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

Ernesto Pasten† Raphael Schoenle‡

March 24, 2014

Abstract

Economies of scope in information processing naturally arise in a rational inattention model of multi-product firms: Processing information is difficult, but once information is internalized, it can be freely used for all decisions it concerns. Monetary shocks concern all pricing decisions in multi-product firms; good-specific shocks, however, concern only a few. Hence, in a model with good-specific shocks, attention to monetary shocks increases as firms produce more goods. Such good-specific shocks are necessary in our model to account for the dispersion of price changes within firms observed in both CPI and PPI data in the U.S.. Our model calibrated to CPI data predicts perfect neutrality of money while calibrated to PPI data it predicts limited non-neutrality.

JEL codes: E3, E5, D8

Keywords: rational inattention, multi-product firms, monetary non-neutrality

*We thank comments by Roland Benabou, Markus Brunnermeier, Paco Buera, Larry Christiano, José de Gregorio, Eduardo Engel, Christian Hellwig, Hugo Hopenhayn, Pat Kehoe, Narayana Kocherlakota, Ben Malin, Virgiliu Midrigan, Juanpa Nicolini, Kristoffer Niemark, 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 Zentrum für Europäische Wirtschaftsforschung (ZEW), Richmond Fed, 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. We thank project coordinator Ryan Ogden for his help and effort 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. Errors and omissions are our own.

†*ernesto.pasten@tse-fr.eu*. Senior Economist, Central Bank of Chile; Research Fellow, Toulouse School of Economics.

‡*schoenle@brandeis.edu*. Assistant Professor, Department of Economics, Brandeis University.

1 Introduction

Rational Inattention Theory (Sims, 1998, 2003) has become an increasingly popular formalization of the idea that limited cognitive ability can reconcile the simplicity of human actions in reality with the complexity of agents in economic models. A prime example in this context – as pointed out in Sims' seminal work – is that firms' prices only slowly respond to monetary policy shocks, which in turn creates macroeconomic inertia. According to the theory, this happens because firms must optimally allocate their limited information processing capacity (their "attention") to reduce observation noise of any aggregate and idiosyncratic shocks. In doing so, they tend to allocate most of their attention to idiosyncratic shocks, and respond slowly to aggregate – in this example: monetary – shocks since aggregate shocks are less volatile than idiosyncratic shocks. Mackowiak and Wiederholt (2009) confirm this result by calibrating a rational inattention model of price setting with monetary and idiosyncratic shocks to US data. They find that such a model indeed yields strong monetary non-neutrality and large aggregate inertia even when the friction of limited cognitive ability is small.

We revisit this result of rational inattention after we relax the usual assumption in macroeconomics that firms price a single good. Instead, we assume that firms price multiple goods. This assumption then directly implies the existence of economies of scope in information processing: information remains available for other decisions at no cost after it has been used once. Thus, when a firm prices more goods, economies of scope predict that a firm should pay more attention to common shocks like monetary and firm-specific shocks, relative to shocks that are specific to a few decisions like good-specific shocks.

These economies of scope are clearly a very general feature of Rational Inattention Theory. They generalize to a wide number of problems as we point out further below, beyond the important topic of pricing. We see our main contribution in demonstrating their quantitative importance in the application to pricing and monetary policy. To do so, we compute the response of prices to a monetary shock in two versions of a rational inattention model of multi-product firms calibrated to match micro moments of prices in the Consumer Price Index (CPI) and the Producer Price Index (PPI) in the US, where we interpret firms as 'stores' or 'good producers'.

We use Mackowiak and Wiederholt (2009) as benchmark. In contrast to previous work, we find a new, quantitatively significant tension between monetary non-neutrality and the size of the friction. This tension is stronger when firms price more goods. The cumulative real effect of a money shock in our model is cut by three when firms price two goods instead of one if the friction is as binding as in the benchmark; money is almost neutral when firms price eight goods or more. Conversely, the severity of the friction quickly exceeds that in the benchmark or other levels from the literature if we aim to keep the degree of monetary non-neutrality constant when firms price more goods. What is an empirically relevant number of goods? The paper discusses evidence on multi-production but, to fix ideas, stores price about $40,000$ goods (FMI, 2010) and producers about 4 goods (Bhattarai and Schoenle (2011)). Given these numbers, we conclude that there is no room for monetary non-neutrality in a rational inattention model for stores. For good producers, the model yields limited monetary non-neutrality and aggregate inertia relative to previous work.

Our key assumptions are that first, a single decision unit within firms prices multiple goods and second, that prices respond to good-specific shocks. Regarding the first assumption, the above evidence suggests that it is unlikely that the number of independent price setters within stores price less than eight goods. For good producers, our interpretation of multi-product firms is consistent with the Bureau of Labor Statistic's (BLS) definition of a firm as a "price-setting unit."[1] Work by Zbaracki et al. (2004) provides additional evidence that indeed a single price setting unit within firms sets all prices although there is some geographical decentralization.

Regarding our second assumption, we document a new empirical fact: There is strong dispersion of price changes within firms. In the CPI data, within-firm dispersion is on average 51.6% of the total dispersion of log non-zero price changes. In the PPI data, this ratio is 59.1%. Such within-firm dispersion in the PPI data is quite remarkable given that firms presumably produce rather homogeneous goods relative to what stores sell. We interpret this fact as evidence that prices respond to good-specific shocks. However, for our results to hold, we only need that good-specific shocks account for a non-zero fraction of the within-firm dispersion of price changes in the data.

Although our main focus is on multi-product firms, we also make the point that the economies of scope play a role even for single-product firms whenever shocks differ in their persistence. The

[1] Given sampling constraints on the number of goods, we in fact consider our estimate as a lower bound.

logic is straight-forward: economics of scope can also accumulate over time. Processing information about more persistent shocks is more attractive for firms because of the relatively high future information content. Does this insight matter quantitatively? In our benchmark monetary and idiosyncratic shocks are equally persistent. We show that this implies a higher serial correlation of log non-zero price changes in the model than in the data. Once we adjust the persistence of idiosyncratic shocks to match the data, our model yields cumulative real effects of a monetary shock about half the size of the benchmark.

What is the exact mechanism driving our results? First, note that our model augments Mackowiak and Wiederholt (2009) by breaking idiosyncratic shocks in multi-product firms up into firm- and good-specific shocks. We continue to allow for monetary shocks. Then, in addition to economies of scope, there are two other forces at work as firms price more goods. The first is that firms must simply pay attention to more shocks. We call this force the "income effect:" it makes the friction more binding given an information processing capacity. The second is the "aggregation effect:" firms have stronger incentives to increase their information processing capacity as they price more goods. Although there is no theory to guide us how to model the aggregation effect, we use a simplified analytical version of our model to show that in any set up in which the friction is equally or less binding as firms price more goods, attention to monetary shocks increases. Conversely, if attention to monetary shocks is constant or decreases as firms price more goods, the friction becomes more binding. Our quantitative exercises using a generalized version of our model confirm these results.

In particular, in our analytical model we also show that even in a situation in which firms spend little attention to monetary shocks to begin with, just a little more attention to these shocks already has a strong effect on reducing monetary non-neutrality. This happens because there is strategic complementarity amongst price setters. We confirm this again in our quantitative exercises: firms' attention to monetary shocks remains a small portion of their total attention although monetary non-neutrality vanishes as firms price more goods.

Such complementarity also shows up when we calibrate our model to PPI moments. This dataset allows us to sort prices by "bins" according to the number of goods firms price. We can then calibrate a "multi-sector" version of our model with firms pricing heterogenous numbers of goods to moments by bins. We find that prices in each bin respond to a monetary shock very similarly – unlike in a homogeneous-firm model. This means that firms pricing many goods have an indirect effect on monetary non-neutrality: they force firms with fewer goods to pay more attention to monetary shocks. Overall, monetary non-neutrality is still limited unless a large friction is assumed.

Underlying our model there are several simplifying assumptions. However, if we relax them, our results remain quite robust. First, sources of information for each type of shock are independent of each other. We show that assuming otherwise decreases within-firm dispersion. Second, the scale of firms increases as they price more goods. Results are identical if firms' scale is preserved. Third, firms' profits for goods that they price are independent of each other. This assumption ignores complementarities among goods that firms exploit. We show that taking into account such complementarities complicates the model but has no effect on our results.

Moreover, there is also a sensitivity issue in the macro and micro predictions of our model that deserves some comments. As mentioned above, firms spend most of their attention on idiosyncratic shocks, so model-predicted moments for prices are quite insensitive to variations in firms' attention to monetary shocks. In contrast, strategic complementarity implies that monetary non-neutrality in the model is highly sensitive to such variation. In our view, this issue invalidates the use of micro moments to calibrate the size of the friction; we hence rely on standards established in the literature. This issue also implies that twisting parameters does not help to undo the effect of multi-production in the model. Multi-production changes firms' attention to monetary shocks almost without affecting the response of prices to idiosyncratic shocks. Any twist in parameters would affect this response and impede the model from matching micro moments in the data.

A common objection to the Rational Inattention Theory merits a final remark: one may think that it gives room to specialization and trade of information. However, this misses the point that decision makers still have to *process* all information they need. The point we are making in this paper has a similar flavor to the objection at first glance, but it is fundamentally different: Once internalized by an agent, information is still available for her (and only for her) after she has used it. This is the source of the economies of scope in information processing. We focus on the monetary application of the theory to show that these economies of scope play a role when firms price multiple goods and shocks have different persistence. Among these, multi-production is very important quantitatively.

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 (2009, 2011), Paciello and Wiederholt (2011) and Matejka 2013), asset pricing (Peng and Xiong (2006)), portfolio choice (Mondria (2010)), 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 (2011)) 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 (2011) and Alvarez and Lippi (2013). 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 not an 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. 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.

Layout. Section 2 displays the model setup and solves it with and without frictions in a simplified analytical version. Section 3 uses this analytical version to theoretically study the effects of multi-production and persistence. Section 4 presents empirical evidence in the literature and in the CPI and PPI data to support the distinctive features of our model. Section 5 calibrates our model to these CPI and PPI moments and obtain all our quantitative results. Section 6 concludes and appendices collect tables, figures, extensions and material omitted in the main text.

2 A Model of Multi-Product, Rationally Inattentive Firms

Our model builds on Mackowiak and Wiederholt (2009) where firms face an information processing constraint of gathering information about nominal aggregate demand – in short, monetary – shocks, firm-specific shocks and good-specific shocks. Firms use this information to price an exogenous number of goods. This section introduces the model and solves it analytically under the assumption of white noise shocks.

2.1 Setup

Consider an economy with a continuum of goods of measure one indexed by $j \in [0,1]$, and a continuum of monopolist firms with 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.

Each good j contributes to the total profits of the producer according to

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

where P_{jt} is the fully flexible 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

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

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

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

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

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

where aggregate prices follow from

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

The total period profit function of firm 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 produced by firm i.

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

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

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 $I(\cdot)$ measures the information flow between observed signals and variables of interest, and $\kappa(N)$ is an exogenous, limited capacity that without loss of generality is assumed to depend on the number N of goods.

The information flow $I(\cdot)$ is a measure of how informative an observed signal is with respect to a given variable. This measure has been proposed by Shannon (1948) and has a complicated functional form that does not need to be specified here except for computational purposes, so we relegate it to the appendix. However, to provide intuition, if one denotes as U_t an arbitrary unobservable variable of interest and as O_t an arbitrary observable signal, and assumes that U_t and O_t are Gaussian i.i.d. processes, then the information flow between U_t and O_t is given by

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

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 the variables of interest.

We also assume that the vector of signals s_{it} may be partitioned into $N+1$ subvectors

$$\left\{s_{it}^{a}, s_{it}^{f}, \{s_{nt}^{z}\}_{n\in\aleph_i}\right\};$$

where each subvector is correlated to one target variable such that $\{q_t, s_{it}^{a}\}$, $\left\{f_{it}, s_{it}^{f}\right\}$ and $\{z_{nt}, s_{nt}^{z}\}_{n\in\aleph_i}$ are independent of each other. Besides, we assume that all variables are Gaussian, jointly stationary and there exists an initial infinite history of signals:

$$s_i^1 = \{s_{i-\infty}, \ldots, s_{i1}\}.$$

These assumptions imply that the information flow is additively separable according to

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

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

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

where

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

is subject to

$$I\left(\{P_t, Y_t\}, \{s_{it}^{a}\}\right) + I\left(\{F_{it}\}, \left\{s_{it}^{f}\right\}\right) + \sum_{n\in\aleph_i} I\left(\{Z_{nt}\}, \{s_{nt}^{z}\}\right) \le \kappa(N)$$

$$\Leftrightarrow \kappa_a + \kappa_f + \sum_{n\in\aleph_i}\kappa_n \le \kappa(N). \tag{9}$$

To abbreviate notation, we denote $I\left(\{P_t, Y_t\}, \{s_{it}^{a}\}\right)$, $I\left(\{F_t\}, \left\{s_{it}^{f}\right\}\right)$ and $I\left(\{Z_{nt}\}, \{s_{nt}^{z}\}\right)$ as κ_a, κ_f and κ_n, where $I\left(\{P_t, Y_t\}, \{s_{it}^{a}\}\right) = I\left(\{Q_t\}, \{s_{it}^{a}\}\right)$ since the only source of aggregate disturbances is Q_t. The absence of nominal rigidities implies that the pricing problem in (8) is static.

