Rational Inattention, Menu Costs, and Multi-Product Firms: Micro Evidence and Aggregate Implications*

Choongryul Yang[a]

[a]*Federal Reserve Board*

Abstract

Using a New Zealand firm-level survey, I show that firms producing more goods have both better information about inflation and more frequent but smaller price changes. To explain these empirical findings, I develop a general equilibrium menu cost model with rationally inattentive multi-product firms. I show that the interaction of nominal and informational rigidities leads to a new selection effect: price adjusters are better informed than non-adjusters. This selection endogenously generates a leptokurtic distribution of desired price changes, which amplifies monetary non-neutrality. Compared to a one-product baseline, the real effects of monetary shocks are 12% smaller in a two-product model.

Keywords: Inflation expectations, Monetary non-neutrality, Rational inattention, Menu costs, Multi-product firms, Economies of scope
JEL classification: E31, E32, E37, E52, E70

1. Introduction

The way firms set their prices and form their expectations has important implications for the transmission of monetary policy shocks. The extensive empirical literature studying detailed microdata on firms' beliefs and pricing behavior has found that many firms are not fully aware of macroeconomic conditions and change their prices infrequently.[1] Economic models have embedded

*This paper is a revised version of the first chapter of my Ph.D. dissertation at UT Austin. I would like to thank Olivier Coibion and Saroj Bhattarai for their guidance and support. I am grateful to Hassan Afrouzi, Chris Boehm, Carlos Carvalho, Stefano Eusepi, Cooper Howes, Andreas Mueller, Ernesto Pasten, Claudia Sahm, Raphael Schoenle, Antonella Tutino, and seminar participants at Fed Board, Boston Fed, Cleveland Fed, Dallas Fed, KC Fed, UT Austin, U of Edinburgh, Santa Clara U, HKUST, HKU, NUS, Sogang University, MEG 2019, ESEWM 2019, CEBRA 2020, ESWC 2020, SEA 2020, and SED 2021 for helpful comments and suggestions. Part of this paper was written at the Federal Reserve Bank of Boston while I was a dissertation intern. I acknowledge the Texas Advanced Computing Center (TACC) at UT Austin for providing HPC resources that have contributed to the simulation results reported within this paper. First version: November 2019; This version: February 2022.

Email address: choongryul.yang@frb.gov (Choongryul Yang)

[1] See Klenow and Malin (2010) and Nakamura and Steinsson (2008, 2013) for comprehensive reviews about micro price stickiness and Candia et al. (2021) for a review about macroeconomic expectations of firms.

realistic features of price setting and firms' expectations formation, such as costly price adjustments and a lack of awareness of economic conditions, to study the ability of monetary policy to stimulate the economy. Moreover, there is a mostly recent literature that has expanded our theoretical and empirical knowledge on multiproduct pricing.[2] This paper contributes to this literature in two primary ways. First, it documents that the firms' product scope is an important determinant of firm-level inattention to macroeconomic conditions: firms with a greater product scope have better information about aggregate economic conditions. Second, it studies aggregate implications using a unified model with menu costs, information friction, and multi-product firms: monetary non-neutrality decreases with the firms' product scope in the economy.

My starting point is to explore the empirical characteristics of multi-product firms regarding their price-setting and information acquisition decisions. Using a survey of New Zealand firms' macroeconomic beliefs, I document that firms' product scopes are systematically related to their attention to macroeconomic conditions and price-setting decisions. First, I find that firms with a greater product scope make systematically smaller errors about recent inflation and are willing to pay more for information about future inflation. This finding implies that firms with more products have incentive to process more information about macroeconomic conditions. To the best of my knowledge, these results are the first empirical evidence documenting differential information acquisition decisions of firms based on the number of products they sell. Second, consistent with the previous empirical literature on multi-product pricing, I find that firms with more products have more frequent but smaller price changes (e.g., Lach and Tsiddon, 1996; Bhattarai and Schoenle, 2014; Parker, 2017).

What are the aggregate implications of the micro-evidence that I show above for monetary non-neutrality? To answer this question, I develop a new model that captures the behavior of firms in the survey. Specifically, I assume that it is costly for firms to process information regarding underlying shocks. This captures the pervasive inattention among firms in the survey. In addition, firms have to pay a fixed menu cost to reset their prices, which leads to the infrequent price changes. I further assume that this fixed cost is independent of how many prices firms change. This assumption introduces economies of scope in price setting: firms with greater product scope change their prices more frequently and by smaller amounts because the average cost of changing prices is lower for them. Lastly, I assume that firms face two types of shocks—idiosyncratic good-specific shocks and aggregate monetary shocks—and the marginal cost of processing information is independent

[2] For example, Lach and Tsiddon (1996), Bhattarai and Schoenle (2014), and Bonomo et al. (2020) empirically study the characteristics of multiproduct pricing. Midrigan (2011) and Alvarez and Lippi (2014) build theoretical models of multi-product pricing with menu costs while Pasten and Schoenle (2016) build a rational inattention model with multi-product firms to study monetary non-neutrality. Multiproduct pricing has also been studied in trade literature (e.g., Bernard et al., 2011; De Loecker, 2011).

of the firms' number of products. This assumption introduces economies of scope in information processing: firms with a greater product scope want to learn more about aggregate monetary shocks because information about them can be used to price all their goods.[3]

I embed this setup of firm decision making into a dynamic general equilibrium model and study the macroeconomic implications for monetary non-neutrality. I focus on two key questions with this general equilibrium model. First, I explore how the interaction between rational inattention and menu costs affects firms' optimal decisions and, therefore, how the economy responds to monetary shocks, regardless of firms' product scopes. Second, I show how firms' product scopes affect monetary non-neutrality through economies of scope in multi-product firms by comparing the macroeconomic dynamics to monetary shocks in the one-good vs. two-good versions of my model.

My first theoretical finding is that the baseline one-good version of the model generates large real effects of monetary shocks that are seven times larger than those in the menu-cost-only model and nearly as large as those in the Calvo sticky price model. The menu-cost-only models have small and short-lived real effects of monetary shocks due to the strong selection effects of price changes: an expansionary monetary shock triggers numerous price increases that originate from far below the average level and offset many price decreases.The extent of the selection effects depends on the underlying distribution of firms' desired price changes. When the distribution is Gaussian, many prices are clustered around the adjustment margins, leading to large price selection effects. To generate strong monetary non-neutrality, previous menu cost models have often assumed that the distribution of idiosyncratic shocks has excess kurtosis, so the majority of desired price changes are near zero while some of them are very far from zero (e.g., Gertler and Leahy, 2008; Midrigan, 2011; Vavra, 2013; Karadi and Reiff, 2019; Baley and Blanco, 2019). However, I show that the interaction between menu costs and rational inattention can *endogenously* generate a distribution of firms' desired price changes with excess kurtosis. This leptokurtic distribution of desired price changes weakens the selection effects of price changes, amplifying the impact response of output to monetary shocks by 23% in my baseline single-product model compared with menu-cost-only models.

The second theoretical finding is that the real effects of monetary shocks decrease by 12% in the two-product model compared with the single-product model. As I discussed above, the two-product model exhibits two types of economies of scope in multi-product firms. First, there are economies of scope in price setting as shown in multi-product menu cost models in Midrigan (2011) and Alvarez

[3]This model nests the menu cost and rational inattention models as special cases. For example, without informational cost, this model coincides with the menu-cost-only models with either single-product firms (e.g., Golosov and Lucas, 2007) or multi-product firms (e.g., Midrigan, 2011; Alvarez and Lippi, 2014). Without menu cost, this model nests the rational-inattention-only models with either single-product firms (e.g., Maćkowiak and Wiederholt, 2009) or multi-product firms (e.g., Pasten and Schoenle, 2016).

and Lippi (2014). Because paying the menu cost allows firms to change the prices of all of their goods simultaneously, there are many small and large price changes in the two-product model. This weakens selection effects of price changes and should tend to amplify the real effects of monetary shocks in the two-product model. Second, as shown in multi-product rational inattention models in Pasten and Schoenle (2016), there are economies of scope in information processing: firms in the two-product model have better information and lower uncertainty about monetary shocks than firms in the single-product model. Because multi-product firms learn about monetary shocks rapidly, this force will tend to reduce monetary non-neutrality. My quantitative analysis shows that cumulative output effects are smaller in the two-product model than in the single-product model. This implies that the scope motive in information processing quantitatively dominates its effect on pricing decisions, thereby leading to reduced effects of nominal shocks on economic activity as we move to a multi-product environment.

These theoretical results are robust to introducing higher numbers of products in the model. I simplify the baseline model by assuming that firms' information acquisition decisions are independent of their price-setting decisions. Although this assumption eliminates the interesting interaction between nominal and informational rigidities, the simplified version of the model shares the core predictions of the baseline model: firms with a greater product scope have better information about aggregate shocks, and the kurtosis of the distribution of price changes increases with the number of products that firms produce. Consistent with the results from the full baseline model, I find that the cumulative output responses in the simplified version of the model decrease with the number of products sold by firms.

The paper is organized as follows. Section 2 empirically evaluates how firms' attentiveness to aggregate inflation and price-setting behavior are related to their number of products. In Section 3, I develop a new menu cost model with rationally inattentive multi-product firms that captures the behavior of firms in the survey, and I study firms' optimal decisions regarding information acquisition and price setting. In Section 4, I extend my model to study the implications for monetary policy transmission. Section 5 concludes the paper.

2. Empirical Evidence

In this section, I empirically explore how firms' product scopes relate to 1) their attentiveness to economic conditions and 2) the frequency and size of their price changes. I use a quantitative survey of firms' inflation expectations in New Zealand, which was conducted in multiple waves among a random sample of firms with broad sectoral coverage.[4]

[4] Several papers use the survey data to characterize how firms form their expectations. For example, Afrouzi (2020) shows that strategic complementarity decreases with competition, and documents that firms with more competitors

2.1. *Firms' Product Scopes in the Survey Data*

I use the second wave of the survey, which was implemented between February and April 2014, to identify firms' number of products. In the survey, firms' managers were asked about the number of products they sell in addition to their main product or product line. The median is 9 products, but it is 7 when firms in retail and wholesale trade sectors are excluded.[5] In the baseline regressions, I exclude these retail and wholesale trade firms because their strategies for pricing and information processing are likely to be different from those of firms in other sectors, such as manufacturing and service industries.[6] In the data, a larger fraction of firms (about 18% of all firms) compared with other studies sell only one product or have one product line.[7] There are two reasons behind the large fraction of single-product firms. First, the firms included in the survey were relatively small, with the average number of employees being about 31 and the largest number being about 600 employees. Second, the survey question is about the number of products or product lines at a firm. As there might be several similar types of products in a product line, this question captures firms' perceptions of the unit of their product scope. In fact, I find that the average of firms' output shares of their main product (or product line) is about 60% excluding single-product firms, implying that firms define their unit of product scope a bit broadly.

2.2. *Number of Products and Attentiveness to Inflation*

I first investigate the relationship between the firms' number of products and their attentiveness to aggregate economic conditions. Firms' attentiveness to aggregate economic conditions, measured by their backcast error in aggregate inflation, is defined to capture firms' knowledge about the current aggregate economy. Given that recent aggregate inflation is largely observable in real time, I define the backcast error as the absolute values of the difference between the actual past 12 months of inflation rate and managers' corresponding beliefs from the survey.[8] As documented in Coibion et al. (2018), firms are not well informed about current aggregate inflation, resulting in 4.5% backcast errors on average.

I find that firms' backcast errors are related to their number of products.[9] Figure 1 shows a

have more certain posteriors about the aggregate inflation. Also, Coibion et al. (2021) evaluate the relation between first-order and higher-order expectations of firms, including how they adjust their beliefs in response to a variety of information treatments. See Coibion et al. (2018) and Kumar et al. (2015) for a comprehensive description of the survey.

[5]Appendix Table G.1 shows the summary statistics regarding firms' number of products by industry.

[6]Including retail and wholesale trade firms in the sample does not change the baseline results that I show later. See, for example, Appendix Table G.2.

[7]For example, Bhattarai and Schoenle (2014) document that 98% of all prices are set by firms with more than one good in the microdata that underlie the calculation of the U.S. PPI.

[8]The CPI is used to calculate the actual past 12 month aggregate inflation. The baseline results are quantitatively similar when I use the GDP deflator or the PPI to calculate the actual inflation rate.

[9]See Appendix Table G.3 for the summary statistics of the firms' backcast error about aggregate inflation by the quartiles of firms' product scopes within different industries.

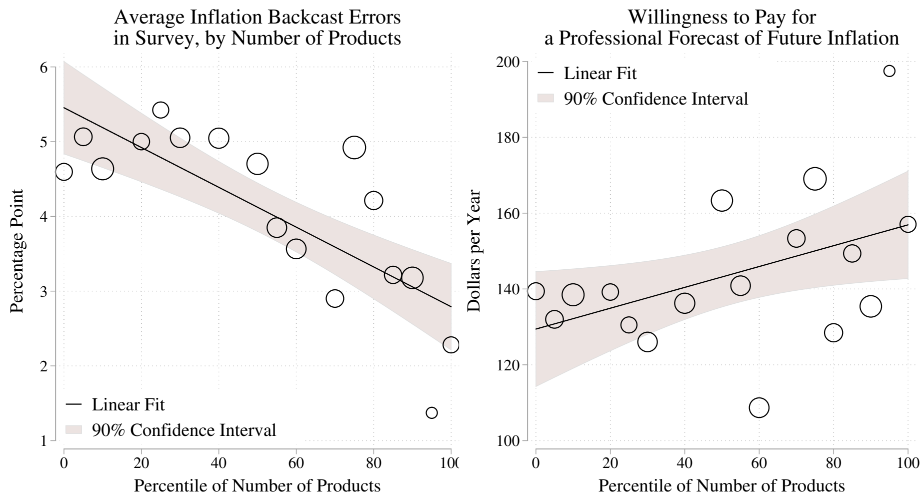


Figure 1: Number of Products and Attentiveness to Aggregate Inflation

Notes: The left panel plots percentile of firms' number of products versus the average of firm backcast errors about past 12 month inflation within each percentile. The right panel plots the percentile of firms' number of products versus the average of willingness to pay for a professional inflation forecast. The willingness to pay is measured from answers to the following question in Wave #4 survey: *"How much would you pay per year to have access to a monthly magazine of professional forecasts of future inflation?"* Black lines are linear fitted lines and shaded areas are 90% confidence intervals. The size of bins represents average size of employment of firms in each percentile.

clear positive correlation between firms' number of products and their attentiveness to aggregate inflation. The left panel shows that firms with a smaller product scope produce larger backcast errors on average. In the right panel, I show the relationship between firms' product scopes and their willingness to pay for professional inflation forecasts from the fourth wave of the survey. The latter is another measure of firms' incentive to be attentive to the aggregate economy. Here, I find a positive correlation between the number of products firms produce and their willingness to pay for professional inflation forecasts, which implies that firms with larger product scopes are likely to pay more attention to aggregate conditions.

One potential concern is that this negative correlation between the number of products firms produce and their knowledge of current aggregate inflation is driven by other firm-level characteristics.[10] To address this concern, I regress firms' inattention to inflation, as measured by 1) their backcast errors about aggregate inflation and 2) their willingness to pay for professional future inflation forecasts on the firms' number of products, controlling for firm-level characteristics such as a log of firm age, a log of total employment, foreign trade share, the firms' number of competitors, their beliefs about price differences relative to their competitors, and the slope of the profit function. The results are shown in Panels A and B of Table 1. Conditional on firm-level observables, firms

[10]As big firms are likely to have a larger product scope and a larger capacity to process information, the negative correlation might stem from the size of the firm rather than its product scope.

Table 1: Number of Products, Knowledge about Aggregate Inflation, and Price Changes

	(1)	(2)	(3)	(4)
Panel A. Dependent variable: Firms' inflation backcast errors				
log(number of products)	-0.314**	-0.207***	-0.580***	-0.252***
	(0.148)	(0.059)	(0.150)	(0.060)
Observations	593	582	448	440
R-squared	0.341	0.801	0.344	0.900
Panel B. Dependent variable: Willingness to pay for professional inflation forecasts				
log(number of products)	5.717**	3.805***	6.965**	3.960**
	(2.565)	(1.186)	(2.635)	(1.643)
Observations	381	368	328	320
R-squared	0.242	0.657	0.289	0.697
Panel C. Dependent variable: Duration of expected next price changes				
log(number of products)	-0.169	-0.205*	-0.276*	-0.474***
	(0.144)	(0.116)	(0.142)	(0.139)
Observations	588	580	440	445
R-squared	0.495	0.604	0.520	0.548
Panel D. Dependent variable: Size of expected next price changes				
log(number of products)	-0.111	-0.318***	-0.113	-0.239***
	(0.066)	(0.089)	(0.079)	(0.081)
Observations	578	578	432	431
R-squared	0.073	0.612	0.075	0.423
Firm-level controls	Yes	Yes	Yes	Yes
Industry fixed effects		Yes		Yes
Manager controls			Yes	Yes

Notes: This table reports results for the Huber robust regression. Dependent variables are the absolute value of firm errors about past 12 month inflation from Wave #1 survey (Panel A), firms' willingness to payment for professional forecasts about future inflation from Wave #4 (Panel B), the duration of expected next price changes from Wave #1 (Panel C) and the (absolute) size of expected next price changes from Wave #1 survey (Panel D). Firm-level controls include log of firms' age, log of firms' employment, foreign trade share, number of competitors, firms' beliefs about price difference from competitors, and the slope of the profit function. Industry fixed effects include dummies for 14 sub-industries excluding retail and wholesale trade sectors. Manager controls include the age of the respondent (each firm's manger), education, income, and tenure at the firm. Sample weights are applied to all specifications. Robust standard errors (clustered at the 3-digit Australian and New Zealand Standard Industrial Classification (ANZ SIC) level) are reported in parentheses. ***, **, * denotes statistical significance at 1%, 5%, and 10% levels respectively.

with a large number of products are likely to make small errors about aggregate inflation and are more willing to pay more for professional inflation forecasts. The significant negative correlation exists after controlling for industry fixed effects. Lastly, as shown in column (3), I report regression results after controlling for managers' characteristics, such as age, education, and income level. After controlling for these manager characteristics, I find a negative correlation between the firms' number of products and their knowledge of attentiveness to aggregate inflation.[11]

2.3. Number of Products and Size and Frequency of Price Changes

In this subsection, I document the relationship between firms' product scope and the frequency and size of their price changes. Firms' managers were asked about the duration and size of their expected next price changes. Conditional on firms' characteristics and their incentives for changing prices, this question quantifies the frequency and size of price changes.

Table 1 shows the relationship between firms' product scope and their pricing behavior. Panel C shows that, conditional on firm-level characteristics, the duration of price changes is negatively related with the number of products. This negative correlation is even stronger when I control for managers' characteristics. Panel D shows that after controlling for industry fixed effects, there is a negative correlation between firms' number of products and the size of their price changes. This confirms the previous findings on multi-product pricing: conditional on the price change, firms with more products change their prices by smaller amounts (e.g., Lach and Tsiddon, 1996; Bhattarai and Schoenle, 2014; Parker, 2017).

2.4. Summary and Relation to Monetary Models

Using the New Zealand firm-level survey data, I show that firms with a greater number of products have both better information about aggregate inflation and more frequent but smaller price changes. Can existing monetary models with multi-product firms explain both empirical findings? In fact, Pasten and Schoenle (2016) show that a rational inattention model with multi-product firms predicts a negative correlation between the firms' number of products and their inattentiveness to aggregate inflation. In the presence of both good-specific shocks and aggregate shocks, multi-product firms want to be more informed about the aggregate shocks, as they will affect the marginal costs of all their products. On the other hand, menu cost models with multi-product firms, such as Midrigan (2011) and Alvarez and Lippi (2014), predict that multi-product firms change their prices more frequently and by less because menu cost technology can introduce economies of scope in price setting.

However, those two models are not capable of explaining both empirical findings simultaneously by assumption. Rational inattention models assume flexible prices to focus on the effects of information rigidity, while menu cost models assume perfect information to focus on the effects of nominal rigidity. Those two assumptions are inconsistent with the empirical evidence that firms change their prices infrequently and are not fully aware of macroeconomic conditions. Moreover,

[11] In Appendix Table G.4, I show that firms' backcast errors in the growth rate of nominal GDP also decrease in their number of products. In the quantitative model in Section 4, I calibrate the information cost parameter to match the slope coefficient of the regression of the backcast errors in nominal GDP growth.

the previous literature finds a contradictory implication of firms' product scopes for monetary non-neutrality. In menu cost models, the output effects of monetary shocks *increase* with the firms' number of products, as selection effects of price changes from menu cost technology decrease with the firms' number of products.[12] In contrast, as shown in Pasten and Schoenle (2016), the multiproduct rational inattention model implies that monetary non-neutrality *decreases* with the number of products because firms with a greater product scope are more informed about aggregate shocks.

In sum, neither model with multi-product firms can account for the empirical relationship between firms' product scopes and their decisions regarding both price setting and information acquisition, and they have contradictory implications regarding firms' product scopes for monetary non-neutrality. This calls for a new model disciplined by the empirical findings from micro data in order to study the macroeconomic implications for monetary non-neutrality.

3. Price Setting with Menu Costs for a Rationally Inattentive Multi-Product Firm

In this section, I develop a menu cost model for a rationally inattentive multi-product firm and study the decision problem of this firm that faces both good-specific shocks and an aggregate monetary shock. The goal of this section is to explore how the firm's decisions regarding information acquisition and price setting are affected by 1) the interaction between menu costs and rational inattention and 2) the firm's product scope .

