Imperfect Common Knowledge and the Information Value of Prices*

Jeffery D. Amato
Bank for International Settlements
jeffery.amato@bis.org

Hyun Song Shin
London School of Economics
h.s.shin@lse.ac.uk

September 12, 2004

Abstract

When economic agents have diverse private information on the fundamentals of the economy, prices may serve as a poor aggregator of this private information. We examine the information value of prices in a monopolistic competition setting that has become standard in the New Keynesian macroeconomics literature. We show that public information has a disproportionate effect on agents' decisions, crowds out private information, and thereby has the potential to degrade the information value of prices. This effect is strongest in an economy with keen price competition. Monetary policy must rely on less informative signals of the underlying cost conditions.

*This paper supersedes the discussion in the first half of our longer paper that circulated under the title "Public and Private Information in Monetary Policy Models". We thank Andy Filardo, Marvin Goodfriend, Nobu Kiyotaki, John Moore, Stephen Morris and Lars Svensson for advice and comments at various stages of the project, and to Herakles Polemarchakis, Roko Aliprantis and an anonymous referee for their helpful comments and guidance. The views are those of the authors and do not necessarily represent those of the BIS. The second author acknowledges support from the U.K. ESRC under grant RES 000220450.

1 Introduction

One of the often-cited virtues of a decentralized economy is the ability of the market mechanism to aggregate the private information of the individual economic agents. Each economic agent has a window on the world. This window is a partial vantage point for the underlying state of the economy. But when all the individual perspectives are brought together, one can gain a much fuller picture of the economy. If the pooling of information is effective, and economic agents have precise information concerning their respective sectors or geographical regions, the picture that emerges for the whole economy would be a very detailed one. When can policy makers rely on the effective pooling of information from individual decisions?

This question is a very pertinent one for the conduct of monetary policy. Central banks that attempt to regulate aggregate demand by adjusting interest rates rely on timely and accurate generation of information on any potential imbalances in the economy. The role of the central bank in this context is of a vigilant observer of events to detect any nascent signs of growing imbalances. In particular, most central banks focus on the development of inflationary pressures. Signs of such pressures can be met by prompt central bank action through the use of monetary policy instruments.

It is possible to make a case that the information value of prices has improved in recent years. In his Jackson Hole paper, Rogoff (2003) notes that in very competitive sectors, like agriculture or semi-conductors, prices are significantly more flexible than in less competitive or highly regulated sectors. To the extent that globalisation and deregulation has increased competition across a wide range of industries, one could argue that firms are operating in a more competitive environment, and hence their prices will behave more like the prices in competitive sectors.

It is not our purpose here to assess the empirical claim that industries have, in fact, become more competitive.[1] Rather, we will assess the conditional statement which claims that if imperfectly competitive economies become more competitive, prices will become more responsive to changes in the fundamentals. Although it is undeniable that very competitive sectors such as agriculture and semi-conductors have very responsive prices, the question is whether the relationship between competition and price responsiveness is a continuous one.

There is some reason for doubting that the relationship between competition and price responsiveness is always continuous. Ball and Romer (1990) show that when firms worry about their market share, they will be reluctant to be first to raise prices (prices exhibit "real rigidities") so that even a small menu cost may induce firms not to change their prices. As the degree of competition increases, the real rigidities will become more severe, and this effect may become more potent.

We will show in this paper that even in the absence of menu costs and other nominal rigidities, greater competition may actually reduce the responsiveness of prices to changes in the underlying fundamentals. Our case is built on the importance of distributed information in decentralised economies where firms have access to information that is local to their region or industry, as well as to publicly available information that is available to firms economy-wide. Firms have their own "window on the world". In such a setting, when firms try to defend their market share, this will entail some degree of second-guessing the pricing strategies of their competitors. Even

[1] The evidence is somewhat mixed. Andersen and Wascher (2001) show that forecasts for inflation were consistently too high relative to actual outcomes in the late 1990s, whereas forecasts of costs did not display unusual behaviour, and suggest a decline in the markup of price. However, direct measurements of markup have not revealed such declines (see Bowman (2003), OECD (2002)).

when there are no nominal rigidities, the outcome of navigating through the higher-order beliefs entailed by the second-guessing of others leads firms to set prices that are far less sensitive to firms' best estimates of the underlying marginal costs.

Our conclusion is that the relationship between competition and price responsiveness is highly discontinuous. As an imperfectly competitive economy becomes more competitive, the price responsiveness falls. In the limit, as markup falls to zero, prices become completely unresponsive to the fundamentals. Thus, it is critical to distinguish between a perfectly competitive economy and the competitive limit of an imperfectly competitive economy.

