Rational Inattention, Multi-Product Firms and the Neutrality of Money

Raphael Schoenle, Economics Department, Brandeis University

Ernesto Pasen, Banco Central de Chile, Toulouse School of Economics

Working Paper Series

2015 | 91

Rational Inattention, Multi-Product Firms and the Neutrality of Money[☆,☆☆]

Ernesto Pasten[a], Raphael Schoenle[b]

[a]*Banco Central de Chile, Toulouse School of Economics*

[b]*Brandeis University*

Abstract

In a quantitative rational inattention model, monetary non-neutrality quickly vanishes as firms price more goods while monetary non-neutrality is strong in a single-product setting under otherwise identical conditions. This result is due to (1) economies of scope that arise naturally in the multi-product setting, where processing information is costly but using already internalized information is free, and (2) good-specific shocks that account for a non-zero fraction of the within-firm dispersion of log price changes, which we document in U.S. data. As a consequence, as firms price more goods, they shift attention from good-specific to common shocks, such as monetary shocks. Aggregate prices then respond much faster to monetary shocks due to strategic complementarity.

Keywords: rational inattention, multi-product firms, monetary non-neutrality
JEL classification: E3, E5, D8

[☆]We thank comments by Daniel Bergstresser, Markus Brunnermeier, Paco Buera, Larry Christiano, José de Gregorio, Eduardo Engel, Christian Hellwig, Hugo Hopenhayn, Pat Kehoe, Oleksiy Kryvtsov, Ben Malin, Virgiliu Midrigan, Juanpa Nicolini, Kristoffer Niemark, Guillermo Ordonez, Felipe Schwartzman, Jean Tirole, Mirko Wiederholt and seminar participants at the Central Bank of Chile, Central European University, CREI, Ente Einaudi, ESSET 2013, the XIV IEF Workshop (UTDT, Buenos Aires), Minneapolis FED, Northwestern, Paris School of Economics, Philadelphia Fed, Princeton, PUC-Chile, Recent Developments in Macroeconomics at ZEW, Richmond Fed, Second Conference on Rational Inattention and Related Theories (Oxford), the 2012 SED Meeting (Cyprus), Toulouse, and UChile-Econ. Pasten thanks the support of the Université de Toulouse 1 Capitole and Christian Hellwig's ERC grant during his stays in Toulouse.

[☆☆]This research was conducted with restricted access to the Bureau of Labor Statistics data. We thank coordinator Ryan Ogden for his help and Miao Ouyang for excellent research assistance. The views expressed herein are those of the authors and do not necessarily represent the position of the Central Bank of Chile or the Bureau of Labor Statistics. All errors or omissions are our own.

Email addresses: `epasten@bcentral.cl` (Ernesto Pasten), `schoenle@brandeis.edu` (Raphael Schoenle)

1. Introduction

Rational Inattention Theory (Sims (1998, 2003)) is an increasingly popular formalization of the idea that limited ability to process information (or "attention") may be behind the simplicity of human actions relative to those of agents in economic models. A prime example – as pointed out in Sims' seminal work – is that prices only respond slowly to monetary shocks because firms allocate most of their attention to highly volatile idiosyncratic shocks. Little attention in turn to less volatile, monetary shocks means high observational noise and a slow response to monetary shocks. This result is confirmed quantitatively by Mackowiak and Wiederholt (2009) who calibrate a rational inattention model of price setting to US data to find large and long-lasting monetary non-neutrality even when the friction is "small."

We revisit this result of rational inattention after we relax the usual assumption in macroeconomics that firms price a single good. In doing so, we also assume that shocks can be both good-specific and firm-specific, in addition to monetary.[1] Then, under these assumptions, our main result emerges: a calibrated model of rationally inattentive, monopolistically competitive firms predicts much milder monetary non-neutrality when firms price multiple goods rather than a single good. This result is particularly strong when firms are interpreted as retailers since empirically, retailers price a large number of goods; but multi-product pricing has a strong effect even for producers who price a much smaller number of goods.

Three factors drive our main result: First, multi-product firms have stronger incentives to pay attention to monetary and firm-specific shocks. The reason lies in economies of scope in information processing: The attention to reduce observation noise is the same for all kinds of shocks, but information about monetary and firm-specific shocks can be used to price all goods. By contrast, the benefit of paying attention to good-specific shocks does not scale up with the number of goods. We call this force "economies of scope in information processing."

Second, a force going in opposite direction is that firms must allocate their limited at-

[1] Adding regional or sectoral shocks would make no difference in our analysis.

tention to more shocks as they price more goods, spreading "thin" their attention. We find that if total attention is held constant, monetary non-neutrality may increase if the number of goods is small but always decreases as this number goes to infinity (so the economies of scope dominate). However, expected profit losses per good due to the friction also increase with more goods. In other words, stronger monetary non-neutrality can only happen as the friction becomes more binding. This is important to keep track of: Once we compare economies for which the friction is equally binding, attention to monetary shocks and monetary neutrality unambiguously increase as firms price more goods. If we were to allow losses to increase in the number of goods, our model would not be internally consistent: Firms would like to split up their pricing decisions into single-good units to minimize total losses.[2]

Third, strategic complementarities amplify these effects. Starting from a situation in which firms pay little attention to monetary shocks, more attention to these shocks has a large effect on reducing monetary non-neutrality. The reason is that under stronger complementarities among competing firms, aggregate prices respond faster to monetary shocks if competitor prices respond faster to these shocks. A corollary of the same effect is that firms pricing a single good respond fast to monetary shocks when they coexist with multi-product firms that respond fast to these shocks.

We then provide empirical support for our key assumptions. First, there is strong evidence that firms indeed price multiple goods. Just to fix ideas, retailers price on average about $40,000$ goods (FMI, 2010) and producers about 4 goods (Bhattarai and Schoenle (2014)). There is also suggestive evidence that firms price their goods in centralized units.[3] To support our assumption of firm- and good-specific shocks, we document a new empirical fact: Within-firm dispersion of log price changes accounts for 51.6% and 59.1% of total cross-sectional

[2]This does not mean that firms would also decentralize their production or commercialization processes.

[3]The Bureau of Labor Statistic's (BLS) defines a firm as a "price-forming unit" in the PPI micro data. In this dataset, only 1.5% of firms price a single good. Further, Zbaracki et al. (2004) present a case study of the pricing process of a firm. They report that all regular prices are decided at headquarters while all sale prices are decided by local managers. At both levels there is a *single* price setting unit for all goods.

dispersion in U.S. Consumer Price Index (CPI) and Producer Price Index (PPI) micro data. Although there are many plausible explanations for this fact, our quantitative results hold as long as good-specific shocks explain a non-zero fraction of this dispersion.

Next, we confirm our theoretical results by calibrating our model. As a benchmark for our calibration, we use the setup of Mackowiak and Wiederholt (2009), which features firms pricing a single good and is calibrated to micro moments from the CPI. Our main twist is to allow for the number of goods to vary and to calibrate our firm- and good-specific shocks to account for the ratio of within-firm to total dispersion of price changes in the data. When firms price two goods, our model yields only one third of the monetary non-neutrality of the benchmark, holding expected per-good losses constant. When firms price eight goods or more, money is almost neutral. Thus, our main result emerges: In a quantitative rational inattention model, monetary non-neutrality quickly vanishes as firms price more goods under the same conditions that lead to strong monetary non-neutrality in a single-good setting. Remarkably, this quantitative result holds although firms' attention to monetary shocks always remains a small portion of their total attention.

Our main result also holds when we calibrate a more realistic heterogeneous-firm model where firms in the economy differ in the number of goods. We use PPI data to calibrate this model since this dataset allows us to compute micro moments after sorting firms into four bins that depend on the number of goods they price.[4] Again, our model yields approximately a third of the monetary non-neutrality of our benchmark, holding expected per-good losses constant. As before, firms spend little attention on monetary shocks, but now we additionally find that prices of all firms (including single-product firms) exhibit very similar impulse responses, another effect of strategic complementarity among firms. We also flip our exercise around to show a general tradeoff between monetary neutrality and the friction: To yield the same monetary non-neutrality as in our benchmark, the cost of the friction has to go

[4] We discuss in the paper the patterns of these moments across bins which for brevity we omit here.

up. In our quantitative exercises, the cost of the friction must exceed the range typically found/assumed in the literature to yield the same monetary non-neutrality as our single-product benchmark.

Our calibration exercises suggest two conclusions: First, since retailers typically price a large number of goods, multi-product pricing can be very important quantitatively for a rational inattention model where firms are interpreted as retailers. Second, we conclude that multi-product pricing is also quite important when firms in the model are interpreted as producers although monetary non-neutrality can still be sizable and our estimate of four goods priced by producers is a lower bound. When we examine a number of extensions and robustness checks, our results continue to hold strongly.

Finally, two side-points worth noting emerge form our calibration exercises. The first point is that economies of scope in information processing do not only matter for multi-product firms. They also matter when shocks have different persistence because processed information depreciates faster for less persistent shocks. When we calibrate idiosyncratic shocks to be less persistent than monetary shocks to match the first-order serial correlation of log price changes in CPI data, monetary non-neutrality is smaller than in our benchmark even when we assume single-product firms. However, the effect is quantitatively less important than multi-product price setting. Our second point is that attention cannot be pinned down from the data because model-predicted moments of prices are very insensitive to variations in firms' attention, while monetary non-neutrality is very sensitive to such variations. This is why we rely on the literature as benchmark to calibrate the size of the friction.[5]

Literature review. We see the economies of scope we highlight in this paper as a general feature of Rational Inattention Theory. Thus our paper is related to all its applications such as monetary economics (Sims (2006), Woodford (2009, 2012), Mackowiak and Wiederholt

[5]While the cost of the friction in Mackowiak and Wiederholt (2015) is smaller in absolute terms than in our benchmark, we conjecture that in relative terms multi-product pricing is also important is their setting.

(2009, 2011), Paciello and Wiederholt (2014) and Matejka (2010), portfolio choice (Mondria (2010)), asset pricing (Peng and Xiong (2006)), rare disasters (Mackowiak and Wiederholt (2011)), consumption dynamics (Luo (2008)), home bias (Mondria and Wu (2010)), the current account (Luo et al. (2012)), discrete choice models (Matejka and McKay (2015)) and search (Cheremukhin et al. (2012)).

Our quantitative work is also complementary to the study of multi-product firms and menu costs, as in Sheshinski and Weiss (1992), Midrigan (2011), Bhattarai and Schoenle (2014) and Alvarez and Lippi (2014). A key result in this literature is that the presence of multi-product firms may increase monetary non-neutrality. We find the opposite because in rational inattention models there is no extensive margin like in menu cost models.

Our empirical work contributes to the literature by providing key moments to calibrate a multi-product rational inattention model of pricing. By contrast, previous empirical work views the data through the lens of menu cost models – for example, Bils and Klenow (2004), Klenow and Kryvtsov (2008) and Nakamura and Steinsson (2008). Finally, Venkateswaran and Hellwig (2009) question the assumption in Mackowiak and Wiederholt (2009) of independent sources of information for each type of shock. We keep this assumption since it yields predictions consistent with the data.

2. Model

Here, we outline our model and present the key elements of its solution when shocks are white noise. We present the fully-fledged model in the online appendix.

Our model is a variation of the economy in Mackowiak and Wiederholt (2009) augmented to allow for multi-product firms, and idiosyncratic shocks broken into firm- and good-specific components. In our economy, each firm $i \in [0, 1/N]$ is the monopolist price setter of N goods whose identity is randomly drawn from the pool of goods $j \in [0, 1]$ and contained in the set $\aleph_i$. Firms are subject to an information processing constraint $\kappa(N)$ on imprecisely observing signals $\left\{s_{it}^{a}, s_{it}^{f}, \{s_{nt}^{z}\}_{n\in\aleph_i}\right\}$ about nominal aggregate demand shocks $Q_t = P_tY_t$, firm-specific shocks $F_{i,t}$ and good-specific shocks $Z_{i,t}$. To get analytical results we assume that all shocks are Gaussian i.i.d.

Firms maximize the expected discounted stream of profits from their N goods by choosing how precisely to observe the respective signals. We show in the appendix that the firm's problem can equivalently be cast up to a second-order approximation as minimizing profit losses from imprecisely observing these signals, by choosing how much of total capacity $\kappa(N)$ to allocate as κ_a, κ_f and $\{\kappa_n\}_{n\in N_i}$ to the observation of each shock. That is, firms solve:

$$\min_{\kappa_a,\kappa_f,\{\kappa_n\}_{n\in\aleph_i}} \frac{\beta}{1-\beta}\frac{|\widehat{\pi}_{11}|}{2}\left[2^{-2\kappa_a}\sigma_{\Delta}^2 N + \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_f}\sigma_f^2 N + \left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2 \sum_{n\in\aleph_i} 2^{-2\kappa_n}\sigma_z^2\right] \quad (1)$$

$$\text{s.t. } \kappa_a + \kappa_f + \sum_{n\in\aleph_i}\kappa_n \leq \kappa(N) \quad (2)$$

where σ_f^2 and σ_z^2 denote the volatility of firm- and good-specific shocks, and σ_Δ^2 the volatility of the compound aggregate variable

$$\Delta_t \equiv p_t + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} y_t \quad (3)$$

that linearly depends on monetary shocks q_t after we guess that $p_t = \alpha q_t$ for $p_t = \int_0^1 p_{jt}dj$.[6] We confirm this guess below. Parameters $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$, $\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}$ and $\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}$ denote the sensitivity of frictionless prices to the log-deviations of real aggregate demand, firm- and good-specific shocks, and $\widehat{\pi}_{11}$ the derivative of profits twice with respect to the good price.

The first-order conditions of this problem are

$$\kappa_a^* = \kappa_f^* + \log_2(x_1) \quad (4)$$

$$\kappa_a^* = \kappa_n^* + \log_2\left(x_2\sqrt{N}\right), \quad \forall n \in \aleph_i \quad (5)$$

[6]Small case notation generically denotes log-deviations from steady-state levels throughout.

for $x_1 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta}{\widehat{\pi}_{14}\sigma_f}$ and $x_2 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta}{\widehat{\pi}_{15}\sigma_z}$. Since we assume that all parameters are the same for all firms and goods, it follows that all firms pay the same attention to monetary and firm-specific shocks, κ_a^* and κ_f^*, and the same attention to all relevant good-specific shocks, $\kappa_n^* = \kappa_z^*$ for all $n \in \aleph_i$ and all i.

In addition, (4) and (5) together with the information capacity constraint imply that

$$\kappa_a^* = \frac{1}{N+2}\left[\kappa(N) + \log_2(x_1) + N\log_2\left(x_2\sqrt{N}\right)\right] \tag{6}$$

if $x_1 x_2^N \in \left[\frac{2^{-\kappa(N)}}{\sqrt{N}}, \frac{2^{(N+1)\kappa(N)}}{\sqrt{N}}\right]$, which ensures that $\kappa_a^* \in [0, \kappa(N)]$.

In words, given N and total capacity $\kappa(N)$, firms pay little attention to monetary shocks when x_1 and/or x_2 are small. A small x_1 results when the ratio of firm to aggregate volatility, $\frac{\sigma_f}{\sigma_\Delta}$, is large, and/or when frictionless prices are very responsive to firm shocks, that is, when $\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}$ is large. Similarly, a small x_2 results when the ratio of good-specific to aggregate volatility, $\frac{\sigma_z}{\sigma_\Delta}$, is large and/or when $\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}$ is large.

After aggregating all prices, the guess $p_t^* = \alpha q_t$ holds for

$$\alpha = \frac{\left(2^{2\kappa_a^*} - 1\right)\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}}{1 + \left(2^{2\kappa_a^*} - 1\right)\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}}. \tag{7}$$

This is the key result of monetary rational inattention models: If firms have unlimited information-processing capacity, $\kappa(N) \to \infty$, they choose infinitely precise signals about monetary shocks, so $\kappa_a^* \to \infty$ and $\alpha \to 1$. Money is fully neutral. In contrast, if $\kappa(N)$ is finite, κ_a^* is finite and thus $\alpha < 1$. Money becomes non-neutral. Monetary non-neutrality is decreasing in κ_a^*. Moreover, for a given $\kappa_a^* < \infty$, monetary non-neutrality is decreasing in $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} > 0$ – the inverse of strategic complementarity in pricing decisions among firms.

Importantly, note the fixed point in the solution for α and κ_a^*: In equation (6), κ_a^* depends on α through σ_Δ, the volatility of the aggregate compound variable defined in (3), which is implicit in x_1 and x_2. In equation (7), α depends on κ_a. This feedback plays a central role in some of our theoretical and quantitative results that come next.