3.1. A Rationally Inattentive Firm's Problem

Consider a multi-product firm that produces N goods, indexed by $j = 1,2,\cdots,N$. The firm sets its price of good j, $p_{j,t}$, to match a (frictionless) optimal price, $p^*_{j,t} = a_{j,t} + m_t$, which is a sum of two components, a good-specific shock ($a_{j,t}$), and an aggregate shock (m_t). I assume that both shocks follow random walk processes: $a_{j,t} = a_{j,t-1} + \varepsilon^a_{j,t}$ with $\varepsilon^a_{j,t} \sim N(0,\sigma^2_a)$ for $j = 1,2,\cdots,N$, and $m_t = m_{t-1} + \varepsilon^m_t$ with $\varepsilon^m_t \sim N(0,\sigma^2_m)$ where $\varepsilon^a_{j,t}$ and ε^m_t are independent and identically distributed.[13] A flow loss in profits is the sum of the distance between its price of each good and the frictionless price, $B\sum_{j=1}^N \left(p_{j,t} - p^*_{j,t}\right)^2$, where B captures the concavity of the profit function with respect to each price.[14]

[12] Alvarez et al. (2016) show that in menu cost models, the cumulative output response to a monetary shock increases in the number of products, N, and is given by $\frac{3N}{N+2}$.

[13] The random walk assumption is common in the menu cost literature, as it simplifies the firm's problem by making it choose its price gaps, which are defined by the difference between the frictionless and the actual prices. See, among others, Barro (1972), Tsiddon (1993), and Alvarez and Lippi (2014).

[14] While I take this characteristic as an assumption, this loss function can be derived as a second order approximation to a twice-differentiable profit function around the non-stochastic steady state.

This firm is rationally inattentive. At the beginning of each period, the firm chooses how precisely it wants to observe its current set of (frictionless) optimal prices subject to a cost of information processing . Formally, at time t, the firm chooses a set of signals about both shocks from a set of available signals, $\mathscr{S}_t = \{\mathscr{S}^a_{j,t}\}_{j=1}^N \cup \mathscr{S}^m_t$, such that $\mathscr{S}^a_{j,t} = \{a_{j,t} + \eta_{j,t}\xi^a_{j,t} : \eta_{j,t} \geq 0,\ \xi^a_{j,t} \sim N(0,1)\}$ for $j = 1,2,\cdots,N$ and $\mathscr{S}^m_t = \{m_t + \eta_{m,t}\xi^m_t : \eta_{m,t} \geq 0,\ \xi^m_t \sim N(0,1)\}$ where $\{\xi^a_{j,t}\}_{j=1}^N$ and ξ^m_t are the firm's rational inattention errors. Let S^{t-1} be the firm's information set at the beginning of period t before it receives new signals about its frictionless optimal prices. Given S^{t-1}, the firm chooses a set of its signals $s^a_{j,t} \in \mathscr{S}^a_{j,t}$ for $j = 1,2,\cdots,N$, and $s^m_t \in \mathscr{S}^m_t$ subject to the cost of information processing. Then, the firm's information set evolves as $S^t = S^{t-1} \cup s_t$ where $s_t = \{\{s^a_{j,t}\}_{j=1}^N, s^m_t\}$. The evolution of information set implies that the firm does not forget information over time. This "no-forgetting constraint" implies that the current information choice has a continuation value and thus the optimal information choice is a solution of a dynamic information acquisition problem.

I assume that the cost of information is linear in Shannon's mutual information function. The firm pays ψ units of its (per-good) revenue for every bit of expected reduction in uncertainty, where uncertainty is measured by entropy. Denote this cost as $\psi\mathscr{I}(s_t;\{p^*_{j,t}\}_{j=1}^N|S^{t-1})$, which will be defined later. At each period, based on its optimal information choice, the firm chooses whether to change its prices. I assume that the firm can change all prices by paying a single fixed cost, θ.[15] Figure 2 shows the timing of events for the firm's problem.

Formally, the firm's problem is as follows:

$$\begin{aligned}\min_{\{\{p_{j,t}\}_{j=1}^N, s_t\}_{t=0}^\infty} \quad & \mathbb{E}\Bigg[\sum_{t=0}^\infty \beta^t \Bigg(\underbrace{B\sum_{j=1}^N \left(p_{j,t} - p^*_{j,t}\right)^2}_{\textbf{loss from suboptimal prices}} + \underbrace{\theta \mathbf{1}_{\{\text{for any } j,\ p_{j,t} \neq p_{j,t-1}\}}}_{\textbf{cost of price changes}} \\ & \qquad + \underbrace{\psi\mathscr{I}\left(s_t;\{p^*_{j,t}\}_{j=1}^N|S^{t-1}\right)}_{\textbf{cost of information processing}} \Bigg) \Bigg| S^{-1}\Bigg] \qquad (1)\\ \text{subject to} \quad & p^*_{j,t} = a_{j,t} + m_t, \quad \forall j = 1,2,\cdots,N, \\ & S^t = S^{t-1} \cup s_t,\ \ S^{-1} \text{ is given,}\end{aligned}$$

where $\mathbf{1}_{\{\text{for any } j,\ p_{j,t} \neq p_{j,t-1}\}}$ is an indicator function that is 1 if it changes any one of its prices.[16]

[15] In Appendix A.1, I discuss evidence of the firm-specific menu costs using the survey data. I also discuss the implications of alternative assumptions on menu costs, such as product-specific menu costs, menu costs that scale linearly, and adding variable costs of price changes. The baseline results are qualitatively robust to these alternative assumptions.

[16] Besides the existence of menu costs, this problem is different from the previous rational inattention models in LQG settings, such as Maćkowiak and Wiederholt (2009) or Pasten and Schoenle (2016), which solve the problem by assuming that the cost of information is not discounted and optimizing at the long-run steady state for the information structure. In contrast, I assume that the firm discounts future costs of information at the same discount rate as its payoffs.

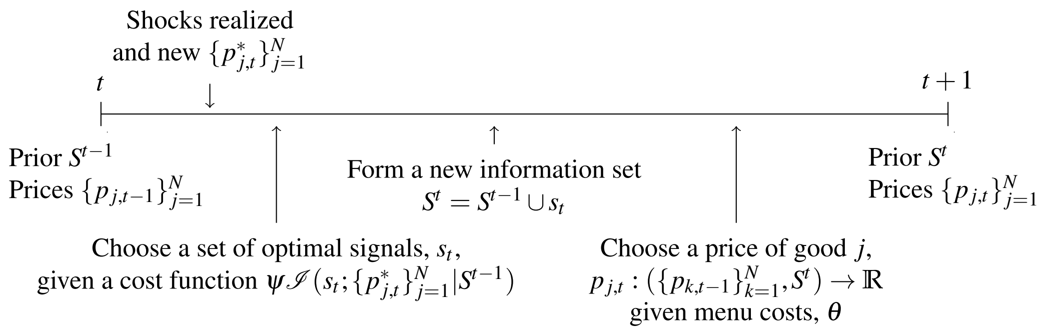


Figure 2: Timing of Events for a Firm's Problem

Notes: This figure shows a sequence of events in each period of the model.

Cost of Information Processing. Let $\mathscr{H}(X|Y)$ be a conditional entropy of a random variable of X given Y. The flow cost of information at t is $\psi \mathscr{I}(s_t; \{p^*_{j,t}\}_{j=1}^N | S^{t-1})$, where $\mathscr{I}(s_t; \{p^*_{j,t}\}_{j=1}^N | S^{t-1}) = \mathscr{H}(\{p^*_{j,t}\}_{j=1}^N | S^{t-1}) - \mathbb{E}[\mathscr{H}(\{p^*_{j,t}\}_{j=1}^N | S^t) | S^{t-1}]$ is the reduction in uncertainty about its (frictionless) optimal prices that the firm experiences by observing the set of signals, s_t, given its prior information set, S^{t-1}, and ψ is the marginal cost of a bit of information.

Let $z^a_{j,t} \equiv var(a_{j,t}|S^t)$ and $z^m_t \equiv var(m_t|S^t)$ be the firm's subjective uncertainty about the j-good-specific shock and about the aggregate shock, respectively. I can rewrite the cost of information processing at time t in terms of $\{z^a_{j,t}\}_{j=1}^N$ and z^m_t:

$$\mathscr{I}\left(s_t; \{p^*_{j,t}\}_{j=1}^N | S^{t-1}\right) = \frac{1}{2}\left(\sum_{j=1}^N \log_2\left(\frac{z^a_{j,t-1} + \sigma_a^2}{z^a_{j,t}}\right) + \log_2\left(\frac{z^m_{t-1} + \sigma_m^2}{z^m_t}\right)\right) \tag{2}$$

where $\{z^a_{j,-1}\}_{j=1}^N$ and z^m_{-1} are given. It follows from the fact that the underlying shocks are independent and the firm observes independent Gaussian signals about them. I also rewrite the no-forgetting constraint, $S^t = S^{t-1} \cup s_t$, in terms of the firm's subjective uncertainty, $0 \leq z^a_{j,t} \leq z^a_{j,t-1} + \sigma_a^2$ for $j = 1, 2, \cdots, N$ and $0 \leq z^m_t \leq z^m_{t-1} + \sigma_m^2$.

This reformulation shows that the cost of information processing is directly related to how much each firm reduces its subjective uncertainty about the good-specific shocks and the aggregate shock, given their prior uncertainty about those shocks. If the marginal cost of information processing, ψ, is zero, the firm would like to choose zero subjective uncertainty about both underlying shocks. As it is costly for the firm to reduce a large amount of uncertainty about the underlying shocks when $\psi > 0$, it optimally chooses to observe less precise signals and to be optimally uncertain about the underlying shocks.

This dynamic nature of firms' information acquisition is the reason that firms are subject to the no-forgetting constraints. See, for instance, Afrouzi and Yang (2021) for a detailed discussion of solutions for the dynamic rational inattention problem in LQG setups.

Recursive Formulation of the Firm's Problem. I reformulate the firm's problem (1) in a recursive form to characterize its optimal decision rules and simulate the model numerically. Let $x_{j,t} = p_{j,t} - \mathbb{E}[p^*_{j,t}|S^t]$ be the firm's *perceived* price gap of product j. Then, the loss from suboptimal prices can be decomposed into two components:

$$\mathbb{E}\left[\left(p_{j,t} - p^*_{j,t}\right)^2 \middle| S^t\right] = \underbrace{z^a_{j,t} + z^m_t}_{\substack{\text{contemporaneous loss from} \\ \textbf{imperfect information}}} + \underbrace{x^2_{j,t}}_{\substack{\text{contemporaneous loss from} \\ \textbf{nominal rigidities}}} . \tag{3}$$

On the one hand, if there is no informational cost, $\psi = 0$, then the firm chooses zero subjective uncertainty and thus there is no contemporaneous loss from imperfect information. In this case, the firm's problem is similar to the problem in a menu cost model with multi-product firms, such as Midrigan (2011). On the other hand, if there is no menu cost, $\theta = 0$, then the firm can always adjust its prices freely and thus will choose zero *perceived* price gaps, $x_{j,t} = 0$ for all j. In this case, the firm's problem is similar to the pioneering work of Pasten and Schoenle (2016), who were the first to consider rational inattention in a setting with multi-product firms.

The perceived price gaps, $\{x_{j,t}\}$, are the firm's choice variables when it changes its prices. However, if the firm does not change its prices, the perceived price gaps are stochastic variables that evolve according to $\mathbf{x}_t \sim N(\mathbf{x}_{t-1}, \Sigma_t)$ where $\mathbf{x}_t = \{x_{1,t}, x_{2,t}, \cdots, x_{N,t}\}'$ and

$$\Sigma_t(j,k) = \begin{cases} z^m_{t-1} + \sigma^2_m - z^m_t & \text{if } j \neq k \\ z^a_{j,t-1} + \sigma^2_a - z^a_{j,t} + z^m_{t-1} + \sigma^2_m - z^m_t & \text{if } j = k. \end{cases} \tag{4}$$

Given (2), (3), and (4), I reformulate the firm's problem (1) in a recursive form with $2N+1$ state variables and occasionally binding constraints:

$$\begin{aligned} V(\{x_{j,-1}\}^N_{j=1}, \{z^a_{j,-1}\}^N_{j=1}, z^m_{-1}) &= \max_{\{\{z^a_j\}^N_{j=1}, z^m\}} \mathbb{E}\Bigg[\max\left\{V^I\left(\{x_j\}^N_{j=1}, \{z^a_j\}^N_{j=1}, z^m\right), V^C\left(\{z^a_j\}^N_{j=1}, z^m\right)\right\} \\ &\qquad - \frac{\psi}{2}\left(\sum_{j=1}^N \log_2\left(\frac{z^a_{j,-1} + \sigma^2_a}{z^a_j}\right) + \log_2\left(\frac{z^m_{-1} + \sigma^2_m}{z^m}\right)\right) \Bigg| S^{-1}\Bigg], \\ &\quad \text{subject to} \quad 0 \leq z^a_j \leq z^a_{j,-1} + \sigma^2_a, \quad \forall j = 1, 2, \cdots, N, \\ &\quad 0 \leq z^m \leq z^m_{-1} + \sigma^2_m, \\ \text{where } V^I\left(\{x_j\}^N_{j=1}, \{z^a_j\}^N_{j=1}, z^m\right) &= -B\sum_{j=1}^N \left(x^2_j + z^a_j + z^m\right) + \beta V\left(\{x_j\}^N_{j=1}, \{z^a_j\}^N_{j=1}, z^m\right) \\ &\quad \text{with } \mathbf{x} \sim N(\mathbf{x}_{-1}, \Sigma), \text{ and} \\ V^C\left(\{z^a_j\}^N_{j=1}, z^m\right) &= \max_{\{y_j\}^N_{j=1}} -B\sum_{j=1}^N \left(y^2_j + z^a_j + z^m\right) - \theta + \beta V\left(\{y_j\}^N_{j=1}, \{z^a_j\}^N_{j=1}, z^m\right). \end{aligned}$$

Here $V^I(\{x_j\}^N_{j=1}, \{z^a_j\}^N_{j=1}, z^m)$ represents the firm's value of not changing its prices. Similarly,

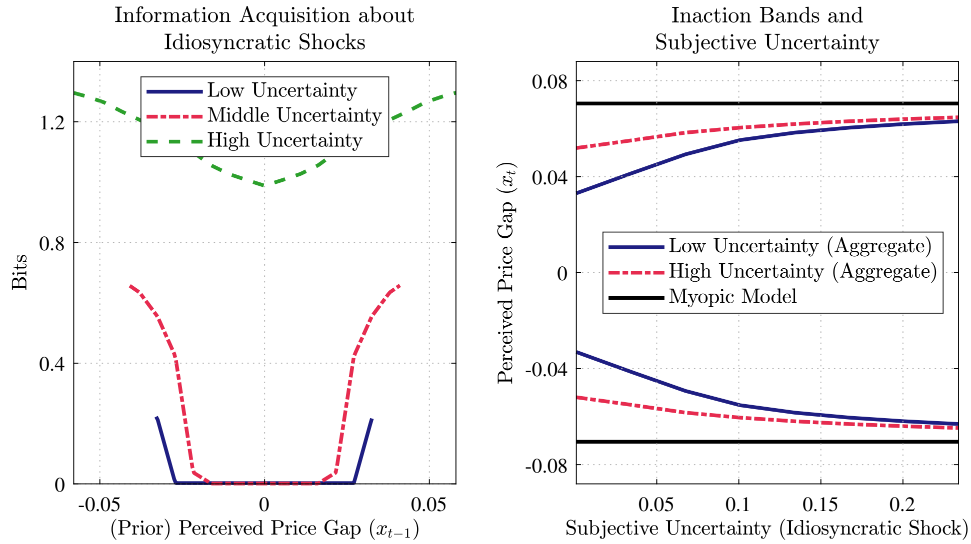


Figure 3: Information Acquisition and Inaction Bands of a Single-Product Firm

Notes: The left panel plots the amounts of information acquisition from a single-product firm's optimal choice of subjective uncertainty about the good-specific shock. The right panel shows inaction bands of a single-product firm as a function of its subjective uncertainty. Different lines represent the inaction bands with different levels of subjective uncertainty about the aggregate shock (z_t^m). Black lines are the inaction bands of a myopic firm whose discount factor is zero. Since this myopic firm does not care about a continuation value of information, the subjective uncertainty is not its state variable, which leads the inaction bands of this firm to be constant

$V^C(\{z_j^a\}_{j=1}^N, z^m)$ is the firm's value of changing its prices.

3.2. Decision Rules

In this section, I describe key properties of the firm's optimal decision rules. First, because of the quadratic objective function and the symmetry of the normal distribution, the value function is also symmetric around the null vector for the perceived price gaps. Second, given the optimal choices of subjective uncertainty about the good-specific shocks and about the aggregate shock, the value function is decreasing in the absolute values of perceived price gaps. These two properties imply that, given optimal choices of subjective uncertainty, the firm chooses to have zero perceived price gaps for all their goods whenever it decides to change its prices by paying the menu cost, θ.

Because the firm's problem is a non-convex optimization problem in its price-setting decision and there are occasionally binding no-forgetting constraints for its choices of subjective uncertainty, it needs to be solved numerically. Using the method of value function iteration, I solve the problems of two types of firms: a single-product firm and a two-product firm. I first investigate how the interaction between menu costs and rational inattention frictions affects the single-product firm by characterizing its optimal information acquisition and price-setting decisions. Then, I show how economies of scope in both price setting and information processing affect the two-product firm by comparing with the single-product firm.

3.2.1. A Single-Product Firm

I first consider a single-product firm's optimal decision rules.[17] Here I drop the j-index because the firm produces only one product.

Optimal Information Acquisition. Optimal policy rules for choosing a firm's subjective uncertainty about the good-specific shock are presented in Figure 3. In particular, the left panel of Figure 3 shows that when its prior uncertainty is low enough and its prior price gap is close to zero, the no-forgetting constraint binds and the firm does not acquire new information. The amount of information acquisition increases in both the firm's prior uncertainty (z^a_{t-1}) and the distance between its current price and the frictionless optimal price ($|x_{t-1}|$). In other words, the firm has a large incentive to collect and process information when the firm is quite uncertain about the realization of the underlying shocks and thinks that it is likely that it will need to change prices. This is because potential losses from mistakes in the price-setting decisions are large if the firm thinks that it is likely that it will need to change its price. The firm could make wrong decisions either by paying the menu cost and changing the price when it was not supposed to do so or by choosing not to change its price when it should.

Price-Setting Decision Given the Optimal Information Choices. After choosing optimal signals about the underlying shocks and forming the new information set, the firm decides whether to change its price, based on the new information set. Due to the fixed menu cost of adjusting prices, the firm adopts *S-s* rules in setting its prices—there are adjustment thresholds (s,S) such that if the firm's *perceived* price gap is greater than S or less than s, it pays the fixed cost and adjusts decreases its price to the frictionless optimal level. The adjustment thresholds are the firm's inaction bands. One interesting feature in this model is that the inaction bands are time-varying. Formally, let $\hat{x}_t$ be the firm's posterior belief about its perceived price gap *after* observing the new optimal signals and *before* changing its price at time t. Then,

$$\hat{x}_t = p_{t-1} - \mathbb{E}[a_t + m_t|S^t] = x_{t-1} - \{\mathscr{K}_t{}^a(s^a_t - \mathbb{E}[a_t|S^{t-1}]) + \mathscr{K}_t{}^m(s^m_t - \mathbb{E}[m_t|S^{t-1}])\},$$

where $\mathscr{K}_t{}^a$ and $\mathscr{K}_t{}^m$ are the optimal Kalman gains for the good-specific and aggregate shocks, respectively. A higher value of the Kalman gains implies that the firm observes more precise signals about the underlying shocks. Given the optimal choice of subjective uncertainty, z^a_t and z^m_t, there exists an adjustment threshold $\tilde{x}(z^a_t, z^m_t) \geq 0$ such that

$$-\tilde{x}_t(z^a_t, z^m_t)^2 + \beta V\left(B\tilde{x}_t(z^a_t, z^m_t), z^a_t, z^m_t\right) = -\theta + \beta V\left(0, z^a_t, z^m_t\right).$$

[17] To illustrate this example, I set $\theta = 0.0074$, $\psi = 0.0035$, $\sigma_a = 0.0168$, and $\sigma_m = 0.0044$. These are the parameters that I calibrate when I solve a general equilibrium single-product model in Section 4.

The firm will change its price if $|\hat{x}_t| > \tilde{x}_t(z_t^a, z_t^m)$. Then, the perceived price gap at the end of period t, x_t, is

$$x_t = \begin{cases} \hat{x}_t & \text{if } |\hat{x}_t| \le \tilde{x}_t(z_t^a, z_t^m) \\ 0 & \text{if } |\hat{x}_t| > \tilde{x}_t(z_t^a, z_t^m). \end{cases}$$

The right panel of Figure 3 shows the inaction bands, $(-\tilde{x}_t(\cdot, z_t^m), \tilde{x}_t(\cdot, z_t^m))$, for the various values of z_t^m. With $\beta > 0$, the inaction bands vary with the subjective uncertainty.[18] When the firm is more uncertain about the underlying shocks, the inaction bands are wider, implying that it is optimal to wait until it gets more information about them.

The main implication of the interaction between menu cost and information friction is that the firm is likely to be more informed about the underlying shocks when it changes prices than when it does not. As will be clear in the next section, in a general equilibrium model with a large number of firms, this interaction leads to a selection in information processing such that price adjusters are more informed than non-adjusters.[19]

3.2.2. A Two-Product Firm

Now, I consider the two-product firm's optimal information acquisition and price setting decision. In fact, the two-product firm shares the same characteristics about its optimal decision rules with the single-product firm that I discussed earlier. However, two interesting motives based on economies of scope emerge in the two-product firm's optimal choices—economies of scope in price changes through the menu cost technology and economies of scope in information processing through the rational inattention friction.