The notion that equilibrium outcomes, particularly prices, are affected by imperfect common knowledge is not new. The macroeconomics literature on the forecasting the forecasts of others begun by Townsend (1983), Phelps (1983) and Sargent (1991) has examined the quantitative impact of "symmetrically uninformed" agents. The issue has recently been revisited by Woodford (2003a) and others[2] in the context of an imperfectly competitive economy. The conclusions drawn from this literature to date have largely relied on numerical simulations of fully-fledged macroeconomic models modified to incorporate private information. The value of such exercises lies in their ability to inform debates on the numerical time series properties of macroeconomic aggregates. For instance, Woodford (2003a) has shown how the combination of strategic behaviour and private information can induce greater persistence in macroeconomic variables in response to shocks relative to a common information benchmark.

[2] See Hellwig (2002), Ui (2003) and Adam (2003). See also Kasa (2000) and Pearlman and Sargent (2002). The latter shows how the problem can sometimes be reduced to the one with common knowledge. Similar issues arise in the context of asset pricing. See Allen et al. (2002), and Bacchetta and Van Wincoop (2002, 2004).

However, the cost of complexity is that it is difficult to isolate the key forces at work. One of our tasks in this paper is to attempt to fill in this gap by presenting a theoretical framework that is simple enough to unpack the precise mechanism at work in the degradation of the information value of prices. For this purpose, we will concentrate on two examples - a simple Gaussian dynamic model, and a static model where we abstract from any intertemporal learning or allocation problems. The virtue of such simple examples is that we can employ arguments that rely on well-known results from elementary probability theory.

We show that when price competition between firms becomes more intense, the aggregate price level becomes extremely unresponsive to the underlying fundamentals. In a dynamic context, prices exhibit a great deal of inertia. Even though firms are rational and form prices based on forward-looking expectations, actual behaviour of prices have the outward signs of adaptive expectations. One of the enduring puzzles that macroeconomists have struggled with is how to explain the the apparent inertia in inflation without resorting to adaptive expectations (see Gali and Gertler (1999)). Our examples suggest that models of distributed information may be a promising line to pursue in tackling this problem.

There are potentially troubling implications of our results for monetary policy. The experience of monetary policy in the 1990s has posed challenges for the view that the central bank can rely on the rate of inflation to guide monetary policy. Even as overall inflation pressures eased during the latter half of the 1990s, economies expanded rapidly under conditions of strong demand, accompanied by surging asset prices. The subsequent downturn in economic activity in the major industrial economies, and especially the United States, has fuelled debates about the information value of goods price inflation as an indicator of overall economic imbalances.[3]

We begin in the next section with a simple Gaussian dynamic price setting model, and then follow up with a general finite state setting in a static context that does not entail any distributional restrictions. We then present a small numerical example. We conclude by discussing the implications of our results for the conduct of monetary policy.

2 Price Inertia

We will be concerned with the pricing rule for firms of the form:

$$q_i = E_i q + \xi E_i \chi \tag{1}$$

where q_i is the (log) price set by firm i, q is the average price across firms, χ is marginal cost (in real terms) — our "fundamental variable" — and ξ is a constant between 0 and 1. The operator E_i denotes the conditional expectation with respect to firm i's information set. Pricing rules of this form have been discussed for some time. Phelps (1983) derived a similar pricing rule in a competitive economy, and compared it to the 'beauty contest' game discussed in Keynes's General Theory (1936), in which the optimal action involves second-guessing the choices of other players. Townsend (1978, 1983) also discussed similar pricing rules.[4]. However, our discussion in this paper has most in common with Woodford (2003a), who has revived interest in pricing rules of this form by showing how they arise naturally in macro models with differentiated goods and imperfectly competitive markets. The parameter ξ is related to the elasticity of substitution between goods, and

[3] See Borio, English and Filardo (2002) for further discussion of the challenges raised for monetary policy frameworks by these experiences.

[4] See Morris and Shin (2002) for a welfare analysis of coordination games that give rise to the beauty contest.

becomes small as the economy becomes more competitive. Appendix A presents an illustrative derivation in a partial equilibrium setting.

Rewrite (1) in terms of nominal marginal cost, defined as $z \equiv \chi + q$, yielding $q_i = (1-\xi) E_i q + \xi E_i z$. Taking the average across firms,

$$q = (1-\xi)\bar{E}q + \xi \bar{E}z \tag{2}$$

where $\bar{E}(\cdot)$ is the "average expectations operator", defined as $\bar{E}(\cdot) \equiv \int E_i(\cdot)di$. Hence,

$$q = \sum_{k=1}^{\infty} \xi (1-\xi)^{k-1} \bar{E}^k z \tag{3}$$

where $\bar{E}^k$ is the k-fold iterated average expectations operator. With differential information, the k-fold iterated average expectations do not collapse to the single average expectation[5].