3. Theoretical Results

Here, we use the model above to show the link between multi-production and monetary non-neutrality. This provides intuition for our main quantitative results in Section 5.

We start with a basic, important result: Having multiple goods by itself is not sufficient to generate any difference in monetary non-neutrality relative to the single-product case.

Proposition 1. *The allocation of attention is invariant to the number N of goods that firms price when they pay no attention to good-specific shocks (either because $\sigma_z = 0$ or $\widehat{\pi}_{15} = 0$) and their information capacity is invariant to N, that is, $\kappa(N) = \kappa$.*

Proof. In Appendix E. ■

This result directly follows from the problem of a firm that sets the prices of N goods and pays no attention to good-specific shocks. Its objective is identical to the single-product case, only scaled by N. Thus, firms' allocation of attention is invariant to N if total attention is also invariant to N. However, if firms indeed pay no attention to good-specific shocks, then the model has the following, empirically strongly counterfactual prediction:

Lemma 1. *Prices set by a multi-product firm that pays no attention to good-specific shocks perfectly co-move.*

Proof. In Appendix Appendix E. ■

This lemma holds true because all goods are only subject to common shocks. In contrast, we document in Section 4.3 that there is strong within-firm dispersion of log price changes in U.S. data. This empirical fact is consistent with the idea that firms' prices react to good-specific shocks. Therefore, we next focus on a subspace of parameters where there is an interior solution for firms' attention to all three types of shocks. Below we show that our most important result holds as long as good-specific shocks account for a non-zero fraction of within-firm dispersion of price changes. Besides, although firm-specific shocks are not essential for our results, they are quantitatively important to match the between-firm dispersion of price changes in the data, which we report in Section 4.

Proposition 2. *For an interior solution of firms' attention and* $\kappa(N) = \kappa$*, attention to monetary shocks* κ_a^* *is decreasing in* N *for* $N < \widehat{N}$ *and increasing for* $N > \widehat{N}$ *where* $\hat{N}$ *solves*

$$\log \hat{N} + \frac{1}{2}\hat{N} = \kappa \log 2 - \log (x_2/x_1) - \log (x_2) - 1.$$

Proof. In Appendix E. ■

This proposition exposes how two opposing forces affect the attention to monetary shocks κ_a^* as N increases. On the one hand, the benefit of reducing profit losses by allocating attention to monetary and firm-specific shocks scales up with N, relative to allocating attention to good-specific shocks. However, the cost of spending attention is the same for all three shocks. This force is what we call "economies of scope in information processing." It creates an incentive for firms to increase attention to monetary and firm-specific shocks as they price more goods. On the other hand, firms that price more goods must allocate their information processing capacity to more shocks. This is because the relevant number of good-specific shocks increases with N. This force creates an incentive to firms to reallocate attention from all shocks to the new good-specific shocks it has to track, "thinning out" attention.

Proposition 2 shows that the latter force is dominant when $N < \widehat{N}$ and the former force when $N > \widehat{N}$. Given the definitions of x_1 and x_2, the threshold $\hat{N}$ increases in the portion of the volatility of frictionless prices due to good-specific shocks. It decreases in the portion due to monetary or firm-specific shocks. The proposition implies the following:

Lemma 2. *Monetary non-neutrality is increasing in* N *as* $N \to \infty$.

This lemma requires no proof. It simply states that the economies of scope are indeed the dominant force in information processing when the number of goods is large enough.

Next, we relax the assumption that we have maintained so far that total capacity is invariant to the number of goods. Instead, we impose discipline through what we call "frictional cost," a measure of profit losses:

Definition 1. *The frictional cost is the expected loss in profits per good that a firm bears due to its limited capacity to process information:*

$$C(N) = \mathbb{E}[\pi - \pi^*] = \frac{|\widehat{\pi}_{11}|}{2}\left[2^{-2\kappa_a^*}\sigma_\Delta^2 + \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_f^*}\sigma_f^2 + \left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_z^*}\sigma_z^2\right] \tag{8}$$

given its optimal allocation of attention $\left(\kappa_a^*, \kappa_f^*, \{\kappa_n^* = \kappa_z^*\}_{n\in\aleph_i}\right)$.

This expression has three components, which make up the expected *per-good* profit loss due to the imprecisely observed monetary shock, firm-specific, and good-specific shock. What will be important in imposing discipline on $\kappa(N)$ is that $C(N)$ is invariant to N except through the effect of N on the allocation of attention to the individual shocks. Before we impose such discipline, we first state a property of $C(N)$ for any functional form of $\kappa(N)$:

Proposition 3. *For an interior solution of firms' attention where κ_a^* is constant or decreasing in N, $C(N)$ is increasing in N.*

Proof. In Appendix E. ■

This property affects how we impose discipline on $\kappa(N)$: If the frictional cost $C(N)$ is increasing in N, firms could decrease their total frictional cost by decentralizing their pricing decisions among N independent price setters who each prices only one good and fully enjoys capacity κ. In short, all firms should be single-product firms. In contrast, in the data most price setting units price multiple goods. Thus, to impose discipline on $\kappa(N)$, we make the conservative assumption that capacity increases with N such that the frictional cost $C(N)$ is invariant to N. This is the minimum capacity such that multi-product firms have no incentives to split up their pricing decisions. This assumption does not impose any functional assumption on the costs of acquiring capacity; it only imposes a "free-entry condition" on the number of goods priced by a given price setter. Thus, our model with an exogenous number of goods priced by firms is equivalent to one in which it is chosen endogenously. This assumption also allows for a clean comparison among firms that price different numbers of goods by holding the severity of the friction constant. Subsection 3.3 discusses an alternative way to discipline to $\kappa(N)$ using the Lagrange multiplier.

Under these assumptions on $C(N)$ and $\kappa(N)$, we next obtain our main proposition:

Proposition 4. *For an interior solution of firms' attention and capacity $\kappa(N)$ such that the frictional cost $C(N)$ is invariant to N, κ_a^* is increasing in N. In particular,*

$$\kappa_a^*(N) = \kappa_a^*(1) + \frac{1}{2}\log_2\left(\frac{N+2}{3}\right) + \log_2\left[\frac{\sigma_\Delta\left(\kappa_a^*(N), \sigma_q\right)}{\sigma_\Delta\left(\kappa_a^*(1), \sigma_q\right)}\right]. \tag{9}$$

Proof. In Appendix E. ■

Proposition 4 presents our main result: the neutrality of money goes up as the number of goods increases, holding the frictional cost constant. This increase in the neutrality of money is reflected above by a widening gap between the attention to monetary shocks in a single-good economy, $\kappa_a^*(1)$, and in an economy with an arbitrary number N of goods, $\kappa_a^*(N)$. Both the second and the third term on the right-hand side increase in N, the latter because the volatility of the aggregate compound variable Δ_t also increases as $\kappa_a^*(N)$ increases.

Does $\kappa(N)$ have a reasonable shape as we keep the frictional cost constant? For brevity we do not elaborate on this point, but one can show that $\kappa(N)$ must be increasing and concave in N. The concavity is due to stronger economies of scope in information processing as N increases, so $\kappa(N)$ must increase less than linearly in N to keep the cost constant.

With our quantitative analysis in mind, the next lemma makes an important observation:

Lemma 3. *Proposition 4 holds for any level of volatility of firm- and good-specific shocks. It also holds if the volatility of these shocks are different among sectors where firms price different number of goods.*

This lemma requires no proof. It is important because empirically good-specific shocks may in principle not be the only source of within-firm price dispersion. However, it says that our main result is robust to variations in the volatility or importance of good-specific shocks as long as they account for a non-zero fraction of within-firm dispersion. We confirm this prediction quantitatively in Section 5.

The following proposition highlights the role of complementarities in our main result:

Proposition 5. *When $\kappa_a^*(1)$ is small, increasing N has large effect on reducing monetary non-neutrality.*

Proof. In Appendix E. ■

The intuition for this result can best be understood from Figure 1 which depicts the fixed point of (6) and (7) that determines $\kappa_a^*(N)$ and α (the responsiveness of aggregate prices to monetary shocks). Equation (6) is drawn in red and its intercept increases in N if the frictional cost is invariant to N. Equation (7) is drawn in blue and is unaffected by N. The equilibrium is where these two curves intersect. The key observation is that the blue line is flatter for low values of α. Therefore, when $\kappa_a^*(1)$ is small, an increase in N pushes up the red to the green line, so a small increment in attention to monetary shocks has a large effect on reducing monetary non-neutrality (increasing α).

In intuitive terms, more attention to monetary shocks increases the responsiveness of individual prices to monetary shocks, so the responsiveness of aggregate prices also increases. Due to strategic complementarity, individual prices are very sensitive to variations in the aggregate price, so individual prices become even more responsive to monetary shocks. This amplification mechanism is stronger when firms' attention to monetary shocks is small.

This proposition is behind an important quantitative result in Section 5, demonstrating the interaction between multi-product firms and complementarities. As is standard in the literature, strong complementarity is needed for a single-product firm economy to generate strong monetary non-neutrality given a "small" friction. In multi-product firm economies, we will show later on that attention to monetary shocks is still a small fraction of total attention; yet monetary non-neutrality is now largely muted due to multi-product firms.

3.1. Heterogeneous Firms

We now augment our model to allow for the coexistence of firms that price different numbers of goods. We do so to provide intuition to another of our main quantitative result obtained when we calibrate our model to PPI data: Even the prices of single-product firms react very quickly to monetary shocks when these firms coexist with multi-product firms.

Consider now that there are G groups of firms such that firms in group $g = 1, ..., G$ price N_g goods. Each group has measure ω_g satisfying $\sum_{g=1}^{G} \omega_g = 1$. The processes for firm- and good-specific shocks are independent for each group, so these shocks still wash out when prices are aggregated. All other parameters are the same for all groups.

In this economy, the solution of κ_a^* is still governed by (6) where we only replace N by N_g. The guess $p_t^* = \alpha q_t$ holds for

$$\alpha = \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} \frac{\sum\limits_{g=1}^{G} \omega_g \left(1 - 2^{-2\kappa_a^*(N_g)}\right)}{1 - \left(1 - \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\right) \sum\limits_{g=1}^{G} \omega_g \left(1 - 2^{-2\kappa_a^*(N_g)}\right)}$$

All our results above also hold; the only modification is that (4) is now $\kappa_a^*(N_g) = \kappa_a^*(1) + \frac{1}{2}\log_2\left(\frac{N_g+2}{3}\right)$. Then, again, the difference in κ_a^* chosen by a multi-product firm and a single-product firm is increasing in N_g.

Lemma 4. *Single-product firms pay more attention to monetary shocks in an economy where they coexist with multi-product firms relative to an economy with only single-product firms. The gap in attention between the two cases is smaller as strategic complementarity is stronger.*

This lemma requires no proof. When single-product firms interact with firms that pay more attention to monetary shocks, the responsiveness of aggregate prices to monetary shocks, α, is higher than in an economy with only single-product firms. The volatility of Δ_t is thus also higher, so single-product firms choose higher attention to monetary shocks. This effect is stronger when strategic complementarity is stronger.

This is an important observation because, in our model calibrated to PPI data, single-product firms pay only slightly less attention to monetary shocks than firms pricing the median number of goods. Therefore, their prices respond to monetary shocks almost as quickly as the prices of multi-product firms. This is partially justified by the strong strategic complementarity we assume which is in line with standard calibrations in the literature.

3.2. Intertemporal Economies of Scope

As a last mechanism that strengthens our quantitative results is the intertemporal dimension of economies of scope in information processing. Without loss of generality, we illustrate the importance of this dimension in the case of a single-good firm. Thus, assume that firms are hit only by monetary and good-specific shocks. Also assume that Δ_t and z_t are $AR(1)$ with persistence ρ_Δ and ρ_z. We present details of the solution in Appendix A.

The first-order conditions of this problem imply that

$$\kappa_a^* + f\left(\rho_\Delta, \kappa_a^*\right) = \kappa_z^* + f\left(\rho_z, \kappa_z^*\right) + \log_2 x \tag{10}$$

where $x \equiv \left|\widehat{\pi}_{11}\right| \sigma_\Delta \sqrt{1-\rho_\Delta^2} / \widehat{\pi}_{15} \sigma_z \sqrt{1-\rho_z^2}$ and $f\left(\rho_h, \kappa_h\right) = \log_2\left(1 - \rho_h^2 2^{-2\kappa_h}\right)$ for $h = a, z$. Then, if ρ_z goes down, there are two opposite effects. The first effect is that paying attention to good-specific shocks becomes less useful for future decisions. This effect is captured by the fact that an increase in $f\left(\rho_z, \kappa_z^*\right)$ implies an increase in κ_a^* relative to κ_z^*. We refer to this as the intertemporal dimension of the economies of scope. The second effect is that lower ρ_z increases the volatility of the exogenous disturbances of good-specific shocks, so x decreases. This effect gives firms incentives to increase κ_z^* relative to κ_a^*.

Which of these effects dominates? There is no clear-cut answer based on theory. However, quantitatively, we find in the next section that the model can only match the average size of price changes in the data if σ_z is decreased as ρ_z is decreased. This means lower monetary non-neutrality as the persistence of idiosyncratic shocks goes down.

3.3. An Alternative to Discipline Information Capacity?

An alternative metric for the severity of the friction is to assume that the shadow price of information processing is the same for firms regardless of the number of goods. We now argue that this approach does not pose a better alternative for studying monetary non-neutrality.

First, if the shadow price of information processing capacity were constant, then this would imply an increase in the neutrality of money as the number of goods increases, like in our main proposition. This effect can be seen directly from the first-order condition with respect to κ_a where we have omitted firm-specific shocks without loss of generality:

$$\frac{\beta}{1-\beta}\frac{|\widehat{\pi}_{11}|}{2}2^{1-2\kappa_a^*}\sigma_\Delta^2 \log(2)\, N = \lambda$$

where λ is the the shadow price of information processing, the Lagrange multiplier (the same condition holds if firm-specific shocks are reintroduced). What drives the result is that an increase in the number of goods N – equivalent to an increase in the scale of firms, just like in the data – implies an increase in attention to monetary shocks κ_a and hence more monetary neutrality when λ is held constant. However, this approach of holding λ constant would not hold profit losses constant, which we had chosen to discipline capacity consistently.

Second, if one were to hold λ/N constant as N changes, this would evidently imply that κ_a^* would be constant. In words, λ/N constant means that the marginal cost of expanding information capacity is higher for firms that price more goods (which are also bigger firms). We hold this assumption a priori to be implausible: It implies, for instance, that buying software to support the pricing process of a firm is more expensive as firms decide more prices (or when their total sales are larger). In addition, a constant λ/N also implies that firms would like to split up their pricing decisions, which, as argued, is inconsistent with empirical evidence documented in the next section.

4. Empirical Regularities on Multi-Product Pricing Behavior

We provide several new empirical regularities in this section about how multi-product firms set prices. We later use these regularities to calibrate our model.

4.1. Data Sources

As our main data source, we use monthly transaction-level micro price data collected by the U.S. Bureau Labor Statistics (BLS) to construct the Consumer Price Index (CPI) and the Producer Price Index (PPI).[7] We generate our results by computing statistics for the whole sample and for four bins. We assign firms to these bins according to their number of goods in the data. Thus, we can track how key statistics change as the number of goods increases. All statistics, including standard deviations, are reported in Table 1, and we describe our detailed data manipulations in an online companion appendix.

4.2. Multi-Product Firms

Based on various sources, we find that retailers sell many goods, while producers sell a much smaller number of goods. On the producer side, counting the number of goods priced by a single firm in the PPI, we find that the median (mean) is 4 (4.13) with a standard deviation of 2.55 goods. Only 1.5% of firms price a single good. These estimates are a lower bound due to sampling constraints, which however will only strengthen our results. An alternative estimate comes from Bernard et al. (2010). They define a product as a category of the five-digit Standard Industrial Classification in the US Manufacturing Census data, which is less narrow than our definition. They report that a firm prices on average 3.5 goods. Using the PPI data we compute moments of pricing for four bins, when firms price a median of 2 (bin 1), 4 (bin 2), 6 (bin 3), and 8 (bin 4) goods.

[7] Nakamura and Steinsson (2008) or Bils and Klenow (2004) describe the CPI data in detail, while for example Bhattarai and Schoenle (2014) describe the PPI data.

