposed to reconcile these observations. For instance, Benigno and Eggertsson (2023) propose a non-linear New Keynesian Phillips curve; Bianchi et al. (2023) attribute the inflation surge to unfunded fiscal shocks. In this paper, I offer an alternative explanation rooted in the framework developed herein.

Specifically, the historically high levels of media coverage and public discourse surrounding supply chain disruptions and inflation likely reduced information frictions, enabling households to more rapidly internalize the rising production costs faced by firms. Supporting this view, Beaudry et al. (2024) show that realized inflation and household inflation expectations moved in close alignment during this period.[21] Pfäuti (2024) show that the inflation surge causes headline inflation to be above the estimated attention threshold. Here, I consider a counterfactual calibration in which shoppers are assumed to have full information, i.e., $\theta = 1$.

[21]Hazell (2024) further demonstrate that other inflation expectation measures—such as those from the Survey of Professional Forecasters, one-year-ahead inflation expectations derived from swaps, and backward-looking adaptive expectations (e.g., the fourth lag of inflation)—do not exhibit the same degree of sensitivity as household expectations.

Figure 8 illustrates the impulse response when information in the model is highly precise. The increase in inflation is almost doubled, which aligns with the drastic inflation surge at the onset of the supply chain disruptions. The output also drops more than 5 times of the baseline level. However, the output did not decline significantly in reality compared to the model's predictions, likely due to the concurrent fiscal stimulus (Bianchi et al., 2023) and labor market tightness (Barnichon and Shapiro, 2024).

The main policy lesson from this application is that the slope of the Phillips curve is endogenous to the level of information available on the consumer side. Conventional Phillips curves, derived from reduced-form assumptions such as infrequent price adjustments and menu costs, are vulnerable to the Lucas Critique. Certain shocks, in particular, may trigger a sequence of public events, media reports, discussion in the public domain, and policy announcements, which significantly increase households' awareness of the inflation, resulting in a steeper slope of Phillips curve. The monetary authority should adopt a more aggressive stance when household inflation expectation is more accurate, as prices are more responsive to the shock in such environments.

6 Empirics

In this section, I provide two pieces of empirical evidences to support the key mechanisms of the model: (i) heightened attention to inflation increases the state-level slope of Phillips curve and (ii) unanticipated inflation increases consumer search activities.

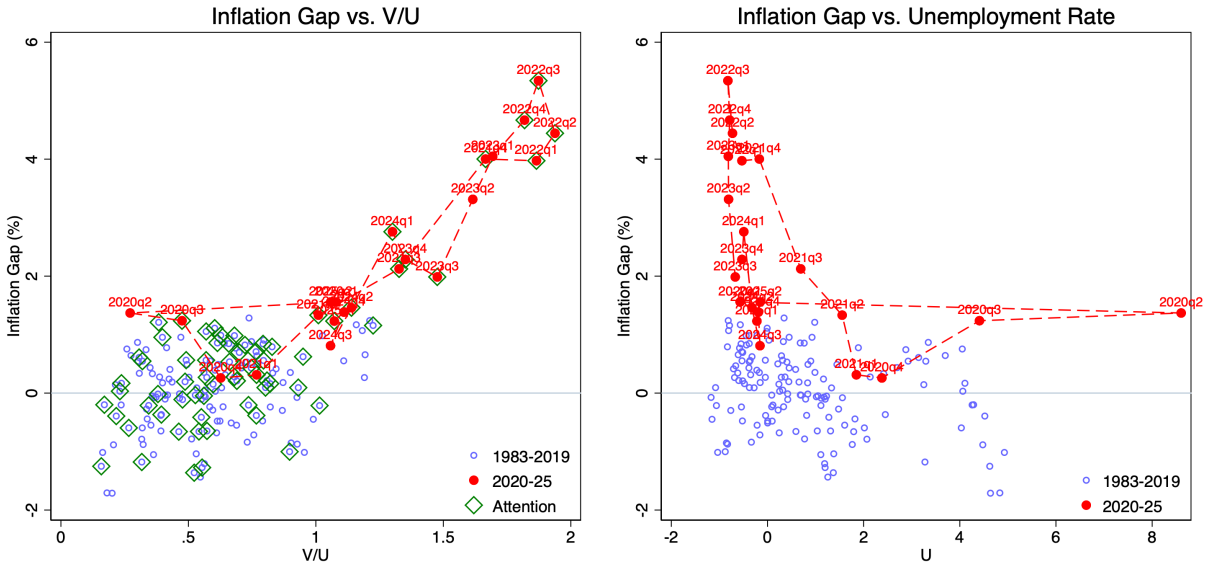
6.1 Inflation Attention and the Phillips Curve

Recent evidence has shown that people are more attentive to inflation in the high-inflation environment (Cavallo et al., 2017; Weber et al., 2025). Importantly, Korenok et al. (2023) and Pfäuti (2024) show that households' attention to inflation increases with inflation only once it exceeds a threshold. Using Google search data and consumer inflation expectation data from the Michigan Survey, they estimate this threshold to be around 3.5–4%. Pfäuti (2024) shows that inflation responses supply shocks become more significant in the high-attention regime. This is an evidence that supports the mechanism of this paper.

This section makes additional progress by estimating how attention to inflation affects the effect of economic slack on core inflation – Phillips curve. Following Pfäuti (2024), I set the threshold at 3.91%. When headline inflation, measured by the CPI, exceeds this threshold, the attention index takes the value of one; otherwise, it is zero. Higher attention should be interpreted as lower information friction following rational inattention literature (Sims, 2003).

Following Ball et al. (2022) and Barnichon and Shapiro (2024), I measure economic slack using the vacancy–unemployment ratio (V/U). Job openings data from 2000 to 2025 come from the Job Openings and Labor Turnover Survey (JOLTS) and are extended back to 1950 using the composite

Figure 9: Inflation Gap and Economic Slack



Notes: The figure shows the relationship between inflation gap and vacancy-unemployment ratio and unemployment gap. The blue circles represent the data points during 1983 and 2019. The red points represent the period 2020-2025. The green squares on the data points indicate the headline inflation is above the 3.91% inflation attention threshold.

help-wanted index constructed by Barnichon (2010).[22] For unemployment, I adjust by excluding temporary layoffs, which were unusually large at the onset of the pandemic.

As a measure of core inflation, I use the weighted median CPI inflation rate published by the Federal Reserve Bank of Cleveland, as in Ball et al. (2022).[23] I then define headline inflation shocks—referred to as *supply shocks*—as the deviation of headline inflation from core inflation. This series is available only from 1983 onward, coinciding with the post-Volcker disinflation period when inflation expectations became more firmly anchored. To isolate the effect of economic slack on inflation, I subtract 10-year inflation expectation provided by the Federal Reserve Bank of Cleveland from core inflation and define the resulting variable as the *inflation gap*. The final sample period is 1983m1-2025m4.

Figure 9 illustrates the relationship between the inflation gap and two measures of slack: the V/U ratio and the unemployment gap.[24] The left panel shows a strong relationship between

[22]Barnichon (2010) splice the Conference Board's print newspaper help-wanted index with the online help-wanted index to build the long historical series.

[23]This measure excludes extreme price changes and mitigates the impact of large swings in specific industries, not just food and energy as in CPIXFE.

[24]The unemployment gap is defined as the unemployment rate minus the noncyclical unemployment rate estimated by the Congressional Budget Office (CBO). Monthly data are aggregated to the quarterly frequency to smooth short-term fluctuations.

the inflation gap and the V/U ratio, particularly during the pandemic when the ratio increased sharply. Importantly, this surge was driven primarily by a rise in vacancies rather than movements in unemployment (Barnichon and Shapiro, 2024; Afrouzi et al., 2024). Consequently, as the right panel demonstrates, the unemployment gap performs much worse as it barely moved after the lockdown.

Before the pandemic, episodes of high headline inflation were concentrated in the 1980s, a period when long-run inflation expectations were also high. As a result, the inflation gap was not necessarily large at that time, even when households were attentive to inflation. In other words, the attention index was not systematically correlated with the inflation gap. This is evident in the left panel: the green squares (indicating periods when the attention index equals one) on top of the blue circles do not follow a clear pattern with inflation gap. The picture changes during the pandemic. With long-run expectations relatively low while actual inflation surged, attention to inflation and the inflation gap became strongly correlated.

Empirical Results – In the Phillips curve regression, I regress the inflation gap over the average of V/U and supply shocks over the current and previous 11 months as a parsimonious way of capturing the lags in the effects of labor market tightness and headline inflation fluctuations. I normalize the forcing variables to have unit variance, so that the coefficients are directly comparable across columns. A larger coefficient thus indicates a larger explanatory power.

Table 5 presents the regression results. Columns (1) and (2) report estimates when only the V/U ratio and supply shocks are included. The coefficient on V/U more than doubles, and the adjusted R^2 rises from 7.6% to 44% once the pandemic period is incorporated. This demonstrates the sharp increase in the explanatory power of V/U, consistent with the raw data patterns shown in Figure 9. By contrast, the passthrough of supply shocks before the pandemic is small, in line with recent evidence.

Columns (3) and (4) add the attention index and its interactions with slack and supply shocks. The coefficients on slack and supply shocks decline somewhat but remain statistically significant. Before the pandemic, heightened attention to inflation does not alter the effect of slack, but it substantially increases the passthrough of supply shocks to the inflation gap—bringing it closer to the magnitude of the slack effect. This suggests that while headline inflation shocks are typically independent of core inflation, heightened attention changes their transmission. This is consistent with the recent survey evidence (Link et al., 2023). The absence of an attention effect on slack before the pandemic is likely due to limited labor market tightness during the 1980s. When pandemic data are included, however, the interaction between attention and slack becomes significant, reflecting the salience of labor market conditions during this period. Overall, higher attention amplifies the effects of both slack and supply shocks on core inflation, yielding a steeper Phillips curve.

Columns (5) and (6) replace V/U with the unemployment gap as the measure of slack. Both the coefficients and adjusted R^2 are considerably smaller, particularly when the pandemic is included,

Table 5: The Effect of Attention to Inflation on the Regional Phillips Curve

	(1)	(2)	(3)	(4)	(5)	(6)
	1983–2019	1983–2025	1983–2019	1983–2025	1983–2019	1983–2025
	V/U	V/U	V/U	V/U	U	U
Slack	0.372**	0.825***	0.302*	0.604***	-0.132	-0.246*
	(0.144)	(0.135)	(0.155)	(0.137)	(0.138)	(0.147)
Supply	-0.000	0.244*	-0.136	-0.114	-0.152	-0.268**
	(0.092)	(0.136)	(0.102)	(0.110)	(0.102)	(0.125)
Attention			0.305**	0.227	0.277**	0.368*
			(0.132)	(0.152)	(0.127)	(0.196)
Attention × Slack			0.219	0.414***	-0.157	-0.078
			(0.176)	(0.123)	(0.142)	(0.185)
Attention × Supply			0.321**	0.593***	0.294**	0.966***
			(0.128)	(0.142)	(0.122)	(0.286)
R-squared	0.080	0.444	0.125	0.523	0.097	0.251
Adj. R-sq	0.076	0.442	0.115	0.518	0.087	0.243
Observations	433	497	433	497	433	497

Notes: V/U denotes ratio of vacancies to unemployed (12 month average). U denotes unemployment gap (12 month average). Supply represents the headline inflation shocks. Attention is the attention index. Newey-West standard errors with 12 lags are reported. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

underscoring the low explanatory power of the unemployment gap. Moreover, the supply shocks have significantly negative impact on core inflation, and heightened attention again has little effect on slack but strongly amplifies the impact of supply shocks—much more so than in the specifications using V/U. This difference arises because the unemployment gap exhibits little variation during periods of high labor market tightness. Appendix C provides robustness checks with alternative thresholds for defining the attention index. Since identification of the interaction between attention and slack is driven primarily by the pandemic episode, the strong significance could also reflect nonlinearities in the effect of slack (Benigno and Eggertsson, 2023). I show that including nonlinear terms does not alter the results, and these terms are not statistically significant once terms on attention are included.

