Optimal Sticky Prices under Rational Inattention*

Bartosz Maćkowiak
Humboldt University Berlin

Mirko Wiederholt
Humboldt University Berlin

First draft: June 2004. This draft: February 2005.

Abstract

In the data, individual prices change frequently and by large amounts. In standard sticky price models, frequent and large price changes imply a fast response of the aggregate price level to nominal shocks. This paper presents a model in which price setting firms optimally decide what to observe, subject to a constraint on information flow. When idiosyncratic conditions are more variable or more important than aggregate conditions, firms pay more attention to idiosyncratic conditions than to aggregate conditions. When we calibrate the model to match the large average absolute size of price changes observed in the data, prices react fast and by large amounts to idiosyncratic shocks, but prices react only slowly and by small amounts to nominal shocks. Nominal shocks have persistent real effects. We use the model to investigate how the optimal allocation of attention and the dynamics of prices depend on the firms' environment.

JEL: E3, E5, D8.

Key words and phrases: rational inattention, sticky prices, real effects of nominal shocks.

*We thank for helpful comments Klaus Adam, Michael Burda, Giancarlo Corsetti, Wouter Den Haan, Andrea Gerali, Frank Smets, Harald Uhlig and seminar participants at the Bank of Italy, Central European University, European Central Bank, European University Institute, Goethe University Frankfurt and Humboldt University Berlin. Bartosz Maćkowiak thanks Chris Sims for conversations about rational inattention. Mirko Wiederholt thanks Ente Einaudi for hospitality. Email: bartosz@wiwi.hu-berlin.de and wiederholt@wiwi.hu-berlin.de.

"An optimizing trader will process those prices of most importance to his decision problem most frequently and carefully, those of less importance less so, and most prices not at all. Of the many sources of risk of importance to him, the business cycle and aggregate behavior generally is, for most agents, of no special importance, and there is no reason for traders to specialize their own information systems for diagnosing general movements correctly." (Lucas, 1977, p. 21)

1 Introduction

In the data, individual prices change frequently and by large amounts. Bils and Klenow (2004) and Klenow and Kryvtsov (2004) study micro data on consumer prices that the U.S. Bureau of Labor Statistics collects to compute the consumer price index. Bils and Klenow find that half of all non-housing consumer prices last less than 4.3 months. Klenow and Kryvtsov find that, conditional on the occurrence of a price change, the average absolute size of the price change is over 13 percent.[1]

At the same time, the aggregate price level responds slowly to monetary policy shocks. A variety of different schemes for identifying monetary policy shocks yield this result (e.g. Christiano, Eichenbaum and Evans (1999), Leeper, Sims and Zha (1996) and Uhlig (2004)). Uhlig (2004) finds that only about 25 percent of the long-run response of the U.S. GDP price deflator to a monetary policy shock occurs within the first year after the shock.

This combination of empirical observations is difficult to explain with standard models of sticky prices. Consider the popular time-dependent model of price setting due to Calvo (1983). The Calvo model can explain why the aggregate price level responds slowly to monetary shocks if: (a) individual firms in the model adjust prices infrequently;[2] or (b) individual firms in the model adjust prices by small amounts. See Woodford (2003) for

[1] The finding that individual prices change frequently and by large amounts is robust to whether temporary price changes reflecting sales are included or not. When Bils and Klenow net out the impact of sales, the median price duration rises from 4.3 to 5.5 months. When Klenow and Kryvtsov net out the impact of sales, the average absolute size of price changes falls from 13.3 to 8.5 percent.

[2] Galí and Gertler (1999) estimate the Calvo model using quarterly aggregate U.S. data. The estimated model implies that a typical firm waits about 5-6 quarters before changing its price.

reasons why firms can find it optimal to adjust prices by small amounts in response to shocks. However, neither (a) nor (b) seems to be true in the data.

Golosov and Lucas (2003) conduct a quantitative experiment with a state-dependent model of price setting. They calibrate a menu cost model with idiosyncratic productivity shocks and monetary shocks to match the frequency and size of price changes reported in Klenow and Kryvtsov (2004). In the calibrated model, the aggregate price level responds quickly to a monetary shock. The reason is that a firm setting a new price in a menu cost model takes into account current values of all shocks. Hence, frequent price adjustment implies a fast response of prices to all shocks, including monetary shocks.

This paper presents a model that can explain why individual prices change frequently and by large amounts and, at the same time, the aggregate price level responds slowly to monetary shocks. We study price setting by firms under "rational inattention" in the sense of Sims (2003). Firms can change prices every period at no cost. The profit-maximizing price depends on the aggregate price level, real aggregate demand and an idiosyncratic state variable (reflecting consumers' tastes or the firm's technology). We let firms decide what to observe. Firms choose the number of signals that they receive every period as well as the stochastic properties of these signals. Firms face the constraint that the information flow between the sequence of signals and the sequence of states of the economy is bounded. Other properties of the signals are up to the firms. In particular, since the state of the economy is multidimensional, firms decide which variables to observe with higher precision. We close the model by specifying exogenous stochastic processes for nominal aggregate demand and the idiosyncratic state variables.

The model makes the following predictions. Firms adjust prices every period and yet impulse responses of prices to shocks are sticky – dampened and delayed relative to the impulse responses under perfect information. The extent of dampening and delay in a particular impulse response depends on the amount of attention allocated to the type of shock. When idiosyncratic conditions are more variable or more important than aggregate conditions, firms pay more attention to idiosyncratic conditions than to aggregate conditions. In this case, price reactions to idiosyncratic shocks are strong and quick, but price reactions to aggregate shocks are dampened and delayed. This can explain why individual prices change frequently and by large amounts while the aggregate price level responds slowly to monetary shocks. In addition, there is a feedback effect. When firms pay little attention to aggregate conditions, the aggregate price level moves little and therefore firms find it optimal to pay even less attention to aggregate conditions. The feedback effect makes the aggregate price level even more sticky.

We calibrate the model to match the average absolute size of price changes reported in Klenow and Kryvtsov (2004). We find that prices react fast and by large amounts to idiosyncratic shocks, but prices react only slowly and by small amounts to nominal shocks. Nominal shocks have persistent real effects. The reason is the following. To match the large average absolute size of price changes observed in the data, idiosyncratic shocks in the model must have a large variance or must be very important for pricing decisions. This implies that firms allocate most of their attention to idiosyncratic conditions.

We use the model to investigate how the optimal allocation of attention and the dynamics of prices depend on the firms' environment. As the variance of nominal aggregate demand increases, the firms' tracking problem becomes more difficult and their profits decrease. Firms react by reallocating attention to aggregate conditions away from idiosyncratic conditions. Firms track both aggregate and idiosyncratic conditions less well. The fall in profits suggests that costs of aggregate instability in the real world may be due to the fact that aggregate instability makes the firms' tracking problem more difficult. As the variance of the idiosyncratic state variables increases, firms reallocate attention to idiosyncratic conditions away from aggregate conditions. Therefore the model predicts that firms operating in more unstable idiosyncratic environments track aggregate conditions less well.

Sims (1998) was the first to propose information flow constraints as a source of inertial behavior. Sims (2003) works out some implications of adding information flow constraints to economic models. The firms' problem of deciding what to observe in our model is similar to the quadratic control problem with an information flow constraint studied in Sims (2003, Section 4). However, there are important differences. One difference is that firms in our model track a multidimensional state of the economy. Thus firms have to decide how to allocate attention across different variables. Another difference is that firms in our model track an endogenous variable – the aggregate price level. This introduces the feedback effect described above.[3]

Our work is also related to the literature on information imperfections and the real effects of monetary shocks. In Lucas (1973), firms observe prices in their markets but not the aggregate price level. Firms rationally misinterpret unexpected inflation for a relative price increase and react by raising output. Monetary shocks have real effects until they become public information. Since changes in monetary policy are published with little delay, it has been argued that the Lucas model cannot explain persistent real effects of monetary policy shocks.

Woodford (2002) points out that there is a difference between public information – information that is available in principle to anyone who chooses to look it up – and the information of which decisionmakers are actually aware. He follows Sims (2003) in arguing that agents have limited capacity to acquire and process information. Woodford uses this idea to motivate a model in which firms observe nominal aggregate demand with idiosyncratic noise. If strategic complementarity in price setting is strong, the real effects of a nominal shock can be large and persistent. While Woodford assumes that firms pay little attention to aggregate conditions, we derive the optimal allocation of attention. This allows us to identify the circumstances under which firms find it optimal to pay little attention to aggregate conditions. Furthermore, this allows us to study how the optimal allocation of attention and the dynamics of prices vary with changes in the firms' environment.[4]

Mankiw and Reis (2002) develop a different model in which firms are imperfectly informed about the state of the economy. Mankiw and Reis assume that every period an exogenous fraction of firms obtains perfect information about all current and past disturbances, while all other firms continue to set prices based on old information. Reis (2004) provides microfoundations for this kind of slow diffusion of information. He assumes that

[3] Moscarini (2003) studies a univariate quadratic control problem with an information flow constraint. In contrast to Sims (2003), Moscarini assumes that the decisionmaker can only meet the information flow constraint by infrequent sampling. Moscarini analyzes the optimal sampling frequency. The information that the decisionmaker receives once he or she samples is given exogenously.

[4] Woodford's (2002) model has been extended in a number of directions. Hellwig (2002) studies the role of public information. Gumbau-Brisa (2003) studies the effects of a Taylor rule. Adam (2004) studies optimal monetary policy.

firms face a fixed cost of obtaining perfect information, implying that firms decide to obtain information infrequently. In Mankiw and Reis (2002) and Reis (2004), prices react with equal speed to all disturbances. In contrast, in our model firms optimally decide to receive more precise information concerning some shocks and less precise information concerning other shocks. Therefore in our model prices react quickly to some shocks but only slowly to other shocks. This can explain the combination of empirical observations that motivate our paper. Note that in a model with a fixed cost of obtaining information, the cost of obtaining information is independent of the stochastic properties of the variables to be tracked. In a model with an information flow constraint, tracking a variable with a higher variance well uses up a larger fraction of the available information flow.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 derives the firms' price setting behavior. In Section 4 we solve a special case of the model analytically. In Sections 5 and 6 we return to the model in its general form. In Section 5 we study the firms' decision problem of what to observe given aggregate variables. In Section 6 we compute the rational expectations equilibrium for a variety of different economies. Section 7 concludes. Appendix A introduces the tools that we use to state the firms' information flow constraint. The remaining appendices contain the proofs of the results used in the main text and details of how to solve the model numerically.

2 The model

2.1 Description of the economy

Consider an economy with a continuum of firms indexed by $i \in [0,1]$. Time is discrete and indexed by t.

Firm i sells a good also indexed by i. Every period $t = 1, 2, \ldots$, the firm sets the price of the good, P_{it}, so as to maximize

$$E_{it}\left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \pi\left(P_{i\tau}, P_{\tau}, Y_{\tau}, Z_{i\tau}\right)\right], \tag{1}$$

where E_{it} is the expectation operator conditioned on the information of firm i in period t, β is a scalar between zero and unity and $\pi\left(P_{it}, P_t, Y_t, Z_{it}\right)$ are real profits of firm i in period

t. The real profits depend on the price set by the firm, P_{it}, the aggregate price level, P_t, real aggregate demand, Y_t, and an idiosyncratic state variable, Z_{it}. The variable Z_{it} reflects consumers' valuation of good i or the firm-specific state of technology. We assume that the function π is twice continuously differentiable and homogenous of degree zero in its first two arguments. Thus real profits only depend on the relative price P_{it}/P_t.[5]

The information of firm i in period t is given by the sequence of all signals that the firm has received up to that point in time

$$s_i^t = \{s_i^1, s_{i2}, ..., s_{it}\}. \tag{2}$$

Here s_{it} denotes the signal that firm i receives in period t. The signal can be vector valued. Furthermore, s_i^1 denotes the sequence of signals that firm i receives in period one. We allow for the possibility that the firm receives a sequence of signals in period one.

The firm can change the price of the good every period at no cost and takes as given the stochastic process for the aggregate price level, $\{P_t\}$, the stochastic process for real aggregate demand, $\{Y_t\}$, and the stochastic process for the idiosyncratic state variable, $\{Z_{it}\}$. Therefore the price setting problem of firm i in period t is a purely static problem

$$\max_{P_{it}} E_{it}[\pi\left(P_{it}, P_t, Y_t, Z_{it}\right)]. \tag{3}$$

The aggregate environment of firms is specified by postulating an exogenous stochastic process for nominal aggregate demand.[6] Let

$$Q_t \equiv P_t Y_t \tag{4}$$

denote nominal aggregate demand and let $q_t \equiv \ln Q_t - \ln \bar{Q}$ denote the log-deviation of nominal aggregate demand from its deterministic trend. We assume that q_t follows a stationary Gaussian process with mean zero and absolutely summable autocovariances.

[5] To give an example, in a standard model of monopolistic competition

$$\pi\left(P_{it}, P_t, Y_t, Z_{it}\right) = Y_t \left(\frac{P_{it}}{P_t}\right)^{1-\theta} - C\left(Y_t \left(\frac{P_{it}}{P_t}\right)^{-\theta}, Y_t, Z_{it}\right),$$

where Y_t is the Dixit-Stiglitz index of real aggregate demand, P_t is the corresponding price index and $Y_t \left(\frac{P_{it}}{P_t}\right)^{-\theta}$ is the demand for good i with $\theta > 1$. Real production costs C depend on the firm's output and may depend on real aggregate demand through factor prices. Here Z_{it} is a productivity index.

[6] This approach is common in the literature. For example, Lucas (1973), Woodford (2002), Mankiw and Reis (2002) and Reis (2004) also postulate an exogenous stochastic process for nominal aggregate demand.