Economies of Scope in Price Changes. Figure 4 shows a two-product firm's information acquisition and price setting behavior in the model simulation. Like a single-product firm, the inaction bands of the two-product firm also depend on its subjective uncertainty. The main difference in the two-product firm's price setting decision is that the price change of one of its products depends on the perceived price gap of the other product. For example, the right upper and lower panels of Figure 4 show that when the perceived price gap of the second product is large, the inaction bands for the first product are narrow. This relationship implies that the timing of price changes within the

[18] The inaction bands in a myopic model ($\beta = 0$) are constant and given by $(-\sqrt{\theta/B}, \sqrt{\theta/B})$ because the firm's subjective uncertainty is no longer a state variable for the firm's problem.

[19] Models with both menu costs and observational costs, such as Alvarez et al. (2011) and Bonomo et al. (2021), also have similar implications. The optimal pricing rule implies that the firm only changes its price when it pays the observational cost, which is an extreme case of the selection in information processing in the sense that the firm has full information when it changes its price while it does not acquire new information at all otherwise. Gorodnichenko (2008) also shows that firms have an incentive to buy an additional signal prior to changing prices in a model with menu costs and endogenous information choice.

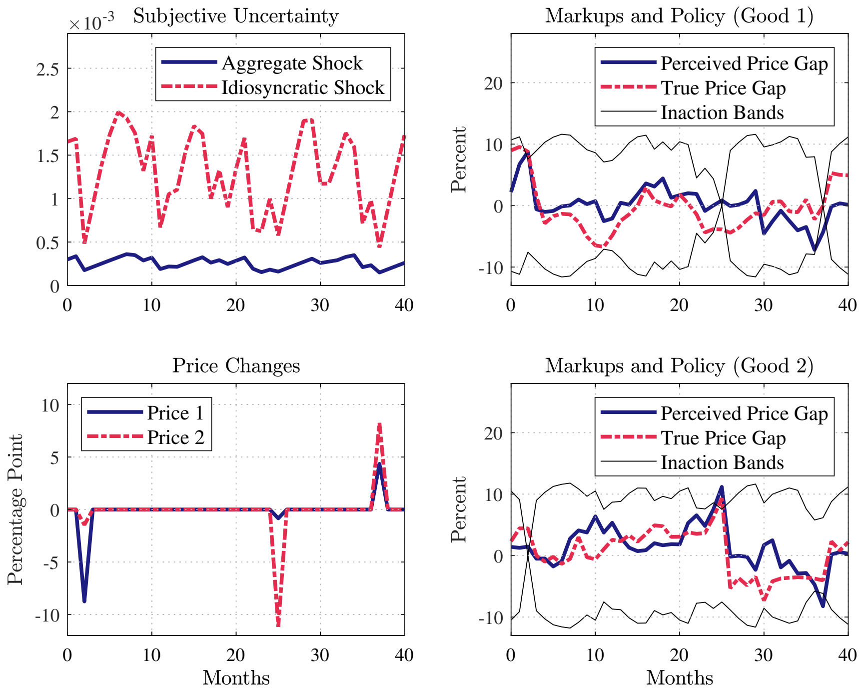


Figure 4: A Two-Product Firm in Model Simulation

Notes: The upper left panel plots a two-product firm's subjective uncertainty about both a good-specific shock (red dash-dot line) and an aggregate shock (blue solid line). The upper right panel plots the firm's perceived price gap (x_t) for good 1, its true price gap for good 1 under perfect information, and inaction bands for good 1. The lower right panel plots the firm's perceived price gap (x_t) for good 2, its true price gap for good 2 under perfect information, and inaction bands for good 2. The firm changes its price when the perceived price gaps are out of the inaction bands. The lower left panel plots these price changes for both goods.

firm is synchronized and, more importantly there are both large and small price changes. This is called economies of scope in price changes from the menu cost technology. If the firm decides to pay the menu cost to change one of its prices, then the price change of the other products is free for the firm. As a result, given the fixed menu cost, the two-product firm is likely to change its prices more frequently than the single-product firm. Moreover, as the additional price change is free and thus there are many small price changes, the two-product firm changes its prices, on average, by a smaller amount than the single-product firm. These results are consistent with empirical findings in Section 2 and confirm the previous multi-product pricing literature that firms producing more goods have more frequent but smaller price changes (e.g., Lach and Tsiddon, 1996; Bhattarai and Schoenle, 2014; Parker, 2017). As highlighted in the menu cost literature, economies of scope in price changes will weaken the selection effects of price changes by letting some price changes be random (e.g., Midrigan 2011; Alvarez and Lippi 2014). The weak selection effects of price changes in an economy with a large number of multi-product firms will then lead to an amplified real effect of monetary shocks.

Economies of Scope in Information Processing. The two-product firm is also different from the single-product firm in terms of its optimal information acquisition about the aggregate shock. In particular, I find that given the same marginal cost of information processing, the two-product firm is more informed about the aggregate shock than the single-product firm. The average subjective uncertainty about the aggregate shock for the two-product firm is 25% smaller than that for the single-product firm. Because the firm's optimal prices for all goods are affected by the aggregate shock, the value of information about the aggregate shock will be higher if the firm produces more products.[20] This result is consistent with my empirical finding that firms that produce more goods have better information about aggregate inflation. As the two-product firm has lower subjective uncertainty about the aggregate shock than the single-product firm, the two-product firm responds more strongly to monetary policy shocks by learning about them more rapidly and therefore changing their prices more rapidly. In the economy with a large number of firms, this scope motive in information processing acts as a strong force to weaken monetary non-neutrality.

In sum, the two-product firm's information acquisition and price-setting decisions show two economies of scope motives that work in opposite directions for monetary non-neutrality. The question is how quantitatively large each scope motive is. To draw the implications of firms' product scopes for monetary non-neutrality in the model with both menu costs and rational inattention friction, we need to extend the model in a general equilibrium setup and discipline it with micro-evidence. This is the goal of the next section.

4. A Dynamic General Equilibrium Model

In this section, I extend the baseline model of Section 3 to a dynamic general equilibrium model. The model is disciplined using micro-evidence of Section 2 and then used for quantitative analysis on the transmission of monetary shocks. The focus of this analysis is twofold. First, I investigate how the interaction between rational inattention and menu costs affects the distribution of firms' desired price changes and the distribution of subjective uncertainty, which are important determinants of monetary non-neutrality. Second, I compare the single- and two-product models to study how multi-product pricing affects monetary non-neutrality through economies of scope in both information processing and price setting.

[20] Pasten and Schoenle (2016) formally prove this relationship under the multi-product rational inattention model. In Appendix B, I revisit the similar relationship under the (dynamic) rational inattention problem without menu costs. In this case, the firm's subjective uncertainty about the aggregate shock satisfies the following FOC: $B \cdot N = \frac{\psi}{2\log 2}\left(\frac{1}{z_t^m} - \beta \frac{1}{z_t^m + \sigma_m^2}\right)$, where z_t^m is decreasing in the number of products, N.

4.1. Environment

I first discuss the representative household problem and then present the firm problem and define the equilibrium.

Households. The representative household consumes a Dixit-Stiglitz aggregate consumption, C_t, of a basket of multiple goods $j \in \{1, 2, \cdots, N\}$ purchased from firms $i \in [0, 1]$, and supplies labor L_t to maximize the expected lifetime utility with a discount factor $\beta \in (0, 1)$.

The representative household's problem is

$$\max_{\{\{C_{i,j,t}\}_{j=1}^{N}, C_t, L_t, B_t\}_{t\geq 0}} \quad \mathbb{E}_0^f \left[\sum_{t=0}^{\infty} \beta^t \log(C_t) - L_t \right],$$

$$\text{subject to} \quad \int \Big(\sum_{j=1}^{N} P_{i,j,t} C_{i,j,t} \Big) di + B_t \leq R_{t-1} B_{t-1} + W_t L_t + \Pi_t, \quad \text{for all } t,$$

where $C_t = \left(\frac{1}{N} \sum_{j=1}^{N} C_{j,t}^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}$ and $C_{j,t} = \left(\int (A_{i,j,t} C_{i,j,t})^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$. Here $\mathbb{E}_t^f[\cdot]$ is the full-information rational expectation operator at time t.[21] B_t is the demand for nominal bonds and R_{t-1} is the nominal interest rate.L_t is the labor supply of the household, W_t is the nominal wage, and Π_t is the aggregate profit from the firms. C_t is the aggregator over the consumption for differentiated goods and $C_{j,t}$ is an aggregator over the consumption of good j. $A_{i,j,t}$ is the quality of the good j produced by firm i. Higher $A_{i,j,t}$ increases the marginal utility of consumption for that good while it also increases the production cost for that good, as I describe below.[22] ε is the constant elasticity of substitution across different firms that produce the same good and γ is the constant elasticity of substitution across different goods.

Firms. There is a measure one of firms, indexed by i, that operate in monopolistically competitive markets. Each firm produces N goods, indexed by j. Firms take wages and demands for their goods as given, and choose their prices $\{P_{i,j,t}\}_{j=1}^{N}$ based on their information set, S_i^t, at that time. After setting their prices, firms hire labor from a competitive labor market and produce the realized level of demand that their prices induce with a production function for good j, $Y_{i,j,t} = \frac{1}{A_{i,j,t}} L_{i,j,t}$, where $L_{i,j,t}$ is firm i's demand for labor for producing good j. Notice that higher quality products require extra labor input. I assume that shocks to $A_{i,j,t}$ are i.i.d. and the log of the j-good specific shock, $a_{i,j,t} \equiv \log(A_{i,j,t})$, follows a random walk: $\log(A_{i,j,t}) = \log(A_{i,j,t-1}) + \varepsilon_{i,j,t}^a$ where $\varepsilon_{i,j,t}^a \sim N(0, \sigma_a^2)$

[21] As the main purpose of this paper is to study the effects of nominal rigidity and rational inattention among firms, I assume that the household is fully informed about all prices and wages.

[22] The assumption that idiosyncratic shocks affect both the cost at which a good is sold and the household's marginal utility for the good is common in the literature to reduce the dimensionality of the state space and the computational burden (see, e.g., Midrigan, 2011; Alvarez and Lippi, 2014; Karadi and Reiff, 2019).

for $j = 1, 2, \cdots, N$.[23] Then, firm i's nominal profit from sales of all goods at prices $\{P_{i,j,t}\}_{j=1}^{N}$ is given by

$$\Pi_{i,t}(\{P_{i,j,t}, A_{i,j,t}, P_{j,t}\}_{j=1}^{N}, W_t, P_t, Y_t) = \sum_{j=1}^{N} (P_{i,j,t} - W_t A_{i,j,t}) A_{i,j,t}^{\varepsilon-1} \left(\frac{P_{i,j,t}}{P_{j,t}}\right)^{-\varepsilon} \left(\frac{P_{j,t}}{P_t}\right)^{-\gamma} Y_t, \tag{5}$$

where Y_t is the aggregate demand.

Firms optimally decide their prices and signals subject to the costs of changing prices and of processing information. First, changing the price entails a fixed cost, $\tilde{\theta}$. I express this cost as a fraction θ of the steady-state frictionless revenue from selling one of N products. Second, firms choose their optimal information set by taking into account the cost of processing information. At the beginning of period t, firm i wakes up with its initial information set, S_i^{t-1}. Then it chooses optimal signals, $s_{i,t}$, from a set of available signals, $\mathscr{S}_{i,t}$, subject to the cost of information, which is linear in Shannon's mutual information function. Denote $\tilde{\psi}$ as the marginal cost of information processing. Again, I express this marginal cost as a fraction ψ of the steady-state frictionless revenue from selling one of N products. Firm i forms a new information set, $S_i^t = S_i^{t-1} \cup s_{i,t}$, and uses it to set its new prices, $\{P_{i,j,t}\}_{j=1}^{N}$.

Firm i chooses a set of signals to observe over time $(s_{i,t} \in \mathscr{S}_{i,t})_{t=0}^{\infty}$ and a pricing strategy that maps the set of its prices at $t-1$ and its information set at t to its optimal price at any given period, $P_{i,j,t} : (\{P_{i,j,t-1}\}_{j=1}^{N}, S_i^t) \to \mathbb{R}$ where $S_i^t = S_i^{t-1} \cup s_{i,t} = S_i^{-1} \cup \{s_{i,\tau}\}_{\tau=0}^{t}$ is the firm's information set at time t. Then, firm i's problem is to maximize the net present value of its life time profits given an initial information set:

$$\max_{\{s_{i,t} \in \mathscr{S}_{i,t}, \{P_{i,j,t}(\{P_{i,k,t-1}\}_{k=1}^{N}, S_i^t)\}_{j=1}^{N}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \Lambda_t \left\{ \Pi_{i,t}(\{P_{i,j,t}, A_{i,j,t}, P_{j,t}\}_{j=1}^{N}, W_t, P_t, Y_t) - \tilde{\theta} \mathbf{1}_{\{\text{for any } j,\ P_{i,j,t} \neq P_{i,j,t-1}\}} - \tilde{\psi} \mathscr{I}(s_{i,t}; \{\{A_{i,j,t}\}_{j=1}^{N}, W_t\} | S_i^{t-1}) \right\} \middle| S_i^{-1} \right]$$

$$\text{subject to} \quad S_i^t = S_i^{t-1} \cup s_{i,t},$$

where $\Lambda_t = \frac{U_{c,t}/P_t}{U_{c,0}/P_0}$ and $\mathscr{I}(s_{i,t}; \{\{A_{i,j,t}\}_{j=1}^{N}, W_t\} | S_i^{t-1})$ is Shannon's mutual information function.

Monetary Policy and Equilibrium. Money supply is equal to nominal spending: $P_t C_t = M_t$, and the log of money supply, $m_t \equiv \log(M_t)$, follows a random walk: $m_t = m_{t-1} + \varepsilon_t^m$ where $\varepsilon_t^m \sim N(0, \sigma_m^2)$ is an independently and identically distributed normal disturbance.[24]

[23] One could alternatively consider either cross-sectional correlation in the good-specific shocks or serial correlation in underlying shocks. I discuss the implications of these alternative assumptions on monetary non-neutrality in Appendix A.4.

[24] For the computational simplicity, I formulate monetary policy in terms of an exogenous rule for money supply, which is popular framework in the literature to study the effects of monetary policy on pricing (e.g., Caplin and Spulber,

An equilibrium consists of an allocation for the representative household, $\Omega^H \equiv \{C_t, \{C_{i,j,t}\}_{j=1}^N, L_t, B_t\}_{t=0}^\infty$, an allocation for firm $i \in [0,1]$, $\Omega_i^F \equiv \{s_{i,t} \in \mathscr{S}_{i,t}, \{P_{i,j,t}, L_{i,j,t}, Y_{i,j,t}\}_{j=1}^N\}_{t=0}^\infty$, a set of prices $\{\{P_{j,t}\}_{j=1}^N, P_t, R_t, W_t\}_{t=0}^\infty$, and a stationary distribution over firms' states such that (i) the household and all firms optimize, (ii) the stationary distribution is consistent with actions, and (iii) all markets clear.

4.2. Computing the Equilibrium

Firms' profit functions (5) imply that without any frictions in price setting and in information processing, firm i's frictionless optimal price of good j, $P_{i,j,t}^*$, is a constant markup over its nominal marginal cost, $P_{i,j,t}^* = \frac{\varepsilon}{\varepsilon-1} W_t A_{i,j,t}$. Let $\mu_{i,j,t} = \bar{\mu} \frac{P_{i,j,t}}{P_{i,j,t}^*}$ be firm i's true price gap for good j, which is a gap between the actual price and the frictionless optimal price. Here $\bar{\mu} = \frac{\varepsilon}{\varepsilon-1}$ is a non-stochastic steady-state level of the price gap. Then, firm i's nominal flow profit at time t can be written as a function of the set of these price gaps, $\sum_{j=1}^N (\mu_{i,j,t} - 1)(\mu_{i,j,t})^{-\varepsilon} (W_t)^{1-\varepsilon} (P_{j,t})^{\varepsilon-\gamma} (P_t)^\gamma Y_t$.

To solve the firm's problem, I take the second-order approximation to the firms' profit function and derive firms' losses from sub-optimal pricing.[25] I assume that the set of available signals, $\mathscr{S}_{i,t}$, has the following properties. First, the firm chooses $N+1$ independent signals for each shock, implying that paying attention to aggregate conditions and paying attention to good-specific idiosyncratic conditions are separate activities. Second, each signal is Gaussian. Third, all noise in signals is idiosyncratic and independent.[26] The second-order approximation reduces the state space of the problem from an entire distribution to its covariance matrix. Moreover, because the signals are Gaussian and the objective function is quadratic, it enables me to focus on a Gaussian posterior. Under these assumptions, each firm's problem is identical to that studied in Section 3.1. I compare two economies with $N = 1$ and $N = 2$.[27] I simulate the economy with a large number of firms for a long period to make sure that the economy reaches a stationary distribution over firms' states.[28]

4.3. Calibration and Parameterization

I set the monthly discount factor to $\beta = 0.96^{(1/12)}$, which implies a real interest rate of 4 percent. I set the elasticity of substitution across firms to be four ($\varepsilon = 4$), which matches the firms' average

1987; Gertler and Leahy, 2008). However, as shown in Afrouzi and Yang (2021), this formulation is identical to a Taylor rule specification where monetary authority targets the growth of the nominal aggregate demand, which can be interpreted as targeting inflation and output growth.

[25] See Appendix C for the detailed derivation of the second-order approximation.

[26] See Appendix A.2 for a discussion on the implications and limitations of these assumptions.

[27] A two-product economy is also considered as the baseline in Midrigan (2011) and Karadi and Reiff (2019). Moreover, in the New Zealand survey data, the average share of the main product's total output value is about 60%, excluding single-product firms. It implies that a two-product firm might be a good benchmark. In Section 4.7, I solve models with any arbitrary number of products under simplifying assumptions.

[28] I describe the computational procedures and simulation algorithm in Appendix D.

markup of 33% in the survey data.[29] Moreover, I assume the elasticity of substitution between goods is the same as that across firms ($\gamma = 4$). However, the value of γ plays little role, as there are no common good-specific shocks in the model.

I calibrate the standard deviation of the log difference in nominal demand, σ_m, to match the standard deviation of the growth rate of nominal GDP in New Zealand, 0.0044. There are three key model parameters that should be calibrated: the size of menu costs (θ), the size of marginal costs of information processing (ψ), and the size of idiosyncratic good-specific shocks (σ_a). I assume that the marginal cost of processing one bit of information in both one-good and two-good versions of the model are the same.[30] I calibrate these three parameters to match the median frequency of price changes (once a year), the median size of absolute price changes (5.76%), and slope of backcast errors in the growth rate of aggregate nominal GDP on the number of products (-0.02) observed in the survey data. The latter is obtained by regressing the firms' backcast errors in the growth rate of nominal GDP on the firms' number of products. The model counterpart measure is calculated by taking the difference between average backcast errors in the growth rate of nominal demand in single- and two-product models. The three moments exactly identify the three key model parameters, and, as Panel A in Table 2 shows, all the targeted moments are well matched.

Panel B of Table 2 shows the calibrated and assigned parameters in both single-product and two-product models. The baseline parameterization implies a menu cost of 0.74 percent of steady-state (per good) revenue in the single-product model. Given the average frequency of price changes, the overall cost of price adjustment in the single-product model is around 0.062 percent of steady-state revenue. Similarly, the overall cost of price adjustment in the two-product model is around 0.058 percent of steady-state revenue. These values are smaller than estimates in the previous literature, which often used U.S. data, as the average size of price changes in New Zealand is small.[31] The standard deviations of the good-specific shocks are around 1.7 percent per month in both models, which are about four times bigger than the standard deviation of the monetary shock. The cost of information processing, $\psi\mathscr{I}(\cdot;\cdot)$, is 0.11 percent of steady-state (per good) revenue. The implied average Kalman gains on the signal about idiosyncratic shocks are 0.27 in both models (see Table 3). This is equivalent to a quarterly gain of 0.60, which is slightly higher than the estimate of 0.50 in Coibion and Gorodnichenko (2015), which uses the U.S. Survey of Professional Forecasters data.

[29] This value is in the middle of 3 and 7, the elasticity of substitution parameters in Midrigan (2011) and Golosov and Lucas (2007), respectively. It directly affects the slopes of profit curves, and thus the estimates of menu costs and the standard deviation of good-specific shocks, without altering the main findings.

[30] As shown in Pasten and Schoenle (2016), there are alternative choices on costs of information acquisition, such as a constant shadow price of information capacity per good or a constant loss per good from imperfect information. In Appendix A.3, I discuss the implications of these alternative assumptions. The baseline results are qualitatively robust to these alternative assumptions.

[31] For example, Levy et al. (1997) find menu costs of 0.7 percent of revenue. Stella (2020) estimates the total cost of changing prices to be between 0.3% and 1.3% of revenues.

Table 2: Model Calibration

	Data	Single-product model	Two-product model
Panel A. Data and model moments			
Median (absolute) size of price changes	0.0576	0.0576	0.0576
Median frequency of price changes	0.0833	0.0833	0.0833
Slope of the backcast error curve	-0.020	-0.020	
Panel B. Calibrated parameters			
Menu cost (θ)		0.0074	0.0281
Information cost (ψ)		0.0035	0.0035
S.D. of idiosyncratic shocks (σ_a)		0.0183	0.0188
S.D. of monetary policy shocks (σ_m)		0.0044	0.0044
Panel C. Assigned parameters			
Time discount factor (β)		0.9966	0.9966
Elasticity of substitution across firms (ε)		4.0	4.0
Elasticity of substitution between goods (γ)			4.0

Notes: The table presents the baseline parameters for the general equilibrium models with single- and two-product firms. Panel A presents moments of the data and simulated series from the single- and two-product models parameterized at the baseline values. To get the slope of the backcast error curve, I regress the absolute value of firm errors about past 12 month nominal GDP growth rate from Wave #4 survey on the number of products each firm produces. Regression results are reported in Appendix Table G.4. Panel B shows the calibrated parameters which match the three key moments shown in Panel A. Panel C shows the assigned parameters. See Section 4.3 for details.