The firm, however, must consider its whole discounted expected stream of profits to allocate its information flow capacity, its "attention", among a set Γ of signals. These signals are restricted to satisfy the above assumptions and must contain no information about future realizations of shocks. If a firm chooses more precise signals about, for instance, $\{P_t, Y_t\}$, then information flow $I(\{P_t, Y_t\}, \{s_{ait}\})$ increases reducing the information capacity allocated to other signals.

We define the equilibrium in this economy as follows:

Definition 1 *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 $\{P_t\}, \{Y_t\}, \{F_{it}\}_{i\in[0,\frac{1}{N}]}$ and $\{Z_{jt}\}_{j\in[0,1]}$, all firms $i \in [0, \frac{1}{N}]$ 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 (4) and (5) for $t \geq 1$.*

Discussion. A profit function $\pi(\cdot)$ that is independent across goods implies that the pricing problem in (8) is independent of N. However, N enters the attention allocation problem through three channels. First, the period objective in (7) sums up the contribution to profits of all goods produced by the firm. This is the source of economies of scope in information processing highlighted in this paper. Second, the firm simply has to pay attention to more good-specific shocks when the firm prices more goods. We label this the "income effect" since it brings to mind a consumer whose basket of goods increases with N. This channel is captured by the left-hand side of (9). Finally, the capacity constraint $\kappa(N)$ in (9) may also depend on N; we call this channel the aggregation effect.

The aggregation effect simply acknowledges that firms may have different capacity to process information when they price a different number of goods. After all, this capacity should be endogenous to firms' internal organization or their investment in information technology. However, there is no theory to guide us how to model this endogenous choice exactly. As we show below, we also cannot calibrate $\kappa(N)$ using micro moments in any straight-forward way. Hence, we take no direct stand on it. Instead, we simply make a variety of alternative assumptions to discipline the aggregation effect, and study the implications of such assumptions.

We make some simplifying assumptions in our model economy in this section: First, we keep the number N of goods produced by all firms constant; Second, we assume that profits $\pi(\cdot)$ are independent across goods produced by the same firm; third, we assume that signals are informative only about one type of shocks. We subsequently relax these assumptions. In section 3, we allow for heterogeneity in N within the same economy, and use this model for calibrations to PPI data in section 5. We relax the other assumptions in Appendix C to find either counterfactual predictions or no substantive effects.

We acknowledge two additional innocuous restrictions: First, in our setup, we do not allow for sectoral or regional shocks. However, these classes of shocks play the same role as our firm-specific shocks: sectoral or regional shocks are common to all decisions of a given price setter but wash out in the aggregate. Second, we implicitly assume that the size of a firm is increasing in N. This comes from the assumption that the measure of goods is fixed as one and the measure of firms is $\frac{1}{N}$. Our analysis goes through unchanged if we preserve the size of firms by assuming that the measure of firms is fixed at 1 and the measure of goods is N.

Finally, we next solve this setup under the assumption that shocks are Gaussian i.i.d. This simplification allows for an analytical solution, but our setup also allows for a more general specification of shocks. We solve this generalized problem in the Appendix B. We use such a solution in our quantitative analysis in section 5.

2.2 Solution for White Noise Shocks

Here, we present the key steps of the solution. As a main result, we derive the expression that relates monetary non-neutrality to information capacity. We present details in Appendix A.

First, when shocks are Gaussian and i.i.d., the firm's problem in (7) and (8) – up to a second-order approximation – is defined as the choice of attention to monetary, firm- and good-specific shocks to minimize the discounted sum of expected losses in profits due to the friction. After some algebra, this problem becomes

$$\min_{\kappa_a,\kappa_f,\{\kappa_n\}_{n\in\aleph_i}} \frac{\beta}{1-\beta}\frac{|\widehat{\pi}_{11}|}{2}\left[2^{-2\kappa_a}\sigma_\Delta^2 N+\left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_f}\sigma_f^2 N+\left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2\sum_{n\in\aleph_i}2^{-2\kappa_n}\sigma_z^2\right] \tag{10}$$

subject to the rational inattention constraint in (9).

In this expression, σ_Δ^2 is the volatility of a compound aggregate variable

$$\Delta_t \equiv p_t + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} y_t \tag{11}$$

that linearly depends on monetary shocks q_t after we guess that the log-deviation of aggregate prices responds linearly to monetary shocks, $p_t = \alpha q_t$. We confirm this guess below. In addition, σ_f^2 and σ_z^2 are respectively the volatility of firm- and good-specific shocks. 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. Parameters $\widehat{\pi}_{11}, \widehat{\pi}_{13}, \widehat{\pi}_{14}$ and $\widehat{\pi}_{15}$ are the derivatives of the marginal effect of the good price on own profits with respect to the good price, real aggregate demand, firm- and good-specific shocks, all evaluated at the non-stochastic steady state.

From the first order conditions of this problem, we obtain

$$\kappa_a^* = \kappa_f^* + \log_2 (x_1), \tag{12}$$

$$\kappa_a^* = \kappa_n^* + \log_2 \left(x_2 \sqrt{N} \right), \ \ \forall n \in \aleph_i \tag{13}$$

for $x_1 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta}{\widehat{\pi}_{14}\sigma_f}$ and $x_2 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta}{\widehat{\pi}_{15}\sigma_z}$. The assumption that all parameters are the same for all firms and goods along with the conditions in (12) and (13) has two implications: first, the attention paid to monetary and firm-specific shocks, κ_a^* and κ_f^*, is the same for all firms; second, the attention paid to good-specific shocks is the same for all goods within all firms, $\kappa_n^* = \kappa_z^*$ for all $n \in \aleph_i$ and all i.

In addition, the conditions in (12) and (13) along with the constraint imply that

$$\kappa_a^* = \frac{1}{N+2} \left[\kappa (N) + \log_2 (x_1) + N \log_2 \left(x_2 \sqrt{N} \right) \right] \tag{14}$$

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, for a given N, the smaller are either capacity $\kappa (N)$ or parameters x_1 and x_2, the smaller is the attention to monetary shocks, or equivalently, the larger is the observation noise of these shocks. A smaller x_1 results from this expression when the volatility σ_f of firm shocks is larger relative to the volatility σ_Δ of the aggregate compound variable in (24) and/or when frictionless prices are more responsive to firm shocks, i.e., when $\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}$ is larger. Similarly, we obtain a smaller x_2 when $\frac{\sigma_z}{\sigma_\Delta}$ is larger and/or when $\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}$ is larger.

Since all firms are identical, the price of any good $n \in \aleph_i$ for any firm i follows

$$p^*_{nt} = \left(1 - 2^{-2\kappa^*_a}\right)(\Delta_t + \varepsilon_{it}) + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\left(1 - 2^{-2\kappa^*_f}\right)(f_{it} + e_{it}) + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}\left(1 - 2^{-2\kappa^*_z}\right)(z_{nt} + \psi_{nt}) \quad (15)$$

where ε_{it}, e_{it} are the realizations of the noisy signals observed by firm i of a monetary shock and a shock to firm i. ψ_{nt} is the realization of the noise of signals of a shock to good n. Aggregating among all goods and firms by using (2), (3) and (5), the log-deviation of the aggregate prices with respect to the steady state is

$$p^*_t = \left(1 - 2^{-2\kappa^*_a}\right)\Delta_t = \left(1 - 2^{-2\kappa^*_a}\right)\left[\alpha + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}(1-\alpha)\right] q_t$$

which confirms the guess $p^*_t = \alpha q_t$ for

$$\alpha = \frac{\left(2^{2\kappa^*_a} - 1\right)\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}}{1 + \left(2^{2\kappa^*_a} - 1\right)\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}}. \quad (16)$$

This is the most important result of the Rational Inattention Theory. If firms have unlimited information-processing capacity, $\kappa(N) \to \infty$, then firms choose infinitely precise signals about monetary shocks (and all other shocks), so $\kappa^*_a \to \infty$ and $\alpha \to 1$. Money is fully neutral. In contrast, if firms have limited information-processing capacity, firms choose signals with finite precision, so κ^*_a is finite and thus $\alpha < 1$. Money becomes non-neutral. The more attention firms pays to idiosyncratic shocks – the higher are either κ^*_f or κ^*_z – the lower is κ^*_a, so monetary non-neutrality is stronger. Moreover, for a given κ^*_a, the stronger is complementarity in pricing decisions among firms – the smaller is $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} > 0$ – the stronger is monetary non-neutrality.

3 Theoretical Results

We next conduct a comparative statics analysis to illustrate the implication of introducing multi-product firms and good-specific shocks into the model. We have mentioned above three forces by which these ingredients affect firms' attention allocation. Before presenting our propositions, we build up intuition for the underlying mechanisms by going through these forces.

The first force is the economies of scope in information processing: The more goods a firm prices, the more pricing decisions can benefit from information that is common to all goods. These economies of scope are captured in the first-order conditions in (12) and (13), which we re-write as

$$\begin{aligned}\kappa_a^* &= \kappa_z^* + \log_2\left(x_2\sqrt{N}\right),\\ \kappa_f^* &= \kappa_z^* + \log_2\left(\frac{x_2}{x_1}\sqrt{N}\right),\end{aligned}$$

after imposing that $\kappa_n^* = \kappa_z^*$ for all $n \in \aleph_i$ and all i. In words, the difference in attention paid by the firm to aggregate and good-specific shocks is increasing in N since $x_2 > 0$ while the difference in attention paid to firm- and good-specific shocks is increasing in N since $x_1, x_2 > 0$.

The second force is the income effect: Firms must pay attention to more good-specific shocks as firms price more goods. This force is captured by the N on the left-hand side of the constraint (9), which we re-write as

$$\kappa_a + \kappa_f + N\kappa_z \leq \kappa(N)$$

after imposing that $\kappa_n^* = \kappa_z^*$ for all $n \in \aleph_i$ and all i. Just as a consumer whose consumption basket expands with N, if $\kappa(N)$ is kept constant, a firm pricing more goods has to distribute its attention amongst more shocks, so its information capacity becomes more binding. Thus, the income effect reduces firms' incentives to allocate attention to all shocks.

The third force is the aggregation effect, which is captured by the unspecified functional form of firms' information capacity $\kappa(N)$ in (9). In the following, we study the interaction of these forces after we make a number of alternative assumptions on the aggregation effect.

We start by assuming away any aggregation effect, $\kappa(N) = \kappa$ and dropping good-specific shocks:

Proposition 1 *If good-specific shocks do not exist, $\sigma_z = 0$, or are irrelevant for pricing decisions, $\pi_{15} = 0$, firms' allocation of attention is invariant to N if $\kappa(N) = \kappa$. Moreover, prices of goods produced by the same firm perfectly commove.*

Proof. 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 (12) and the constraint $\kappa_a + \kappa_f = \kappa$:

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

which is constant in N. Moreover, the optimal pricing rule in (15) reduces to

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

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

Intuitively, firms can equally exploit the economies of scope in information processing by paying attention to either monetary or firm-specific shocks for any N. Further, there is no income effect since the number of shocks hitting firms is constant in N. Since $\kappa\left(N\right) = \kappa$, firms' constraint is invariant to N and thus N only affects the scale of the firms' objective.

Proposition 1 is useful for benchmarking. When shocks affect the whole firm, a multi-product firm allocates its attention exactly as a single-product firm. However, the model in this case also predicts no dispersion of price changes within firms because prices do not respond to any good-specific disturbance. As anticipated in the introduction and documented in section 4, this prediction is strongly counterfactual according to both CPI data and PPI data. Hence, we focus on a setup with good-specific shocks.[2] None of our results below in this section relies on a specific process for these good-specific shocks but only on an interior allocation of firms' attention to all shocks. The latter ensures that prices respond to some extent to all three kinds of shocks.

Next, we keep further assume that the responsiveness α of aggregate prices to a monetary shock is exogenously constant while continuing to assume that there is no aggregation affect.