The log of the aggregate price level is defined as

$$\ln P_t \equiv \int_0^1 \ln P_{it} di. \tag{5}$$

One obtains the same equation in a standard model of monopolistic competition after a log-linearization.[7]

The idiosyncratic environment of firms is specified by postulating an exogenous stochastic process for the idiosyncratic state variables. Let $z_{it} \equiv \ln Z_{it} - \ln \bar{Z}$ denote the log-deviation of the idiosyncratic state variable of firm i from its deterministic trend. We assume that the processes $\{z_{it}\}$, $i \in [0,1]$, are mutually independent and independent of $\{q_t\}$. Furthermore, we assume that the z_{it}, $i \in [0,1]$, follow a common stationary Gaussian process with mean zero and absolutely summable autocovariances. Since the z_{it}, $i \in [0,1]$, for given t are mutually independent and identically distributed random variables with mean zero and finite variance, we have[8]

$$\int_0^1 z_{it} di = 0. \tag{6}$$

One could close the model by making an assumption about the information that firm i obtains in period t. That is what is typically done in the literature.[9] In contrast, we would like to capture the fact that firms can decide what to observe. Therefore we let each firm choose the stochastic process for the signal s_{it}. We refer to this choice as the choice of the information system. Formally, we assume that firm i solves in period zero

$$\max_{\{s_{it}\} \in \Gamma} E\left[\sum_{t=1}^{\infty} \beta^t \pi\left(P_{it}^*, P_t, Y_t, Z_{it}\right)\right], \tag{7}$$

subject to

$$P_{it}^* = \arg\max_{P_{it}} E[\pi\left(P_{it}, P_t, Y_t, Z_{it}\right) | s_i^t], \tag{8}$$

[7] In a standard model of monopolistic competition, the aggregate price level is defined as $P_t \equiv \left(\int_0^1 P_{it}^{1-\theta} di\right)^{\frac{1}{1-\theta}}$. Log-linearizing this equation around any point with the property that all the P_{it} are equal yields equation (5).

[8] See Uhlig (1996), Theorem 2.

[9] For example, the perfect information case obtains when $s_{it} = (P_t, Y_t, Z_{it})'$ for all i, t. In a signal-extraction model, s_{it} would equal the variables of interest plus exogenous noise. In an information-delay model, $s_{it} = (P_{t-n}, Y_{t-n}, Z_{it-n})'$ for some $n > 0$. In a sticky information model, $s_{it} = (P_1, ..., P_t, Y_1, ..., Y_t, Z_{i1}, ..., Z_{it})'$ with some probability ρ and $s_{it} = s_{it-1}$ with probability $1 - \rho$.

and the information flow constraint

$$\mathcal{I}\left(\left\{P_t\right\},\left\{Y_t\right\},\left\{Z_{it}\right\};\left\{s_{it}\right\}\right)\leq\kappa. \tag{9}$$

The firm can choose the stochastic process for the signal from the set Γ defined below. The firm chooses the stochastic process for the signal so as to maximize the expected discounted sum of profits, taking into account how the choice of the information system affects the price setting behavior in future periods. We follow Sims (2003) in assuming that agents have limited ability to acquire and process information. The information flow constraint (9) imposes an upper bound on the information flow between the sequence of signals and the sequence of states of the economy. The information flow between stochastic processes is defined in Appendix A.

The information flow constraint implies that the firm cannot decide to observe the multidimensional state of the economy perfectly in every period. The firm can decide to observe some variables with a higher precision than other variables, as long as the total information flow does not exceed the parameter κ.[10] This formalizes the idea that agents can focus on some variables at the cost of not focusing on other variables. See the remark by Lucas (1977) quoted at the beginning of this paper.

The definition of the set Γ captures additional assumptions about how firms acquire and process information. The set Γ is defined as the set of all stochastic processes that have the following four properties. First, the signals contain no information about future innovations to nominal aggregate demand and future innovations to the idiosyncratic state variables. Second,

$$s_{it}=\left(s_{1it},s_{2it}\right)', \tag{10}$$

where

$$\left\{s_{1it}\right\},\left\{P_t\right\},\left\{Y_t\right\}\text{ are independent of }\left\{s_{2it}\right\},\left\{Z_{it}\right\}. \tag{11}$$

The vector of signals that firm i receives in period t consists of a first set of signals concerning aggregate conditions, s_{1it}, and a second set of signals concerning idiosyncratic conditions, s_{2it}. A signal concerning aggregate conditions contains no information about idiosyncratic

[10] In the model, the ability of a firm to acquire and process information is exogenous. It is straightforward to extend the model by specifying a cost function for κ and letting each firm choose the optimal κ.

conditions and vice versa. This assumption captures the idea that acquiring information about aggregate conditions and acquiring information about idiosyncratic conditions are two separate activities.[11] Third,

$$\{s_{1it}, s_{2it}, p_t, y_t, z_{it}\} \text{ is a stationary Gaussian vector process,} \tag{12}$$

where p_t denotes the log-deviation of the aggregate price level from its deterministic trend and y_t denotes the log-deviation of real aggregate demand from its deterministic trend. It seems reasonable that in a Gaussian economy Gaussian signals are optimal. In special cases, this is easy to prove. In the general case, we have not proved this yet. Therefore we state (12) as an assumption.[12] Fourth, all noise in signals is idiosyncratic. Thus we follow Sims (2003) and Woodford (2002) in supposing that the critical bottleneck is not the public availability of information but instead the limited ability of private decisionmakers to acquire and process correctly all relevant information.

Finally, we make a simplifying assumption. We assume that firms receive a long sequence of signals in period one after having chosen the information system in period zero

$$s_i^1 = \{s_{i-\infty}, \ldots, s_{i1}\}. \tag{13}$$

This assumption implies that the price set by a firm follows a stationary process. This simplifies the analysis.[13]

[11] We think that most signals in the real world approximately satisfy the independence assumption (11). Consider a manager who has to set a price. There is no signal in the real world that would simply tell the manager the optimal price. The manager has to collect different pieces of information (he may delegate this task). For example, the manager may read about the aggregate state of the economy in a financial newspaper. To give another example, the manager may commission a marketing report about tastes of customers. Reading a financial newspaper gives the manager very little information about whether customers like a particular good, what production of the good would cost and whether competitors might produce the good more cheaply. A marketing report gives the manager very little information about the aggregate state of the economy.

[12] Note that assumption (12) can only be satisfied when $\{p_t, y_t, z_{it}\}$ is a stationary Gaussian vector process. We will verify that this is true in equilibrium. Similarly, condition (11) can only be satisfied when $\{P_t\}$ and $\{Y_t\}$ are independent of $\{Z_{it}\}$. Again we will verify that this is true in equilibrium.

[13] One can show that observing a long sequence of signals in period one does not change the information flow.

2.2 Equilibrium

An equilibrium of the model are stochastic processes for the signals, $\{s_{it}\}$, for the prices, $\{P_{it}\}$, for the aggregate price level, $\{P_t\}$, and for real aggregate demand, $\{Y_t\}$, such that:

1. Each firm $i \in [0,1]$ chooses the stochastic process for the signal optimally in period $t = 0$ and sets the price for the good that it sells optimally in periods $t = 1, 2, \ldots$, taking as given $\{P_t\}$, $\{Y_t\}$ and $\{Z_{it}\}$.

2. In every period $t = 1, 2, \ldots$ and in each state of nature, the aggregate price level is given by (5) and real aggregate demand satisfies (4).

3 Price setting behavior

In this section, we look at the firms' price setting behavior for a given choice of the information system.

The first-order condition for optimal price setting by firm i in period t is

$$E[\pi_1 \left(P_{it}^*, P_t, Y_t, Z_{it}\right) | s_i^t] = 0, \tag{14}$$

where π_1 denotes the derivative of the function π with respect to its first argument. In order to obtain a closed-form solution for the price set by the firm, we work with a log-quadratic approximation to the profit function around the solution of the non-stochastic version of the model.

The solution of the non-stochastic version of the model is as follows. Suppose that $Q_t = \bar{Q}$ for all t and $Z_{it} = \bar{Z}$ for all i, t. In this case, there is no uncertainty and all firms solve the same price setting problem. Therefore, in equilibrium, it has to be true that

$$\pi_1 \left(P_t, P_t, Y_t, \bar{Z}\right) = 0. \tag{15}$$

Multiplying by $P_t > 0$ yields[14]

$$\pi_1 \left(1, 1, Y_t, \bar{Z}\right) = 0. \tag{16}$$

[14] Recall that the profit function π is homogeneous of degree zero in its first two arguments. Therefore the function π_1 is homogeneous of degree minus one in its first two arguments.

The last equation characterizes equilibrium real aggregate demand, denoted $\bar{Y}$.[15] The equilibrium aggregate price level, denoted $\bar{P}$, is given by

$$\bar{P} = \frac{\bar{Q}}{\bar{Y}}. \tag{17}$$

Next we compute the log-quadratic approximation to the profit function around the non-stochastic solution of the model. Let $x_t \equiv \ln X_t - \ln \bar{X}$ denote the log-deviation of a variable from its value at the non-stochastic solution. Note that $X_t = \bar{X}e^{x_t}$ and define the function $\hat{\pi}$ via $\hat{\pi}\left(p_{it}, p_t, y_t, z_{it}\right) = \pi\left(\bar{P}e^{p_{it}}, \bar{P}e^{p_t}, \bar{Y}e^{y_t}, \bar{Z}e^{z_{it}}\right)$. Compute a second-order Taylor approximation to the function $\hat{\pi}$ around the point $(p_{it}, p_t, y_t, z_{it}) = (0,0,0,0)$. This yields the log-quadratic approximation to the profit function

$$\begin{aligned}\tilde{\pi}\left(p_{it}, p_t, y_t, z_{it}\right) = {} & \hat{\pi}\left(0,0,0,0\right) + \hat{\pi}_1 p_{it} + \hat{\pi}_2 p_t + \hat{\pi}_3 y_t + \hat{\pi}_4 z_{it} \\ & + \frac{\hat{\pi}_{11}}{2} p_{it}^2 + \frac{\hat{\pi}_{22}}{2} p_t^2 + \frac{\hat{\pi}_{33}}{2} y_t^2 + \frac{\hat{\pi}_{44}}{2} z_{it}^2 \\ & + \hat{\pi}_{12} p_{it} p_t + \hat{\pi}_{13} p_{it} y_t + \hat{\pi}_{14} p_{it} z_{it} \\ & + \hat{\pi}_{23} p_t y_t + \hat{\pi}_{24} p_t z_{it} + \hat{\pi}_{34} y_t z_{it},\end{aligned} \tag{18}$$

where $\hat{\pi}_1$, for example, denotes the derivative of the function $\hat{\pi}$ with respect to its first argument evaluated at the point $(p_{it}, p_t, y_t, z_{it}) = (0,0,0,0)$. It is straightforward to show that $\hat{\pi}_1 = 0$, $\hat{\pi}_{11} < 0$ and $\hat{\pi}_{12} = -\hat{\pi}_{11}$.

After the log-quadratic approximation to the profit function, the solution to the price setting problem of firm i in period t is[16]

$$p_{it}^* = E[p_t|s_i^t] + \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} E[y_t|s_i^t] + \frac{\hat{\pi}_{14}}{|\hat{\pi}_{11}|} E[z_{it}|s_i^t]. \tag{19}$$

The log of the price set by firm i in period t is a linear function of the conditional expectation of the log of the aggregate price level, the conditional expectation of the log of real aggregate demand and the conditional expectation of the log of the idiosyncratic state variable.

For comparison, the solution to the price setting problem of firm i in period t under perfect information is

$$p_{it}^f = p_t + \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} y_t + \frac{\hat{\pi}_{14}}{|\hat{\pi}_{11}|} z_{it}. \tag{20}$$

[15] Here we assume that equation (16) has a unique solution.

[16] Take the derivative of $E\left[\tilde{\pi}\left(p_{it}, p_t, y_t, z_{it}\right)|s_i^t\right]$ with respect to p_{it}, set the derivative equal to zero and solve for p_{it}. Recall that $\hat{\pi}_1 = 0$, $\hat{\pi}_{11} < 0$ and $\hat{\pi}_{12} = -\hat{\pi}_{11}$. This yields equation (19).

Whenever the price (19) differs from the price (20) there is a loss in profits due to imperfect information. More precisely, the period t loss in profits due to imperfect information is

$$\tilde{\pi}\left(p_{it}^{f}, p_t, y_t, z_{it}\right) - \tilde{\pi}\left(p_{it}^{*}, p_t, y_t, z_{it}\right) = \frac{|\hat{\pi}_{11}|}{2}\left(p_{it}^{f} - p_{it}^{*}\right)^2. \tag{21}$$

The firm can affect this loss by choosing the information system.

Before we turn to the choice of the information system, two additional observations will be helpful. First, let us define $\Delta_t \equiv p_t + \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} y_t$. The following equations show that the variable Δ_t summarizes all the firm would like to know about aggregate conditions. The prices (19) and (20) can be expressed as

$$p_{it}^{*} = E[\Delta_t | s_i^t] + \frac{\hat{\pi}_{14}}{|\hat{\pi}_{11}|} E[z_{it} | s_i^t], \tag{22}$$

and

$$p_{it}^{f} = \Delta_t + \frac{\hat{\pi}_{14}}{|\hat{\pi}_{11}|} z_{it}. \tag{23}$$

Second, computing the integral over all i of the price (20) and using equation (6) as well as $y_t = q_t - p_t$ yields the following expression for the aggregate price level under perfect information

$$p_t^{f} = \left(1 - \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}\right) p_t + \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} q_t. \tag{24}$$

The fixed point of this mapping is the equilibrium aggregate price level under perfect information. Assuming $\hat{\pi}_{13} \neq 0$, the unique fixed point is

$$p_t^{f} = q_t. \tag{25}$$

Hence, the equilibrium aggregate price level under perfect information moves one for one with nominal aggregate demand.

4 Analytical solution when exogenous processes are white noise

In this section, we solve the model under the assumption that log-deviations of nominal aggregate demand and log-deviations of the idiosyncratic state variables follow white noise processes. In this special case, the model can be solved analytically. We illustrate the main mechanisms of the model with the help of this simple example. Afterwards, we solve the model under more realistic assumptions concerning the exogenous processes.

In this section, we assume that q_t follows a white noise process with variance $\sigma_q^2 > 0$ and the z_{it}, $i \in [0,1]$, follow a common white noise process with variance $\sigma_z^2 > 0$. We guess that in equilibrium

$$p_t = \alpha q_t, \tag{26}$$

and

$$y_t = (1-\alpha)\, q_t, \tag{27}$$

where $\alpha \in [0,1]$. The guess will be verified.

Suppose that firm i can choose among signals of the form

$$s_{1it} = \Delta_t + \varepsilon_{it}, \tag{28}$$

$$s_{2it} = z_{it} + \psi_{it}, \tag{29}$$

where $\{\varepsilon_{it}\}$ and $\{\psi_{it}\}$ are Gaussian white noise processes with variances σ_ε^2 and σ_ψ^2, respectively. The processes $\{\varepsilon_{it}\}$ and $\{\psi_{it}\}$ are mutually independent and independent of $\{q_t\}$ and $\{z_{it}\}$. We think of the noise in the signals as reflecting observational errors and processing errors. The firm can reduce the amount of noise in a particular signal by devoting more attention to that variable. Choosing among signals of the form (28) – (29) is more restrictive than choosing a stochastic process for the signal from the set Γ. Later we will prove the following result (see Proposition 3). When Δ_t and z_{it} follow white noise processes, there exist optimal signals of the form (28) – (29) and all optimal signals imply the same price setting behavior. Therefore, in the special case analyzed in this section, restricting the firms' choice to signals of the form (28) – (29) does not change the equilibrium of the model.