6.2 Unanticipated Inflation and Search Activities

In this section, I provide evidence on the effect of unanticipated inflation on search activities. I first briefly describe the data, then propose a new measure of search activity, and lastly I present the results.

Data – The data source is the NielsenIQ Consumer Panel Data set.[25] The sample period is 2004Q1 - 2019Q4.[26] NielsenIQ tracks the shopping behavior of average 55,000 households every year. Each household uses in-home scanners to record purchases. Households also record any deals used that may affect the price. These households represent a demographically balanced sample of households in 49 states and about 3,000 counties in the United States. Each household stays in the panel for 9 years on average. The dataset has over 1,000 NielsenIQ-defined product modules that are organized into over 100 product groups and covers around 30% of total US expenditures on food and beverages and 2% of the total household consumption.[27]

The dataset contains information about each shopping trip the household takes, such as the retailer, the spending on each product defined as a barcode, the product module and group that the product belongs to, and the date of transaction. Moreover, the data includes households' demographic information such as age, education, employment, marital status, which are updated annually. I aggregate the dataset to the household-quarter level.

Measurement of Search Activities – The search protocol described and characterized in the model simplifies the actual shopping process. In practice, shoppers tend to make the majority of their purchases at a primary retail store, while visiting other stores for specific needs. For instance, during the sample period, shoppers allocate an average of 65% of their total spending on a given product group to the store they visit the most frequently. Additionally, shoppers often purchase multiple items in a single trip, and most stores offer a wide range of product groups. As a result, search activities may manifest as a reallocation of spending across stores for the same product group. For example, a shopper may initially buy milk from store A and cheese from store B but, after a shock, begin purchasing both milk and cheese from store A while reducing cheese purchases from store B. This suggests that simply counting the number of trips and the number of distinct stores shopper visits may not fully capture the search effort, even though it is the definition of a search in the model.

To account for the empirical patterns of shopping, I construct the measure of search effort as follows. First, for each consumer and each product group, I find the store she visits most frequently in the entire sample period. I call that store the *routine store*. Second, if she does not buy from the routine store or she buys from multiple stores for this product group, I record the total spending that she spends on the stores other than the routine store. Repeat this process for all the product

[25] Researcher's own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[26] I do not use the data after 2019 because NielsenIQ changes the organization of the dataset since 2021. Building a concordance linking product modules before and after 2021 is out of scope of this paper.

[27] For further discussion of the NielsenIQ data, see Broda and Weinstein (2010) and Jaravel (2019).

groups. Note, that households may have different routine store for each product group. Finally, I define the non-routine share of spending as the ratio of all the non-routine spending over all groups and the total spending in the given quarter. The non-routine share captures the *extra* search effort to look for the products that shoppers usually buy from the routine store. Figure C.1 in Appendix C shows that the time trend of the non-routine share of spending declines from 40% to 34%. In addition, it is highly correlated with the average shopping time derived from American Time Use Survey.[28] This makes us confident that this measure captures correctly the search activities.

I also construct other two measures of search activities. The first is the number of trips each household takes in a quarter. Second, I measure the number of distinct retail stores visited by consumers in a quarter. Both measures captures some aspects of our primary measure. Table 1 presents the summary statistics for these search measures. On average, consumers' non-routine share of spending is 36%. They make about 2 shopping trips per week, visit approximately 6 distinct retailers per quarter. The variance of these measures is substantial. This is consistent with our model in which search is primarily driven by idiosyncratic match utility shocks and productivity shocks, which may have large variances.

Table 6: Descriptive Statistics of Household Search Behavior

	Mean	Median	S.D.	10th Percentile	90th Percentile
Non-routine Share(%)	35.9	33.8	24.7	4.0	70.7
Number of Trips	23.3	21	13.8	8	42
Number of Distinct Retailers	6.2	6	3.27	2	11

Notes: The table reports summary statistics for key household search behavior variables in the sample. *Non-routine Share* is the share of non-routine spending in a quarter. *Number of Trips* is the total number of shopping trips per quarter per household. *Number of Retailers* is the distinct retailers visited by a household per quarter.

Unanticipated Inflation – I use quarter-to-quarter inflation for food and beverages as my main measure of inflation. I assume that consumers use historical data to forecast the current inflation. In particular, consumers estimate a simple OLS regression of inflation on its four lags. The residual from this regression is the unanticipated inflation, which has zero mean and a standard deviation of 60 bp in the sample period. I focus on inflation for food and beverages because the majority of NielsenIQ products fall into this category. In addition, I consider headline inflation. The unanticipated inflation in this case has zero mean and a standard deviation of 49 bp. I normalize the two series of unanticipated inflation, so their units are standard deviation in our sample.

Empirical Specification – To assess the changes in search behavior after an unanticipated inflation shock, our baseline specification is:

$$y_{it+1} = \lambda_i + \beta\tilde{\pi}_t + X_{it} + e_{it} \tag{48}$$

[28]The time spent on shopping declines from 18.5 hours per month to 16 hours per month.

Table 7: Non-routine Share of Spending and Unanticipated Inflation

Dep. var.: Non-routine Share	(1)	(2)	(3)	(4)
Unanticipated inflation (F&D)	0.122***		0.115***	0.104***
	(0.026)		(0.026)	(0.026)
Unanticipated inflation (overall)		0.118***		
		(0.011)		
Number of trips			0.044***	-0.087***
			(0.004)	(0.005)
Number of distinct stores				0.883***
				(0.016)
Observations	3,439,497	3,439,497	3,439,497	3,439,497
Household fixed effect	√	√	√	√
Household varying effect	√	√	√	√

Notes: The table reports the estimates in specification (48). Each observation is at the household×quarter level covering from 2004 Q1 to 2019 Q4. The coefficient represents the corresponding change in different measures of search behavior after a standard deviation increase in unanticipated inflation. Consumer fixed and time-varying effects are controlled. Standard errors are clustered at the household level. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

where t is time; i represents household. λ_i is the household fixed effect. $\tilde{\pi}_t$ is the unanticipated inflation. β is the coefficient of interest. It measures the magnitude of the correlation between the unanticipated inflation and the consumers' search behavior. y_{it+1} is the non-routine share of spending in the next quarter. I use next-period value for two reasons. First, it avoids reverse causality because inflation and consumer search behavior are jointly determined in theory. Second, it may take time for consumers to change their shopping habits. X_{it} is the time-varying household controls.[29] As pointed out by Aguiar and Hurst (2007), these variables have large affect on the pattern of shopping behavior.

Table 7 reports the estimates. Column (1) shows that a one–standard deviation (60 bp) increase in unanticipated food and beverage inflation raises the non-routine spending share by 12.2 bp. This implies consumers respond to higher prices by intensifying search and shifting purchases away from their routine stores; the implied elasticity is about 20% (12.2/60). Column (2) uses unanticipated overall inflation and finds a comparable effect: a 49 bp increase is associated with an 11.8 bp rise in the non-routine share—roughly 24% (11.8/49). This similarity is noteworthy given that the NielsenIQ data primarily cover food and beverages, where one might expect stronger sensitivity to category-specific inflation. A plausible interpretation is that higher overall inflation erodes real wages for sampled households, prompting broader substitution toward cheaper outlets and goods. The third and fourth columns introduce controls for the number of shopping trips

[29]These controls include household age, employment, education, marital status, gender if single, race, having children or not, and household i's total spending at time t.

and the number of distinct stores visited. The coefficients decrease only slightly, suggesting these variables capture a small portion of search effort. This indicates that consumers allocate most of their search efforts to substituting products within the same categories across the stores they have already visited on the same trip, rather than increasing trips or visiting new stores.

The estimated coefficient may be biased downward for several reasons. First, consumers only record purchases from stores included in the NielsenIQ dataset, which predominantly covers large retail stores. As a result, our measure may not fully capture the non-routine share of spending if consumers switch to stores not included in the dataset or to online purchases. Second, substantial substitution within a product group could contribute to the bias. For example, consumers may trade down to lower-quality goods within the same store (Jaimovich et al., 2019) after an increase in inflation. Finally, the inflation shock may not be entirely passed on to retail prices, as suppliers may absorb part of the shock in their wholesale costs. This incomplete passthrough could further dampen the observed relationship between inflation and non-routine spending. To find passthrough of inflation shocks to retail prices is beyond the scope of this paper.

Overall, these two pieces of evidence support key mechanisms: (i) more attention to inflation increases the slope of Phillips curve and (ii) as prices rise after an unanticipated inflation shock, consumers are incentivized to search for alternatives.

7 Conclusion

This paper develops a new framework for understanding monetary non-neutrality, driven entirely by consumer-side frictions. At the heart of the model is the information asymmetry about nominal marginal costs between consumers and firms. Specifically, the framework integrates a heterogeneous firm block and incomplete consumer information into a standard sequential search model. When consumers observe a price increase, they attribute it to adverse productivity shocks rather than increases in nominal wages, prompting them to search for outside options. Firms, in turn, internalize this consumer search behavior, limiting the passthrough of cost changes to prices. The framework also accommodates aggregate supply shocks, providing a toolbox for analyzing a wide range of shocks as in the standard New Keynesian model.

I further present a dynamic general equilibrium model that incorporates both habit formation and wage rigidity. To match the extremely gradual adjustment in household inflation expectations observed in the data, the estimated model features significant information frictions. These frictions amplify monetary non-neutrality and generate much dampened inflation responses. The key determinant of information friction in estimation is the relative impulse responses of CPI and cumulative inflation expectation. For aggregate productivity shocks, the inflation path in the model displays a sharp initial increase followed by a rapid decline in contrast with the predictions of standard Calvo- or menu cost-based New Keynesian models. Overall, the model successfully

replicates key empirical features of impulse responses to both monetary and productivity shocks. Future research can investigate the effect of fiscal policy in this model.

This paper also provides preliminary empirical validation of the model's key mechanisms. I show heightened inflation attention can amplify the passthrough of headline inflation fluctuations and the effect of labor market tightness. Additionally, using large-scale consumer panel data, I document that unanticipated inflation significantly increases consumers' search intensity, consistent with the model's predictions. Future research can investigate the passthrough of inflation shocks to retail prices and the pricing complementarity between retailers

Several further topics of inquiry are left for future research. First, although this paper focuses on final goods markets, the framework can be extended to any market with many sellers and buyers, such as upstream and downstream firms in supply chains. An interesting extension would be embedding this model into production network models. Second, applying the model to the labor market could yield valuable insights. Workers' incomplete information about the average posted wage could influence their job-search decisions as documented in the recent literature. Finally, to better calibrate the model and assess the quantitative importance of the mechanism, incorporating both firm- and consumer-side frictions into the model is necessary.

Bibliography

Afrouzi, H., Blanco, A., Drenik, A., and Hurst, E. (2024). A theory of how workers keep up with inflation. Working Paper 33233, National Bureau of Economic Research.

Aguiar, M. and Hurst, E. (2007). Life-cycle prices and production. *American Economic Review*, 97(5):1533–1559.

Amiti, M., Itskhoki, O., and Konings, J. (2014). Importers, exporters, and exchange rate disconnect. *American Economic Review*, 104(7):1942–1978.