On the retailer side, the median (mean) number of goods sampled from a single CPI outlet is 1.39 (2.05) with a standard deviation of 2.03 goods.[8] In these data, 87% (75%) of outlets have less than 3 (2) goods. Given that outlets tend to be retailers, we consider that CPI data do not provide a reliable, realistic estimate of the number of goods. We therefore report moments by bins for information only, and use the whole sample for calibration. A more plausible estimate of the number of goods priced by retailers comes from the Food Marketing Institute (FMI) 2010 Report.[9] The FMI reports an average of $38,718$ items per retailer.[10] Similar evidence comes from Rebelo et al. (2010) who use data from one particular retailer that prices approximately 60,000 items. What we take away from these various sources is that retailers sell many goods.

Importantly, what matters for our model is not only that there are many goods per firm but also who sets prices. We take the view in this paper that there is one single price setter per firm. The fact that our data has firms defined as "price-forming units" is consistent with this idea that one unit has to process all relevant information as well as the fact that decision power in firms tends to be centralized. Further evidence may be found in a case study by Zbaracki et al. (2004) which reports that all regular prices are decided at headquarters while all sales prices are decided by local managers in small geographical areas. At both levels there is a single decision unit setting prices for all goods. An extreme alternative view could be that every single good is priced completely independently within firms – without any shared information processing – which seems rather far-fetched in our opinion. However, we fully acknowledge that such other interpretations of the data are possible and go for the one that is theoretically and quantitatively more interesting.

[8] The median is not integer because for the following reason: First, we compute for each outlet over time its mean number of goods. Due to exit and entry, this may not be an integer. Second, we take the median or mean across firms. The same reasoning applies to the PPI data.

[9] The FMI is an industry association that represents $1,500$ food retailers and wholesalers in the U.S.. The members are large multi-store chains, regional firms and independent supermarkets, retailers and drug stores with a combined annual sales volume of $680 billion. http://www.fmi.org/about-us/who-we-are

[10] http://www.fmi.org/research-resources/supermarket-facts

4.3. Are There Good-Specific Shocks?

While we cannot observe good-specific shocks or quantify their variance, we can observe the behavior of good-specific prices. It turns out that their behavior is consistent with the existence of good-specific shocks, our second crucial modeling assumption.

We compute the ratio of the within-firm dispersion relative to the total cross-sectional dispersion of log non-zero price changes.[11] Specifically, we compute

$$r = \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{i=1}^{I_t}\sum_{n\in\aleph_i}\left(\Delta p_{nt} - \overline{\Delta p}_{it}\right)^2 \Big/ \sum_{i=1}^{I_t}\sum_{n\in\aleph_i}\left(\Delta p_{nt} - \overline{\Delta p}_{t}\right)^2\right]$$

where $\overline{\Delta p}_{it}$ is the mean absolute size of non-zero log price changes $\Delta p_{nt} \equiv p_{nt} - p_{nt-1}$ across all goods sampled for firm i at time t and $\overline{\Delta p}_{t}$ is the grand total mean.[12]

In the PPI data, we find that this ratio r is non-zero and increasing as firms price more goods, from 36.5% (for bin 1) to 72.4% (for bin 4). In the full PPI sample, 59.1% of the total dispersion is due to within-firm variance. In the full CPI sample, similarly 51.6% is due to within-firm dispersion. This result also holds when we take into account sales prices: Then, the ratio becomes 56.5%. Sales have no systematic impact on the ratio.

We take this within-firm dispersion as evidence of prices responding to good-specific shocks. Although this may not be the only explanation, our theoretical and quantitative results only rely on the assumption that prices respond to some extent to these shocks.[13]

[11] In ANOVA terminology, this is the ratio of the SSW to the SST.

[12] An alternative way to measure relative dispersion is to compute, by bin, the ratio of the average firm variance to the overall variance. This includes Bessel correction factors of the kind $N-1$. We have done this and we find that our results are both qualitatively and quantitatively robust. The trends with the number of goods in particular are unaffected.

[13] Further evidence suggestive of good-specific shocks comes from Golosov and Lucas Jr. (2007) and Nakamura and Steinsson (2008). They point out that feeding only aggregate shocks into models leads to difficulties generating the high empirical frequency of negative price changes and the large magnitudes of micro price changes. However, their analysis does not focus on multi-product firms and leaves open the possibility that simply firm-specific shocks are highly volatile. Thus, we additionally document within-firm dispersion.

4.4. Statistics for Calibration

Here, we present several additional statistics that allow us to calibrate the allocation of attention in our quantitative exercise. First, we compute the average size of absolute non-zero price changes. This will help us pin down the magnitude of equilibrium price changes in our calibration. We denote this statistic as $|\overline{\Delta p}|$. Labeling time as t, firms as i and goods produced by firm i at time t as $n \in \aleph_{it}$,

$$|\overline{\Delta p}| = \frac{1}{I}\sum_{i=1}^{I}\left[\frac{1}{N_i}\sum_{n\in\aleph_i}\left[\frac{1}{T_n}\sum_{t=1}^{T_n}|\Delta p_{nt}|\right]\right]$$

where $\Delta p_{nt} \equiv p_{nt} - p_{nt-1}$ is the non-zero log price change for good n, T_n is the total number of periods for which inflation for good n can be computed, N_i is the number of goods of firm i in the sample, and I is the total number of firms in the sample.

In the CPI data, the mean (median) absolute size of regular price changes is 11.3% (9.6%), according to Klenow and Kryvtsov (2008). Our own computation gives us 11.01% (8.42%).[14] If we take into account sales, this number becomes somewhat larger. In the PPI data, the mean absolute size of price changes for the whole sample is 7.8%. For bins 1 to 4, the magnitudes are as follows: 8.5%, 7.9%, 6.8%, and 6.5%. As the number of goods increases, the magnitude of price changes becomes smaller.

Second, we compute a measure of intertemporal economies of scope. We do so by calculating the serial correlation of price changes, denoted by ρ for the whole sample and by ρ_k for bins $k \in (1, 2, 3, 4)$. We obtain this statistic by computing median quantile estimates of an $AR(1)$ coefficient for $\Delta p_{n,k,t}$, conditional on non-zero price changes, such that $\hat{\rho}_k = argmin_{\rho_k} E[|\Delta p_{n,k,t} - \rho_k \Delta p_{n,k,t-1}|]$. We find a median estimate of -0.29 in the CPI sample. Bils and Klenow (2004) estimate a comparable first-order serial correlation of -0.05.[15]

[14] The difference is due to our focus on outlets as unit of analysis, which changes the aggregation approach.

[15] Bils and Klenow (2004) compute their estimate as the average of $AR(1)$ coefficients for inflation of 123 categories in the CPI data. They include sales and zero price changes, between 1995 and 1997. We differ in our methodology and by focusing on the period from 1989 to 2009. Qualitatively, both approaches give the

If we take sales into account, our estimate becomes more negative due to the nature of sales. In the PPI data, our estimate of the AR(1) coefficient is -0.04. It ranges from -0.05 in bin 1 to -0.03 in bin 4. All coefficients are statistically highly significant.

Finally, we compute the cross-sectional variance $\widetilde{\sigma}$ of price changes. We use this as an additional moment when we discuss our calibration of information capacity since the latter cannot be directly measured. This statistic is defined as

$$\widetilde{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{i=1}^{I_t}\sum_{n\in\aleph_i}\left(\Delta p_{nt} - \overline{\Delta p}_t\right)^2 / \left(\sum_{i=1}^{I_t} N_{it} - 1\right)\right]$$

where $\overline{\Delta p}_t$ is the average of non-zero absolute log price changes Δp_{nt} of all goods sampled at time t, N_{it} is the total number of goods sampled for firm i at time t, I_t is the total number of firms at time t, and T is the total number of periods in our data. We find that the cross-sectional variance is 3.51% (2.65%) in the full PPI (CPI) sample. If we consider sales in the CPI, price changes are again more dispersed. There is no clear trend in the PPI data.

4.5. Robustness: Number of Goods or Firm Size?

One concern might be that firm size, not the number of goods, is driving our empirical results. However, we can explicitly control for firm size when constructing the above statistics. We do so by conditioning our firm-level statistics on the number of employees, which we take as our measure of size. We obtain the necessary employment data directly from the PPI database where it is recorded at resampling every two to five years.

We find very strong evidence that our key statistics and trends are robust to controlling for firm size. Our most important moment concerns the within-firm dispersion ratio. When we control for firm size, the within-firm dispersion ratio remains positive, in each bin and in the full sample. This corroborates our second main assumption, the existence of good-specific shocks. The ratio also increases monotonically, as summarized in Table D.8 in the Online

same results.

Appendix. Similar results hold for the absolute size, the persistence and the cross-sectional dispersion of price changes. Overall, these findings strongly suggest that the number of goods per firm is crucial in determining key modeling statistics.

5. Quantitative Results

This section reports our results obtained after calibrating a version of our model that allows for a general specification of the stochastic processes of the shocks. This general problem and its numerical solution are presented in Appendix B.

5.1. The Quantitative Importance of Multiproduct Firms

First, we quantitatively confirm our main theoretical result from Proposition 4: multi-product pricing has a large effect on monetary non-neutrality. We show this by contrasting an economy with only single-product firms and one with only multi-product firms when the severity of the friction is the same in both economies, measured by the frictional cost per good.

Thus, to start, we calibrate our model exactly as our benchmark (Mackowiak and Wiederholt (2009)). We drop the firm-specific shocks ($\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} = 0$) and calibrate parameters in the pricing rules to be $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} = 0.15$ and $\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} = 1$.[16] Nominal aggregate demands shocks (in short, monetary shocks) are assumed to be $AR(1)$ with $\rho_q = 0.95$ and $\sigma_q = 2.68\%$ to fit the estimates using quarterly GNP detrended data spanning 1959:1–2004:1. We assume idiosyncratic shocks to be as persistent as monetary shocks with volatility $\sigma_z = 11.8\sigma_q$ to match the 9.6% of median absolute non-zero log price changes per good in the U.S. CPI data. To get a numerical solution these processes are approximated by $MA(20)$ processes with parameters decreasing linearly. Information processing capacity is assumed to be $\kappa(1) = 3$.

[16] If we drop good-specific shocks instead of firm-specific shocks, the result in Proposition 1 is obtained, so multi-product pricing has no effect on monetary non-neutrality.

For single-product firms, we replicate exactly the solution in Mackowiak and Wiederholt (2009). Firms' attention is $\kappa_a^*(1) = 0.09$ to monetary shocks and $\kappa_z^*(1) = 2.91$ to idiosyncratic shocks.[17] This yields large and long-lasting monetary non-neutrality. Figure 2 depicts this result graphically, showing the response of prices after a 1% innovation to nominal demand. Prices under rational inattention (the blue line) absorb only 2.8% of the innovation on impact. Their response remains sluggish relative to the response of frictionless prices (black line) for all 20 periods and their cumulated response is only 22% of the cumulated response of frictionless prices. The per-good cost of the friction is 0.21% of steady state revenues $\overline{Y}$, which is considered "small."

We then compute the response of prices when firms price $N = 2, 4$ and 8 goods. Calibration of σ_z must be adjusted in each case to match the target moment of micro prices. We vary total information capacity targeting a per-good frictional cost of $0.21\%\overline{Y}$. This corresponds directly to Proposition 4. The responses of aggregate prices in these cases are depicted in Figure 2.

We find that even in the case of $N = 2$ monetary non-neutrality is largely reduced. For $N = 2$ (in red), $\kappa_a^*(2) = 0.36$ and $\kappa_z^*(2) = 2.92$, prices absorb 15% of the innovation on impact, their response remains sluggish only for 7 periods (the output deviation is less than 5% of the 1% innovation thereafter) and the cumulative response is 74% of the frictionless response. In short, monetary non-neutrality is cut by three. For $N = 4$, $\kappa_a^*(2) = 0.58$ and $\kappa_z^*(2) = 2.90$, prices absorb 15% of the innovation on impact, prices remain sluggish for 4 periods and their cumulative response is 86% of the frictionless response. For $N = 8$, $\kappa_a^*(2) = 0.9$ and $\kappa_z^*(2) = 2.87$, prices absorb 49% of the money shock on impact to become almost neutral after 2 periods. Note that in all these cases attention to monetary shocks is

[17]In our numerical algorithm, we use a tolerance of 2% for convergence, exactly as in Mackowiak and Wiederholt (2009). We keep this criterion for comparability with Mackowiak and Wiederholt (2009) in the following sections, but from Section 5.4 on, we replace it with a tighter tolerance of 0.01%. If we use the tighter convergence criterion in this and the next sections, we obtain even starker predictions from introducing multi-product firms.

only a small fraction of the firm's total capacity; yet aggregate responses are very different.

Recall that retailers, the relevant pricing units in the CPI data, price a large number of goods, much larger than 8. We thus conclude that the multi-product feature of the price setter can be very relevant quantitatively in a rational inattention model where price setters are meant to be retailers. In particular, our model predicts almost perfect monetary neutrality when these retailers price a realistic number of goods. When retailers price a single good, the model by contrast yields strong monetary non-neutrality.

5.2. Robustness

Here, we verify that Proposition 4 continues to hold quantitatively when we allow for less persistent idiosyncratic shocks and for the existence of both good- and firm-specific shocks, as implied by the data. Our main result is also robust to substantial variations in the relative importance of the two shocks, as described in Lemma 3.

First, we find that allowing for less persistent idiosyncratic shocks increases the neutrality of money. This result is due to intertemporal economies of scope in information processing; its intuition is explained in Section 3.2. Focusing on the case $N = 1$, we calibrate the persistence of idiosyncratic shocks to match the -0.05 serial correlation of price changes in the CPI reported by Bils and Klenow (2004) and alternatively our own computation (-0.29, see Table 1). While methodologically different,[18] in both cases the persistence of idiosyncratic shocks has to be substantially less than for monetary shocks.[19] We continue to target the per-good frictional cost to be $0.21\%\overline{Y}$. We present details in the appendix.

Figure 3 depicts the responses of prices to a 1% innovation of the monetary shock. In

[18] Bils and Klenow (2004) compute this statistic by averaging the coefficient of $AR(1)$ regressions for inflation of 123 categories in the CPI data, including sales and zero price changes, between 1995 and 2007. We compute the coefficient from an $AR(1)$ quantile regressions for non-zero inflation of each item in the CPI data, excluding sales and zero price changes, between 1989 and 2009. Our computation is consistent with the other statistics we report.

[19] In the first case, we set z_{jt} to follow an $MA(5)$ with $\sigma_z = 10.68\sigma_q$. In the second case, we set z_{jt} to follow a $MA(1)$ with coefficient 0.33 and $\sigma_z = 9.74\sigma_q$.

both cases, the price response on impact is 7% of the shock and output is within 5% of the frictionless again after 12 periods. The cumulative response of prices is 52% of the frictionless case. This is substantially larger than the 22% cumulative response of our benchmark calibration for $N = 1$.

As robustness check, we calibrate our model to match three instead of only one moment from micro data: the 9.6% average size of price changes, the -0.05 serial correlation of price changes (results are almost identical if we target on -0.29 serial correlation), and the 51.6% of the contribution of within-firm dispersion of price changes to total cross-sectional dispersion (shown in Table 1). As explained above, this last moment motivates the introduction of firm- and good-specific shocks. In particular, we assume that both types of idiosyncratic shocks are equally persistent and that good-specific shocks account for all within-firms dispersion of price changes. The per-good frictional cost is $0.21\%\overline{Y}$.

Our main result regarding multi-product pricing and monetary non-neutrality remains almost identical to the above. When $N = 4$, prices absorb 30% of the monetary innovation on impact, their response remains sluggish for only 4 periods and their cumulative response is 86% of the frictionless response. For $N = 8$, prices absorb 52% of the monetary shock to become almost fully neutral after 2 periods.

To deal with the concern that good-specific shocks may not be the only source of within-firms dispersion of price changes, we calibrate the volatility of good-specific shocks targeting a within dispersion ratio of only 10%. Second, we repeat our calibration with a 75% target. We hold all other calibration targets fixed. Again, our results remain almost identical, shown in Figure 4. The reason is, as explained in Lemma 3, that our results do not depend on good-specific shocks being important, but only on the assumption that prices respond to some extent to these shocks (so firms pay positive attention to them).

5.3. Producers as Price Setters

We now take the view of interpreting price setters in the model as good producers. For this we must calibrate our model to the same three moments used above but now for PPI data. Further, since we have moments for four bins from the PPI (with firms pricing a median of 2, 4, 6 and 8 goods), we use a version of our model in Section 3.1 with firms pricing a heterogeneous number of goods that allows for a general specification of shocks.