Proposition 2 *If $\kappa\left(N\right) = \kappa$ and α is exogenously constant, firms' attention κ_a^* to monetary shocks is*

[2]The appendix solves for an extension of the model that specifies a common signal for all good-specific shocks that affect a given firm. We show that Proposition 1 holds. This result gives ground to our assumption that signals provide information about only one type of shocks.

increasing in N *for* $N > \hat{N}$ *and* $\kappa_a^* \in [0, \kappa]$*, where* $\hat{N}$ *solves*

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

Proof. α constant implies that $x_1 \equiv \frac{|\hat{\pi}_{11}|\sigma_\Delta}{\hat{\pi}_{14}\sigma_f}$ and $x_2 \equiv \frac{|\hat{\pi}_{11}|\sigma_\Delta}{\hat{\pi}_{15}\sigma_z}$ are also constant. $\hat{N}$ solves $\frac{\partial \kappa_a^*}{\partial N} = 0$ for the interior solution of (14) after setting $\kappa(N) = \kappa$. ■

Proposition 2 states that the economies of scope in information processing dominate the income effect when $N > \hat{N}$. Note that a constant α allows us to abstract away from the feedback between firms' allocation of attention and the responsiveness of aggregate prices to shocks. We introduce such feedback in the next proposition. Here, since $N \in \mathbb{N}$, $\log 2 < 0$ and $x_1, x_2 > 0$, $\hat{N} \geq 1$ only holds if x_2/x_1 and/or x_2 are small enough. Since $x_1 \equiv \frac{|\hat{\pi}_{11}|\sigma_\Delta}{\hat{\pi}_{14}\sigma_f}$ and $x_2 \equiv \frac{|\hat{\pi}_{11}|\sigma_\Delta}{\hat{\pi}_{15}\sigma_z}$, x_2 is small either when $\frac{\sigma_\Delta}{\sigma_z}$ is small, that is, the volatility of good-specific shocks is high relative to the compound aggregate variable Δ_t, or when $\frac{\hat{\pi}_{15}}{|\hat{\pi}_{11}|}$ is large, that is, frictionless prices are highly responsive to good-specific shocks. Similarly, x_2/x_1 is small either when $\frac{\sigma_\Delta}{\sigma_f}$ is small or when $\frac{\hat{\pi}_{15}/|\hat{\pi}_{11}|}{\hat{\pi}_{14}/|\hat{\pi}_{11}|}$ is high, that is, frictionless prices are highly responsive to good-specific shocks relative to firm shocks.

To introduce endogeneity of α, we assume $\kappa(N)$ is unrestricted. This allows us to state a general result.

Proposition 3 *The endogeneity of* α *amplifies the effect of* N *on* κ_a^**. This amplification is stronger when the complementarity in pricing decisions is stronger, that is, when* $\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} \leq 1$ *is smaller.*

Proof. From equation (16), α is increasing in κ_a for $\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} \leq 1$, so Δ_t in (11) and thus σ_Δ are also increasing in κ_a^*; hence x_1 and x_2 are increasing in α. Besides, the interior solution of κ_a^* in (14) is increasing in x_1 and x_2; hence κ_a^* is increasing in α. As a result, the effect of N on κ_a^* in (14) gets amplified by the endogeneity of α captured in (16). According to (16), α is more increasing in κ_a as $\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}$ is smaller, therefore this amplification effect is stronger. ■

Proposition 3 states that firms' optimal attention to monetary shocks (κ_a^*) and the degree of monetary non-neutrality (α) are jointly determined by a fixed point that solves equations (14) and (16). Visually, Figure 1 draws these two equations in the space (α, κ_a). The interior solution of (14) is drawn in red, while (16) is drawn in blue. In addition, the upper bounds of $\kappa_a \in [0, \kappa(N)]$ and

$\alpha \in [0,1]$ are represented by dashed lines. Equilibrium α is denoted as α_1^*.

Equation (16) is invariant to N but the intercept of (14) may decrease or increase while its slope decreases in N. The green line in Figure 1 depicts the case of a higher intercept of (14) as N increases, so equilibrium α is now α_2^*. As a result, the effect of N on κ_a^* in Proposition 2 is amplified by an indirect effect of κ_a^* on σ_Δ^2 in the same direction.

A key observation is that (16) is more flattened out for intermediate values of α than for high and low α. Hence, an increase in κ_a has a large effect on α when α is at an intermediate level. This result is stronger when $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$ is smaller, that is, complementarity in pricing decisions is stronger. The reason is that (16) is flattened out for intermediate values of α as $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$ is smaller. This result plays a crucial role in our quantitative analysis in section 5 when a small increase in firms' attention to monetary shocks yields a large reduction of monetary non-neutrality.

For our next result we assume that the aggregation effect is such that firms' attention to monetary shocks is invariant to N, that is $\kappa_a^*(N) = \bar{\kappa}_a$. This implies that $\kappa(N)$ is increasing for $N < \hat{N}$ and decreasing for $N > \hat{N}$, with $\hat{N}$ defined in Proposition 2. The next proposition shows that this assumption is equivalent to assuming that the friction is more binding as N increases as measured by two alternative measures. One is the frictional cost, which is defined as the expected loss in profits due to the friction per-good and unit of time,

$$C(\kappa_a, \kappa_f, \kappa_n) = \frac{|\widehat{\pi}_{11}|}{2}\left[2^{-2\kappa_a}\sigma_\Delta^2 + \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_f}\sigma_f^2 + \left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_n}\sigma_z^2\right] \tag{17}$$

and the second is the shadow price of information-processing capacity, which is equal to the Lagrange multiplier of the constraint in (9).

Proposition 4 *When $\kappa(N)$ is set such that $\kappa_a^*(N) = \bar{\kappa}_a$, the friction is increasingly more binding, either measured by the frictional cost or the shadow price of information-processing capacity.*

Proof. Using the first order conditions in (12) and (13) and $\kappa_a^*(N) = \bar{\kappa}_a$, the frictional cost in (17) becomes

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

and the shadow price of information-processing capacity

$$\lambda\left(N\right)=-\frac{\beta}{1-\beta}\left|\widehat{\pi}_{11}\right|\log\left(2\right)2^{-2\bar{\kappa}_{a}}\sigma_{\Delta}^{2}N$$

both of which increase linearly with N since σ_{Δ}^{2} is also invariant to N because $\kappa_{a}(N)=\bar{\kappa}_{a}$. ■

The next proposition extends this result to any case in which firms' attention to aggregate shocks decreases in the number N of goods that firms price.

Proposition 5 *Any specification of the model such that $\kappa_{a}^{*}\left(N\right)$ decreases in N implies that the frictional cost and the shadow price of information-processing capacity increase in N.*

Proof. Proposition 4 states that the friction is increasingly binding with N for $N\leq\hat{N}$ even when $\kappa\left(N\right)$ is increased to preserve a constant κ_{a}^{*}. This is an lower bound for the case that this proposition refers to. ■

Proposition 4 and 5 predict that the size of the friction must increase to yield constant or increasing monetary non-neutrality when firms price more goods. We confirm this prediction in our quantitative exercises in section 5. As a matter of fact, in section 5 we find that the severity of the friction must be much higher in our calibrated model with multi-product firms than in a comparable model of single-product firms such that both models yield the same monetary non-neutrality. This severity of the friction must be also higher than what is assumed in alternative models or what has been found in empirical studies.

A natural alternative assumption to discipline the aggregation effect is to assume that the severity of the friction is invariant to the number of goods that firms price. Since we use two measurements of friction, the exact condition we impose takes two alternative forms. In the first, we set the aggregation effect such that the frictional cost is invariant to N at the optimal allocation of attention, that is, the expected loss in profits due to the friction per good is the same in equilibrium for any firm regardless of the number of pricing decision it takes. In the second, the shadow price of information processing capacity is invariant to N at the optimal allocation of attention, that is, firms' incentives to increase their information processing capacity is the same in equilibrium for any firm regardless of the number of pricing decisions it takes.

Both of these alternative assumptions imply a concave relationship between information processing capacity and N. This is because firms can increasingly exploit the economies of scope in information processing as they take more pricing decisions. The next proposition states our main result for this case.

Proposition 6 *When the capacity function $\kappa(N)$ is restricted to keep the friction equally binding for any N, firms' attention $\kappa_a^*(N)$ to monetary shocks unambiguously increases in N. When the frictional cost is invariant to N, $\kappa_a^*(N)$ becomes*

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

When the shadow price of information processing is invariant to N, $\kappa_a^*(N)$ becomes

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

Proof. These expressions follow from the definition of the frictional cost and the shadow price of information capacity along the optimal conditions for the allocation of attention. ■

This proposition provides the intuition for our main quantitative results. For a given extent of the friction, the model underestimates firms' attention to monetary shocks – and thus overstates monetary non-neutrality – under the assumption of single-product firms. This result holds for any specification of the model, even if the specification varies with N. It is only important to have the volatility of monetary shocks σ_q be invariant to N and an interior solution hold for firms' allocation of attention. In our quantitative exercises we pin down these parameters directly from the data.

The specification of the aggregation effect that leads to Proposition 6 has two attractive features relative to those previously imposed: The first is that it provides natural discipline to compare among firms that price a different number of goods. The second is that it allows for internal consistency. In our model, we assume that the number of pricing decisions that firms take is exogenous. If the aggregation effect is such that frictional cost increased in the number of goods priced by a firm, then firms would have incentives to delegate their pricing decisions to smaller decision units. Hence, if the number of pricing decisions were endogenous, our economy would collapse to one in which all goods are priced by single-product firms. Of course, there are more reasons for firms to produce a basket of goods than exploiting economies of scope in information processing. These other reasons are absent in our model due to the assumption that the contribution to profits of each good is independent of the number of goods and which goods a firm produces. We relax this assumption in Appendix C. We find that Proposition 6 still applies: Decision units pay more attention to monetary shocks when they prices more goods. Along the same line, we find that given the friction and the degree of strategic complementarity in the economy, firms that produce more goods pay more attention to monetary shocks.

A similar argument applies to the shadow price of information-processing capacity. In our model, firms's information processing capacity is exogenous. However, if the shadow price of such a capacity were increasing, firms pricing more goods would have larger incentives to invest in this capacity. By contrast, assuming that the shadow price of information processing is invariant to N is equivalent to assuming that the cost of building information capacity does not depend on the number of decisions that can benefit from the processed information.

3.1 Heterogeneous Firms

We now modify our model to allow for the coexistence of firms that price a heterogeneous number of goods. This model produces the same qualitative results as above. However, it allows us to perform a more realistic calibration using moments from PPI data which we compute according to the number of goods per firm.

Thus, consider G groups of firms such that firms in group $g = 1, ..., G$ produce 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 parameters are the same for all groups.

For a given guess $p_t^* = \alpha q_t$, the solution of κ_a^* is still represented by (14), only replacing N by

N_g. This guess is now confirmed for

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

We find that Propositions 1 to 5 continue to hold in this setup. Proposition 6 is modified to

$$\kappa_a^*\left(N_g\right) = \kappa_a^*\left(1\right) + \frac{1}{2}\log_2\left(\frac{N+2}{3}\right), \text{and}$$

$$\kappa_a^*\left(N_g\right) = \kappa_a^*\left(1\right) + \frac{1}{2}\log_2\left(N\right),$$

which is still increasing in N_g although there are two differences with respect to the above result. First, $\kappa_a^*(1)$ is now the attention paid to monetary shocks by single-product firms in an economy with firms pricing different numbers of goods – before it was firms' attention in an economy populated only by single-product firms. Second, the last term on the right-hand side of the equation for $\kappa_a^*\left(N_g\right)$ in Proposition 4 is zero since the volatility σ_Δ now is common to all firms. This modification implies that the effect of N on $\kappa_a^*\left(N_g\right)$ conditional on $\kappa_a^*(1)$ is now less steep. However, $\kappa_a^*(1)$ is higher as the average number of goods priced by a firm increases. This is because $\kappa_a^*(1)$ is increasing in σ_Δ and σ_Δ is higher when the average number of goods priced by a firm is higher.

3.2 Persistent Shocks

We now turn to study the temporal dimension of economies of scope. This dimension affects the allocation of attention when one type of shocks become more persistent than another. For illustrative purposes, we display results for a special case with only monetary and good-specific shocks while allowing for possibly different persistence in the shocks. Appendix A solves a full version of our model that allows for a partial analytical solution with multi-product firms and *persistent* monetary and idiosyncratic shocks broken into firm- and good-specific shocks. The key assumption to get a partial analytical solution in both the full and the simple model is that the endogenous aggregate compound variable Δ_t and good-specific shocks z_t follow $AR(1)$ processes with persistency ρ_Δ and ρ_z respectively.