Amiti, M., Itskhoki, O., and Konings, J. (2019). International shocks, variable markups, and domestic prices. *The Review of Economic Studies*, 86(6):2356–2402.

Anderson, S. P., Palma, A. D., and Thisse, J.-F. (1987). The ces is a discrete choice model? *Economics Letters*, 24(2):139–140.

Anderson, S. P. and Renault, R. (1999). Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model. *The RAND Journal of Economics*, 30(4):719–735.

Angeletos, G.-M. and Huo, Z. (2021). Myopia and anchoring. *American Economic Review*, 111(4):1166–1200.

Angeletos, G.-M. and La'O, J. (2013). Sentiments. *Econometrica*, 81(2):739–779.

Angeletos, G.-M. and La'o, J. (2010). Noisy business cycles. *NBER Macroeconomics Annual*, 24(1):319–378.

Angeletos, G.-M. and Lian, C. (2023). Dampening general equilibrium: Incomplete information and bounded rationality. In Handbook of Economic Expectations, V. ., editor, *Handbook of Economic Expectations*, pages 613–645. Elsevier.

Auclert, A., Bardóczy, B., Rognlie, M., and Straub, L. (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, 89(5):2375–2408.

Auclert, A., Rigato, R., Rognlie, M., and Straub, L. (2024). New pricing models, same old phillips curves? *The Quarterly Journal of Economics*, 139(1):121–186.

Auclert, A., Rognlie, M., and Straub, L. (2020). Micro jumps, macro humps: Monetary policy and business cycles in an estimated hank model. Technical Report 26647, National Bureau of Economic Research.

Ball, L. M., Leigh, D., and Mishra, P. (2022). Understanding u.s. inflation during the covid era. *Brookings Papers on Economic Activity*, 53(2):1–54.

Baqaee, D. R. and Burstein, A. (2023). Welfare and output with income effects and taste shocks. *The Quarterly Journal of Economics*, 138(2):769–834.

Baqaee, D. R. and Farhi, E. (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics*, 135(1):105–163.

Baqaee, D. R., Farhi, E., and Sangani, K. (2024). The supply-side effects of monetary policy. *Journal of Political Economy*, 132(4).

Barnichon, R. (2010). Building a composite help-wanted index. *Economics Letters*, 109(3):175–178.

Barnichon, R. and Shapiro, A. H. (2024). Phillips meets beveridge. *Journal of Monetary Economics*, 148(Supplement):n/a.

Bauer, M. D. and Swanson, E. T. (2023). A reassessment of monetary policy surprises and high-frequency identification. *NBER Macroeconomics Annual*, 37(1):87–155.

Beaudry, P., Hou, C., and Portier, F. (2024). The dominant role of expectations and broad-based supply shocks in driving inflation. Working Paper 32322, National Bureau of Economic Research.

Begenau, J., Farboodi, M., and Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, 97:71–87.

Benedek, D., De Mooij, R. A., Keen, M., and Wingender, P. (2015). Estimating vat pass through. IMF Working Paper 15/214, International Monetary Fund.

Benigno, P. and Eggertsson, G. B. (2023). It's baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve. Working Paper 31197, National Bureau of Economic Research.

Bianchi, F., Faccini, R., and Melosi, L. (2023). A fiscal theory of persistent inflation. *The Quarterly Journal of Economics*, 138(4):2127–2179.

Binetti, A., Nuzzi, F., and Stantcheva, S. (2024). People's understanding of inflation. Working Paper 32497, National Bureau of Economic Research. Revision Date June 2024.

Broda, C. and Weinstein, D. E. (2010). Product creation and destruction: Evidence and price implications. *American Economic Review*, 100(3):691–723.

Buettner, T. and Madzharova, B. (2021). Unit sales and price effects of preannounced consumption tax reforms: Micro-level evidence from european vat. *American Economic Journal: Economic Policy*, 13(3):103–134.

Burdett, K. and Judd, K. L. (1983). Equilibrium price dispersion. *Econometrica*, 51(4):955–969.

Bénabou, R. and Gertner, R. (1993). Search with learning from prices: Does increased inflationary uncertainty lead to higher markups? *The Review of Economic Studies*, 60(1):69–93.

Candia, B., Coibion, O., and Gorodnichenko, Y. (2023). The macroeconomic expectations of firms. In Manski, C. F. and Thomas, F., editors, *Handbook of Economic Expectations*, pages 321–353. Elsevier, Amsterdam.

Cavallo, A., Cruces, G., and Perez-Truglia, R. (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics*, 9(3):1–35.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, volume 1, Part A of *Handbook of Macroeconomics*, pages 65–148. Elsevier, Amsterdam.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Christiano, L. J., Eichenbaum, M. S., and Trabandt, M. (2016). Unemployment and business cycles. *Econometrica*, 84(4):1523–1569.

Coibion, O. and Gorodnichenko, Y. (2015). Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1):197–232.

Coibion, O., Gorodnichenko, Y., and Ropele, T. (2020). Inflation expectations and firm decisions: New causal evidence. *Quarterly Journal of Economics*, 135(1):165–219.

Coibion, O., Gorodnichenko, Y., and Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6):1579–1617.

D'Acunto, F., Malmendier, U., Ospina, J., and Weber, M. (2021). Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, 129(5):1615–1639.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2020). Changing business dynamism and productivity: Shocks versus responsiveness. *American Economic Review*, 110(12):3952–3990.

Del Negro, M., Giannoni, M. P., and Patterson, C. (2024). The forward guidance puzzle. *American Economic Journal: Macroeconomics*, 16(2):1–38.

DellaVigna, S. and Gentzkow, M. (2019). Uniform pricing in u.s. retail chains. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Diamond, P. A. (1971). A model of price adjustment. *Journal of Economic Theory*, 3(2):156–168.

D'Acunto, F., Hoang, D., Paloviita, M., and Weber, M. (2023). Iq, expectations, and choice. *The Review of Economic Studies*, 90(5):2292–2325.

D'Acunto, F., Malmendier, U., and Weber, M. (2022). Gender roles produce divergent economic expectations. *Proceedings of the National Academy of Sciences*, 119(32):e2204252119.

Erceg, C. J., Henderson, D. W., and Levin, A. T. (2000). Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics*, 46(2):281–313.

Eyster, E., Madarasz, K., and Michaillat, P. (2015). Preferences for fair prices, cursed inferences, and the nonneutrality of money. CEP Discussion Paper dp1325, Centre for Economic Performance, London School of Economics.

Farboodi, M. and Veldkamp, L. (2021). A model of the data economy. NBER Working Paper 28427, National Bureau of Economic Research, Cambridge, MA. Revised June 2022.

Farhi, E. and Werning, I. (2019). Monetary policy, bounded rationality, and incomplete markets. *American Economic Review*, 109(11):3887–3928.

Friedman, M. (1968). The role of monetary policy. *The American Economic Review*, 58(1):1–17.

Gaballo, G. and Paciello, L. (2021). Spending allocation under nominal uncertainty: a model of effective price rigidity. CEPR Discussion Paper DP16101, Centre for Economic Policy Research (CEPR).

Gagliardone, L. and Gertler, M. (2023). Oil prices, monetary policy and inflation surges. Working Paper 31263, National Bureau of Economic Research.

Gagliardone, L., Lenzu, S., and Tielens, J. (2025). Anatomy of the phillips curve: Micro evidence and macro implications. *American Economic Review*. Forthcoming.

García-Schmidt, M. and Woodford, M. (2019). Are low interest rates deflationary? a paradox of perfect-foresight analysis. *American Economic Review*, 109(1):86–120.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Goldberg, P. K. and Hellerstein, R. (2013). A structural approach to identifying the sources of local currency price stability. *The Review of Economic Studies*, 80(1):175–210.

Gopinath, G., Gourinchas, P.-O., Hsieh, C.-T., and Li, N. (2011). International prices, costs, and markup differences. *American Economic Review*, 101(6):2450–2486. †.

Hazell, J. (2024). Comment on "the dominant role of expectations and broad based supply shocks in driving inflation". In Leahy, J. V., Eichenbaum, M. S., and Ramey, V. A., editors, *NBER Macroeconomics Annual 2024, Volume 39*. University of Chicago Press. This is a preliminary draft and may not have been subjected to the formal review process of the NBER. This page will be updated as the chapter is revised.

Hazell, J., Herreño, J., Nakamura, E., and Steinsson, J. (2022). The slope of the phillips curve: Evidence from u.s. states. *Quarterly Journal of Economics*, 137(3):1299–1344.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Hulten, C. R. (1973). Divisia index numbers. *Econometrica*, 41(6):1017–1025.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of Economic Studies*, 45(3):511–518.

Jaimovich, N., Rebelo, S., and Wong, A. (2019). Trading down and the business cycle. *Journal of Monetary Economics*, 102:96–121.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics*, 134(2):715–783.

Kaplan, G. and Menzio, G. (2015). The morphology of price dispersion. *International Economic Review*, 56(4):1207–1238.

Kimball, M. S. (1995). The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking*, 27(4, Part 2):1241–1277.

Klenow, P. J. and Willis, J. L. (2016). Real rigidities and nominal price changes. *Economica*, 83(331):443–472.

Kleshchelski, I. and Vincent, N. (2009). Market share and price rigidity. *Journal of Monetary Economics*, 56(3):344–352.

Korenok, O., Munro, D., and Chen, J. (2023). Inflation and attention thresholds. *Review of Economics and Statistics*, pages 1–28.

Kumar, S., Afrouzi, H., Coibion, O., and Gorodnichenko, Y. (2015). Inflation targeting does not anchor inflation expectations: Evidence from firms in new zealand. NBER Working Paper 21814, National Bureau of Economic Research.

L'Huillier, J.-P. (2020). Consumer imperfect information and endogenous price rigidity. *American Economic Journal: Macroeconomics*, 12(2):94–123.

Link, S., Peichl, A., Roth, C., and Wohlfart, J. (2023). Attention to the macroeconomy. `https://ssrn.com/abstract=4669798` or `http://dx.doi.org/10.2139/ssrn.4669798`. SSRN Working Paper.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory*, 4(2):103–124.

Maćkowiak, B. and Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803.

Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328.

Matějka, F. (2015). Rigid pricing and rationally inattentive consumer. *Journal of Economic Theory*, 158:656–678. Part B.

Mongey, S. and Waugh, M. E. (2024). Discrete choice, complete markets, and equilibrium. Staff Report 656, Federal Reserve Bank of Minneapolis.

Morris, S. and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5):1521–1534.

Nakamura, E. and Steinsson, J. (2011). Price setting in forward-looking customer markets. *Journal of Monetary Economics*, 58(3):220–233.

Nakamura, E. and Zerom, D. (2010). Accounting for incomplete pass-through. *The Review of Economic Studies*, 77(3):1192–1230.

Pfäuti, O. (2024). The inflation attention threshold and inflation surges. Technical Report 2024-02, EMPCT Working Paper Series, University of Texas at Austin. First version: August 2023; Revised August 2024.

Phelps, E. S. (1967). Phillips curves, expectations of inflation, and optimal unemployment over time. *Economica*, 34(135):254–281.

Phelps, E. S. (1969). The new microeconomics in inflation and employment theory. *American Economic Review*, 59(2):147–160.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2 of *Handbook of Macroeconomics*, pages 71–162. Elsevier, Amsterdam.

Rebelo, S., Santana, M., and Teles, P. (2024). Behavioral sticky prices. Working Paper 32214, National Bureau of Economic Research. Revised October 2024.

Romer, C. D. and Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084.

Rosolia, A. (2024). Do firms act on their inflation expectations? another look at italian firms. *Journal of Political Economy Macroeconomics*, 2(4).