Our main finding is that prices of all firms have very similar responses to monetary shocks regardless of the bin they belong to, shown in Figure 5. This is due to the effect of strategic complementarity in pricing decisions, as explained in Section 3.1. The response of aggregate prices is 16.7% of the monetary shock on impact, and the cumulative response of aggregate prices is 75% of the frictionless price response. Recall that firms absorbed 2.8% of the shock on impact and 22% cumulatively in the pure single-good economy. Therefore, we conclude that even though producers price a much smaller number of goods than retailers, monetary non-neutrality is still sizable in a calibrated rational inattention model where multiple-good price setters are producers. Multi-good price setting is still very important quantitatively since monetary non-neutrality is much smaller than in an identical single-product economy.

5.4. The Importance of Strategic Complementarity

Monetary non-neutrality is strongest when we study strategic complementarity in a version of our model calibrated to PPI data. In particular, we make two points. To make the first, assume that there is a single-product firm in this economy which has near-zero weight in aggregate prices (in the data the weight is about 1%). What we find is in line with Lemma 4: The price of the single-product firm has a very similar response to prices of multi-product firms. This is due to strategic complementarity, as explained in Section 3.1. Therefore, we conclude that multi-good price setting is important for the responsiveness of prices to monetary shocks even for single-product firms when they coexist with multi-product firms.

Our second point has to do with the effect of reducing strategic complementarity in the model. In particular, we increase $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$ from 0.15 to 0.85. This modification has two effects: On the one hand, as the discussion of Proposition 5 implies, an increase in attention to monetary shocks has a milder effect on reducing monetary non-neutrality when $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$ is higher. On the other hand, for a given level of attention to monetary shocks, monetary non-neutrality is lower when the extent of complementarities is lower. This result comes from equation (7). Our calibrated model now has prices absorbing 23% of the monetary shock on impact, which is higher than the 16.7% when $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} = 0.15$. Also, there are almost no real effects after only 6 periods now, and the cumulated response of prices is 84% that of frictionless prices. We conclude that decreasing strategic complementarity does not help in a multi-product setting to generate stronger monetary non-neutrality.

5.5. Calibrating Information Capacity

In the analysis so far, we have pinned down total capacity $\kappa(N)$ by imposing a per-good frictional cost of 0.21% of steady state revenues. As we argued, this is one reasonable way to discipline capacity because it is scale-invariant and it is consistent with multi-product pricing within the model. Here we make three points about alternative calibrations for $\kappa(N)$.

First, we repeat our exercises in Section 5.2 but hold the Lagrange multiplier on the capacity constraint constant across goods: $\lambda(N) = \bar{\lambda}$. Just as Section 3.3 predicts, we then find that the cumulative price response increases from 52% to 83% of the frictionless response, which Table 4 summarizes. Monetary neutrality is extremely high. This is because total capacity $\kappa(N)$ increases such that firms' attention to good specific shocks is invariant to N. Due to economies of scope, this means that firms' attention to monetary and firm-specific shocks increases with N. Then, strategic complementarity implies that a little more attention to monetary shocks means much less monetary non-neutrality.

The second point highlights why we do not pin down $\kappa(N)$ directly from the micro data. This is because the predicted micro price moments of our model are almost invariant to small variations of $\kappa(N)$, but the macro predictions of our model are highly sensitive to such variations. To see this, we solve our model from Section 5.2 for $N = 2$ and $N = 4$ on a grid of $\kappa(N)$. Tables 5 and 6 show that predicted micro moments are very similar to each other.

A final point is that, since we cannot calibrate $\kappa(N)$ directly from data, we may ask instead about the implications for monetary non-neutrality in the model when we vary the per-good frictional cost. We use our calibrated model from Section 5.2 for this purpose. We find that an increase of monetary non-neutrality by a factor of 2 (3) is associated with an increase in the friction by a factor of approximately 2 (3) as well. We illustrate this trade-off for a wider range of frictional cost in Figure 6. To yield the same monetary non-neutrality as in our benchmark, we need a frictional cost of 1.9% of steady state revenues, much higher than the 0.21% in our bechmark, the 0.32% obtained by Midrigan (2012) for a menu cost model with firms pricing two goods, or the 0.23% of revenues computed by Zbaracki et al (2004) as "informational and managerial cost" of changing prices.

6. Conclusion

Our results show that multi-product pricing can have a big quantitative effect on monetary non-neutrality in a model of rationally inattentive firms. In particular, we find that under the same calibration that yields strong monetary non-neutrality when firms price a single good, monetary non-neutrality almost vanishes when firms price eight goods or more. This result is robust to several robustness checks, and our model assumptions are consistent with evidence from CPI and PPI micro data.

Two directions for future work directly follow from the results: First, while there is price stickiness in the data, rational inattention pricing models do not feature price stickiness. To fully match the data, rational inattention models need to allow for such price stickiness. Second, firms make many decisions besides pricing. Our main mechanism of economies of scope should equally apply there, and potentially yield new implications.

References

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

Bernard, A., Redding, S., Schott, P., 2010. Multi-product firms and product switching. American Economic Review 100, 70–97.

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.

Bils, M., Klenow, P.J., 2004. Some evidence on the importance of sticky prices. Journal of Political Economy 112, 947–985.

Cheremukhin, A., Restrepo-Echavarria, P., Tutino, A., 2012. The assignment of workers to jobs with endogenous information selection. Meeting Papers 164. Society for Economic Dynamics.

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

Klenow, P.J., Kryvtsov, O., 2008. State-dependent or time-dependent pricing: Does it matter for recent u.s. inflation? The Quarterly Journal of Economics 123, 863–904.

Luo, Y., 2008. Consumption dynamics under information processing constraints. Review of Economic Dynamics 11, 366–385.

Luo, Y., Nie, J., Young, E.R., 2012. Robustness, information–processing constraints, and the current account in small open economies. Journal of International Economics 88, 104–120.

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

Mackowiak, B., Wiederholt, M., 2015. Business cycle dynamics under rational inattention. Forthcoming, Review of Economic Studies.

Mackowiak, B.A., Wiederholt, M., 2011. Inattention to Rare Events. CEPR Discussion Papers 8626. C.E.P.R. Discussion Papers.

Matejka, F., 2010. Rationally Inattentive Seller: Sales and Discrete Pricing. Working Paper wp408. CERGE-EI.

Matejka, F., McKay, A., 2015. Rational inattention to discrete choices: A new foundation for the multinomial logit model. American Economic Review 105, 272–98.

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

Mondria, J., 2010. Portfolio choice, attention allocation, and price comovement. Journal of Economic Theory 145, 1837–1864.

Mondria, J., Wu, T., 2010. The puzzling evolution of the home bias, information processing and financial openness. Journal of Economic Dynamics and Control 34, 875–896.

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

Paciello, L., Wiederholt, M., 2014. Exogenous Information, Endogenous Information, and Optimal Monetary Policy. Review of Economic Studies 81, 356–388.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. Journal of Financial Economics 80, 563–602.

Rebelo, S., Jaimovich, N., Eichenbaum, M., 2010. Reference Prices and Nominal Rigidities. 2010 Meeting Papers 1049. Society for Economic Dynamics.

Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423, 623–656.

Sheshinski, E., Weiss, Y., 1992. Staggered and synchronized price policies under inflation: The multiproduct monopoly case. Review of Economic Studies 59, 331–59.

Sims, C.A., 1998. Stickiness. Carnegie-Rochester Conference Series on Public Policy 49, 317–356.

Sims, C.A., 2003. Implications of rational inattention. Journal of Monetary Economics 50, 665–690.

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

Venkateswaran, V., Hellwig, C., 2009. Setting the right prices for the wrong reasons. Journal of Monetary Economics 56, S57–S77.

Woodford, M., 2009. Convergence in macroeconomics: Elements of the new synthesis. American Economic Journal: Macroeconomics 1, 267–79.

Woodford, M., 2012. Inflation Targeting and Financial Stability. NBER Working Papers 17967. National Bureau of Economic Research, Inc.

Zbaracki, M.J., Ritson, M., Levy, D., Dutta, S., Bergen, M., 2004. Managerial and customer costs of price adjustment: Direct evidence from industrial markets. The Review of Economics and Statistics 86, 514–533.

7. Tables and Figures

Table 1: Multi-Product Firms and Moments from CPI and PPI data

CPI	1-3 Goods	3-5 Goods	5-7 Goods	>7 Goods	All
# goods, mean	1.47	3.89	6.02	10.82	2.05
# goods, median	1.00	3.85	6.00	9.00	1.39
Absolute size of price changes	10.87%	11.64%	11.69%	12.55%	11.01%
	(0.03%)	(0.09%)	(0.15%)	(0.11%)	(0.03%)
Within ratio of $\lvert\Delta p\rvert$	20.9%	55.8%	62.8%	79.0%	51.6%
	(0.3%)	(0.4%)	(0.4%)	(0.4%)	(0.6%)
Cross-sectional variance	1.93%	2.65%	3.60%	2.85%	2.65%
	(0.52%)	(0.70%)	(0.89%)	(0.50%)	(0.31%)
Serial correlation	-0.248	-0.307	-0.334	-0.355	-0.291
	(0.0008)	(0.0013)	(0.0022)	(0.0015)	(0.0006)
PPI					
# goods, mean	2.19	4.02	6.03	10.25	4.13
# goods, median	2	4	6	8	4
Absolute size of price changes	8.5%	7.9%	6.8%	6.5%	7.8%
	(0.13%)	(0.09%)	(0.14%)	(0.16%)	(0.10%)
Within ratio of $\lvert\Delta p\rvert$	36.5%	54.6%	67.2%	72.4%	59.1%
	(0.7%)	(0.6%)	(0.8%)	(1.0%)	(0.6%)
Cross-sectional variance	3.72%	3.60%	2.91%	3.64%	3.51%
	(0.20%)	(0.19%)	(0.15%)	(0.22%)	(0.10%)
Serial correlation	$-.050$	$-.057$	$-.033$	$-.032$	$-.043$
	(0.0024)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Share of total employment	25.0%	27.7%	16.0%	31.3%	100%

NOTE: We compute the above statistics using the monthly micro price data underlying the PPI and CPI. The time periods are from 1998 through 2005, and 1998 through 2009, respectively. We compute all statistics for firms with less than 3 goods (bin 1), with 3-5 goods (bin 2), with 5-7 goods (bin 3), >7 goods (bin 4), and the full sample. First, we compute the time-series mean of the number of goods per firm. We then report the mean (median) number of goods across all firms. Second, we start by computing the time-series mean of the absolute value of log price changes for each good in a firm. We take the median across goods within each firm, then report means across firms. Standard errors across firms are given in brackets. Third, we compute the monthly within dispersion ratio as the ratio of two statistics: first, the sum of squared deviations of the absolute value of individual, non-zero log price changes from their average within each firm, summed across firms; second, the sum of squared deviations of the absolute value of individual, non-zero log price changes from their cross-sectional average. We then report the time-series mean. Standard errors across monthly means are given in brackets. Fourth, we estimate the first-order auto-correlation coefficient of non-zero price changes using a median quantile regression. Fifth, we compute the monthly cross-sectional variance of absolute log price changes and then report standard errors of this monthly statistic. Finally, we compute the share of employment relative to total employment in each category at the time of re-sampling in 2005.

Table 2: Multi-Product Firms and Within-Firm Dispersion Ratio, Robustness

CPI	1-3 Goods	3-5 Goods	5-7 Goods	>7 Goods	All
Within ratio of Δp					
Mean	8.8%	32.8%	45.6%	64.7%	35.9%
Median	9.2%	32.7%	44.5%	62.0%	35.2%
Std. Error	(0.2%)	(0.3%)	(0.4%)	(0.5%)	(0.6%)
Within ratio of Δp, sales					
Mean	25.9%	57.6%	64.6%	81.4%	56.5%
Median	26%	57.8%	64.8%	82.4%	58.7%
Std. Error	(0.2%)	(0.2%)	(0.3%)	(0.2%)	(0.4%)
PPI					
Within ratio of Δp					
Mean	18.4%	31.2%	44.3%	54.0%	38.1%
Median	18.1%	30.1%	44.4%	53.3%	37.4%
Std. Error	(0.7%)	(0.9%)	(1.1%)	(1.0%)	(0.7%)

NOTE: We compute the above statistics using the monthly micro price data underlying the PPI and CPI. The time periods are from 1998 through 2005, and 1998 through 2009, respectively. We compute all statistics for firms with less than 3 goods (bin 1), with 3-5 goods (bin 2), with 5-7 goods (bin 3), >7 goods (bin 4), and the full sample. For the first 3 rows of the CPI and PPI panels, we compute the monthly within dispersion ratio as the ratio of two statistics: first, the sum of squared deviations of the individual log price changes, *including* zeros, from their average within each firm, summed across firms; second, the sum of squared deviations of individual log price changes, including zeros, from their cross-sectional average. We also compute the ratio for all non-zero price changes in the CPI, but include sales. This is summarized in rows 4-6 in the CPI panel. We then report the time-series mean and medians. Standard errors across monthly means are given in brackets.

Table 3: Moments from the PPI and the Model

	1-3 Goods	3-5 Goods	5-7 Goods	>7 Goods	All
Absolute size of price changes, data	8.5%	7.9%	6.8%	6.5%	7.8%
Absolute size of price changes, model	*8.5%*	*8.5%*	*8.5%*	*8.5%*	*8.5%*
Serial correlation, data	$-.050$	$-.057$	$-.033$	$-.032$	$-.043$
Serial correlation, model	$-.050$	$-.050$	$-.050$	$-.050$	$-.050$
Within dispersion ratio, data	36.5%	54.6%	67.2%	72.4%	59.1%
Within dispersion ratio, model	*36.5%*	*54.5%*	*60.5%*	*63.5%*	*53.8%*

NOTE: We report moments predicted by the model in Section 5.4 in italics. We contrast them with the moments from the data presented in Table 1.

Table 4: Value of Information Capacity and the Number of Goods

	$N = 1$	$N = 2$	$N = 4$	$N = 8$
$\lambda(N)$	3.3348	3.3348	3.3348	3.3348
Absolute size of price changes	9.62%	9.60%	9.60%	9.60%
Serial correlation	-0.291	-0.291	-0.291	-0.291
Within-firm dispersion ratio	0.00%	50.12%	51.59%	51.58%
Cross-sectional variance	7.26%	7.25%	7.23%	7.25%
$\kappa_a(N)$	0.1935	0.2606	0.4429	0.6867
Cumulated price response	51.81%	53.48%	72.05%	82.70%
(rel. to frictionless prices)				
Loss	0.21%	0.20%	0.24%	0.21%

NOTE: We calibrate our model with homogeneous firms to moments for the whole sample of CPI data as we vary N. Firms' information processing capacity is calibrated such that its shadow price is invariant to N.

Table 5: Moments from the CPI and the Model, N=2

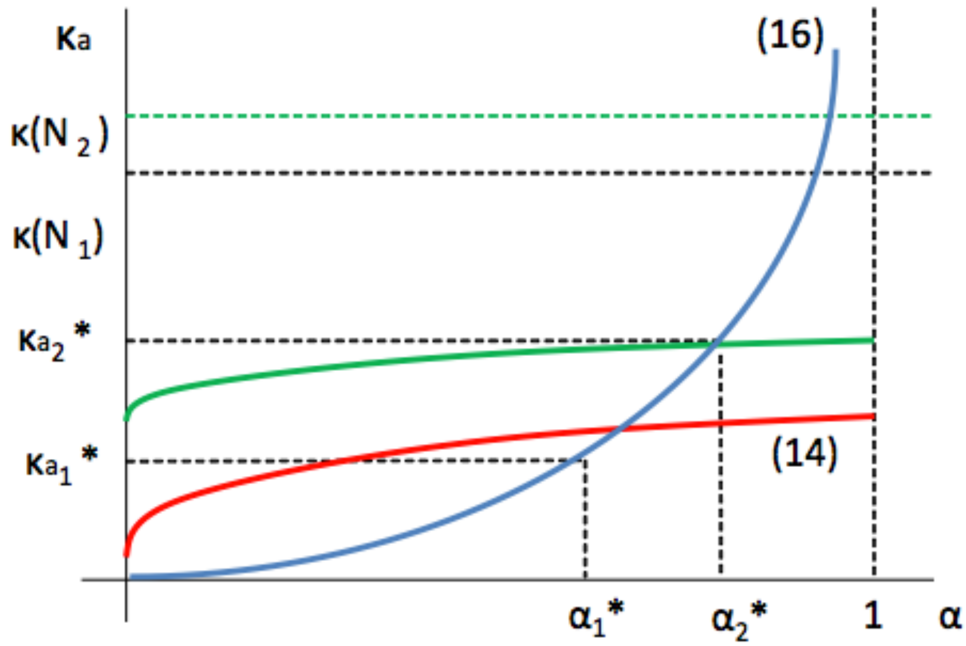
	data	$\kappa = 5$	$\kappa = 6$	$\kappa = 7$	$\kappa = 8$	$\kappa = 9$	$\kappa = 10$	$\kappa = 30$
Abs. size of price changes	9.6%	9.61%	9.65%	9.67%	9.70%	9.70%	9.73%	9.75%
Serial correlation	-0.29	-0.291	-0.290	-0.290	-0.290	-0.289	-0.288	-0.289
Within-firm var. ratio	51.6%	50.12%	50.04%	50.01%	50.01%	50.01%	50.04%	50.15%
Cross-sectional variance	2.65%	7.22%	7.28%	7.31%	7.32%	7.33%	7.34%	7.36%
$\kappa_a^*(2)$		0.219	0.309	0.473	0.676	0.920	1.212	8.123
Cumulated price response		51.67%	57.82%	71.97%	80.73%	86.14%	90.02%	97.98%
(rel. to frictionless prices)								

NOTE: As discussed in section 5.5, the table shows moments computed from the data and their counterparts generated by the model for N=2 using different values for firms' capacity to process information.