Since firms are exposed to only two shocks, there is only one first order condition:

$$\kappa_a^* + f(\rho_\Delta, \kappa_a^*) = \kappa_z^* + f(\rho_z, \kappa_z^*) + \log_2 x\sqrt{N}$$

where $x \equiv \frac{|\hat{\pi}_{11}|\sigma_\Delta\sqrt{1-\rho_\Delta^2}}{\hat{\pi}_{15}\sigma_z\sqrt{1-\rho_z^2}}$ and $f(\rho_h, \kappa_h) = \log_2\left(1-\rho_h^2 2^{-2\kappa_h}\right)$ for $h = a, z$.

The term $\sqrt{N}$ on the right hand side of this equation still captures the economies of scope in information processing of multi-product firms. In this subsection we do not vary N but we decrease ρ_z. This is what we do in section 5 when we depart from our benchmark, where serial correlation of monetary and idiosyncratic shocks are the same, to a calibration in which serial correlation of idiosyncratic shocks is much smaller than that of monetary shocks.

A decrease of ρ_z has two effects. First, σ_z decreases relative to σ_Δ which cancels out with the direct effect of ρ_z on the parameter x. Therefore, this force has no effect on (κ_a^*, κ_z^*). In our quantitative exercises in section 5 we need to adjust the volatility of innovations in a given shock when we adjust the persistence of the shock to match a measure of empirical variability of price changes. The second force is the temporal dimension of the economies of scope, which is captured by the increase in the term $f(\rho_z, \kappa_z^*)$. As ρ_z decreases, paying attention to Δ_t is more useful for future pricing decisions rather than to z_{jt}. This force increases κ_a^* relative to κ_z^*. As a result, a decrease in ρ_z increases κ_a^*. There is less monetary non-neutrality.

4 Empirical Results

This section provides empirical evidence for our assumptions of our analysis of firms: the multi-product nature of firms and the response of prices to good-specific disturbances. We also report several new statistics from the CPI and PPI data. We later use these to calibrate our model, interpreting firms as either "stores" or "good producers."

4.1 Data Description

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).[3] 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.

Our CPI data span the time period from 1988 to 2009, containing approximately $125,782$ outlets. An outlet corresponds to a non-producing retailer or, in colloquial terms, a store. Our main manipulation of the CPI data is to exclude sales and zero price changes. The exclusion of sales is common in studies of price setting: sales are usually considered practices of firms that are not necessarily related to the business cycles (for instance, see Guimaraes and Sheedy (2011)). We exclude zero price changes for two reasons: First, this is more consistent with our model, in which prices are fully flexible. Second, Mackowiak and Wiederholt (2009), which we use as benchmark in section 5, calibrate their model to statistics excluding sales and zero price changes.

Our PPI sample spans the time period from 1998 to 2005, containing approximately $28,575$ firms. A "firm" is typically a goods producer which is defined as a decision unit that prices a given number of goods. Again, we exclude zero price changes, but we do not control for sales. This practice is sufficiently less common for firms in this dataset to leave results unchanged.[4]

4.2 Multi-Product Firms and Good-Specific Shocks

Multi-product firms. In our CPI sample, the median (mean) number of goods sampled from a single outlet is 1.39 (2.05) with a standard deviation of 2.03 goods.[5] In these data, 87% (75%) of outlets have less than 3 (2) goods in the sample. Given that outlets are usually stores, we conclude

[3] Nakamura and Steinsson (2008) or Bils and Klenow (2004) describe the CPI data in detail, while for example Bhattarai and Schoenle (2011) describe the PPI data.

[4] As work by Nakamura and Steinsson (2008) has shown, sales are not a factor in determining the behavior of PPI prices. We have computed our statistics excluding sales prices, and have found no significantly different results.

[5] 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.

that the CPI data does not provide a reliable estimate of the number of goods that a single outlet prices. We therefore also consider any moments computed by bin from this dataset as useless for calibrations; we only use moments computed for the whole sample. However, we nonetheless present the empirical moments by bins in the CPI data as complementary evidence.

To obtain a more accurate estimate of the number of goods in stores, we use the Food Marketing Institute (FMI) 2010 Report. The FMI is an industry-based institution that represents $1,500$ food retailers and wholesalers in the U.S.. Their members are large multi-store chains, regional firms and independent supermarkets, stores and drug stores of a large variety of classes with a combined annual sales volume of 680billion.[6] The FMI reports an average of 38,718 items per store.[7] Additional evidence is provided by Rebelo et al. (2010) who use data from one particular store that prices about 60,000 items. In any case, as anticipated in the introduction, the quantitative effect of multi-production is so strong in our exercises in section 5 that we do not need to pin down a precise estimate of the number of goods priced by stores. It suffices to establish that this number is large.

In contrast to CPI data, we do consider the number of goods priced by firms in the PPI as valid. This allows us to relate that number to changes in key price statistics and calibrate our model accordingly. To estimate this number, we follow Bhattarai and Schoenle (2011) by counting the number of goods per firm has. We consider this as a lower bound due to sampling constraints. However, what is important for the representativeness of trends across bins is that there is a monotonic mapping from the actual number of goods to our sample, as Bhattarai and Schoenle (2011) discuss in detail. Moreover, our data contain substantial variation in the number of goods per firm such that we consider any bin-specific statistic as quite reliable. The median (mean) number of goods per firm is 4 (4.13) with a standard deviation of 2.55 goods. The median number of goods per firm are 2 (bin 1), 4 (bin 2), 6 (bin 3), and 8 (bin 4).

A rough alternative estimate for the number of goods priced by a firm 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

[6] http://www.fmi.org/about-us/who-we-are. For a detailed list of stores included, see http://www.fmi.org/about-us/our-members

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

average number of products produced by a multi-product firm is 3.5.

A final remark is that although the multi-product nature of firms is a well-established empirical fact, we can only indirect verify that prices are set by multi-product decision units and not individually by separate agents for each good. First, we know that data are carefully collected by the BLS such that a firm is a "price-forming unit." Second, Zbaracki et al. (2004) conduct a case study that documents in great detail the decision process of setting prices of a productive firm. They report that the firm they study has multiple products, whose regular prices are decided at headquarters; sales prices are set by local managers in small geographical areas. However, at both levels there is a single decision unit setting prices for the whole portfolio of goods. We consider this as evidence that prices are indeed set by multi-product decision units.

Within-firm dispersion of price changes. We construct a measure of the ratio of within-firm dispersion of log non-zero price changes relative to total cross-sectional dispersion of log non-zero price changes. In ANOVA terminology, this is the ratio of the SSW to the SST. We denote this statistics as follows by r:

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

where $\overline{\pi}_{it}$ is the mean absolute size of log price changes across all goods sampled for firm i at time t and $\bar{\pi}_t$ is the grand total mean.[8]

Computation of this statistic leads to our most important empirical result. In the CPI data, 51.6% of the cross-sectional dispersion of log price changes is due to within-firm dispersion. In the PPI data, this ratio is increasing as firms produce more goods, from 36.5% (for bin 1, where firms produce between 1 and 3 goods) to 72.4% (for bin 4, where firms produce more than 7 goods). In the full PPI sample, 59.1% of the total variance is due to within-firm variance.

In the remainder of the paper, we sometimes refer to this finding as "imperfect co-movement of price changes." We interpret this result as evidence of the existence of good-specific shocks that

[8] 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.

firms must take into account in their pricing decisions. As we anticipate in the introduction and in our theoretical analysis, our results are almost invariant to the assumption of what fraction of this dispersion is due to good-specific shocks.

4.3 Aggregate and Idiosyncratic Inflation Components

Additional evidence for the importance of aggregate and idiosyncratic shocks as a function of the number of goods comes from an extension of the analysis in Boivin et al. (2009) to the multi-product dimension of firms. The idea is to separate, for each bin, price changes into a common aggregate component and a component that is specific to that bin. Then, one can study the behavior of these two components as a function of the number of goods in the data.

To do so, we first compute a monthly bin-specific, average good-level inflation rate $\pi_{b,t} = \frac{1}{N_b} \sum_i \log\left(\frac{p_{i,b,t}}{p_{i,b,t-1}}\right)$ where N_b denotes the number of goods i per bin b. We then split each bin-specific inflation rate into a common and an idiosyncratic component:

$$\pi_{b,t} = \lambda_b C_t + \epsilon_{b,t} \tag{18}$$

where we have used the common aggregate factors C_t that Boivin et al. (2009) construct for the U.S. PPI. The estimated residuals $\epsilon_{b,t}$ in turn represent inflation that is specific to each bin.

We then extend the basic analysis in Boivin et al. (2009) by separately recording the R^2 from estimating eqn.(18) for each bin, and computing - also by bin - the unconditional volatility and persistence of an AR(13) of each of the two components.

We find that the explanatory power of the aggregate component has a low level, but increases from 6.4% to 9% as the number of goods increases. The estimated persistence exhibits opposing trends for the two components: it increases from 0.5 to 0.8 for the aggregate component while it decreases from 0.39 to 0.24 for the idiosyncratic component. The unconditional volatility of each component shows no clear trend with the number of goods. However, the volatility of the idiosyncratic component is larger by a factor of approximately 3, consistent with a low R^2 due to the aggregate component. We summarize these results in Table 3.

4.4 Complementary Statistics

Absolute size of price changes. We compute the average size of absolute price changes. denoting this statistic as $|\overline{\pi}|$. Labelling time as t, firms as i and goods produced by firm i at time t as $n \in \aleph_{it}$,

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

where $\pi_{nt} \equiv p_{nt} - p_{nt-1}$ is non-zero inflation 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. Thus, we first compute for each good the magnitude of price changes, conditional on non-zero price changes. Second, we compute firm-level averages. Finally, we take the mean across all firms in the full sample, or within one of the bins.

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%).[9] 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%. This trend shows that the multi-product nature of price changes is strongly related to the size of price changes: as the number of goods increases, the magnitude of price changes becomes smaller. Various robustness checks in the PPI data, reported in Bhattarai and Schoenle (2011), leave this result unchanged.

Serial correlation of price changes. We denote this statistic 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 for the parameter of an $AR(1)$ coefficient for $\pi_{n,k,t}$, conditional on non-zero price changes. We compute the median quantile regression by estimating the following specification:

$$\hat{\rho}_k = argmin_{\rho_k} E[|\pi_{n,k,t} - \rho_k \pi_{n,k,t-1}|]$$

We find that the median estimate of the AR(1) coefficient is -0.29 in the CPI sample. Bils and Klenow (2004) estimate a comparable first-order serial correlation of -0.05.[10] In the PPI data, our

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

[10] 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

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.

Cross-sectional dispersion of price changes. This statistic is denoted as $\widetilde{\sigma}$ and defined as

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

where $\overline{\pi}_t$ is the average of non-zero absolute log price changes π_{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. As Table 1 shows, the cross-sectional dispersion is 3.51% (2.65%) in the full PPI (CPI) sample. There is no clear trend in the PPI data.

5 Quantitative Results

We use these above moments to calibrate a generalized version of our model that allows for persistent monetary and idiosyncratic shocks. The latter are split into firm- and good-specific shocks.[11] Our goal is to explore the quantitative effect of the economies of scope of information processing: first, due to the multi-product nature of firms, and second due to different persistence of monetary and idiosyncratic shocks.

5.1 Baseline Calibration

We start by replicating the results of Mackowiak and Wiederholt (2009) for an economy of single-product firms. We subsequently use this as a benchmark, to which we add calibration targets from the data and incorporate the multi-product nature of the firm. We nest the calibration of Mackowiak and Wiederholt (2009) in our model by setting

$$N=1;\kappa\left(1\right)=3;\frac{\widehat{\pi}_{13}}{\left|\widehat{\pi}_{11}\right|}=0.15;\frac{\widehat{\pi}_{14}}{\left|\widehat{\pi}_{11}\right|}=0;\frac{\widehat{\pi}_{15}}{\left|\widehat{\pi}_{11}\right|}=1.$$

by focusing on the period from 1989 to 2009. Qualitatively, both approaches give the same results.

[11] We discuss the problem of the firm and its numerical solution algorithm in the appendix.

First, setting capacity $\kappa(1) = 3$ implies a small frictional cost of 0.21% of firms' steady state revenues. Second, the complementarity in pricing decisions $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} = 0.15$ is in the lower bound of the range suggested by Woodford (2003). It is also exactly what Mackowiak and Wiederholt (2009) assume. Third, they refer to idiosyncratic shocks as firm-specific in their model. However, since their model has only single-product firms, idiosyncratic shocks in their model are indistinguishably firm-specific or good-specific shocks in our model. Therefore, to replicate their exercise, we must shut down one of them. We begin our analysis assuming that these shocks are good-specific, $\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} = 0$. Fourth, we set idiosyncratic and monetary shocks to be equally important for profits, so $\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} = 1$. This parameter enters in the model through $x_2 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta}{\widehat{\pi}_{15}\sigma_z}$, so setting $\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} = 1$ implies that $\frac{\sigma_\Delta}{\sigma_z}$ must be pinned down from the data.