The variables p_t, y_t and Δ_t are perfectly correlated and the variables Δ_t, z_{it}, s_{1it} and s_{2it} follow white noise processes. In this case, the information flow constraint (9) becomes

$$\frac{1}{2}\log_2\left(\frac{\sigma_\Delta^2}{\sigma_\varepsilon^2}+1\right)+\frac{1}{2}\log_2\left(\frac{\sigma_z^2}{\sigma_\psi^2}+1\right) \leq \kappa. \tag{30}$$

See Appendix B. The information flow constraint places a restriction on the signal-to-noise ratios, $\sigma_\Delta^2/\sigma_\varepsilon^2$ and σ_z^2/σ_ψ^2. When the information flow constraint is binding, the firm faces a trade-off: Increasing one signal-to-noise ratio requires reducing the other signal-to-noise ratio.

Let $\kappa_1 = \frac{1}{2}\log_2\left(\frac{\sigma_\Delta^2}{\sigma_\varepsilon^2}+1\right)$ denote the information flow allocated to aggregate conditions and let $\kappa_2 = \frac{1}{2}\log_2\left(\frac{\sigma_z^2}{\sigma_\psi^2}+1\right)$ denote the information flow allocated to idiosyncratic conditions. A given allocation of the information flow (κ_1 and κ_2) is associated with the following variances of noise

$$\sigma_\varepsilon^2 = \frac{\sigma_\Delta^2}{2^{2\kappa_1}-1}, \tag{31}$$

$$\sigma_\psi^2 = \frac{\sigma_z^2}{2^{2\kappa_2}-1}. \tag{32}$$

These variances of noise imply the following price setting behavior

$$\begin{aligned} p_{it}^* &= \frac{\sigma_\Delta^2}{\sigma_\Delta^2+\sigma_\varepsilon^2}s_{1it} + \frac{\hat{\pi}_{14}}{|\hat{\pi}_{11}|}\frac{\sigma_z^2}{\sigma_z^2+\sigma_\psi^2}s_{2it} \\ &= \left(1-2^{-2\kappa_1}\right)\left(\Delta_t+\varepsilon_{it}\right) + \frac{\hat{\pi}_{14}}{|\hat{\pi}_{11}|}\left(1-2^{-2\kappa_2}\right)\left(z_{it}+\psi_{it}\right), \end{aligned} \tag{33}$$

where the first equality follows from a standard linear projection argument and the second equality follows from equations (31) – (32). This price setting behavior is associated with the following expected losses in profits due to imperfect information

$$\begin{aligned} & E\left[\sum_{t=1}^{\infty}\beta^t\left\{\tilde{\pi}\left(p_{it}^f,p_t,y_t,z_{it}\right)-\tilde{\pi}\left(p_{it}^*,p_t,y_t,z_{it}\right)\right\}\right] \\ =& \sum_{t=1}^{\infty}\beta^t\frac{|\hat{\pi}_{11}|}{2}E\left[\left(p_{it}^f-p_{it}^*\right)^2\right] \\ =& \frac{\beta}{1-\beta}\frac{|\hat{\pi}_{11}|}{2}\left\{2^{-2\kappa_1}\sigma_\Delta^2+\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 2^{-2\kappa_2}\sigma_z^2\right\}, \end{aligned} \tag{34}$$

where the first equality follows from equation (21) and the second equality follows from equation (23) and equations (31) – (33).

Therefore the optimal allocation of the information flow (the optimal allocation of attention) is the solution to the strictly convex minimization problem

$$\min_{\kappa_1\in[0,\kappa]}\frac{\beta}{1-\beta}\frac{|\hat{\pi}_{11}|}{2}\left\{2^{-2\kappa_1}\sigma_\Delta^2+\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 2^{-2(\kappa-\kappa_1)}\sigma_z^2\right\}. \tag{35}$$

Assuming $\hat{\pi}_{14} \neq 0$, the unique solution to this problem is

$$\kappa_1^* = \begin{cases} \kappa & \text{if } x \geq 2^{2\kappa} \\ \frac{1}{2}\kappa + \frac{1}{4}\log_2(x) & \text{if } x \in \left[2^{-2\kappa}, 2^{2\kappa}\right] \\ 0 & \text{if } x \leq 2^{-2\kappa} \end{cases}, \tag{36}$$

where $x \equiv \sigma_\Delta^2 / \left(\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 \sigma_z^2\right)$. Hence the firm's optimal choice of the information system is to observe the signals (28) – (29) with variances of noise (31) – (32) and optimal allocation of the information flow given by equation (36).

The information flow allocated to aggregate conditions, κ_1^*, is increasing in x – the ratio of the variance of the perfect information price due to aggregate shocks divided by the variance of the perfect information price due to idiosyncratic shocks. See equation (23). When idiosyncratic conditions are more variable or more important than aggregate conditions, the firm pays more attention to idiosyncratic conditions than to aggregate conditions, $\kappa_1^* < (1/2)\kappa < \kappa_2^*$.[17] In this case, price reactions to idiosyncratic shocks are strong, but price reactions to aggregate shocks are weak. See equation (33). This can explain why individual prices change by large amounts and, at the same time, individual prices react only weakly to aggregate shocks.

Computing the integral over all i of the price (33) yields the following expression for the aggregate price level

$$p_t^* = \left(1 - 2^{-2\kappa_1^*}\right)\Delta_t, \tag{37}$$

where κ_1^* is given by equation (36). The equilibrium aggregate price level is the fixed point of the mapping between the guess (26) and the actual law of motion (37). Assuming $\hat{\pi}_{13} > 0$, the unique fixed point is

$$p_t^* = \begin{cases} \frac{\left(2^{2\kappa}-1\right)\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}}{1+\left(2^{2\kappa}-1\right)\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}} q_t & \text{if } \lambda \geq 2^{-\kappa} + \left(2^{\kappa} - 2^{-\kappa}\right)\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} \\ \left(1 - 2^{-\kappa}\lambda^{-1}\right) q_t & \text{if } \lambda \in \left[2^{-\kappa}, 2^{-\kappa} + \left(2^{\kappa} - 2^{-\kappa}\right)\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}\right] \\ 0 & \text{if } \lambda \leq 2^{-\kappa} \end{cases}, \tag{38}$$

[17] More precisely, $\kappa_1^* < (1/2)\kappa < \kappa_2^*$ if and only if $x < 1$. The reason for $x < 1$ can be that idiosyncratic conditions are more variable than aggregate conditions ($\sigma_z^2 > \sigma_\Delta^2$), or that idiosyncratic conditions are more important for the pricing decision than aggregate conditions ($|\hat{\pi}_{14}/\hat{\pi}_{11}| > 1$), or both.

where $\lambda \equiv \sqrt{\left(\frac{\hat{\pi}_{13}}{\hat{\pi}_{14}}\right)^2 \frac{\sigma_q^2}{\sigma_z^2}}$. The extent to which the aggregate price level moves with nominal aggregate demand is increasing in λ. The reason is the optimal allocation of attention. When the idiosyncratic state variable has a higher variance or is more important than nominal aggregate demand, firms pay more attention to idiosyncratic conditions than to aggregate conditions. This makes prices react little to innovations in nominal aggregate demand. In addition, there is a feedback effect. When firms pay little attention to aggregate conditions, the aggregate price level moves little and therefore firms find it optimal to pay even less attention to aggregate conditions. Formally, the smaller the variance of p_t, the smaller the variance of $\Delta_t = \left(1 - \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}\right) p_t + \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} q_t$ and the smaller is the attention allocated to aggregate conditions. The feedback effect is stronger the smaller is $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$.

The feedback effect involving the optimal reallocation of attention is new in the literature. To illustrate its quantitative importance, consider a simple example. Suppose that $\sigma_q^2 = \sigma_z^2 = 10$, $(\hat{\pi}_{13}/|\hat{\pi}_{11}|) = 0.15$, $(\hat{\pi}_{14}/|\hat{\pi}_{11}|) = 1$ and $\kappa = 3$. If all other firms set the perfect information price, then $p_t = q_t$ and $\sigma_\Delta^2 = \sigma_q^2 = \sigma_z^2$. In this case, the optimal allocation of attention for an individual firm would be fifty-fifty, $\kappa_1 = (1/2)\,\kappa = \kappa_2$. In equilibrium, the variance of p_t is smaller than the variance of q_t implying $\sigma_\Delta^2 < \sigma_q^2 = \sigma_z^2$. Therefore, in equilibrium, firms allocate only 20% of their attention to aggregate conditions.

Finally, if λ is very small or very large, the equilibrium allocation of attention is a corner solution. If λ is very small, firms allocate no attention to aggregate conditions and the aggregate price level equals its deterministic trend at each point in time. If λ is very large, firms allocate all attention to aggregate conditions.

It is straightforward to compute equilibrium real aggregate demand from equation (38) and the equation $y_t = q_t - p_t$.

5 The firms' decision of what to observe

Next we show how to solve the model in the general case when the log-deviations of nominal aggregate demand and the log-deviations of the idiosyncratic state variables follow stationary Gaussian moving average processes. In this section, we focus on the firms' choice of the information system given the aggregate behavior of the economy. In the next section, we derive the rational expectations equilibrium processes for the aggregate price level and real aggregate demand. We guess that in equilibrium

$$\{p_t\} \text{ and } \{y_t\} \text{ are independent of } \{z_{it}\}, \forall i \in [0,1], \tag{39}$$

and

$$\{p_t, y_t\} \text{ is a stationary Gaussian vector process.} \tag{40}$$

These guesses will be verified in the next section.

The firm chooses the information system so as to maximize the expected discounted sum of profits (7).

Lemma 1 *(Expected discounted sum of profits) Let the profit function be given by (18) and suppose that* (39) – (40) *hold. Then*

$$E\left[\sum_{t=1}^{\infty}\beta^t \pi\left(P_{it}^*, P_t, Y_t, Z_{it}\right)\right] = E\left[\sum_{t=1}^{\infty}\beta^t \tilde{\pi}\left(p_{it}^f, p_t, y_t, z_{it}\right)\right] - \frac{\beta}{1-\beta}\frac{|\hat{\pi}_{11}|}{2} E\left[\left(p_{it}^f - p_{it}^*\right)^2\right]. \tag{41}$$

Proof. See Appendix C. ■

The expected discounted sum of profits equals the expected discounted sum of profits under perfect information (the first term on the right-hand side) minus the expected discounted sum of losses in profits due to imperfect information (the second term on the right-hand side). The expected discounted sum of losses in profits is increasing in the mean squared difference $E\left[\left(p_{it}^f - p_{it}^*\right)^2\right]$. Thus the firm chooses the information system so as to minimize this mean squared difference.

The firm has to respect the information flow constraint (9).

Lemma 2 *(Information flows) Suppose that* (39) – (40) *hold. Then*

$$\begin{aligned} \mathcal{I}\left(\{P_t\}, \{Y_t\}, \{Z_{it}\}; \{s_{it}\}\right) &= \mathcal{I}\left(\{p_t\}, \{y_t\}; \{s_{1it}\}\right) + \mathcal{I}\left(\{z_{it}\}; \{s_{2it}\}\right) && (42) \\ &\geq \mathcal{I}\left(\{\Delta_t\}; \{s_{1it}\}\right) + \mathcal{I}\left(\{z_{it}\}; \{s_{2it}\}\right) && (43) \\ &\geq \mathcal{I}\left(\{\Delta_t\}; \left\{\hat{\Delta}_{it}\right\}\right) + \mathcal{I}\left(\{z_{it}\}; \{\hat{z}_{it}\}\right), && (44) \end{aligned}$$

where $\hat{\Delta}_{it} \equiv E\left[\Delta_t | s_{1i}^t\right]$ *and* $\hat{z}_{it} \equiv E[z_{it}|s_{2i}^t]$*. If* $\{s_{1it}\}$ *and* $\{s_{2it}\}$ *are univariate processes, then inequality* (44) *holds with equality. If* $s_{1it} = \Delta_t + \varepsilon_{it}$ *where* $\{\varepsilon_{it}\}$ *is a stochastic process independent of* $\{p_t\}$*, then inequality* (43) *holds with equality.*

Proof. See Appendix D. ■

Equality (42) says that the information flow between the signals and the states of the economy equals the information flow between the first set of signals and aggregate conditions plus the information flow between the second set of signals and idiosyncratic conditions. This result follows from the independence assumption (11) and implies that one can make statements of the sort: "The firm allocates X percent of the information flow to aggregate conditions and *1-X* percent of the information flow to idiosyncratic conditions." Inequality (43) states that the signals concerning aggregate conditions contain weakly more information about the aggregate price level and real aggregate demand than they contain about the variable Δ_t alone. The relationship holds with equality when the signals concerning aggregate conditions contain information about Δ_t only. Inequality (44) states that the signals contain weakly more information than the conditional expectations computed from the signals. The relationship holds with equality when the signals are scalars.

Lemma 1, Lemma 2 and equations (22) – (23) imply that the firm's problem of choosing the information system can be stated in the following way.

Proposition 1 *(The problem of firm i) Let the profit function be given by (18) and suppose that* (39) – (40) *hold. Then the firm's problem of choosing the information system can be stated as*

$$\min_{\{(s_{1it},s_{2it})'\}\in\Gamma}\left\{E\left[\left(\Delta_t-\hat{\Delta}_{it}\right)^2\right]+\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 E\left[\left(z_{it}-\hat{z}_{it}\right)^2\right]\right\}, \tag{45}$$

subject to

$$\mathcal{I}\left(\{p_t\},\{y_t\};\{s_{1it}\}\right)+\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right)\leq\kappa. \tag{46}$$

Proof. See Appendix E. ■

The firm's problem of choosing the information system looks similar to the problem studied in Sims (2003, Section 4). There the decisionmaker chooses a process for Y_t to track X_t with loss $E\left[(X_t-Y_t)^2\right]$ subject to a constraint on the information flow between the two processes. However, there are important differences between the two problems. First, in Sims (2003) the same variables appear in the objective function and in the information flow constraint. In contrast, the objective function (45) depends on conditional expectations, $\hat{\Delta}_{it}=E\left[\Delta_t|s_{1i}^t\right]$ and $\hat{z}_{it}=E[z_{it}|s_{2i}^t]$, whereas the information flow constraint (46) applies to the underlying signals, s_{1it} and s_{2it}. Second, the problem of the firm is a collection of two quadratic control problems with a single information flow constraint. Thus the firm has to decide how to allocate the available information flow across the problem of tracking aggregate conditions and the problem of tracking idiosyncratic conditions.[18] Third, the firm tracks an endogenous variable, Δ_t. This introduces the feedback effect.

