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

Xiaojie Liu†

January 21, 2025

Click here to view the latest version

Abstract

This paper develops a model of monetary non-neutrality driven by information asymmetry between consumers and firms about nominal marginal costs in a sequential search framework. With only consumer-side frictions, this approach is distinguished from the standard one that relies on firm-side pricing frictions. Consumers' value of search is determined by their information about the price index, and firm's elasticity of demand depends on the perceived relative price. The passthrough of aggregate shocks to prices is therefore incomplete. 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. I further show that the output gap is proportional to the nowcast error of price index in the Phillips curve. Despite its parsimonious nature, the calibrated dynamic general equilibrium model can generate substantial monetary non-neutrality. Consistent with the mechanism, higher inflation is associated empirically with more active consumer search.

Keywords: Monetary non-neutrality, Phillips curve, Search, Information frictions, Information asymmetry, Passthrough, Higher-order beliefs

*I am grateful for comments and discussions with Alireza Tahbaz-Salehi, Lawrence Christiano, Sara Moreira, George-Marios Angeletos, Benjamin Jones, Alessandro Pavan, Harry Pei, Diego Känzig, Chen Lian and Huo Zhen, Ryungha Oh, Diego Huerta.

†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 known to have large real effects on the economy in the short run. Both output and inflation decline following an unexpected increase in interest rate. This pattern has been repeatedly uncovered in the empirical literature.[1] Most of existing theories explain this phenomenon by focusing on frictions on the firm side. Models of price stickiness posit that price adjustments are infrequent due to either exogenous factors (Taylor, 1980; Calvo, 1983) or fixed costs (Mankiw, 1985; Golosov and Lucas, 2007). Another theory which dates at least back to Phelps (1969) and Lucas (1972) suggests that firms set prices based on incomplete information about aggregate shocks. Reis (2006) and Alvarez et al. (2016) further argue that costs of acquiring and processing information contribute to price rigidity.

Growing evidence suggests that consumers face more severe frictions than firms. For instance, consumers have misperception about inflation (Binetti et al., 2024; Candia et al., 2023) and pay particular attention to salient prices (Kumar et al., 2015; D'Acunto et al., 2021). Moreover, Kaplan and Menzio (2015) and Kaplan et al. (2019) document a large price dispersion for an identical product even within the same market and week. This firms' market power reveals consumer-side frictions that prevent them from finding the cheapest product. These evidence motivates two central questions: Can we micro-found monetary non-neutrality using only the consumer-side frictions? How do different consumer-side frictions affect the transmission of monetary policy?

To address these questions, I develop a new monetary model that places consumer-side frictions at the center. Consumers face two frictions: (i) information friction about firms' nominal marginal costs and (ii) search frictions on the goods market. On the other hand, firms have full information about the model economy and set the prices flexibly. The main mechanism operates as follows. Following a positive monetary shock, the nominal marginal costs increase. Firms tend to increase prices. However, shoppers with incomplete information about the costs attribute much of this price increase to firm-specific adverse shocks. Marginal shoppers who are initially indifferent between

[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.

purchasing and searching are now incentivized to seek outside options. To dissuade search, firms compress markups, thereby limiting the passthrough of the monetary shock.

To show the core mechanism, I start with a static partial equilibrium model. The model features shoppers, firms, and a monetary authority. The monetary authority sets the nominal wage. A monetary shock is modeled as a shock to the nominal wage. The shopper's problem follows the sequential search literature (Wolinsky, 1986; Anderson and Renault, 1999). Specifically, shoppers search sequentially and randomly with free recall. Shoppers must incur a search cost to visit a firm and learn the price. I extend the framework from the literature on consumer search in two ways. First, firms are heterogeneous in productivity. Second, shoppers have incomplete information about the nominal wage. These two features lead to shoppers' rational confusion between the aggregate (nominal wage) and idiosyncratic (firm productivity) components of the cost. In particular, when observing a price, shoppers who believe the nominal wage is lower will interpret the price as indicative of encountering a firm with lower productivity and vice versa.

More specifically, each shopper receives a signal about the nominal wage before shopping. The search decision hinges on the value of an additional search, which in turn depends on each shopper's perceived price distribution. Shoppers construct their own perceived price distributions using posterior belief of the nominal wage and the knowledge imposed by the rational expectations equilibrium. The optimal search strategy is characterized by a threshold rule: if the value of a good exceeds the threshold, the shopper proceeds with the purchase and stop searching. This threshold decreases in the signal, since a lower perceived nominal wage implies a higher likelihood of receiving lower price quote.

I begin by characterizing the full-information rational-expectation equilibrium, where the only friction is consumers' search costs. I demonstrate that, under full information about the nominal wage, monetary policy is neutral. The intuition is that although shoppers lack knowledge of individual firms' prices—due to the limitations imposed by search frictions—they understand that, in equilibrium, all prices fully adjust to changes in the nominal wage. As a result, relative prices remain unchanged, and search decisions are not affected. The key takeaway of this result is that the search friction alone is not sufficient to generate monetary non-neutrality. We need the interplay of these two frictions. The situation changes in the incomplete-information rational-expectation equilibrium. I show that, to the first-order approximation, the elasticity of demand increases with the perceived relative price, defined as the ratio between the actual firm's price and the average expectation of price index. Due to the information friction, the average expectation of the change in price index is typically dampened compared to the actual one. As a result, firms behave as if they are competing against other firms that set lower prices. In response, firms lower their prices, which compresses the markup and limits the passthrough of the shock. The intuition is very similar to Lucas (1972) where agents confuse the increase in nominal prices with the increase in relative prices, expect that here this mechanism is operated completely on the household side.

Next, I characterize the aggregate total passthrough, which is the passthrough from the shock to the price index. The total passthrough is composed of two fundamental passthroughs that can be derived from the optimal pricing strategy in the firm's problem (Amiti et al., 2019): passthrough of marginal cost shocks (own-cost passthrough) and passthrough of competitors' prices (cross-price passthrough). I then present the main theorem of this paper: (i) the aggregate total passthrough is generically incomplete and (ii) high-productivity firms contribute more to the incompleteness, and (iii) the aggregate total passthrough decreases with both frictions, with one friction amplifies the effect of the other.

To gain more insights behind the main mechanism, Consider an extreme case where shoppers have no information about the nominal wage. In this scenario, a nominal wage shock resembles an idiosyncratic shock to firms, and the aggregate total passthrough reduces to the aggregate own-cost passthrough. When shoppers possess some information, firms start to respond to other firms' prices perceived by shoppers. The aggregate response of firms can be understood as occurring in iterative rounds. In each round, firms incorporate part of the additional change in the price index from the previous round into their prices, with the passthrough attenuated by shoppers' information frictions. Iterating ad infinitum, the aggregate total passthrough ultimately represents the infinite sum of these rounds of cross-price passthroughs, with each successive round increasingly attenuated, leading to the incomplete passthrough.

I close the model in general equilibrium by endogenizing the nominal wage and assuming that the monetary authority sets nominal GDP. I assume a big household where there is a single worker and a continuum of shoppers. The worker has full information, earns wage, and distributes spending across shoppers. Shoppers still have incomplete information. In this way, the monetary shock can affect both nominal marginal costs and nominal aggregate spending, yet remain unknown to shoppers. I then characterize the Phillips curve, where the output gap is proportional to the nowcast error of inflation. Unlike the modern Phillips curve literature, which emphasizes the role of firms' expectations in determining the slope, this model introduces a mechanism through which household expectations affect firms' pricing decisions. Furthermore, the more precise information households have, the steeper the Phillips curve is. We then achieve a state-dependent Phillips curve.

I conclude the model section by incorporating the main mechanism from the static model into a dynamic general equilibrium model. In the model, monetary policy sets the nominal interest rate. Shocks are persistent. Shopper's problem is repeated static. In each period, they receive a signal about the inflation and learn about the shock over time. I present a three-equation system that describes the joint dynamics of aggregate consumption, inflation and interest rate.

I then 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 sufficient statistic that summarizes the "deep" parameters related to the search friction. I take its value from Amiti et al. (2019) and Gopinath et al. (2011). Next, I estimate the impulse responses of inflation and inflation expectation following main inflation shocks and main business cycle shocks (Angeletos et al. (2020a)). I calibrate the information friction such that the time needed for the nowcast error shrinking to zero in the model is similar to its empirical counterpart.

Despite the model's parsimonious nature, the calibrated model can generate substantial monetary non-neutrality comparable to a standard New-Keynesian (NK) model with Calvo parameter equal to 0.7. That is, 70% firms cannot adjust prices in each period. The new insight in the dynamic model is that since the output response depends on the gap between the actual inflation and inflation nowcast, rapid learning can close this gap before the shock fully dissipates, leading to endogenously reduction in the persistence of the output response.

Finally, I present empirical evidence that supports my mechanism. I use the 2006–2019 NielsenIQ Consumer Panel data, which includes approximately 55,000 households annually, with each household participating in the panel for an average of 30 quarters. The search behavior observed in the data demonstrates path-dependence and a "one-stop" shopping pattern. To better capture search efforts, I first calculate the spending in a given product category that does not occur in the store the consumer visits most frequently. I then derive the non-routine share of total spending by summing this spending across all product categories. The average non-routine share of spending is about 35%. I find that one standard deviation (51 bp) increase in unanticipated inflation for food and drinks leads to a 26.5 bp increase in the non-routine share of spending, which is highly statistically significant. The direction of the response is consistent with the model prediction. The magnitude of response is modest (about a half) and I discuss several reasons that could potentially bias the estimate downward.

Related Literature – This paper contributes to three strands of literature. The first studies the price stickiness and monetary non-neutrality by focusing on the consumer-side mechanisms. The prominent paper in this literature is Matějka (2015). He shows that firms set discrete prices as consumers "hate" price fluctuations, which increases their costs of attention. My paper has the similar flavor in the sense that the price stickiness originates from the strategic interaction between consumers and firms. However, my paper has different mechanisms and investigates these mechanisms in general equilibrium. Gabaix and Graeber (2024) and Rebelo et al. (2024) propose behavioral theories of price stickiness. My paper is within the realm of rational expectations.

Second, it adds to the research on the role of search frictions in explaining price stickiness. Based on Burdett and Judd (1983), Head et al. (2012) show that price stickiness can result from the mixed pricing strategy. As nominal price increases, profit can still be maximized despite a fall in real price. However, in their model, the money is neutral. Similarly, Burdett and Menzio (2018) incorporate same mechanism into a menu-cost model, where a broader range of optimal prices leads to larger price adjustments. Benabou (1988) shows that when monopolistic competition arises from costly consumer search, the inaction region in a menu-cost model expands with increasing search costs. Gaballo and Paciello (2021) show a model where consumers are motivated to leave the monopolistic firm and find lower prices in a separate market where firms have perfect competition, when the inflation rises. The price is sticky because some demand shifts to the low-price market. However, my paper does not rely on the separation of markets, and firms are heterogeneous in productivity and compete monopolistically in a unified market. This paper advances this second stream by integrating search models a là Wolinsky (1986) and Anderson and Renault (1999) with a heterogeneous firm block, a monetary general equilibrium framework, and incomplete information on the consumer side.

Third, this paper contributes to the literature on the role of consumer-side information frictions in explaining monetary non-neutrality. Bénabou and Gertner (1993), L'Huillier (2020) and L'Huillier and Zame (2022) focus on the role of individual prices consumers observe during the search as revealing the information about aggregate shocks. Especially, L'Huillier (2020) consider a signaling game between a monopolistic firm and consumers. The price of that firm is rigid if the pooling equilibrium in which firms' profits are maximized is selected. The price is not rigid in the separating equilibrium. In my paper, I specifically shut down this signaling effect. Therefore, my mechanism is different from L'Huillier (2020). This paper also broadly contributes to the literature on information frictions and the transmission of monetary policy. For example, Angeletos and La'O (2013) model aggregate demand fluctuations driven by sentiment in beliefs. Venkateswaran (2014) examines an incomplete-information version of the Diamond–Mortensen–Pissarides model. Angeletos and Lian (2018) show that incomplete information games among consumers can mitigate the forward guidance puzzle.

Outline − The rest of the paper is organized as follows. Section 1 presents a static model and establishes the main results on the passthrough. Section 2 presents the dynamic model and the calibration results. Section 3 shows empirical evidences. The last section concludes. Appendix A contains some of the proofs omitted from the text. Appendix B contains the proof of the dynamic model and additional calibration details. Appendix C contains details of empirical setup and additional empirical evidence.

1 Static Model

I develop a macroeconomic model with (i) information asymmetry between shoppers and firms about nominal marginal costs and (ii) search frictions on the goods market. I proceed in two steps. First, I present a static model to explain the core mechanism. Second, I present a full-fledged dynamic general equilibrium framework in the next section. I first consider static partial equilibrium model in which there are firms, shoppers and a monetary authority and then close the model in the general equilibrium.

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 States, Strategies and Distributions

At the beginning of each period, Nature draws the aggregate state w from a given distribution Φ_w, idiosyncratic states for each firm a_k from Φ_a, idiosyncratic noisy signals for each shopper x_i from Φ_x about the aggregate state w. $\Phi_w, \Phi_x, \Phi_a, G$ are the common knowledge for all agents. Each firm produces a good and selects the optimal pricing strategy p^* that maps its state $\{a_k, w\}$ to the price of its good, i.e., $p_k = p^*(a_k, w)$, where $p^* : \mathcal{R}^2 \to \mathcal{R}$. The optimal pricing strategy monotonically decreases in a_k and increases in w. Variation in a_k across firms induces an endogenous price distribution $F(p|w)$, which has a density distribution $f(p|w)$. Let $p^{*-1}(p,w)$ denote the inverse mapping from p_k to a_k given w. It is straightfoward to see that p^{*-1} monotonically decreases in a_k and increases in w. The price distribution conditional on w is given by,

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

In the rational expectations equilibrium, shoppers know the equilibrium pricing strategy p^*. Conditional on knowing p^* and w, the shoppers are able to derive the price distribution. However, they have incomplete information about w. Their posterior belief is denoted by $H(w|x)$. The shopper's perceived price distribution conditional on x, $f(p|x)$, is given by,

$$f(p|x) = \int f(p|w)h(w|x)dw \tag{2}$$

The above equation has two implications. First, when $x = w$, the perceived and objective price distributions coincide. Second, if $h(w|x_i)$ first-order stochastically dominates (FOSD) $h(w|x_j)$, 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.

Moreover, the expected idiosyncratic state a_k after observing the price p_k is given by,

$$E(a_k|x) = \int p^{*-1}(p_k,w)h(w|x)dw \tag{3}$$

Similarly, it is obvious to show that $E(a_k|x_i) < E(a_k|x_j)$ when $h(w|x_i)$ FOSD $h(w|x_j)$. A shopper who believes a lower nominal wage perceives lower a_k.

1.2 Setup

Now, I state the model. Time is discrete and infinite $t \in \mathbf{N}$. The timeline is as follows. Within a period, the monetary authority first sets the nominal wage. Shoppers do not work. Instead, 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,

$$Y_k = A_k L_k \tag{4}$$

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 $\mathcal{N}(0, \sigma_a^2)$. 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 for now.

Monetary Authority − The monetary authority draws the log nominal wage w from the normal distribution $\mathcal{N}(\bar{w}, \sigma_w^2)$. Monetary shock is defined as $\hat{w} = w - \bar{w}$.

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,

$$X_i = W \exp\left(\sigma_x \varepsilon_{xi} - \frac{\sigma_x^2}{2}\right) \tag{5}$$

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. It is easy to show that the mean of X_i is W. The shopper treats X_i as signal about w. The expected log nominal wage is then given by,

$$E(w|x_i) = \theta x_i + (1 - \theta)\bar{w} \tag{6}$$

where $\theta = \frac{\sigma_x^{-2}}{\sigma_w^{-2} + \sigma_x^{-2}}$. Let $H(w|x_i)$ denote the posterior belief about the nominal wage with density $h(w|x_i)$. Based on the Bayes' rule, it follows $\mathcal{N}(E(w|x_i), (\sigma_w^{-2} + \sigma_z^{-2})^{-1})$. Shoppers construct their perceived price distribution according to (2).