To obtain σ_Δ, we also follow Mackowiak and Wiederholt (2009). We estimate an $AR(1)$ process for GNP quarterly data spanning 1959:1–2004:1 to obtain the volatility and persistence of q_t, $\sigma_q = 2.68\%$ and $\rho_q = .95$. Then, for computational simplicity, we approximate this process by an $MA(20)$:

$$q_t = \sum_{k=0}^{20} \left(1 - \frac{k}{20}\right) v_{t-k} \tag{19}$$

where $v_t \sim N(0,1)$ and coefficients linearly decrease with the order of lags up to 20 lags. Hence an innovation in nominal aggregate demand dies out after 21 periods. Given the process for q_t, the compound aggregate variable Δ_t also follows a $MA(20)$:

$$\Delta_t = \sum_{k=0}^{20} \left[\left(1 - \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\right) \alpha_k + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} \left(1 - \frac{k}{20}\right) \rho_q \right] v_{t-k}$$

where $\{\alpha_k\}$ are the parameters of the guessed process of aggregate prices, which is also $MA(20)$:

$$p_t = \sum_{k=0}^{20} \alpha_k v_{t-k} \tag{20}$$

such that $\{\alpha_k\}$ are found in equilibrium. We provide a detailed explanation in the appendix.

For idiosyncratic volatility σ_z, we assume that these shocks follow an $MA(20)$ similar to (19) with an adjusted scale of coefficients to match the 9.6% average absolute per-good inflation reported by Klenow and Kryvtsov (2008) for CPI data in the US. This implies $\sigma_z = 11.8\sigma_q$.

We then replicate the results in Mackowiak and Wiederholt (2009):[12] Firms' attention is $\kappa_a^*(1) = 0.09$ to monetary shocks and $\kappa_z^*(1) = 2.91$ to idiosyncratic shocks. This yields large and long-lasting monetary non-neutrality. Figure 2 depicts the response of aggregate prices after an innovation of 1% in q_t. The black line draws the response of frictionless prices; this response inherits the process assumed for q_t in (19). The blue line draws the response of aggregate prices under rational inattention. On impact, prices absorb only 2.8% of the innovation in q_t. Their response remains sluggish relative to the response of frictionless prices for 20 periods (the output deviation is less than 5% of the shock thereafter) and the cumulated response of prices is only 22% of the cumulated response of frictionless prices. As anticipated above, the frictional cost is 0.21% of the firm's steady state quarterly real revenue $\overline{Y}$. This cost is conventionally considered small. It would give little incentive to firms to increase their information capacity if such decision were endogenous. Importantly, these single-product results confirm Sims' statement about the ability of the Rational Inattention model to generate large macroeconomic effects even with a small friction.

5.2 Multi-Product Firms

We now extend the baseline calibration, allowing firms to produce $N > 1$ goods. Again, we target an average absolute size of price changes of 9.6% and a frictional cost of $0.21\%\overline{Y}$ per good. What is the right choice of N? We know from Section 4 that the CPI data does not provide sufficient variation to calibrate N. However, indirect estimates indicate $N \approx 40,000$ (see the Food Marketing Institute's 2010 report). Since the effect of multi-production is already very strong for $N = 2, 4$ and 8, we report results for these cases only.

We find that already for a two-good firm, monetary non-neutrality is cut by three in magnitude and duration. We illustrate in Figure 2 in red the resulting response of prices to a monetary shock, when for $N = 2$ firms' attention is $\kappa_a^*(2) = 0.36$ and $\kappa_z^*(2) = 2.92$. Strikingly, the response of prices is almost identical to the frictionless price response after 7 periods (the output deviation is less than 5% of the shock thereafter), and their cumulated response is 74% that of frictionless

[12] 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 5.5 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.

prices. Prices absorb 29% of the innovation in q_t on impact.

We also show the response of prices to an aggregate shock for a four-good firm in Figure 2 in green. We find that $\kappa_a^*(4) = 0.58$ and $\kappa_z^*(4) = 2.90$. Prices absorb 15% of the shock on impact. Overall, their response closely follows that of frictionless prices after 4 quarters with almost no real effects thereafter. Their cumulated response is 86% that of frictionless prices. For $N = 8$, in magenta in figure 2, results are even stronger: $\kappa_a^*(8) = 0.90$ and $\kappa_z^*(8) = 2.87$, prices absorb 49% of the shock on impact, the output deviation is less than 5% of the shock after 2 quarters and the cumulated response of prices is 93% that of frictionless prices.

Note that all these result holds when a firm's attention to monetary shocks is only a small fraction of the firm's total capacity. The strong effect is due to complementarity in pricing decisions, as stated by Proposition 3. Given these results, we find it uninformative to report results for $N = 40,000$: According to a standard rational inattention model calibrated to stores, money is almost fully neutral.

5.3 Serial Correlation of Price Changes

We now calibrate the persistence of idiosyncratic shocks z_{jt} to match the persistence observed in the CPI data. Again, we find that this substantially reduces monetary non-neutrality. What do we calibrate good-level persistence to? So far we have followed Mackowiak and Wiederholt (2009) by assuming that z_{jt} is as persistent as q_t. Instead, we pick a much lower value for the persistence parameter. According to Bils and Klenow (2004), the first-order serial correlation of per-good inflation is -0.05. Our own computation is -0.29 (see Table 1). Both computations are methodologically different,[13] but both suggest that idiosyncratic shocks are substantially less persistent than monetary shocks.

We therefore set z_{jt} to follow an MA process for which the coefficients linearly decrease with the order of lags, as for q_t in equation (19). However, to match the -0.05 first-order serial correlation, z_{jt} must follow a $MA(5)$. We must also set $\sigma_z = 10.68\sigma_q$ to match the average absolute per

[13] 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.

good inflation. To generate a -0.29 first-order serial correlation, z_{jt} must follow a $MA(1)$ with coefficient 0.33 and $\sigma_z = 9.74\sigma_q$. We also keep targeting $0.21\%\overline{Y}$ of per-good frictional cost.

We focus on results for the case $N = 1$. Figure 3 summarizes the response of prices to a 1% innovation in q_t. The black and blue lines show the response of frictionless prices, and prices under rational inattention for the benchmark calibration. The red line draws the response of prices calibrated to a -0.05 serial correlation. We find that $\kappa_a^*(1) = 0.20$ and $\kappa_z^*(1) = 2.81$. This implies that the response of prices on impact is 7% of the shock and the deviation of output is less than 5% of the shock after 12 periods, and the cumulated response of prices is 52% of the frictionless price response. The green line shows the response of prices calibrated to -0.29 serial correlation of price changes. Now $\kappa_a^*(1) = 0.19$ and $\kappa_z^*(1) = 2.66$. This implies that the response of prices on impact is 7% of the shock, the output deviation is less than 5% of the shock after 12 periods, and the cumulated response of prices is 52% that of frictionless prices.[14]

We conclude that the monetary non-neutrality predicted by the model is substantially reduced even for $N = 1$ when we depart from our benchmark by calibrating the model to match the serial correlation of good-level price changes found in the data. The intuition is the same as in our theoretical analysis in section 3.2: Incentives to pay attention to a shock are lower for a less persistent shock: the processed information about the realization of this shock is less useful for future pricing decisions. This is the temporal dimension of the economies of scope in information processing we referred to in the introduction and in section 3.2.

5.4 Within-Firm Dispersion of Price Changes

We now introduce firm-specific shocks. This is necessary in order to match the observed imperfect co-movement of price changes observed within firms. Matching this additional target is only possible in a model with all three kinds of shocks: firm-specific, good-specific and aggregate (in our case, monetary) shocks. Again, we find that monetary non-neutrality quickly vanishes as N increases.

To perform this exercise, we target the ratio of within-firm dispersion to total cross-sectional

[14] If we used OLS instead of quantile regressions to estimate an $AR(1)$ process for price changes, the estimate would be -0.22. The implied calibration results are very similar to those reported here.

dispersion of non-zero absolute price changes in the CPI. In our exercises above with no firm-specific shocks, this statistic was 50% for $N = 2$, 86% for $N = 4$ and 93% for $N = 8$. Our target now is a ratio of 51.6%, as shown in Table 1.

To choose the relative volatility of firm-specific and good-specific shocks, we assume that $\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} = 1$; that is, firm-specific shocks f_{it} have the same weight in firms' profits as aggregate and good-specific shocks, q_t and z_{nt}.[15] For any number of goods N, we must set the process of firm-specific and good-specific shocks to follow an $MA(1)$ with parameter 0.33 to match -0.29 serial correlation of per-good inflation. The total volatility of these shocks to match 9.6% average absolute per-good inflation must be $9.75\sigma_q$ for $N = 2$, $11.7\sigma_q$ for $N = 4$, and $11.18\sigma_q$ for $N = 8$. To match the 51.6% ratio of within-firm dispersion, the calibration of σ_f/σ_z is also specific to the number of goods. For $N = 2$, we set $\sigma_f = 0$ since the highest within-firm dispersion ratio we can generate is 50%, so results for this case are the same as in section 5.3. We set $\sigma_f = 1.37\sigma_z$ for $N = 4$ and $\sigma_f = 1.90\sigma_z$ for $N = 8$. Finally, we calibrate $\kappa\,(N)$ to yield a $0.21\%\overline{Y}$ per-good frictional cost as in our previous exercises.

We find that for a four-good firm, the allocation of attention becomes $\kappa_a^*\,(4) = 0.61$, $\kappa_f^*\,(4) = 3.27$ and $\kappa_z^*\,(4) = 2.05$ respectively for monetary, firm-specific and good-specific shocks. This implies that aggregate prices absorb 30% of a monetary shock on impact. The deviation of output is less than 5% of the shock after 4 periods, and the cumulated response of prices is 87% of the frictionless price' response. For an eight-good firm, $\kappa_a^*\,(8) = 0.96$, $\kappa_f^*\,(8) = 3.85$ and $\kappa_z^*\,(4) = 1.90$, and prices absorb 52% on impact. The deviation of output is less than 5% of the shock after 2 periods, and the cumulated price response is 94% of the frictionless price response.

In a nutshell, our model still predicts that monetary non-neutrality is negible for price setters deciding prices for more than eight goods, such as stores.

For robustness, we re-do the analysis for $N = 4$ but calibrate the volatility of good-specific shocks to match a within-firm dispersion of price changes of only 10%. This number is much smaller than what we find in the data, leaving room for alternative explanations for this empirical fact. Figure 4 displays the results. The response of prices to a monetary shock for $N = 4$ is almost identical to the case in which all within-firm dispersion of price changes stems from good-specific

[15] Similarly than for x_2, $\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}$ enters in the model's predictions through $x_1 \equiv \frac{|\widehat{\pi}_{11}|\sigma_\Delta}{\widehat{\pi}_{14}\sigma_f}$.

shocks. As anticipated in the introduction and in our theoretical analysis in section 3, our results do not depend on the exact importance of good-specific shocks. Instead, what matters is that prices respond to good-specific shocks to some extent.

5.5 Calibration to PPI Data

Here, we calibrate our model introduced at the end of section 3. It allows for heterogeneity in the number of goods across firms, and persistent monetary, firm- and good-specific shocks. We concentrate on the ability of the model to replicate moments from the data for our four bins and the monetary non-neutrality predicted once the model is calibrated to these moments. The PPI data naturally lends itself to calibrate this model because of the variation in the number of goods in that data. Since this is also a completely different dataset than the one used by Mackowiak and Wiederholt (2009), we no longer use their work as benchmark. However, we keep our target of 0.21% of steady state revenues for the frictional cost. We also use a modified numerical solution algorithm from the one used by Mackowiak and Wiederholt (2009). Our algorithm is more precise but also finds a solution with less attention to monetary shocks, so it goes against our results.

We calibrate our model with four groups of firms (each group pricing a given number of goods) to match moments by bins in the PPI sample. As discussed in section 4, our PPI sample has sufficient variation in N. This allows us to compute moments conditional on various N. In our model, we assume that there are four groups of firms, producing $2, 4, 6$ and 8 goods – the median number of goods per firm in each bin. The relative weights of these groups in the model economy are the shares of total employment in each bin. We keep our baseline calibration except for the processes of firm-specific and good-specific shocks. As above, we calibrate these processes to match three statistics: the mean absolute size of non-zero log price changes, the dispersion ratio of per-good inflation, and the first-order serial correlation of non-zero log price changes.