The following proposition presents a procedure for solving the firm's problem of choosing the information system.

Proposition 2 *(Solving the problem of firm i) Let the profit function be given by (18) and suppose that* (39) – (40) *hold. Then a stochastic process for the signal obtained by the following two-step procedure is an optimal information system.*

1. *Derive stochastic processes* $\left\{\hat{\Delta}_{it}^*\right\}$ *and* $\{\hat{z}_{it}^*\}$ *that solve*

$$\min_{\{\hat{\Delta}_{it}\},\{\hat{z}_{it}\}} \left\{ E\left[\left(\Delta_t - \hat{\Delta}_{it}\right)^2\right] + \left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 E\left[(z_{it} - \hat{z}_{it})^2\right] \right\}, \tag{47}$$

subject to

$$\mathcal{I}\left(\{\Delta_t\};\left\{\hat{\Delta}_{it}\right\}\right) + \mathcal{I}\left(\{z_{it}\};\{\hat{z}_{it}\}\right) \leq \kappa, \tag{48}$$

$$\left\{\Delta_t, \hat{\Delta}_{it}\right\} \text{ and } \{z_{it}, \hat{z}_{it}\} \text{ are independent,} \tag{49}$$

$$\left\{\Delta_t, \hat{\Delta}_{it}, z_{it}, \hat{z}_{it}\right\} \text{ is a stationary Gaussian vector process.} \tag{50}$$

2. *Show that there exist signals of the form*

$$s_{1it} = \Delta_t + \varepsilon_{it}, \tag{51}$$

$$s_{2it} = z_{it} + \psi_{it}, \tag{52}$$

that have the property

$$\hat{\Delta}_{it}^* = E\left[\Delta_t | s_{1i}^t\right], \tag{53}$$

$$\hat{z}_{it}^* = E[z_{it} | s_{2i}^t], \tag{54}$$

where $\{\varepsilon_{it}\}$ *and* $\{\psi_{it}\}$ *are idiosyncratic stationary Gaussian moving average processes that are mutually independent and independent of* $\{p_t\}$, $\{y_t\}$ *and* $\{z_{it}\}$.

[18] Sims (2003) only considers multivariate tracking problems within the simplified recursive framework of Section 5 of his paper.

Proof. See Appendix F. ■

The first step consists of solving a standard constrained minimization problem. This is explained in detail in Appendix H. The second step amounts to inverting a signal extraction problem. Instead of computing conditional expectations for given signals, we search for signals that generate certain processes as conditional expectations.

The processes $\left\{\hat{\Delta}_{it}^*\right\}$ and $\{\hat{z}_{it}^*\}$ have standard properties of a linear projection.

Proposition 3 *(Properties of a solution) A solution to the program* $(47)-(50)$ *satisfies*

$$E\left[\Delta_t - \hat{\Delta}_{it}^*\right] = 0, \tag{55}$$

$$E\left[z_{it} - \hat{z}_{it}^*\right] = 0, \tag{56}$$

and, for all $k = 0, 1, 2, \ldots,$

$$E\left[\left(\Delta_t - \hat{\Delta}_{it}^*\right)\hat{\Delta}_{it-k}^*\right] = 0, \tag{57}$$

$$E\left[\left(z_{it} - \hat{z}_{it}^*\right)\hat{z}_{it-k}^*\right] = 0. \tag{58}$$

Proof. See Appendix G. ■

The expected "forecast errors" are zero and the "forecast errors" are orthogonal to all current and past $\hat{\Delta}_{it}^*$ and $\hat{z}_{it}^*$. This suggests that there exist signals that have the property $(53)-(54)$. We will always verify numerically that such signals exist. Furthermore, Proposition 3 implies that, when Δ_t and z_{it} follow white noise processes, then $\hat{\Delta}_{it}^*$ and $\hat{z}_{it}^*$ also follow white noise processes. In this case, restricting the firm's choice to signals with i.i.d. noise does not change the equilibrium of the model. We used this result in Section 4.

6 Numerical solutions when exogenous processes are serially correlated

In this section we show numerical solutions of the model. We compute the solutions as follows. First, we make a guess concerning the stochastic process for the aggregate price level. Second, we solve for the optimal information system of an individual firm. Namely, we derive the stochastic processes $\left\{\hat{\Delta}_{it}^*\right\}$ and $\{\hat{z}_{it}^*\}$ and we show that there exist signals of the form $(51)-(52)$ that have the property $(53)-(54)$. See Proposition 2 and Appendix H.

Third, we compute the individual prices from equation (22) and the aggregate price level from equation (5). Fourth, we compare the stochastic process for the aggregate price level that we obtain to our guess. We update the guess until a fixed point is reached.

6.1 The benchmark economy

See Table 1 for the parameter values of the benchmark economy. The ratio $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ determines the sensitivity of individual prices to real aggregate demand, y_t. This is a standard parameter in models with monopolistic competition. Woodford (2003) recommends a value between 0.1 and 0.15. In the benchmark economy we set $(\hat{\pi}_{13}/|\hat{\pi}_{11}|) = 0.15$. Later we show how changes in $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ affect the solution.

The ratio $(\hat{\pi}_{14}/|\hat{\pi}_{11}|)$ determines the sensitivity of individual prices to the idiosyncratic state variable, z_{it}. Since changes in the value of $(\hat{\pi}_{14}/|\hat{\pi}_{11}|)$ have the same effects on equilibrium as changes in the variance of the idiosyncratic state variable, we normalize $(\hat{\pi}_{14}/|\hat{\pi}_{11}|)$ to one and we only calibrate the variance of z_{it}.

We calibrate the stochastic process for q_t using quarterly U.S. nominal GNP data from 1959:1 to 2004:1.[19] We take the natural log of the data and detrend the data by fitting a second-order polynomial in time. We then estimate the equation $q_t = \rho q_{t-1} + \nu_t$, where q_t is 100 times the deviation of the natural log of nominal GNP from its fitted trend. The estimate of ρ that we obtain is, after rounding off, 0.95 and the standard deviation of the error term is 1. This implies the moving average representation $q_t = \sum_{l=0}^{\infty} \rho^l \nu_{t-l}$. Since with geometric decay shocks die out after a very large number of periods and computing time is fast increasing with the number of lags, we approximate the estimated process by a process that dies out after twenty periods: $q_t = \sum_{l=0}^{20} a_l \nu_{t-l}$, $a_0 = 1$ and $a_l = a_{l-1} - 0.05$, for all $l = 1, ..., 20$.[20]

We calibrate the stochastic process for z_{it} so as to make the model match the average absolute size of price changes in the data. Recall that Bils and Klenow (2004) find that the median firm changes its price about every 4 months. Furthermore, Klenow and Kryvtsov (2004) find that, conditional on the occurrence of a price change, the average absolute size of

[19] The source are the National Income and Product Accounts of the United States.

[20] For the benchmark parameter values, we also solved the model without applying the approximation. We set $q_t = \sum_{l=0}^{80} \rho^l \nu_{t-l}$. While computing time was many times larger, the results were affected little.

the price change is 13.3% or 8.5% (depending on whether sales are included or excluded). We know from the analytical solution that a larger variance of the idiosyncratic state variable makes the aggregate price level more sticky. We also know that under rational inattention compared to perfect information a larger variance of the idiosyncratic state variable is required to generate a given average absolute size of price changes. We take a conservative approach and choose the standard deviation of z_{it} such that the average absolute size of price changes under perfect information is 8.5% per period.[21] This yields a standard deviation of z_{it} that is ten times the standard deviation of q_t.[22]

We set the parameter that bounds the information flow to $\kappa = 3$ bits. Our choice is motivated by two considerations. First, $\kappa = 3$ is sizable compared to the amount of uncertainty in the model. If firms in the model wanted to, they could track aggregate conditions extremely well.[23] Second, with this value of κ the model predicts a negligible difference between the price set by a firm under rational inattention and the profit-maximizing price. We find this prediction realistic.

Table 1 and Figures 1-2 summarize the results for the benchmark economy. The average absolute size of price changes is 8.2% per period. Firms allocate 94% of their attention to idiosyncratic conditions. This optimal allocation of attention implies the following price setting behavior. Figure 1 shows the impulse response of the price set by firm i to an innovation in the idiosyncratic state variable. Comparing the price reaction under rational inattention (the line with squares) to the price reaction under perfect information (the line with points), we see that under rational inattention the price reaction to idiosyncratic shocks is almost as strong and fast as under perfect information. The line with crosses is the impulse response of the price set by firm i to noise in the signal concerning idiosyncratic conditions.

Figure 2 shows the impulse response of the price set by firm i to an innovation in nominal aggregate demand. Comparing the price reaction under rational inattention (the line with

[21] Recall that one period in the model is one quarter.

[22] We assume the same rate of decay in the z_{it} process as in the q_t process.

[23] To illustrate this point, consider a simple example. Suppose that q_t was a white noise process with variance 10, which is the variance of q_t in the data. Then allocating 3 bits of information flow to tracking q_t implies that the variance of q_t conditional on the signal is 0.15. Thus the variance is reduced by 98.5%.

squares) to the price reaction under perfect information (the line with points), we see that under rational inattention the price reaction to nominal shocks is dampened and delayed. Note that, since all firms choose the same stochastic process for the signal, the line with squares is also the impulse response of the aggregate price level to an innovation in nominal aggregate demand. The aggregate price level responds weakly and slowly to innovations in nominal aggregate demand. The reasons are the following. Since idiosyncratic conditions are more variable than aggregate conditions, firms allocate most of their attention to idiosyncratic conditions. In addition, there is the feedback effect. When firms pay little attention to aggregate conditions, the aggregate price level moves little and therefore firms find it optimal to pay even less attention to aggregate conditions. As a result, the equilibrium aggregate price level under rational inattention differs markedly from the equilibrium aggregate price level under perfect information. Finally, the line with crosses in Figure 2 is the impulse response of the price set by an individual firm to noise in the signal concerning aggregate conditions.[24]

The effect of an innovation in nominal aggregate demand on real aggregate demand equals the difference between the perfect-information impulse response in Figure 2 and the rational-inattention impulse response in Figure 2. It is apparent that the real effect of an innovation in nominal aggregate demand is persistent.

Figures 3-4 show simulated price series. Figure 3 shows a sequence of prices set by an individual firm under rational inattention (diamonds) and the sequence of prices that the firm would have set if it had had perfect information (crosses). Firms in the benchmark economy track the profit-maximizing price extremely well. Figure 4 shows sequences of aggregate price levels. The equilibrium aggregate price level under rational inattention

[24] The reader interested in the impulse response of inflation to an innovation in nominal aggregate demand should note the following. In the benchmark economy, the peak response of inflation occurs on impact. Below we conduct experiments in which the impulse response of the aggregate price level becomes more dampened and delayed than in the benchmark economy. In these experiments, the impulse response of inflation becomes hump-shaped. See the experiment with larger variance of the idiosyncratic state variable (section 6.3) and the experiment with more strategic complementarity in price setting (section 6.4). We read the evidence from structural VARs as indicating clearly that the aggregate price level responds slowly to a monetary policy shock. We read the evidence as less conclusive regarding whether the impulse response of inflation to a monetary policy shock is hump-shaped (see Uhlig (2004)).

(diamonds) differs markedly from the equilibrium aggregate price level under perfect information (crosses). The reason is the optimal allocation of attention in combination with the feedback effect.

In the benchmark economy, prices react fast and by large amounts to idiosyncratic shocks, but prices react only slowly and by small amounts to nominal shocks. Thus the model can explain why individual prices change frequently and by large amounts and, at the same time, the aggregate price level responds slowly to monetary shocks. To match the large average absolute size of price changes observed in the data, idiosyncratic shocks in the model must have a large variance or must be very important for pricing decisions. This in turn implies that firms in the model allocate most of their attention to idiosyncratic conditions.

We turn to examining how changes in parameter values affect the optimal allocation of attention and the dynamics of the economy.

6.2 Increasing the variance of nominal aggregate demand

In Table 2 and Figure 5 we show what happens when the variance of nominal aggregate demand increases. Firms reallocate attention to aggregate conditions away from idiosyncratic conditions (κ_1^* increases). Firms track both aggregate and idiosyncratic conditions less well. Profits decrease. The real effects of changes in nominal aggregate demand increase. The fall in profits suggests that costs of aggregate instability in the real world may be due to the fact that aggregate instability makes the firms' tracking problem more difficult.

These predictions differ from the Lucas (1973) model. In the Lucas model, an increase in the variance of nominal aggregate demand implies that prices that firms observe become more precise signals of nominal aggregate demand and less precise signals of idiosyncratic conditions. Therefore firms in the Lucas model track nominal aggregate demand better and idiosyncratic conditions worse. The real effects of changes in nominal aggregate demand become smaller.

6.3 Increasing the variance of the idiosyncratic state variable

In Table 2 and Figure 6 we show what happens when the variance of the idiosyncratic state variable increases. Firms reallocate attention to idiosyncratic conditions away from aggregate conditions (κ_1^* decreases). Firms track both idiosyncratic and aggregate conditions less well. The reaction of the aggregate price level to a nominal shock becomes more dampened and delayed.

The model predicts that firms operating in more unstable idiosyncratic environments allocate less attention to aggregate conditions, and therefore respond more slowly to aggregate shocks. This result is consistent with the empirical finding of Bils, Klenow and Kryvtsov (2003) according to which firms that change prices relatively frequently react *more slowly* to monetary policy shocks than firms that change prices relatively infrequently. The finding of Bils, Klenow and Kryvtsov is difficult to reconcile with other models of sticky prices.

The reader may wonder whether these predictions continue to hold in a model with an endogenous κ. Suppose that firms can choose the information flow, κ, facing an increasing, strictly convex cost function, $C(\kappa)$. Now consider again the effects of increasing the variance of the idiosyncratic state variable. The marginal value of information about idiosyncratic conditions increases. Therefore firms choose a higher κ and the marginal cost of information increases. This implies that the marginal value of information about aggregate conditions has to increase as well – the information flow allocated to aggregate conditions has to fall. Hence, both idiosyncratic and aggregate conditions get tracked less well.