Table 6: Moments from the CPI and the Model, N=4

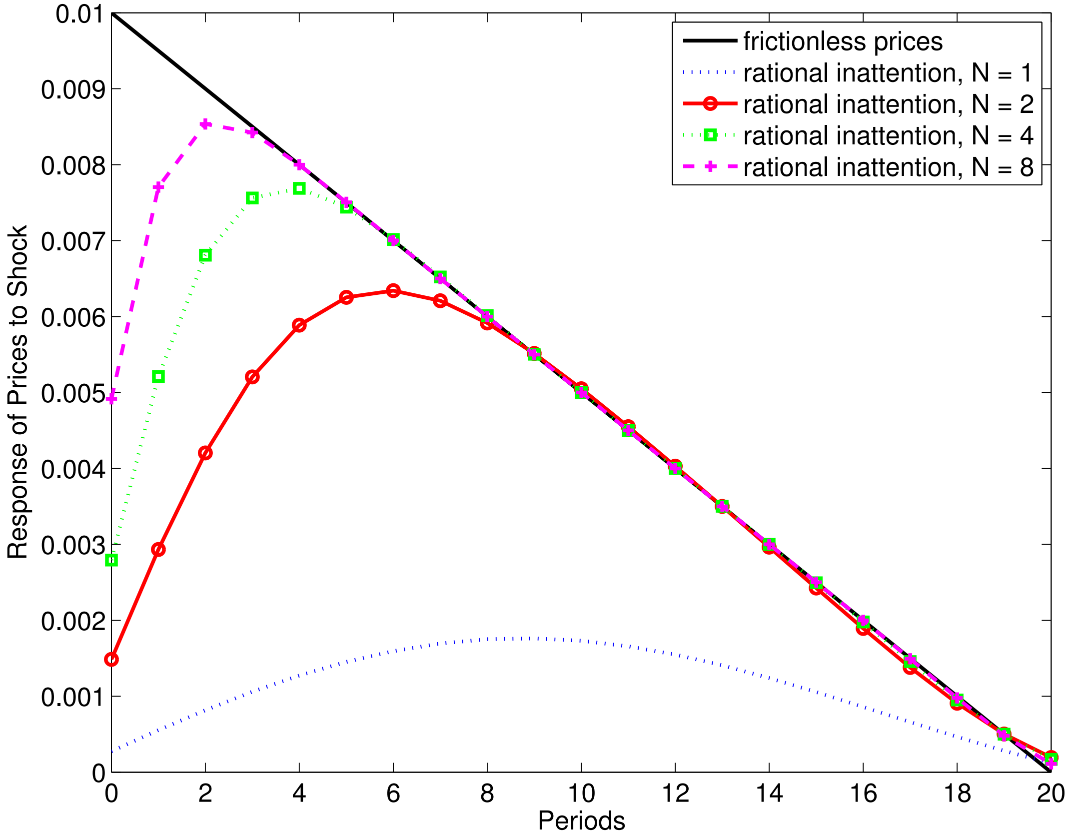
	data	$\kappa = 10$	$\kappa = 11$	$\kappa = 12$	$\kappa = 13$	$\kappa = 14$	$\kappa = 15$	$\kappa = 30$
Abs. size of price changes	9.60%	9.50%	9.54%	9.58%	9.60%	9.62%	9.66%	9.74%
Serial correlation	-0.291	-0.292	-0.2908	-0.291	-0.2911	-0.2901	-0.2895	-0.2893
Within-firm var. ratio	51.60%	50.99%	51.27%	51.53%	51.77%	51.85%	51.91%	52.11%
Cross-sectional variance	2.65%	7.16%	7.21%	7.24%	7.25%	7.28%	7.29%	7.35%
$\kappa_a^*(4)$		0.31	0.37	0.44	0.52	0.62	0.72	3.31
Cumulated price response		60.17%	64.74%	70.50%	75.32%	79.33%	82.60%	98.46%
(rel. to frictionless prices)								

NOTE: As discussed in section 5.5, the table shows moments computed from the data and their counterparts generated by the model for N=2 using different values for firms' capacity to process information.



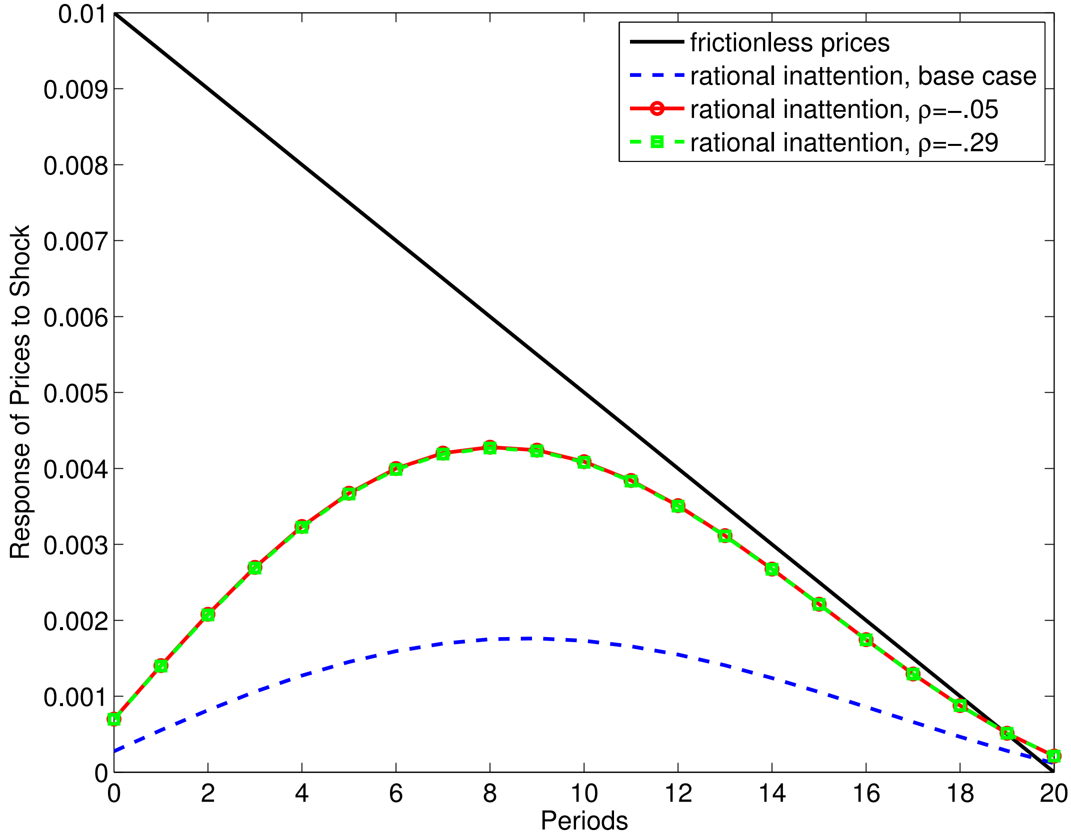
NOTE: The figure illustrates the fixed point problem of attention allocation given by equations (14) and (16). Equation (14) is drawn in red, while equation (16) is drawn in blue. Equation (16) is invariant to N, but N affects the drift and slope of equation (14). Under conditions described in Proposition 2 the drift of equation (14) is increasing in N. An upwards shift of this function is represented in green.

Figure 1: Equations (6) and (7) in the space (α, κ_a)



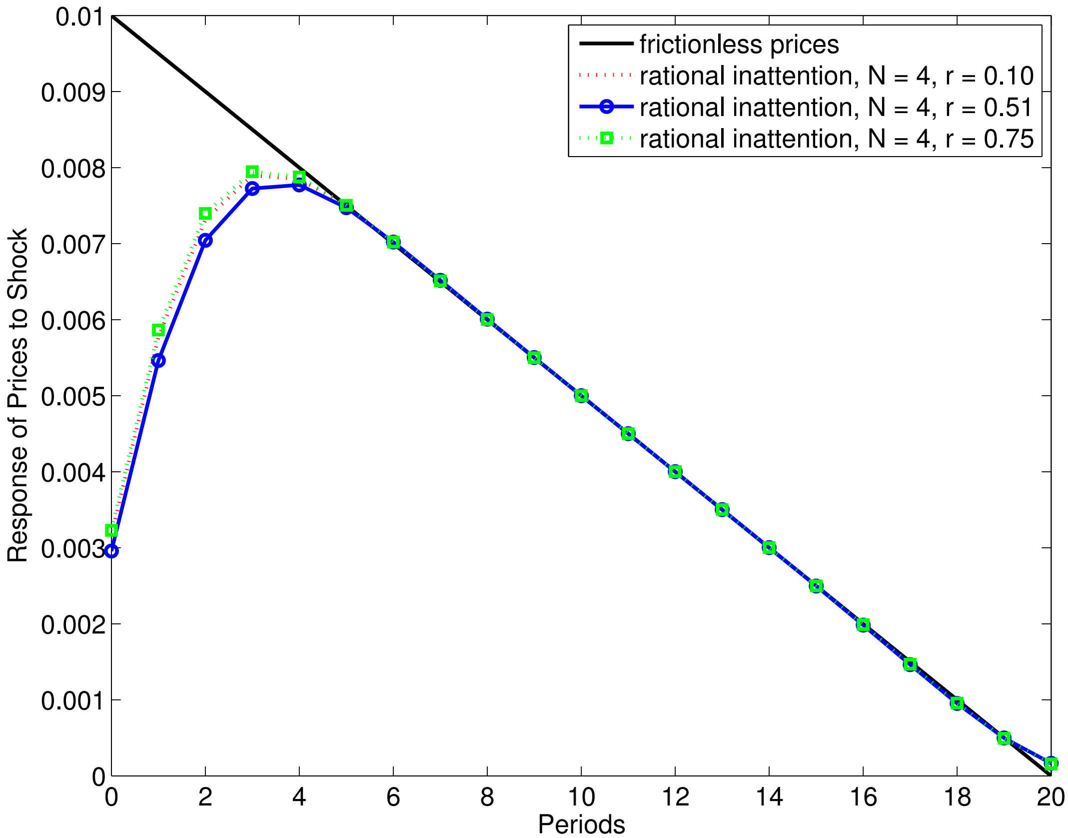
NOTE: We illustrate the response of prices to a 1% monetary shock as we vary N in our model calibrated to moments from the CPI data. The black line is for frictionless prices, the dashed blue line is for the benchmark of rationally inattentive prices with N=1, the red line with circles is for rationally inattentive prices with N=2, the dashed green line with squares is for rationally inattentive prices with N=4, and the dashed magenta line with dots is for is for rationally inattentive prices with N=8. The response of prices quickly becomes closer to that of frictionless prices as N increases. Details are given in sections 5.1 and 5.2.

Figure 2: Response of Prices to a 1% Impulse in q_t for Section 5.1



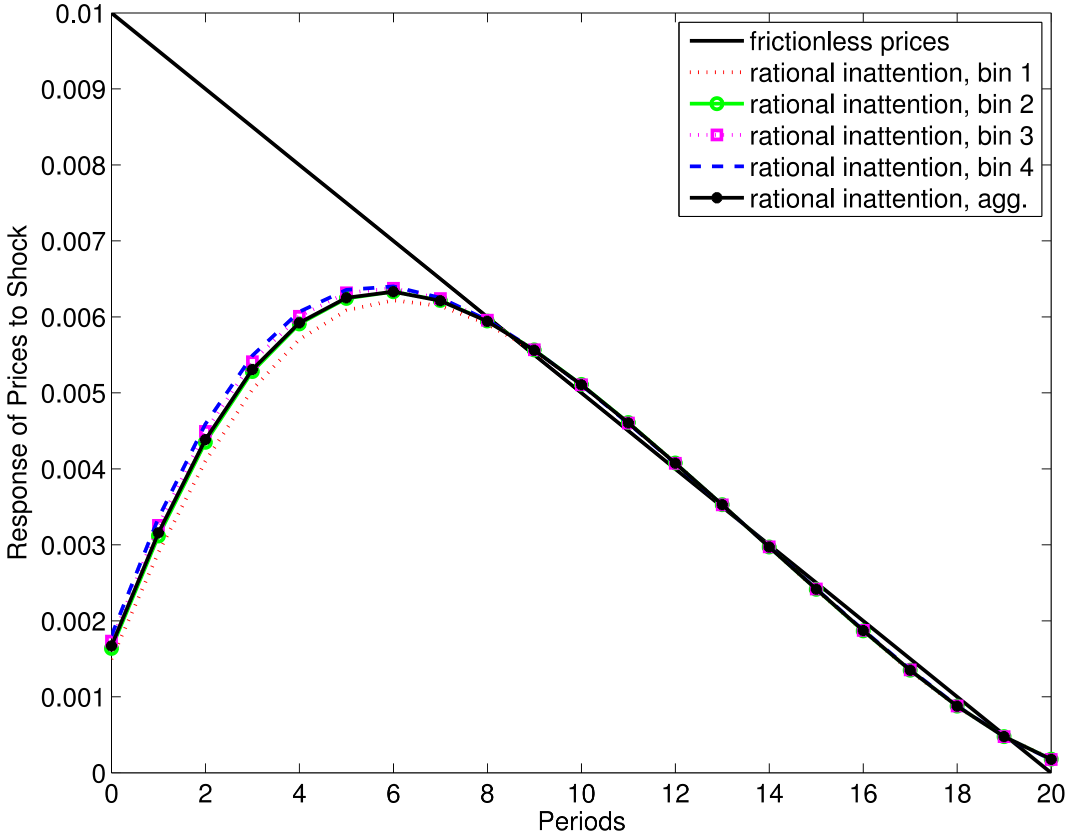
NOTE: We illustrate the response of prices to a 1% monetary shock as we vary the persistence of idiosyncratic shocks in our model calibrated to moments from the CPI data. The black line is for frictionless prices, the dashed blue line is for our benchmark with highly persistent idiosyncratic shocks, the red line with circles is for rationally inattentive prices that have serial correlation of -0.05, the dashed green line with squares is for rationally inattentive prices that have serial correlation of -0.29. Section 5.2 contains further details.