We start our analysis by asking if one set of parameters can explain trends in the data as the number of goods increases. To do so, we assume that the processes for firm-specific and good-specific shocks are the same in all bins. This means that we calibrate these two shock processes to match the three targets for bin 1 only. We then assume that the processes for the other bins follow the same calibrated processes. We report the model-predicted moments in italics in Table 4 and contrast them with the moments from the data. The model fails to account for the observed moments in other bins. While the average absolute size of log price changes and the serial correlation of non-zero log price changes are invariant to N in the model, they are decreasing in the data. Moreover, the ratio of within-firm dispersion is less increasing in N in the model than in the data.

Next, we calibrate firm-specific and good-specific shocks independently for each group to match the statistics for all bins. Figure 4 illustrates the response of prices to an aggregate shock. The response of prices is not very different across bins. This is due to strategic complementarity in pricing decisions within and across bins. Aggregate prices absorb on impact 16.7% of the monetary shock, output is less than 5% higher than its steady state after 7 periods, and the cumulated response of aggregate prices is 75% of the frictionless price response. This leads us to conclude that monetary shocks have a strong effect on impact when we interpret firms in the model as goods producers, while the persistence of these shocks is limited.

5.6 The Role of Complementarity in Pricing

What is the role of strategic complementary in our calibrated model? We examine this aspect by varying the parameters that govern strategic complementarity in the model. We do so by increasing $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$ from 0.15 to 0.85. This modification has two effects. On the one hand, an increase in firms' attention to monetary shocks has a milder effect on reducing monetary non-neutrality when $\frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}$ is higher. This comes from Proposition 3. On the other hand, for a given level of firms' attention to monetary shocks, monetary non-neutrality is lower when the extent of complementarities is lower. This result comes from equation (16).

Overall, our model calibrated to all bins in the PPI data implies the following: Aggregate prices absorb 23% of the aggregate shock on impact, there are almost no real effects after only 6 periods, and the cumulated response of prices is 84% that of frictionless prices.

5.7 Calibrating Information Capacity

One important parameter in our model is the total information capacity $\kappa(N)$. So far, we have pinned it down by imposing a constant frictional cost of 0.21% per good in terms of expected profit losses. We now depart from this assumption: Since a direct attempt at calibrating $\kappa(N)$ reveals a high insensitivity to moments from the micro data, we instead use the shadow value of information-procession capacity to pin down $\kappa(N)$. As an alternative to calibrate the information friction, we also target the friction implied by the menu costs in Midrigan (2011).

First, we attempt to calibrate $\kappa(N)$ directly from the data by targeting an additional moment: The total cross-sectional dispersion of absolute log price changes in the CPI sample. We then take our model from section 5.4 for $N = 2$ as well as $N = 4$ and solve it for a grid of $\kappa\,(N)$. We report results in Tables 5 and 6.[16] We find that we are unable to pin down $\kappa(N)$ through this approach. The reason is that the predictions of the model regarding moments that can be contrasted with the micro data are highly insensitive to changes in $\kappa(N)$. At the same time, the predictions regarding monetary non-neutrality are highly sensitive to changes in $\kappa(N)$. This invalidates this approach of calibrating $\kappa(N)$. While we could relax the assumption of Gaussian shocks to match our additional moment, this would would add new parameters to calibrate and would not solve the problem illustrated here.

Indeed, as the above suggests, the trade-off between monetary non-neutrality and the severity of the friction is quite substantial: When we calibrate our model from section 5.4 to the PPI moments of the median four-good firm, 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 the friction in Figure 6. Monetary non-neutrality is monotonically increasing in the friction.

For these reasons, we choose to pin down $\kappa(N)$ by keeping the shadow value of information capacity constant across goods. That is, we keep the Lagrange multiplier on the capacity constraint (9) constant across goods: $\lambda(N) = \bar{\lambda}$. We implement this in our model from section 5.4, again targeting an average of log absolute price changes of 9.6%, a serial correlation of -0.291 and a within-firm variance ratio of 51.6%. We find that the cumulated price response lies in between 52% and 83% of the frictionless price response while the friction ranges between 0.20% and 0.24%. Monetary neutrality is extremely high. It increases with the number of goods, exactly as predicted

[16] Similar results hold when we try to pin down the lagrange multiplier $\lambda(N)$ this way, so we do not choose to display these tables.

by our model.

Finally, as an alternative to pin down the information capacity, we force our model of section 5.5 to generate the same frictional cost as in the two-good menu cost model of Midrigan (2011) calibrated using the distribution of prices for a given store. There, the cost of the friction is 0.34% of steady state revenues. If we force our model, calibrated to PPI data, to generate this level of cost, aggregate prices absorb 8.44% of the shock on impact. The deviation of output then is less than 5% of the shock after 16 periods, and the cumulated response of aggregate prices is 48% that of frictionless prices.

From these sensitivity tests on calibration the friction $\kappa(N)$, we conclude that there is a trade-off between the strength of the friction and monetary non-neutrality. We cannot calibrate it using micro moments, however: the predicted relevant moments are insensitive to $\kappa(N)$.

6 Conclusion

In this paper, we have explored the impact of monetary policy on the real economy in a model of rational inattention that accounts for the multi-product nature of firms and introduces good-specific shocks. The real impact of monetary policy is much lower than in a less realistic model in which firms produce only one good. This result is due to economies of scope in information processing: As firms produce more goods, the return to gathering information on common monetary, rather than good-specific shocks increases. When we calibrate our multi-product firm model to U.S. CPI data, we find that monetary policy has minimal real effects. Calibrating our model to PPI data, in which firms price a much smaller number of goods, suggests only limited non-neutrality and aggregate inertia.

There is also a temporal dimension in which these economies of scope operate in the model: shocks can differ in their persistence. The returns to gathering information about one type of shock increase as this shock is more persistent. Therefore, this temporal dimension implies less monetary non-neutrality in the model when idiosyncratic shocks are less persistent than aggregate shocks, as the CPI and PPI data suggest. Although the effect is substantial, this temporal dimension is quantitatively less important than the multi-product dimension.

References

ALVAREZ, F. AND F. LIPPI (2013): "Price Setting with Menu Cost for Multi-Product Firms," Forthcoming, Econometrica.

BERNARD, A., S. REDDING, AND P. SCHOTT (2010): "Multi-Product Firms and Product Switching," *American Economic Review*, 100, 70–97.

BHATTARAI, S. AND R. SCHOENLE (2011): "Multiproduct Firms and Price-setting: Theory and Evidence from U.S. Producer Prices," Globalization and Monetary Policy Institute Working Paper 73, Federal Reserve Bank of Dallas.

BILS, M. AND P. J. KLENOW (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112, 947–985.

BOIVIN, J., M. P. GIANNONI, AND I. MIHOV (2009): "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data," *American Economic Review*, 99, 350–84.

CHEREMUKHIN, A., P. RESTREPO-ECHAVARRIA, AND A. TUTINO (2012): "The Assignment of Workers to Jobs with Endogenous Information Selection," 2012 Meeting Papers 164, Society for Economic Dynamics.

GUIMARAES, B. AND K. D. SHEEDY (2011): "Sales and Monetary Policy," *American Economic Review*, 101, 844–76.

KLENOW, P. J. AND O. KRYVTSOV (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., J. NIE, AND E. R. YOUNG (2012): "Robustness, Information–Processing Constraints, and the Current Account in Small Open Economies," *Journal of International Economics*, 88, 104–120.

MACKOWIAK, B. AND M. WIEDERHOLT (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769–803.

MACKOWIAK, B. A. AND M. WIEDERHOLT (2011): "Inattention to Rare Events," CEPR Discussion Papers 8626, C.E.P.R. Discussion Papers.

MATEJKA, F. AND A. MCKAY (2011): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," CERGE-EI Working Papers wp442, The Center for Economic Research and Graduate Education - Economic Institute, Prague.

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. AND T. WU (2010): "The Puzzling Evolution of the Home Bias, Information Processing and Financial Openness," *Journal of Economic Dynamics and Control*, 34, 875–896.

NAKAMURA, E. AND J. STEINSSON (2008): "Five Facts about Prices: A Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 123, 1415–1464.

PACIELLO, L. AND M. WIEDERHOLT (2011): "Exogenous Information, Endogenous Information and Optimal Monetary Policy," EIEF Working Papers Series 1104, Einaudi Institute for Economic and Finance (EIEF).

PENG, L. AND W. XIONG (2006): "Investor Attention, Overconfidence and Category Learning," *Journal of Financial Economics*, 80, 563–602.

REBELO, S., N. JAIMOVICH, AND M. EICHENBAUM (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. AND Y. WEISS (1992): "Staggered and Synchronized Price Policies under Inflation: The Multiproduct Monopoly Case," *Review of Economic Studies*, 59, 331–59.

SIMS, C. A. (2006): "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review*, 96, 158–163.

VENKATESWARAN, V. AND C. HELLWIG (2009): "Setting the Right Prices for the Wrong Reasons," *Journal of Monetary Economics*, 56, S57–S77.

WOODFORD, M. (2003): "Optimal Interest-Rate Smoothing," *Review of Economic Studies*, 70, 861–886.

——— (2009): "Convergence in Macroeconomics: Elements of the New Synthesis," *American Economic Journal: Macroeconomics*, 1, 267–79.

——— (2012): "Inflation Targeting and Financial Stability," NBER Working Papers 17967, National Bureau of Economic Research, Inc.

ZBARACKI, M. J., M. RITSON, D. LEVY, S. DUTTA, AND M. BERGEN (2004): "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *The Review of Economics and Statistics*, 86, 514–533.

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 variance ratio of $\|\Delta p\|$	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 variance ratio of $\|\Delta p\|$	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 variance 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 Variance Ratio, Robustness

CPI	1-3 Goods	3-5 Goods	5-7 Goods	>7 Goods	All
Within variance 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%)
PPI					
Within variance 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. We compute the monthly within variance 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 then report the time-series mean and medians. Standard errors across monthly means are given in brackets.

Table 3: Aggregate and Idiosyncratic Inflation Components by Number of Goods

	R^2	Persistence		Volatility in %	
Number of Goods		Aggregate	Idiosyncratic	Aggregate	Idiosyncratic
$1-3$ Goods	6.41%	0.50	0.39	0.15%	0.59%
$3-5$ Goods	7.29%	0.76	0.35	0.19%	0.67%
$5-7$ Goods	7.20%	0.76	0.21	0.18%	0.65%
>7 Goods	9.04%	0.80	0.24	0.18%	0.55%

NOTE: As explained in the text, we fist compute bin-specific inflation $\pi_{b,t}$ as the average of item-level log price changes for each bin. We next estimate $\pi_{b,t} = \lambda_b C_t + \epsilon_{b,t}$ where C_t denotes the common component estimated by Boivin et al. (2009) and $\epsilon_{b,t}$ the idiosyncratic component. The R^2 statistic measures the fraction of the variance of $\pi_{b,t}$ explained by the common component; persistence is based on an estimated AR processes with 13 lags.

Table 4: 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-firm variance ratio, data	36.5%	54.6%	67.2%	72.4%	59.1%
Within-firm variance ratio, model	*36.5%*	*54.5%*	*60.5%*	*63.5%*	*53.8%*

NOTE: We report moments predicted by the model in section 5.5 in italics. We contrast them with the moments from the data presented in Table 1.