6.4 Changing the degree of strategic complementarity in price setting

The third and fourth example in Table 2 and Figure 7 show what happens when the ratio $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ changes.[25] As $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ decreases, the impulse response of the aggregate price level becomes more dampened and delayed. The reason is the following. Under rational inattention, the aggregate price level is less variable than nominal aggregate demand. Thus decreasing $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ lowers the variance of $\Delta_t = \left(1 - \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}\right) p_t + \frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|} q_t$. Firms react by

[25] It is common in the literature to refer to the ratio $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ as a measure of the degree of strategic complementarity in price setting, where a smaller value of $(\hat{\pi}_{13}/|\hat{\pi}_{11}|)$ corresponds to a larger degree of strategic complementarity in price setting.

reallocating attention to idiosyncratic conditions away from aggregate conditions.

6.5 The effects of serial correlation

Decreasing the serial correlation of nominal aggregate demand (holding constant its variance) leads to a fall in profits, because the firms' tracking problem becomes more difficult. This suggests that there is a payoff from "interest rate smoothing" by central banks. We obtained ambiguous predictions concerning the effect of a decrease in the serial correlation of nominal aggregate demand (holding constant its variance) on the allocation of attention. We found that the marginal return from allocating attention to aggregate conditions may go up or down. The reason is that decreasing the serial correlation of nominal aggregate demand makes firms track aggregate conditions less well (for a given allocation of attention), but also lowers the improvement in tracking that can be achieved by reallocating attention to aggregate conditions.[26]

6.6 Optimal signals

We always verify numerically that there exist signals of the form (51) – (52) that have the property (53) – (54). Figures 8 and 9 present optimal signals for the benchmark economy, by plotting the parameters of the moving average representations of Δ_t, ε_{it}, z_{it} and ψ_{it}. A common assumption in the literature is that signals are equal to the true state plus exogenous i.i.d. noise. We always find optimal signals that have the structure "true state plus a moving average noise process". However, only in some cases we find optimal signals that have the structure "true state plus i.i.d. noise". For example, the optimal idiosyncratic signal depicted in Figure 9 has the form "true state plus i.i.d. noise", but the optimal aggregate signal shown in Figure 8 does not.[27]

[26] We obtained the same results when we changed the serial correlation of the idiosyncratic state variable.

[27] Note that optimal signals are not unique. For example, applying any one-sided linear filter to the signals depicted in Figures 8 and 9 yields new optimal signals. The reason is that applying a one-sided linear filter changes neither the conditional expectations computed from the signals nor the information flow.

7 Conclusions and further research

That individual prices move frequently and by large amounts in the data does *not* imply that the aggregate price level must react fast to monetary policy shocks. When idiosyncratic conditions are more variable or more important than aggregate conditions, rationally inattentive firms optimally allocate more attention to idiosyncratic conditions than to aggregate conditions. As a result, prices react fast and by large amounts to idiosyncratic shocks, but prices react only slowly and by small amounts to nominal shocks. Innovations in nominal aggregate demand have persistent real effects.

In standard sticky price models, frequent and large price changes imply a *fast* response of the aggregate price level to nominal shocks. In our model, frequent and large price changes imply a *slow* response of the aggregate price level to nominal shocks. The *same* empirical observation on the frequency and size of individual price changes leads to the *opposite* aggregate prediction. Therefore our model can simultaneously explain the micro and the macro evidence.

Our model makes several testable predictions that we plan to compare to data. For example, according to the model, firms operating in more unstable idiosyncratic environments react more slowly to nominal shocks.

The model can be extended in a variety of directions. For example, the model in its current form abstracts from physical costs of repricing. This implies that prices in the model change every period. It would be interesting to add menu costs. This is likely to increase the real effects of nominal disturbances even further.[28]

Furthermore, it will be interesting to develop a richer general equilibrium model with rational inattention and compare its predictions to, for example, Altig, Christiano, Eichenbaum and Linde (2005), Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2003).

[28] See Dotsey, King and Wolman (1999) for a general equilibrium model with menu costs.

A Quantifying information flows

This appendix introduces the tools that we use to quantify information flows. We borrow the tools from Shannon's (1948) information theory. For a textbook on information theory, see Cover and Thomas (1991). For an application in economics, see Sims (2003).

In economics the payoff of a decisionmaker often depends on the realization of a random variable. One can quantify the uncertainty by using the concept of *entropy*. The entropy of a random variable is a measure of the uncertainty of the random variable. The entropy $H(X)$ of a random variable X with density function $p(X)$ is defined by $H(X) = -E[\log_2 p(X)]$. Entropy is measured in bits. For example, the entropy of a normally distributed random variable X with variance σ^2 is

$$H(X) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right).$$

In this simple example, entropy is a strictly increasing function of the variance.[29]

The definition of entropy extends to random vectors. In the definition of entropy, simply replace the density function by the joint density function. For example, applying the definition of entropy to a set of random variables $X_1, ..., X_T$ that have a multivariate normal distribution with covariance matrix Ω_{XX} yields

$$H(X_1, ..., X_T) = \frac{1}{2}\log_2[(2\pi e)^T \det \Omega_{XX}]. \tag{59}$$

The entropy of the random vector depends on the number of random variables and on their covariance matrix. A larger determinant of the covariance matrix implies a larger entropy. For given variances, the entropy is largest when the random variables are uncorrelated.

In economics a decisionmaker often observes a random vector that is correlated with the random vector of interest. One can quantify the conditional uncertainty by using the concept of *conditional entropy*. For example, suppose that a decisionmaker is interested in $X_1, ..., X_T$ and observes $Y_1, ..., Y_T$, where $X_1, ..., X_T$ and $Y_1, ..., Y_T$ have a multivariate normal distribution with covariance matrix Ω. Then the entropy of $X_1, ..., X_T$ conditional

[29] The definition of entropy can be derived from axioms – requirements that a "reasonable" measure of uncertainty should satisfy (see e.g. Ash (1990)). Moreover, entropy arises as the answer to a number of natural questions in communication theory and statistics (see e.g. Cover and Thomas (1991)).

on $Y_1, ..., Y_T$ is

$$H(X_1, ..., X_T \mid Y_1, ..., Y_T) = \frac{1}{2}\log_2\{(2\pi e)^T \det[\Omega_{XX} - \Omega_{XY}\Omega_{YY}^{-1}\Omega_{YX}]\}. \tag{60}$$

The expression in square brackets is the covariance matrix of $X_1, ..., X_T$ conditional on $Y_1, ..., Y_T$.

Now one can quantify the amount of information that one random vector contains about another random vector. *Mutual information* is the reduction in the uncertainty of one random vector due to the knowledge of another random vector. The mutual information between $X_1, ..., X_T$ and $Y_1, ..., Y_T$ is

$$I(X_1, ..., X_T; Y_1, ..., Y_T) = H(X_1, ..., X_T) - H(X_1, ..., X_T \mid Y_1, ..., Y_T). \tag{61}$$

It is also straightforward to quantify the information flow between stochastic processes. Let $X_1, ..., X_T$ denote the first T elements of the stochastic process $\{X_t\}$. Let $Y_1, ..., Y_T$ denote the first T elements of the stochastic process $\{Y_t\}$. The processes $\{X_t\}$ and $\{Y_t\}$ can be vector processes. The information flow between the processes $\{X_t\}$ and $\{Y_t\}$ can be defined by

$$\mathcal{I}(\{X_t\}; \{Y_t\}) = \lim_{T\to\infty}\frac{1}{T}I(X_1, ..., X_T; Y_1, ..., Y_T). \tag{62}$$

The information flow between stochastic processes is the average amount of information per unit of time that one stochastic process contains about another stochastic process. The limit in (62) exists when the processes $\{X_t\}$ and $\{Y_t\}$ are jointly stationary.

In the Gaussian case, an analytical expression exists for the information flow. If $\{X_t\}$ and $\{Y_t\}$ are univariate, jointly stationary, jointly Gaussian processes with absolutely summable autocovariance matrices then

$$\mathcal{I}(\{X_t\}; \{Y_t\}) = -\frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2[1 - \mathcal{C}_{X,Y}(\omega)]\,d\omega, \tag{63}$$

where $\mathcal{C}_{X,Y}(\omega)$ is the coherence between the processes $\{X_t\}$ and $\{Y_t\}$ at frequency ω. This follows from equations $(59)-(62)$ and the asymptotic properties of determinants of Toeplitz matrices. See Cover and Thomas (1991, pp. 273-274), Gray (2002, pp. 62-63) or Sims (2003). Note that the coherence lies between zero and one, $0 \leq \mathcal{C}_{X,Y}(\omega) \leq 1$ for all ω. It follows that the information flow in (63) is bounded below by zero and is unbounded above.

B Information flow constraint in the white noise case

Assumptions (10) – (11) imply

$$\mathcal{I}\left(\{P_t\},\{Y_t\},\{Z_{it}\};\{s_{it}\}\right)=\mathcal{I}\left(\{p_t\},\{y_t\};\{s_{1it}\}\right)+\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right).$$

This general result is proved below. See Lemma 2. Furthermore, equations (26) – (27) imply that $\{p_t\}$ and $\{y_t\}$ can be calculated from $\{\Delta_t\}$ and vice versa. It follows that

$$\mathcal{I}\left(\{p_t\},\{y_t\};\{s_{1it}\}\right)=\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right).$$

The signal concerning aggregate conditions is given by equation (28). Equation (63) applies

$$\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right)=-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2\left[1-\mathcal{C}_{\Delta,s_{1i}}(\omega)\right]d\omega,$$

where $\mathcal{C}_{\Delta,s_{1i}}(\omega)$ is the coherence between the processes $\{\Delta_t\}$ and $\{s_{1it}\}$ at frequency ω. The processes $\{\Delta_t\}$ and $\{s_{1it}\}$ are white noise processes. Therefore the coherence simply equals the squared correlation coefficient and

$$\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right)=-\frac{1}{2}\log_2\left(1-\rho^2_{\Delta,s_{1i}}\right).$$

Using (28) yields

$$\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right)=\frac{1}{2}\log_2\left(\frac{\sigma^2_\Delta}{\sigma^2_\varepsilon}+1\right).$$

The same arguments yield

$$\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right)=\frac{1}{2}\log_2\left(\frac{\sigma^2_z}{\sigma^2_\psi}+1\right).$$

The information flow constraint becomes

$$\frac{1}{2}\log_2\left(\frac{\sigma^2_\Delta}{\sigma^2_\varepsilon}+1\right)+\frac{1}{2}\log_2\left(\frac{\sigma^2_z}{\sigma^2_\psi}+1\right)\leq\kappa.$$

C Proof of lemma 1

First, when the profit function is given by equation (18), the expected discounted sum of profits equals

$$\begin{aligned}
E\left[\sum_{t=1}^{\infty}\beta^{t}\pi\left(P_{it}^{*},P_{t},Y_{t},Z_{it}\right)\right] &= E\left[\sum_{t=1}^{\infty}\beta^{t}\tilde{\pi}\left(p_{it}^{*},p_{t},y_{t},z_{it}\right)\right] \\
&= E\left[\sum_{t=1}^{\infty}\beta^{t}\tilde{\pi}\left(p_{it}^{f},p_{t},y_{t},z_{it}\right)\right]-E\left[\sum_{t=1}^{\infty}\beta^{t}\tilde{\pi}\left(p_{it}^{f},p_{t},y_{t},z_{it}\right)\right] \\
&\quad +E\left[\sum_{t=1}^{\infty}\beta^{t}\tilde{\pi}\left(p_{it}^{*},p_{t},y_{t},z_{it}\right)\right] \\
&= E\left[\sum_{t=1}^{\infty}\beta^{t}\tilde{\pi}\left(p_{it}^{f},p_{t},y_{t},z_{it}\right)\right]-E\left[\sum_{t=1}^{\infty}\beta^{t}\frac{|\hat{\pi}_{11}|}{2}\left(p_{it}^{f}-p_{it}^{*}\right)^{2}\right],
\end{aligned}$$

where the last equality follows from equation (21). Second, the difference between the price set under perfect information (20) and the price set under imperfect information (19) equals

$$p_{it}^{f}-p_{it}^{*}=p_{it}^{f}-E\left[p_{it}^{f}|s_{i}^{t}\right].$$

The joint normality of p_{it}^{f} and $s_{i}^{t}=\{s_{i}^{1},s_{i2},...,s_{it}\}$ implies that the conditional expectation equals the linear projection. The joint stationarity of p_{it}^{f} and $s_{i}^{t}=\{s_{i}^{1},s_{i2},...,s_{it}\}$ and assumption (13) imply that the linear projection coefficients are independent of t

$$p_{it}^{f}-p_{it}^{*}=p_{it}^{f}-\left[\mu+\alpha\left(L\right)s_{it}\right],$$

where μ is a constant and $\alpha\left(L\right)$ is an infinite order vector lag polynomial. Hence, $p_{it}^{f}-p_{it}^{*}$ follows a stationary process. Thus $E\left[\left(p_{it}^{f}-p_{it}^{*}\right)^{2}\right]$ does not depend on t and

$$E\left[\sum_{t=1}^{\infty}\beta^{t}\frac{|\hat{\pi}_{11}|}{2}\left(p_{it}^{f}-p_{it}^{*}\right)^{2}\right]=\frac{\beta}{1-\beta}\frac{|\hat{\pi}_{11}|}{2}E\left[\left(p_{it}^{f}-p_{it}^{*}\right)^{2}\right].$$

D Proof of lemma 2

First, since $\{P_{t}\},\{Y_{t}\},\{Z_{it}\}$ can be calculated from $\{p_{t}\},\{y_{t}\},\{z_{it}\}$ and vice versa, we have

$$\mathcal{I}\left(\{P_{t}\},\{Y_{t}\},\{Z_{it}\};\{s_{it}\}\right)=\mathcal{I}\left(\{p_{t}\},\{y_{t}\},\{z_{it}\};\{s_{it}\}\right).$$