A shopper buys and consumes only one good.[2] The utility that shopper i gains from consuming good k is given by,

$$\log \frac{X_i}{P_k} + \frac{1}{\lambda}\epsilon_{ik} \tag{7}$$

where $\epsilon_{ik} \sim G$ is match utility between shopper i and good k. G is triple continuously differentiable and its density function is g. It captures idiosyncratic consumer preferences for certain goods over others. ϵ_{ik} are i.i.d across firms and shoppers.[3] Also, following the literature, I assume that G is log-concave. Note that some commonly used distribution functions are log-concave, e.g., normal

[2]In Section 2.3, I extend the model by allowing shoppers to get access to a bunch of goods by incurring one search cost.

[3]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.

distribution, uniform distribution, Gumbel distribution.[4] 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,

$$y_{ik} = -p_k + \frac{1}{\lambda}\epsilon_{ik} \tag{8}$$

According to (8), the distribution of value of drawing a random good is a convolution of the perceived price distribution and the distribution of match utility,

$$\psi(y|x) = \int f(u-y|x)\lambda g(\lambda u)du \tag{9}$$

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. 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. I refer to the latter as search without commitment.[5] 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 the nominal wage from individual prices they observe during the search.*

This assumption is plausible if the idiosyncratic variations of the productivity shocks in prices are way larger than the aggregate variations of the nominal wage shocks. Therefore, even when a shopper consistently encounters firms charging high prices, she attributes this to bad luck rather than an increased nominal wage. I make this assumption for tractability. First, optimal prices are non-linear in productivity and nominal wage, which makes the linear Bayesian updating infeasible.

[4]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.

[5]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.

Second, if shoppers' information sets depend on the whole search history, it leads to exploding states and the model becomes intractable. As I will show, The main mechanism remains valid as long as information friction exists. It is independent of how shoppers acquire information. In Section 2, I assume that shoppers receive a signal about the current inflation before searching as an alternative way to incorporate learning from shopping.

I now state the shopper's problem. Let v_{ir} denote the maximum value of previously visited firms in rth round for shopper i. In particular, each shopper is assigned with a good for free, i.e., $v_{i1} = y_{i1}$. I define v_{ir} for $r > 1$ as follows,

$$v_{ir} = \max\{v_{ir-1}, y_{ir}\} \tag{10}$$

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,

$$U(v|x) = \max\left\{v, -\kappa + U(v|x)\int_{-\infty}^{v} \psi(y|x)dy + \int_{v}^{\infty} U(y|x)\psi(y|x)dy\right\} \tag{11}$$

The maximum represents that the shopper can either receive the maximum value v until this round and stop searching more, 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 (10). The value function U is stationary only when Assumption 1 holds. Otherwise, with the information sets expanding over the rounds of search, the value function U should be indexed by the search round.

The shopper's problem is solved in two steps. First, she finds the U function that solves the functional equation (11). Second, she keeps sampling firms until v first exceeds the expression to the right of the comma in (11).

Partial Equilibrium – The equilibrium concept is Perfect Bayesian Nash equilibrium (PBNE). 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.[6] Formally, I define the equilibrium as follows:

[6] In the standard search literature, prices are bounded and consumers know the range of prices. Therefore, they are able to detect the off-equilibrium prices. The regulation of off-equilibrium belief in the case where consumers observe prices not in this range is needed.

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.*
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 (11).*
3. *Nominal wage and cash 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 w.*

1.3 Equilibrium Characterization

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

Characterization of Optimal Strategies – I first characterize the search strategy. The 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 is unique for each x. It is determined by,*

$$v^*(x) = -\frac{\kappa}{1-\Psi(v^*(x)|x)} + \frac{\int_{v^*(x)}^{\infty} y\psi(y|x)dy}{1-\Psi(v^*(x)|x)} \tag{12}$$

where $\Psi(p|x)$ is the perceived distribution of the value of a random draw. In addition, $v^*(x)$ decreases in x.

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,

$$U(v|x) = \max\{v, v^*(x)\} \tag{13}$$

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

The result extends the standard threshold rule in the literature (Weitzman, 1979; Wolinsky, 1986) by incorporating both endogenous price distribution and incomplete information. In these models, firms are homogeneous and shoppers have correct belief about the optimal price. Here, firms are heterogeneous and shoppers have incomplete information about the price distribution. Proposition 1 shows that under Assumption 1, the optimal search strategy still follows the threshold rule.[7] The threshold decreases in the shopper's information set x. In particular, a higher signal about the nominal wage leads to a lower threshold. Intuitively, a shopper who perceives a lower nominal wage believes the price distribution is shifted leftward. As a result, the increased perceived likelihood of encountering a low-price firm raises the threshold.

To make clear how price distribution and match utility distribution affect the threshold, the following Corollary shows an alternative way to solve the threshold.

Corollary 1. *The threshold $v^*(x)$ is given by,*

$$\int \int_{\lambda(v^*(x)+p)}^{\infty} \big(\frac{1}{\lambda}\epsilon - p - v^*(x)\big) g(\epsilon) d\epsilon f(p|x) dp = \kappa \tag{14}$$

where $f(p|x)$ is the perceived price distribution.

The left-hand side represents the expected additional benefit of additional search at the threshold, i.e., when shopper's state is $v^*(x_i)$. From Proposition 1, she is indifferent between sampling

[7]One can show that the threshold rule holds when shoppers can learn from individual prices, only if variance of productivity shocks is sufficiently larger than the variance of monetary shocks. Rothschild (1974) shows that the consumers still follow a threshold rule if the difference between these two prices is smaller than their informational content. The threshold in this case depends on the whole history of price observations.

another firm and stopping searching. 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 additional utility obtained in this case is $\max\{\frac{1}{\lambda}\epsilon_{ik} - p_k - v^*(x_i), 0\}$. The threshold is achieved when the expected additional utility is equal to the search cost.

Aggregation – I now show the aggregation of optimal search decisions. In particular, I show the expenditure allocation across firms and the resulting profits. According to the optimal search strategy, the shopper only purchases the good k if $\frac{1}{\lambda}\epsilon_{ik} - p_k > v^*(x)$. The shopper's probability of purchasing from firm k in each round is $1 - G\big(\lambda(v^*(x) + p_k)\big)$. Since learning from shopping is prohibited, this probability is the same for all rounds of search. The probability of any shopper purchasing from any firm in each search given w is given by,

$$\rho = \int \int \big(1 - G\big(\lambda(v^*(x) + p)\big)\big) d\Phi_x(x) f(p|w) dp \tag{15}$$

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. It is straightforward to show that 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,

$$\omega_k = \frac{1}{\rho} \int X\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x) \tag{16}$$

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 given by,

$$\pi_k = \frac{1}{\rho} \int X\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x) \frac{1}{P_k}(P_k - \frac{W}{A_k}) \tag{17}$$

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 (17) with respect to prices. The following proposition presents the optimal pricing strategy,

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,*

$$P_k = \mu_k \frac{W}{A_k}; \quad \mu_k = \frac{e_k}{e_k - 1} \tag{18}$$

The elasticity of demand e_k is determined by,

$$e_k = \lambda \frac{\int X g\big(\lambda(v^*(x) + p_k)\big) d\Phi_x(x)}{\int X\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x)} + 1 \tag{19}$$

Proof. See Appendix A. ■

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

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. Similar to Corollary 1, it is given by,

$$\int \int_{\lambda(v^*(w)+p)}^{\infty} \big(\frac{1}{\lambda}\epsilon - p - v^*(w)\big) g(\epsilon) d\epsilon f(p|w) dp = \kappa \tag{20}$$

where $f(p|w)$ is the actual price distribution. It is easy to show that firm's profit is given by,

$$\pi_k = \frac{1}{\rho}\Big(1 - G\big(\lambda(v^*(w) + p_k)\big)\Big)\frac{W}{P_k}(P_k - \frac{W}{A_k}) \tag{21}$$

where $\rho = \int \big(1 - G(\lambda(v^*(w) + p))\big) f(p|w) dp$. In this case, the distribution of thresholds is reduced to a single value $v^*(w)$. The optimal pricing strategy is given by,

$$P_k = \frac{e_k}{e_k - 1}\frac{W}{A_k} \tag{22}$$

$$e_k = \lambda \frac{g\big(\lambda(v^*(w) + p_k)\big)}{1 - G\big(\lambda(v^*(w) + p_k)\big)} + 1 \tag{23}$$

Firm's market power is determined by two factors. Larger λ implies more competition on prices, inducing higher elasticity.[8] In addition, the search friction naturally gives rise to monopoly power (Diamond, 1971). In particular, 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

[8] 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 $\lambda + 1$.

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 that setting a higher price motivates marginal shoppers to search, while extracting more profit from the infra-marginal shoppers. 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, high-productivity firms will set lower price and have higher markup.

Now, I present an important property of $v^*(w)$. It is a crucial step in proving the existence of the equilibrium. This property is also useful to understand the effect of search frictions on other equilibrium objects I will discuss in Section 1.4.

Lemma 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 (23). The Lemma show that the direct effect outweighs the indirect effect. I now establish the existence of the full-information steady-state 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. It implies that higher price induces higher elasticity and lower markup. Therefore, the individual prices are uniquely determined by (22) and (23). Second, from Lemma 1, the threshold $v^*(w)$ is uniquely determined by (20) for the given price distribution that is derived from optimal pricing strategy. The equilibrium is then the fixed point of the reservation value and the price distribution. Note that there always exist the equilibria in which firms charge sufficiently high prices and shoppers do not search. However, there is only one equilibrium in which shoppers search actively. From now on, I refer to the full-information equilibrium where $w = \bar{w}$ as the full-information steady state.

Next, I present the first main result about monetary non-neutrality.

Theorem 2. *In the full information equilibrium, monetary policy is neutral. Specifically, nominal wage 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 nominal wage shock, although consumers do not know individual prices of each firm, they understand the relative prices are unchanged and thus there is no gain from searching more or less. In other words, search friction alone is not sufficient to generate monetary non-neutrality. We need the interplay of both search and information frictions.

Approximate Optimal Strategies – Both the threshold shown in Proposition 1 and the pricing strategy shown in Proposition (2) are highly non-linear. Following the literature, I consider the first-order approximation around full-information steady state given its existence. In particular, the monetary authority draws monetary shocks 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.

To proceed, I first define the passthrough from the monetary shock $\hat{w}$ to prices in equilibrium. The first-order approximation to the optimal pricing strategy p^* is given by,

$$p^*(a_k, w) = p^*(a_k, \bar{w}) + p^*_w(a_k, \bar{w})\hat{w} \tag{24}$$

I refer to $\varphi_k = p^*_w(a_k, \bar{w})$ as the total passthrough, as it reflects the sum of two passthroughs which I will show in Section 1.4. Shoppers know the distribution of total passthrough as they know $p^*, \bar{w}$ and Φ_a in equilibrium.

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, Hulten, 1978) offers a simple non-parametric formula for the change in price index. If preference is homothetic and stable[9], which

[9]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 17, the demand for firm k is proportional to the aggregate income $W = \int X_i di$, which implies homotheticity. In addition,

is the case here, the log change in the price index is the expenditure share-weighted log changes in all the prices.

$$\hat{p} = \Phi\hat{w} \tag{25}$$

where $\Phi = \int \varphi_k \bar{\omega}_k dk$ and $\bar{\omega}_k = \frac{1}{\rho}\big(1 - G\big(\lambda(v^*(\bar{w}) + p_k)\big)\big)$ is the expenditure share in the full-information steady state. Therefore, the aggregate total passthrough Φ measures the passthrough from the nominal wage shock to the price index. In particular, shoppers know Φ as they know the distribution of φ_k and $\bar{\omega}_k$. Therefore, the average expectation of the change in price index is $\bar{E}\hat{p} = \Phi\bar{E}\hat{w} = \theta\hat{p}$. For larger information friction, θ is closer to zero, indicating more dampening response of $\bar{E}\hat{p}$ compared to $\hat{p}$.

The following proposition presents a first-order approximation to the optimal search strategy in (14) and the optimal pricing strategy in (19).

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 threshold* $v^*(x)$ *is given by,*

$$v^*(x) = v^*(\bar{w}) - E(\hat{p}|x) \tag{26}$$

where Φ *is the aggregate total passthrough.*
*[**Part 2**] The elasticity of demand is given by,*

$$e_k = \lambda \frac{g\big(\lambda(v^*(w) + \bar{p}_k + \hat{p}_k - \bar{E}\hat{p})\big)}{1 - G\big(\lambda(v^*(w) + \bar{p}_k + \hat{p}_k - \bar{E}\hat{p})\big)} + 1 \tag{27}$$

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 threshold under incomplete information is equal to the threshold under full-information steady state, adjusted downward the expected change in the price index. Consistent with Proposition 1, the threshold is negatively related to x. The result takes a step forward by showing that under the first-order approximation, all the changes in perceived price distribution following the shock, as denoted in (24), for shoppers with information set x is encapsulated by $E(\hat{p}|x)$. To see the intuition, remember $E(\hat{p}|x) = \Phi E(\hat{w}|x)$. Suppose the total

there is no other factors, such as taste shocks, state of nature, other than relative prices and income, that can affect the demand.

passthroughs are uniform for all firms, i.e., $\varphi_k = \varphi_0$, then $\Phi = \varphi_0$. An increase in the perceived nominal wage shifts the entire price distribution uniformly to the right, as occurs in the full-information equilibrium where $\varphi_k = 1, \forall k$. If, however, the distribution of total passthrough is non-uniform, the shape of price distribution also changes after the shock. In particular, the negative covariance between expenditure share and total passthrough induces lower aggregate total passthrough, thereby lowering the average threshold by less. Furthermore, the average change of thresholds is equal to $\bar{E}\hat{p} = \theta\hat{p}$, which implies that the adjustment in threshold is smaller due to information friction, leading to higher thresholds, i.e., pickier consumers, compared to the full-information benchmark.

The second part of the proposition characterizes the elasticity of demand when thresholds are on average higher than the full-information benchmark. On the first order, the distribution of thresholds and covariance between cash in hand and thresholds vanish to zero and only $\bar{E}\hat{p}$ is retained, which is inherited from the average change in thresholds. The result shows that the elasticity depends on the perceived relative price, $\hat{p}_k - \bar{E}\hat{p}$. Due to the information friction, it is larger than the actual relative price. Consequently, firms behave as though they are competing with rivals setting prices lower than their actual levels, prompting them to reduce their own prices. This increases the elasticity of demand, which in turn decreases markups and leads to incomplete passthroughs of monetary shocks to prices. This mechanism will be examined in greater detail in Section 1.4. The intuition is closely related to Lucas (1972), where agents mistake increases in nominal prices for increases in relative prices. However, in this context, the mechanism operates entirely on the household side.

1.4 Characterization of Passthroughs in Equilibrium

I now characterize passthroughs in the equilibrium. I present the main finding: the aggregate passthrough of a money supply shock to the price index is generically incomplete. The total passthrough is composed of two fundamental passthroughs: the own-cost passthrough and the cross-price passthrough. Decomposing the total passthrough into these two elements is crucial to understand the intuition behind the main finding. Following Amiti et al. (2019), I define both passthroughs using markup elasticities.

Lemma 2. *The price responds to both firm's own cost shocks and competitors' prices,*

$$\hat{p}_k = \gamma_k(\hat{w} - \hat{a}_k) + \xi_k\hat{p} \tag{28}$$

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

$$\gamma_k = (1 - \frac{d\log\mu_k}{dp_k}\Big|_{\hat{w}=0})^{-1}; \xi_k = \frac{d\log\mu_k}{dp}\Big|_{\hat{w}=0}\gamma_k \tag{29}$$

where μ_k is the markup. The total passthrough of each firm is given by,

$$\varphi_k = \gamma_k + \Phi\xi_k \tag{30}$$

The aggregate total passthrough Φ is given by,

$$\Phi = \frac{\Gamma}{1 - \Xi} \tag{31}$$

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 19. The nominal wage shock acts as both a shock to marginal costs and a shock to competitors' prices. Accordingly, the total passthrough is defined as the combination of these two components. By integrating both sides, the aggregate total passthrough is derived. Importantly, the own-cost passthrough γ does not depend on the information friction θ when evaluated at the steady state. Firms pass part of their idiosyncratic cost shocks to prices regardless of the aggregate shocks and the associated information friction. In contrast, the cross-price passthrough ξ depends on how shoppers perceive the price index.

I now state our main results on passthroughs under incomplete information. To push the results on monotone comparative statics as far as possible, I assume that the distribution of the match utility, G, follows the Gumbel distribution.