Figure 3: Response of Prices to a 1% Impulse in q_t for Section 5.2



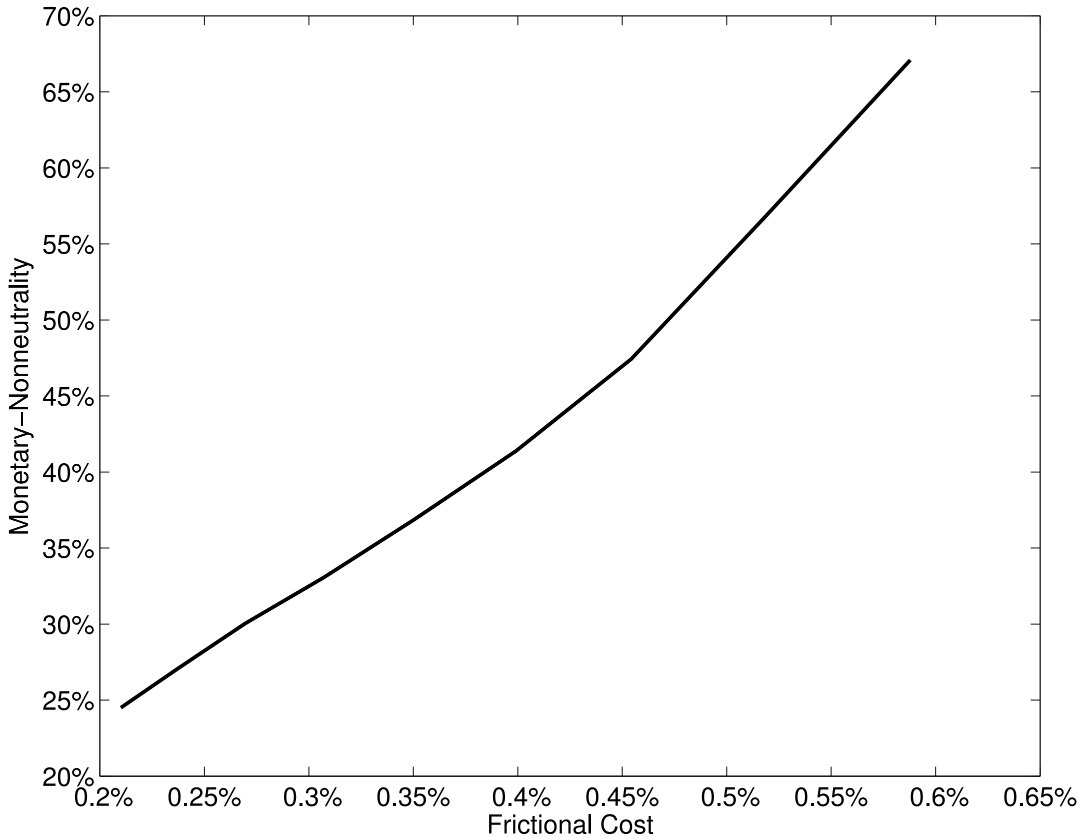
NOTE: We illustrate the response of prices to a 1% monetary shock for $N = 4$ as we vary the extent of within-firm log non-zero price dispersion. The blue line with circles denotes the impulse response for a 51.6% within-firm dispersion ratio, the red doted line the response for a 10% ratio, and the green line with squares the response for a 75% ratio. The black line is for frictionless prices. Section 5.3 contains further details.

Figure 4: Impulse Response of Prices under Differing Within-Firm Dispersion for Section 5.2



NOTE: We illustrate the response of prices to a 1% monetary shock as we vary N in our model calibrated to moments of the PPI data by bins. The black line is for frictionless prices, the dashed blue line is for rationally inattentive prices in bin 1, the red line with circles is for rationally inattentive prices in bin 2, the dashed green line with squares is for rationally inattentive prices in bin 3, and the dashed magenta line with dots is for is for rationally inattentive prices in bin 4, and the black solid line with dots is for aggregate rationally inattentive prices. Section 5.4 contains further details.

Figure 5: Response of Prices to a 1% Impulse in q_t for Section 5.3



NOTE: We illustrate the relationship between monetary non-neutrality, measured as the cumulative response of rational inattentive prices relative to frictionless prices, and the frictional cost of as we vary firms' information processing capacity in our model calibrated to moments of the PPI data. Section 5.5 contains further details.

Figure 6: Trade-Off between Monetary Non-Neutrality and Frictional Cost

Appendix A. The Problem of the Firm in Our Analytical Model

Appendix A.1. Full Model

Consider an economy with a continuum of goods of total measure one indexed by $j \in [0,1]$, and a continuum of monopolist firms with total measure $\frac{1}{N}$ indexed by $i \in \left[0, \frac{1}{N}\right]$ for $N \in \mathbb{N}$. Each firm i prices N goods which are randomly drawn without replacement from the set of goods. Denote by $\aleph_i$ the set that collects the identity of the N goods produced by firm i.

Firms are meant to be pricing decision units, or "price setters", so we focus on modelling firms' pricing decisions.[20] Each good j contributes to the total profits of its price setter according to

$$\pi \left(P_{jt}, P_t, Y_t, F_{it}, Z_{jt}\right), \tag{A.1}$$

where P_{jt} is the price of good j, P_t is the aggregate price, Y_t is real aggregate demand, and F_{it} and Z_{jt} are two idiosyncratic, exogenous random variables, the former specific to firm i and the latter specific to good j, all at time t. The function $\pi(\cdot)$ is assumed to be independent of which and how many goods the firm prices, twice continuously differentiable and homogenous of degree zero in the first two arguments. Idiosyncratic variables F_{it} and Z_{jt} satisfy

$$\int_0^{\frac{1}{N}} f_{it} di = 0, \tag{A.2}$$

$$\int_0^1 z_{jt} dj = 0, \tag{A.3}$$

where small case notation generically denotes log-deviations from steady-state levels. Hence, f_{it} and z_{jt} have direct interpretation as firm- and good-specific shocks.

Nominal aggregate demand Q_t is assumed to be exogenous and stochastic satisfying

$$Q_t = P_t Y_t, \tag{A.4}$$

where aggregate prices follow from

$$p_t = \int_0^1 p_{jt} dj. \tag{A.5}$$

[20] We use "firms" and "price setters" indistinguishably through the paper.

The total period profit function of price setter i is

$$\sum_{n\in\aleph_i} \pi\left(P_{nt}, P_t, Y_t, F_{it}, Z_{nt}\right),$$

which sums up the contribution to profits of all goods priced by firm i.

The key assumption of rational inattention models is that price setters are constrained in the "flow of information" that they can process at every period t:

$$I\left(\left\{Q_t, F_{it}, \{Z_{nt}\}_{n\in\aleph_i}\right\}, \{s_{it}\}\right) \leq \kappa\left(N\right)$$

where $Q_t, F_{it}, \{Z_{nt}\}_{n\in\aleph_i}$ are variables of interest for firm i that are not directly observable, s_{it} is the vector of signals that firm i actually observes, the function $I\left(\cdot\right)$ measures the information flow between observed signals and variables of interest, and $\kappa\left(N\right)$ is an exogenous, limited capacity that without loss of generality is assumed to depend on the number N of goods.

The information flow $I\left(\cdot\right)$ is a measure the informational containt of an observable variable respect to an unobservable variable.[21] This measure has been proposed by Shannon (1948) and does not need to be specified here except for computational purposes, so we relegate it to the appendix.

We also assume that the vector of signals s_{it} may be partitioned into $N+1$ subvectors $\left\{s_{it}^a, s_{it}^f, \{s_{nt}^z\}_{n\in\aleph_i}\right\}$. Each subvector is correlated to one target variable such that $\{q_t, s_{it}^a\}$, $\left\{f_{it}, s_{it}^f\right\}$ and $\{z_{nt}, s_{nt}^z\}_{n\in\aleph_i}$. Besides, we assume that all variables are Gaussian, jointly stationary and there exists an initial infinite history of signals. Under these assumptions the information flow is additively separable:

$$I\left(\left\{Q_t, F_{it}, \{Z_{nt}\}_{n\in\aleph_i}\right\}, \{s_{it}\}\right) = I\left(\{Q_t\}, \{s_{it}^a\}\right) + I\left(\{F_{it}\}, \left\{s_{it}^f\right\}\right) + \sum_{n\in\aleph_i} I\left(\{Z_{n\tau}\}, s_{nt}^z\right).$$

[21] To provide intuition, if one denotes as U_t and O_t respectively as arbitrary unobservable and observable Gaussian i.i.d. random variables, the information flow between them is

$$I\left(\{U_t\}, \{O_t\}\right) = \frac{1}{2}\log_2\left(\frac{1}{1-\rho_{U,O}^2}\right), \tag{A.6}$$

which is increasing in $|\rho_{U,O}|$, the absolute correlation between U_t and O_t. Hence, a given information flow pins down the observation noise of U_t respect to O_t.

Hence, the problem of the firm i may be represented as

$$\max_{\{s_{it}\}\in\Gamma} \mathbb{E}_{i0}\left[\sum_{t}^{\infty}\beta^{t}\left\{\sum_{n\in\aleph_i}\pi\left(P_{nt}^{*},P_t,Y_t,F_{it},Z_{nt}\right)\right\}\right] \tag{A.7}$$

with

$$P_{nt}^{*}=\arg\max_{P_{nt}}\mathbb{E}\left[\pi\left(P_{nt},P_t,Y_t,F_{it},Z_{nt}\right)\mid s_{it}\right] \tag{A.8}$$

and subject to

$$\kappa_a+\kappa_f+\sum_{n\in\aleph_i}\kappa_n\leq\kappa\left(N\right). \tag{A.9}$$

where κ_a, κ_f and κ_n respectively denote $I\left(\{Q_t\},\{s_{it}^{a}\}\right)$, $I\left(\{F_t\},\left\{s_{it}^{f}\right\}\right)$ and $I\left(\{Z_{nt}\},\{s_{nt}^{z}\}\right)$. Since prices are flexible, the pricing problem in (A.8) is static. The firm, however, must consider its whole discounted expected stream of profits to choose signals in the set Γ of signals that satisfy the above assumptions. The constraint in (A.9) implies a trade-off for the firm: Increasing the precision of signals about, say, Q_t, forces it to decrease the precision of signals about F_t and $\{Z_{nt}\}_{n\in\aleph_i}$.

The equilibrium in this economy is defined as follows:

Definition 2. *An equilibrium is a collection of signals $\{s_{it}\}$, prices $\{P_{jt}\}$, the aggregate price level $\{P_t\}$ and real aggregate demand $\{Y_t\}$ such that*

1. *Given $\{P_t\},\{Y_t\},\{F_{it}\}_{i\in\left[0,\frac{1}{N}\right]}$ and $\{Z_{jt}\}_{j\in[0,1]}$, all firms $i\in\left[0,\frac{1}{N}\right]$ choose the stochastic process of signals $\{s_{it}\}$ at $t=0$ and the price of goods they produce, $\{P_{nt}\}_{n\in\aleph_i}$ for $t\geq 1$.*
2. *$\{P_t\}$ and $\{Y_t\}$ are consistent with equations (A.4) and (A.5) for $t\geq 1$.*

We now refer to some simplifying assumptions in our model. First, firms' size scales up with the number N of goods they price. This is broadly consistent with empirical evidence from US data (Bernard et al. (2010); Bhattarai and Schoenle (2014)). Section 3 discusses some implications of this assumption. Second, contribution to profits $\pi\left(\cdot\right)$ of goods priced by a firm are independent of each other. Appendix C relaxes this assumption to find no qualitative differences in results. Third, signals are informative only about one type of shock. Also discussed in Appendix C, this assumption helps the model to capture the within-firms dispersion of log price changes we document for US data (reported in Section 4). Fourth, we take the number N of goods firms price and firms' information processing capacity $\kappa\left(N\right)$ as exogenous. In Section 3 we impose assumptions on $\kappa\left(N\right)$ to make our results consistent with those in a model where N and $\kappa\left(N\right)$ are firms' choices.

White-Noise Shocks

We start by computing the frictionless non-stochastic steady state in this economy. Let $\bar{Q}$, $\bar{F}_i = \bar{F}\ \forall i$ and $\bar{Z}_j = \bar{Z}\ \forall j$ be the steady state level of these variables. Without frictions, it must hold that

$$\pi_1\left(1, 1, Y_t, \bar{F}, \bar{Z}\right) = 0,$$

which follows from the optimality of prices. This equation solves for the steady-state level of real aggregate demand $\bar{Y}$, and equation (A.4) for the steady-state aggregate price level $\bar{P} = \bar{Q}/\bar{Y}$.

A second-order approximation of the problem of firm i around this steady-state is

$$\max_{\{p_{nt}\}_{n\in\aleph_i}} \sum_{n\in\aleph_i} \left\{ \begin{array}{c} \widehat{\pi}_1 p_{nt} + \frac{\widehat{\pi}_{11}}{2} p_{nt}^2 + \widehat{\pi}_{12} p_{nt} p_t + \widehat{\pi}_{13} p_{nt} y_t + \widehat{\pi}_{14} p_{nt} f_{it} + \widehat{\pi}_{15} p_{nt} z_{nt} \\ + \textit{terms independent of } p_{nt}. \end{array} \right\}$$

with $\widehat{\pi}_1 = 0$, $\widehat{\pi}_{11} < 0$, $\widehat{\pi}_{12} = -\widehat{\pi}_{11}$ and all parameters identical for all goods and all firms.

The optimal frictionless pricing rule for each good $n \in \aleph_i$ for all i is

$$p_{nt}^{\diamondsuit} = p_t + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} y_t + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} f_{it} + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} z_{nt} \equiv \Delta_t + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} f_{it} + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} z_{nt} \tag{A.10}$$

where the compound variable Δ_t collects aggregate variables.

Since this is a linear pricing rule, the optimal price of good $n \in \aleph_i$ of an arbitrary firm i that solves (A.8) is

$$p_{nt}^{*} = \mathbb{E}\left[\Delta_t \mid s_{it}^{a}\right] + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} \mathbb{E}\left[f_{it} \mid s_{nt}^{f}\right] + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} \mathbb{E}\left[z_{jt} \mid s_{nt}^{z}\right]. \tag{A.11}$$

given the signal structure $\left\{s_{it}^{a}, s_{it}^{f}, s_{nt}^{z}\right\}$. We must solve now for firms' optimal choice of signals. To do so, we recast the firms' problem up to second order as the minimization the discounted sum of firms' expected loss in profits due to the friction (the "frictional costs" hereafter) for all goods produced by the firm:

$$\sum_{t=1}^{\infty} \beta^t \sum_{n\in\aleph_i} \left\{ \frac{|\widehat{\pi}_{11}|}{2} \mathbb{E}\left[\left(p_{nt}^{\diamondsuit} - p_{nt}^{*}\right)^2\right] \right\} \tag{A.12}$$

We assume now that shocks q_t, f_{it} and z_{jt} are white noise, with variances σ_q^2, σ_f^2 for any firm $i \in \left[0, \frac{1}{N}\right]$ and σ_z^2 for any good $j \in [0, 1]$. This assumption allows us to obtain analytical solution.[22]

Given this assumption, we guess that the log-deviation of aggregate prices respond linearly to a monetary shock, $p_t = \alpha q_t$, so the compound aggregate variable Δ_t is given by

$$\Delta_t = \left[\alpha + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\left(1-\alpha\right)\right] q_t. \tag{A.13}$$

In addition, signals chosen by firm $i \in \left[0, \frac{1}{N}\right]$ are restricted to have the structure

$$\begin{aligned} s_{it}^a &= \Delta_t + \varepsilon_{it}, \\ s_{it}^f &= f_{it} + e_{it}, \\ s_{nt}^z &= z_{nt} + \psi_{nt}, \end{aligned}$$

where $\sigma_{\varepsilon i}^2$, σ_{ei}^2 and $\sigma_{\psi n}^2$ are the variance of noise ε_{it}, e_{it} and ψ_{nt}.[23]

Therefore, given signals $\left\{s_{it}^a, s_{it}^f, s_{nt}^z\right\}$, the optimal pricing rule (A.11) solves as

$$p_{nt}^* = \frac{\sigma_\Delta^2}{\sigma_\Delta^2 + \sigma_{\varepsilon i}^2} s_{it}^a + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\frac{\sigma_f^2}{\sigma_f^2 + \sigma_{ei}^2} s_{it}^f + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}\frac{\sigma_z^2}{\sigma_z^2 + \sigma_{\psi n}^2} s_{nt}^z.$$

Replacing $p_{nt}^{\diamondsuit}$ and p_{nt}^* in (A.12) and using the functional form of information flow in (A.6) because shocks are Gaussian white noise, the problem of the firm becomes

$$\min_{\kappa_a, \kappa_f, \{\kappa_n\}_{n\in\aleph_i}} \frac{\beta}{1-\beta}\frac{|\widehat{\pi}_{11}|}{2}\left[2^{-2\kappa_a}\sigma_\Delta^2 N + \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_f}\sigma_f^2 N + \left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2 \sum_{n\in\aleph_i} 2^{-2\kappa_n}\sigma_z^2\right] \tag{A.14}$$

subject to

$$\kappa_a + \kappa_f + \sum_{n\in\aleph_i}\kappa_n \leq \kappa\left(N\right). \tag{A.15}$$

[22] The appendix relaxes this assumption and presents the numerical algorithm used to solve for it. We use this general problem to obtain our quantitative results in Section 5.