Let us now embed the price setting decision in a dynamic context. Time is indexed by $t \in \{1, 2, \cdots\}$, and suppose that z follows an AR(1) Gaussian process $\{z_t\}$ where

$$z_t = a + \phi z_{t-1} + \eta_t$$

η_t is Gaussian noise, and $0 < \phi < 1$. The unconditional expectation of z_t is $\mu = a/(1-\phi)$. There is a continuum of firms, and none of them ever observe the true value of the fundamentals z_t. Instead, at date t, firm i observes the realization of the signal

$$x_{it} = z_t + \varepsilon_{it}$$

where ε_{it} is normal with mean zero and variance σ_ε^2. The noise terms $\{\varepsilon_{it}\}$ are independent across i and across t. The information set of firm i at date t is

$$\{x_{i1}, x_{i2}, \cdots, x_{it}\}$$

[5] See Morris and Shin (2002)

and must form beliefs on current, future and past realizations of z_t based on these signals alone. However, the firms know that z_t has mean μ, and so can utilize this information also. We may regard the unconditional mean μ as being the sole piece of public information. All other information is private to that firm alone.

Let $E_{it}(.)$ denote the expectation conditional on firm i's information set at date t. From the formula for conditional expectations of jointly normal random variables, we have

$$E_{it}\begin{bmatrix} z_1 \\ \vdots \\ z_t \end{bmatrix} = \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix} + V_{zx}V_{xx}^{-1}\begin{bmatrix} x_{i1}-\mu \\ \vdots \\ x_{it}-\mu \end{bmatrix} \tag{4}$$

where V_{zx} is the matrix of covariances where the (s,u)th entry is the covariance between z_s and x_{iu}, and V_{xx} is the covariance matrix for $(x_{i1}, x_{i2}, \cdots, x_{it})$. We can write (4) as

$$E_{it}\begin{bmatrix} \mu \\ z_1 \\ \vdots \\ z_t \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_{1t} & & & \\ \vdots & & V_{zx}V_{xx}^{-1} & \\ c_{tt} & & & \end{array}\right]\begin{bmatrix} \mu \\ x_1 \\ \vdots \\ x_t \end{bmatrix} \tag{5}$$

where

$$\begin{bmatrix} c_{1t} \\ \vdots \\ c_{tt} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} - V_{zx}V_{xx}^{-1}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{6}$$

are the weights placed on the unconditional mean μ in forming conditional beliefs on $\{z_s\}$. Taking the average of (5) across the continuum of firms, we have

$$\bar{E}_t\begin{bmatrix} \mu \\ z_1 \\ \vdots \\ z_t \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_{1t} & & & \\ \vdots & & V_{zx}V_{xx}^{-1} & \\ c_{tt} & & & \end{array}\right]\begin{bmatrix} \mu \\ z_1 \\ \vdots \\ z_t \end{bmatrix} \tag{7}$$

Let us denote:

$$B_t \equiv \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline c_{1t} & & & \\ \vdots & & V_{zx}V_{xx}^{-1} & \\ c_{tt} & & & \end{array}\right]$$

The matrix B_t corresponds to the average belief operator $\bar{E}_t$, and satisfies the following property.

Lemma 1 *B_t is the transition matrix of a Markov chain over $\{\mu, z_1, \cdots, z_t\}$ where μ is the sole absorbing state. All other states are transient.*

To prove this lemma, note from (6) that the rows of B_t must sum to one. We need to show that $V_{zx}V_{xx}^{-1}$ has non-negative entries but its largest eigenvalue is strictly less than one. To begin, note that since the noise terms $\{\varepsilon_{it}\}$ are independent across t, we have $\text{Cov}(x_{it}, z_s) = \text{Cov}(z_t, z_s)$. Thus, V_{zx} is equal to V_{zz}, the covariance matrix for $\{z_t\}$. Since V_{zz} is symmetric and positive definite, it can be diagonalized as

$$V_{zz} = E\Lambda E'$$

where Λ is the diagonal matrix of strictly positive eigenvalues $\lambda_1, \cdots, \lambda_t$. Meanwhile, note that

$$\begin{aligned} V_{xx} &= V_{zz} + \sigma_\varepsilon^2 I \\ &= E\Lambda E' + \sigma_\varepsilon^2 I \\ &= E\left(\Lambda + \sigma_\varepsilon^2 I\right) E' \end{aligned}$$

so that V_{xx} can also be diagonalized in terms of the set of eigenvectors E. In other words, V_{zz} and V_{xx} are diagonalizable within the same linear basis. Then,

$$V_{zz}V_{xx}^{-1} = E \begin{bmatrix} \frac{\lambda_1}{\lambda_1+\sigma_\varepsilon^2} & & \\ & \ddots & \\ & & \frac{\lambda_t}{\lambda_t+\sigma_\varepsilon^2} \end{bmatrix} E'$$

Hence, all the eigenvalues of $V_{zz}V_{xx}^{-1}$ are positive, but strictly less than one. This proves the lemma.