Theorem 3 (Aggregate Total Passthrough under Incomplete Information)**.** *Under incomplete information about the nominal wage, the aggregate total passthrough has following properties,*

1. ***[Incompleteness]*** $\Phi < 1$
2. ***[Composition]*** *φ_k decreases in productivity; $\bar{\omega}_k$ increases in productivity.*
3. ***[MCS on κ]*** *Φ decreases in search friction κ given θ. Limits: $\lim_{\kappa\to\infty} \Phi = 0$; $\lim_{\kappa\to 0} \Phi = 1$*

   [MCS on θ] *Φ increases in information friction θ given κ. Limits: $\lim_{\theta\to 1} \Phi = 1$; $\lim_{\theta\to 0} \Phi = \Gamma$*

Proof. See Appendix A. ■

This is the main theorem of this paper. It establishes that the total passthrough is generically incomplete if there exists any information friction. The full-information case is a knife-edge case. I first explain the result on incompleteness. The key equation to understand the intuition is:

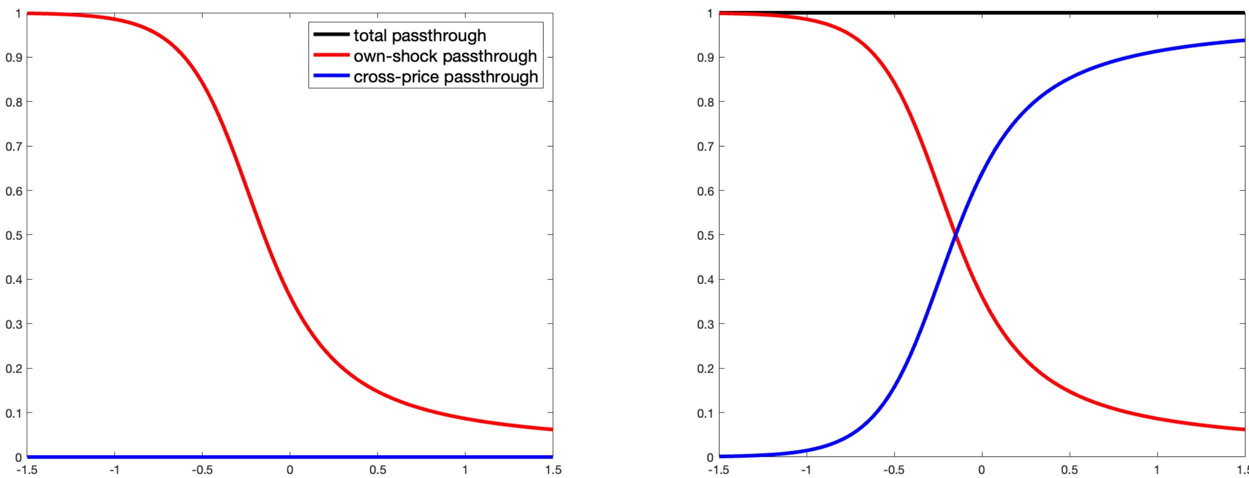
$$1 - (\gamma_k + \xi_k) = (1 - \gamma_k)(1 - \theta) > 0 \tag{32}$$

It shows that the difference between the sum of the passthroughs, $\gamma_k + \xi_k$, and one can be decomposed into two terms. The first term captures the role of the search friction. Recall that the own-cost passthrough is not related to the information friction. The second term captures the role of information asymmetry about the nominal wage between consumers and firms. They are both positive. 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 total passthrough is always one for firms of any productivity. This result aligns with Theorem 2, where I prove the exact result without relying on the first-order approximation. The decomposition shows the contribution of each passthrough to the total passthrough for firms with different productivities. Notably, Amiti et al. (2019) establishes that the total passthrough is one for broad preferences including nested CES and first-order Kimball demand family as well as for the broad homothetic families of demand considered in Matsuyama and Ushchev (2017). 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 total 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 actual change in the price index to their prices. In this case, the change in the nominal wage is as if an idiosyncratic cost shock to firms. In the intermediate case when information is incomplete, the increase in cross-price passthrough is still not sufficient to offset the decrease in own-shock passthrough, leading to incomplete total passthroughs.

Moreover, the individual total passthrough is defined as $\varphi_k = \gamma_k + \Phi\xi_k$. The effect of strategic complementarity is further dampened by the aggregate total passthrough as $\Phi < 1$. This induces a even smaller individual total passthrough. Intuitively, the fact that shoppers understand that aggregate total passthrough is incomplete reduces the expected change in price index for any given

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 total passthrough. The left and right panels show passthroughs in the incomplete and complete information cases, respectively.

nominal wage shock. Firms then reduce prices in response, resulting in even lower passthrough. Therefore, the incompleteness of the total 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).

I leave the discussion of composition and comparative statics results to the later section. Now, I focus on clarify the main insight behind the incompleteness result.

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. In this section, 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 (2003a), and Angeletos and Lian (2023). The following proposition presents a way to translate the above economy into a beauty contest game.

Proposition 4. *The model economy can be expressed as a beauty contest game as follows,*

$$\hat{p} = \Gamma \hat{w} + (1 - \Gamma)\bar{E}\hat{p} \tag{33}$$

where $\bar{E}\hat{p}$ represents shoppers' average expectation of $\hat{p}$. In the rational expectations equilibrium, the change in price index can be express as the infinite sum of higher-order beliefs of $\hat{w}$,

$$\hat{p} = \Gamma \sum_{h=0}^{\infty} (1-\Gamma)^h \bar{E}^h \hat{w} \tag{34}$$

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

$$\Phi = \Gamma \sum_{h=0}^{\infty} (1-\Gamma)^h \theta^h \tag{35}$$

Proof. See Appendix A. ■

The proposition 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 standard beauty contest game between firms (Woodford, 2003a), where strategic complementarity is in terms of *firms'* beliefs about the price index.[10]

In the first round, when the nominal wage shock occurs, assume no firm has yet responded to the monetary shock and that firms expect other firms not to adjust their prices. Thus, 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), where the rounds of thinking extend to $k \to \infty$, representing the gradual adjustment process toward a rational equilibrium.

Moreover, the information friction attenuates the strategic complementarity in pricing. It attenuates more the high-order strategic complementarity, since the higher-order beliefs of shoppers

[10] To write the game into the infinite sum of higher-order beliefs as in (34), 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 know the common knowledge of rationality among shoppers. 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. 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.

are more anchored to the prior. For θ close to zero (meaning a sufficiently large departure from common knowledge of the shock), the aggregate total 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.

Symmetry – In this section, I give interpretations about the symmetry of incomplete passthrough for both positive and negative monetary shocks.

Here's a revised version of your text for improved clarity and flow:

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, balancing the additional profits from infra-marginal shoppers with the benefits of attracting new shoppers. In this case, the appropriate interpretation is that firms limit the passthrough to extract greater profits from infra-marginal consumers.

Firm Heterogeneity in Passthroughs – To discuss the second part of Theorem 3, I first complete the picture by showing the cross-sectional properties of three types of individual total passthroughs and their limit results. I define passthroughs on the primitive: $\varphi_k = \varphi(a_k)$, $\gamma_k = \gamma(a_k)$ and $\xi_k = \xi(a_k)$.

Proposition 5. *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, high-productivity firms set lower prices, attract more shoppers, and choose higher markups. Similar to oligopolistic CES models, the lower bound of markups is determined by the degree of substitutability 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).

Furthermore, because information frictions dampen only the cross-price passthrough, high-productivity firms experience a more pronounced decline in cross-price passthrough and, therefore, total passthrough. Coupled with their high expenditure share, these firms disproportionately contribute to the incompleteness of aggregate total passthrough. According to the limit results, the total passthrough is bounded below by the response of $\bar{E}\hat{p}$ to $\hat{w}$, i.e., $\Phi\theta$. This implies that $\hat{p}_k - \bar{E}\hat{p}$ is always positive, resulting in uniformly higher perceived relative prices. Consequently, following a positive monetary shock, shoppers are more likely to search and markups decrease for all firms.

Supply Side Effect of Monetary Shock – The positive monetary shock induces more search under incomplete information, with larger information frictions leading to even greater search intensity. Increased search raises the utility cost due to higher total search effort. However, shoppers are more likely to be drawn to high-productivity, low-price firms, leading to an expansion in demand for these firms and a contraction for low-productivity firms. This reallocation improves allocative efficiency and boosts aggregate productivity (Baqaee et al., 2024). In general equilibrium with finite labor elasticity, this amplifies the output response while dampening the inflation response following a monetary shock as shown in Proposition 11.

Monotone Comparative Statics – To understand the results of comparative statics of the aggregate total passthroughs in Theorem 3, I establish that, all else equal, (i) the total passthrough is larger when the search cost is smaller for all firms and (ii) the total passthrough is larger when information friction is smaller for all firms.

Proposition 6. *Let equilibrium passthrough distributions $\gamma(a;\kappa), \xi(a;\kappa), \varphi(a;\kappa)$ for given κ. Given $\sigma_z > 0$, for $\kappa_2 > \kappa_1$,*

1. *Own-cost passthrough: $\forall a, \gamma(a;\kappa_2) < \gamma(a;\kappa_1)$; $\lim_{\kappa\to\infty} \gamma(a) = 0$; $\lim_{\kappa\to 0} \gamma(a) = 1$*
2. *Cross-price passthrough: $\forall a, \xi(a;\kappa_2) > \xi(a;\kappa_1)$; $\lim_{\kappa\to\infty} \xi(a) = \theta$; $\lim_{\kappa\to 0} \xi(a) = 0$*
3. *Total passthrough: $\forall a, \varphi(a;\kappa_2) < \varphi(a;\kappa_1)$; $\lim_{\kappa\to\infty} \varphi(a) = 0$; $\lim_{\kappa\to 0} \varphi(a) = 1$*

Proof. See Appendix A. ■

This proposition establishes that, all else equal, total passthrough is lower in an economy with higher search costs for any given productivity. Therefore, the search friction can amplify the effect of the information asymmetry on total passthroughs. As we know from Lemma 1, higher search cost lowers the thresholds, which results in lower elasticity of demand. Higher markups are more sensitive to the change in relative prices, leading to lower own-cost passthrough. Based on (32), lower own-cost passthrough implies larger gap between the sum of two fundamental passthroughs and one, resulting in lower total passthrough. In addition, the limit results implies that the total passthrough can vary from zero to one, implying the possibility of large degree of monetary non-neutrality.

Proposition 7. *Let equilibrium passthrough distributions $\gamma(a;\theta), \xi(a;\theta), \varphi(a;\theta)$ for given θ. For $\theta_2 > \theta_1$,*

1. *Own-cost passthrough: $\forall a, \gamma(a;\theta_2) = \gamma(a;\theta_1)$*
2. *Cross-price passthrough: $\forall a, \xi(a;\theta_2) > \xi(a;\theta_1)$; $\lim_{\theta\to 1} \xi(a) = 1 - \gamma(a)$; $\lim_{\theta\to 0} \xi(a) = 0$*
3. *Total passthrough: $\forall a, \varphi(a;\theta_2) > \varphi(a;\theta_1)$; $\lim_{\theta\to 1} \varphi(a) = 1$; $\lim_{\theta\to 0} \varphi(a) = \gamma(a)$*

Proof. See Appendix A. ■

This proposition highlights that, all else equal, the total passthrough increases at any given productivity as information becomes more precise. Intuitively, if shoppers are more aware of changes in price index, firms will be more responsive to changes in competitors' prices. Higher cross-price passthrough pushes up the total passthrough. In contrast, the own-cost passthrough is irrelevant to the information frictions. 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 shows that the passthrough from changes in the nominal wage to price index is generically incomplete. The incompleteness is mostly contributed by high-productivity firms. The aggregate passthrough is lower when there is more search frictions and information frictions. Proposition 4 shows that the main insight lies on the attenuation effect of the information friction on strategic complementarity. Propositions 5-7 yield more disaggregated predictions of three types of passthroughs for individual firms and additional limit results.

1.5 General Equilibrium

In this section, 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, but 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. In addition, there is no security markets where shoppers can trade claims that are contingent on the search process and final choices.[11]

The timeline is as follows. The period is divided into morning and afternoon. In the morning, the monetary authority sets the nominal GDP, and the worker makes decision on labor supply and aggregate consumption spending they send to shoppers. In the afternoon, firms post prices, and each shopper searches sequentially and decides where to make the purchase. They are constrained by the amount of cash allocated to them.

I state the worker's problem. The worker maximizes the expected value that she will obtain from shoppers who will shop and consume in the afternoon, net of the disutility associated with

[11] 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.

labor supply.

$$\begin{aligned}
&\max_{X,L} \int \left(\log \frac{X_i}{P_k} + \frac{1}{\lambda}\epsilon_{ik} \right) di - L \\
\text{s.t.} \quad & X = WL + \Pi \\
& X_i = X \exp\left(\sigma_x \varepsilon_{xi} - \frac{\sigma_x^2}{2}\right)
\end{aligned}$$

where X is the aggregate consumption spending. The division of X among shoppers is random, where $\varepsilon_{xi} \sim \mathcal{N}(0,1)$ and $\int X_i di = X$. L is the labor effort, W is the nominal wage, and Π is firms' total nominal profit. P_k and ϵ_{ik} are the price and the match utility of the firm that shopper i accepts in the search process. The uncertainty of final choice of firms is washed out on the household level. In other words, for the worker, the ex-ante sum of shoppers' value is equal to the ex-post one.

Due to the log utility, the consumption spending X_i only shifts the level of the value and does not affect the search decisions. Therefore, the worker and shopper problems can be separated. The first-order conditions imply $X = W$. Since $X = PC$, it implies

$$\hat{p} + \hat{c} = \hat{w} \tag{36}$$

I assume that shoppers only learn about the aggregate nominal wage W from X_i and they search sequentially. The shopper problem is the same as before.

Aggregate Supply Shocks – The framework can also incorporate the aggregate supply shocks. Here, we consider aggregate supply shocks as aggregate productivity shocks. In particular, the firm's productivity has two components,

$$\log A_k = \log A + \sigma_a \varepsilon_{ak} \tag{37}$$

where A is the aggregate productivity shock. Let $a = \log A$. It draws from $\mathcal{N}(0, \sigma_A^2)$. I only consider one aggregate shock a time. Specifically, when analyzing aggregate supply shocks, the nominal wage is common knowledge to all agents. Shoppers also 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,

$$p^*(a_k, w, a) = p^*(a_k, w - a) \tag{38}$$

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 total passthrough Φ. The value of Φ depends on the information friction in this case.

Phillips Curve – The following proposition characterizes the Phillips curve.

Proposition 8. *The Phillips curve is given by,*

$$\hat{p} - \bar{E}\hat{p} = \frac{\Gamma}{1-\Gamma}(\hat{c} - \hat{a}) \tag{39}$$

where $\hat{a}$ is the aggregate productivity shock and Γ is the aggregate own-cost passthrough.

Proof. See Appendix A. ■

This proposition shows that the output gap is proportional to shoppers' nowcast error of inflation. It is well-documented that average expectation of inflation under-react to aggregate shocks, even among professional forecasters (Coibion and Gorodnichenko, 2015a; Bordalo et al., 2020; Angeletos et al., 2020b). When information frictions are larger, inflation expectations deviate further from actual inflation, resulting in a larger output gap. Conversely, smaller information frictions reduce the output gap. However, once household inflation expectations are accounted for, the slope of the Phillips curve becomes constant, provided that search frictions remain unchanged. This result aligns well with findings in Coibion and Gorodnichenko (2015b) and Beaudry et al. (2024), who argue that controlling for household inflation expectations helps explain the "missing disinflation" during the Great Recession and the hyperinflation observed in the post-pandemic era. Moreover, Hazell (2024) emphasize that controlling for other types of inflation expectations—such as those from the Survey of Professional Forecasters, inflation expectations derived from TIPS, or the Blue Chip Survey—is insufficient, highlighting the pivotal role of household inflation expectations. Therefore, our model provides a micro-foundation for why household inflation expectation is crucial in explaining these puzzles in the data.