[23] Mackowiak and Wiederholt (2009) show that this structure of signals is optimal. This result is not affected by the modifications to their model introduced here.

where

$$\begin{aligned}\kappa_a &\equiv \frac{1}{2}\log_2\left(\frac{\sigma_\Delta^2}{\sigma_{\varepsilon i}^2}+1\right);\\ \kappa_f &\equiv \frac{1}{2}\log_2\left(\frac{\sigma_f^2}{\sigma_{ei}^2}+1\right);\\ \kappa_n &= \log_2\left(\frac{\sigma_z^2}{\sigma_{\psi n}^2}+1\right).\end{aligned}$$

which gives the representation presented in Section 2.

Persistent Shocks

We now solve for a simplified version of our model that allows for persistent shocks and keeps at least partial closed solution. Assume that the process of q_t is such that Δ_t is $AR(1)$ with persistence ρ_Δ. Idiosyncratic shocks f_{it} and z_{jt} are also $AR(1)$ respectively with persistence ρ_f for all i and ρ_z for all j. The starting guess is now

$$p_t = \sum_{l=0}^{\infty} \alpha_l v_{t-l}, \tag{A.16}$$

where $\{v_{t-l}\}_{l=0}^{\infty}$ is the history of nominal aggregate demand innovations.

The firms' problem may be cast in two stages. In the first stage, firms choose

$$\begin{aligned}&\min_{\hat{\Delta}_{it},\{\hat{z}_{nt}\}_{n\in\aleph_i}} \sum_{n\in\aleph_i}\left\{\sum_{t=1}^{\infty}\beta^t\frac{|\widehat{\pi}_{11}|}{2}\mathbb{E}\left[\left(p_{nt}^{\diamondsuit}-p_{nt}^*\right)^2\right]\right\}\\ \to\ &\min_{\hat{\Delta}_{it},\{\hat{z}_{nt}\}_{n\in\aleph_i}} \frac{\beta}{1-\beta}\frac{|\widehat{\pi}_{11}|}{2}\left\{\begin{array}{c}\mathbb{E}\left(\Delta_t-\hat{\Delta}_{it}\right)^2 N+\left(\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\right)^2\mathbb{E}\left(f_{it}-\hat{f}_{it}\right)^2 N\\ +\left(\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}\right)^2\sum_{n\in\aleph_i}\mathbb{E}\left(z_{nt}-\hat{z}_{nt}\right)^2\end{array}\right\}\end{aligned}$$

subject to

$$\begin{aligned}I\left(\left\{\Delta_t,\hat{\Delta}_{it}\right\}\right) &\le \kappa_a,\\ I\left(\left\{f_{it},\hat{f}_{it}\right\}\right) &\le \kappa_f,\\ I\left(\{z_{nt},\hat{z}_{nt}\}\right) &\le \kappa_n,\quad for\ n\in\aleph_i\\ \kappa_a+\sum_{n\in\aleph_{ie}}\kappa_n &\le \kappa(N)\end{aligned}$$

For the second stage, firms choose the signals that deliver $\hat{\Delta}_{it}^*, \{\hat{z}_{nt}^*\}_{n\in\aleph_{we}}$. As in Appendix

B, we omit this stage. Our representation for the firm's problem follows from a result in Mackowiak and Wiederholt (2009): The solution of

$$\min_{b,c} \mathbb{E}\left(U_t - O_t\right)^2$$

where U_t is an unobservable and O_t is an observable variable, subject to

$$\begin{aligned} U_t &= \rho U_{t-1} + a u_t, \\ O_t &= \sum_{l=0}^{\infty} b_l u_{t-l} + \sum_{l=0}^{\infty} c_l \varepsilon_{t-l}, \\ \kappa &\geq I\left(\{U_t, O_t\}\right) \end{aligned}$$

yields

$$\mathbb{E}\left(U_t - O_t^*\right)^2 = \sigma_T^2 \frac{1-\rho^2}{2^{2\kappa}-\rho^2}.$$

Therefore, firms' problem may be represented as

$$\min_{\kappa_a,\kappa_f,\{\kappa_n\}_{n\in\aleph_i}} \frac{\beta}{1-\beta}\frac{|\widehat{\pi}_{11}|}{2}\left[\frac{1-\rho_\Delta^2}{2^{2\kappa_a}-\rho_\Delta^2}N\sigma_\Delta^2 + \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 \frac{1-\rho_f^2}{2^{2\kappa_f}-\rho_f^2}N\sigma_f^2 + \left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2 \sum_{n\in\aleph_i}\frac{1-\rho_z^2}{2^{2\kappa_n}-\rho_z^2}\sigma_z^2\right]$$

subject to

$$\kappa_a + \kappa_f + \sum_{n\in\aleph_i}\kappa_n \leq \kappa(N).$$

This problem is identical to that solved in Section 3 for $\rho_\Delta = \rho_z = 0$. Its first order conditions are

$$\begin{aligned} \kappa_a^* + f\left(\rho_\Delta, \kappa_a^*\right) &= \kappa_f^* + f\left(\rho_f, \kappa_f^*\right) + \log_2 \tilde{x}_1 \\ \kappa_a^* + f\left(\rho_\Delta, \kappa_a^*\right) &= \kappa_z^* + f\left(\rho_z, \kappa_z^*\right) + \log_2 \tilde{x}_2\sqrt{N} \end{aligned}$$

where $\tilde{x}_1 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta\sqrt{1-\rho_\Delta^2}}{\widehat{\pi}_{14}\sigma_f\sqrt{1-\rho_f^2}}$, $\tilde{x}_2 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta\sqrt{1-\rho_\Delta^2}}{\widehat{\pi}_{15}\sigma_z\sqrt{1-\rho_z^2}}$ and $f\left(\rho,\kappa\right) = \log_2\left(1-\rho^2 2^{-2\kappa}\right)$.

The function $f\left(\rho,\kappa\right)$ is weakly negative and increasing in κ, so the difference in attention to aggregate and good-specific shocks, $\kappa_a^* - \kappa_z^*$, is still increasing in N. As before, the difference $\kappa_a^* - \kappa_f^*$ remains invariant to N. The function $f\left(\rho,\kappa\right)$ is also decreasing in $|\rho|$. Hence, a decrease in persistence of idiosyncratic shocks ρ_f and ρ_z implies an increase of κ_a^* relative to κ_f^* and κ_z^* if $\sigma_z\sqrt{1-\rho_z^2}$ and $\sigma_f\sqrt{1-\rho_f^2}$ are held constant.

Appendix B. The Problem of the Firm for a General Structure of Shocks

This appendix displays the analytical representation of firms' problem in the setup of Section 5 and explains the numerical algorithm applied to solve it. This appendix adapts to our setup a similar presentation by Mackowiak and Wiederholt (2009). Assume that firms are exposed to three types of shocks:

$$q_t = \sum_{l=0}^{\infty} a_l v_{t-l},$$

$$f_{it} = \sum_{l=0}^{\infty} b_l \xi_{t-l},$$

$$z_{jt} = \sum_{l=0}^{\infty} c_l \zeta_{t-l},$$

where q_t is a nominal aggregate demand shock (interpreted as a "monetary" shock), f_{it} is a shock idiosyncratic to each firm $i \in \left[0, \frac{1}{N}\right]$, z_{jt} is a shock idiosyncratic to each good $j \in [0,1]$, and $\{v_{t-l}, \xi_{t-l}, \zeta_{t-l}\}_{l=0}^{\infty}$ are innovations following Gaussian independent processes.

We guess that the log-deviation of aggregate prices follows

$$p_t = \sum_{l=0}^{\infty} \alpha_l v_{t-l}$$

which, given the definition of Δ_t in (A.10) and $y_t = q_t - p_t$, implies

$$\Delta_t = \left(1 - \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\right) \sum_{l=0}^{\infty} \alpha_l v_{t-l} + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} \sum_{l=0}^{\infty} a_l v_{t-l} \equiv \sum_{l=0}^{\infty} d_l v_{t-l} \tag{B.1}$$

The problem of firm $i \in \left[0, \frac{1}{N}\right]$ has two stages. In the first stage, firms must choose conditional expectations for Δ_t , f_{it} and $\{z_{nt}\}_{n\in\aleph_i}$ to minimize the deviation of prices with respect to frictionless optimal prices subject to the information capacity constraint:

$$\min_{\hat{\Delta}_{it}, \{\hat{z}_{nt}\}_{n\in\aleph_i}} \sum_{n\in\aleph_i} \left\{ \sum_{t=1}^{\infty} \beta^t \frac{|\widehat{\pi}_{11}|}{2} \mathbb{E}\left[\left(p_{nt}^{\diamondsuit} - p_{nt}^{*}\right)^2 \right] \right\}$$

which is equivalent to

$$\min_{\hat{\Delta}_{it},\widehat{f}_{it},\{\hat{z}_{nt}\}_{n\in\aleph_i}}\left\{\begin{array}{c}\mathbb{E}\left[\left(\Delta_t-\widehat{\Delta}_{it}\right)^2\right]N+\left(\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\right)^2\mathbb{E}\left[\left(f_{it}-\widehat{f}_{it}\right)^2\right]N\\+\left(\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}\right)^2\sum_{n\in\aleph_i}\mathbb{E}\left[\left(z_{nt}-\widehat{z}_{nt}\right)^2\right]\end{array}\right\}$$

subject to the process of Δ_t, f_{it} and $\{z_{nt}\}_{n\in\aleph_i}$ and the information capacity constraint

$$I\left(\Delta_t,\widehat{\Delta}_{it}\right)+I\left(f_{it},\widehat{f}_{it}\right)+\sum_{n\in\aleph_{we}}I\left(z_{nt},\widehat{z}_{nt}\right)\leq\kappa\left(N\right).$$

The function $I(\cdot)$ is the information flow. For instance, this function for Δ_t takes the form:

$$I\left(\Delta_t,\widehat{\Delta}_{it}\right)\equiv-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2\left[1-C_{\Delta_t,\widehat{\Delta}_{it}}\left(\omega\right)\right]d\omega$$

where $C_{\Delta_t,\widehat{\Delta}_{it}}(\omega)$ is called coherence function, which is defined as follows. Let us describe the conditional expectations $\widehat{\Delta}_{it}$ as

$$\widehat{\Delta}_{it}=\sum_{l=0}^{\infty}g_l v_{t-l}+\sum_{l=0}^{\infty}h_l\varepsilon_{t-l},$$

then

$$C_{\Delta,\widehat{\Delta}_{we}}\left(\omega\right)\equiv\frac{\frac{G\left(e^{-i\omega}\right)G\left(e^{i\omega}\right)}{H\left(e^{-i\omega}\right)H\left(e^{i\omega}\right)}}{\frac{G\left(e^{-i\omega}\right)G\left(e^{i\omega}\right)}{H\left(e^{-i\omega}\right)H\left(e^{i\omega}\right)}+1},$$

where $G\left(e^{i\omega}\right)=g_0+g_1e^{i\omega}+g_2e^{i2\omega}+...$ and $H\left(e^{i\omega}\right)=h_0+h_1e^{i\omega}+h_2e^{i2\omega}+...$

If the conditional expectations $\widehat{f}_{it}$ and $\{\widehat{z}_{nt}\}_{n\in\aleph_i}$ are described by

$$\begin{aligned}\widehat{f}_{it}^{*}&=\sum_{l=0}^{\infty}r_l\xi_{t-l}+\sum_{l=0}^{\infty}s_l\varepsilon_{t-l},\\\widehat{z}_{nt}^{*}&=\sum_{l=0}^{\infty}w_{nl}\zeta_{t-l}+\sum_{l=0}^{\infty}x_{nl}e_{nt-l}\text{ for }n\in\aleph_i.\end{aligned}$$

Then the problem may be represented as

$$\min_{g,h,r,s,\{w_n,x_n\}_{n\in\aleph_i}}\left\{\begin{array}{c}\left[\sum_{l=0}^{\infty}\left(d_l-g_l\right)^2+\sum_{l=0}^{\infty}h_l^2\right]N+\left(\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\right)^2N\left[\sum_{l=0}^{\infty}\left(b_l-r_l\right)^2+\sum_{l=0}^{\infty}s_l^2\right]\\+\left(\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}\right)^2\sum_{n\in\aleph_{we}}\left[\sum_{l=0}^{\infty}\left(c_l-w_{nl}\right)^2+\sum_{l=0}^{\infty}x_{nl}^2\right]\end{array}\right\}$$

$$s.t.\;\; I\left(\Delta_t, \widehat{\Delta}_{it}\right) + I\left(f_{it}, \widehat{f}_{it}\right) + \sum_{n \in \aleph_{we}} I\left(z_{nt}, \widehat{z}_{nt}\right) \leq \kappa\left(N\right)$$

where $g, h, r, s, \{w_n, x_n\}_{n\in\aleph_i}$ represent vectors of coefficients. The first order conditions for g and h are

$$\begin{aligned} g_l &: 2\left(d_l^* - g_l^*\right) N = -\frac{\mu_a}{4\pi \log\left(2\right)} \int_{-\pi}^{\pi} \frac{\partial \log\left[1 - C_{\Delta,\widehat{\Delta}_{we}^*}\left(\omega\right)\right]}{\partial g_l} d\omega, \\ h_l &: 2h_l^* N = \frac{\mu_a}{4\pi \log\left(2\right)} \int_{-\pi}^{\pi} \frac{\partial \log\left[1 - C_{\Delta,\widehat{\Delta}_{we}^*}\left(\omega\right)\right]}{\partial h_l} d\omega \end{aligned}$$

where μ_a is the Lagrangian multiplier. Similar conditions must be satisfied by r^* and s^* and by $\{w_n^*, x_n^*\}_{n\in\aleph_i}$ but without N.

The second stage of the problem is to obtain optimal signals structures that deliver $\widehat{\Delta}_{it}^* = \widehat{\Delta}_{it}\left(\kappa_a^*\left(N\right), N\right)$ and $\widehat{z}_{nt}^* = \widehat{z}_{nt}\left(\kappa_n\left(N\right), N\right)$. Since we are interested in the aggregate implications of the model, we do not solve this part.

Numerically, we truncate the memory of all processes to 20 lags, which is the same order assumed for the MA process for q_t. Then we start from a guess for α to compute d, we find $g^*, h^*, r^*, s^*, \{w_n, x_n\}_{n\in\aleph_i}$ by using the Levenberg-Marquardt algorithm to solve the system of first-order conditions plus the information flow constraint after imposing symmetry in $\{w_n, x_n\}_{n\in\aleph_i}$. With these vectors, we compute $I\left(\Delta_t, \widehat{\Delta}_{it}\right) = \kappa_a^*\left(N\right)$, $I\left(f_{it}, \widehat{f}_{it}\right) = \kappa_f^*\left(N\right)$ and $I\left(z_{nt}, \widehat{z}_{nt}\right) = \kappa_z^*(N)$ and the vector α. We use this α as guess for a new iteration upon convergence on α.

Appendix C. Extensions

This appendix relaxes some expositional assumptions made in the set up studied in the main text. These extensions yield no substantive changes to our conclusions or counterfactual predictions.

Common Signals

In the main text we have assumed that there exists an independent signal for each good-specific idiosyncratic shock. We relax this assumption and instead we assume that there exists a signal

$$s_{it}^z = \sum_{n\in\aleph_i} z_{nt} + \psi_{it}.$$

In words, firms receive only one common signal regarding all its good-specific shocks. Under this assumption, we are in the same situation as in Proposition 1, where firms' attention to aggregate shocks is inviariant in the number of goods that this firm produces, but its prices perfectly comove. This latter result is clear from observing the form of optimal prices under rational inattention:

$$p^*_{nt} = \frac{\sigma^2_\Delta}{\sigma^2_\Delta + \sigma^2_{\varepsilon i}} s^a_{it} + \frac{\sigma^2_f}{\sigma^2_f + \sigma^2_{ei}} s^f_{it} + \frac{\sigma^2_z}{N\sigma^2_z + \sigma^2_{\psi i}} s^z_{it}$$

which only responds to aggregate and firm-specific components.

Interdependent Profits

We now depart from our assumption in the main text that firms' pricing decisions are independent of each other. We model the 'interdependence' in pricing decisions within firms by assuming that the contribution to profits of a given good $n \in \aleph_i$ to its pricing firm i is now

$$\pi \left(P_{nt}, P_t, Y_t, F_{it}, Z_{nt}, \{P_{-nt}\}_{-n \in \aleph_i} \right).$$

Our notation remains identical to the main text for aggregate prices P_t, real aggregate demand Y_t, firm-specific shocks F_{it} and good-specific shocks Z_{nt}. The novelty comes in the last argument, $\{P_{-nt}\}_{-n \in \aleph_i}$, which represents the prices set by firm i for all its produced goods but good n.