Rotemberg, J. J. (1982). Sticky prices in the united states. *Journal of Political Economy*, 90(6):1187–1211.

Rotemberg, J. J. (2005). Customer anger at price increases, changes in the frequency of price adjustment, and monetary policy. *Journal of Monetary Economics*, 52(4):829–852.

Rotemberg, J. J. (2010). Altruistic dynamic pricing with customer regret. *Scandinavian Journal of Economics*, 112(4):646–672.

Rotemberg, J. J. (2011). Fair pricing. *Journal of the European Economic Association*, 9(5):952–981. `https://www.jstor.org/stable/41305841`.

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

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.

Stantcheva, S. (2024). Why do we dislike inflation? *Brookings Papers on Economic Activity*, 2024(Spring). Forthcoming or published in Spring 2024 issue.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Weber, M., Candia, B., Afrouzi, H., Ropele, T., Lluberas, R., Frache, S., Meyer, B. H., Kumar, S., Gorodnichenko, Y., Georgarakos, D., Coibion, O., Kenny, G., and Ponce, J. (2025). Tell me something i don't already know: Learning in low- and high-inflation settings. *Econometrica*, 93(1):229–264.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica*, 47(3):641–654.

Werning, I. (2023). Expectations and the rate of inflation. NBER Working Paper 31428, National Bureau of Economic Research.

Wolinsky, A. (1986). True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, 101(3):493–511.

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

Yotzov, I., Bloom, N., Bunn, P., Mizen, P., and Thwaites, G. (2024). The speed of firm response to inflation. NBER Working Paper 32731, National Bureau of Economic Research.

A Omitted Proofs in Static Model

Proof of Proposition 1:

Proof. We consider the case where the first draw does not need the search cost. Shoppers are randomly assigned a firm for free in the first round of search. This guarantees that they always participate the market. I prove that for given x, if $\int U(v|x)\psi(y|x)dy = -\infty$, then $v^*(x) = -\infty$; if $\int U(v|x)\psi(y|x)dy > -\infty$, there exists a unique threshold defined as follows.

$$v^*(x) = -\frac{\kappa}{1-\Psi(v^*(x)|x)} + \frac{\int_{v^*(x)}^{\infty} y\psi(y)dy}{1-\Psi(v^*(x)|x)} \tag{A.1}$$

First, given x, I consider two limits. When $v \to -\infty$,

$$-\kappa + U(v|x)\int_{-\infty}^{v}\psi(y|x)dy + \int_{v}^{\infty}U(y|x)\psi(y|x)dy = -\kappa + \int U(y|x)\psi(y|x)dy$$

Then, if $\int U(v|x)\psi(y|x)dy > -\infty$, $U(v|x) > v$. Intuitively, when the agent has the worst value whatsoever, i.e., $v \to -\infty$, the agent should choose to continue searching. In the other limit where $v \to \infty$,

$$-\kappa + U(v|x)\int_{-\infty}^{v}\psi(y|x)dy + \int_{v}^{\infty}U(y|x)\psi(y|x)dy = -\kappa + U(v|x)$$

Then, $U(v|x) = v$. Intuitively, when the agent has the best value whatsoever, the agent should stop search because she cannot get any other good that offers better value.

Second, I claim that if $U(v|x) > v$, then $U'(v|x) = 0$. To prove the claim, suppose there exist values of v on its support (at least at $-\infty$) such that $U(v) > v$. Then, $U(v)$ satisfies,

$$U(v|x) = -\kappa + U(v|x)\int_{-\infty}^{v}\psi(y|x)dy + \int_{v}^{\infty}U(y|x)\psi(y|x)dy \tag{A.2}$$

Equivalently,

$$U(v|x) = -\frac{\kappa}{1-\Psi(v|x)} + \frac{\int_{v}^{\infty}U(y|x)\psi(y|x)dy}{1-\Psi(v|x)} \tag{A.3}$$

It is easy to show that

$$U'(v|x) = \frac{\psi(v|x)}{(1-\Psi(v|x))^2}\left\{\int_{v}^{\infty}U(y|x)\psi(y|x)dy - U(v|x)(1-\Psi(v|x)) - \kappa\right\} = 0 \tag{A.4}$$

The second equation holds due to (A.3). Then, it is easy to see that $U(v|x)$ is a constant in $(-\infty, v^*]$, which is equal to $v^*(x)$, and then $U(v) = v$ in (v^*, ∞). v^* is unique. In addition, for $v < v^*$,

$$-\kappa + U(v|x)\int_{-\infty}^{v}\psi(y|x)dy + \int_{v}^{\infty}U(y|x)\psi(y|x)dy = -\kappa + v^*(x)\Psi(v^*(x)) + \int_{v^*(x)}^{\infty}y\psi(y|x)dy = v^*(x) \tag{A.5}$$

The search problem in (10) is therefore simplified to

$$U(v|x) = \max\{v, v^*(x)\} \tag{A.6}$$

This implies that the value of an additional search does not depend on the state v. No matter what state the shopper has, she always compare the state with $v^*(x)$.

Now, we prove the second part: $v^*(x)$ decreases in x. First, notice that from (A.5), the threshold can be written alternatively by the following,

$$\int \int_{\lambda(v^*(x)+p)}^{\infty} (\frac{1}{\lambda}\epsilon - p - v^*(x))g(\epsilon)d\epsilon f(p|x)dp = \kappa \tag{A.7}$$

For $x_1 < x_2$, $f(p|x_1)$ is FOSD over $f(p|x_2)$. Since the inner integral in LHS decreases in p, that implies

$$v^*(x_1) > v^*(x_2)$$

We consider the case in which $\int U(v|x)\psi(y|x)dy = -\infty$ for any x. This happens when firms charge arbitrarily high prices, which implies $\psi(y|x)$ is a Dirac function at $y = -\infty$. Then, it is optimal to accept any firm in the first round and the resulting $v^*(x) = -\infty, \forall x$. This is the case considered in Diamond (1971). ■

Proof of Proposition 2:

Proof. We take first-order condition of firm's profit in (15) with respect to P_k. It is easy to show that we can always express the pricing strategy as in (16) and define the elasticity of demand as:

$$e_k = -\frac{\partial \log D(P_k)}{\partial \log P_k} = -\frac{\partial \log \Big(\int X(1 - G(\lambda(v^*(x) + p_k)))d\Phi_x(x) \Big)}{\partial p_k} + 1 \tag{A.8}$$

■

Proof of Theorem 1: I first prove a lemma that is useful for proving Theorem 1.

Lemma A-1. *The full-information threshold $v^*(w)$ decreases with the search cost, κ.*

Proof. See Appendix A. ■

The Lemma establishes a one-to-one correspondence between the full-information threshold and the search cost. A lower search cost directly raises the value of additional search. But at the same time, it reduces this value indirectly by increasing prices through (20). The Lemma show that the direct effect outweighs the indirect effect.

Proof. We first define the markup elasticity,

$$\Lambda_k = -\frac{d \log \mu_k}{dp_k} = \frac{1}{e_k(e_k - 1)} \frac{\partial e_k}{\partial p_k}$$

Since G is log-concave, $\frac{\partial e_k}{\partial p_k} > 0$. Therefore, the markup elasticity is positive. We can further write out the definition,

$$\Lambda_k = \frac{\frac{g'}{g}\frac{1-G}{g} + 1}{\frac{1}{\lambda}\frac{1-G}{g} + 1}$$

Here, I omit the argument of functions for simplicity. The argument is $\lambda(v^*(w) + p_k)$. Note that full-information equilibrium price is a function of $v^*(w)$ and a_k. I now rewrite the LHS of (A.7) in terms of the integral over productivity distribution, which is exogenous,

$$\int \int_{\lambda(v^*(w)+p^*(a;v^*(w)))}^{\infty} (\frac{1}{\lambda}\epsilon - p^*(a; v^*(w)) - v^*(w))g(\epsilon)d\epsilon\phi_a(a)da = \kappa \tag{A.9}$$

where $\phi_a(a)$ is the pdf of productivity distribution. Fix $p^*(a; v^*(w))$, the LHS is decreasing in $v^*(w)$. However, higher $v^*(w)$ also decreases optimal prices. To know the net effect of these two forces, we need to derive $\frac{\partial p^*(a_k;v^*(w))}{\partial v^*(w)}$. From now on, we use $x_k(v^*(w))$ to denote $x(v^*(w), a_k)$ for any variable x.

$$\frac{\partial p_k(v^*(w))}{\partial v^*(w)} = \frac{\partial \log \mu_k(v^*(w))}{\partial v^*(w)} = -\frac{1}{e_k(v^*(w))(e_k(v^*(w)) - 1)}\frac{\partial e_k(v^*(w))}{\partial v^*(w)} \tag{A.10}$$

It is easy to show that

$$\frac{\partial e_k(v^*(w))}{\partial v^*(w)} = \lambda^2\frac{g'(1-G) + g^2}{(1-G)^2}\left(1 + \frac{\partial p_k(v^*(w))}{\partial v^*(w)}\right) \tag{A.11}$$

Combine the above two equations, we have:

$$\frac{\partial p_k(v^*(w))}{\partial v^*(w)} = -\frac{\Lambda_k}{1 + \Lambda_k} \tag{A.12}$$

Since $\Lambda_k > 0$, $\frac{\partial p_k(v^*(w))}{\partial v^*(w)} \in (-1, 0)$. This implies that $\frac{\partial(v^*(w)+p_k(v^*(w)))}{\partial v^*(w)} \in (0, 1)$. Therefore, the LHS of (A.7) decreases in $v^*(w)$. ■

Now, I prove Theorem 1.

Proof. First, notice that since G is log-concave, μ_k decreases with P_k. From $P_k = \mu_k\frac{W}{A_k}$, we know that there is a unique solution for each optimal price P_k given the full-information threshold $v^*(w)$. Then, from Lemma A-1, we know that there exists a unique solution $v^*(w)$ to A.7 given that the prices are computed through the first-order conditions as in (20). The equilibrium in which shoppers search actively exists and is unique because $v^*(w)$ exists and is unique. ■

Remarks on Computation of the Full-information Equilibrium:

Remark 1 (Remark on Computing Steady-State Equilibrium)**.** *The proof of Theorem 1 provides insights on the computational method for the full-information equilibrium.*

1. *First guess a $v^*(w)$*

2. *Calculate the optimal price distribution given $v^*(w)$*

3. *Plug guessed $v^*(w)$ and derived price distribution and plug it into 18 and check if LHS is equal to the given search cost κ.*

4. *Increases guessed $v^*(w)$ if LHS is larger than search cost, according to Lemma A-1. Vice versa.*

5. *Loop the procedure 1-4 until the difference between the LHS and the RHS of 18 is smaller than the given tolerance*

Proof of Theorem 2:

Proof. Suppose the nominal wage increases from w to w'. Notice $\Delta w = w' - w$ does not need to be small. We guess that all the optimal prices increase proportionally, i.e., $p'_k = p_k + \Delta w$. Then the price distribution $f(p|w)$ shifts to the right and becomes $f(p - \Delta w|w')$. The threshold is determined by,

$$\int \int_{\lambda(v^*(w')+p)}^{\infty} (\frac{1}{\lambda}\epsilon - p - v^*(w'))g(\epsilon)d\epsilon f(p - \Delta w|w')dp = \kappa \tag{A.13}$$

Let $z = p - \Delta w$. Then we can rewrite the LHS of the above equation,

$$\int \int_{\lambda(v^*(w')+\Delta w+z)}^{\infty} (\frac{1}{\lambda}\epsilon - z - (\Delta w + v^*(w')))g(\epsilon)d\epsilon f(z|w')dz = \kappa \tag{A.14}$$

This implies that $v^*(w') = v^*(w) - \Delta w$. According to (20), the elasticity of demand is given by,

$$\begin{aligned} e'_k &= \lambda \frac{g(\lambda(v^*(w') + p'_k))}{1 - G(\lambda(v^*(w') + p'_k))} + 1 \\ &= \lambda \frac{g(\lambda(v^*(w) - \Delta w + p_k + \Delta w))}{1 - G(\lambda(v^*(w) - \Delta w + p_k + \Delta w))} + 1 \\ &= e_k \end{aligned}$$

Since the elasticity of demand does not change, the optimal prices is given by,

$$p'_k = \log(\frac{e_k}{e_k - 1}) + w' - a_k = p_k + w' - w = p_k + \Delta w \tag{A.15}$$

Therefore, we verify the guess that the optimal prices increase proportionally with the nominal wage. Based on Theorem 1, the equilibrium is unique. Therefore, the above constructed equilibrium is the only equilibrium when the nominal wage is w'. ■

Proof of Proposition 3: Here, I provide separate proofs for two parts of Proposition 3. I start with the proof of the first part.