Another way to write (39) is as follows,

$$\hat{p} = \frac{\Gamma}{1-\Gamma}\frac{1}{1-\theta}(\hat{c} - \hat{a}) \tag{40}$$

where θ is the average information friction about the particular shock of interest. This formulation emphasizes the slope of Phillips curve is endogenous to the information households have and thus state-dependent. The slope steepens when shoppers have more information about the shock. In the limit, the curve is vertical. The lower bound of the slope is governed by aggregate own-cost passthrough as the information approaches zero.

2 Dynamic Model

I now present the full-fledged dynamic general equilibrium model. The dynamic model will translate the monetary shock as a shock to the nominal GDP to a shock to the interest rate. It also features persistent shocks and slow learning about the underlying shock for shoppers.

The timeline is as follows. In each period, 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. The worker has full information and shoppers have incomplete information about aggregate shocks.

Firm – Let A_{kt} denote the firm's productivity

$$\log A_{kt} = \log A_t + \eta_{kt} \tag{41}$$

where η_{kt} is the idiosyncratic productivity. It follows,

$$\eta_{kt} = \rho_a \eta_{kt-1} + \sigma_a \varepsilon_{akt} \tag{42}$$

where the idiosyncratic productivity shock follows $\varepsilon_{akt} \sim \mathcal{N}(0,1)$. Also, A_t is the aggregate productivity. Denote $a_t = \log A_t$. It follows an AR(1) process,

$$a_t = \rho_A a_{t-1} + \sigma_A \varepsilon_{At} \tag{43}$$

where ε_{At} is the shock to aggregate productivity. It follows $\varepsilon_{At} \sim \mathcal{N}(0,1)$.

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

$$i_t = \phi \hat{\pi}_t + \upsilon_{mt} \tag{44}$$

where υ_{mt} follows,

$$\upsilon_{mt} = \rho_m \upsilon_{mt-1} + \sigma_m \varepsilon_{mt} \tag{45}$$

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

Worker – The worker maximizes the expected discounted utility with discount factor $\beta \in (0,1)$ and period utility defined over the sum of values which shoppers will obtain in the afternoon. Worker can save in risk-free bonds B_t (in zero net supply) that pay an interest rate of R_t.

$$\max_{B_t, X_t, L_t} E_0 \sum_{t=0}^{\infty} \beta^t \left(\int \left(\log \frac{X_{it}}{P_{kt}} + \frac{1}{\lambda} \epsilon_{ikt} \right) di - L_t \right)$$

$$\text{s.t.} \quad X_t + B_t = W_t L_t + R_{t-1} B_{t-1} + \Pi_t$$

$$X_{it} = X_t \exp\left(\sigma_x \varepsilon_{xit} - \frac{\sigma_x^2}{2}\right)$$

where X_t is the total consumption expenditure transferred to all the shoppers. The division of the consumption expenditure among shoppers is random and i.i.d. across shoppers and time, where $\varepsilon_{xit} \sim \mathcal{N}(0,1)$ and $\int X_{it} di = X_t$. L_t is the labor effort, W_t is nominal wage, and Π_t is total nominal profits of firms. Again, the worker's problem can be simplified due to the log utility.

Shopper – The shopper solves the static problem in (11) period by period given the history of signals. I make the simplifying assumption that shoppers observe the past price level but do not extract information from it. Following Mondria et al. (2021) and Angeletos and Huo (2021), this assumption can be interpreted as a form of bounded rationality or inattention. It is made for two reasons. First, measures of household price level expectation is not available. The assumption transforms the Phillips curve in (39) into a relationship that directly links the output gap to the nowcast error of inflation, rather than the log deviation of price index. This allows us to use the inflation expectation in the Michigan Survey of Consumers to calibrate the model. Second, it also allows for a direct comparison of the impulse responses generated by our dynamic general equilibrium model with those produced by the standard NK model.

Besides the signal X_{it}, I assume that each shopper receives a noisy signal about $\hat{\pi}_t$ at the beginning of each period to accommodate learning from observed prices. Since these signals are about the endogenous variables, they can be applied to both monetary and supply shocks.

Dynamic Equilibrium – I present the proposition that describes the dynamics of consumption, inflation and interest rate in equilibrium.

Proposition 9. *The equilibrium dynamics of* $\{\hat{\pi}_t, \hat{c}_t, i_t\}$ *is described by the following system of three equations:*

$$\hat{c}_t = E_t\hat{c}_{t+1} - (i_t - E_t\hat{\pi}_{t+1})$$
$$\hat{\pi}_t - \bar{E}^s\hat{\pi}_t = \frac{\Gamma}{1-\Gamma}(\hat{c}_t - \hat{a}_t)$$
$$i_t = \phi\hat{\pi}_t + v_{mt}$$

where Γ *is the aggregate own-cost passthrough and* $\bar{E}^s$ *is the shoppers' average expectation.*

Proof. See Appendix B. ■

This system is only different from the standard NK model in terms of the Phillips curve. First, since the worker has full information and decides the aggregate consumption spending and bond decisions, the standard Euler equation applies. Second, the monetary non-neutrality is determined by how accurately and fast shoppers learn about the inflation. In Appendix B, I discuss the dynamics of the system if we retain the Phillips curve in (39) combined with a price-level targeting monetary rule.

2.1 Calibration

In this section, I present a calibration of the model. I proceed in two steps. First, I calibrate the aggregate own-cost passthrough and deep parameters that are related to the search friction. Second, I calibrate the information friction based on the empirical evidence.

Aggregate own-cost passthrough – 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 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 does not depend on the information friction. Therefore, we can separate the calibration of these two frictions.

Amiti et al. (2019) show that the aggregate own-cost passthrough is around 0.5 in the universe of Belgian manufacturing firms. Gopinath et al. (2011) use a retail chain database which contains information on wholesale costs and demonstrate substantial variation in own-cost passthrough estimates, with a median of around 0.5 for U.S. stores and 0.25 for Canadian stores. Own-cost passthrough estimates are often biased upward because, without controlling for all competitors' prices, the estimates may capture strategic complementarity effects. Specifically, other firms respond simultaneously to both the correlated cost shocks and change in that firm's price. Therefore, I pick $\Gamma = 0.3$ which is closed to the lower end of the their estimates.[12]

Although the deep parameters related to the search friction are not relevant for impulse responses, it will still be valuable to check the sanity of the implied values of these parameters. 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 one of the extension, I generalize the model to the setup where one firm sell multiple products and there is no search friction within the store. I show that retailers charge the same markup for all products they sell. In addition, the idiosyncratic productivity dispersion σ_a and persistence are set to be consistent with Decker et al. (2020), i.e., $\sigma_a = 0.3$, $\rho_a = 0.6$. The following table shows the baseline calibration for deep parameters that do not relate to the information friction.

Table 1: Baseline calibration of the model

Parameter	Description	Value
κ	Search cost	0.22
λ	Relative importance of match utility	5.51

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

[12]Interestingly, Amiti et al. (2019) also show that the sum of own-cost and cross-price passthroughs cannot be rejected from being one. This seems to indicate that the passthrough from the shock to prices is close to one in the model. However, search and information frictions are more prevalent in non-tradable industries that directly engage with household, such as retail and broad service sectors. In contrast, the buyer-seller relationships in tradable industries usually involves firms on both sides and firms are generally better informed about prices compared to households. Their relationships are often governed by contracts (Gopinath et al., 2011).

Table 2: Model fit

	Moment	Model	Data	Source
M1	Average markup	1.67	1.67	DellaVigna and Gentzkow (2019)
M2	Average own-cost passthrough	0.3	0.3	Amiti et al. (2019); Gopinath et al. (2011)

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

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 two implications. First, although the targeted elasticity of demand is 2.5, the implied $\lambda + 1$ is about 6.5. Remember $\lambda + 1$ represents the elasticity of demand when $\kappa \to 0$ as shown in Proposition 6. 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 contact per shopper ρ^{-1} equal to 1.26. This suggests a relatively large search cost and about 80% shoppers make the purchase on the first search. Therefore, Assumption 1 is more plausible, as shoppers only visit a limited number of firms.

Information friction − In our Phillips curve, the output gap is proportional to household nowcast error of inflation. Understanding how households learn about inflation is therefore critical. However, households may draw on different sources of information about inflation depending on the nature of the shocks. For instance, Kumar et al. (2015) shows that households are more attentive to salient prices, such as oil prices. Similarly, D'Acunto et al. (2021) find that households learn inflation from shopping in the retail sector. Moreover, Candia et al. (2023) document that households may confuse demand shocks with supply shocks, and Liu and Zhang (2024) highlight that a specific signal structure is necessary to generate such confusion.

To ensure the robustness of inflation learning in different scenarios, I consider four types of shocks. The first two are the main inflation shock and the main business cycle shock from Angeletos et al. (2020a). The main inflation shock is identified by maximizing its contribution to the business-cycle variation in inflation. A key feature of this shock is its minimal impact on real quantities and zero impact on TFP, making it akin to cost-push shocks in DSGE models. Since this shock affects only inflation without influencing real variables, household learning relies solely on observing inflation itself, capturing a "generic" way of learning about inflation, e.g., a noisy signal about inflation. The main business cycle shock is identified by maximizing its contribution to the business-cycle variation in unemployment. It accounts for the large bulk of the business-cycle comovements in unemployment, hours worked, output, consumption, and investment. It reveals how households learn about inflation when real quantities are also affected. Additionally, I include two other shocks: oil news shocks (Känzig, 2021) and monetary shocks (Bauer and Swanson, 2023). They are more cleanly identified. But as I will show, they are also special in terms of inflation learning.

In Michigan Survey of Consumers, consumers are asked their inflation estimate in the next 12 months.[13] However, I find that the consensus inflation expectation can be explained more, measured by R-squared, by the current inflation than the 1-year ahead inflation in the sample from 1978M1 to 2024M1. Indeed, Figure 9 in Appendix B shows that the R-squared reaches the maximum in the second month in the entire sample and monotonically decreases in horizon j for a variety of specifications in the sample excluding Volcker period and pandemic period.[14] Therefore, the survey inflation expectation is better served as the proxy for household inflation nowcast than 1-year ahead inflation expectation.

The main empirical strategy is to estimate impulse responses of inflation and household inflation expectation using the local projection method a là Jordà (2005). The specification is,

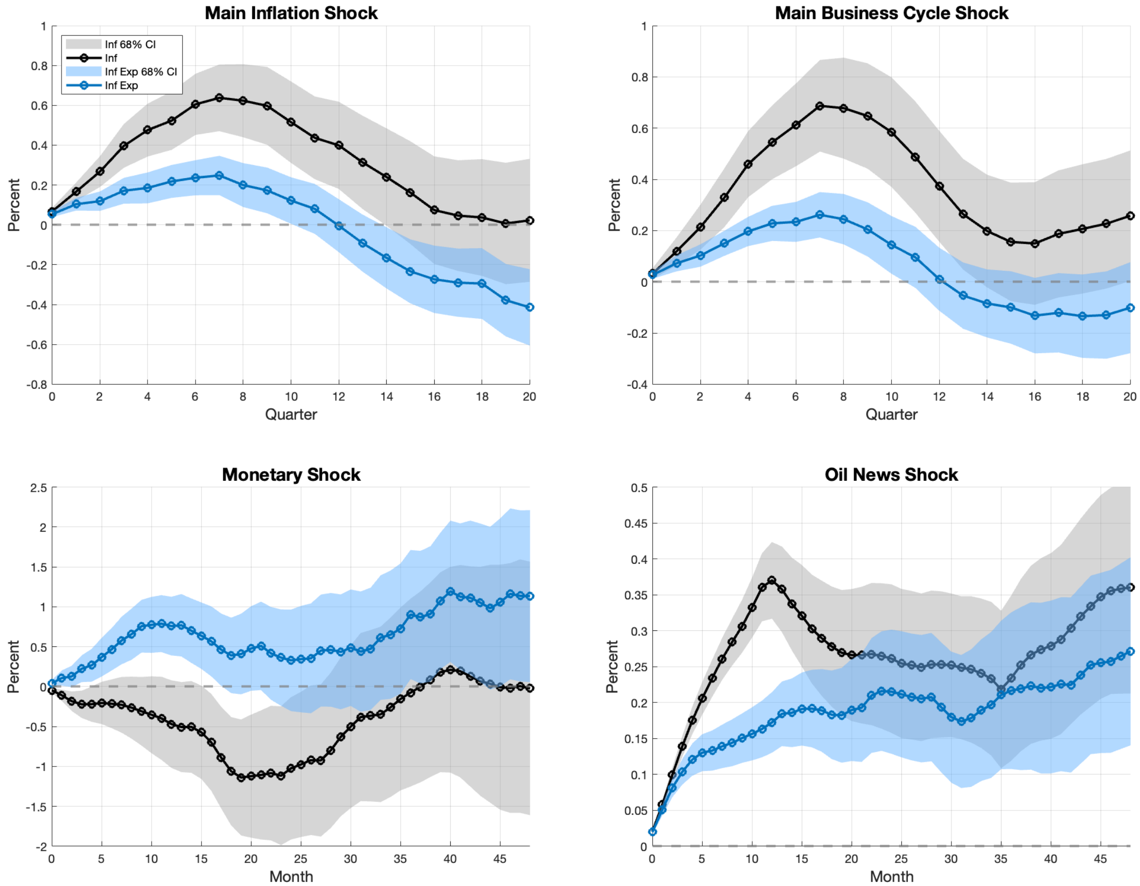
$$y_{t+h} = \alpha_h + \beta_h \varepsilon_t + \mathbf{\Gamma}' \mathbf{X}_t + u_{t+h} \tag{46}$$

where $\{\beta_h\}_{h=0}^{H}$ trace out the dynamic responses of the outcome. y_t is the cumulative inflation π_t and the cumulative consensus inflation expectation $\bar{E}_t \pi_t$ based on Michigan Survey of Consumers. X_t is a vector of controls. For main inflation shock and main business cycle shock, we run quarterly regression. The controls are 4 lags of shock itself, inflation, survey inflation expectation, 1-year government bond rate, log index of industrial production, and unemployment. For monetary shock and oil news shock, we run monthly regression. The controls are 12 lags of above variables. I construct standard errors for the coefficients that are heteroskedasticity and autocorrelation robust (HAC). All reported error bands are 68% confidence intervals. In Appendix B, I show the robustness of the impulse responses across different empirical specifications.

[13]The questions were: "During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?" and "By what percent do you expect prices to go up, on the average, during the next 12 months?"

[14]The specifications include OLS, polynomial regression, and non-parametric regression.

Figure 2: Impulse Responses of Inflation and Household Inflation Expectation



Notes: Dynamic responses: outcomes and forecasts following four types of shocks. The sample period is Q3 1979–Q4 2017. The main inflation shock and main business cycle shock are in quarterly frequency. The monetary shock and oil news shock are in monthly frequency. The black line represents the impulse responses of inflation and the blue line represents the impulse responses of inflation expectation. The shaded areas are 68% confidence intervals based on heteroskedasticity and autocorrelation robust standard errors and 12 lags. The x-axis denotes quarters or months from the shock, starting at 0. The y-axis denotes percent.

Figure 2 presents the impulse responses of cumulative inflation and inflation expectations with a 68% confidence interval. Focusing first on the case of the main inflation shock, the key takeaway is that learning from inflation occurs slowly. Inflation expectation consistently undershoots actual inflation, with the gap between cumulative responses widening over the first two years. In the third year, the gap narrows slightly as actual inflation overshoots into negative territory, a pattern also documented in Angeletos et al. (2020b). From the third year onward, the two responses become parallel, indicating that inflation expectations and actual inflation eventually converge. In the case of the main business cycle shock, the pattern is remarkably similar, although the inflation response is more pronounced and exhibits a greater negative overshoot in the third year. These patterns suggest that households may adopt a consistent approach to learning about inflation across these two scenarios.

The bottom-right panel presents the impulse responses following an oil news shock. The gap widens during the first year. Over the subsequent six months, the gap narrows as actual inflation undergoes a significant negative overshoot. Beyond this point, the two responses converge and overlap. This observation is consistent with Kumar et al. (2015), which highlights that households generally pay greater attention to oil prices. As a result, inflation driven by oil news shocks is more readily and rapidly recognized by households, facilitating a faster convergence between the two variables. To the left, I display the responses following a monetary shock. Here, inflation decreases while inflation expectations rise after a contractionary shock, suggesting that households may confuse monetary shocks with supply shocks. Liu and Zhang (2024) propose a model in which agents must infer whether the underlying shock is demand or supply based on endogenous signals. Agents may perceive positive inflation if they believe the shock is a negative supply shock as opposed to a contractionary monetary shock. As a result, to match such empirical impulse responses, we need additional bells and whistles in the model. I leave it to the future research.