4.4. Simulation

In this section, I emphasize two distributional characteristics that will be important to understand the transmission of monetary shocks in both models. First, I show that selection in information processing endogenizes a leptokurtic distribution of desired price changes, which acts as a force to weaken selection effects of price changes. Second, multi-product firms value more information about the monetary policy shock than the single-product firms.

Selection in Information Processing and Endogenous Leptokurtic Distribution of Price Gaps. Table 3 shows an important characteristic about firms' optimal information choices. The second and third rows compare the average Kalman gain of firms that adjust their prices to those of firms that do not adjust their prices. The Kalman gain represents how much weight firms put on new information relative to their prior estimates.[32] Thus, the average Kalman gains can be interpreted as the average degree of firms' attentiveness to the underlying shocks. I find that there is a selection in information processing: price adjusters are more informed about both idiosyncratic and aggregate shocks than

[32]When firms' signals are perfectly telling about the true shocks, the Kalman gain is 1. When firms optimally choose not to be perfectly informed due to the information cost, the Kalman gain is less than 1.

Table 3: Average Kalman Gains in Models

	Single-product model		Two-product model	
Average Kalman gains	Signal about good-specific shocks ($\mathscr{K}_t^a$)	Signal about monetary shocks ($\mathscr{K}_t^m$)	Signal about good-specific shocks ($\mathscr{K}_t^a$)	Signal about monetary shocks ($\mathscr{K}_t^m$)
All firms	0.267	0.089	0.279	0.131
- Price adjusters	0.611	0.257	0.653	0.259
- Price non-adjusters	0.235	0.074	0.244	0.120

Notes: The table presents average Kalman gains across firms in both baseline models. Column (1) and (3) show the average Kalman gains on the signal about the good-specific shocks in the single-product model and those in the two-product model, respectively. Column (2) and (4) show the average Kalman gains on the signal about the monetary policy shock in the single-product model and those in the two-product model, respectively.

price non-adjusters. In particular, the price adjusters' average Kalman gains in the signals about idiosyncratic shocks are about three times bigger than those of price non-adjusters. These findings are true in both the single- and the two-product models, implying that selection in information processing operates regardless of the number of products in the model.

This selection in information processing is due to the interaction between firms' optimal information and pricing decisions. As shown in Section 3.1, firms' optimal information acquisition policies are affected by their beliefs about their price gaps. If a firm believes that its price is far away from its optimal level, mistakes in pricing decisions would be very costly, which make the firm process more information about the shocks to reduce potential losses. After the realization of shocks, this firm is likely to be a price adjuster because its prior price gap is close to the inaction bands. For this reason, on average, the price adjusters in the economy are better informed about the underlying shocks than the price non-adjusters.[33]

This selection mechanism in information processing *endogenously* generates a leptokurtic distribution of firms' perceived desired price changes.[34] Figure 5 shows the distributions of the perceived and true price gaps in the single-product model. First, the blue line is a prior distribution about firms' perceived price gaps. At the beginning of the period, all firms believe that their prices are within their inaction bands and there are many zero perceived price gaps. This implies that the *prior* distribution of firms' perceived price gaps is concentrated around zero and has a high kurtosis. After being hit by Gaussian idiosyncratic shocks, the distribution of true price gaps is Gaussian (red dashed line). If firms have perfect information about their true optimal prices, their pricing decisions would be based on their true price gaps. As the distribution is Gaussian, as in the single-product

[33]In Appendix F, I show that in the New Zealand survey data, firms with greater uncertainty are more likely to delay their price changes.

[34]If a firm has a price gap of x % and it is free to change its price, then it would change by $-x$%. I use "price gaps" and "desired price changes" interchangeably.

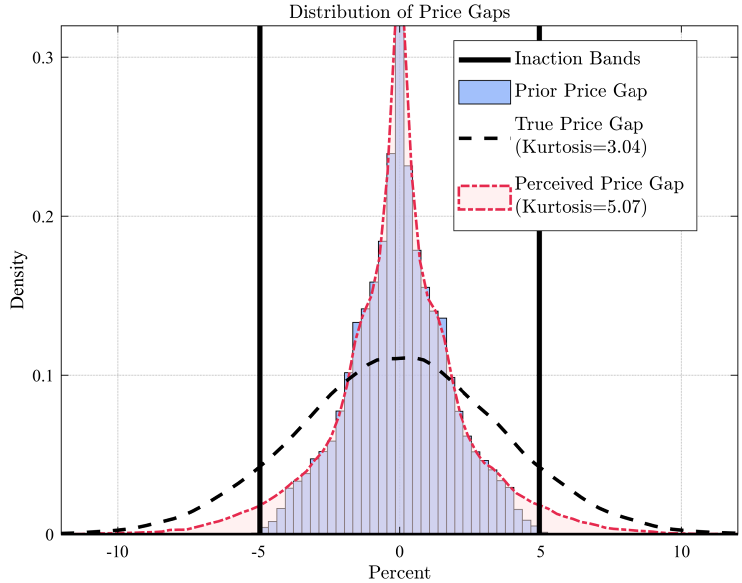


Figure 5: Distributions of True and Perceived Price Gaps in the Single-Product Model

Notes: This figure plots distributions of price gaps in the baseline single-product model. Black lines are the average of inaction bands across firms in the model. At the beginning of period, before the realization of shocks, all firms believe that their price is within the inaction bands. Blue bar graph shows the distribution of firms' *prior* about their price gaps ($p_{i,j,t-1} - \mathbb{E}_{t-1}[p^*_{i,j,t}|S^{t-1}]$) at the beginning of period. After the Gaussian shocks realized, firms' marginal costs change, and thus their true price gap ($p_{i,j,t-1} - p^*_{i,j,t}$) also changes. Black dashed line shows the distribution of these true price gaps. Firms choose their optimal signals about the shocks and form a new posterior about their (frictionless) optimal price. Then, the *posterior* of perceived price gap is $p_{i,j,t-1} - \mathbb{E}_t[p^*_{i,j,t}|S^t]$. Red dash-dot line shows the distribution of these perceived price gaps.

menu-cost-only model, there would be large selection effects of price changes: an expansionary monetary shock triggers many large price increases, but it offsets a mass of large price decreases.

However, in the model with both menu costs and informational costs, firms are rationally inattentive about their true optimal prices. Firms all choose their optimal Gaussian signals and update their estimates of price gaps, but they do not do so in the same way. Firms that think that their price gap is well within their inaction bands and that think it is unlikely that they will need to change prices have little incentive to collect much new information: they choose to remain quite uninformed and update the estimates of their price gaps with a large weight on their (imprecise) priors. In contrast, firms that think they are close to the boundaries of their inaction regions have a high incentive to collect information and therefore choose to become more informed. Because the distribution of priors of firms' perceived price gaps is very concentrated around zero, this selection in information processing makes the distribution of posteriors of the perceived gaps *leptokurtic*.[35]

This leptokurtic distribution implies a small selection effect of price changes because the

[35] See Appendix Figure G.1 for the endogenous leptokurtic distribution in the two-product model.

rationally inattentive firms' pricing decisions are based on their posterior of perceived price gaps.[36] Thus, the endogenous leptokurtic distribution will act as a strong force to amplify monetary non-neutrality in the general equilibrium model. Previous menu cost models often assume exogenously a leptokurtic distribution of shocks (e.g., Gertler and Leahy, 2008; Midrigan, 2011; Vavra, 2013; Karadi and Reiff, 2019; Baley and Blanco, 2019). Unlike these studies, due to selection in information processing, my baseline model can generate the leptokurtic distribution *endogenously* even if the distribution of shocks is Gaussian.[37]

Value of Information about the Aggregate Shock. Optimal attention allocation implies that firms have an incentive to allocate their attention toward more volatile shocks (e.g. Maćkowiak and Wiederholt, 2009). The first row of Table 3 shows, in both single- and two-product models, that the average Kalman gains for the idiosyncratic shocks are larger than those for the aggregate shock because the latter is less volatile.

However, the amount of information processing about the aggregate shock is different in the single- and the two-product models. The value of information about the aggregate shock is higher for the two-product firms than the single-product firms, as the firms' ideal prices for all goods are affected by the aggregate shock. The Kalman gains for aggregates shock are about twice as large in the two-product model than in the single-product model. Consistent with the implication of multi-product rational inattention models without menu costs in Pasten and Schoenle (2016), this implies that the two-product firms will be more responsive to monetary shocks than the single-product firms as they are more informed about it.

4.5. Real Effects of Monetary Policy Shocks

In this subsection, I take the calibrated models and hit them with a one-standard-deviation shock to monetary policy. Figure 6 shows the output response in the one-good and the two-good versions of the model. I also show the impulse responses in the standard menu cost model with single-product firms and in the Calvo sticky price model.[38]

The output response to a monetary policy shock in the standard menu cost model is small and short-lived. The half-life of output response is only two months. This is a well-known fact in this model: there are large selection effects of price changes, which act as a strong force to reduce monetary non-neutrality (e.g., Golosov and Lucas, 2007). In the single-product model with both menu costs and rational inattention however, the real effects of monetary shocks are large and

[36] In Appendix E, I compare the distributions of actual price changes in the baseline models.

[37] In Appendix F, I show that the empirical distribution of desired price changes is leptokurtic.

[38] In the appendix, I show how the real effects of monetary shocks in my baseline models change when I shut down one friction at a time. Appendix Figures G.2 and G.3 show the output responses in the menu-cost-only models and in the rational-inattention-only models with single- and two-product firms, respectively.

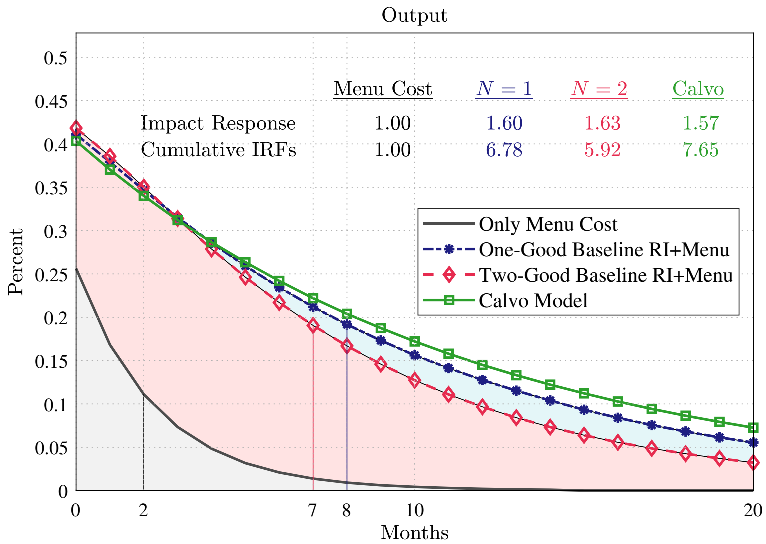


Figure 6: Impulse Response of Output to a One S.D. Monetary Shock

Notes: This figure plots impulse responses of output to a one standard deviation monetary shock. Cumulative IRFs refers to area under the responses of output. I normalize both impact and cumulative output response in the menu-cost-only model with single-product firms as one.

persistent. The impact response increases by 60% and the cumulative output responses, which are defined as the area under the output responses, are about seven times larger than those in the menu-cost-only model. In fact, this large real effect is comparable to that in the Calvo sticky price model.

However, the large real effects are reduced in the two-product model. The output responses in the two-product model are 12% smaller than those in the single-product model, and the half-life of output responses also decreases from eight months to seven months.[39]

4.6. Inspecting the Mechanisms

In this subsection, I investigate the key mechanisms behind the results of monetary non-neutrality in the baseline model. To this end, I start from the single-product menu-cost-only model, such as Golosov and Lucas (2007), and consider counterfactual models by adding core elements of my baseline model. I discuss five main mechanisms; three of them have been studied in the previous literature while two of them are new in this paper.[40]

[39]Interestingly, as shown in Appendix Figure G.4, the implied kurtosis of the distribution of price changes is higher in the two-product model than in the single-product model. This suggests that in this model with both rational inattention and menu costs, the ratio of kurtosis to the frequency of price changes might not be a sufficient statistic for the output response to a monetary shock, which is derived by Alvarez et al. (2016).

[40]Appendix Figure G.5 shows how each counterfactual model is related to the underlying mechanisms.

Endogenous Leptokurtic Distribution. The first model (1A) is the single-product menu-cost-only model in Golosov and Lucas (2007). In this model, firms have perfect information about both idiosyncratic and monetary shocks. This model implies small and short-lived real effects of monetary shocks due to large selection effects of price changes (black solid line in Appendix Figure G.6). For comparison with other counterfactual models, I normalize the impact output response and the cumulative output response in this model to 1.

In the next model (1B), I assume that the single-product firms have perfect information about the monetary shock but are rationally inattentive to their good-specific shock. Because firms choose their optimal signals about the good-specific shock, selection in information processing about the idiosyncratic shock makes the distribution of firms' desired prices endogenously leptokurtic. This leads to small price selection effects because there is only a small fraction of firms around the inaction bands. Thus, the output responses in this counterfactual model are larger than those in the menu-cost-only model. As shown in Appendix Figure G.6, the impact output effect in model (1B) increases by 23% compared with that in model (1A).[41] This mechanism is new in the literature; the interaction between menu costs and rational inattention generates the endogenous leptokurtic distribution of desired price changes, which amplifies monetary non-neutrality in a non-trivial way.

Imperfect Information about Monetary Policy Shocks. Next, I assume that single-product firms are not only rationally inattentive to the good-specific shock, but informationally constrained about monetary shocks. I assume that each firm is *exogenously* given a signal about monetary shocks, which has the same precision as the average precision in the baseline single-product model (1D). In other words, firms in this counterfactual economy have the same degree of attentiveness as do firms in the baseline model where all information choices are endogenous. This counterfactual model (1C) captures the role of imperfect information about monetary shocks for monetary non-neutrality, which has been studied in the literature (e.g., Woodford, 2003; Maćkowiak and Wiederholt, 2009). Appendix Figure G.6 shows that it has the most important role for amplifying the real effects of monetary shocks. The output effects are seven times bigger in this model than in the menu-cost-only model.

Selection in Information Processing about Monetary Policy Shocks. Now, I assume that the single-product firms choose their optimal signal about the monetary shocks rather than receive an exogenous signal. This model (1D) is my baseline single-product model. The comparison of this model to model (1C) captures the role of selection in information processing about the monetary shocks. Firms in both models (1C) and (1D), *on average*, acquire the same amount of information about the

[41] As firms have perfect information about monetary shocks and the aggregate demand follows a random walk, they immediately observe and respond to the shocks if their prices are around the adjustment margins, which also makes the real effects of monetary shocks in this counterfactual model short-lived.

monetary shocks. While all firms have the same information about the monetary shocks in the model (1C), price adjusters choose to have better information than non-adjusters in the single-product baseline model (1D). This implies that firms that change their prices following the monetary shocks in the baseline model will adjust more strongly and learn quickly about the shocks compared with the price-adjusting firms in model (1C). Thus, the real effects of monetary shocks are smaller in this baseline model (1D) than (1C). The output responses in the baseline single-product model are 20 % smaller than those in model (1C). This mechanism is also new in the literature.

Economies of Scope in Price Setting and Information Processing. Lastly, I consider the baseline two-product model. The two-product model entails economies of scope motives in price setting and information processing, studied in Midrigan (2011) and Pasten and Schoenle (2016), respectively. Both economies of scope motives work in the opposite directions for monetary non-neutrality. The output responses decrease by 12% in the two-product model compared with the single-product model, implying that in the calibrated model, economies of scope in information processing act as a strong force to reduce monetary non-neutrality.

4.7. Models with a Large Number of Products

In this subsection, I show that the implication of multi-product pricing for monetary non-neutrality can be extended to the models with an arbitrary large number of products. The main computational challenge for solving the baseline model with more than two products is that the number of state variables double with an additional good produced by the firms and firms' optimal information choice problem is subject to occasionally binding constraints.[42] To simplify the analysis, I make two assumptions. First, I assume that firms choose how much to process information about the underlying shocks *as if* they do not face menu costs. Second, given menu costs, firms choose their prices based on that information, but they are *myopic* in the sense that they do not care about the continuation value of their current pricing decisions. One limitation of making these assumptions is that it eliminates the interesting interaction between rational inattention and menu costs because all firms choose to have the same information set about the underlying shocks. However, as it simplifies the model analysis by eliminating state variables but keeps the core of the baseline model, I analyze the implications of multi-product pricing for monetary non-neutrality under these assumptions.[43]

Figure 7 shows the cumulative output responses to a monetary shock in the simplified models with various numbers of products. I calibrate each model with a different number of products to

[42]The number of state variables is increasing linearly with the number of products as $2N+1$ and the number of choice variables is increasing linearly as $N+1$. In addition, the most computationally burdensome part is to compute firms' expected future values. When the firm does not change prices, its price gaps are stochastic variables that are jointly normally distributed with a mean vector, $\mathbf{x}_{-1}$, which is the firm's state variable, and a covariance matrix, Σ, which is its choice variable. Standard approximation methods for the transition probability of states, like Tauchen (1986), are not applicable, as approximation errors are large. I compute the expected value of the firms' value functions

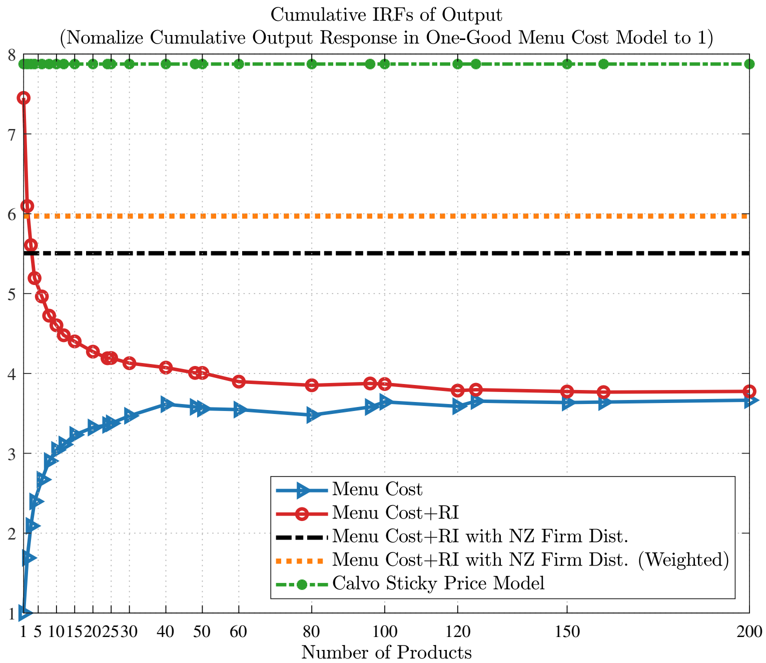


Figure 7: Cumulative Output Responses and Number of Products in the Simplified Models

Notes: This figure plots cumulative output responses in the simplified models with different number of products. "Menu Cost+RI" refers to the model with both menu costs and rational inattention. Red line shows the cumulative output responses in the only menu cost models with different number of products and blue line shows those in the models with both menu costs and rational inattention with different number of products. Black (orange) line shows the cumulative output response in the model with both menu costs and rational inattention when I match the unweighted (employment-weighted) empirical distribution of firms' number of products. See Section 4.7 for details.

have the same size and frequency of price changes.[44] I normalize the cumulative output response in the single-product menu-cost-only model to one. In the menu-cost-only models (blue line with triangles), the effects increase with the number of products. In the models with both rational inattention and menu costs (red line with circles), the real effects decrease with the number of products but converge to the menu-cost-only models with a large number of products. As the number of products increases in the model, firms' subjective uncertainty about monetary policy shocks decreases and converges to zero, implying firms have almost perfect information about the monetary policy shocks.

I also consider the model that exactly matches the distribution of number of products across

using an explicit numerical integration technique.

[43]For example, as shown in Appendix Figure G.7, the backcast errors in the growth rate of nominal GDP decrease with the number of products. This relationship stems from the economies of scope in information processing in rational inattention models with multi-product firms. Moreover, kurtosis of the distribution of price changes increases with the number of products and converges to a value of three, which is consistent with the implications of menu cost models with multi-product firms.

[44]The calibration strategy implies that each model with a different number of products is observationally equivalent in the sense that it has the same price statistics observed in the data.

firms in the New Zealand survey data. I consider both employment-weighted (black dash-dot line) and unweighted versions (orange dotted line). The cumulative output effects in these models are similar to those in the two- or three-product economy, implying that our baseline analysis with the two-product model could be an empirically relevant benchmark. In sum, the main implication on the relationship between monetary non-neutrality and firms' product scopes can be extended to the model with a large number of products.

5. Conclusion

Understanding the nature of firms' expectations formation and price setting behavior has been an active area of research in monetary economics. In this paper, using a firm-level survey from New Zealand, I find that firms with a greater number of products have both better information about aggregate inflation and more frequent but smaller price changes.