Optimal frictionless prices now solve

$$P^\diamondsuit_{nt} = \arg\max_{P_{nt}} \mathbb{E} \left[\sum_n \pi \left(P^*_{nt}, P_t, Y_t, F_{it}, Z_{nt}, \left\{ P^*_{-nt} \right\}_{-n \in \aleph_i} \right) \right]$$

This problem is identical to the one in the main text with the exception that optimal frictionless prices must take into account their effect on the contribution to profits of all goods produced by the same firm. The optimality of prices implies that in steady state prices must solve

$$\pi_1 \left(1, 1, Y_t, \bar{F}, \bar{Z}, \{1\}_{-n \in \aleph_i} \right) + (N-1)\, \pi_6 \left(1, 1, Y_t, \bar{F}, \bar{Z}, \{1\}_{-n \in \aleph_i} \right) = 0;$$

which implicitly assumes equal marginal effect of the price of any good on other good's profits.

A second order approximation of the total profits function is

$$\begin{gathered}\left(\widehat{\pi}_1+\widehat{\pi}_6\left(N-1\right)\right)p_{nt}+\frac{1}{2}\left(\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right)p_{nt}^2+\left(\widehat{\pi}_{12}+\widehat{\pi}_{62}\left(N-1\right)\right)p_{nt}p_t\\ +\left(\widehat{\pi}_{13}+\widehat{\pi}_{63}\left(N-1\right)\right)p_{nt}y_t+\left(\widehat{\pi}_{14}+\widehat{\pi}_{64}\left(N-1\right)\right)p_{nt}f_{it}+\widehat{\pi}_{15}p_{nt}z_{nt}\\ +\sum_{-n\in\aleph_i}\widehat{\pi}_{65}p_{nt}z_{-nt}+2\sum_{-n\in\aleph_i}\widehat{\pi}_{16}p_{nt}p_{-nt}\\ +\ terms\ independent\ of\ p_{nt}.\end{gathered}$$

Hence, the optimal frictionless price solves

$$\begin{aligned}p_{nt}^{\diamondsuit} &= \frac{\widehat{\pi}_{12}+\widehat{\pi}_{62}\left(N-1\right)}{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}p_t+\frac{\widehat{\pi}_{13}+\widehat{\pi}_{63}\left(N-1\right)}{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}y_t+\frac{\widehat{\pi}_{14}+\widehat{\pi}_{64}\left(N-1\right)}{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}f_{it}+\\ &\quad \frac{\widehat{\pi}_{15}}{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}z_{nt}+\sum_{-n\in\aleph_i}\frac{\widehat{\pi}_{65}}{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}z_{-nt}+\sum_{-n\in\aleph_i}\frac{2\widehat{\pi}_{16}}{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}p_{-nt}^{\diamondsuit}.\end{aligned}$$

The interdependence between profit functions has two implications on optimal frictionless prices. First, frictionless prices respond to all good-specific shocks that hit a given firm. Second, frictionless prices respond to other prices set by the same firm. If we represent this linear pricing rule by

$$p_{nt}^{\diamondsuit}=b_0p_t+b_1y_t+b_2f_{it}+b_3z_{nt}+b_4\sum_{-n\in\aleph_i}z_{-nt}+b_5\sum_{-n\in\aleph_i}p_{-nt}^{\diamondsuit},$$

then a reduced form of this rule is

$$p_{nt}^{\diamondsuit}=\frac{1}{1-\left(N-1\right)b_5}\left[\begin{array}{c}b_0p_t+b_1y_t+b_2f_{it}+\left(b_3-\frac{(N-1)b_5(b_3-b_4)}{1+b_5}\right)z_{nt}\\ +\left(b_4+\frac{b_5(b_3-b_4)}{1+b_5}\right)\sum_{-n\in\aleph_i}z_{-nt}\end{array}\right]$$

with a short-hand representation as

$$p_{nt}^{\diamondsuit}=a_0p_t+a_1y_t+a_2f_{it}+a_3z_{nt}+a_4\sum_{-n\in\aleph_i}z_{-nt}.$$

Note that a_0, a_1, a_2, a_3 and a_4 are functions of N. Further, to obtain neutrality of frictionless prices,

$$a_0=1$$

and to ensure that $a_1>0$, parameters must satisfy

$$1-\left(N-1\right)b_5\equiv\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|-2\left(N-1\right)\widehat{\pi}_{16}>0.$$

Turning to solve for optimal prices under rational inattention, we start by computing the second-order approximation for

$$\sum_{n,-n\in\aleph_i}\left\{\widetilde{\pi}\left(p_{nt}^{\diamondsuit},p_t,y_t,f_{it},z_{nt},\left\{p_{-nt}^{\diamondsuit}\right\}_{-n\in\aleph_i}\right)-\widetilde{\pi}\left(p_{nt}^{*},p_t,y_t,f_{it},z_{nt},\left\{p_{-nt}^{*}\right\}_{-n\in\aleph_i}\right)\right\}$$

which solves

$$\frac{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}\left(N-1\right)\right|}{2}\sum_{n\in\aleph_i}\left(p_{nt}^{\diamondsuit}-p_{nt}^{*}\right)^2-\widehat{\pi}_{16}\sum_{n\in\aleph_i}\sum_{-n\in\aleph_i}\left(p_{nt}^{\diamondsuit}-p_{nt}^{*}\right)\left(p_{-nt}^{\diamondsuit}-p_{-nt}^{*}\right).$$

Guessing $p_t=\alpha q_t$, defining $\Delta_t\equiv p_t+a_1y_t$, imposing

$$p_{nt}^{*}=\frac{\sigma_{\Delta}^2}{\sigma_{\Delta}^2+\sigma_{\varepsilon i}^2}s_{it}^a+a_2\frac{\sigma_f^2}{\sigma_f^2+\sigma_{ei}^2}s_{it}^f+a_3\frac{\sigma_z^2}{\sigma_z^2+\sigma_{\psi n}^2}s_{nt}^z+a_4\sum_{-n\in\aleph_i}\frac{\sigma_z^2}{\sigma_z^2+\sigma_{\psi n}^2}s_{nt}^z,$$

and using the definitions of κ_a,κ_f and $\{\kappa_n\}_{n\in\aleph_i}$, the problem of a decision unit taking $\widetilde{N}$ pricing decisions within a firm that sells (or produces) N goods is

$$\min_{\kappa_a,\kappa_f,\{\kappa_n\}_{n\in\aleph_i}}\left\{\begin{array}{c}\frac{\left|\widehat{\pi}_{11}+\widehat{\pi}_{66}(N-1)\right|}{2}\left[\left(2^{-2\kappa_a}\sigma_{\Delta}^2+a_2^2 2^{-2\kappa_f}\sigma_f^2\right)\widetilde{N}+\left(a_3^2+\left(N-1\right)a_4^2\right)\sum_{n\in\aleph_i}2^{-2\kappa_n}\sigma_z^2\right]\\-\widehat{\pi}_{16}\left(N-1\right)\left[\left(2^{-2\kappa_a}\sigma_{\Delta}^2+a_2^2 2^{-2\kappa_f}\sigma_f^2\right)\widetilde{N}+\left(2a_3a_4+a_4^2\left(N-2\right)\right)\sum_{n\in\aleph_i}2^{-2\kappa_n}\sigma_z^2\right]\end{array}\right\}.$$

subject to

$$\kappa_a+\kappa_f+\sum_{n\in\aleph_i}\kappa_n\leq\kappa\left(\widetilde{N}\right)$$

We make the distinction between $\widetilde{N}$ and N because firms now can have a portfolio of goods N they sell/produce but a pricing unit within firms may price only $\widetilde{N}$ of them. As in the main text, we focus on pricing units. A pricing unit is endowed by information capacity $\kappa\left(\widetilde{N}\right)$ which, as in the main text, may depend on the number $\widetilde{N}$ of prices that this decision unit must set. To do so, a decision unit must take into account the cross effects of all prices set within the firm, which is captured by the optimal pricing rules for $p_{nt}^{\diamondsuit}$ and p_{nt}^{*} obtained above.

The first-order conditions for the allocation of attention are now

$$\begin{aligned}\kappa_a^{*}&=\kappa_f^{*}+\log_2\left(\widetilde{x}_1\left(N\right)\right),\\\kappa_a^{*}&=\kappa_n^{*}+\log_2\left(\widetilde{x}_2\left(N\right)\sqrt{\widetilde{N}}\right),\quad\forall n\in\aleph_i.\end{aligned}$$

The economies of scope in information processing are captured by $\sqrt{\widetilde{N}}$ in the second condition. The interdependence of profits introduced here are captured in $\widetilde{x}_1(N)$ and $\widetilde{x}_2(N)$, which in the main text are parameters and here are functions of N:

$$\begin{aligned}\widetilde{x}_1 &\equiv \frac{\sigma_\Delta}{a_2\sigma_f},\\ \widetilde{x}_2(N) &\equiv \left[\frac{\left(\frac{|\widehat{\pi}_{11}+\widehat{\pi}_{66}(N-1)|}{2}-\widehat{\pi}_{16}(N-1)\right)\frac{\sigma_\Delta^2}{\sigma_z^2}}{\frac{|\widehat{\pi}_{11}+\widehat{\pi}_{66}(N-1)|}{2}\left(a_3^2+(N-1)a_4^2\right)-\widehat{\pi}_{16}(N-1)\left(2a_3a_4+a_4^2(N-2)\right)}\right]^{\frac{1}{2}}\end{aligned}$$

We then follow a similar logic than in Proposition 4. We discipline $\kappa\left(\widetilde{N}\right)$ by assuming that the information capacity of decision units depends on the number $\widetilde{N}$ of decisions they take such that they have no incentives to merge or delegate their pricing decisions. This assumption is equivalent to assume that the frictional cost per good produced in a firm that produces N goods is independent of the number $\widetilde{N}$ of decisions taken by decision units within the firm. Under this assumption, we can establish that

$$\kappa_a^*\left(\widetilde{N};N\right)=\kappa_a^*(1;N)+\frac{1}{2}\log_2\left(\frac{\widetilde{N}+2}{3}\right)+\frac{1}{2}\log_2\left(\frac{\sigma_\Delta^2\left(\widetilde{N};N\right)}{\sigma_\Delta^2(1;N)}\right).$$

This expression is identical to Proposition 6, but its interpretation is more subtle. In an economy where firms sell/produce N goods, the attention paid to aggregate shocks is increasing in the number $\widetilde{N}$ of pricing decisions that single decision units must take within firms. As in the main text, this result highlights the importance of economies of scope in information processing on the aggregate predictions of the rational inattention model. In the literature, these economies of scope are shut down by the assumption that firms produce only one good and decide only one price.

Finally, we drop the distinction between N and $\widetilde{N}$, that is, $N=\widetilde{N}$, to produce a version of proposition 3. This assumption is consistent with the evidence that a single decision unit prices all goods in the firm's portfolio of goods. If we arrange parameters such that firms' attention to monetary shocks is invariant in N, $\kappa_a^*(N)=\overline{\kappa}_a$, then the frictional cost is

$$C_n(N)=\left(\frac{|\widehat{\pi}_{11}+\widehat{\pi}_{66}(N-1)|}{2}-\widehat{\pi}_{16}(N-1)\right)(N+2)\,2^{-2\overline{\kappa}_a}\sigma_\Delta^2$$

which is increasing in N unless $\widehat{\pi}_{16} > 0$ is high enough. If this is the case, then the term

$$\frac{\left|\widehat{\pi}_{11} + \widehat{\pi}_{66}\left(N-1\right)\right|}{2} - \widehat{\pi}_{16}\left(N-1\right)$$

is decreasing in N. This has two implications. The first is that strategic complementarity in pricing (a_1) is increasing in N. As in the main text, the complementarity in pricing is deduced from aggregate data, so it should remain constant in our quantitative exercises. The second is that per-good expected profits of the firm falls as N increases. This contradicts our assumption that the number of produced goods by firms is exogenous and the observation that firms produce multiple goods.

Therefore, we conclude that in a relevant parametrization of the model the frictional cost must be increasing as N increases to keep $\kappa_a^*\left(N\right)$ invariant to N.

Appendix D. Robustness of Key Statistics

We summarize in this section the variation of additional statistics of the CPI when we include sales, complementing Table 1. We also show that our key statistics are robust to controlling for firm size. We discuss this in detail in Section 4.5.

Table D.7: Multi-Product Firms and Moments from CPI, Including Sales

CPI	1-3 Goods	3-5 Goods	5-7 Goods	>7 Goods	All
Absolute size of price changes	10.10%	14.00%	12.77%	18.59%	14.54%
	(0.04%)	(0.11%)	(0.2%)	(0.14%)	(0.04%)
Cross-sectional variance	3.31%	4.57%	5.40%	4.67%	4.33%
	(0.53%)	(0.89%)	(0.65%)	(0.51%)	(0.35%)
Serial correlation	−0.4741	−0.5447	−0.5118	−0.566	−0.526
	(0.0005)	(0.0007)	(0.0010)	(0.0007)	(0.0003)

NOTE: We compute the above statistics using the monthly micro price data underlying the CPI, exactly as in Table 1. Here, we additionally consider sales price changes.

Table D.8: Multi-Product Firms and Moments from PPI, Controlling for Size

CPI	1-3 Goods	3-5 Goods	5-7 Goods	>7 Goods	All
Within dispersion ratio of $\|\Delta p\|$	16.10%	25.70%	34.64%	41.20%	29.41%
	(0.70%)	(0.89%)	(1.11%)	(0.97%)	(0.67%)
Absolute size of price changes	7.83%	7.28%	6.39%	6.38%	7.35%
	(0.13%)	(0.10%)	(0.19%)	(0.28%)	(0.81%)
Cross-sectional variance	2.20%	2.60%	2.26%	3.48%	2.64%
	(0.12%)	(0.14%)	(0.11%)	(0.19%)	(0.08%)
Serial correlation	−0.0551	−0.0735	−0.0643	−0.0734	−0.042
	(0.0014)	(0.0002)	(0.0002)	(0.0002)	(0.0009)

NOTE: We compute the above statistics using the monthly micro price data underlying the PPI, exactly as in Table 1. Here, we additionally control for the number of employees. In the case of the within-firm ratio, we filter out firm size from individual price changes and then proceed as before. In all other cases, we filter out firm size from the firm-level statistics.

Appendix E. Proofs

Proof of Proposition 1

When $\sigma_z = 0$ or $\pi_{15} = 0$, firms ignore signals s_t^z regarding firm-specific shocks, so $\kappa_z^* = 0$. Then κ_a^* is obtained from combining the condition in (4) and the constraint $\kappa_a + \kappa_f = \kappa$:

$$\kappa_a^* = \frac{1}{2}\left[\kappa + \log_2(x_1)\right].$$

which is constant in N.

Proof of Lemma 1

the optimal pricing rule reduces to

$$p_{nt}^* = \left(1 - 2^{-2\kappa_a^*}\right)(\Delta_t + \varepsilon_{it}) + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\left(1 - 2^{-2\kappa_f^*}\right)(f_{it} + e_{it})$$

which only varies with aggregate or firm-specific disturbances Δ_t, f_{it}, ε_{it} and e_{it}.

Proof of Proposition 2

$\hat{N}$ solves $\frac{\partial \kappa_a^*}{\partial N} = 0$ for the interior solution of (6) after setting $\kappa(N) = \kappa$ and assuming that α and thus σ_Δ are constant in N. This is true when $\frac{\partial \kappa_a^*}{\partial N} = 0$. Further, since α is increasing in κ_a^* and κ_a^* is increasing in α, $\hat{N}$ is unique.

Proof of Proposition 3

Given the first order conditions in (4) and (5) the frictional cost in (8) becomes

$$C_n(N) = \frac{|\widehat{\pi}_{11}|}{2} 2^{-2\kappa_a(N)} \sigma_\Delta^2 (N+2)$$

which is increasing in N if κ_a^* is invariant or decreasing in N.

Proof of Proposition 4

Given the first order conditions in (4) and (5) the frictional cost in (8) becomes

$$C_n(N) = \frac{|\widehat{\pi}_{11}|}{2} 2^{-2\kappa_a(N)} \sigma_\Delta^2 (N+2)$$

The result follows after equating $C(N)$ and $C(1)$, and solving for $\kappa_a(N)$ while noticing that σ_Δ^2 is also a function of N through α.

Proof of Proposition 5

From α in (7),

$$\frac{\partial \alpha}{\partial (2^{2\kappa_a})} = \frac{\left(\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\right)^{-1}}{\left[\left(\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\right)^{-1} + (2^{2\kappa_a} - 1)\right]^2}$$

which is positive and decreasing in $\kappa_a \geq 0$.