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

	data	$\kappa = 5$	$\kappa = 6$	$\kappa = 7$	$\kappa = 8$	$\kappa = 9$	$\kappa = 10$	$\kappa = 30$
absolute 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%
(relative to frictionless prices)								

NOTE: As discussed in section 5.7, 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	$\kappa = 10$	$\kappa = 11$	$\kappa = 12$	$\kappa = 13$	$\kappa = 14$	$\kappa = 15$	$\kappa = 30$
absolute 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%
(relative to frictionless prices)								

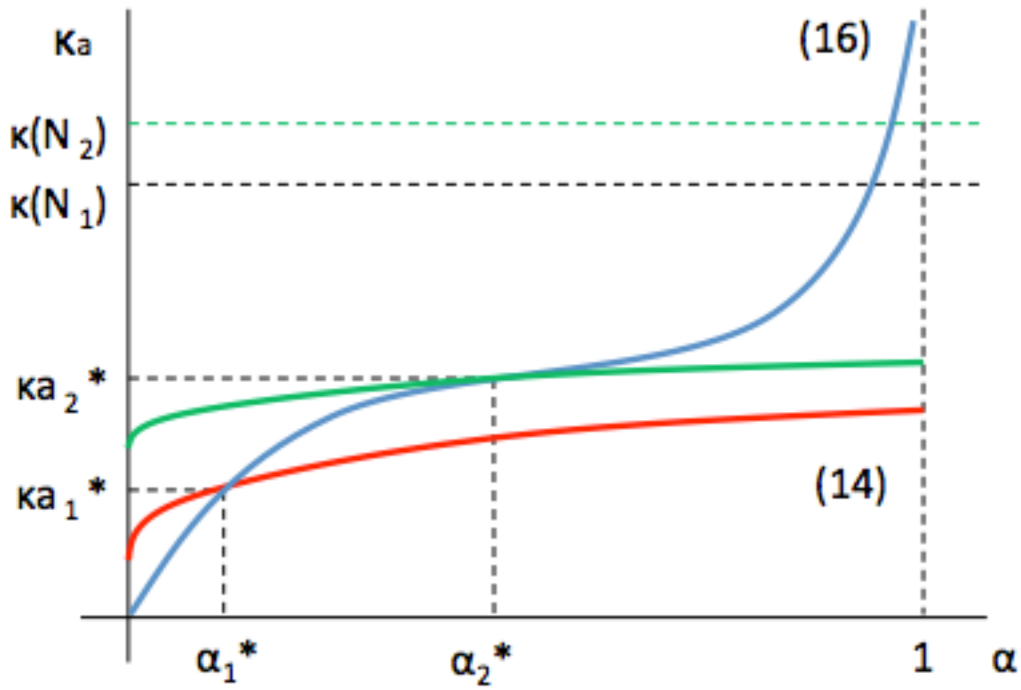
NOTE: As discussed in section 5.7, the table shows moments computed from the data and their counterparts generated by the model for N=2 using different values for firms' capacity to process information.

Table 7: 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 variance 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%
(relative to frictionless prices)				
loss	0.21%	0.20%	0.24%	0.21%

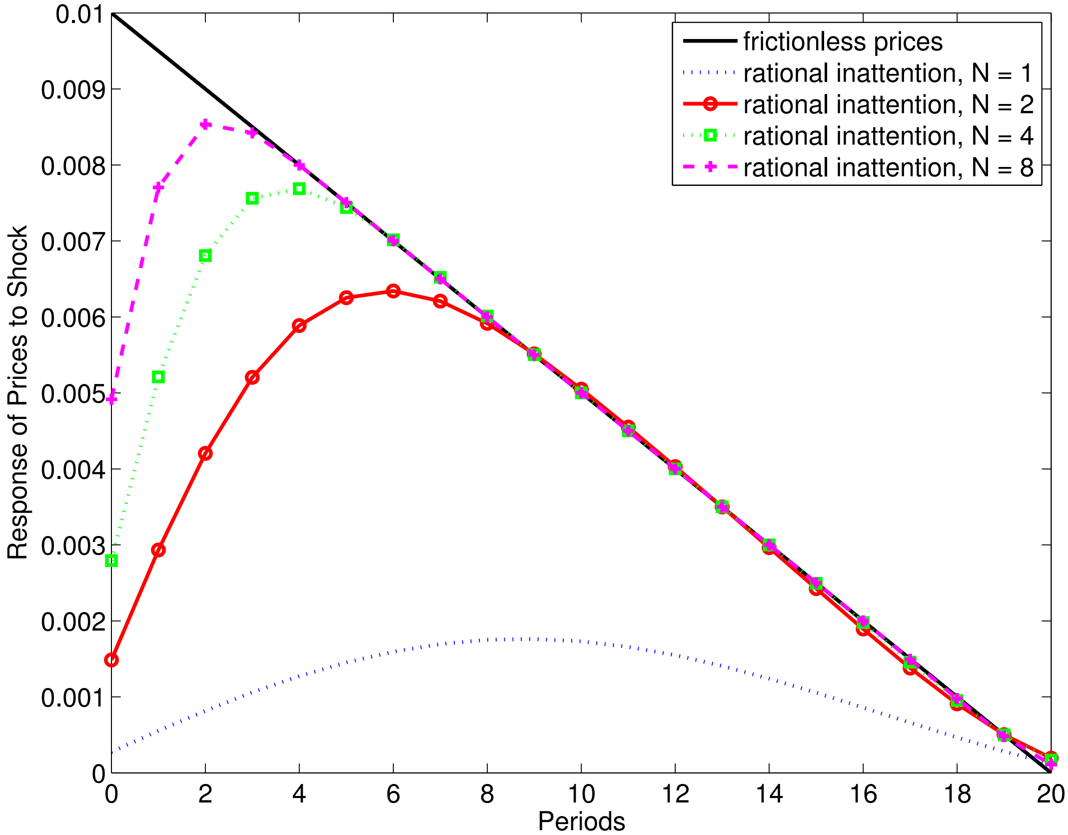
NOTE: We calibrate our model with homogeneous firms to moments for the whole sample of CPI data as we vary N. Firms' information processing capacity is calibrated such that its shadow price is invariant to N.

Figure 1: Equations (14) and (16) in the space (α, κ_a)



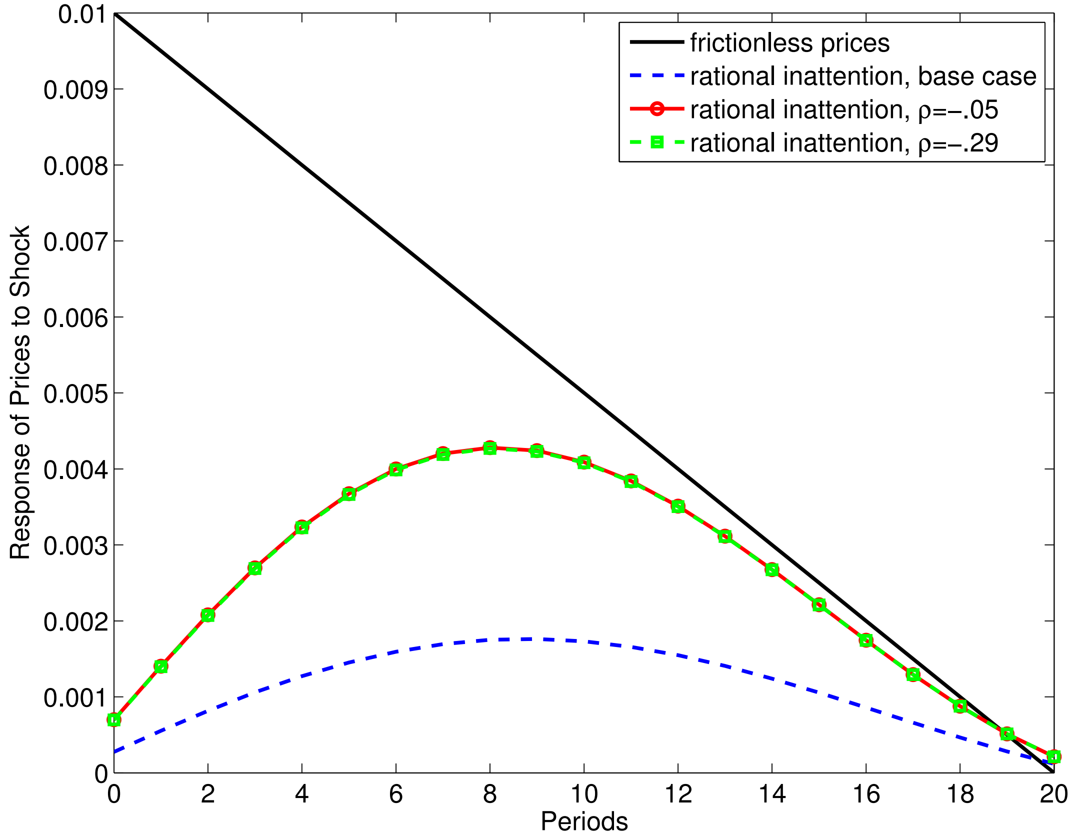
NOTE: The figure illustrates the fixed point problem of attention allocation given by equations (14) and (16). Equation (14) is drawn in red, while equation (16) is drawn in blue. Equation (16) is invariant to N, but N affects the drift and slope of equation (14). Under conditions described in Proposition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

Figure 2: Response of prices to a 1% impulse in q_t for sections 5.1 and 5.2



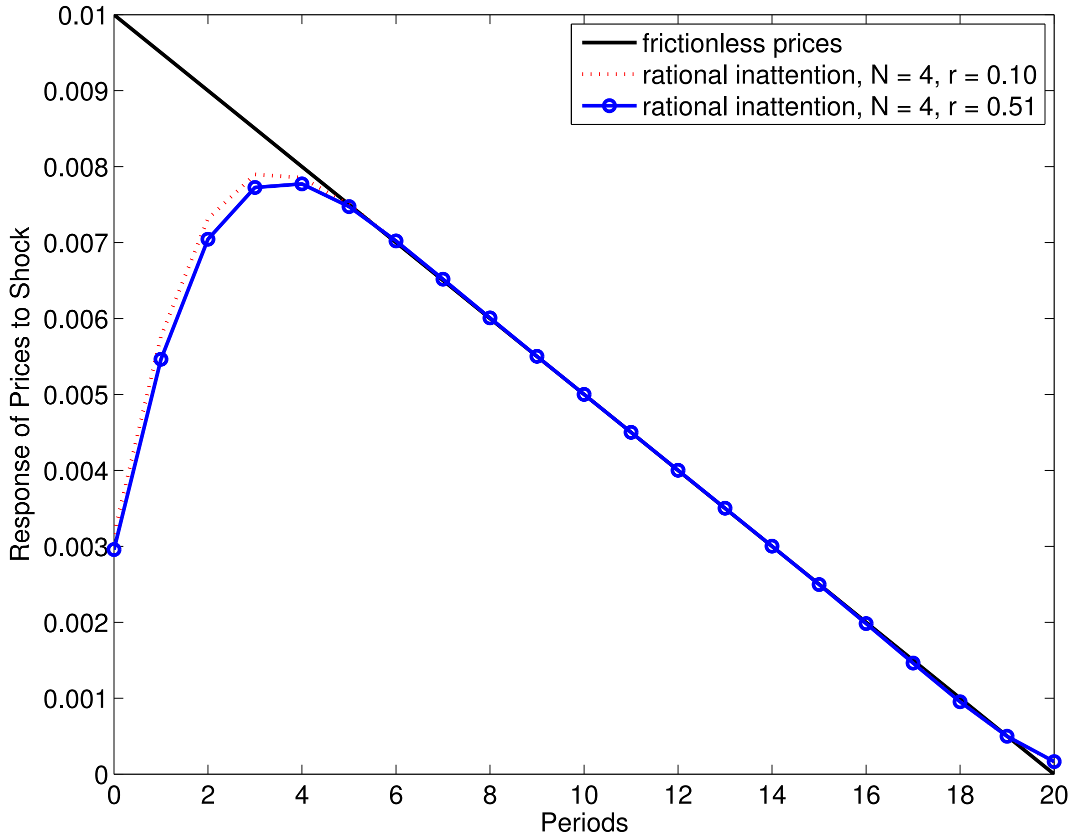
NOTE: We illustrate the response of prices to a 1% monetary shock as we vary N in our model calibrated to moments from the CPI data. The black line is for frictionless prices, the dashed blue line is for the benchmark of rationally inattentive prices with N=1, the red line with circles is for rationally 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 3: Response of prices to a 1% impulse in q_t for section 5.3



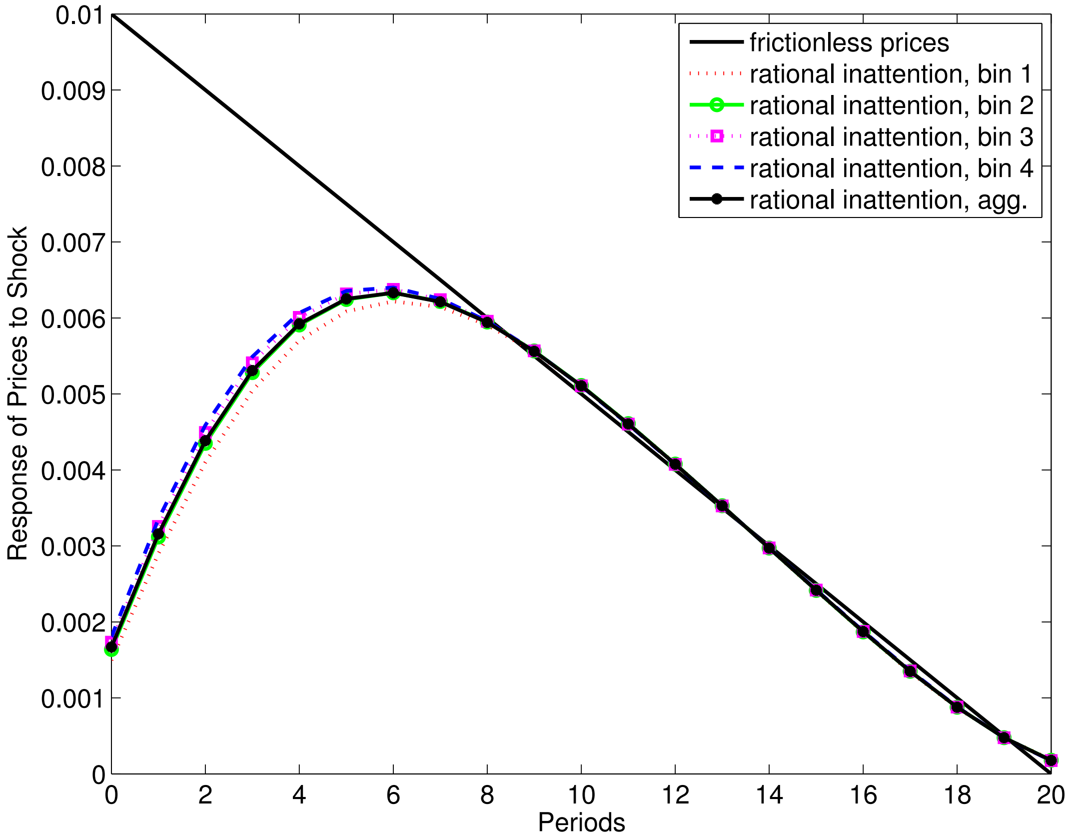
NOTE: We illustrate the response of prices to a 1% monetary shock as we vary the persistence of idiosyncratic shocks in our model calibrated to moments from the CPI data. The black line is for frictionless prices, the dashed blue line is for our benchmark with highly persistent idiosyncratic shocks, 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 5.3 contains further details.