Overall, we are confident that households learn inflation in a pretty consistent way over the business cycle. Therefore, I calibrate the signal-to-noise ratio of the signal about the inflation so that two impulse responses of the inflation and the inflation nowcast coincide approximately three years after the shock, which is a conservative calibration. At the same time, I assume that X_{it} completely uninformative, i.e., $\sigma_{xit} \to \infty$.[15]

2.2 Impulse Responses of the Calibrated Model

In this section, I discuss the impulse responses of the calibrated model. The calibrated model is computed using the frequency-domain methods based on Rondina and Walker (2021), Huo and Takayama (2023), and Han et al. (2022). The Taylor rule parameter is calibrated as $\phi = 1.5$. Figure 3 presents the results. The first row shows the impulse responses of output, interest rate, inflation and inflation nowcast following a 100 basis-point interest rate cut based on the calibrated model. The second row presents the impulse responses following a 100 basis-point reduction in

[15] One can easily find other combinations of the calibration of these two shocks to match the empirical findings. The calibration results are similar.

the aggregate productivity. In the third column, the blue line represents the inflation and the red line represents the inflation nowcast. Our theory predicts that the gap between these two lines is proportional to the output response. In addition, I compare our model with the standard NK model with the Phillips curve as follows,

$$\hat{\pi}_t = \kappa(\hat{y} - \hat{a}) + \beta E_t \pi_{t+1} \tag{47}$$

where $\kappa = (1 - \beta\psi)(1 - \psi)/\psi$ and ψ is the probability of a firm not being able to adjust its price, which we call Calvo parameter. Applying $E_t \pi_{t+1} = \rho \pi_t$,

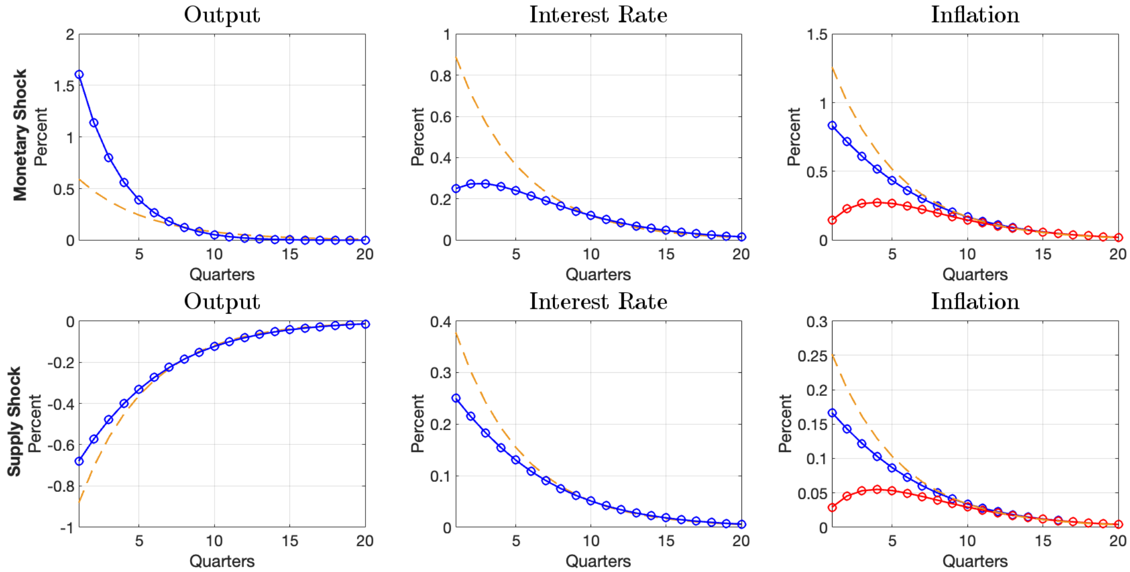
$$\hat{\pi}_t = \frac{1}{1 - \beta\rho} \kappa(\hat{y} - \hat{a}) \tag{48}$$

To show how big the impact of the information friction and learning channel, I impose the slope of Phillips curve as $\kappa = \frac{\Gamma}{1-\Gamma}$. Since $\Gamma = 0.3$, the implied value for Calvo parameter is 52.5%. The dashed orange line plots the impulse responses of the standard NK model with Calvo parameter 52.5% following two kinds of shocks.

There are three key implications about the monetary non-neutrality from the calibration analysis. First, with exogenously calibrated parameters Γ and $\{\theta_t\}$, the model exhibits substantial monetary non-neutrality. A 100 basis-point cut in the interest rate results in an initial output increase of approximately 160 basis points. This response is comparable to that in a standard NK model where the probability of not adjusting price being around 70%. This probability aligns with the value that is usually calibrated from a DSGE model (Christiano et al., 2005 and Smets and Wouters, 2007). This suggests that the degree of monetary non-neutrality generated by our mechanism is sufficient to explain and match the macro-level impulse-response evidence.

Second, the monetary non-neutrality generated by the standard NK model is three times smaller than that in our model when $\kappa = \frac{\Gamma}{1-\Gamma}$. The reason is that in the survey, household inflation expectation lags far behind the actual inflation. When we calibrate this evidence into the model instead of leveraging the knowledge imposed by the rational expectations, we achieve much larger monetary non-neutrality. To understand the magnitude of such difference, notice that the coefficient $1/(1 - \beta\rho)$ in (48) is the counterpart of $1/(1 - \theta)$ in our model. In practice, the discount factor β is close to one and the persistence of the shock is $\rho = 0.8$. Then, $1/(1 - \beta\rho) = 5$. In contrast, the implied information friction in the initial period from our calibration is $1/(1 - \theta) = 1.2$, making the slope of the Phillips curve about four times smaller in our model.

Figure 3: Impulse Responses of Inflation, Output and Interest Rate in the Calibrated Economy



Notes: The figure plots the impulse responses of output, interest rate, inflation and inflation nowcast following a 100 basis-point interest rate cut and a 100 basis-point decrease in the aggregate productivity based on the calibrated model. In the third column, the blue line represents the inflation and the red line represents the inflation nowcast. The dashed orange line represents the impulse responses generated from a standard NK model with the Calvo parameter 52.5%, i.e., 52.5% firms cannot adjust prices in each period. The x-axis is quarters. The y-axis is percent.

Third, since the output response depends on the gap between actual inflation and the inflation nowcast, rapid learning can close this gap before the shock fully dissipates. Consequently, the persistence of the output response is endogenously smaller than the persistence of the monetary shock. This is consistent with the empirical evidence. As shown in Figure 3 of Bauer and Swanson (2023), following a monetary shock, industrial production reaches its trough in a year, while the CPI continues to decline and approximately reaches its trough after 40 months. In addition, Figure 8 in Appendix B shows the impulse responses of unemployment rate and inflation following a main business cycle shock. The unemployment responses peaks earlier than inflation responses. The theory based on household expectation and slow learning provides an endogenous explanation for less persistent output responses.

Now, I turn to supply shock. Following a 100 basis-point reduction in the aggregate productivity is rather not surprising, the output declines and the inflation increases by less than 100 basis points. This occurs because prices remain below the full-information benchmark, which sustains aggregate demand above the natural output level and the higher labor supply in equilibrium. The persistence of the output response inherits the shock's persistence, when the gap between inflation and inflation nowcast narrows to zero. This differs from the case of a monetary shock, as the aggregate productivity shock directly impacts output.

Overall, our mechanism can generate substantial monetary non-neutrality, comparable in magnitude to the calibrated Calvo parameter commonly used in DSGE literature. The model sheds light on how the survey data can be directly linked to the output gap. The persistence of the output response is endogenously determined, with faster learning leading to lower persistence.

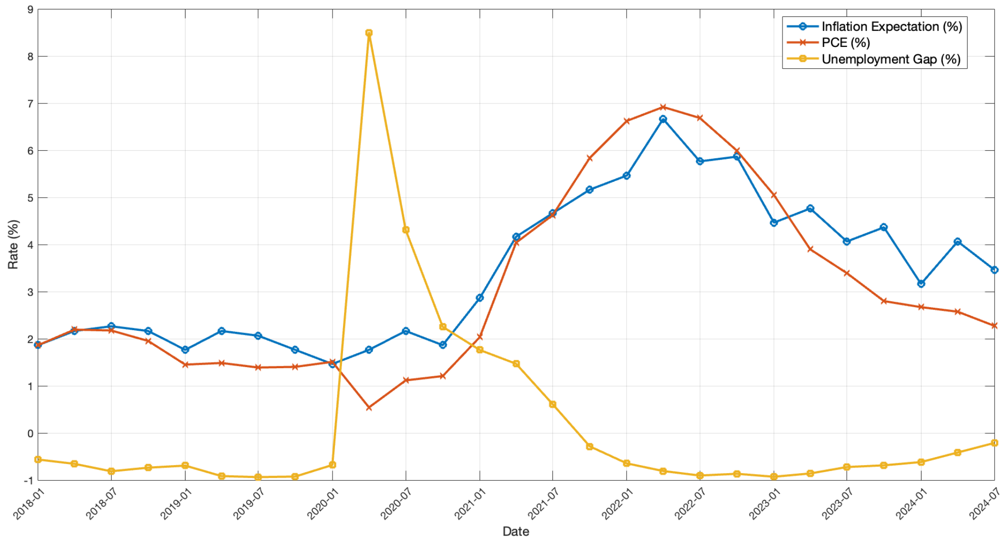
Applications – I consider two applications of the model: (i) anticipated VAT reforms, and (ii) post-pandemic steepened slope of Phillips curve.

First, Buettner and Madzharova (2021) find a rapid passthrough of anticipated VAT-induced cost changes to prices within four months—two months before and two months after the reform as shown in their Figure 4. This finding contradicts the predictions of the standard NK model, which suggests that firms would begin raising prices at least one year prior to the reform. In contrast, my model predicts no price changes before the reform and full passthrough upon implementation, as these tax reforms are common knowledge to consumers, resulting in minimal price stickiness. In practice, additional factors influence price dynamics: intertemporal substitution of durable goods tends to elevate prices before the reform, while delays in cost passthrough within the supply chain can defer price increases until after implementation.

Second, two puzzles have emerged in the dynamics of inflation and unemployment rate in the post-pandemic era. First, the supply-induced inflation surged immediately after the pandemic and remained persistent, contrary to the prediction of a flat Phillips curve estimated using pre-COVID data (Hazell et al., 2022). Second, despite the Fed's aggressive interest rate hikes beginning in August 2022, demand-driven inflation, caused by fiscal stimulus and labor market tightness (Blanchard and Bernanke, 2023), declined rapidly without a substantial rise in unemployment.

Theories have been proposed to address these puzzles. Benigno and Eggertsson (2023) introduces a non-linear New Keynesian Phillips curve in which labor becomes significantly more costly when the vacancy-to-unemployed ratio surpasses a critical threshold. However, as supply shortages have eased, this framework struggles to explain the second puzzle. Instead, I propose a unified explanation for both puzzles through the information channel in our model. The historically high media coverage and public discussion of supply chain disruptions, elevated inflation, and labor

Figure 4: Post-Pandemic Unemployment Rate, Inflation and Fed Funds Rate



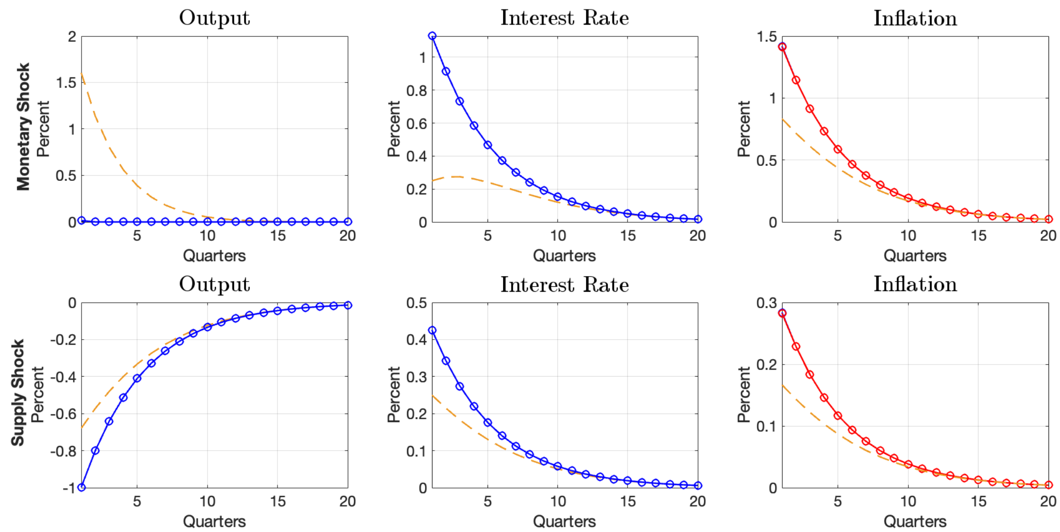
Notes: The figure shows the time series of unemployment gap measured by unemployment rate minus the noncyclical rate of unemployment (NROU), PCE inflation and mean inflation expectation based on Michigan survey from Jan 2018 to July 2024. The x-axis is date. The y-axis is rate.

shortages likely reduced information asymmetry, enabling households to learn more quickly about the rising costs faced by firms.

Figure 4 illustrates the dynamics of the unemployment gap, PCE inflation, and inflation expectations based on the Michigan Survey of Consumers. To eliminate the level effect, I shift inflation expectations to coincide with actual inflation in January 2018 (D'Acunto et al., 2022). It is evident that inflation expectations closely track actual inflation during this period. In contrast, the unemployment gap remains small and stable, consistent with our theory that a small difference between inflation and inflation expectations results in a small output gap. This aligns with the findings of Beaudry et al. (2024), who argue that household inflation expectations are crucial for explaining recent inflation dynamics. Moreover, it is specifically household inflation expectations that matter.

Figure 5 illustrates the responses of inflation and output when information in the model is highly precise. The orange dashed line represents the impulse responses based on our benchmark calibrated model. In both types of shocks, inflation and inflation expectations closely overlap, as observed in Figure 4. First, the second row presents the impulse responses following a supply shock. The increase in inflation is approximately double that of the benchmark case, which aligns with

Figure 5: Impulse Responses of Inflation and Household Inflation Expectation



Notes: The figure shows the impulse responses of output, interest rate, inflation and inflation nowcast following a 100 basis-point decrease in the aggregate productivity in a calibrated model with a large signal-to-noise ratio. The orange dashed line represents our benchmark calibrated model. In the third column, the blue line represents the inflation and the red line represents the inflation nowcast. The x-axis is quarters. The y-axis is percent.

the drastic inflation surge at the onset of the supply chain disruptions. However, the output and unemployment rate did not decline significantly in reality compared to the model's predictions, likely due to the concurrent fiscal stimulus. Second, the first row of Figure 5 shows the impulse responses to a demand shock under precise information. The output response is negligible, while inflation declines as the interest rate rises, which explains the second puzzle.

Policy Implication – 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. This has important implications for the conduct of monetary policy. 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.

2.3 Extensions

I discuss two extensions. In the first 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 10. *Each firm charges same markup over all the products it sells.*

$$P_{kj} = \frac{e_k}{e_k - 1} \frac{W}{A_{kj}} \tag{49}$$

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

$$P_k = \frac{e_k}{e_k - 1} \frac{W}{A_k}; e_k = \lambda \frac{\int X\big(g\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x)}{\int X\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x))} + 1 \tag{50}$$

where P_k is the CES price index of P_{kj}. The passthrough of product-level productivity shocks a_{kj} increases toward one when the number of products increases.

This extension 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.[16] 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.

In the second extension, I generalize the setup to accommodate finite labor supply elasticity. The details of setup and equilibrium are delegated to Appendix B. I present the following result.

Proposition 11. *The Phillips curve, when the elasticity of labor supply is η, is given by,*

$$\hat{p} = (1 + \eta) \frac{\Phi}{1 + \eta\upsilon - \Phi} (\hat{c} - \frac{1 + \eta - \eta\upsilon}{1 + \eta} \hat{a}) \tag{51}$$

where υ is defined in the Appendix B.