Applying the definition of the information flow (62) yields

$$\mathcal{I}\left(\{p_t\},\{y_t\},\{z_{it}\};\{s_{it}\}\right)=\lim_{T\rightarrow\infty}\frac{1}{T}I\left(p^T,y^T,z_i^T;s_i^T\right),$$

where $p^T\equiv(p_1,\ldots,p_T)$, $y^T\equiv(y_1,\ldots,y_T)$, $z_i^T\equiv(z_{i1},\ldots,z_{iT})$ and $s_i^T\equiv\left(s_i^1,s_{i2},\ldots,s_{iT}\right)$. The assumption (10) implies

$$I\left(p^T,y^T,z_i^T;s_i^T\right)=I\left(p^T,y^T,z_i^T;s_{1i}^T,s_{2i}^T\right),$$

equation (61) implies

$$I\left(p^T,y^T,z_i^T;s_{1i}^T,s_{2i}^T\right)=H\left(p^T,y^T,z_i^T\right)-H\left(p^T,y^T,z_i^T|s_{1i}^T,s_{2i}^T\right),$$

and conditional entropy equals

$$H\left(p^T,y^T,z_i^T|s_{1i}^T,s_{2i}^T\right)=H\left(p^T,y^T,z_i^T,s_{1i}^T,s_{2i}^T\right)-H\left(s_{1i}^T,s_{2i}^T\right).$$

See, for example, Cover and Thomas (1991), p. 230, equation 9.33. We arrive at

$$I\left(p^T,y^T,z_i^T;s_i^T\right)=H\left(p^T,y^T,z_i^T\right)-H\left(p^T,y^T,z_i^T,s_{1i}^T,s_{2i}^T\right)+H\left(s_{1i}^T,s_{2i}^T\right).$$

The entropy of independent random variables or independent random vectors equals the sum of the entropies. See, for example, Cover and Thomas (1991), p. 232, equation 9.59. Therefore assumption (11) implies

$$\begin{aligned}I\left(p^T,y^T,z_i^T;s_i^T\right) &= H\left(p^T,y^T\right)+H\left(z_i^T\right)-H\left(p^T,y^T,s_{1i}^T\right)-H\left(z_i^T,s_{2i}^T\right)\\ &\quad +H\left(s_{1i}^T\right)+H\left(s_{2i}^T\right).\end{aligned}$$

The last equation can also be expressed as

$$I\left(p^T,y^T,z_i^T;s_i^T\right)=I\left(p^T,y^T;s_{1i}^T\right)+I\left(z_i^T;s_{2i}^T\right).$$

Dividing by T on both sides and taking the limit as $T\rightarrow\infty$ yields

$$\mathcal{I}\left(\{p_t\},\{y_t\},\{z_{it}\};\{s_{it}\}\right)=\mathcal{I}\left(\{p_t\},\{y_t\};\{s_{1it}\}\right)+\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right).$$

Second, since $\{p_t\},\{y_t\}$ can be calculated from $\{p_t\},\{\Delta_t\}$ and vice versa, we have

$$\mathcal{I}\left(\{p_t\},\{y_t\};\{s_{1it}\}\right)=\mathcal{I}\left(\{p_t\},\{\Delta_t\};\{s_{1it}\}\right).$$

Applying the definition of the information flow (62) yields

$$\mathcal{I}\left(\{p_t\},\{\Delta_t\};\{s_{1it}\}\right)=\lim_{T\to\infty}\frac{1}{T}I\left(p^T,\Delta^T;s_{1i}^T\right).$$

Equation (61) implies

$$I\left(p^T,\Delta^T;s_{1i}^T\right)=H\left(p^T,\Delta^T\right)-H\left(p^T,\Delta^T|s_{1i}^T\right).$$

The terms on the right-hand side can be expressed as

$$\begin{aligned}H\left(p^T,\Delta^T\right)&=H\left(\Delta^T\right)+H\left(p^T|\Delta^T\right),\\ H\left(p^T,\Delta^T|s_{1i}^T\right)&=H\left(\Delta^T|s_{1i}^T\right)+H\left(p^T|\Delta^T,s_{1i}^T\right).\end{aligned}$$

See, for example, Cover and Thomas (1991), p. 230, equation 9.33. We arrive at

$$I\left(p^T,\Delta^T;s_{1i}^T\right)=H\left(\Delta^T\right)+H\left(p^T|\Delta^T\right)-H\left(\Delta^T|s_{1i}^T\right)-H\left(p^T|\Delta^T,s_{1i}^T\right).$$

The last equation can also be expressed as

$$I\left(p^T,\Delta^T;s_{1i}^T\right)=I\left(\Delta^T;s_{1i}^T\right)+I\left(p^T;s_{1i}^T|\Delta^T\right).$$

Finally,

$$I\left(p^T;s_{1i}^T|\Delta^T\right)\geq 0,$$

with equality if and only if p^T and s_{1i}^T are conditionally independent given Δ^T. See, for example, Cover and Thomas (1991), p. 232, first corollary to theorem 9.6.1. Hence,

$$\mathcal{I}\left(\{p_t\},\{\Delta_t\};\{s_{1it}\}\right)\geq\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right),$$

with equality if p^T and s_{1i}^T are conditionally independent given Δ^T for all T.

Third, applying the definition of the information flow (62) yields

$$\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right)=\lim_{T\to\infty}\frac{1}{T}I\left(z_i^T;s_{2i}^T\right).$$

Equation (61) implies

$$I\left(z_i^T;s_{2i}^T\right)=H\left(z_i^T\right)-H\left(z_i^T|s_{2i}^T\right).$$

Furthermore, since $\hat{z}_i^T=\left(\hat{z}_{i1},\ldots,\hat{z}_{iT}\right)$ can be calculated from s_{2i}^T, we have

$$H\left(z_i^T|s_{2i}^T\right)=H\left(z_i^T|s_{2i}^T,\hat{z}_i^T\right).$$

In addition, since conditioning reduces entropy, we have

$$H\left(z_i^T|s_{2i}^T,\hat{z}_i^T\right)\leq H\left(z_i^T|\hat{z}_i^T\right).$$

See, for example, Cover and Thomas (1991), p. 232, second corollary to theorem 9.6.1. We arrive at

$$I\left(z_i^T;s_{2i}^T\right)\geq I\left(z_i^T;\hat{z}_i^T\right).$$

Dividing by T on both sides and taking the limit as $T\to\infty$ yields

$$\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right)\geq\mathcal{I}\left(\{z_{it}\};\{\hat{z}_{it}\}\right).$$

The same arguments yield

$$\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right)\geq\mathcal{I}\left(\{\Delta_t\};\left\{\hat{\Delta}_{it}\right\}\right).$$

Next, suppose that $\{s_{1it}\}$ is a univariate process. Then

$$\hat{\Delta}_{it}=\mu_1+\alpha_1(L)\,s_{1it},$$

where μ_1 is a constant and $\alpha_1(L)$ is an infinite order lag polynomial. See proof of Lemma 1. Thus $\left\{\hat{\Delta}_{it}\right\}$ is obtained from $\{s_{1it}\}$ by applying a one-sided linear filter (and possibly adding a constant). Standard results on linear filters imply

$$\mathcal{C}_{\Delta,\hat{\Delta}_i}(\omega)=\mathcal{C}_{\Delta,s_{1i}}(\omega),$$

where $\mathcal{C}_{\Delta,\hat{\Delta}_i}(\omega)$ denotes the coherence between the processes $\{\Delta_t\}$ and $\left\{\hat{\Delta}_{it}\right\}$ at frequency ω. This result in combination with equation (63) yields

$$\mathcal{I}\left(\{\Delta_t\};\left\{\hat{\Delta}_{it}\right\}\right)=\mathcal{I}\left(\{\Delta_t\};\{s_{1it}\}\right).$$

The same arguments yield that, if $\{s_{2it}\}$ is a univariate process, then

$$\mathcal{I}\left(\{z_{it}\};\{\hat{z}_{it}\}\right)=\mathcal{I}\left(\{z_{it}\};\{s_{2it}\}\right).$$

E Proof of proposition 1

The objective function (45) follows from Lemma 1, equations (22) – (23) and the orthogonality of $\Delta_t-\hat{\Delta}_{it}$ and $z_{it}-\hat{z}_{it}$. The information flow constraint (46) follows from equation (42).

F Proof of proposition 2

First, when the profit function is given by (18) and (39) – (40) hold, the firms' problem of choosing the information system can be stated as the program (45) – (46). See Proposition 1. The objective function (45) and the objective function (47) are identical. Furthermore, the constraint (46) implies the inequality (48). See Lemma 2. In addition, assumption (11) implies condition (49) and assumptions (12) – (13) imply condition (50). Hence, a solution to the program (45) – (46) cannot make the firm better off than a solution to the program (47) – (50).

Second, signals of the form (51) – (52) are an element of the set Γ. Furthermore, signals of the form (51) – (52) imply that inequalities (43) – (44) hold with equality. Hence, if signals of the form (51) – (52) have the property (53) – (54), then they are a solution to the program (45) – (46).

G Proof of proposition 3

First, the mean of the process $\left\{\hat{\Delta}_{it}\right\}$ affects $E\left[\left(\Delta_t - \hat{\Delta}_{it}\right)^2\right]$ but does not affect the information flow $\mathcal{I}\left(\{\Delta_t\};\left\{\hat{\Delta}_{it}\right\}\right)$. See equation (63). Therefore a solution to the program (47) – (50) has to satisfy

$$E\left[\hat{\Delta}^*_{it}\right] = E\left[\Delta_t\right].$$

The same arguments yield that a solution to the program (47) – (50) has to satisfy

$$E\left[\hat{z}^*_{it}\right] = E\left[z_{it}\right].$$

Second, a solution to the program (47) – (50) has to satisfy, for all $k = 0, 1, 2, \ldots$,

$$E\left[\left(\Delta_t - \hat{\Delta}^*_{it}\right)\hat{\Delta}^*_{it-k}\right] = 0.$$

Take a process $\left\{\hat{\Delta}'_{it}\right\}$ that does not have this property. Formally, for some $k \in \{0, 1, 2, \ldots\}$,

$$E\left[\left(\Delta_t - \hat{\Delta}'_{it}\right)\hat{\Delta}'_{it-k}\right] \neq 0.$$

Then one can define a new process $\left\{\hat{\Delta}''_{it}\right\}$ as follows

$$\hat{\Delta}''_{it} = \left(1 + \alpha L^k\right)\hat{\Delta}'_{it},$$

where L is the lag operator and α is the projection coefficient in the linear projection of $\Delta_t - \hat{\Delta}'_{it}$ on $\hat{\Delta}'_{it-k}$. The new process has the property

$$\mathcal{I}\left(\{\Delta_t\}; \left\{\hat{\Delta}''_{it}\right\}\right) = \mathcal{I}\left(\{\Delta_t\}; \left\{\hat{\Delta}'_{it}\right\}\right),$$

because applying a one-sided linear filter to a stochastic process does not change the information flow. See proof of Lemma 2. Furthermore, the new process has the property

$$E\left[\left(\Delta_t - \hat{\Delta}''_{it}\right)^2\right] < E\left[\left(\Delta_t - \hat{\Delta}'_{it}\right)^2\right].$$

Thus the process $\left\{\hat{\Delta}'_{it}\right\}$ cannot be a solution to the program (47) – (50). It follows that a solution has to satisfy, for all $k = 0, 1, 2, \ldots$,

$$E\left[\left(\Delta_t - \hat{\Delta}^*_{it}\right)\hat{\Delta}^*_{it-k}\right] = 0.$$

The same arguments yield that a solution has to satisfy, for all $k = 0, 1, 2, \ldots$,

$$E\left[\left(z_{it} - \hat{z}^*_{it}\right)\hat{z}^*_{it-k}\right] = 0.$$

H Numerical solution procedure

Let the moving average representations for q_t and z_{it} be given by

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

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

where $\{\nu_t\}$ and $\{\xi_{it}\}$ are Gaussian white noise processes with unit variance. We make a guess concerning the stochastic process for the aggregate price level

$$p_t = \sum_{l=0}^{\infty} c_l \nu_{t-l}. \tag{64}$$

Applying Proposition 2 , we solve the following constrained optimization problem

$$\min_{d,f,g,h}\left\{E\left[\left(\Delta_t - \hat{\Delta}_{it}\right)^2\right] + \left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 E\left[\left(z_{it} - \hat{z}_{it}\right)^2\right]\right\},$$

subject to

$$\left\{-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2\left[1-\mathcal{C}_{\Delta,\hat{\Delta}_i}(\omega)\right]d\omega\right\}+\left\{-\frac{1}{4\pi}\int_{-\pi}^{\pi}\log_2\left[1-\mathcal{C}_{z_i,\hat{z}_i}(\omega)\right]d\omega\right\}\leq\kappa,$$

with

$$\Delta_t=\left(1-\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}\right)\sum_{l=0}^{\infty}c_l\nu_{t-l}+\frac{\hat{\pi}_{13}}{|\hat{\pi}_{11}|}\sum_{l=0}^{\infty}a_l\nu_{t-l},$$

$$\hat{\Delta}_{it}=\sum_{l=0}^{\infty}d_l\nu_{t-l}+\sum_{l=0}^{\infty}f_l\eta_{it-l},$$

$$\hat{z}_{it}=\sum_{l=0}^{\infty}g_l\xi_{it-l}+\sum_{l=0}^{\infty}h_l\zeta_{it-l},$$

where $\{\eta_{it}\}$ and $\{\zeta_{it}\}$ are Gaussian white noise processes with unit variance that are mutually independent and independent of $\{\nu_t\}$ and $\{\xi_{it}\}$. Here we make use of equation (63) to express information flow as a function of coherence.

Consider, as an example, the choice of the g_l and h_l, for all $l=0,1,\ldots$. The following simplifications are helpful. Observe that in the objective

$$\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2E\left[(z_{it}-\hat{z}_{it})^2\right]=\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2\left[\sum_{l=0}^{\infty}(b_l-g_l)^2+\sum_{l=0}^{\infty}h_l^2\right],$$

and in the constraint

$$\mathcal{C}_{z_i,\hat{z}_i}(\omega)=\frac{\left[\frac{G\left(e^{-i\omega}\right)G\left(e^{i\omega}\right)}{H\left(e^{-i\omega}\right)H\left(e^{i\omega}\right)}\right]}{\left[\frac{G\left(e^{-i\omega}\right)G\left(e^{i\omega}\right)}{H\left(e^{-i\omega}\right)H\left(e^{i\omega}\right)}\right]+1},$$

where the polynomials $G\left(e^{i\omega}\right)$ and $H\left(e^{i\omega}\right)$ are defined as $G\left(e^{i\omega}\right)\equiv g_0+g_1e^{i\omega}+g_2e^{i2\omega}+\ldots$ and $H\left(e^{i\omega}\right)\equiv h_0+h_1e^{i\omega}+h_2e^{i2\omega}+\ldots$. The first-order condition with respect to g_l for any l is

$$\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 2(b_l-g_l)=-\frac{\mu}{4\pi\ln(2)}\int_{-\pi}^{\pi}\frac{\partial\ln\left[1-\mathcal{C}_{z_i,\hat{z}_i}(\omega)\right]}{\partial g_l}d\omega,$$

where μ is the Lagrange multiplier. The first-order condition with respect to h_l for any l is

$$\left(\frac{\hat{\pi}_{14}}{\hat{\pi}_{11}}\right)^2 2h_l=\frac{\mu}{4\pi\ln(2)}\int_{-\pi}^{\pi}\frac{\partial\ln\left[1-\mathcal{C}_{z_i,\hat{z}_i}(\omega)\right]}{\partial h_l}d\omega.$$

We obtain a system of nonlinear equations in d, f, g, h and μ that we solve numerically.