Proof of Part 1. First, we can rewrite 11 in terms of integrating over the exogenous productivity distribution and posterior belief about nominal wage,

$$\int\int\int_{\lambda(v^*(x)+p^*(a,\hat{w}))}^{\infty}(\frac{1}{\lambda}\epsilon - p^*(a,\hat{w}) - v^*(x))g(\epsilon)d\epsilon\phi_a(a)dah(\hat{w}|x)d\hat{w} = \kappa \tag{A.16}$$

where $v^*(x)$ is implicitly determined by the above equation. On the first order, the posterior belief of the nominal wage collapses to a Dirac function at $E(w|x)$. Also, combining the passthrough $\varphi(a)$, the above equation becomes,

$$\int\int_{\lambda(v^*(x)+\bar{p}_k+\varphi(a)E(\hat{w}|x))}^{\infty}(\frac{1}{\lambda}\epsilon - \bar{p}(a) - \varphi(a)E(\hat{w}|x) - v^*(x))g(\epsilon)d\epsilon\phi_a(a)da = \kappa \tag{A.17}$$

Take the total derivative on $E(\hat{w}|x)$ on both sides,

$$\int\Big(\varphi(a) + \frac{dv^*(x)}{dE(\hat{w}|x)}\Big)(1 - G(\lambda(v^*(x) + \bar{p}(a) + \varphi(a)E(\hat{w}|x))))\phi_a(a)da = 0$$

On the first order, it can be written as follows,

$$\frac{dv^*(x)}{dE(\hat{w}|x))} = -\int\varphi(a)\bar{\omega}(a)\phi_a(a)da = -\int\varphi_k\bar{\omega}_k dk \tag{A.18}$$

where $\bar{\omega}(a)$ is the expenditure share of firms with productivity a in the full-information steady state. Then, the threshold, on the first order, is given by,

$$v^*(x) = v^*(\bar{w}) - \frac{\partial v^*(x)}{\partial v^*(x)}v^*(x) = v^*(\bar{w}) - \Phi E(\hat{w}|x)$$

Recall $\hat{p} = \Phi\hat{w}$. We have the result. ■

Proof of Part 2 . First, plug the result in Part 1 into the elasticity of demand in (17), the elasticity becomes,

$$e_k = \lambda\frac{\int Xg(\lambda(v^*(\bar{w}) - \Phi E(\hat{w}|x) + p_k))d\Phi_x(x)}{\int X(1 - G(\lambda(v^*(\bar{w}) - \Phi E(\hat{w}|x) + p_k)))d\Phi_x(x)} + 1 \tag{A.19}$$

with some abuse of notation, Φ is the aggregate correlated-shock passthrough and $\Phi_x(x)$ is the cdf of information sets. Let $y = \lambda\Phi(E(\hat{w}|x) - \bar{E}(\hat{w}))$ and its pdf is $\phi_y(y)$, which is a Gaussian distribution with mean zero and standard deviation $\sigma = \lambda\Phi\theta\sigma_s$, where σ_s^2 is the variance of distribution of $E(\hat{w}|x) - \bar{E}(\hat{w})$ and is proportional to $\sigma_{\hat{w}}^2$. Further, we denote $z = \lambda(v^*(w) - \Phi\bar{E}(\hat{w}) + p_k)$. One should think of z as mean and y is the error caused by idiosyncratic noises in private signals. The elasticity is rewritten as follows,

$$e_k = \lambda\frac{\int Xg(z+y)\phi_y(y)dy}{\int X(1 - G(z+y))\phi_y(y)dy} + 1 \tag{A.20}$$

It is easy to show that the first-order approximation to the ratio is equivalent to separately approximating numerator and denominator and then combining them. Following this result, we first expand the numerator.

$$g(z+y) = g(z) + g'(z)y + \frac{g''(z)}{2}y^2 + \mathcal{O}(y^3)$$

Substitute this expansion into the integral and also notice that $X \propto \exp(y + \bar{w})$,

$$\begin{aligned}\int \exp(y)g(z+y)\phi_y(y)dy &= \int (1+y)\Big(g(z) + g'(z)y + \frac{g''(z)}{2}y^2 + \mathcal{O}(y^3)\Big)\phi_y(y)dy \\ &= g(z) + \int \Big((g'(z) + \frac{g''(z)}{2})y^2 + \mathcal{O}(y^3)\Big)\phi_y(y)dy \\ &= g(z) + (g'(z) + \frac{g''(z)}{2})\sigma^2 + \mathcal{O}(\sigma^3)\end{aligned}$$

Therefore, $\int g(z+y)\phi_y(y)dy \to g(z)$ on the order of σ_w^2, which is second-order term. Similarly, $\int (1 - G(z+y))\phi_y(y)dy \to 1 - G(z)$ on the order of σ_w^2. On the other hand, the average expectation of monetary shock $\bar{E}(\hat{w})$ approaches zero on the order of σ_w. Therefore, on the first order, the noises coming from the idiosyncratic noises converge to zero faster than the mean. As a result, elasticity, on the first order, is given by,

$$e_k = \lambda \frac{g(z)}{1 - G(z)} + 1 = \lambda \frac{g(\lambda(v^*(w) - \Phi\bar{E}(\hat{w}) + p_k))}{1 - G(\lambda(v^*(w) - \Phi\bar{E}(\hat{w}) + p_k))} + 1 \tag{A.21}$$

■

Proof of Lemma 1:

Proof. First, log price is given by,

$$p_k = \log \mu_k + w - a_k$$

Total differentiate on both sides,

$$\hat{p}_k = \frac{\partial \log \mu_k}{\partial p_k}\hat{p}_k + \frac{\partial \log \mu_k}{\partial p}\hat{p} + \hat{w} \tag{A.22}$$

We further define $\gamma_k = (1 - \frac{d\log\mu_k}{dp_k}\Big|_{\hat{w}=0})^{-1}$ and $\xi_k = \frac{d\log\mu_k}{dp}\Big|_{\hat{w}=0}\gamma_k$. Then, we have,

$$\hat{p}_k = \gamma_k\hat{w} + \xi_k\hat{p} \tag{A.23}$$

Since $\hat{p} = \Phi\hat{w}$, the correlated-shock passthrough for individual firm satisfies,

$$\varphi_k = \gamma_k + \Phi\xi_k$$

Integrate on both sides,

$$\Phi = \frac{\Gamma}{1 - \Xi}$$

where $\Gamma = \int \gamma_k\bar{\omega}_k dk$, $\Xi = \int \xi_k\bar{\omega}_k dk$. ■

Proof of Theorem 3: I first prove the "Incompleteness" part of the Theorem. Then I prove an important lemma and finally I prove the comparative statics results.

Proof. Notice that for any function $f(w)$, the first-order approximation to its derivative is,

$$f'(w) = f'(\bar{w}) + f''(\bar{w})\hat{w} \tag{A.24}$$

Then, $f'(w)|_{\hat{w}=0} = f'(\bar{w})$ on the first order. I claim that the approximation gives the same result if we reverse the order of the operations. Let's first take the first-order approximation to $f(w)$,

$$f(w) = f(\bar{w}) + f'(\bar{w})\hat{w} \tag{A.25}$$

Then if we take derivative w.r.t. $\hat{w}$ and then make $\hat{w} = 0$, we get the same result.

Now, following this result, $\frac{\partial e_k}{\partial p_k}\Big|_{\hat{w}=0}$ and $\frac{\partial e_k}{\partial p}\Big|_{\hat{w}=0}$ can be computed by first first-order approximating e_k as we did in Proposition 3 and then take derivative to $\hat{p}_k$ and $\hat{p}$. For simplicity, I omit the argument of the functions, in particular, $F = F(v^*(\bar{w}) + \bar{p}_k)$ for any function F. It is straightforward to show that, on the first order,

$$\frac{\partial e_k}{\partial p_k}\Big|_{\hat{w}=0} = \lambda^2 \frac{g'(1-G) + g^2}{(1-G)^2} \tag{A.26}$$

We can derive that γ_k does not depend on θ. In addition, we have

$$\frac{\partial e_k}{\partial p}\Big|_{\hat{w}=0} = -\theta\lambda^2 \frac{g'(1-G) + g^2}{(1-G)^2} \tag{A.27}$$

Then, we have $\frac{\partial \log \mu_k}{\partial p} = -\theta \frac{\partial \log \mu_k}{\partial p_k}$. According to Lemma 1, we have

$$\xi_k = \theta(1-\gamma_k) \tag{A.28}$$

Substitute back to Lemma 1,

$$\hat{p}_k = \gamma_k(\hat{w} - \hat{a}_k) + \theta(1-\gamma_k)\hat{p} \tag{A.29}$$

Integrate on both sides,

$$\hat{p} = \Gamma\hat{w} + \theta(1-\Gamma)\hat{p} \tag{A.30}$$

Or equivalently,

$$\hat{p} = \Gamma\hat{w} + (1-\Gamma)E^c\hat{p} \tag{A.31}$$

■

Proof of Corollary 1:

Proof. The aggregation in (A.30) requires every firm is individually rational, so they know (A.29). Suppose every shopper believe that firms and other shoppers are rational. Then every shopper believes (A.30) holds. Shoppers' average expectation of $\hat{p}$ satisfies,

$$\bar{E}\hat{p} = \Gamma\bar{E}\hat{w} + (1-\Gamma)\bar{E}^2\hat{p} \tag{A.32}$$

where $\bar{E}^2[\cdot] = \bar{E}[\bar{E}[\cdot]]$ denotes the second-order belief. Iterating ad infinitum, the change in the actual price index $\hat{p}$ can be expressed in terms of the higher-order beliefs of the monetary shock $\hat{w}$:

$$\hat{p} = \Gamma\sum_{h=0}^{\infty}(1-\Gamma)^h\bar{E}^h\hat{w} \tag{A.33}$$

"Iterating ad infinitum" amounts to imposing common knowledge of rationality. The first iteration requires that shoppers know that firms and other shoppers are rational, the second iteration requires that shoppers know that others know they are rational and firms are rational, and so on. ■

Proof of Proposition 4 and 5: Before proving Proposition 4 and 5, I first state and prove the following lemma.

Lemma A-2. *Let* $\Lambda_k = -\frac{d\log\mu_k}{dp_k}\Big|_{\hat{w}=0}$ *denote the steady-state markup elasticity. When G is Gumbel distribution, it has the following property,*

$$\frac{d\Lambda(y)}{dy} < 0$$

In addition, $\Lambda(y)\to 0$ *as* $y\to\infty$ *and* $\Lambda(y)\to\infty$ *as* $y\to-\infty$.