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

From Proposition 5, shoppers are more likely to search for all firms. They are more likely to be drawn to high-productivity, low-price firms, leading to an expansion in demand for these firms and a contraction for low-productivity firms. The parameter υ captures the first-order effect of demand reallocation on the aggregate productivity. To understand the intuition, suppose the effect of demand reallocation is absent, i.e., $\upsilon = 0$. Then $\hat{p} = (1+\eta)\frac{\Phi}{1-\Phi}(\hat{c} - \hat{a})$ and the price index response is larger if $\eta > 0$, because with finite labor elasticity, the real marginal cost increases. Now consider a monetary shock (similarly for a supply shock) with $\upsilon > 0$. Compared to the $\upsilon = 0$ case, the price index response is smaller because demand reallocation toward high-productivity firms effectively raises aggregate productivity, mitigating the rise in the price index and enhancing the output response.

3 Empirics

In this section, I provide empirical support for the key mechanism of the model: unanticipated inflation increases search activities. To support this logic, we must establish two ingredients in the data: (i) the presence of information frictions and (ii) the correct direction of the reaction of search behavior. The former is addressed in Section 2.1. Here, we focus on the second component.

In particular, I utilize a detailed consumer panel dataset that includes information on households' shopping trips, spending, store choices, and demographic characteristics to examine whether higher inflation is associated with increased measures of search behavior. My contribution to the literature lies in developing a novel measure of search activities that more accurately captures search efforts and aligns more closely with the predictions of standard sequential search models.

3.1 Search Activities and Unanticipated Inflation

I first briefly describe the data, then propose our measure of search activity, and lastly I present the results.

Data – The data source is the NielsenIQ Consumer Panel Data set.[17] The sample period is 2006 Q1 - 2019 Q4. NielsenIQ tracks the shopping behavior of average 55,000 households every

[17]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.

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 30 quarters 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 all expenditure on goods in the CPI.[18] 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. I only consider households live in the Metropolitan Statistical Area.[19]

Measurement − 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

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

[19] People living in the countryside may exhibit different search behaviors compared to those in urban areas. They are more likely to engage in "one-stop" shopping due to limited store availability. Their lower search activity may not indicate a lack of willingness to search but rather reflect physical constraints imposed by the environment.

all the product groups. 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 10 in Appendix B shows the time trend of the non-routine share of spending. It is highly correlated with the average shopping time derived from American Time Use Survey. 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 35.6%. They make about 3.5 shopping trips per week, visit approximately 12 different 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 substantial variance.

Table 3: Descriptive Statistics of Household Search Behavior

	Mean	S.D.	10th Percentile	90th Percentile	Observations
Non-routine Share(%)	35.6	25.2	3.23	71.1	2,846,354
Number of Trips	40.5	29.3	11	80	2,846,354
Number of Distinct Retailers	11.7	7.27	4	22	2,846,354

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 – Following the literature, I assume that consumers use historical data to forecast the current inflation. In particular, consumers estimate a simple OLS regression of inflation on four lags of the inflation for food and drinks and unemployment rate. The residual from this regression is the unanticipated inflation, which has a mean of -7.8 bp and a standard deviation of 51 bp in the sample period. I focus on inflation for food and drinks because the majority of NielsenIQ products fall into this category.[20] Additionally, I consider the measure based on the

[20] I do not use the main inflation shock since it is not correlated with the inflation in food and drinks. If any, there is insignificant negative relationship between these two time series. See Appendix C for details.

inflation for overall goods and services. The unanticipated inflation in this case has a mean of -7.5 bp and a standard deviation of 76 bp. I normalize the two series of unanticipated inflation, so their units are standard deviation in our sample. Appendix C provides further details of regression and discusses several robustness checks.

Impact of Unanticipated Inflation on Consumers' Search Behavior – To assess the changes in search behavior after an unanticipated inflation shock, our baseline specification is:

$$y_{it+1} = \lambda_i + \beta\tilde{\pi}_t + X_{it} + e_{it} \tag{52}$$

where t is time; i represents consumer. λ_i is the consumer 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 consumer controls. These controls include consumer age, employment, education, marital status, having children or not, and consumer i's total spending in time t. As pointed out by Aguiar and Hurst (2007), these variables have large affect on the pattern of shopping behavior.

The results are presented in Table 4. The first column indicates that one standard deviation (51 bp) increase in unanticipated food and drink inflation leads to a 26.5 bp increase in the non-routine share of spending. This suggests that consumers respond to higher prices by engaging in more active search, shifting purchases to stores outside their routine stores. The magnitude of the response is modest, which is about a half (26.5 bp/51 bp). Given the information friction estimated in (46), small magnitude may indicates large search frictions.

The second column uses the unanticipated overall inflation. It shows a much smaller increase in the non-routine share and not statistically significant. The effect is smaller because the NielsenIQ data primarily covers food and beverages, and search behavior is more sensitive to inflation in these sectors. The third and fourth columns introduce controls for the number of shopping trips and the number of distinct stores visited. The coefficients decrease only slightly, suggesting these variables capture a very limited portion of search effort. This suggests that consumers allocate most of their search efforts to substituting products within the same categories across the stores they have already visited, rather than increasing trips or visiting new stores.

Table 4: Non-routine Share of Spending and Unanticipated Inflation

Dep. var.: Non-routine Share	(1)	(2)	(3)	(4)
Unanticipated inflation (F&D)	0.265***		0.253***	0.234***
	(0.023)		(0.023)	(0.023)
Unanticipated inflation (overall)		0.022		
		(0.018)		
Number of trips			0.033***	-0.028***
			(0.002)	(0.003)
Number of distinct stores				0.327***
				(0.011)
Observations	2,660,735	2,660,735	2,660,735	2,660,735
Consumer fixed effect	√	√	√	√
Consumer varying effect	√	√	√	√

Notes: The table reports the estimates in specification (52). Each observation is at the consumer×quarter level covering from 2006 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 consumer level. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

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. However, we do not observe the time spent in each shopping trip. 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 pass-through could further dampen the observed relationship between inflation and non-routine spending.

Overall, this evidence supports a key aspect of the main mechanism: as prices rise after an aggregate shock, consumers are incentivized to search for alternatives. The evidence indicates the response of search activities to unanticipated inflation is statistically significant. There are potentially two reasons for the magnitude of the response being modest. One is that the search friction is large, and another is that the estimate is biased downwards.

4 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 NK model. Future research can investigate the effect of fiscal policy in this model.

This paper further presents a Phillips curve that relates the output gap to household nowcast error of inflation. It emphasizes the role of household expectations in determining the slope of Phillips curve. After calibrating the model to moments drawn from the literature and empirical evidence, the dynamic model generates substantial monetary non-neutrality, highlighting the importance of consumer-side mechanisms in macroeconomic dynamics. In addition, empirical findings support for the model's key mechanism. I also document a secular decline in search activity and connect this trend to the observed secular decline in the slope of the Phillips curve.

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. Given the parallels in the literature between search behavior in goods and labor markets, this extension might be straightforward. Workers' incomplete information regarding the average posted wage could influence their job-search decisions, prompting firms to adjust wage-setting and potentially generating wage stickiness. Finally, to better calibrate the model and assess the quantitative importance of the mechanism, a micro-foundation for search costs is necessary. A spatial and industrial organization model would be a promising candidate.

Bibliography

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

Alvarez, F. E., Lippi, F., and Paciello, L. (2016). Monetary shocks in models with inattentive producers. *The Review of Economic Studies*, 83(2):421–459. Published: 29 October 2015.

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., Collard, F., and Dellas, H. (2020a). Business-cycle anatomy. *American Economic Review*, 110(10):3030–3070.

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

Angeletos, G.-M., Huo, Z., and Sastry, K. A. (2020b). Imperfect macroeconomic expectations: Evidence and theory. *NBER Macroeconomics Annual*, 35(1):1–481.

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

Angeletos, G.-M. and Lian, C. (2018). Forward guidance without common knowledge. *American Economic Review*, 108(9):2477–2512.

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.

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

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.

Benabou, R. (1988). Search, price setting and inflation. *The Review of Economic Studies*, 55(3):353–376.

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.

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.

Blanchard, O. J. and Bernanke, B. S. (2023). What caused the us pandemic-era inflation? Working Paper 31417, National Bureau of Economic Research.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2020). Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–2782.

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.

Burdett, K. and Menzio, G. (2018). The (q,s,s) pricing rule. *The Review of Economic Studies*, 85(2):892–928.

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.

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

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.

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.

Coibion, O. and Gorodnichenko, Y. (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–2678.

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

D'Acunto, F., Gorodnichenko, Y., and Coibion, O. (2022). The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives*, 36(3):157–184.

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.

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.

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

Gabaix, X. and Graeber, T. (2024). The complexity of economic decisions. Working Paper 33109, National Bureau of Economic Research.

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).

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.

Golosov, M. and Lucas, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, 115(2):171–199.

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. †.

Han, Z., Tan, F., and Wu, J. (2022). Analytic policy function iteration. *Journal of Economic Theory*, 200:105395.

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.

Head, A., Liu, L. Q., Menzio, G., and Wright, R. (2012). Sticky prices: A new monetarist approach. *Journal of the European Economic Association*, 10(5):939–973.

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.

Huo, Z. and Takayama, N. (2023). Rational expectations models with higher-order beliefs. *Available at SSRN 3873663*.

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

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182.

Känzig, D. R. (2021). The macroeconomic effects of oil supply news. *American Economic Review*, 111(4):1092–1125.

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

Kaplan, G., Menzio, G., Rudanko, L., and Trachter, N. (2019). Relative price dispersion: Evidence and theory. *American Economic Journal: Microeconomics*, 11(3):68–124.

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

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.

Liu, X. and Zhang, D. R. (2024). Confusion, phillips curves and de-anchored inflation. Working paper, Northwestern.

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

L'Huillier, J.-P. and Zame, W. R. (2022). Optimally sticky prices: Foundations. *Journal of Economic Dynamics Control*, 141:104397.

Mankiw, N. G. (1985). Small menu costs and large business cycles: A macroeconomic model of monopoly. *Quarterly Journal of Economics*, 100(2):529–537.

Matsuyama, K. and Ushchev, P. (2017). Beyond ces: Three alternative classes of flexible homothetic demand systems. Research Paper WP BRP 172/EC/2017, Higher School of Economics. Available at SSRN: `https://ssrn.com/abstract=3015279` or `http://dx.doi.org/10.2139/ssrn.3015279`.

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

Mondria, J., Vives, X., and Yang, L. (2021). Costly interpretation of asset prices. *Management Science*, 68(1):210–227.

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

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.

Reis, R. (2006). Inattentive producers. *The Review of Economic Studies*, 73(3):793–821. Published: 01 July 2006.

Rondina, G. and Walker, T. B. (2021). Confounding dynamics. *Journal of Economic Theory*, 196:105251. Available at: `https://doi.org/10.1016/j.jet.2021.105251`.

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

Rothschild, M. (1974). Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy*, 82(4):689–711.

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

Taylor, J. B. (1980). Aggregate dynamics and staggered contracts. *Journal of Political Economy*, 88(1):1–23.

Venkateswaran, V. (2014). Heterogeneous information and labor market fluctuations. Available at SSRN: https://ssrn.com/abstract=2687561 or http://dx.doi.org/10.2139/ssrn.2687561.

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

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

Woodford, M. (2003a). 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.

Woodford, M. (2003b). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, NJ.

Appendix

A Proofs and Derivations 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.

$$v^*(x) = -\frac{\kappa}{1-\Psi(v^*(x)|x)} + \frac{\int_{v^*(x)}^{\infty} y\psi(y)dy}{1-\Psi(v^*(x)|x)} \tag{53}$$

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 search cost is lower than the value of search when the agent has the worst value whatsoever, 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,

$$U(v|x) = -\kappa + U(v|x)\int_{-\infty}^{v}\psi(y|x)dy + \int_{v}^{\infty}U(y|x)\psi(y|x)dy \tag{54}$$

Equivalently,

$$U(v|x) = -\frac{\kappa}{1-\Psi(v|x)} + \frac{\int_{v}^{\infty}U(y|x)\psi(y|x)dy}{1-\Psi(v|x)} \tag{55}$$

It is easy to show that

$$U'(v|x) = \frac{\psi(v|x)}{(1-\Psi(v|x))^2}\left\{\int_{v}^{\infty}U(y|x)\psi(y|x)dy - U(v|x)(1-\Psi(v|x)) - \kappa\right\} = 0 \tag{56}$$

The second equation holds due to (55). 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 $v < v^*$,

$$-\kappa + U(v|x)\int_{-\infty}^{v}\psi(y|x)dy + \int_{v}^{\infty}U(y|x)\psi(y|x)dy = -\kappa + v^*(x)\Psi(v^*(x)) + \int_{v^*(x)}^{\infty}y\psi(y|x)dy = v^*(x) \tag{57}$$

The search problem in (11) is therefore simplified to

$$U(v|x) = \max\{v, v^*(x)\} \tag{58}$$

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 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 $v^*(x) = -\infty, \forall x$.

Now, we prove the second part: $v^*(x)$ decreases in x. First, notice that the threshold is determined alternatively by the following,

$$\int\int_{\lambda(v^*(x)+p)}^{\infty}\big(\frac{1}{\lambda}\epsilon - p - v^*(x)\big)g(\epsilon)d\epsilon f(p|x)dp = \kappa \tag{59}$$

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

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

■

Proof of Proposition 2:

Proof. We take first-order condition of firm's profit in (17) with respect to P_k. It is easy to show that we can always express the pricing strategy as in (18) and define the elasticity of demand as:

$$e_k = -\frac{\partial \log D(P_k)}{\partial \log P_k} = -\frac{\partial\Big(\int X\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big)d\Phi_x(x)\Big)}{\partial p_k} + 1 \tag{60}$$

■

Proof of Lemma 1:

Proof. We are interested in the case in which consumers search actively. 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 equilibrium price is a function of $v^*(w)$ and a_{kt}. I now rewrite the LHS of (20) in terms of the integral over productivity distribution, which is exogenous,

$$\int\int_{\lambda(v^*(w)+p^*(v^*(w),a))}^{\infty}\big(\frac{1}{\lambda}\epsilon - p^*(v^*(w),a) - v^*(w)\big)g(\epsilon)d\epsilon\phi_a(a)da = \kappa \tag{61}$$

where $\phi_a(a)$ is the pdf of productivity distribution. Fix $p^*(v^*(w),a)$, 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 $\frac{\partial p(v^*(w),a_k)}{\partial v^*(w)}$. From now on, we use $x_k(v^*(w))$ to denote $x(v^*(w),a_k)$ for any variable x.

$$\frac{\partial p_k(v^*(w))}{\partial v^*(w)} = \frac{\partial mu_k(v^*(w))}{\partial v^*(w)} = -\frac{1}{e_k(v^*(w))(e_k(v^*(w))-1)}\frac{\partial e_k(v^*(w))}{\partial v^*(w)} \tag{62}$$

It is easy to show that

$$\frac{\partial e_k(v^*(w))}{\partial v^*(w)} = \lambda^2\frac{g'(1-G)+g^2}{(1-G)^2}\left(1+\frac{\partial p_k(v^*(w))}{\partial v^*(w)}\right) \tag{63}$$

Combine the above equation with (62), we have:

$$\frac{\partial p_k(v^*(w))}{\partial v^*(w)} = -\frac{\Lambda_k}{1+\Lambda_k} \tag{64}$$

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 (20) decreases in $v^*(w)$. ■

Proof of 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 1, we know that there exists a unique solution $v^*(w)$ to 20 given that the prices are computed through the first-order conditions as in (23). The equilibrium in which shoppers search actively exists and is unique because $v^*(w)$ exists and is unique.