References

[1] Adam, Klaus (2004): "Optimal Monetary Policy with Imperfect Common Knowledge." Discussion paper, European Central Bank.

[2] Altig, David, Lawrence J. Christiano, Martin Eichenbaum and Jesper Linde (2005): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." NBER working paper 11034.

[3] Ash, Robert B. (1990): "Information Theory." Dover Publications, New York.

[4] Bils, Mark and Peter J. Klenow (2004): "Some Evidence on the Importance of Sticky Prices." Journal of Political Economy, 112, 947-985.

[5] Bils, Mark, Peter J. Klenow and Oleksiy Kryvtsov (2003): "Sticky Prices and Monetary Policy Shocks." Federal Reserve Bank of Minneapolis Quarterly Review, 27(1), 2-9.

[6] Calvo, Guillermo A. (1983): "Staggered Prices in a Utility-Maximizing Framework." Journal of Monetary Economics, 12, 383-398.

[7] Christiano Lawrence J., Martin Eichenbaum and Charles L. Evans (1999): "Monetary Policy Shocks: What Have We Learned and to What End?" In "Handbook of Macroeconomics", edited by John B. Taylor and Michael Woodford, Elsevier, New York.

[8] Christiano, Lawrence J., Martin Eichenbaum and Charles Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." Forthcoming, Journal of Political Economy.

[9] Cover, Thomas M. and Joy A. Thomas (1991): "Elements of Information Theory." John Wiley and Sons, New York.

[10] Dotsey, Michael, Robert G. King and Alexander L. Wolman (1999): "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output." Quarterly Journal of Economics, 114, 655-690.

[11] Galí, Jordi and Mark Gertler (1999): "Inflation Dynamics: A Structural Econometric Analysis." Journal of Monetary Economics, 44, 195-222.

[12] Golosov, Mikhail and Robert E. Lucas, Jr. (2003): "Menu Costs and Phillips Curves." NBER working paper 10187.

[13] Gray, Robert M. (2002): "Toeplitz and Circulant Matrices: A Review." Discussion paper, Stanford University, http://www-isl.stanford.edu/~gray/toeplitz.pdf.

[14] Gumbau-Brisa, Fabia (2003): "Inflation Persistence and Learning in General Equilibrium." Discussion paper, Harvard University.

[15] Hellwig, Christian (2002): "Public Announcements, Adjustment Delays and the Business Cycle." Discussion paper, University of California at Los Angeles.

[16] Klenow, Peter J. and Oleksiy Kryvtsov (2004): "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" Discussion paper, Stanford University and Bank of Canada.

[17] Leeper, Eric M., Christopher A. Sims, and Tao Zha (1996): "What Does Monetary Policy Do?" Brookings Papers on Economic Activity, 1996:2, 1-63.

[18] Lucas, Robert E. Jr. (1973): "Some International Evidence on Output-Inflation Trade-Offs." American Economic Review, 63, 326-334.

[19] Lucas, Robert E. Jr. (1977): "Understanding Business Cycles." Carnegie-Rochester Conference Series on Public Policy, 5, 7-29.

[20] Mankiw, N. Gregory and Ricardo Reis (2002): "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." Quarterly Journal of Economics, 117, 1295-1328.

[21] Moscarini, Giuseppe (2003): "Limited Information Capacity as a Source of Inertia." Forthcoming, Journal of Economic Dynamics and Control.

[22] Reis, Ricardo (2004): "Inattentive Producers." Discussion paper, Harvard University.

[23] Shannon, Claude E. (1948): "A Mathematical Theory of Communication." The Bell System Technical Journal, 27, 623-656.

[24] Sims, Christopher A. (1998): "Stickiness." Carnegie-Rochester Conference Series on Public Policy, 49, 317-356.

[25] Sims, Christopher A. (2003): "Implications of Rational Inattention." Journal of Monetary Economics, 50, 665-690.

[26] Smets, Frank and Raf Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." Journal of European Economic Association, 1, 1123-1175.

[27] Uhlig, Harald (1996): "A Law of Large Numbers for Large Economies," Economic Theory, 8, 41-50.

[28] Uhlig, Harald (2004): "What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." Discussion paper, Humboldt University in Berlin.

[29] Woodford, Michael (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy." In P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds: "Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps." Princeton University Press, Princeton.

[30] Woodford, Michael (2003): "Interest and Prices. Foundations of a Theory of Monetary Policy." Princeton University Press, Princeton and Oxford.

Table 1: Parameters and main results for the benchmark economy

Parameters	Interpretation
$(\widehat{\pi}_{13}/\left\|\widehat{\pi}_{11}\right\|) = 0.15$	Determines the sensitivity of prices to real aggregate demand y_t
$(\widehat{\pi}_{14}/\left\|\widehat{\pi}_{11}\right\|) = 1$	Determines the sensitivity of prices to the idiosyncratic state variable z_{it}
$q_t = \sum_{l=0}^{20} a_l \nu_{t-l}, \nu_t \sim N(0,1)$, with $a_0 = 1, a_l = a_{l-1} - 0.05, l = 1, ..., 20$	The MA representation of nominal aggregate demand q_t
$z_{it} = \sum_{l=0}^{20} b_l \xi_{it-l}, \xi_{it} \sim N(0,1)$, with $b_0 = 10, b_l = b_{l-1} - 0.5, l = 1, ..., 20$	The MA representation of the idiosyncratic state variable z_{it}
$\kappa = 3$	The upper bound on the information flow
Main results	**Interpretation**
8.2%	The average absolute size of price changes per period
$\kappa_1^* = 0.19, \kappa_2^* = 2.81$	94% of attention allocated to the idiosyncratic state
$E\left[\left(\Delta_t - \hat{\Delta}_{it}^*\right)^2\right] = 0.39$	Expected loss from imperfect tracking of Δ_t
$\left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 E\left[(z_{it} - \widehat{z}_{it}^*)^2\right] = 2.1$	Expected loss from imperfect tracking of z_{it}

Table 2: Varying parameter values	
Changes in parameter values relative to the benchmark economy in Table 1	**Changes in results**
$a_0 = 50, a_l = a_{l-1} - 2.5, l = 1, ..., 20$ Larger variance of nominal aggregate demand	The average absolute size of price changes per period is 35% κ_1^* increases to 76% of κ $E\left[\left(\Delta_t - \hat{\Delta}_{it}^*\right)^2\right] = 75.6, \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 E\left[\left(z_{it} - \widehat{z}_{it}^*\right)^2\right] = 54$
$b_0 = 12, b_l = b_{l-1} - 0.6, l = 1, ..., 20$ Larger variance of the idiosyncratic state variable	The average absolute size of price changes per period is 10% κ_1^* decreases to 4% of κ $E\left[\left(\Delta_t - \hat{\Delta}_{it}^*\right)^2\right] = 0.44, \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 E\left[\left(z_{it} - \widehat{z}_{it}^*\right)^2\right] = 2.7$
$\left(\widehat{\pi}_{13}/\left\|\widehat{\pi}_{11}\right\|\right) = 0.1$ More strategic complementarity in price setting	The average absolute size of price changes per period is 8.2% κ_1^* decreases to 5% of κ $E\left[\left(\Delta_t - \hat{\Delta}_{it}^*\right)^2\right] = 0.31, \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 E\left[\left(z_{it} - \widehat{z}_{it}^*\right)^2\right] = 1.9$
$\left(\widehat{\pi}_{13}/\left\|\widehat{\pi}_{11}\right\|\right) = 0.3$ Less strategic complementarity in price setting	The average absolute size of price changes per period is 8.2% κ_1^* increases to 9% of κ $E\left[\left(\Delta_t - \hat{\Delta}_{it}^*\right)^2\right] = 0.62, \left(\frac{\widehat{\pi}_{14}}{\widehat{\pi}_{11}}\right)^2 E\left[\left(z_{it} - \widehat{z}_{it}^*\right)^2\right] = 2.3$

Figure 1: Impulse responses of an individual price to an innovation in the idiosycratic state variable, benchmark economy

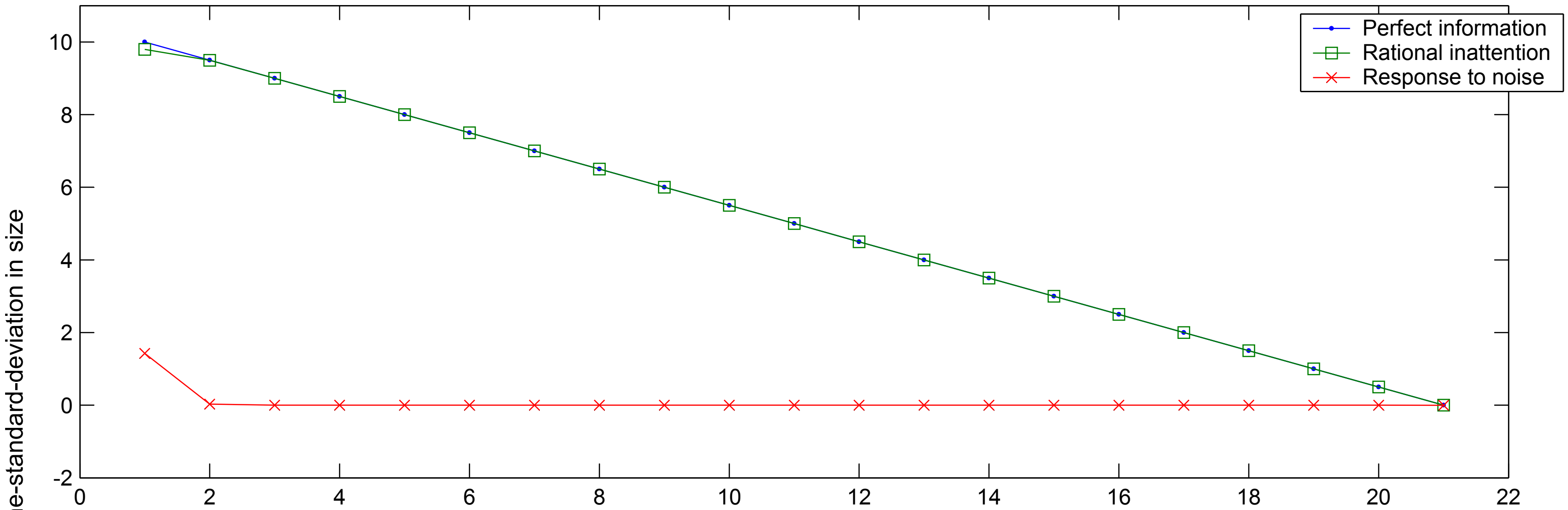
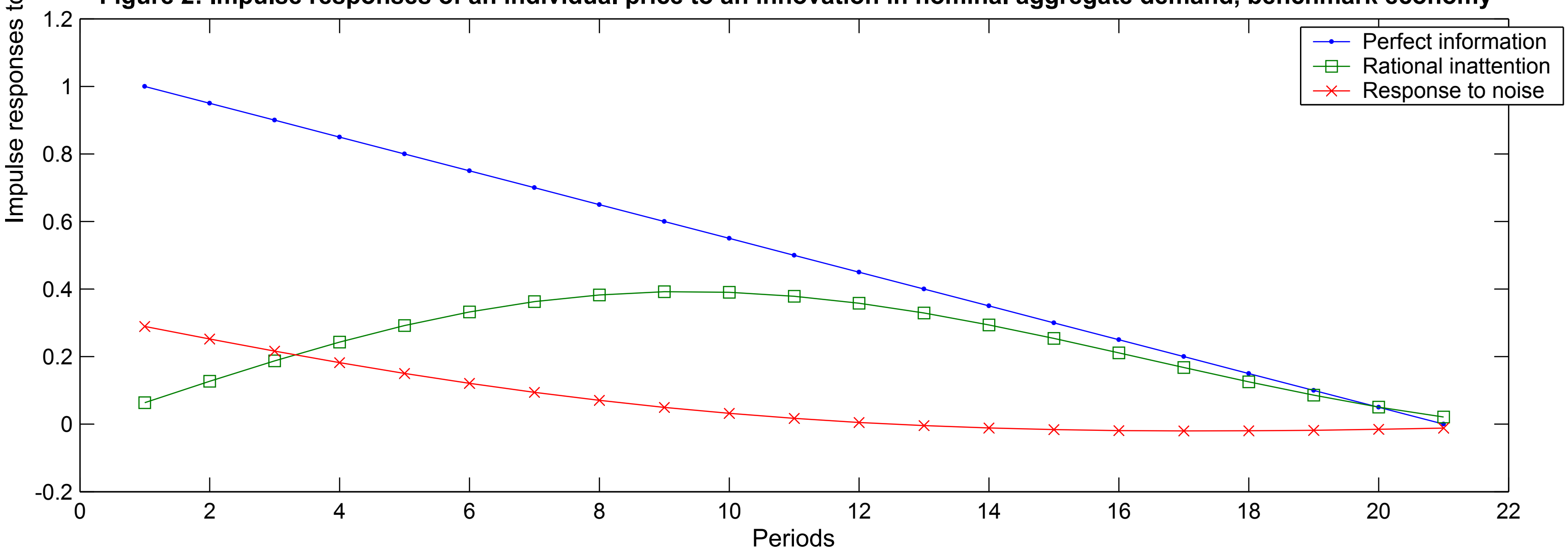


Figure 2: Impulse responses of an individual price to an innovation in nominal aggregate demand, benchmark economy