Figure 4: Impulse response of prices under differing within-firm dispersion for section 5.4



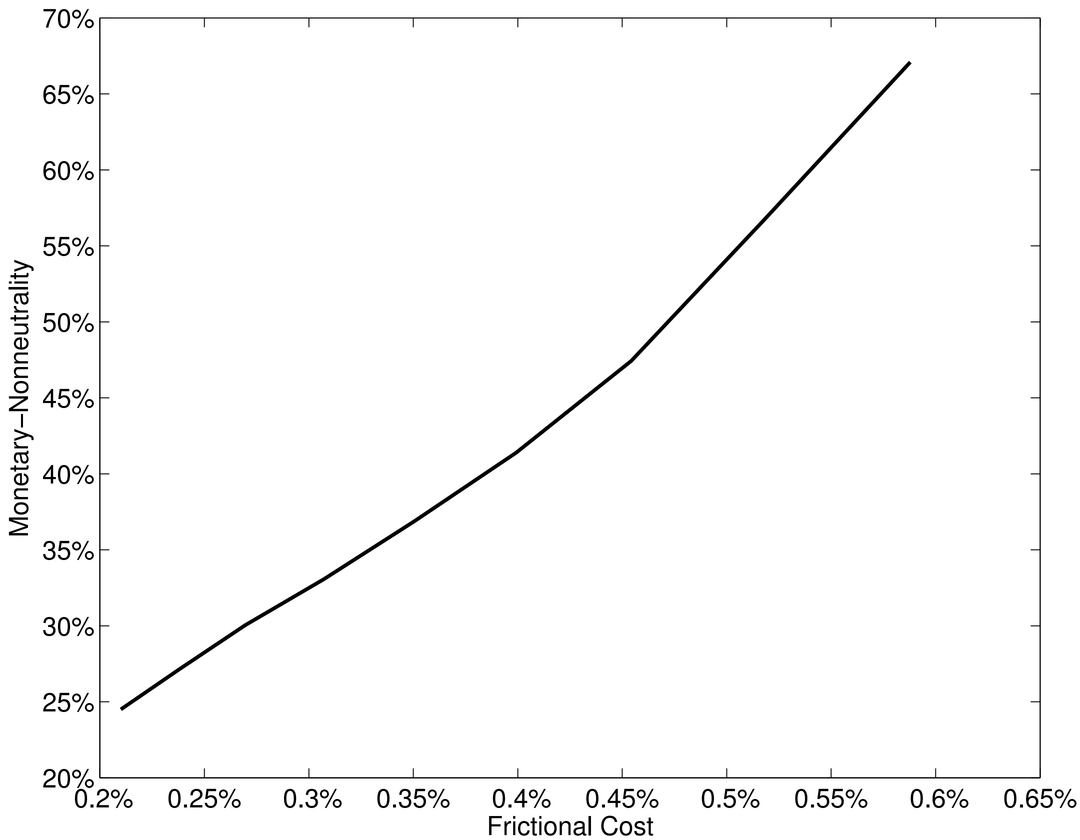
NOTE: We illustrate the response of prices to a 1% monetary shock for $N = 4$ as we vary the extent of within-firm log non-zero price dispersion. The blue line with circles denotes the impulse response for a 51% within-firm dispersion, the red doted line the response for a 10% within-firm dispersion. The black line is for frictionless prices. Section 5.4 contains further details.

Figure 5: Response of prices to a 1% impulse in q_t for section 5.5



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 5.5 contains further details.

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



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 5.7 contains further details.

A APPENDIX: The Problem of the Firm in our Analytical Model

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 (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

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

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

$$p_{nt}^{\diamondsuit} = p_t + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} y_t + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} f_{it} + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} z_{nt} \equiv \Delta_t + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} f_{it} + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} z_{nt} \tag{21}$$

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 (8) is

$$p_{nt}^{*} = \mathbb{E}\left[\Delta_t \mid s_{it}^{a}\right] + \frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|} \mathbb{E}\left[f_{it} \mid s_{nt}^{f}\right] + \frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|} \mathbb{E}\left[z_{jt} \mid s_{nt}^{z}\right]. \tag{22}$$

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:

$$\sum_{t=1}^{\infty} \beta^t \sum_{n\in\aleph_i} \left\{ \frac{|\widehat{\pi}_{11}|}{2} \mathbb{E}\left[\left(p_{nt}^{\diamondsuit} - p_{nt}^{*}\right)^2\right] \right\} \tag{23}$$

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.[17]

Given this assumption, we guess that the log-deviation of aggregate prices respond linearly to

[17] 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.

a monetary shock, $p_t = \alpha q_t$, so the compound aggregate variable Δ_t is given by

$$\Delta_t = \left[\alpha + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}(1-\alpha)\right] q_t. \tag{24}$$

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}.[18]

Therefore, given signals $\left\{s_{it}^a, s_{it}^f, s_{nt}^z\right\}$, the optimal pricing rule (22) 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 (23) and using the functional form of information flow in (6) because shocks are Gaussian white noise, the problem of the firm becomes

$$\min_{\kappa_a, \kappa_f, \{\kappa_n\}_{n \in \aleph_i}} \frac{\beta}{1-\beta} \frac{|\widehat{\pi}_{11}|}{2} \left[2^{-2\kappa_a} \sigma_\Delta^2 N + \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 2^{-2\kappa_f} \sigma_f^2 N + \left(\frac{\widehat{\pi}_{15}}{\widehat{\pi}_{11}}\right)^2 \sum_{n \in \aleph_i} 2^{-2\kappa_n} \sigma_z^2\right] \tag{25}$$

subject to

$$\kappa_a + \kappa_f + \sum_{n \in \aleph_i} \kappa_n \le \kappa(N). \tag{26}$$

where

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

[18] 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.

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 persistency ρ_Δ. Idiosyncratic shocks f_{it} and z_{jt} are also $AR\,(1)$ respectively with persistency ρ_f for all i and ρ_z for all j. The starting guess is now

$$p_t = \sum_{l=0}^{\infty} \alpha_l v_{t-l}, \tag{27}$$

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

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

$$\rightarrow \min_{\hat{\Delta}_{it},\{\hat{z}_{nt}\}_{n\in\aleph_i}} \frac{\beta}{1-\beta} \frac{|\widehat{\pi}_{11}|}{2} \left\{ \begin{array}{c} \mathbb{E}\left(\Delta_t - \hat{\Delta}_{it}\right)^2 N + \left(\frac{\widehat{\pi}_{14}}{|\widehat{\pi}_{11}|}\right)^2 \mathbb{E}\left(f_{it} - \hat{f}_{it}\right)^2 N \\ + \left(\frac{\widehat{\pi}_{15}}{|\widehat{\pi}_{11}|}\right)^2 \sum_{n\in\aleph_i} \mathbb{E}\left(z_{nt} - \hat{z}_{nt}\right)^2 \end{array} \right\}$$

subject to

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

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 Sims (2003): 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

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

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 $\tilde{x}_1\equiv\frac{|\widehat{\pi}_{11}|\sigma_\Delta\sqrt{1-\rho_\Delta^2}}{\widehat{\pi}_{14}\sigma_z\sqrt{1-\rho_f^2}}$, $\tilde{x}_2\equiv\frac{|\widehat{\pi}_{11}|\sigma_\Delta\sqrt{1-\rho_\Delta^2}}{\widehat{\pi}_{15}\sigma_z\sqrt{1-\rho_z^2}}$ 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 persistency of idiosyncratic shocks ρ_f and ρ_z implies an increase of κ_a^* relative to κ_f^* and κ_z^*.

B APPENDIX: 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 (2007). 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 (21) and $y_t = q_t - p_t$, implies

$$\Delta_t = \left(1 - \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|}\right) \sum_{l=0}^{\infty} \alpha_l v_{t-l} + \frac{\widehat{\pi}_{13}}{|\widehat{\pi}_{11}|} \sum_{l=0}^{\infty} a_l v_{t-l} \equiv \sum_{l=0}^{\infty} d_l v_{t-l} \tag{28}$$

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

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

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

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

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

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

where $G\left(e^{i\omega}\right) = g_0 + g_1 e^{i\omega} + g_2 e^{i2\omega} + ...$ and $H\left(e^{i\omega}\right) = h_0 + h_1 e^{i\omega} + h_2 e^{i2\omega} + ...$

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

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

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

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

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

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.

C APPENDIX: 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:

$$p_{nt}^{*} = \frac{\sigma_{\Delta}^2}{\sigma_{\Delta}^2 + \sigma_{\varepsilon i}^2} s_{it}^{a} + \frac{\sigma_f^2}{\sigma_f^2 + \sigma_{ei}^2} s_{it}^{f} + \frac{\sigma_z^2}{N\sigma_z^2 + \sigma_{\psi i}^2} s_{it}^{z}$$

which only responds to aggregate and firm-specific components.

Interdependent Profits

We now depart from our assumption in the main text that firms are decision units but not production units. We now assume that firms' production or commercialization processes are integrated such that the pricing decision of goods produced by a single firm are interdependent. We capture this 'interdependence' by assuming that the contribution to profits of a given good $n \in \aleph_i$ to its producing 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

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

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

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

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

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

which solves

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

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

$$p^{*}_{nt}=\frac{\sigma^2_\Delta}{\sigma^2_\Delta+\sigma^2_{\varepsilon i}}s^a_{it}+a_2\frac{\sigma^2_f}{\sigma^2_f+\sigma^2_{ei}}s^f_{it}+a_3\frac{\sigma^2_z}{\sigma^2_z+\sigma^2_{\psi n}}s^z_{nt}+a_4\sum_{-n\in\aleph_i}\frac{\sigma^2_z}{\sigma^2_z+\sigma^2_{\psi n}}s^z_{nt},$$

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 produces N goods is

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

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 are both production units and decision units. A firm that has an integrated productive process for its N goods may still decide to keep separated pricing processes such that a single decision unit decides $\widetilde{N}<N$ prices. A decision 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 $p^{\diamondsuit}_{nt}$ and p^{*}_{nt} obtained above.

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

$$\begin{aligned}\kappa^*_a &= \kappa^*_f+\log_2\left(\widetilde{x}_1\left(N\right)\right),\\ \kappa^*_a &= \kappa^*_n+\log_2\left(\widetilde{x}_2\left(N\right)\sqrt{\widetilde{N}}\right),\ \ \forall n\in\aleph_i.\end{aligned}$$

As in the main text, 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 $\widetilde{x}_1\left(N\right)$ and

$\widetilde{x}_2(N)$, which in the main text are parameters and here are functions of N:

$$\widetilde{x}_1 \equiv \frac{\sigma_\Delta}{a_2\sigma_f},$$

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

We then follow a similar logic than in Proposition 6. 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 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 4. 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 at the optimum is

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

and the shadow price of information-processing capacity is

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

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

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

is decreasing in N. However, this expression also governs the complementarity in pricing (a_1), so this complementarity would be increasing in N. As in the main text, the complementarity in pricing is deduced from aggregate data.

In addition, if this expression is decreasing in N, then the 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.