I build a general equilibrium menu cost model with rationally inattentive multi-product firms to study the aggregate implications for monetary non-neutrality. In this model, the interaction between nominal and informational rigidities leads to a novel selection effect of information processing: price adjusters have better information about the underlying shocks than non-adjusters. This new selection effect leads to a fat-tail distribution of firms' desired price changes and thus weakens selection effects of price changes. As a result, the real effects of monetary shocks in the single-product model are nearly as large as those in the Calvo model. Finally, the output responses in the two-product model are smaller than those in the single-product model due to the strong economies of scope motive in information processing.

References

Afrouzi, H., 2020. Strategic inattention, inflation dynamics and the non-neutrality of money. mimeo.

Afrouzi, H., Yang, C., 2021. Dynamic rational inattention and the Phillips curve. mimeo.

Alvarez, F., Le Bihan, H., Lippi, F., 2016. The real effects of monetary shocks in sticky price models: A sufficient statistic approach. American Economic Review 106, 2817–51.

Alvarez, F., Lippi, F., 2014. Price setting with menu cost for multiproduct firms. Econometrica 82, 89–135.

Alvarez, F., Lippi, F., Paciello, L., 2011. Optimal price setting with observation and menu costs. The Quarterly Journal of Economics 126, 1909–1960.

Baley, I., Blanco, A., 2019. Firm uncertainty cycles and the propagation of nominal shocks. American Economic Journal: Macroeconomics 11, 276–337.

Barro, R.J., 1972. A theory of monopolistic price adjustment. The Review of Economic Studies 39, 17–26.

Bernard, A.B., Redding, S.J., Schott, P.K., 2011. Multiproduct Firms and Trade Liberalization. The Quarterly Journal of Economics 126, 1271–1318.

Bhattarai, S., Schoenle, R., 2014. Multiproduct firms and price-setting: Theory and evidence from u.s. producer prices. Journal of Monetary Economics 66, 178–192.

Bonomo, M., Carvalho, C., Garcia, R., Malta, V., 2021. Persistent monetary non-neutrality in an estimated model with menu Costs and partially costly information. mimeo.

Bonomo, M., Carvalho, C., Kryvtsov, O., Ribon, S., Rigato, R., 2020. Multi-product pricing: Theory and evidence from large retailers in Israel. mimeo.

Candia, B., Coibion, O., Gorodnichenko, Y., 2021. The macroeconomic expectations of firms. mimeo.

Caplin, A.S., Spulber, D.F., 1987. Menu costs and the neutrality of money. The Quarterly Journal of Economics 102, 703–725.

Coibion, O., Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review 105, 2644–78.

Coibion, O., Gorodnichenko, Y., Kumar, S., 2018. How do firms form their expectations? new survey evidence. American Economic Review 108, 2671–2713.

Coibion, O., Gorodnichenko, Y., Kumar, S., Ryngaert, J., 2021. Do you know that i know that you know...? higher-order beliefs in survey data. The Quarterly Journal of Economics 136, 1387–1446.

De Loecker, J., 2011. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. Econometrica 79, 1407–1451.

Denti, T., 2015. Games with unrestricted information acquisition. Technical Report. Mimeo.

Fisher, T.C., Konieczny, J.D., 2000. Synchronization of price changes by multiproduct firms: evidence from canadian newspaper prices. Economics Letters 68, 271–277.

Gertler, M., Leahy, J., 2008. A phillips curve with an Ss foundation. Journal of Political Economy 116, 533–572.

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

Gorodnichenko, Y., 2008. Endogenous information, menu costs and inflation persistence. Working Paper 14184. National Bureau of Economic Research.

Jung, J., Kim, J.H.J., Matějka, F., Sims, C.A., 2019. Discrete actions in information-constrained decision problems. The Review of Economic Studies 86, 2643–2667.

Karadi, P., Reiff, A., 2019. Menu costs, aggregate fluctuations, and large shocks. American

Economic Journal: Macroeconomics 11, 111–46.

Klenow, P.J., Malin, B.A., 2010. Microeconomic evidence on price-setting., in: Friedman, B.M., Woodford, M. (Eds.), Handbook of Monetary Economics. Elsevier. volume 3, pp. 231–284.

Kumar, S., Afrouzi, H., Coibion, O., Gorodnichenko, Y., 2015. Inflation targeting does not anchor inflation expectations: Evidence from firms in new zealand. Brookings Papers on Economic Activity 2015, 151–225.

Lach, S., Tsiddon, D., 1996. Staggering and synchronization in price-setting: Evidence from multiproduct firms. The American Economic Review 86, 1175–1196.

Lach, S., Tsiddon, D., 2007. Small price changes and menu costs. Managerial and Decision Economics 28, 649–656.

Letterie, W., Nilsen, Ø.A., 2016. Price changes-stickiness and internal coordination in multiproduct firms. CESifo Working Paper Series 5701. CESifo Group Munich.

Levy, D., Bergen, M., Dutta, S., Venable, R., 1997. The magnitude of menu costs: Direct evidence from large U.S. supermarket chains. The Quarterly Journal of Economics 112, 791–825.

Maćkowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. American Economic Review 99, 769–803.

Matějka, F., 2015. Rationally inattentive seller: Sales and discrete pricing. The Review of Economic Studies 83, 1125–1155.

Midrigan, V., 2011. Menu costs, multiproduct firms, and aggregate fluctuations. Econometrica 79, 1139–1180.

Nakamura, E., Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. The Quarterly Journal of Economics 123, 1415–1464.

Nakamura, E., Steinsson, J., 2013. Price rigidity: Microeconomic evidence and macroeconomic implications. Annual Review of Economics 5, 133–163.

Parker, M., 2017. Price-setting behaviour in new zealand. New Zealand Economic Papers 51, 217–236.

Pasten, E., Schoenle, R., 2016. Rational inattention, multi-product firms and the neutrality of money. Journal of Monetary Economics 80, 1–16.

Steiner, J., Stewart, C., Matějka, F., 2017. Rational inattention dynamics: Inertia and delay in decision-making. Econometrica 85, 521–553.

Stella, A., 2020. Economies of scope in price setting: A moment inequalities estimation. Journal of Monetary Economics 110, 50–61.

Tauchen, G., 1986. Finite state markov-chain approximations to univariate and vector autoregressions. Economics Letters 20, 177–181.

Tsiddon, D., 1993. The (mis)behaviour of the aggregate price level. The Review of Economic Studies 60, 889–902.

Vavra, J., 2013. Inflation dynamics and time-varying volatility: New evidence and an ss interpretation. The Quarterly Journal of Economics 129, 215–258.

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

Appendix For Online Publication

Appendix A. Discussion on Model Assumptions

In this appendix, I discuss two key assumptions in my baseline model with both menu costs and rational inattention. First, I provide some evidence of firm-specific menu costs from the survey data. Second, I discuss implications and limitations of the assumptions about the set of available information.

Appendix A.1. Alternative Assumptions on Menu Costs

In the baseline two-good version of the model, economies of scope in price setting emerge from the existence of firm-specific menu costs. Previous literature found ample evidence of the firm-specific menu costs for multi-product firm. For example, recent work by Stella (2020) and Letterie and Nilsen (2016) directly estimate various types of adjustment costs for the multi-product firms and find that there are sizable component of costs from firm-specific menu costs.[45] In the New Zealand survey data, I also find some evidence of firm-specific menu costs. Managers were asked about how typical it is to synchronize price reviews and price changes across multiple products sold by their firms. They report that on average 75% of their price changes and price review decisions are synchronized within their firms.

While the firm-specific fixed cost implies perfect within-firm synchronization of price changes, the data show that firms synchronize their price changes partially. Bonomo et al. (2020) also find partial synchronization using Israel retail price data and show that even small departures from full synchronization in menu costs models substantially weaken monetary non-neutrality. This implies that introducing a product-specific menu cost in my baseline model would weaken economies of scope in price setting. In this case, the real effects of monetary shocks in the two-good version of the model will be much smaller than those in the single-product model.

In the baseline models, I calibrate the size of menu costs in the single- and two-product models to match the size and the frequency of prices changes in the New Zealand data. Alternatively, one could consider menu costs that scale linearly in the number of products. Given the calibrated parameters for the two-product model, I linearly scale down the size of menu costs for the single-product model, which is slightly bigger than the baseline calibration. As shown in Appendix Figure G.8 (green line), under this alternative calibration, the output response to a monetary policy shock is very similar to the output response in the single-product baseline model. Also, in the simplified

[45] Bhattarai and Schoenle (2014) and Lach and Tsiddon (2007) test the implications of menu cost models with a single fixed menu cost and find that micro price data support the existence of economies of scope for multi-product firms. Moreover, Lach and Tsiddon (1996) Fisher and Konieczny (2000), and Midrigan (2011) find that price changes within multi-product firms are highly synchronized, suggesting the existence of firm-specific menu costs.

version of the model, I consider this alternative menu cost assumption of linear scaling. Appendix Figure G.9 shows that the cumulative output responses in the menu-cost-only models increase in the number of products, while those in the model with both rational inattention and menu costs decrease in the number of products. Notice that compared to the simplified version of the model in Section 4.7, monetary non-neutrality in the model with many products is smaller since a larger size of menu costs makes the selection effects of price changes bigger.

Lastly, one could have both fixed and variable costs of price changes. For example, using Norwegian price data, Letterie and Nilsen (2016) estimate the economically significant size of fixed and variable costs of firms' price changes. In particular, they find that the size of quadratic price adjustment cost is quite larger than the size of fixed menu cost. Motivated by this finding, I consider the simplified version of models with both fixed and quadratic price adjustment costs. I calibrate the ratio of quadratic adjustment costs and fixed menu costs to be similar to the empirical estimate in Letterie and Nilsen (2016). Appendix Figure G.10 shows that my main results are qualitatively robust to adding the variable cost of price changes in the model.

Appendix A.2. Set of Available Information.

In my main model analysis, I assume that the set of available signals has three properties. First, the firm chooses $N+1$ independent signals for each shock, implying that paying attention to aggregate conditions and paying attention to good-specific idiosyncratic conditions are separate activities. Although this assumption is often made in the rational inattention literature, such as Maćkowiak and Wiederholt (2009) and Pasten and Schoenle (2016), it might be suboptimal for firms to choose to observe independent signals. In fact, Afrouzi and Yang (2021) show that in LQG setting rational inattention models (without menu costs), the number of signals that firms choose to observe are no more than the number of actions. In my model, the number of actions is N as firms choose N prices for their goods, implying that firms might waste their resources to observe additional signals.

Second, I assume that every signal is Gaussian. Gaussian signals are optimal when the underlying shocks are Gaussian and firms' objective function is quadratic. However, in my model, firms' objective is not quadratic due to menu costs. Recent rational inattention literature considers models with general objective functions with some assumptions of a simple stochastic process, a static setup or finite actions and states (Matějka, 2015; Jung et al., 2019; Steiner et al., 2017). Solving fully non-linear dynamic problems under rational inattention is computationally demanding as firms' state variable is an infinitely dimensional object if the shocks are continuously distributed. The assumption of Gaussian signals is for tractability.

Third, I assume that all noise in signals is idiosyncratic and independent. This assumption is without loss of generality since I consider Shannon's mutual information as the cost of information

(Denti, 2015; Afrouzi, 2020).

Appendix A.3. Alternative Assumptions on Informational Costs

In the baseline model, I assume that the cost of information is linear in the Shannon's mutual information. Pasten and Schoenle (2016) consider alternative ways of modeling rational inattention costs, such as a constant shadow price of information capacity per good and a constant loss per good from rational inattention friction. As a robustness analysis, I first assume that the loss per good from rational inattention is constant regardless of the number of products that firms produce. Given the calibrated parameters for the two-product model, I change the marginal cost of information for the single-product model to have the same per good loss from the rational inattention with that in the two-product model. The output response to a monetary shock in this alternative model is show in Appendix Figure G.8 (purple line). Compared to the baseline single-product model, this alternative assumption on the informational costs implies that the firms are less informed about both idiosyncratic and aggregate shocks, leading to a more persistent and bigger real effect of monetary shocks.

Second, I assume a constant shadow price of information capacity per good. This means that the marginal cost of Shannon's mutual information is linearly scaled in the number of products. To this end, given the calibrated parameters for the two-product model, I scale down the marginal cost of information for the single-product model. The output response in this alternative model is shown in Appendix Figure G.8 (orange line). In this case, the output response in the two-product model is bigger than in the single-product model because, unlike our baseline calibration, the economies of scope motive in price setting dominates the economies of scope motive in information processing. However, this alternative way of modeling informational cost implies that firms' attentiveness to aggregate economy is smaller in the two-product model than in the single-product model, which is not consistent with my empirical finding.

Appendix A.4. Correlated Shocks

In my baseline model, I assume that both good-specific and aggregate shocks are independently identically distributed. We can think of alternative assumptions of correlated shocks. First, one can assume that there is a correlation across good-specific shocks. A simple extreme example would be introducing a firm-specific shock which simultaneously affects the marginal costs of all products. In this case, this perfectly correlated shock can generate synchronization of price changes. Due to this correlation, learning about the firm-specific shock would be informative for frictionless optimal prices for all products. This makes the firm allocate more attention to this shock, and weakens the economies of scope motive in information processing for monetary shocks. The same logic applies when we consider a partial correlation across good-specific shocks.

Also, one can introduce serial correlation for the underlying shocks. As discussed in Pasten and Schoenle (2016), depending on the relative persistence of good-specific and aggregate shocks, there would be intertemporal economies of scope motive. On the one hand, if the persistence of good-specific shock decreases, it would be less useful to pay attention to the good-specific shock for their future decision. In particular, this force is getting stronger when we think of dynamic rational inattention problem. On the other hand, lower persistence increases the (unconditional) volatility of the exogenous disturbances of good-specific shock, which leads to a strong incentive to allocate attention to the good-specific shock. However, Pasten and Schoenle (2016) shows that, quantitatively, this class of the model can only match the average size of price changes in the data if the volatility of each good-specific shock is negatively correlated with its persistence. This suggests that when the persistence of good-specific shock decreases, we will have smaller monetary non-neutrality.

Appendix B. A Rationally Inattentive Firm's Problem without Menu Costs ($\theta = 0$)

In this appendix, I solve a rationally inattentive firm's problem without menu costs. This problem is similar to the one studied in Pasten and Schoenle (2016) with one main difference. They solve the problem by assuming that the cost of information is not discounted and optimizing at the long-run steady-state for information structure. Here, I assume that the firm discounts future costs of information at the same discount rate as their payoffs and solve the dynamic information acquisition problem. This setup is also similar to Afrouzi and Yang (2021) that study the dynamic multivariate rational inattention problem. One difference is that I assume that the set of available signals are partitioned into two subsets, one for signals about idiosyncratic shocks and the other for signals about aggregate shocks, that are independent each other.

Without menu costs, the firm can change its prices and collapses the price gaps to zero whenever it wants, i.e. there is no micro rigidity in price setting. In this case, the firm's prior price gaps are no longer its state variables, and thus the firm's problem is deterministic. Then, in a recursive formation, the firm's problem is:

$$
\begin{aligned}
V\left(\{z^a_{j,-1}\}_{j=1}^N, z^m_{-1}\right) = \max_{\{\{z^a_j\}_{j=1}^N, z^m\}} \; & -B\sum_{j=1}^N \left(z^a_j + z^m\right) + \beta V\left(\{z^a_j\}_{j=1}^N, z^m\right) \\
& -\frac{\psi}{2}\left(\sum_{j=1}^N \log_2\left(\frac{z^a_{j,-1}+\sigma_a^2}{z^a_j}\right) + \log_2\left(\frac{z^m_{-1}+\sigma_m^2}{z^m}\right)\right) \\
\text{s.t.} \qquad & 0 \le z^a_j \le z^a_{j,-1} + \sigma_a^2, \quad \forall j = 1, 2, \cdots, N, \\
& 0 \le z^m \le z^m_{-1} + \sigma_m^2,
\end{aligned}
$$

Notice that with $\psi > 0$, the constraints $z^a_j \ge 0$ and $z^m_j \ge 0$ will not bind. The first order necessary conditions are:

$$
\begin{aligned}
\partial z_j^a: & \; -B+\frac{\psi}{2\log 2}\frac{1}{z_j^a}+\beta V_{z_j^a}\left(\{z_j^a\}_{j=1}^N, z^m\right)-\phi_j=0, \;\; \forall\, j\in\{1,2,\cdots,N\},\\
\partial z^m: & \; -BN+\frac{\psi}{2\log 2}\frac{1}{z^m}+\beta V_{z^m}\left(\{z_j^a\}_{j=1}^N, z^m\right)-\phi_m=0,\\
& V_{z_{j,-1}^a}\left(\{z_{j,-1}^a\}_{j=1}^N, z_{-1}^m\right)=-\frac{\psi}{2\log 2}\frac{1}{z_{j,-1}^a+\sigma_a^2} \;\; \forall\, j\in\{1,2,\cdots,N\},\\
& V_{z_{-1}^m}\left(\{z_{j,-1}^a\}_{j=1}^N, z_{-1}^m\right)=-\frac{\psi}{2\log 2}\frac{1}{z_{-1}^m+\sigma_m^2},
\end{aligned}
$$

and complementarity slackness conditions where $\{\phi_j\}_{j=1}^N$ and ϕ_m are Lagrangian multipliers for no-forgetting constraints. The no-forgetting constraints will bind when the marginal cost of information processing is high enough. Here, I focus on interior solutions where the constraints do not bind.

When the standard deviation of idiosyncratic good-specific shocks is the same, then subjective uncertainty about idiosyncratic shocks is also the same across all goods. Then, the optimal subjective uncertainty about both shocks satisfies:

$$
\begin{aligned}
B=&\frac{\psi}{2\log 2}\left(\frac{1}{z_j^a}-\beta\frac{1}{z_j^a+\sigma_a^2}\right), \;\; \forall\, j\in\{1,2,\cdots,N\}\\
B\cdot N=&\frac{\psi}{2\log 2}\left(\frac{1}{z^m}-\beta\frac{1}{z^m+\sigma_m^2}\right).
\end{aligned} \tag{B.1}
$$

Several interesting characteristics emerge. First, the firm's optimal subjective uncertainty is constant while it is time-varying with menu costs. Second, subjective uncertainty increases in the size of marginal cost of information processing, ψ, and the size of shocks, σ_a^2, while it decreases in the slope of profit function , B, and the time preference parameter, β. Third, optimal subjective uncertainty about idiosyncratic shocks is independent of the number of products that the firm produces. Fourth, optimal subjective uncertainty about aggregate shocks is *decreasing* in the number of products ($\frac{\partial z^m(N)}{\partial N}<0$) and $\lim_{N\to\infty} z^m(N)=0$.

Appendix B.1. Real Effects of Monetary Policy Shocks

Let $\tilde{z}^m=\frac{z^m}{\sigma_m^2}$ be firms' subjective uncertainty relative to the variance of monetary policy shocks. Then I can rewrite Equation (B.1) as:

$$
\frac{1}{\tilde{z}^m}-\beta\frac{1}{\tilde{z}^m+1}=\sigma_m^2\frac{BN}{\psi}\left(2\log 2\right) \tag{B.2}
$$

Given parameters, firms' subjective uncertainty about monetary shocks decreases in their number of products.