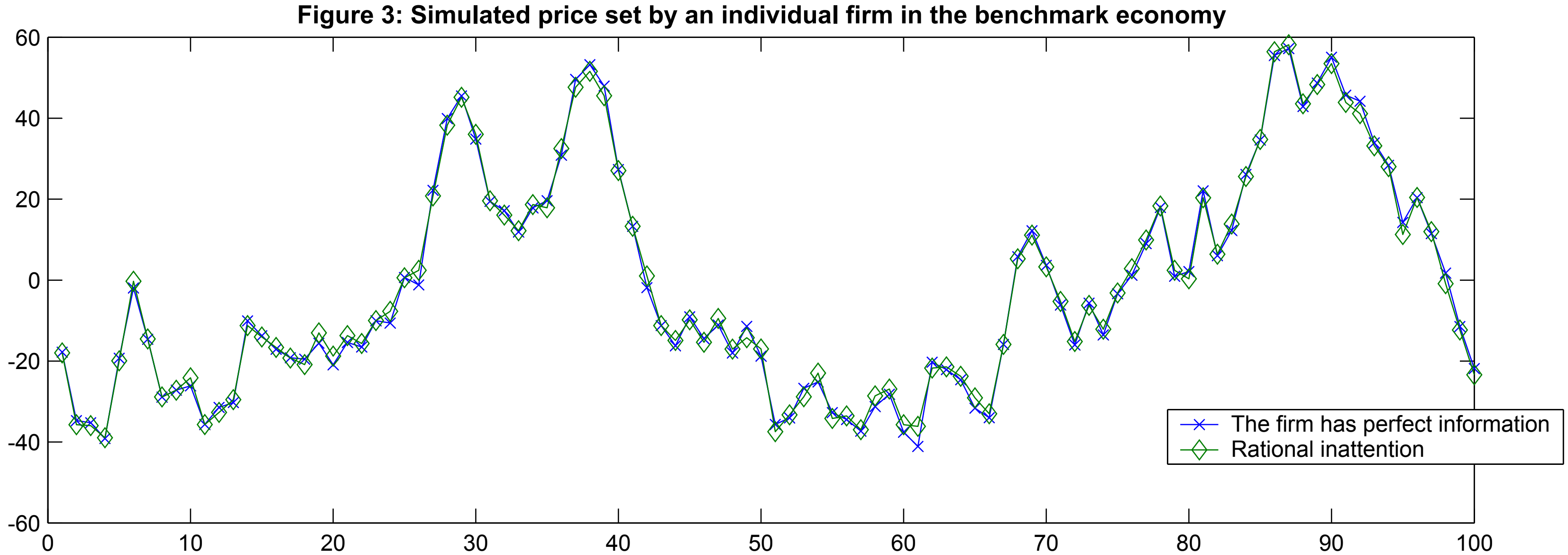


Figure 3: Simulated price set by an individual firm in the benchmark economy
60
40
20
0
-20
-40
-60
0
10
20
30
40
50
60
70
80
90
100
The firm has perfect information
Rational inattention

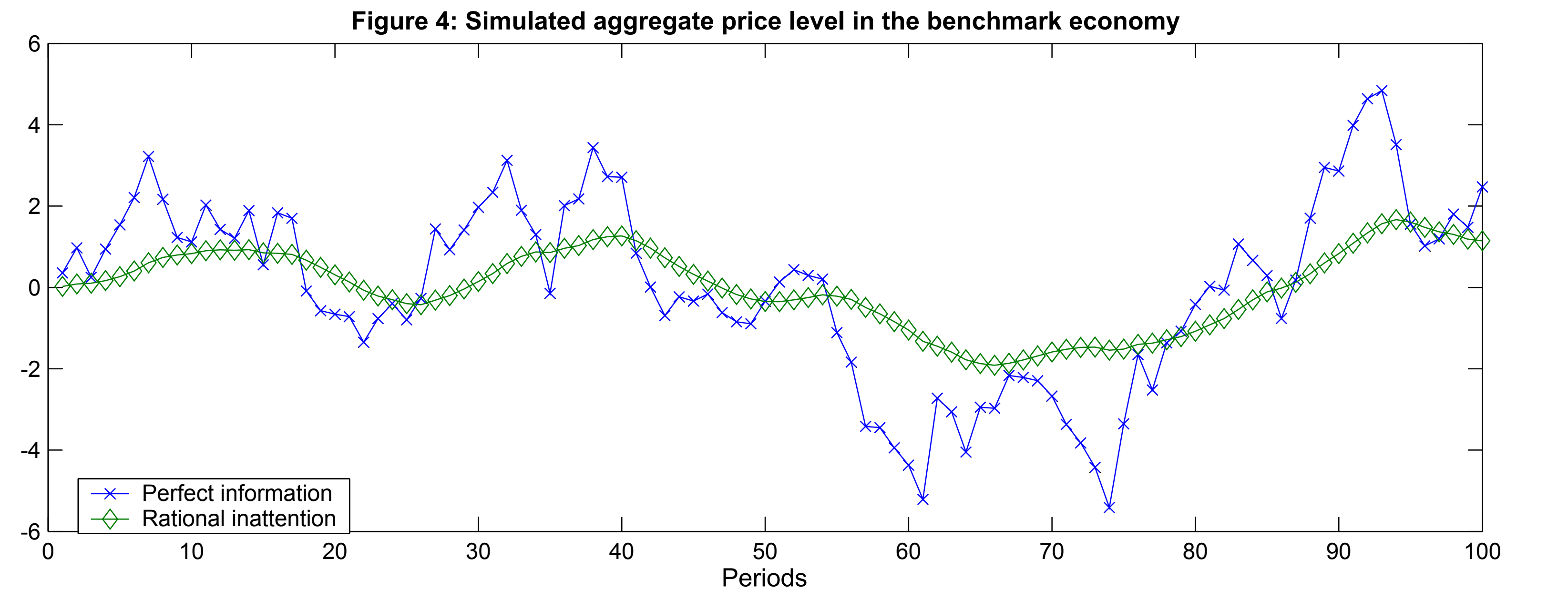


Figure 4: Simulated aggregate price level in the benchmark economy
6
4
2
0
-2
-4
-6
0
10
20
30
40
50
60
70
80
90
100
Periods
Perfect information
Rational inattention

Figure 5: Impulse responses of an individual price to an innovation in the idiosycratic state variable

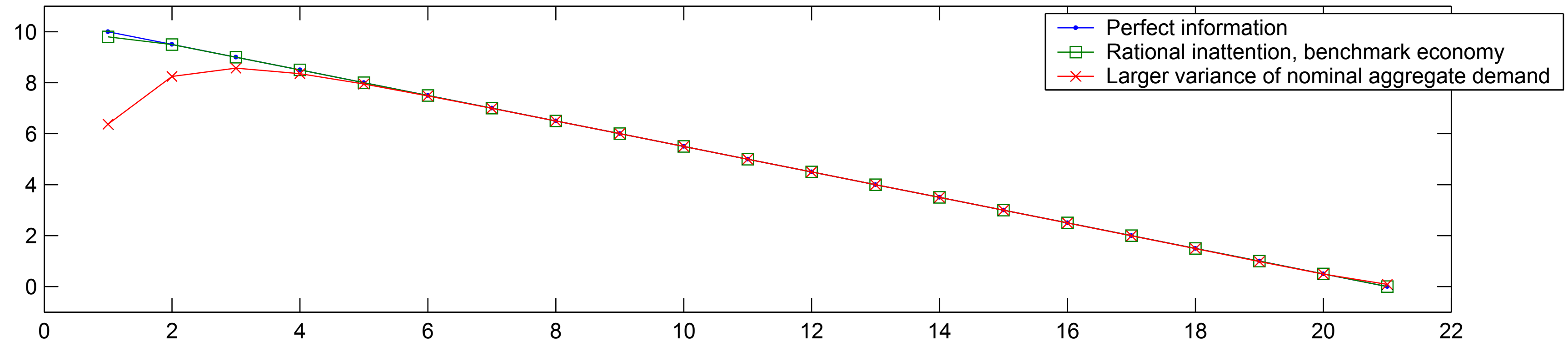


Figure 6: Impulse responses of the aggregate price level to an innovation in nominal aggregate demand

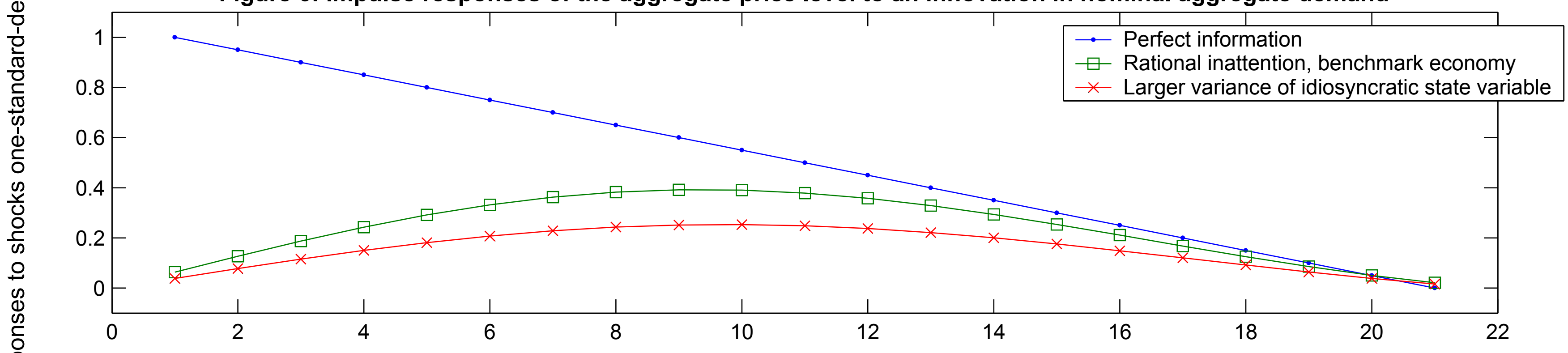


Figure 7: Impulse responses of the aggregate price level to an innovation in nominal aggregate demand

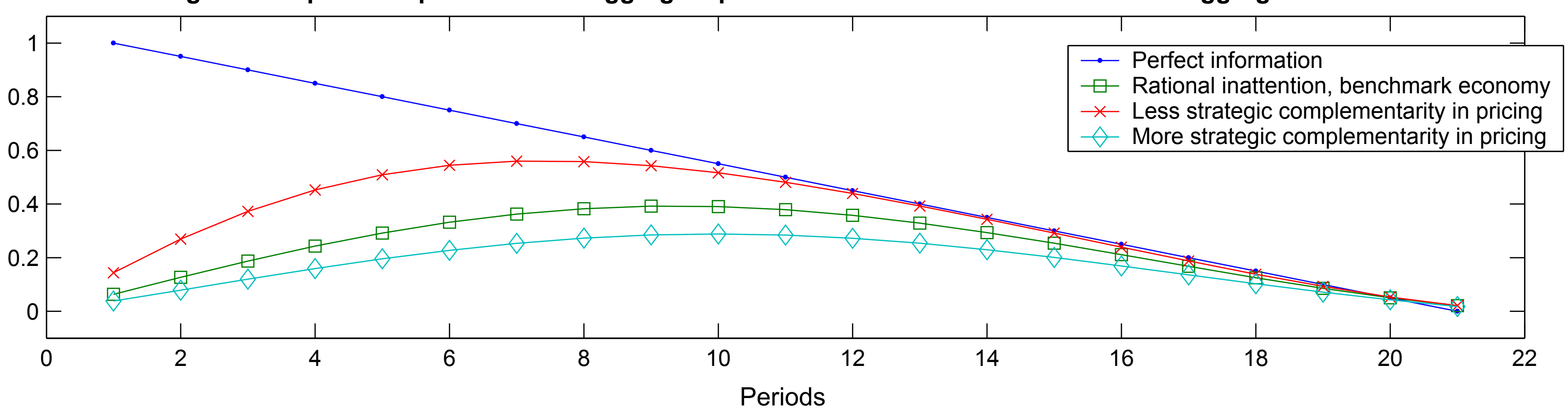


Figure 8: An optimal aggregate signal, benchmark economy

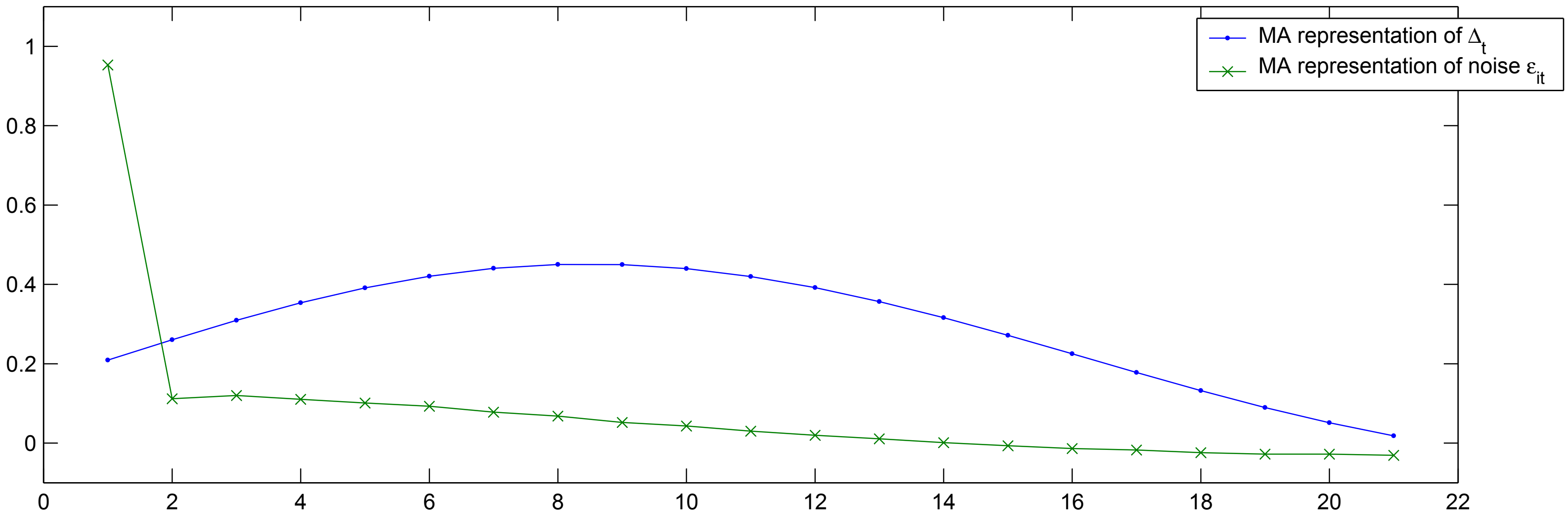


Figure 9: An optimal idiosyncratic signal, benchmark economy