Proof. To simplify notation, I denote $y_k = \lambda(v^*(\bar{w}) + \bar{p}_k)$. I assume that G follows standard type-I extreme-value distribution, i.e., $G(y) = \exp(\exp(-y))$, $g(y) = \exp(-y-\exp(-y))$. It also follows that $\frac{g'}{g} = e^{-y} - 1$.

First, since G is log-concave, g is also log-concave. This implies the following property,

$$(\frac{g'}{g})' < 0; \quad (\frac{1-G}{g})' < 0$$

Additionally, using G is Gumbel, I have $\frac{1-G}{g}\to 1$ as $y\to\infty$ and therefore $\frac{1-G}{g} > 1$.

Second, we write out markup elasticity,

$$\Lambda = \frac{\frac{g'}{g}\frac{1-G}{g}+1}{\frac{1}{\lambda}\frac{1-G}{g}+1} \tag{A.34}$$

Then the derivative of markup elasticity w.r.t y is given by,

$$\begin{aligned}\frac{d\Lambda(y)}{dy} &= \left(\frac{g'}{g}\right)'\left(\frac{1-G}{g}\right)^2\frac{1}{\lambda} - \left(\frac{1-G}{g}\right)'\frac{1}{\lambda} + \left(\frac{g'}{g}\frac{1-G}{g}\right)' \\ &= -e^{-y}\left(\frac{1-G}{g}\right)^2\frac{1}{\lambda} + \frac{1}{\lambda}(1+\frac{g'}{g}\frac{1-G}{g}) + \left(\frac{g'}{g}\frac{1-G}{g}\right)' \\ &< -e^{-y}\left(\frac{1-G}{g}\right)\frac{1}{\lambda} + \frac{1}{\lambda}(1+(e^{-y}-1)\frac{1-G}{g}) + \left(\frac{g'}{g}\frac{1-G}{g}\right)' \\ &= \frac{1}{\lambda}\left(1-\frac{1-G}{g}\right) + \left(\frac{g'}{g}\frac{1-G}{g}\right)'\end{aligned}$$

The inequality is due to $\frac{1-G}{g} > 1$. The first term of the last equation is negative. We need to deal with the second term. Notice that

$$\left(\frac{g'}{g}\frac{1-G}{g}\right)' = \left(\frac{g'}{g}\right)'\left(\frac{1-G}{g}\right) + \left(\frac{g'}{g}\right)\left(\frac{1-G}{g}\right)'$$

It is straightforward that when $y < 0$, $\frac{g'}{g} = e^{-y} - 1 > 0$. Then, $\left(\frac{g'}{g}\frac{1-G}{g}\right)' < 0$. We now focus on the case where $y > 0$.

$$\begin{aligned}\left(\frac{g'}{g}\frac{1-G}{g}\right)' &= -e^{-y}\frac{1-G}{g} + (1-e^{-y}) - (1-e^{-y})^2\frac{1-G}{g} \\ &< -e^{-y}\frac{1-G}{g} + (1-e^{-y}) - (1-e^{-y})\frac{1-G}{g} \\ &= (1-e^{-y}) - \frac{1-G}{g} \\ &< 0\end{aligned}$$

The first inequality is due to $1 - e^{-y} \in (0,1)$ when $y > 0$. The second inequality is due to $\frac{1-G}{g} > 1$. Combine all together, we have

$$\frac{d\Lambda(y)}{dy} < 0$$

For limit results, first recall $\frac{1-G}{g} \to 1$ as $y \to \infty$. Plug into (A.34), we get $\Lambda \to 0$ when $y \to \infty$. Second, since $\frac{1-G}{g} \to \infty$ when $y \to -\infty$, $\Lambda \approx \lambda g'/g \to \infty$ since $g'/g \to \infty$. ■

Next, we turn to the result on comparative statics.

Proof of Proposition 4: First, notice that to show φ_k decreases in a_k, it is sufficient to show that γ_k decreases in a_k. From Lemma 1, $\gamma_k = (1+\Lambda_k)^{-1}$. Since Λ_k decreases in P_k based on Lemma A-2, it implies γ_k increases in P_k. Since P_k decreases in a_k, it implies γ_k decreases in a_k.

Since the markup μ_k decreases in p_k, it increases in a_k. It is easy to show that $\lim_{a\to-\infty} e_k = 1 - \lambda\frac{g'}{g} = 1 + \lambda$ and $\lim_{a\to\infty} e_k = 1$. Then, $\lim_{a\to-\infty}\mu_k = \frac{\lambda+1}{\lambda}$ and $\lim_{a\to\infty}\mu_k = \infty$.

Also, we know $\xi_k = \theta(1-\gamma_k)$, so ξ_k increases in a_k. Also, in the limit when $a \to \infty$, $p \to -\infty$. Then $\Lambda \to \infty$ and $\gamma \to 0$, $\xi \to \theta$. In the opposite limit, $p \to \infty$, and $\Lambda \to 0$, implying $\gamma \to 1$ and $\xi \to 0$. ■

Proof of Proposition 5: I first consider the comparative statics on κ. From Lemma A-1, an increase in the search cost κ implies lower threshold $v^*(\bar{w})$. Since the steady-state markup elasticity decreases in $v^*(\bar{w})$, γ_k decreases for all k. As a result, φ_k decreases in κ for all k. When $\kappa \to 0$, $v^*(\bar{w}) \to \infty$ and therefore $\Lambda_k \to 0$, resulting in $\gamma_k \to 1$ for any k. This implies $\varphi_k = 1$ for any k. When $\kappa \to \infty$, $v^*(\bar{w}) \to -\infty$, and therefore $\Lambda_k \to \infty$, resulting in $\gamma_k \to 0$ for any k. This implies $\varphi_k = 0$ for any k. Then, I consider the comparative statics on θ. An increase in the information friction, which means a decrease in θ, induces lower φ_k. When $\theta \to 0$, $\Phi = \Gamma$. When $\theta \to 1$, $\varphi_k = 1$. ■

Proof of Proposition 6:

Proof. The profit of a firm that has N products indexed by j is given by,

$$\pi_k = \frac{1}{\rho} \sum_{j=1}^{N} \int X(1 - G(\lambda(v^*(x) + p_k))) d\Phi_x(x) \frac{1}{P_{jk}} (P_{jk} - \frac{W}{A_{jk}}) (\frac{P_{jk}}{P_k})^{-\sigma} \tag{A.35}$$

where $P_k = (\sum_{j=1}^{N} P_{jk}^{1-\sigma})^{\frac{1}{1-\sigma}}$. Denote the total expenditure share on firm k as $\omega_k = \omega(P_k)$, then,

$$\pi_k = \frac{1}{\rho} \sum_{j=1}^{N} \omega(P_k) \frac{1}{P_{jk}} (P_{jk} - \frac{W}{A_{jk}}) (\frac{P_{jk}}{P_k})^{-\sigma} \tag{A.36}$$

Firm k will internalize the effect of change in P_{jk} on overall price index P_k of the firm. The first-order condition w.r.t. P_{jk} is given by,

$$P_{jk} = \mu_k \frac{W}{A_{jk}} = \frac{\sigma}{\sigma - 1 - P_k^{-1} \tilde{\pi}_k (e_k - 1 + \sigma - 1)} \frac{W}{A_{jk}} \tag{A.37}$$

where e_k is the same as in (17). Notice that the markup is the same across all the products within the firm. Also, we have $\tilde{\pi}_k = (1 - \mu_k) P_k$. Plug in the above equation, we have,

$$\mu_k = \frac{e_k}{e_k - 1} \tag{A.38}$$

The passthrough of productivity shock to a given product j is proportional to the expenditure share of this product over all products on the market.

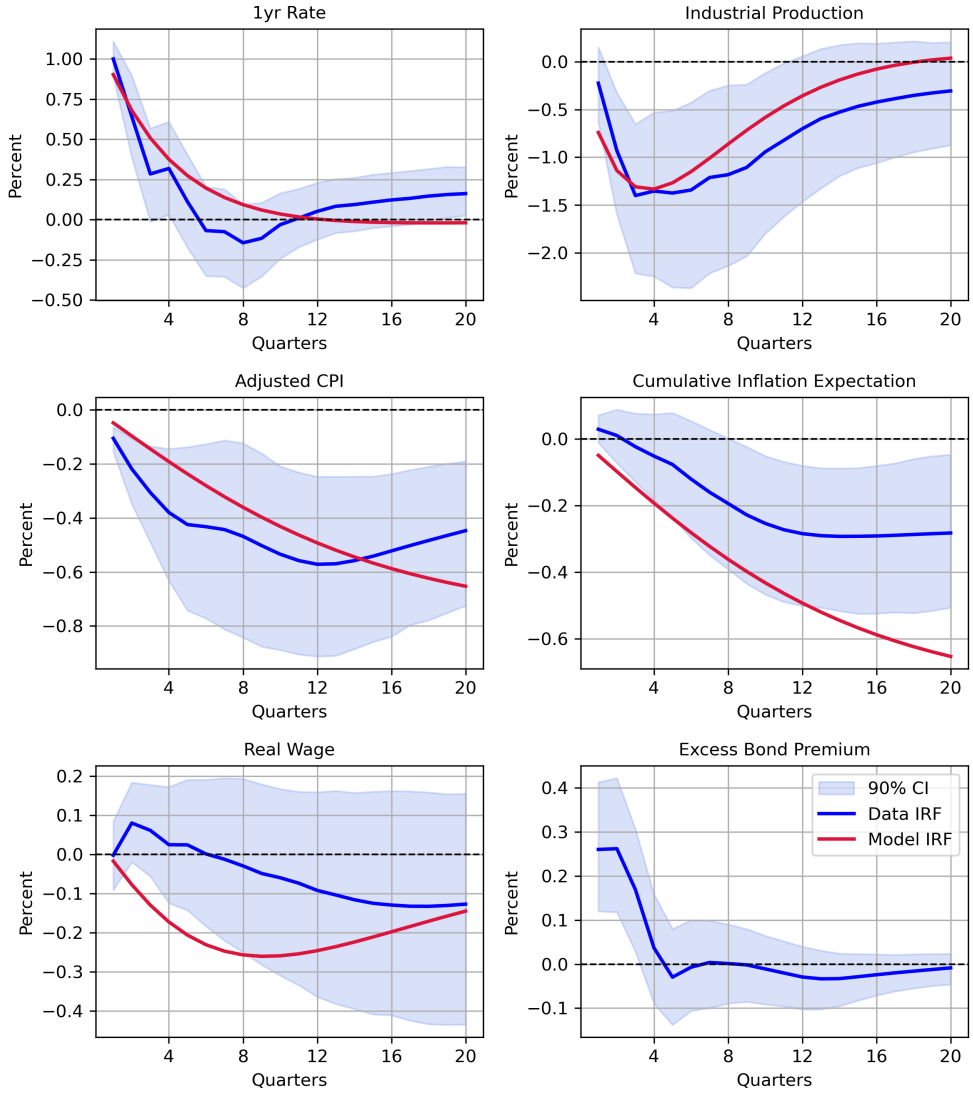
$$\gamma_{jk} = \frac{\omega_{jk}}{\omega_k} \gamma_k \tag{A.39}$$

Therefore, when $N = 1$, $\gamma_{jk} = \gamma_k$; when $N \to \infty$, $\gamma_{jk} \to 0$. ■

B Additional Details on Calibration and Estimation

Estimated Calvo model: I present the estimated Calvo model where the wage rigidity and the persistence of the monetary policy shock are fixed at the values reported in Table 1. The estimated Calvo parameter is 0.9. Figure B.1 shows the result. The model yields a good fit for the inflation and output responses. However, it fails to capture the response of inflation expectations.