Higher-order average beliefs at date t are then determined by the iterated application of the average belief matrix B_t. The inertia of higher order beliefs is reflected in the weight given to realizations of z_s in the distant past and to the ex ante mean μ. We have

$$B_t^k = \begin{bmatrix} 1 & 0 \\ \left(\sum_{i=0}^{k-1}\left(V_{zx}V_{xx}^{-1}\right)^i\right)c & \left(V_{zx}V_{xx}^{-1}\right)^k \end{bmatrix}$$

Since the eigenvalues of $V_{zx}V_{xx}^{-1}$ are strictly less than one, $\left(V_{zx}V_{xx}^{-1}\right)^k \to 0$ as k becomes large. The unconditional mean μ is the unique absorbing state of the Markov chain. As the order of iterated beliefs increases, all the weight becomes concentrated on μ and away from the subsequent realizations of $\{z_s\}$. Thus, we have proved:

Theorem 2 *As* $k \to \infty$, $\bar{E}_t^k \begin{bmatrix} \mu \\ z_1 \\ \vdots \\ z_t \end{bmatrix} \to \begin{bmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix}$.

The unconditional mean μ in our example is the sole piece of public information among the firms - it is the sole piece of information that is common knowledge among the firms. In general, any signal that is publicly observed by the firms, and hence common knowledge among them, will be an absorbing state in the Markov chain representation.

The implications of theorem 2 are profound for an economy with firms that face fierce competition. From equation (3), we know that the average price level is a weighted average of higher order average expectations of nominal marginal cost. As the parameter ξ becomes small, it is the higher order average expectations that receive more and more weight. However, we know from theorem 2 that higher order average expectations contain very little information value in that the recent realizations of the fundamentals $\{z_t\}$ have very little impact on price. In the limit as $\xi \to 0$, we approach the competitive limit of the imperfectly competitive economy. From (3), whatever is the history of the economy and the current value of z_t is, we have

$$q_t \to \mu \text{ as } \xi \to 0$$

In this economy, the price level is completely unresponsive to changes in the fundamentals. The price level is held fixed at μ, the unconditional mean of nominal marginal cost. Thus, in spite of the lack of any nominal rigidities, and in spite of rational pricing setting behaviour of the firms, there is a great deal of inertia.

3 General Static Economy

We now examine a more general example with a finite number of states, but a one shot pricing problem for firms. We maintain the pricing rule

$$q_i = (1 - \xi) E_i q + \xi E_i z$$

If nominal marginal cost were common knowledge among the firms, the unique equilibrium price would then be given by $q_i = z$, for all firms i. We call this outcome the *perfect information benchmark*.

However, suppose that firms have imperfect information. In particular, we will suppose that firm i observes a noisy signal z_i of z given by

$$z_i = z + \varepsilon_i$$

where the noise term ε_i is i.i.d. across firms. Faced with this uncertainty, each firm must estimate the price charged by the other firms from his own estimate of z. We will first examine a special case in which firms have identical information concerning the underlying fundamentals.

3.1 Pooled Information Benchmark

By the *pooled information benchmark*, we refer to the hypothetical situation in which firms pool their information, so that all firms have access to all the information available to the set of firms taken as a whole. Among other things, firms end up having identical information in this case.

When there are a large number of firms each with conditionally independent signals, this case can be considered an approximation to the perfect information case. If we denote by $\mathcal{I}_i$ the information set of firm i, then the pooled information benchmark endows each firm with the information set given by the union

$$\bigcup_i \mathcal{I}_i$$

The information set of firm i includes all those random variables that firm i observes - such as its own characteristics. These signals then define a partition over the overall state space in the standard way. We will make use of both formalisms below.

Let us use the shorthand $E_*(\cdot)$ for the conditional expectation with respect to the pooled information set $\bigcup_i \mathcal{I}_i$. When firms have access to pooled information, the equilibrium pricing rule gives

$$q = E_*(z) \tag{8}$$

Thus, average price is given by the mark-up times the expected average nominal marginal costs across firms. An increase in marginal cost is reflected in a one-for-one increase in average prices. Thus, in the pooled information benchmark, an outside observer (such as the central bank) can make good inferences on the underlying cost conditions. We will now contrast this with the general case in which firms have differential information.