Now, the size of price changes for good j is given by:

$$
\begin{aligned}
\Delta p_{i,j,t}=&\mathscr{K}^a(N)\left(a_{i,j,t-1}-a_{i,j,t-1|t-1}+\varepsilon_{i,j,t}+\eta_{i,j,t}\right)\\
&+\mathscr{K}^m(N)\left(m_{t-1}-m_{i,t-1|t-1}+\varepsilon_{m,t}+\eta_{i,m,t}\right),
\end{aligned}
$$

where

$$a_{i,jt} - a_{i,j,t|t} = (1 - \mathscr{K}^a(N)) \left(a_{i,j,t-1} - a_{i,j,t-1|t-1} + \varepsilon_{i,j,t} \right) - \mathscr{K}^a(N) \eta_{i,j,t},$$
$$m_t - m_{i,t|t} = (1 - \mathscr{K}^m(N)) \left(m_{t-1} - m_{i,t-1|t-1} + \varepsilon_{m,t} \right) - \mathscr{K}^m(N) \eta_{i,m,t},$$

and Kalman gains are:

$$\mathscr{K}^a(N) = \frac{1}{1 + \tilde{z}_j^a(N)}, \;\; \mathscr{K}^m(N) = \frac{1}{1 + \tilde{z}^m(N)}.$$

Let $p_{j,t} = \int p_{i,j,t} di$. Let $p_{j,t}^{\text{NoMP}}$ and $p_{j,t}^{\text{MP}}$ be price level at time t without and with monetary shocks, respectively. Then, since all noise in signals, $\eta_{i,j,t}$, is idiosyncratic and independent, we have

$$\begin{aligned} p_{j,t}^{\text{NoMP}} &= p_{j,t-1}^{\text{NoMP}} + \Delta p_{j,t}^{\text{NoMP}} \\ &= p_{j,t-1}^{\text{NoMP}} + \left(\mathscr{K}^a(N) \int \left(a_{i,j,t-1} - a_{i,j,t-1|t-1} \right) di + \mathscr{K}^m(N) \int \left(m_{t-1} - m_{i,t-1|t-1} \right) di \right) \end{aligned}$$

and

$$\begin{aligned} p_{j,t}^{\text{MP}} &= p_{j,t-1}^{\text{MP}} + \Delta p_{j,t}^{\text{MP}} \\ &= p_{j,t-1}^{\text{MP}} + \left(\mathscr{K}^a(N) \int \left(a_{i,j,t-1} - a_{i,j,t-1|t-1} \right) di + \mathscr{K}^m(N) \int \left(m_{t-1} - m_{i,t-1|t-1} \right) di + \varepsilon_{m,t} \right) \end{aligned}$$

Then, Notice that by symmetry across goods, we have $p_t = p_{j,t}$ for all j. Define an impulse response of aggregate price to a monetary shock as the gap between the prices with and without the monetary shock, that is,

$$IRF_t^P = p_t^{\text{MP}} - p_t^{\text{NoMP}}.$$

Assume at time 0, there is a monetary shock, $\varepsilon_{m,0}$. Then,

$$\begin{aligned} IRF_0^P &= \mathscr{K}^m(N) \varepsilon_{m,0} \\ IRF_1^P &= \mathscr{K}^m(N) \varepsilon_{m,0} + \mathscr{K}^m(N) \left(1 - \mathscr{K}^m(N) \right) \varepsilon_{m,0} \\ IRF_2^P &= \mathscr{K}^m(N) \varepsilon_{m,0} + \mathscr{K}^m(N) \left(1 - \mathscr{K}^m(N) \right) \varepsilon_{m,0} + \mathscr{K}^m(N) \left(1 - \mathscr{K}^m(N) \right)^2 \varepsilon_{m,0} \\ &\vdots \\ IRF_t^P &= \mathscr{K}^m(N) \left(\frac{1 - \left(1 - \mathscr{K}^m(N) \right)^{t+1}}{1 - \left(1 - \mathscr{K}^m(N) \right)} \right) \varepsilon_{m,0} \\ &= \left(1 - \left(1 - \mathscr{K}^m(N) \right)^{t+1} \right) \varepsilon_{m,0} \end{aligned}$$

and output responses are given by

$$\begin{aligned} IRF_t^Y &= \varepsilon_{m,0} - IRF_t^P \\ &= \left(1 - \mathscr{K}^m(N) \right)^{t+1} \varepsilon_{m,0}. \end{aligned}$$

Let a cumulative impulse response of output as a function of the number of product, $M(N)$ be the area under the impulse response function of output. Then,

$$\begin{aligned}\mathscr{M}(N) &= \int_0^\infty IRF_t^Y dt = \int_0^\infty (1-\mathscr{K}^m(N))^{t+1}\varepsilon_{m,0}dt \\ &= -\frac{(1-\mathscr{K}^m(N))}{\log(1-\mathscr{K}^m(N))}\varepsilon_{m,0} \\ &= -\frac{\left(\frac{\tilde{z}_m(N)}{1+\tilde{z}_m(N)}\right)}{\log\left(\frac{\tilde{z}_m(N)}{1+\tilde{z}_m(N)}\right)}\varepsilon_{m,0}.\end{aligned}$$

where $\tilde{z}_m(N)$ is a solution of Equation (B.2) as a function of N. Notice that

$$\frac{\partial \mathscr{M}(N)}{\partial N} = -\underbrace{\frac{\frac{\partial \tilde{z}^m(N)}{\partial N}}{(1+\tilde{z}^m(N))^2}}_{<0} \times \underbrace{\frac{1}{\log\left(\frac{\tilde{z}_m(N)}{1+\tilde{z}_m(N)}\right)}}_{<0} \times \underbrace{\left(1-\frac{1}{\log\left(\frac{\tilde{z}_m(N)}{1+\tilde{z}_m(N)}\right)}\right)}_{>0} < 0.$$

Although N is an arbitrary integer number, here I assume $\mathscr{M}(N)$ is continuously differentiable with respect to N. Appendix Figure G.11 shows that both subjective uncertainty about monetary shocks and cumulative responses of output to monetary shocks decreases in number of products that firms produce.

Appendix C. Quadratic Approximation to a Firm's Profit Function

Firm i produces N products indexed by j in monopolistic competitive markets. Its demand for good j is given by

$$Y_{i,j,t} = A_{i,j,t}^{\varepsilon-1}\left(\frac{P_{i,j,t}}{P_{j,t}}\right)^{-\varepsilon}\left(\frac{P_{j,t}}{P_t}\right)^{-\gamma} Y_t,$$

where P_t is the price of aggregate output Y_t, $P_{j,t}$ is the price of good j, γ is the constant elasticity of substitution across different firms that produce the same good, and ε is the constant elasticity of substitution across different goods. Then, the firm's profit function is

$$\begin{aligned}\Pi_{i,t} &= \sum_{j=1}^{N}(P_{i,j,t}-W_tA_{i,j,t})Y_{i,j,t} \\ &= \sum_{j=1}^{N}(P_{i,j,t}-W_tA_{i,j,t})A_{i,j,t}^{\varepsilon-1}\left(\frac{P_{i,j,t}}{P_{j,t}}\right)^{-\varepsilon}\left(\frac{P_{j,t}}{P_t}\right)^{-\gamma}Y_t,\end{aligned}$$

where

$$P_t = \left(\frac{1}{N}\sum_{j=1}^{N}P_{j,t}^{1-\gamma}\right)^{\frac{1}{1-\gamma}}, \quad P_{j,t} = \left(\int_0^1 (P_{i,j,t})^{1-\varepsilon}di\right)^{\frac{1}{1-\varepsilon}}.$$

Define firm i's markup for good j, $\mu_{i,j,t} = \frac{P_{i,j,t}}{W_tA_{i,j,t}}$. Then, the profit function can be written as a function of the firm's markups for each good:

$$\begin{aligned}\Pi_{i,t} &= \sum_{j=1}^{N} \left(P_{i,j,t} - W_t A_{i,j,t}\right) A_{i,j,t}^{\varepsilon-1} \left(\frac{P_{i,j,t}}{P_{j,t}}\right)^{-\varepsilon} \left(\frac{P_{j,t}}{P_t}\right)^{-\gamma} Y_t \\ &= \sum_{j=1}^{N} \left(\mu_{i,j,t} - 1\right) \left(\mu_{i,j,t}\right)^{-\varepsilon} \left(W_t\right)^{1-\varepsilon} \left(P_{j,t}\right)^{\varepsilon-\gamma} \left(P_t\right)^{\gamma} Y_t.\end{aligned}$$

Let $R_{i,j,t}$ be the revenue from good j:

$$\begin{aligned}R_{i,j,t} &= P_{i,j,t} A_{i,j,t}^{\varepsilon-1} \left(\frac{P_{i,j,t}}{P_{j,t}}\right)^{-\varepsilon} \left(\frac{P_{j,t}}{P_t}\right)^{-\gamma} Y_t \\ &= \mu_{i,j,t} \left(\mu_{i,j,t}\right)^{-\varepsilon} \left(W_t\right)^{1-\varepsilon} \left(P_{j,t}\right)^{\varepsilon-\gamma} \left(P_t\right)^{\gamma} Y_t.\end{aligned}$$

A second order approximation to the profit function around the optimal frictionless markup, $\mu_j^* = \frac{\varepsilon}{\varepsilon-1}$, yields

$$\begin{aligned}\Pi\left(\left\{\mu_{j,t}\right\}_{j=1}^{N}\right) &= \Pi\left(\left\{\mu_j^*\right\}_{j=1}^{N}\right) + \frac{1}{2} \sum_{j=1}^{N} \left.\frac{\partial^2 \Pi_t}{\partial \mu_{j,t}^2}\right|_{\left\{\mu_{j,t}=\mu_j^*\right\}_{j=1}^{N}} \left(\frac{\mu_{j,t} - \mu_j^*}{\mu_j^*}\right)^2 \left(\mu_j^*\right)^2 \\ &= \Pi\left(\left\{\mu_j^*\right\}_{j=1}^{N}\right) + \frac{1}{2} \sum_{j=1}^{N} \left.\frac{\partial^2 \Pi_t}{\partial \mu_{j,t}^2}\right|_{\left\{\mu_{j,t}=\mu_j^*\right\}_{j=1}^{N}} \left(\hat{\mu}_{j,t}\right)^2 \left(\mu_j^*\right)^2,\end{aligned}$$

where $\hat{\mu}_{j,t} = \log\left(\mu_{j,t}\right) - \log\left(\mu_j^*\right)$ is the realized markup-gap. Then, given the CES demand and the constant returns to scale technology, we can express the expected losses that arise from frictions (both nominal and informational) relative to the frictionless case, expressed as a fraction of per-good revenue:

$$\mathscr{L} \equiv \mathbb{E}\left[\left.\frac{\Pi\left(\{\mu_{j,t}\}_{j=1}^{N}\right) - \Pi\left(\left\{\mu_j^*\right\}_{j=1}^{N}\right) - \tilde{\theta}\mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} - \tilde{\psi}\mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^{N},W_t\}|S_i^{t-1})}{R\left(\mu_j^*\right)}\right| S^{t-1}\right]$$

$$= \mathbb{E}\left[\frac{1}{2}\frac{1}{R\left(\mu_j^*\right)}\sum_{j=1}^{N}\left.\frac{\partial^2\Pi_t}{\partial\mu_{j,t}^2}\right|_{\left\{\mu_{j,t}=\mu_j^*\right\}_{j=1}^{N}}(\hat{\mu}_{j,t})^2\left(\mu_j^*\right)^2\right.$$

$$\left. - \theta\mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} - \psi\mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^{N},W_t\}|S_i^{t-1})\right| S^{t-1}\right]$$

$$= \mathbb{E}\left[\frac{1}{2}\frac{\Pi\left(\left\{\mu_j^*\right\}_{j=1}^{N}\right)}{R\left(\mu_j^*\right)}\sum_{j=1}^{N}\frac{\left(\mu_j^*\right)^2\left.\frac{\partial^2\Pi_t}{\partial\mu_{j,t}^2}\right|_{\left\{\mu_{j,t}=\mu_j^*\right\}_{j=1}^{N}}}{\Pi\left(\left\{\mu_j^*\right\}_{j=1}^{N}\right)}(\hat{\mu}_{j,t})^2\right.$$

$$\left. - \theta\mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} - \psi\mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^{N},W_t\}|S_i^{t-1})\right| S^{t-1}\right],$$

where

$$\left.\frac{\partial^2\Pi_t}{\partial\mu_{j,t}^2}\right|_{\{\mu_{j,t}=\mu_j^*\}_{j=1}^{N}} = \varepsilon\left(\mu_j^*\right)^{-\varepsilon-2}\left[(\varepsilon+1)\left(\mu_j^*-1\right)-2\mu_j^*\right](\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y}$$

$$= -\varepsilon\left(\mu_j^*\right)^{-\varepsilon-2}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y},$$

$$\Pi\left(\{\mu_j^*\}_{j=1}^{N}\right) = \sum_{j=1}^{N}\left(\mu_j^*-1\right)\left(\mu_j^*\right)^{-\varepsilon}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y},$$

and

$$R\left(\mu_j^*\right) = \left(\mu_j^*\right)^{1-\varepsilon}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y}.$$

Notice that I express the cost of change price, $\tilde{\theta}$, as a fraction θ of the steady state frictionless revenue from selling one of N products, that is $\tilde{\theta} = \theta R(\mu_j^*)$. Similarly, the marginal cost of information processing is $\tilde{\psi} = \psi R(\mu_j^*)$.

Then, the loss function is

$$
\begin{aligned}
\mathscr{L} = \mathbb{E}\Bigg[& \frac{1}{2}\frac{\Pi\left(\left\{\mu_j^*\right\}_{j=1}^N\right)}{R\left(\mu_j^*\right)}\sum_{j=1}^N \frac{\left(\mu_j^*\right)^2 \left.\frac{\partial^2 \Pi_t}{\partial \mu_{j,t}^2}\right|_{\left\{\mu_{j,t}=\mu_j^*\right\}_{j=1}^N}}{\Pi\left(\left\{\mu_j^*\right\}_{j=1}^N\right)}\left(\hat{\mu}_{j,t}\right)^2 \\
& - \theta \mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} - \psi \mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^N, W_t\}|S_i^{t-1}) \Big| S^{t-1}\Bigg] \\
= -\mathbb{E}\Bigg[& \varepsilon \frac{1}{2}\left(\frac{\sum_{j=1}^N \left(\mu_j^*-1\right)\left(\mu_j^*\right)^{-\varepsilon}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y}}{\left(\mu_j^*\right)^{1-\varepsilon}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y}}\right) \\
& \times \frac{\sum_{j=1}^N \left(\mu_j^*\right)^{-\varepsilon}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y}}{\sum_{j=1}^N \left(\mu_j^*-1\right)\left(\mu_j^*\right)^{-\varepsilon}(\bar{W})^{1-\varepsilon}(\bar{P}_j)^{\varepsilon-\gamma}(\bar{P})^{\gamma}\bar{Y}}\left(\hat{\mu}_{j,t}\right)^2 \\
& + \theta \mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} + \psi \mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^N, W_t\}|S_i^{t-1}) \Big| S^{t-1}\Bigg].
\end{aligned}
$$

Now, assume $\varepsilon = \gamma$ (or by symmetry across product industry due to there are no common industry specific shocks). Then, the second order approximation of the firm's profit function is

$$
\mathscr{L} = \mathbb{E}\left[-\varepsilon\frac{1}{2}\left(\frac{1}{\mu_j^*}\right)\sum_{j=1}^N \left(\hat{\mu}_{i,j,t}\right)^2 + \theta \mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} + \psi \mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^N, W_t\}|S_i^{t-1}) \Big| S^{t-1}\right].
$$

Let $p_{i,j,t}^* = \log(W_t) + \log(A_{i,j,t})$ be the log deviation of (frictionless) optimal price of good j from its non-stochastic steady state. Then, we define firm i's true price gap of good j as

$$
\hat{\mu}_{i,j,t} = p_{i,j,t} - p_{i,j,t}^*,
$$

where $p_{i,j,t}$ is the log deviation of the price of good j from its non-stochastic steady state.[46] Then, I derive a second order approximation of firms' loss function:

$$
\begin{aligned}
\mathscr{L} &= \mathbb{E}\left[-\varepsilon\frac{1}{2}\left(\frac{1}{\mu_j^*}\right)\sum_{j=1}^N \left(p_{i,j,t} - p_{i,j,t}^*\right)^2 + \theta \mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} + \psi \mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^N, W_t\}|S_i^{t-1}) \Big| S^{t-1}\right] \\
&= -\mathbb{E}\left[B\sum_{j=1}^N \left(p_{i,j,t} - p_{i,j,t}^*\right)^2 + \theta \mathbf{1}_{\{\text{for any } j,\ p_{i,j,t}\neq p_{i,j,t-1}\}} + \psi \mathscr{I}(s_{i,t};\{\{A_{i,j,t}\}_{j=1}^N, W_t\}|S_i^{t-1}) \Big| S^{t-1}\right],
\end{aligned}
$$

where $B = \frac{\varepsilon-1}{2}$ is a slope of profit curve.

[46] The true price gap $\hat{\mu}_{i,j,t}$ can be also written as a markuk gap, $\hat{\mu}_{i,j,t} = \log(\mu_{i,j,t}/\mu_j^*)$, which is a the log deviation of the current markup to the non-stochastic steady state markup $\mu_j^* = \frac{\varepsilon}{\varepsilon-1}$.

Appendix D. Computational Procedures for the Two-Product Model

I use the method of value function iteration to solve the two-product firms' optimization problem. There are 5 state variables for the problem: prior perceived price gap for two products, subjective uncertainty about each good-specific shock, and prior subjective uncertainty about monetary policy shocks. Since this problem is non-convex optimization problem with occasionally binding constraints, it should b e solved numerically.

The most computationally burdensome part is to compute firms' expected future values since 1) tomorrow's perceived price gaps are stochastic variables when firms do not change prices today, and 2) the distribution of these price gaps has a mean vector $(x_{1,-1}, x_{2,-1})'$, which is firms' state variable, and a covariance matrix, Σ, which is firms' choice variable. Standard approximation methods for the transition probability of states, such as Tauchen approximation method, are not applicable since the approximation errors are quite large. I compute expected value of the firms' value functions using Gauss-Legendre quadrature which is an explicit numerical integration technique.

I solve the firms' problem and the value function and the optimal policy functions using the following procedure:

1. Construct grids for individual state variables, such as prior of perceived price gaps for each product, $x_{1,-1}$, $x_{2,-1}$, prior subjective uncertainty about two good-specific shocks, $z^a_{1,-1}$, $z^a_{2,-1}$, and prior subjective uncertainty about monetary shocks, z^m_{-1}. I use 21 grids for $x_{1,-1}$, $x_{2,-1}$, and 16 grids for $z^a_{1,-1}$, $z^a_{2,-1}$, and z^m_{-1}. The ranges of $x_{1,-1}$ and $x_{2,-1}$, are $[-1.5\sqrt{\theta/B}, 1.5\sqrt{\theta/B}]$ where θ is the size of menu costs and B is the slope of firms' profit curve.[47] More grid points are assigned around inaction bands. $z^a_{1,-1}$, $z^a_{2,-1}$, and z^m_{-1} are equally spaced in the range of [0 0.004].
2. Construct the abscissas, $\{\tilde{x}_i\}_{i=1}^{N_q}$, and weights, $\{\tilde{w}_i\}_{i=1}^{N_q}$, of the Gauss-Legendre quadrature with $N_q = 500$ points.
3. Solve the individual value functions at each grid point. In this step, I obtain the optimal decision rules for subjective uncertainty about both good-specific shocks and monetary shocks,

$$g^a_1(x_{1,-1}, x_{2,-1}, z^a_{1,-1}, z^a_{2,-1}, z^m_{-1}),\ g^a_2(x_{1,-1}, x_{2,-1}, z^a_{1,-1}, z^a_{2,-1}, z^m_{-1}),\ g^m(x_{1,-1}, x_{2,-1}, z^a_{1,-1}, z^a_{2,-1}, z^m_{-1}),$$

   and the value function, $V(x_{1,-1}, x_{2,-1}, z^a_{1,-1}, z^a_{2,-1}, z^m_{-1})$. The detailed steps are as follows:
   (a) Make an initial guess for the value functions, V_0, for all grid points.
   (b) Solve firms' optimization problem and compute V_1. Notice that the problem can be

[47] In fact $[-\sqrt{\frac{\theta}{B}}, \sqrt{\frac{\theta}{B}}]$ is the inaction bands for myopic firms ($\beta = 0$). In this regard, I have more conservative ranges of grids for prior price gaps.

written:

$$
\begin{aligned}
&V_1\left(x_{1,-1},x_{2,-1},z^a_{1,-1},z^a_{2,-1},z^m_{,-1}\right) = \\
&\quad \max_{\{z^a_1,z^a_2,z^m\}} -B\left(z^a_1+z^a_2+2z^m\right)-\frac{\psi}{2}\left(\sum_{j=1}^{2}\log_2\left(\frac{z^a_{j,-1}+\sigma^2_a}{z^a_j}\right)+\log_2\left(\frac{z^m_{-1}+\sigma^2_m}{z^m}\right)\right) \\
&\qquad +\int_{(x_1,x_2)}\max\Big\{-B(x_1^2+x_2^2)+\beta V_0\left(x_1,x_2,z^a_1,z^a_2,z^m\right), \\
&\qquad\qquad -\theta+\beta V_0\left(0,0,z^a_1,z^a_2,z^m\right)\Big\}dF\left((x_1,x_2);(x_{1,-1},x_{2,-1}),\Sigma\right) \\
&\text{s.t.}\qquad 0\le z^a_j\le z^a_{j,-1}+\sigma^2_a,\quad \forall j=1,2 \\
&\qquad\quad 0\le z^m\le z^m_{-1}+\sigma^2_m \\
&\qquad\quad \Sigma_t(j,k)=\begin{cases} z^m_{-1}+\sigma^2_m-z^m & \text{if } j\ne k \\ z^a_{j,-1}+\sigma^2_a-z^a_j+z^m_{-1}+\sigma^2_m-z^m & \text{if } j=k. \end{cases}
\end{aligned}
$$

where $F\left((x_1,x_2);(x_{1,-1},x_{2,-1}),\Sigma\right)$ is a joint normal distribution with mean $(x_{1,-1},x_{2,-1})$ and covariance matrix Σ.

(c) If V_0 and V_1 are close enough for each grid point, and go to the next step. Otherwise, update the value functions ($V_0 = V_1$), and go back to (a).

(d) Simulate the model with a large number of firms to obtain a stationary distribution, $G(x_1,x_2,z^a_1,z^a_2,z^m)$, over firm states $(x_1,x_2,z^a_1,z^a_2,z^m)$. Simulation algorithm is described in Appendix D.1.

(e) Compute aggregate variables.

Appendix D.1. Simulation Algorithm for the Two-Product Model

I simulate the two-good version of the baseline model with 100,000 firms for 5,000 periods.