Figure B.1: Empirical and Calvo Model Impulse Responses



Notes: The figure plots the impulse responses for 4 variables following a 100 bp shock on the 1-year government bond rate. The sample period is 1978M1–2019M12. The blue lines represent the empirical impulse responses and the red lines represent the Calvo model counterparts. The shaded areas are 90% confidence intervals derived from wild bootstrap. The red line represents the impulse responses implied in Section 4. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

Details of Section 5.4: Without information frictions, the Matrix $\mathcal{M}^y$ is given by,

$$\mathbf{M}^y = (1-\beta)\begin{bmatrix} \beta^0 & \beta^1 & \beta^2 & \cdots \\ \beta^0 & \beta^1 & \beta^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \beta^0 & \beta^1 & \beta^2 & \cdots \end{bmatrix}$$

Figure B.2: Monetary Transmission Channels in the model with Workers' Information Friction



Notes: For all three panels, the solid line plots the baseline impulse responses at the estimated parameters in Table 4. In Panel A, the dashed line represents impulse responses when $\theta = 1$, holding all other parameters at their estimated values. In Panel B, the dashed line represents impulse responses when $\theta_w = 1$. In Panel C, the dashed line represents impulse responses when $\lambda_w = 0$.

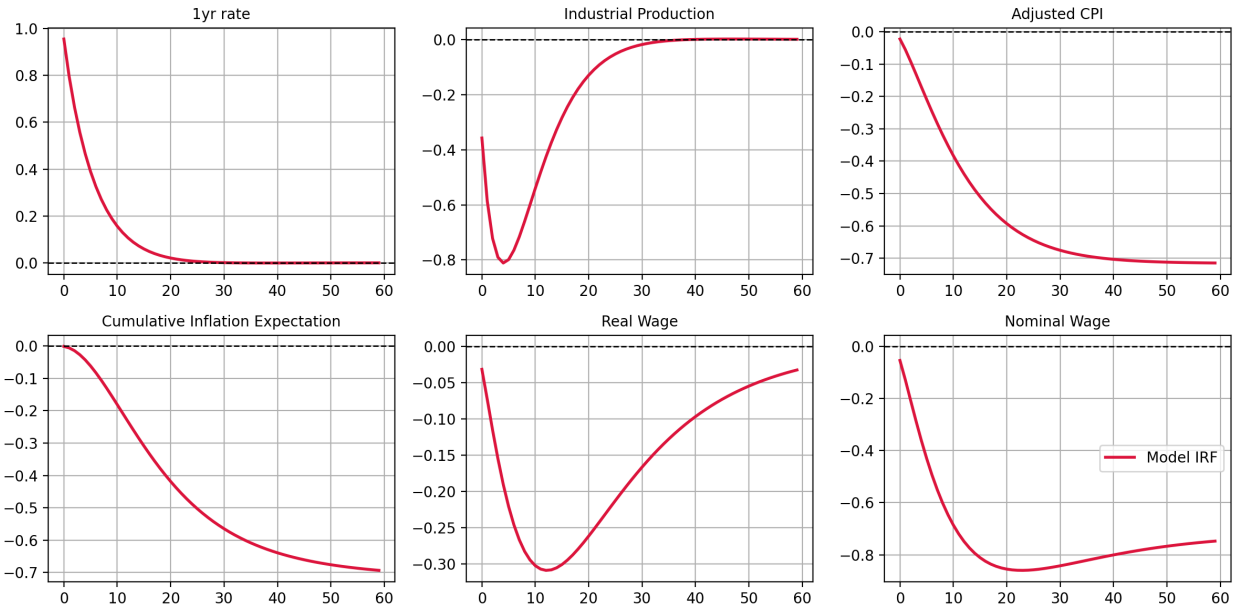
Similarly, $\mathbf{M}^r$ can be derived using the property of the forward guidance in the standard Euler equation, which is given by,

$$\mathbf{M}^r = \left(\mathbf{I} - (1-\beta) \begin{bmatrix} \beta^0 & \beta^1 & \dots & \beta^{T-1} \\ \beta^0 & \beta^1 & \dots & \beta^{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta^0 & \beta^1 & \dots & \beta^{T-1} \end{bmatrix} \right) \cdot \begin{bmatrix} -\frac{1}{\sigma} & -\frac{1}{\sigma} & \dots & -\frac{1}{\sigma} \\ 0 & -\frac{1}{\sigma} & \dots & -\frac{1}{\sigma} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -\frac{1}{\sigma} \end{bmatrix}$$

Following Auclert et al. (2020), I modify the matrices to incorporate information frictions. Figure B.2 shows the counterfactuals when we shut down each of the frictions. In particular, Panel A shows the experiment where I shut down the information friction regarding the CPI. Panel B shows the case where the information friction in the Euler equation is shut down. Figure B.3 shows the longer-horizon impulse responses up to 60 quarters. The CPI response is 20 bps larger than the baseline as the persistence of monetary shock is much higher.

Other Robustness Checks – Figure B.4 shows the first robustness check where I use the raw CPI instead of adjusted CPI. The estimated persistence now increases to 0.82, which is in the ballpark of the estimates in the literature, and the information friction is 0.07, implying the average duration of information friction is 14 quarters.

Figure B.3: Longer-Horizon Model Impulse Responses



Notes: The figure plots the impulse responses for the longer horizons. The x-axis denotes quarters from the shock, starting at 0. The y-axis denotes percent.

Figure B.5 shows the second robustness check where I use the monthly Brave-Butters-Kelley real GDP instead of the Industrial Production Index. The response of GDP index is much smaller than IP index, inducing a even smaller persistence of monetary shock, which is estimated to be 0.61. The wage stickiness is also smaller, 0.83, as the shock is less pronounced. However, the estimated information friction is 0.053, almost the same as the result of the simple back-of-the-envelop calculation.

Overall, although the persistence of monetary shock and the degree of wage stickiness vary across specifications, the estimate of information friction is robust. The average duration of information friction is between 14 quarters to 18 quarters.

Details of Section 5.5: In this section, I specify the dynamic general equilibrium model for aggregate productivity shocks. Monetary authority's, worker's and labor union's problems are the same.

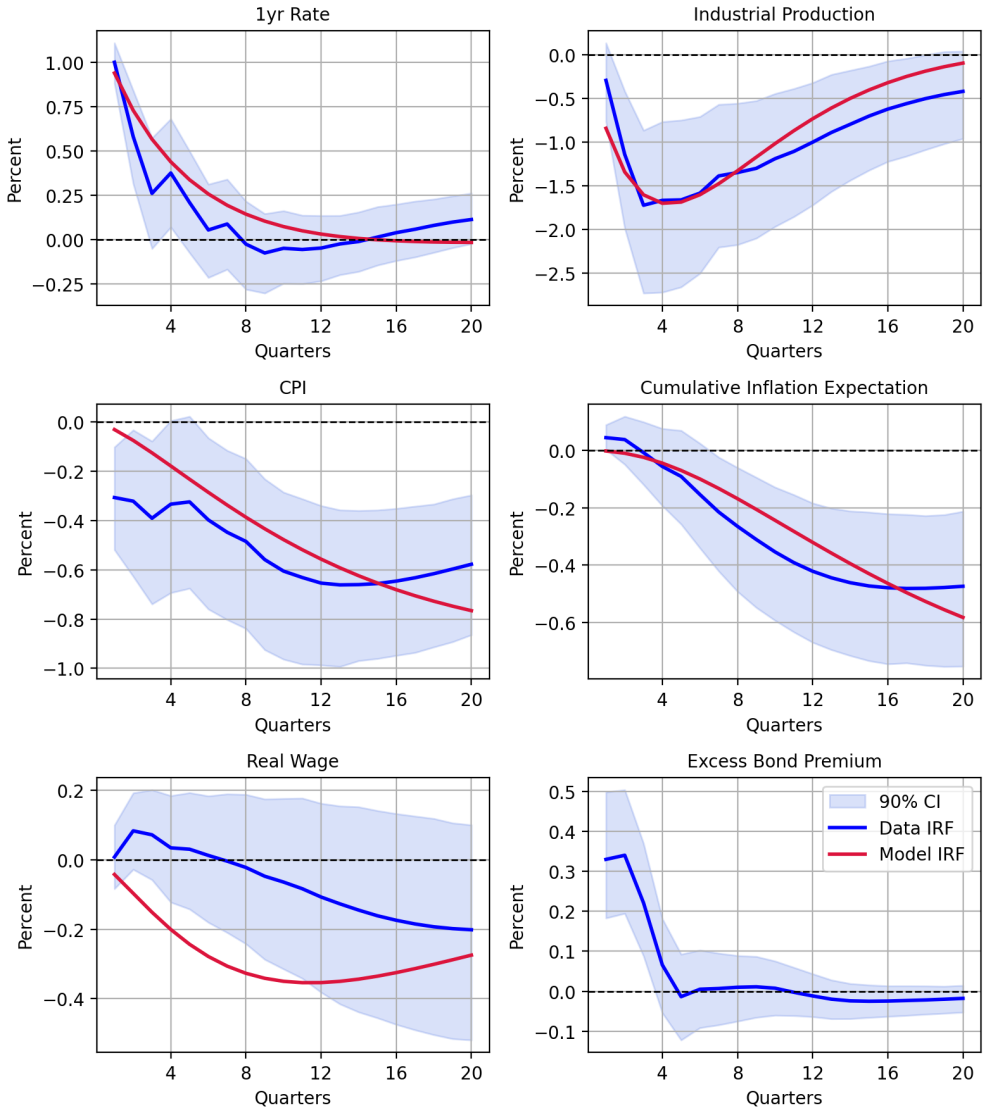
Firm – The firm's productivity A_{kt} is composed by two components.

$$\log A_{kt} = \log A_t + \log B_{kt} \tag{B.1}$$

where the aggregate productivity follows AR(1) process,

$$\log A_t = \rho_a \log A_t + \sigma_a \varepsilon_{at} \tag{B.2}$$

Figure B.4: Robustness: CPI



Notes: I use the raw CPI instead of adjusted CPI.

where aggregate productivity shocks follow $\varepsilon_{at} \sim \mathcal{N}(0,1)$; ρ_a is the persistence.