As Diamond (1971) points out famously, there are always a continuum of equilibria where shoppers do not search and firms charge very high prices, i.e., $v^*(w) = -\infty$ and $P_k = \infty, \forall k$. ■

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 into 20 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 1. Vice versa.*
5. *Loop the procedure 1-4 until the difference between the LHS and the RHS of 20 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,

$$\int \int_{\lambda(v^*(w')+p)}^{\infty} \big(\frac{1}{\lambda}\epsilon - p - v^*(w')\big) g(\epsilon) d\epsilon f(p - \Delta w|w') dp = \kappa \tag{65}$$

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

$$\int \int_{\lambda(v^*(w')+\Delta w+z)}^{\infty} \big(\frac{1}{\lambda}\epsilon - z - (\Delta w + v^*(w'))\big) g(\epsilon) d\epsilon f(z|w') dz = \kappa \tag{66}$$

This implies that $v^*(w') = v^*(w) - \Delta w$. According to (23), the elasticity of demand is given by,

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

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

$$p'_k = \log(\frac{e_k}{e_k - 1}) + w' - a_k = p_k + w' - w = p_k + \Delta w \tag{67}$$

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 14 in terms of integrating over the exogenous productivity distribution,

$$\int\int\int_{\lambda(v^*(x)+p^*(a,w))}^{\infty} \big(\frac{1}{\lambda}\epsilon - p^*(a,w) - v^*(x)\big) g(\epsilon) d\epsilon \phi_a(a) da h(w|x) dw = \kappa \tag{68}$$

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 shopper's expected price conditional on x as shown in (**??**), on the first order, the above equation becomes,

$$\int\int_{\lambda(v^*(x)+\bar{p}_k+\varphi(a)E(\hat{w}|x))}^{\infty} \big(\frac{1}{\lambda}\epsilon - \bar{p}_k - \varphi(a)E(\hat{w}|x) - v^*(x)\big) g(\epsilon) d\epsilon \phi_a(a) da = \kappa \tag{69}$$

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

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

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

$$\frac{dv^*(x)}{dE(\hat{w}|x))} = -\int \varphi(a)\bar{\omega}(a)\phi_a(a) da = -\int \varphi_k \bar{\omega}_k dk \tag{70}$$

where $\bar{\omega}(a)$ is the expenditure share of firms with productivity a in the full-information equilibrium where $w = \bar{w}$. 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 (19), the elasticity becomes,

$$e_k = \lambda \frac{\int Xg\big(\lambda(v^*(\bar{w}) - \Phi E(\hat{w}|x) + p_k)\big) d\Phi_x(x)}{\int X\big(1 - G\big(\lambda(v^*(\bar{w}) - \Phi E(\hat{w}|x) + p_k)\big)\big) d\Phi_x(x)} + 1 \tag{71}$$

with some abuse of notation, Φ is the aggregate total passthrough and $\Phi_x(x)$ is the cdf of information sets. Let $y = \lambda\Phi\big(E(\hat{w}|x) - \bar{E}(\hat{w})\big)$ and its pdf is $\phi_y(y)$, which is a Gaussian distribution with mean zero and standard deviation $\sigma = \lambda\Phi\theta\sigma_s$. Further, we denote $z = \lambda(v^*(w) - \Phi\bar{E}(\hat{m}) + p_k)$. The elasticity is rewritten as follows,

$$e_k = \lambda \frac{\int Xg(z+y)\phi_y(y)dy}{\int X\big(1 - G(z+y)\big)\phi_y(y)dy} + 1 \tag{72}$$

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 $X \propto \exp(y + \bar{w})$,

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

Therefore, $\int g(z+y)\phi_y(y)dy \rightarrow g(z)$ on the order of σ_s^2, which is second-order term. Similarly, $\int \big(1 - G(z+y)\big)\phi_y(y)dy \rightarrow 1 - G(z)$ on the order of σ_s^2. On the other hand, the average expectation of nominal wage shock $\bar{E}(\hat{w})$ approaches zero on the order of σ_s. Therefore, the elasticity, on the first order, is given by,

$$e_k = \lambda \frac{g(z)}{1 - G(z)} + 1 = \lambda \frac{g(\lambda(v^*(w) - \Phi\bar{E}(\hat{w}) + p_k))}{1 - G(\lambda(v^*(w) - \Phi\bar{E}(\hat{w}) + p_k))} + 1 \tag{73}$$

■

Proof of Lemma 2:

Proof. First, log price is given by,

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

Total differentiate on both sides,

$$\hat{p}_k = \frac{\partial \log \mu_k}{\partial p_k}\hat{p}_k + \frac{\partial \log \mu_k}{\partial p}\hat{p} + \hat{w} \tag{74}$$

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,

$$\hat{p}_k = \gamma_k \hat{w} + \xi_k \hat{p} \tag{75}$$

Since $\hat{p} = \Phi \hat{w}$, the total 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 of Incompleteness . First, notice that for any function $f(w)$, the first-order approximation to its derivative is,

$$f'(w) = f'(\bar{w}) + f''(\bar{w})\hat{w} \tag{76}$$

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 $f(w)$,

$$f(w) = f(\bar{w}) + f'(\bar{w})\hat{w} \tag{77}$$

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, $z = z(v^*(\bar{w}) + \bar{p}_k)$ for any function z. It is straightforward to show that, on the first order,

$$\frac{\partial e_k}{\partial p_k}\Big|_{\hat{w}=0} = \lambda^2 \frac{g'(1-G) + g^2}{(1-G)^2} \tag{78}$$

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

$$\frac{\partial e_k}{\partial p}\Big|_{\hat{w}=0} = -\theta\lambda^2 \frac{g'(1-G) + g^2}{(1-G)^2} \tag{79}$$

Then, we have $\frac{\partial \log \mu_k}{\partial p} = -\theta \frac{\partial \log \mu_k}{\partial p_k}$. According to Lemma 2, we have

$$\xi_k = \theta(1 - \gamma_k) \tag{80}$$

I claim that $\varphi_k < 1, \forall k$ and notice that to prove it, it is sufficient to prove $\gamma_k + \xi_k < 1, \forall k$. We can write out $\gamma_k + \xi_k$,

$$1 - (\gamma_k + \xi_k) = (1 - \gamma_k)(1 - \theta) \tag{81}$$

Since $-\frac{d \log \mu_k}{dp_k} > 0$, $\gamma_k < 1$. ■

Before turning to the result on "Composition", I first prove the following lemma,

Lemma A-1. *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, we denote $g(y_k) = g(\lambda(v^*(\bar{w}) + \bar{p}_k))$ and $G(y_k) = G(\lambda(v^*(\bar{w}) + \bar{p}_k))$. We 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, we have $\frac{1-G}{g} \to 1$ as $y \to \infty$ and therefore $\frac{1-G}{g} > 1$.

Second, we write out markup elasticity,

$$\Lambda = \frac{\frac{g'}{g}\frac{1-G}{g} + 1}{\frac{1}{\lambda}\frac{1-G}{g} + 1} \tag{82}$$

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 (82), we get $\Lambda \to 0$ when $y \to \infty$. Second, since $\frac{1-G}{g} \to \infty$ when $y \to -\infty$, $\Gamma \approx \lambda g'/g \to \infty$ since $g'/g \to \infty$. ■

Next, we turn to the result on composition and comparative statics.

Proof of Composition and MCS. First, notice that to show φ_k decreases in a_k, it is sufficient to show that γ_k decreases in a_k according to (81).

From Lemma 2, $\gamma_k = (1+\Lambda_k)^{-1}$. Since Λ_k decreases in P_k, it implies γ_k increases in P_k. Since P_k decreases in a_k, it implies γ_k decreases in a_k.

Next, we prove the comparative statics results. We first consider the comparative statics on κ. From Lemma 1, an increase in the search cost κ implies lower threshold $v^*(\bar{w})$. According to Lemma A-1, the steady-state markup elasticity decreases in $v^*(\bar{w})$. This implies a uniform decrease in γ_k. As a result, φ_k decreases in κ for any k according to (81). 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, we consider the comparative statics on θ. An increase in the information friction, which means a decrease in θ, induces lower φ_k, according to (81). When $\theta \to 0$, $\Phi = \Gamma$ according to Lemma 2. When $\theta \to 1$, $\varphi_k = 1$ according to (81). ■

Proof of Proposition 4:

Proof. From (80), and Lemma 2, we have,

$$\hat{p}_k = \gamma_k \hat{w} + (1 - \gamma_k)\theta \hat{w} \tag{83}$$

$$= \gamma_k \hat{w} + (1 - \gamma_k)\bar{E}\hat{w} \tag{84}$$

Aggregate, we have

$$\hat{p} = \Gamma \hat{w} + (1 - \Gamma)\bar{E}\hat{w} \tag{85}$$

The aggregation requires every firm is individually rational, so they know (84). Suppose every shopper believe that firms and other shoppers are rational. Then every shopper believes the above condition holds. Shoppers' average expectation of $\hat{p}$ satisfies,

$$\bar{E}\hat{p} = \Gamma \bar{E}\hat{w} + (1 - \Gamma)\bar{E}^2\hat{p} \tag{86}$$

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 $\hat{w}$:

$$\hat{p} = \Gamma \sum_{h=0}^{\infty} (1 - \Gamma)^h \bar{E}^h \hat{w} \tag{87}$$

"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 5:

Proof. 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 $\lim_{a \to \infty} \mu_k = \infty$.

From Lemma 2, $\gamma_k = (1 + \Lambda_k)^{-1}$. Also, we know $\xi_k = \theta(1 - \gamma_k)$. Since γ_k decreases in a_k as shown in the proof of Theorem 3, 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 6:

Proof. See Proof of Theorem 3. ∎

Proof of Proposition 7:

Proof. See Proof of Theorem 3. ∎

Proof of Proposition 39:

Proof. In the case of monetary shock, recall the labor supply condition,

$$\hat{p} + \hat{c} = \hat{w} \tag{88}$$

Combine with the definition of the aggregate passthrough,

$$\hat{p} = \frac{\Phi}{1 - \Phi} \hat{c} \tag{89}$$

Plug in the condition in Proposition 4, we have

$$\hat{p} = \frac{\Gamma}{1 - \Gamma} \frac{1}{1 - \theta} \hat{c} \tag{90}$$

In the case of aggregate supply shock, two equations are as follows,

$$\hat{p} + \hat{c} = \hat{w}$$

$$\hat{p} = \Phi(\hat{w} - \hat{a})$$

Substitute $\hat{w}$, we achieve $\hat{p} = \frac{\Phi}{1-\Phi}(\hat{c} - \hat{a})$. Again, leveraging Proposition 4, we have the result.

∎

B Proofs and Calibration in Dynamic Model

Proof of Proposition 9:

Proof. We can rewrite worker's problem as:

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

It is easy to show that the standard Euler equation holds. Recall we have,

$$\hat{p}_t - \bar{E}\hat{p}_t = \frac{\Gamma}{1-\Gamma}(\hat{c}_t - \hat{a}_t) \tag{91}$$

Since we assume that households know $\hat{p}_{t-1}$, we have:

$$\begin{aligned}
(\hat{p}_t - \hat{p}_{t-1}) - (\bar{E}\hat{p}_t - \hat{p}_{t-1}) &= \frac{\Gamma}{1-\Gamma}(\hat{c}_t - \hat{a}_t) \\
\hat{\pi}_t - \bar{E}\hat{\pi}_t &= \frac{\Gamma}{1-\Gamma}(\hat{c}_t - \hat{a}_t)
\end{aligned}$$

■

Price-level Targeting – In this section, we show the dynamics with price-level targeting.

Proposition 12. *The equilibrium dynamics of $\{\hat{p}_t, c_t, i_t\}$ is described by the following system of equations:*

$$\begin{aligned}
\hat{c}_t &= E_t \hat{c}_{t+1} - (i_t - E_t \hat{\pi}_{t+1}) \\
\hat{p}_t - \bar{E}^s \hat{p}_t &= \frac{\Gamma}{1-\Gamma}(\hat{c}_t - \hat{a}_t) \\
i_t &= \phi_\pi \hat{p}_t + \varepsilon_{mt}
\end{aligned}$$

where Γ is the aggregate passthrough.

Figure 6 shows the impulse responses generated from the system in the above proposition. The dashed orange line again represents the responses of the standard NK model with Calvo Parameter of 70%. We see the results are similar, except the output responses are smaller, which is consistent with Woodford (2003b).

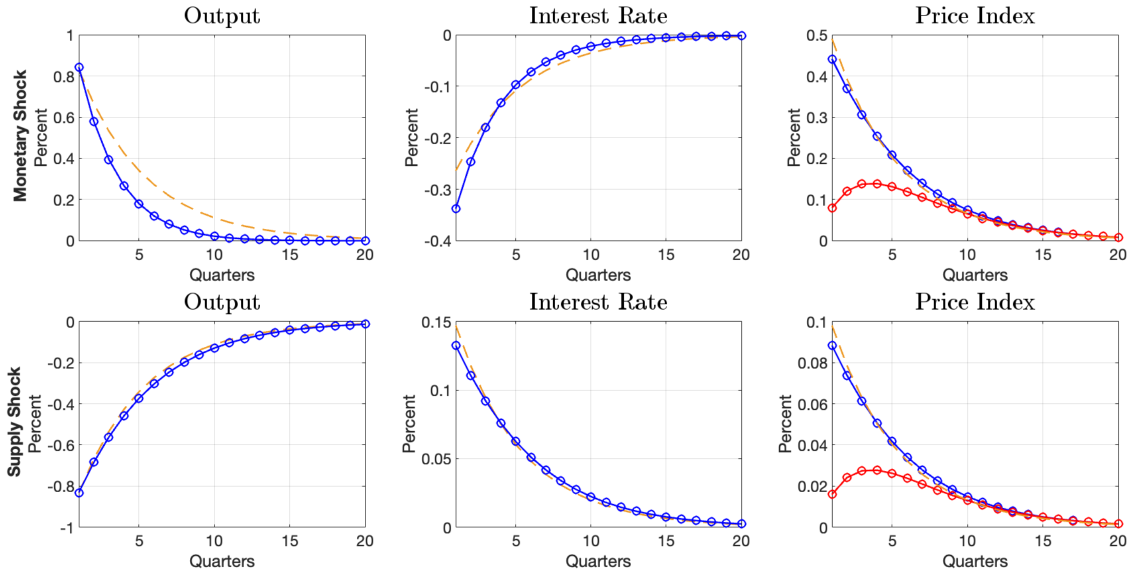


Figure 6: Dynamic Responses with Price-level Targeting

Dynamic Model without a Big-Household Assumption In this section, I follow the tradition by adding noises to the observables in household budget constraints to prevent them from knowing the underlying shock. By doing so, we do not need to divide household into a worker and a continuum of shoppers. I only briefly describes the model. Most of the setup is similar to the model laid out in the main text. There is a continuum of shoppers indexed by $i \in I$. In the morning, they make consumption-saving and labor supply decisions. In the afternoon, they go shopping with given consumption expenditure that they have allocated to consumption in the morning.

In the morning, the shopper maximizes the following discounted utility:

$$\max_{B_{it},X_{it},L_{it}} E_0 \sum_{t=0}^{\infty} \beta^t \left(\log \frac{X_{it}}{P_{kt}} + \frac{1}{\lambda}\epsilon_{ikt} - L_t \right)$$

$$\text{s.t.} \quad X_{it} + B_{it} = W_{it}L_{it} + R_{it-1}B_{it-1} + \Pi_{it}$$

where X_{it} is the consumption expenditure left for the shopping in the afternoon. Notice that W_{it} and R_{it} are shopper-specific and I assume $w_{it} = w_t + \zeta_{it}$, $i_{it} = i_t + u_{it}$, $\pi_{it} = \pi_t + v_{it}$, where i_t is the log-deviation of R_t following the convention. These shocks $\zeta_{it}, u_{it}, v_{it}$ are IID across i and t, independent of one another, and independent of any other random variable in the economy. Shocks follow normal distribution with their respective variances. They "noise up" the information that each shopper can extract from the available market signals. I assume that shoppers only treat the current-period price index $\hat{p}_t$ as their signal to align with the information structure in the main text.

In the afternoon, shoppers have the same search problem. The following proposition characterize the three-equation system,

Proposition 13. *The equilibrium dynamics of $\{\hat{p}_t, c_t, i_t\}$ is described by the following system of equations:*

$$
\begin{aligned}
\hat{c}_t &= -\sigma \left\{ \sum_{k=1}^{\infty} \beta^{k-1} \bar{E}_t^s \left[i_{t+k} - \hat{\pi}_{t+k+1} \right] \right\} + \frac{1-\beta}{\beta} \left\{ \sum_{k=1}^{\infty} \beta^k \bar{E}_t^s \left[\hat{c}_{t+k} \right] \right\} \\
\hat{p}_t - \bar{E}^s \hat{p}_t &= \frac{\Gamma}{1-\Gamma} (\hat{c}_t - \hat{a}_t) \\
i_t &= \phi_\pi \hat{p}_t + \varepsilon_{mt}
\end{aligned}
$$

where Γ is the aggregate passthrough.