1. Set initial $x_{i,1,t-1}$, $x_{i,2,t-1}$, $z^a_{i,1,t-1}$, $z^a_{i,2,t-1}$, and $z^m_{i,t-1}$.
2. Generate random numbers for the shocks $\varepsilon^m_t \sim N(0,\sigma^2_m)$, $\varepsilon^a_{i,1,t} \sim N(0,\sigma^2_a)$, and $\varepsilon^a_{i,2,t} \sim N(0,\sigma^2_a)$
3. Find $z^m_{i,t}$, $z^a_{i,1,t}$, and $z^a_{i,2,t}$, given policy functions,

$$
\begin{aligned}
g^a_1\left(x_{i,1,t-1},x_{i,2,t-1},z^a_{i,1,t-1},z^a_{i,2,t-1},z^m_{i,t-1}\right)\\
g^a_2\left(x_{i,1,t-1},x_{i,2,t-1},z^a_{i,1,t-1},z^a_{i,2,t-1},z^m_{i,t-1}\right)\\
g^m\left(x_{i,1,t-1},x_{i,2,t-1},z^a_{i,1,t-1},z^a_{i,2,t-1},z^m_{i,t-1}\right).
\end{aligned}
$$

4. Calculate standard deviations of signal noises and Kalman gains from

$$
\begin{aligned}
z^m_{i,t} &= \left(1-\mathcal{K}^m_{i,t}\right)\left(z^m_{i,t-1}+\sigma^2_m\right)\\
z^a_{i,1,t} &= \left(1-\mathcal{K}^a_{i,1,t}\right)\left(z^a_{i,1,t-1}+\sigma^2_1\right)\\
z^a_{i,2,t} &= \left(1-\mathcal{K}^a_{i,2,t}\right)\left(z^a_{i,2,t-1}+\sigma^2_1\right)
\end{aligned}
$$

and

$$\mathscr{K}_{i,t}^{m}=\frac{z_{i,t-1}^{m}+\sigma_m^2}{z_{i,t-1}^{m}+\sigma_m^2+\eta_{i,m,t}^2},\ \mathscr{K}_{i,1,t}^{a}=\frac{z_{i,1,t-1}^{a}+\sigma_1^2}{z_{i,1,t-1}^{a}+\sigma_1^2+\eta_{i,1,t}^2},\ \mathscr{K}_{i,2,t}^{a}=\frac{z_{i,2,t-1}^{a}+\sigma_2^2}{z_{i,2,t-1}^{a}+\sigma_2^2+\eta_{i,2,t}^2}.$$

Then

$$\eta_{i,m,t}^2=\frac{z_{i,t}^{m}\left(z_{i,t-1}^{m}+\sigma_m^2\right)}{z_{i,t-1}^{m}+\sigma_m^2-z_{i,t}^{m}},\ \eta_{i,1,t}^2=\frac{z_{i,1,t}^{a}\left(z_{i,1,t-1}^{a}+\sigma_1^2\right)}{z_{i,1,t-1}^{a}+\sigma_1^2-z_{i,1,t}^{a}},\ \eta_{i,2,t}^2=\frac{z_{i,2,t}^{a}\left(z_{i,2,t-1}^{a}+\sigma_1^2\right)}{z_{i,2,t-1}^{a}+\sigma_1^2-z_{i,2,t}^{a}}$$

5. Generate random numbers for signal noises $\xi_{i,m,t}\sim\mathscr{N}\left(0,\eta_{i,m,t}^2\right)$, $\xi_{i,1,t}\sim\mathscr{N}\left(0,\eta_{i,1,t}^2\right)$, $\xi_{i,2,t}\sim\mathscr{N}\left(0,\eta_{i,2,t}^2\right)$.
6. Calculate the perceived gap(markup) **after observing their signals** at t

$$\begin{aligned}
x_{i,1,t}&=x_{i,1,t-1}-\left[\mathscr{K}_{i,t}^{m}\left(s_{i,t}^{m}-m_{i,t-1|t-1}\right)+\mathscr{K}_{i,1,t}^{a}\left(s_{i,1,t}^{a}-a_{i,1,t-1|t-1}^{a}\right)\right]\\
&=x_{i,1,t-1}-\left[\mathscr{K}_{i,t}^{m}\left(m_{t-1}-m_{i,t-1|t-1}+\varepsilon_t^m+\xi_{i,m,t}\right)\right.\\
&\qquad\left.+\mathscr{K}_{i,1,t}^{a}\left(a_{i,1,t-1}^{a}-a_{i,1,t-1|t-1}^{a}+\varepsilon_{i,1,t}^{a}+\xi_{i,1,t}\right)\right]
\end{aligned}$$

$$\begin{aligned}
x_{i,2,t}&=x_{i,2,t-1}-\left[\mathscr{K}_{i,t}^{m}\left(s_{i,t}^{m}-m_{i,t-1|t-1}\right)+\mathscr{K}_{i,2,t}^{a}\left(s_{i,2,t}^{a}-a_{i,2,t-1|t-1}^{a}\right)\right]\\
&=x_{i,2,t-1}-\left[\mathscr{K}_{i,t}^{m}\left(m_{t-1}-m_{i,t-1|t-1}+\varepsilon_t^m+\xi_{i,m,t}\right)\right.\\
&\qquad\left.+\mathscr{K}_{i,2,t}^{a}\left(a_{i,2,t-1}^{a}-a_{i,2,t-1|t-1}^{a}+\varepsilon_{i,2,t}^{a}+\xi_{i,2,t}\right)\right]
\end{aligned}$$

where

$$\begin{aligned}
a_{i,1,t}-a_{i,1,t|t}&=\left(1-\mathscr{K}_{i,1,t}^{a}\right)\left(a_{i,1,t-1}-a_{i,1,t-1|t-1}\right)+\varepsilon_{i,1,t}^{a}-\mathscr{K}_{i,1,t}^{a}\left(\varepsilon_{i,1,t}^{a}+\xi_{i,1,t}\right)\\
a_{i,2,t}-a_{i,2,t|t}&=\left(1-\mathscr{K}_{i,2,t}^{a}\right)\left(a_{i,2,t-1}-a_{i,2,t-1|t-1}\right)+\varepsilon_{i,2,t}^{a}-\mathscr{K}_{i,2,t}^{a}\left(\varepsilon_{i,2,t}^{a}+\xi_{i,2,t}\right)\\
m_t-m_{i,t|t}&=\left(1-\mathscr{K}_{i,t}^{m}\right)\left(m_{t-1}-m_{i,t-1|t-1}\right)+\varepsilon_t^m-\mathscr{K}_{i,t}^{m}\left(\varepsilon_t^m+\xi_{i,m,t}\right)
\end{aligned}$$

with given $a_{i,1,-1}-a_{i,1,-1|-1}=0$, $a_{i,2,1}-a_{i,2,-1|-1}=0$, and $m_{-1}-m_{i,-1|-1}=0$.

7. Price changes: for $j\in\{1,2\}$,

$$\Delta p_{i,j,t}=\begin{cases}0 & \text{if } -\theta+\beta V\left(0,0,z_{i,1,t}^{a},z_{i,2,t}^{a},z_{i,t}^{m}\right)\\ & \qquad\leq-\left[\left(x_{i,1,t}\right)^2+\left(x_{i,2,t}\right)^2\right]+\beta V\left(x_{i,1,t},x_{i,2,t},z_{i,1,t}^{a},z_{i,2,t}^{a},z_{i,t}^{m}\right)\\ -x_{i,j,t} & \text{if } -\theta+\beta V\left(0,0,z_{i,1,t}^{a},z_{i,2,t}^{a},z_{i,t}^{m}\right)\\ & \qquad>-\left[\left(x_{i,1,t}\right)^2+\left(x_{i,2,t}\right)^2\right]+\beta V\left(x_{i,1,t},x_{i,2,t},z_{i,1,t}^{a},z_{i,2,t}^{a},z_{i,t}^{m}\right)\end{cases}$$

8. True markup: for $j\in\{1,2\}$,

$$\begin{aligned}
\mu_{i,j,t}&=p_{i,j,t}-m_t-a_{i,j,t}\\
&=\Delta p_{i,j,t}+x_{i,j,t-1}-\left(m_{t-1}-m_{i,t-1|t-1}\right)-\left(a_{i,j,t-1}-a_{i,j,t-1|t-1}\right)-\varepsilon_t^m-\varepsilon_{i,j,t}^{a}
\end{aligned}$$

where $a_{i,j,-1}-a_{i,j,-1|-1}=0$ and $m_{-1}-m_{i,-1|-1}=0$.

Appendix E. Distribution of Price Changes

Appendix Figure G.4 shows the distribution of price changes in the single-product model and in a two-product model. As a comparison, I also plot the distribution of price changes in the menu-cost-only model with single-product firms (yellow bar). All three models are calibrated to match the same frequency and size of price changes. In the baseline single-product model (blue bar), there are no small price changes because price changes occur when firms believe that their price is outside of their inaction bands. However, the kurtosis of the distribution in the baseline model is higher than that in the menu-cost-only model, and there is a relatively small fraction of firms around the inaction bands in the baseline model compared with the menu-cost-only model.[48] Notice that firms in the baseline model have different inaction bands depending on their subjective uncertainty, while firms in the menu-cost-only model have the same inaction bands (black vertical line). The heterogeneity in firms' subjective uncertainty makes the distribution of price changes in the baseline model more dispersed than that in the menu-cost-only model.

In contrast to the single-product model, the baseline two-product model generates both small and large price changes. When a two-product firm believes that one of its prices is far away from its perceived optimal price, the firm pays a fixed menu cost to change its price. As additional price changes are free after paying this menu cost, the firm also changes the price of the other product, even if it is still close to the perceived optimal price. Thus, the economy with two-product firms can have a large fraction of small price changes and a higher kurtosis of the price change distribution.[49] This motive based on economies of scope in menu cost technology weakens selection effects of price changes, which act as a strong force to reduce monetary non-neutrality in a standard menu cost model such as Golosov and Lucas (2007).

Appendix F. Additional Evidence on Models Predictions

In this appendix, I show two additional kinds of evidence that support the key predictions of the baseline model. First, I show using the New Zealand survey data that the empirical distribution of desired price changes has a fat tail. Second, I show that in the survey data, firms with greater uncertainty are more likely to delay their price changes.

[48] The red vertical lines are the average inaction bands across all firms in the baseline single-product model. Because both models have the same frequency and size of price changes, the average inaction bands in both models are similar.

[49] Standard menu cost models with two-product firms can also generate small price changes through the same economies of scope motive (e.g. Midrigan, 2011; Bhattarai and Schoenle, 2014; Alvarez and Lippi, 2014). However, the baseline two-product model has a more dispersed distribution of price changes than the menu cost only models with two-product firms because, again, firms' optimal inaction bands are a function of their subjective uncertainty. Appendix Figure G.12 shows a comparison of the distribution of price changes in the baseline two-product model with that in the menu cost only model with two-product firms.

Appendix F.1. Evidence on the Leptokurtic Distribution of Desired Price Changes.

One key result of this paper is to show that selection in information processing endogenously leads to a fat-tail distribution of desired price changes. I find that this result is empirically consistent with what we observe in the survey data. In the second wave of the survey data, firms' managers were asked how much they would like to change the price of their main product if it was free to change its price in three months. The answer gives firms' desired price changes in three months. To construct a model-consistent measure of desired price changes, I define an inflation-adjusted desired price changes as the gap between the desired price changes and their inflation expectations in three months. The left panel of Appendix Figure G.13 shows that the distribution of desired price changes has a cluster near zero, while some desired price changes are very far away from zero. The distribution of desired price changes exhibits kurtosis around 5, implying that the survey supports the fat-tail distribution of desired price changes. The baseline model with both rational inattention and menu costs endogenously captures this distributional characteristic without an assumption that the distribution of good-specific shocks is leptokurtic.[50]

Appendix F.2. Subjective Uncertainty and the Duration of Price Changes.

Another characteristic of firms' pricing rule is that firms' optimal inaction bands depend on their subjective uncertainty about the underlying shocks. When firms are more uncertain, the inaction bands are wider, implying that the wait-and-see effects are present in firms' optimal price-setting decision. I directly test this implication using the New Zealand survey data. Firms asked to assign probabilities (from 0 to 100) to the different outcomes for growth rates of unit sales of their main product over the next 12 months. I calculate the standard deviation—which is a measure of firms' subjective uncertainty—surrounding firms' sales forecast using the implied probability distribution. The right panel of Appendix Figure G.13 shows that the firms are shorter duration of next price changes when they are less uncertain about their future sales. I regress the duration of firms' expected next price changes on their subjective uncertainty about future sales growth. Appendix Table G.5 shows that firms that have greater uncertainty expect a longer duration before their next price changes. This finding is consistent with the prediction of the baseline model. When firms are more uncertain about their fundamentals, they are reluctant to change their prices. Instead, firms want to wait until they acquire more information to resolve their own uncertainty about their fundamentals.

[50] The previous literature on menu costs often assumes that this fat-tail distribution of idiosyncratic shocks weakens selection effects of price changes. Midrigan (2011) supports this assumption by providing evidence of excess kurtosis in the distribution of markup gaps in U.S. retail data.

Appendix G. Appendix Tables and Figures

Appendix Table G.1: Summary Statistics for Number of Products

Industries	Obs.	Mean	Median	Std. Dev.	Max.
Total	712	67.4	9	234.2	2115
Total without Retail/Wholesale Trade	627	9.55	7	8.47	48
– Manufacturing	278	9.57	8	7.75	39
– Professional and Financial Services	276	7.95	7	6.09	35
– Other Services	37	14.49	13	11.63	48
– Construction and Transportation	36	8.42	5	8.92	40

Notes: This table reports summary statistics for firms' number of products by sectors. The number of products of each firm is measured from answers to the following question in the second wave of New Zealand Firms' Expectation Survey: "*In addition to your main product or product line, how many other products do you sell?*" See Coibion et al. (2018) for details about the survey data. Moments are calculated using sampling weights.

Appendix Table G.2: Number of Products and Knowledge about Aggregate Inflation (All Industries)

	(1)	(2)	(3)	(4)
Panel A. Dependent variable: Inflation backcast errors				
log(number of products)	-0.405***	-0.102***	-0.205***	-0.057*
	(0.069)	(0.030)	(0.063)	(0.033)
Observations	673	658	506	495
R-squared	0.229	0.835	0.273	0.896
Panel B. Dependent variable: Willingness to pay for professional inflation forecasts				
log(number of products)	-1.968	2.359*	-2.945*	3.871***
	(1.782)	(1.268)	(1.532)	(1.162)
Observations	438	434	378	372
R-squared	0.102	0.626	0.171	0.662
Firm-level controls	Yes	Yes	Yes	Yes
Industry fixed effects		Yes		Yes
Manager controls			Yes	Yes

Notes: This table reports results for the Huber robust regression. Dependent variables are the absolute value of firm errors about past 12 month inflation from Wave #1 survey (Panel A) and firms' willingness to payment for professional forecaster's forecasts about future inflation from Wave #4 (Panel B). Firm-level controls include log of firms' age, log of firms' employment, foreign trade share, number of competitors, firms' beliefs about price difference from competitors, and the slope of the profit function. Industry fixed effects include dummies for 17 sub-industries. Manager controls include the age of the respondent (each firm's manger), education, income, and tenure at the firm. Sample weights are applied to all specifications. Robust standard errors (clustered at the 3-digit ANZ SIC level) are reported in parentheses. ***, **, * denotes statistical significance at 1%, 5%, and 10% levels respectively.

Appendix Table G.3: Summary Statistics for Inflation Backcast Errors by Industries

Industries	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	N	Mean (S.D.)	N	Mean (S.D.)	N	Mean (S.D.)	N	Mean (S.D.)
Total	1-5	5.01 (4.56)	6-9	5.83 (4.93)	10-16	3.89 (5.20)	>16	2.48 (2.99)
Total without Retail/Wholesale	1-4	5.23 (4.10)	5-8	6.30 (5.03)	9-15	4.40 (5.57)	>15	3.54 (3.62)
− Manufacturing	1-5	1.52 (2.19)	6-9	1.70 (1.86)	10-15	2.25 (2.74)	>15	2.46 (2.56)
− Professional and Financial Services	1-4	6.42 (3.17)	5-8	6.16 (4.65)	9-13	7.00 (3.57)	>13	5.51 (3.71)
− Other Services	1-6	2.29 (1.97)	7-15	0.72 (0.46)	16-22	0.76 (0.52)	>22	0.90 (0.51)
− Construction and Transportation	1-3	7.46 (5.06)	4-5	7.38 (4.55)	6-9	10.82 (5.34)	>9	7.36 (7.59)

Notes: This table reports summary statistics for firms' (absolute) backcast errors about aggregate inflation by quartiles of the distribution of the number of products in each industry. The backcast errors are the absolute value of firm errors about past 12 month inflation from Wave #1 survey. Moments are calculated using sampling weights.

Appendix Table G.4: Number of Products and Knowledge about Nominal GDP Growth

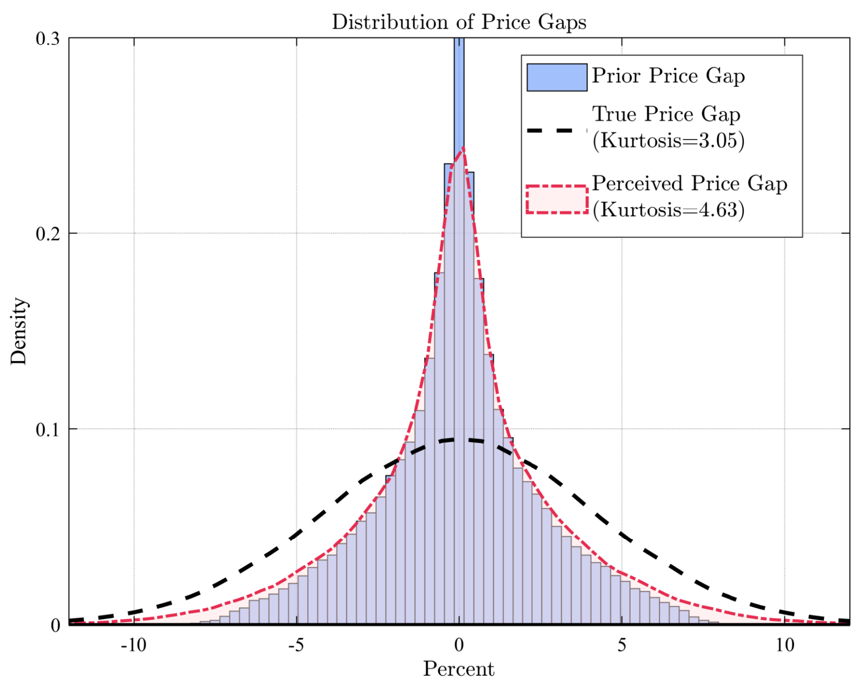
	(1)	(2)	(3)	(4)
Dependent variable: Backcast errors about nominal GDP growth rate				
Number of products	-0.041*** (0.012)	-0.020*** (0.007)	-0.035*** (0.011)	-0.017* (0.009)
Observations	390	378	334	326
R-squared	0.375	0.615	0.412	0.610
Firm-level controls	Yes	Yes	Yes	Yes
Industry fixed effects		Yes		Yes
Manager controls			Yes	Yes

Notes: This table reports results for the Huber robust regression. Dependent variable is the absolute value of firm errors about the growth rate of nominal GDP from Wave #4 survey. Firms' perceived growth rate of nominal GDP are calculated by taking the summation of firms' belief about current inflation and the real GDP growth rate in New Zealand. Firm-level controls include log of firms' age, log of firms' employment, foreign trade share, number of competitors, firms' beliefs about price difference from competitors, and the slope of the profit function. Industry fixed effects include dummies for 14 sub-industries excluding retail and wholesale trade sectors. Manager controls include the age of the respondent (each firm's manger), education, income, and tenure at the firm. Sample weights are applied to all specifications. Robust standard errors (clustered at the 3-digit ANZ SIC level) are reported in parentheses. ***, **, * denotes statistical significance at 1%, 5%, and 10% levels respectively.

Appendix Table G.5: Subjective Uncertainty and Expected Duration of Price Changes

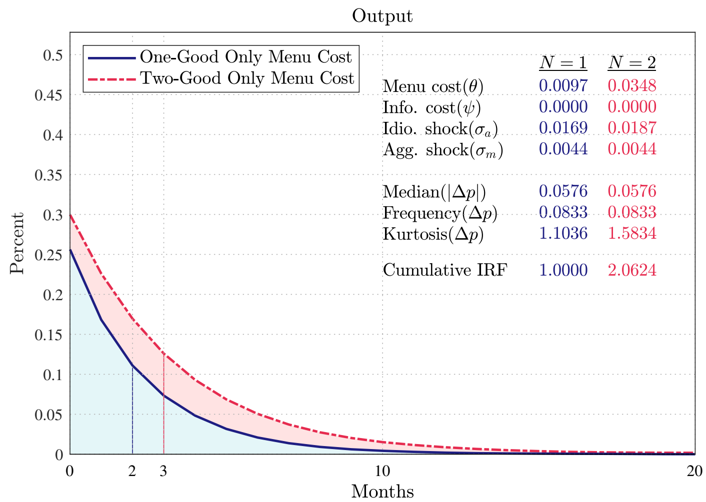
	(1)	(2)	(3)	(4)
Dependent variable: Duration of expected next price changes				
Standard deviation of the growth rate of sales over the next 12 months	0.132***	0.104**	0.153***	0.160**
	(0.039)	(0.042)	(0.054)	(0.065)
Observations	583	591	443	442
R-squared	0.323	0.697	0.342	0.436
Firm-level controls	Yes	Yes	Yes	Yes
Industry fixed effects		Yes		Yes
Manager controls			Yes	Yes

Notes: This table reports results for the Huber robust regression. Dependent variable is the duration of expected next price changes from Wave #2. The regressor is the standard deviation implied by the reported probability distribution for the growth rates of unit sales of firms' main product over the next 12 months. Firm-level controls include log of firms' age, log of firms' employment, the number of competitors, and log of firms' number of products. Industry fixed effects include dummies for 14 sub-industries excluding retail and wholesale trade sectors. Manager controls include the age of the respondent (each firm's manger), education, income, and tenure at the firm. Sample weights are applied to all specifications. Robust standard errors (clustered at the 3-digit ANZ SIC level) are reported in parentheses. ***, **, * denotes statistical significance at 1%, 5%, and 10% levels respectively.