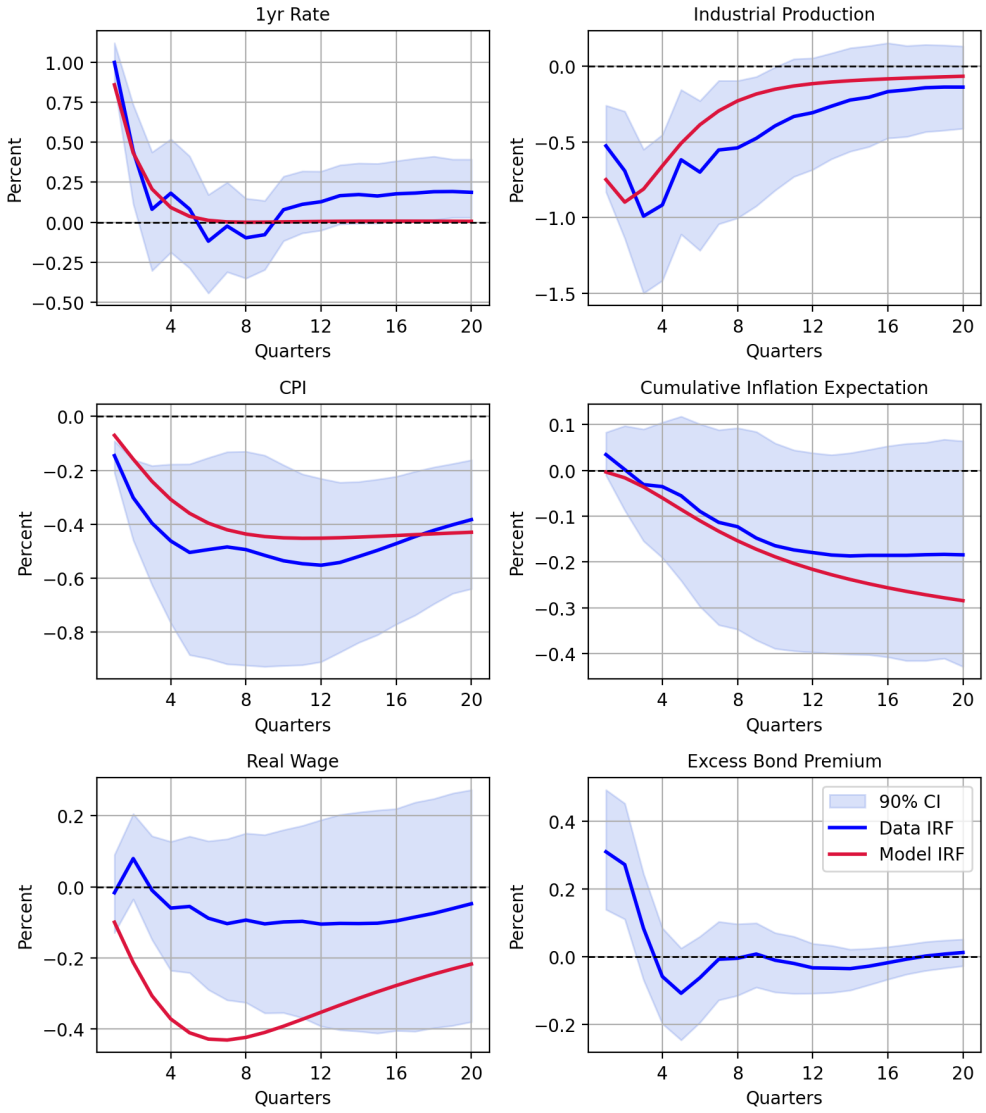
$$\log B_{kt} = \rho_b \log B_{kt-1} + \sigma_b \varepsilon_{akt} \tag{B.3}$$

where idiosyncratic productivity shocks follow $\varepsilon_{akt} \sim \mathcal{N}(0,1)$; ρ_b is the persistence.

Shopper – With productivity shocks, the nominal marginal cost is changed. The Pricing equation can be re-written as,

$$\hat{\boldsymbol{p}} = \Gamma(\boldsymbol{I} - (1-\Gamma)\boldsymbol{\Theta})^{-1}(\widehat{\boldsymbol{w}} - \widehat{\boldsymbol{a}}) \tag{B.4}$$

Figure B.5: Robustness: GDP



Notes: I use the Brave-Butters-Kelley Real Gross Domestic Product Index instead of the Industrial Production Index.

The generalized Phillips curve is written as,

$$\hat{\boldsymbol{\pi}} = \underbrace{(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{\Psi}\,(\boldsymbol{I} - \boldsymbol{\Psi})^{-1}}_{\boldsymbol{K}}(\hat{\boldsymbol{w}}^r - \hat{\boldsymbol{a}}) \tag{B.5}$$

Reallocation and Aggregate Productivity – The aggregate production function is:

$$\hat{y}_t = \hat{l}_t - \hat{a}_t + (\hat{v}_t - \hat{p}_t) \tag{B.6}$$

Following the same method as computing $\hat{p}_t$, $\hat{v}_t$ is given by,

$$\hat{\boldsymbol{v}} = \underbrace{\tilde{\Gamma}(\boldsymbol{I} - (1-\tilde{\Gamma})\boldsymbol{\Theta})^{-1}}_{\tilde{\boldsymbol{\Psi}}}(\hat{\boldsymbol{w}} - \hat{\boldsymbol{a}}) \tag{B.7}$$

where $\tilde{\Gamma} = \int \bar{\alpha}_k \gamma_k dk$ and $\tilde{\Gamma} > \Gamma$. Then, the aggregate production function is given by,

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{l}} - \hat{\boldsymbol{a}} + (\tilde{\boldsymbol{\Psi}} - \boldsymbol{\Psi})(\hat{\boldsymbol{w}} - \hat{\boldsymbol{a}}) \tag{B.8}$$

Equilibrium – In the limit of small monetary shock, the equilibrium dynamics of variables $\{\hat{\pi}_t, \hat{p}_t, \hat{w}^r_t, \hat{w}_t, \hat{v}_t, \hat{c}_t, \hat{i}_t\}$ is governed by equations (37), (41), (B.8), (34), (B.7) and two equations that connect nominal and real variables: $\hat{w}_t = \hat{w}^r_t + \hat{p}_t$ and $\hat{\pi}_t = \hat{p}_t - \hat{p}_{t-1}$.

C Additional Results on Empirics

Additional Results of Section 6.1: Table C.1 reports the robustness results. Columns (1) and (2) replicate the baseline estimates. Columns (3) and (4) add nonlinear terms of the V/U ratio. The coefficients on the interaction between the attention index and V/U remain unchanged, while the nonlinear terms themselves are statistically insignificant. This indicates that nonlinearities in slack cannot account for the result, reinforcing the interpretation that it is driven by attention to inflation rather than nonlinear effects of labor shortages. Columns (5) and (6) present results using an alternative attention threshold of 3.21% as in Pfäuti (2024). The findings are broadly similar, with one exception: in the pre-pandemic period, supply shocks have a negative impact on core inflation. This likely reflects the role of monetary policy actively responding to these shocks.

Details of Section 6.2 I show the secular decline in both non-routine share of spending and shopping time. To complement the analysis based on NielsenIQ Consumer Panel, I also use data from the 2003–2019 waves of the American Time Use Survey (ATUS) conducted by BLS. The ATUS sample is drawn from individuals exiting the Current Population Survey (CPS). Each wave is based on 24-hour time diaries, where respondents report their activities from the previous day in detailed time intervals. The ATUS includes over 400 detailed time-use categories, providing rich insights into daily activity patterns. Specifically, I use the time each respondent reports in the category "obtaining goods" as a proxy for shopping time.[30] Figure C.1 shows that the non-routine share of spending and average shopping time are trending down from early 2000s and stay relatively stable in 2010s. Both time series are highly correlated. This evidence boosts our confidence of the relevance of our measure of search activity. Different from ATUS, the NielsenIQ datasets link spending and the search measure.

[30] I do not use the ATUS in the previous section because the data lacks a panel structure. Household characteristics are very important in explaining shopping patterns (Aguiar and Hurst, 2007).

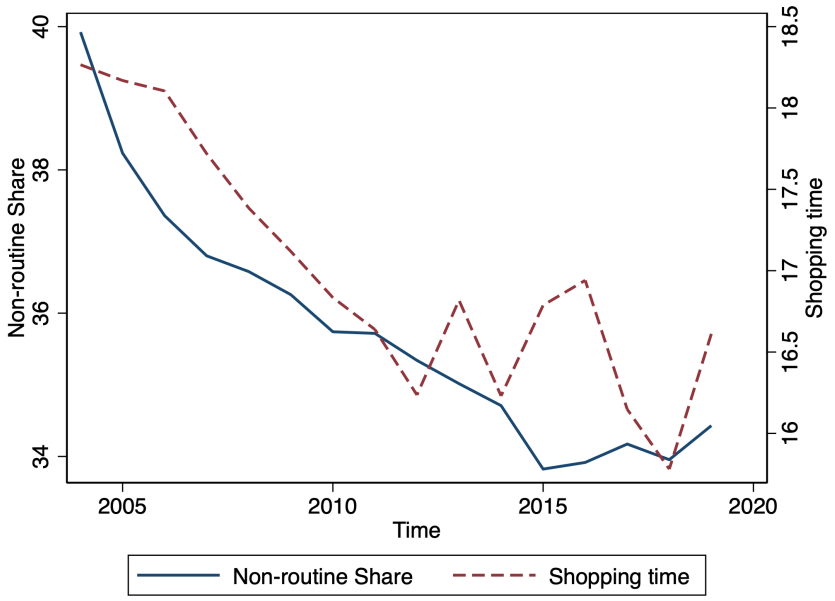
Table C.1: The Effect of Attention to Inflation on the Regional Phillips Curve

	(1)	(2)	(3)	(4)	(5)	(6)
	1983–2019	1983–2025	1983–2019	1983–2025	1983–2019	1983–2025
V/U	0.302*	0.604***	0.072	0.362***	0.261*	0.482***
	(0.155)	(0.137)	(0.265)	(0.117)	(0.155)	(0.122)
V/U^2			0.090	0.122		
			(0.141)	(0.123)		
V/U^3			0.199	0.021		
			(0.183)	(0.037)		
Supply	-0.136	-0.114	-0.120	-0.133	-0.200**	-0.160
	(0.102)	(0.110)	(0.101)	(0.099)	(0.099)	(0.107)
Attention	0.305**	0.227	0.349**	0.267**	0.147	0.187
	(0.132)	(0.152)	(0.136)	(0.135)	(0.126)	(0.124)
Attention $\times$ V/U	0.219	0.414***	0.271	0.312***	0.213	0.523***
	(0.176)	(0.123)	(0.198)	(0.093)	(0.178)	(0.102)
Attention $\times$ Supply	0.321**	0.593***	0.312**	0.619***	0.434***	0.600***
	(0.128)	(0.142)	(0.126)	(0.148)	(0.115)	(0.124)
Constant	0.049	0.181	-0.016	0.024	0.035	0.145
	(0.107)	(0.117)	(0.132)	(0.120)	(0.110)	(0.114)
R-squared	0.125	0.523	0.138	0.564	0.125	0.522
Adj. R-sq	0.115	0.518	0.124	0.558	0.115	0.517
Observations	433	497	433	497	433	497

Notes: V/U denotes ratio of vacancies to unemployed (12 month average). U denotes unemployment gap (12 month average). Supply represents the headline inflation shocks. Attention is the attention index. Newey-West standard errors with 12 lags are reported. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

Lastly, I show that the other two measures of search activities also increase significantly after unanticipated inflation shock. Similar to the main results, the responses of measures to inflation shocks to food and beverage are much larger than shocks to overall inflation.

Figure C.1: Time Trend of Non-Routine Share of Spending and Shopping Time



Notes: The figure plots the average non-routine share of spending and shopping time over time. I use the American Time Use Survey (2003-2019) to compute the average shopping time.

Table C.2: Other Measures of Search Activities and Unanticipated Inflation

Dep. var.:	#Trips	#Trips	#Store	#Store
Unanticipated inflation (F&D)	0.095***		0.034***	
	(0.012)		(0.029)	
Unanticipated inflation (overall)		0.035***		0.020***
		(0.005)		(0.001)
Observations	3,439,497	3,439,497	3,439,497	3,439,497
Household fixed effect	√	√	√	√
Household varying effect	√	√	√	√

Notes: The table reports the estimates in specification (48) using number of trips and distince retailers. Each observation is at the household×quarter level covering from 2004 Q1 to 2019 Q4. The coefficient represents the corresponding change in different measures of search behavior after a standard deviation increase in unanticipated inflation. Consumer fixed and time-varying effects are controlled. Standard errors are clustered at the household level. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.