The first equation represents incomplete-information Euler equation. Under incomplete information, households need to form expectations of future aggregate consumption and inflation, which in turn depends on how other households respond to the shock. This forms a dynamic beauty contest game as described in Angeletos and Lian (2018). The second equation presents the Phillips Curve as in the main text. The third equation represents the price-level targeting rule. It guarantees that this system will converges to its steady state as the shock dissipates. Otherwise, if we use the inflation-targeting rule, the price level will never return to its original value.

Proof of Proposition 10:

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

$$
\pi_k = \frac{1}{\rho} \sum_{j=1}^{N} \int X\big(1 - G\big(\lambda(v^*(x) + p_k)\big)\big) d\Phi_x(x) \frac{1}{P_{jk}} (P_{jk} - \frac{W}{A_{jk}}) (\frac{P_{jk}}{P_k})^{-\sigma} \tag{92}
$$

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,

$$
\pi_k = \frac{1}{\rho} \sum_{j=1}^{N} \omega(P_k) \frac{1}{P_{jk}} (P_{jk} - \frac{W}{A_{jk}}) (\frac{P_{jk}}{P_k})^{-\sigma} \tag{93}
$$

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,

$$
P_{jk} = \mu_k \frac{W}{A_{jk}} = \frac{\sigma}{\sigma - 1 - P_k^{-1} \tilde{\pi}_k (e_k - 1 + \sigma - 1)} \frac{W}{A_{jk}} \tag{94}
$$

where e_k is the same as in (19). 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,

$$\mu_k = \frac{e_k}{e_k - 1} \tag{95}$$

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.

$$\gamma_{jk} = \frac{\omega_{jk}}{\omega_k}\gamma_k \tag{96}$$

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

Proof of Proposition 11:

Proof. I generalize the static general equilibrium model to accommodate finite labor supply elasticity.

Due to finite labor supply elasticity, the nominal wage is now endogenously determined by the labor market clearing. The log-linearized labor supply condition for household i is $\eta\hat{l}_i = \hat{w}_i - \hat{x}_i$. Integrate and get aggregate labor supply,

$$\eta\hat{l} = \hat{w} - \hat{x} \tag{97}$$

The resource constraint is $C = AL$. However, the average productivity changes after the shock due to demand reallocation cross firms. For instance, after a monetary shock, shoppers search more actively and are more likely to be captured by the high-productivity and low-price firms. The aggregate productivity is given by,

$$\bar{A} = \left[\frac{\frac{1}{\exp(a)}\frac{1}{\exp(p(a,w))}\omega(a,w)d\Phi_a(a)}{\frac{1}{\exp(p(a,w))}\omega(a,w)d\Phi_a(a)}\right]^{-1} \tag{98}$$

On the first order, the aggregate productivity has two components. One is the shock to the aggregate productivity and another is the allocation-induced productivity change:

$$\hat{\bar{a}} = \hat{a} + \upsilon(\hat{w} - \hat{a})$$

where υ is the first-order effect of change in the nominal wage and aggregate productivity on the aggregate productivity. Therefore, combining resource constraint, we have,

$$\hat{l} = \hat{c} - (\hat{a} + \upsilon(\hat{w} - \hat{a})) \tag{99}$$

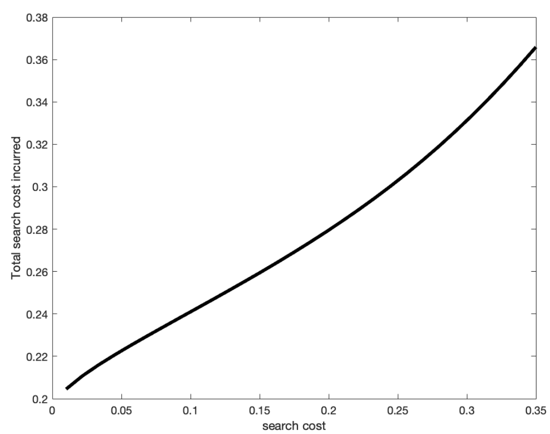
Combine $\hat{p} = \Phi(\hat{w} - \hat{a})$ and substitute out $\hat{w}$,

$$\hat{p} = (1+\eta)\frac{\Phi}{1+\eta\upsilon - \Phi}(\hat{c} - \frac{1+\eta-\eta\upsilon}{1+\eta}\hat{a}) \tag{100}$$

Suppose the effect of demand reallocation is absent, i.e., $\upsilon = 0$. Then $\hat{p} = (1+\eta)\frac{\Phi}{1-\Phi}(\hat{c} - \hat{a})$ and the price index response is larger if $\eta > 0$ because with finite labor elasticity, the real marginal cost increases. Now consider a monetary shock (similarly for a supply shock) with $\upsilon > 0$. Compared to the $\upsilon = 0$ case, the price index response is smaller because demand reallocation toward high-productivity firms effectively raises aggregate productivity, mitigating the rise in the price index. ■

Figure 7 – As we see, the total search cost increases with the search cost, even though the average number of searches goes down.

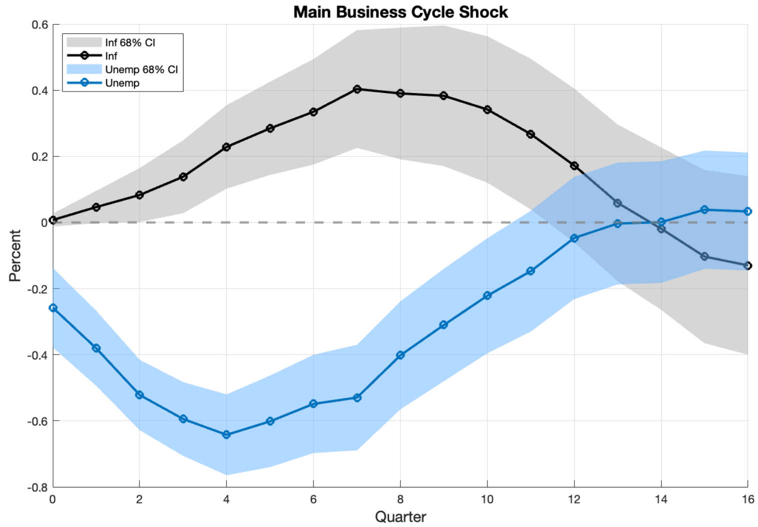
Figure 7: Total search cost and search cost



Notes: The figure plots the total search cost versus the search cost κ.

Figure 8 – I show how inflation and unemployment rate respond to the main business cycle shocks in our sample in Figure 8.

Figure 8: Impulse responses of inflation and Unemployment Rate



Notes: The figure plots the impulse responses of unemployment rate and inflation following a main business cycle shock.The sample period is Q3 1979–Q4 2017. The black line represents the impulse responses of inflation and the blue line represents the impulse responses of unemployment rate. The shaded areas are 68% confidence intervals based on heteroskedasticity and autocorrelation robust standard errors and 12 lags. The x-axis denotes months from the shock, starting at 0. The y-axis denotes percent.

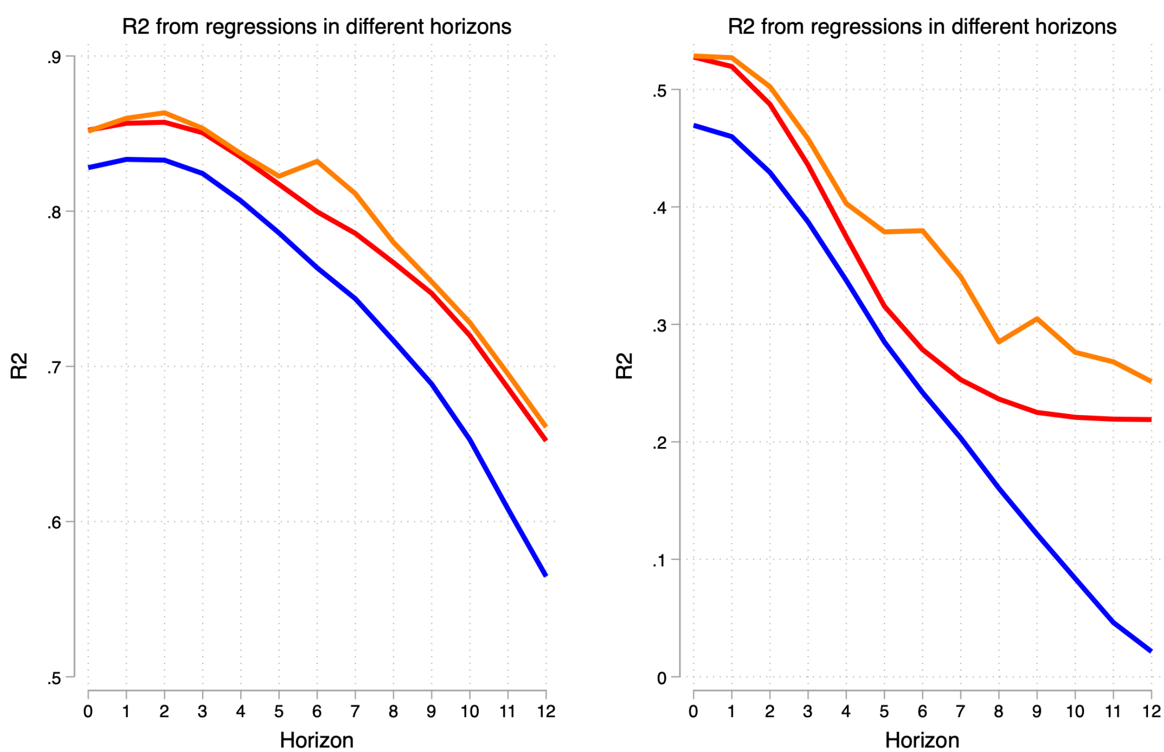
Survey Inflation Expectation – I show that inflation expectation based on Michigan Survey of Household serves better as the inflation nowcast instead of the inflation forecast. I run the following regression,

$$y_t = \beta f(x_{t+h}) + \epsilon_t \tag{101}$$

where y_t is the survey inflation expectation at month t; x_{t+h} is the actual inflation at month $t + h$. f is the function we pick for different specifications: OLS, polynomial, non-parametric. We are interested in the R2 of this regression, i.e., how much variation in survey inflation expectation can be explained by the actual inflation in the future horizons. If households respond by forecasting the inflation in the next year, we should expect highest R2 when $h = 12$. In addition, the R2 will increase over the horizons. However, Figure 9 shows the opposite. In the left panel, we use the whole sample ranging from 1978M1 to 2024M1. The blue, red and orange lines represent OLS, polynomial, non-parametric specifications respectively. We find that the R2 is largest when $h = 2$ or $h = 3$, and it is decreasing after $h = 3$. In the right panel, we use the shorter sample ranging from 1984M1 to 2019M12. This short sample cuts off the high-inflation episodes. The

R2s are significantly lower in this sample, consistent with the rational inattention model in which households pay more attention to inflation when variance of inflation is larger. The R2 is largest for the current inflation. Combining both panels, we are confident that survey inflation expectation serves as a good proxy for inflation nowcast. The possible explanation is that households may not understand what the question means. They just report their guesses of current inflation.

Figure 9: R2 in Different Samples and Horizons

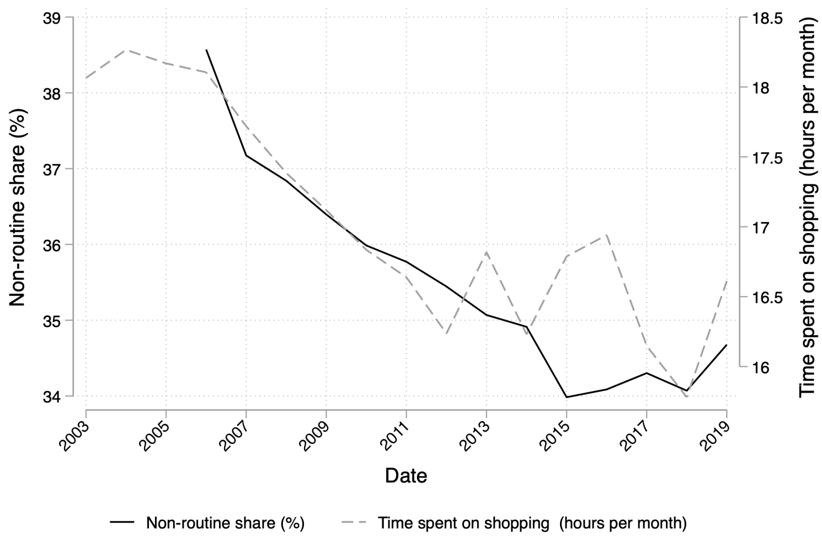


Notes: The blue line represents the R-square using OLS. The red line uses polynomial regression to third order. The orange line uses the nonparametric regression with Epanechnikov kernel. The left panel uses the whole sample: 1978M1 - 2024M1. The right panel uses the subsample: 1984M1 - 2019M12.

C Empirics on Information Asymmetry

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

Figure 10: 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.

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.[21] Figure 10 shows the results.

I first show that the main inflation shock exhibits an insignificant negative correlation with inflation in food and drinks. Regressing inflation in food and drinks on the main inflation shocks and four lags of food and drinks inflation over the sample period 1979 Q3 to 2017 Q4 yields a coefficient of -0.12 with a p-value of 0.864. In contrast, using overall inflation instead of food and drinks inflation results in a coefficient of 0.3, which is highly significant. This suggests that the main inflation shocks capture most of the variation in overall inflation but not in the inflation of food and drinks. Therefore, the main inflation shock is not appropriate for our empirical analysis based on NielsenIQ consumer panel.

Now, I show the robust checks if we assume alternative models households use to derive the predicted inflation. Here, I consider the case where households only use 4 lags of inflation in food

[21] 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).

and drinks to predict inflation. The standard deviation of the unanticipated inflation is 52 bp. The following table shows the result.

Table 5: Non-routine Share of Spending and Unanticipated Inflation

Dep. var.: Non-routine Share	(1)	(2)	(3)	(4)
Unanticipated inflation (F&D)	0.190***		0.181***	0.162***
	(0.021)		(0.021)	(0.021)
Unanticipated inflation (overall)		0.075***		
		(0.017)		
Number of trips			0.033***	-0.028***
			(0.002)	(0.003)
Number of distinct stores				0.327***
				(0.011)
Observations	2,660,735	2,660,735	2,660,735	2,660,735
Consumer fixed effect	√	√	√	√
Consumer varying effect	√	√	√	√

Notes: The table reports the estimates in the robustness check. Each observation is at the consumer×quarter level covering from 2006 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 consumer level. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

The table shows that the basic results do not change. The non-routine share of total spending still increases significantly after an unanticipated inflation. After controlling number of trips and number distinct stores, the coefficients only fall a little. The key difference is that now the magnitude of the response is a bit smaller. This can happen when unemployment rates control away some variations that are not related to the change in the search activities.

Now, I consider the case where households use 4 lags of inflation in food and drinks, unemployment rates, 1-year government bond rates. This specification assumes a lot of knowledge about the macro economy households possess. The mean of the residual, measuring the unanticipated inflation, is 2.2 bp, and the standard deviation is 49 bp.

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

Dep. var.: Non-routine Share	(1)	(2)	(3)	(4)
Unanticipated inflation (F&D)	0.087***		0.078***	0.057*
	(0.023)		(0.023)	(0.023)
Unanticipated inflation (overall)		-0.066***		
		(0.018)		
Number of trips			0.033***	-0.028***
			(0.002)	(0.003)
Number of distinct stores				0.328***
				(0.011)
Observations	2,660,735	2,660,735	2,660,735	2,660,735
Consumer fixed effect	√	√	√	√
Consumer varying effect	√	√	√	√

Notes: The table shows the estimates for robustness check. Each observation is at the consumer×quarter level covering from 2006 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 consumer level. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

Now the magnitude of responses are much smaller. The 49 bp increase in the unanticipated inflation only triggers 9 bp increase in the non-routine share. It is 2.5 times smaller than our baseline. We should interpret this result as the following. The relatively larger response in the baseline result is largely due to the fact that households do not know the government bond rate or the response of the monetary authority. In theory, knowing perfectly the interest rate is sufficient to back out the underlying aggregate shock. However, in reality, households know little about the interest rate as revealed by Michigan survey.