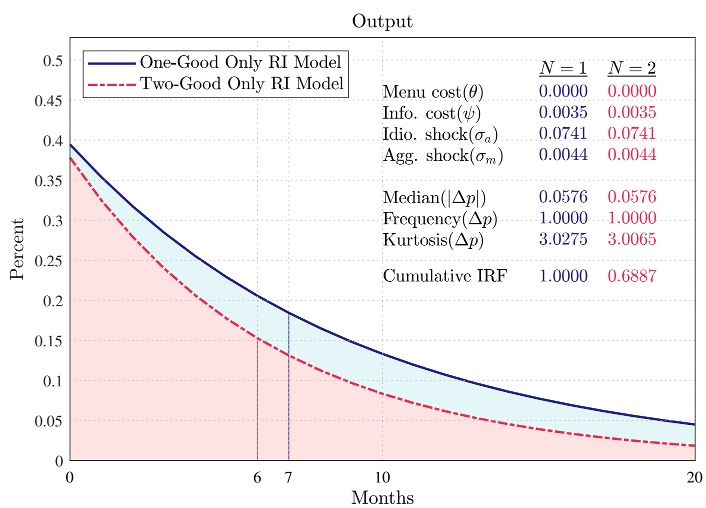
Appendix Figure G.1: Distribution of True and Perceived Price Gaps in the Two-Product Model

Notes: This figure plots distributions of price gaps in the two-good version of the baseline model. Blue bar graph shows the distribution of firms' *prior* about their price gaps at the beginning of period. Black dashed line shows the distribution of firms' true price gaps after their Gaussian shocks realized. Firms choose their optimal signals about the shocks and form a new posterior about their (frictionless) optimal price. Red dash-dot line shows the distribution of the *posterior* of perceived price gaps.



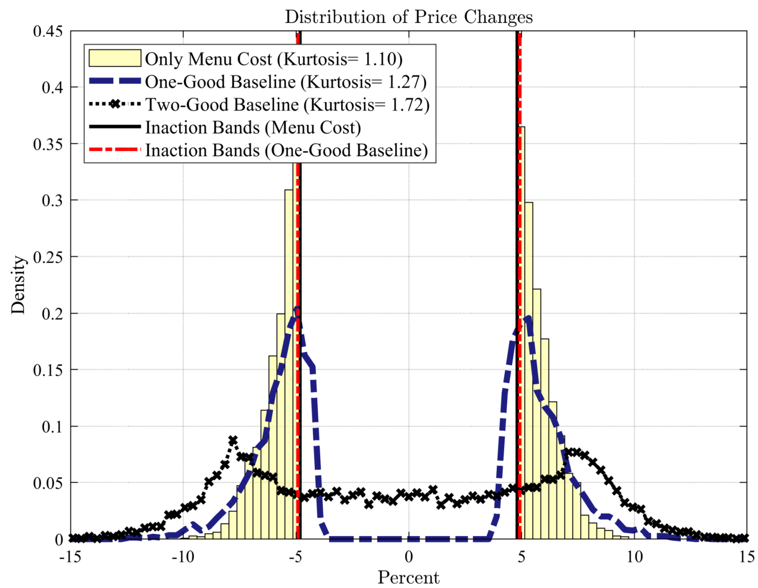
Appendix Figure G.2: IRFs of Output in the Menu Costs Models with Perfect Information

Notes: This figure plots impulse responses of output to a one standard deviation monetary shock in the one-good and two-good versions of menu cost models with perfect information. Cumulative IRFs refers to area under the responses of output. I normalize the cumulative output response in the one-good version of the model as one.



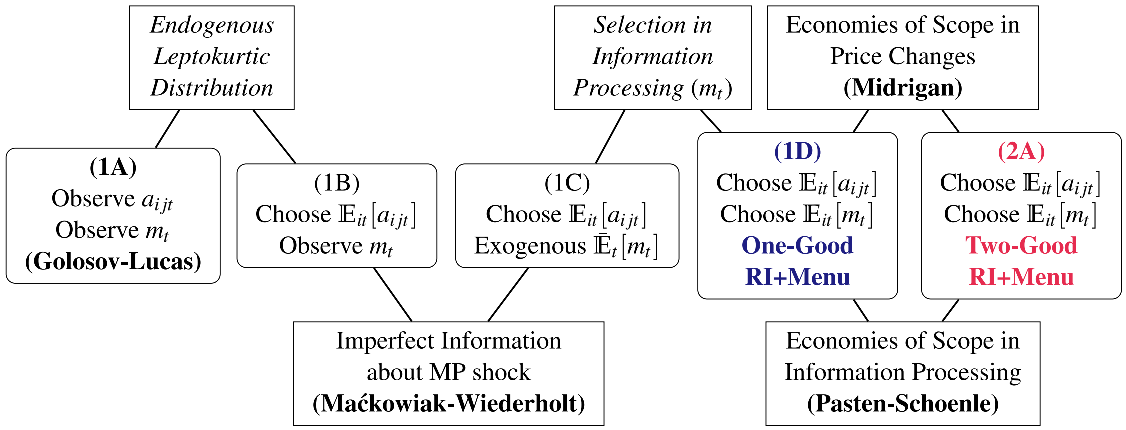
Appendix Figure G.3: IRFs of Output in the Rational Inattention Models without Menu Costs

Notes: This figure plots impulse responses of output to a one standard deviation monetary shock in the one-good and two-good versions of rational inattention models without menu costs. Cumulative IRFs refers to area under the responses of output. I normalize the cumulative output response in the one-good version of the model as one.



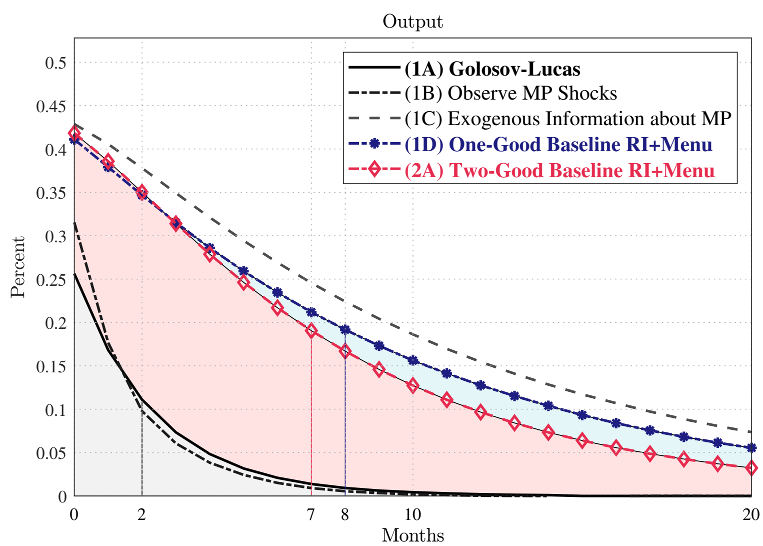
Appendix Figure G.4: Distributions of Price Changes

Notes: This figure plots the distribution of price changes in the single-product menu cost model with perfect information (yellow bar), that in the baseline single-product model (blue dashed line), and that in the baseline two-product model (black line with cross markers). Black vertical lines are the inaction bands for firms in the menu cost model with perfect information. In this model, every firms have the same inaction bands. Red vertical dash-dot lines are the average of inaction bands across firms in the baseline single-product model.



Appendix Figure G.5: Counterfactuals and Model Mechanism

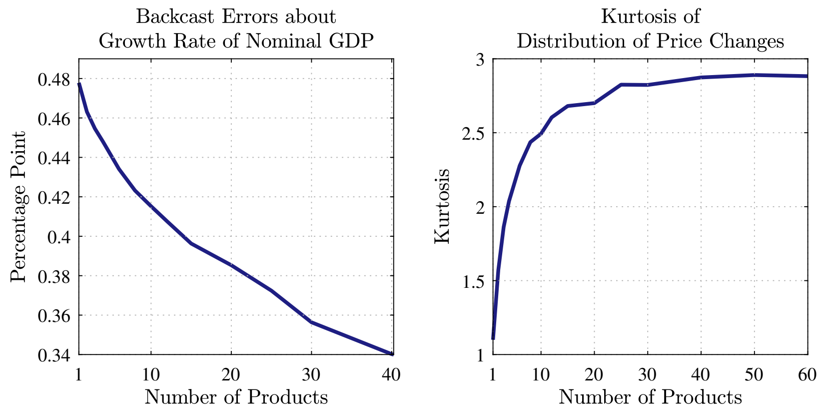
Notes: This figure shows counterfactual models and the implied model mechanisms. Model (1A), (1B), (1C), and (1D) are single-product menu cost models with different assumptions about firms' information set. In model (1A), firms have full-information about both idiosyncratic and monetary shocks. Firms in model (1B) have perfect information about the monetary shock, but choose their optimal signal about the idiosyncratic shock. All firms in model (1C) are given the same exogenous signal about the monetary shock, while they choose their optimal signal about the idiosyncratic shock. Model (1D) is the baseline single-product model where all firms choose their optimal signals about both shocks. Model (2A) is the baseline two-product model. See Section 4.6 for details.



	(1A)	(1B)	(1C)	(1D)	(2A)
Impact Response	**1**	1.23	1.67	**1.60**	**1.63**
Cumulative IRFs	**1**	1.11	7.95	**6.78**	**5.92**

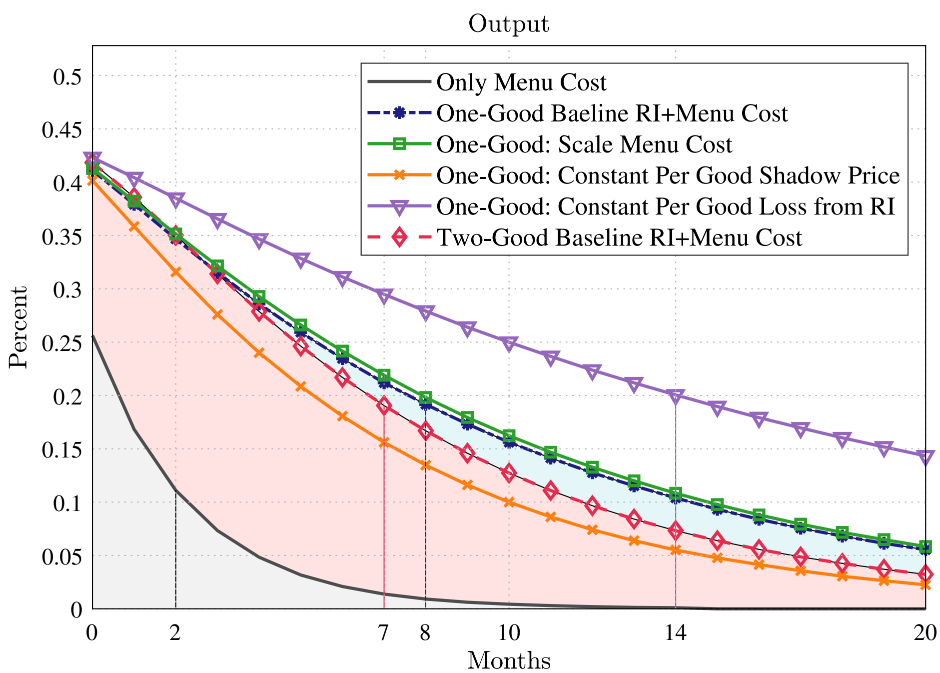
Appendix Figure G.6: Impulse Responses of Output in Counterfactual Models

Notes: This figure plots output responses to a one standard deviation monetary shock in counterfactual models described in Appendix Figure G.5. Cumulative IRFs refers to area under the output responses. I normalize both impact and cumulative output response in the menu-cost-only model with single-product firms as 1. See Section 4.6 for details.



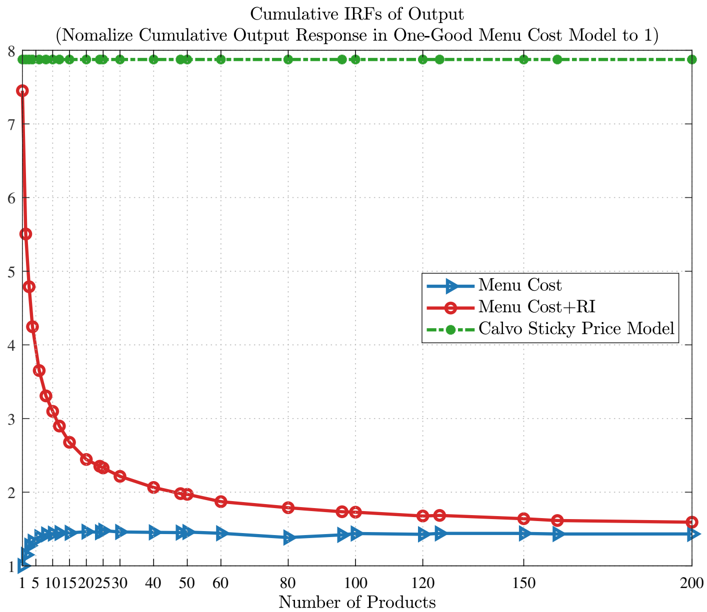
Appendix Figure G.7: Backcast Errors and Kurtosis of Price Changes in the Simplified Models

Notes: This figure shows shows the backcast errors of firms about the growth rate of nominal GDP (left panel) and the kurtosis of distribution of price changes (right panel) in the simplified version of the baseline models with different numbers of products sold by firms. See Section 4.7 for details.



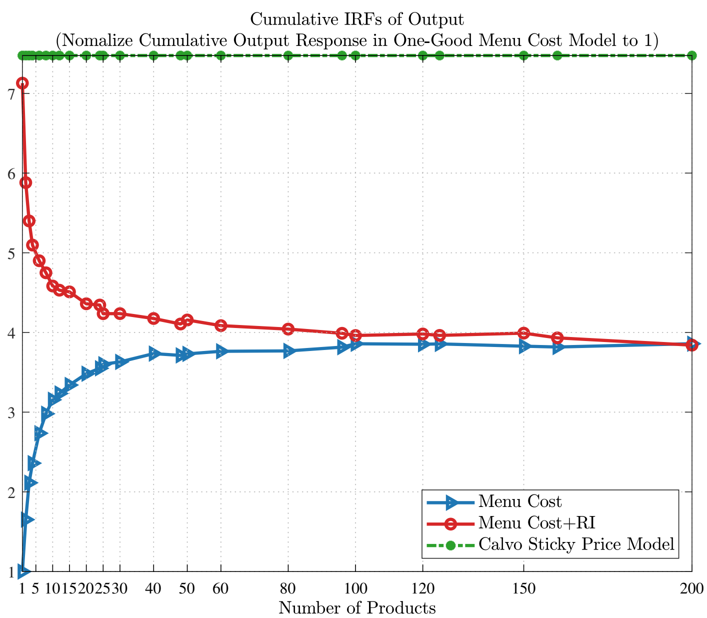
Appendix Figure G.8: Impulse Response of Output to a Monetary Shock under Alternative Assumptions

Notes: This figure plots impulse responses (IRF) of output to a one standard deviation monetary shock. Black solid line is the IRF in the menu-cost-only model. Blue line with circles is the IRF in the baseline single-product model with both menu cost and rational inattention. Green line with cubes is the IRF in the single-product model with both menu cost and rational inattention where the size of menu cost is scaled down linearly. Purple line with triangles is the IRF in the single-product model with both menu cost and rational inattention where the loss per good from rational inattention is the same as that in the two-product model. Orange line is the IRF in the single-product model with both menu cost and rational inattention where the shadow price of information processing is scaled down linearly. Lastly, red line with diamonds is the IRF in the baseline two-product model. See Appendix A.1 and Appendix A.3 for details.



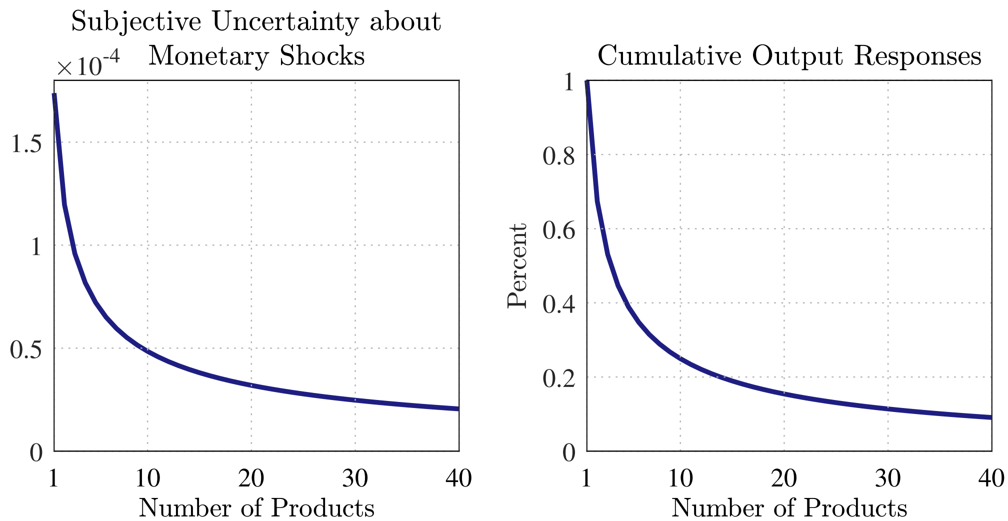
Appendix Figure G.9: Cumulative Output Responses and Number of Products in the Simplified Models When Menu Costs Scale Linearly

Notes: This figure plots cumulative output responses in the simplified models when menu costs scale linearly with the number of products. "Menu Cost+RI" refers to the model with both menu costs and rational inattention. Red line shows the cumulative output responses in the only menu cost models with different number of products and blue line shows those in the models with both menu costs and rational inattention with different number of products. See Appendix A.1 for details.



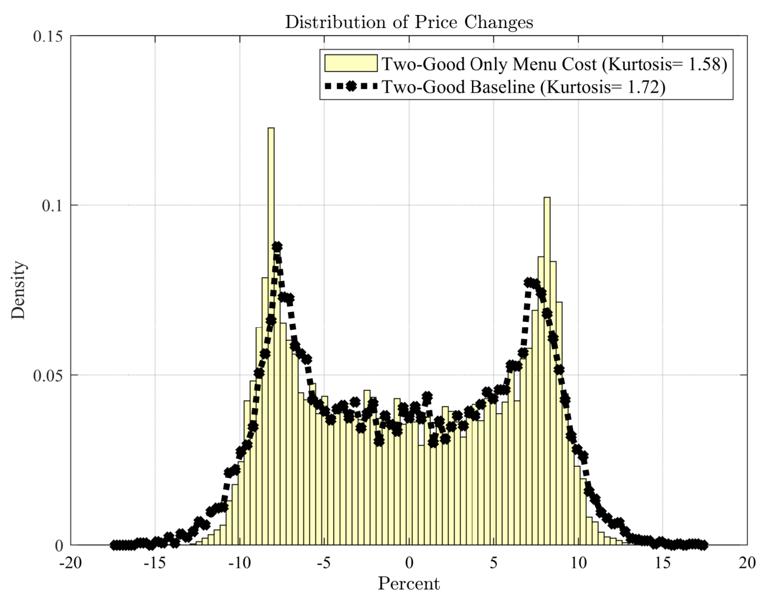
Appendix Figure G.10: Cumulative Output Responses and Number of Products in the Simplified Models with Variable Menu Costs

Notes: This figure plots cumulative output responses in the simplified models with different number of products when I include variable menu costs. "Menu Cost+RI" refers to the model with both menu costs and rational inattention. Red line shows the cumulative output responses in the only menu cost models with different number of products and blue line shows those in the models with both menu costs and rational inattention with different number of products. See Appendix A.1 for details.



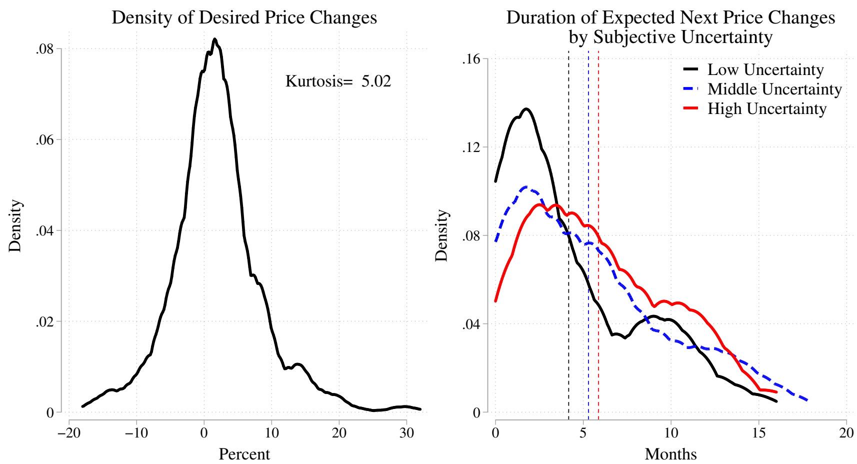
Appendix Figure G.11: Subjective Uncertainty and Cumulative Output Responses in Rational Inattention Only Models with Different Numbers of Products

Notes: The left panel plots firms' subjective uncertainty about monetary shocks in the rational inattention only model with different numbers of products. The right panel plots cumulative output responses to a one standard deviation monetary shock in the rational inattention only models with different numbers of products sold by firms. The cumulative output response in the single-product model is normalized to one.



Appendix Figure G.12: Distribution of Price Changes in the Two-Product Models

Notes: This figure plots the distribution of price changes in the two-product menu cost model with perfect information (yellow bar) and that in the two-good version of the baseline model (black line with circle markers).



Appendix Figure G.13: Distribution of Desired Price Changes and Duration of Expected Price Changes

Notes: This left panel shows the distribution of desired price changes in the second wave of the survey data. Firms' managers in the survey were asked how much they would like to change the price of their main product if it was free to change its price in three months. The answer gives firms' desired price changes in three months. I define the inflation-adjusted desired price changes as the gap between firms' desired price changes in three months and their three-month ahead inflation expectations. The right panel shows the distribution of firms' expected duration of their next price changes by their degree of subjective uncertainty. The subjective uncertainty is measured by the standard deviation implied by the reported probability distribution for the growth rates of unit sales of firms' main product over the next 12 months. See Appendix F for details.