3.2 Equilibrium

To avoid unnecessary technical details, we will assume that the firms can be partitioned into a finite number N of equally-sized subclasses, where firms in each subclass are identical, and commonly known to be so. We will also assume that the random variables z and $\{\varepsilon_i\}$ take on finitely many possible values. We define a *state* ω to be an ordered tuple:

$$\omega \equiv (z, \varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N)$$

that specifies the outcomes of all random variables of relevance. We will denote by Ω the *state space* that consists of all possible states. The state space is finite given our assumptions.

There is a known *prior density* ϕ over the state space Ω that is implied by the joint density over z and the idiosyncratic terms ε_i. The prior is shared by all firms, and represents the commonly shared assessment of the likelihood of various outcomes before the realization of marginal cost z. However, once the firm observes its own signal z_i, it makes inferences on the economy from its knowledge of the joint density of the random variables by forming the conditional density over the state space based on the realization of its own signal z_i. Equivalently, firm i's information partition over Ω is generated by the equivalence relation $\sim_i$ over Ω, where $\omega \sim_i \omega'$ if and only if the realization of z_i is the same at ω and ω'.

Some matrix notation is useful at this point. Index the state space Ω by the set $\{1, 2, \cdots, |\Omega|\}$. We will use the convention of denoting a random variable $f : \Omega \to \mathbb{R}$ as a *column vector* of length $|\Omega|$, while denoting any probability density over Ω as a *row vector* of the same dimension. Thus, from here on, the prior density ϕ will be understood to be a row vector of length $|\Omega|$. We will denote by $b_i(k)$ the row vector that gives the posterior density for firm i at the state indexed by k. By gathering together the conditional densities across all states for a particular firm i, we can construct the matrix of posterior probabilities for that firm. Define the matrix B_i as the matrix whose kth row is given by the posterior density at the state indexed by k. That is

$$B_i \equiv \begin{bmatrix} - & b_i(1) & - \\ - & b_i(2) & - \\ & \vdots & \\ - & b_i(|\Omega|) & - \end{bmatrix}$$

We note one important general property of this matrix. We know that the average of the rows of B_i weighted by the prior probability of each state must be equal to the prior density itself. This is just the consequence of the consistency between the prior density and the posterior densities. In our matrix notation, this means that

$$\phi = \phi B_i \tag{9}$$

for all firms i, so that ϕ is a fixed point of the mapping defined by B_i. More specifically, note that B_i is a stochastic matrix in the sense that it is a matrix of non-negative entries where each row sums to one. Hence, it is associated with a Markov chain defined on the state space Ω. Then equation (9) implies that the prior density ϕ is an *invariant distribution* over the states for this Markov chain. We will make much use of this property in what follows. This formalization of differential information environments in terms of Markov chains follows Shin and Williamson (1996) and Samet (1998).

For any random variable $f : \Omega \to \mathbb{R}$, denote by $E_i f$ the conditional expectation of f with respect to i's information. $E_i f$ is itself a random variable, and so we can denote it as a column vector whose kth component is the conditional expectation of firm i at the state indexed by k. In terms of our matrix notation, we can write:

$$E_i f = B_i f$$

As well as the conditional expectation of any particular firm, we will also be interested in the average expectation across all firms. Define $\bar{E} f$ as

$$\bar{E} f = \frac{1}{N} \sum_{i=1}^{N} E_i f$$

$\bar{E} f$ is the random variable whose value at state ω gives the average expectation of f at that state. The matrix that corresponds to the average expectations operator $\bar{E}$ is simply the average of the conditional belief matrices $\{B_i\}$, namely

$$B \equiv \frac{1}{N} \sum_{i=1}^{N} B_i$$

Then, for any random variable f, the average expectation random variable $\bar{E} f$ is given by the product Bf. Since Bf is itself a random variable, we can define

$$B^2 f \equiv BBf$$

as the average expectation of the average expectation of f. Iterating further, we can define $B^k f$ as the kth order iterated average expectation of f. Then, the equilibrium pricing rule (1) can be expressed in matrix form as

$$q_i = \xi B_i z + (1 - \xi) B_i q$$

Here, q_i is a column vector whose jth entry is the price of firm i at the jth state, and z is the column vector whose jth component is the (true) marginal cost at the jth state. Taking the average across firms, we have

$$q = \xi Bz + (1 - \xi) Bq \tag{10}$$

By successive substitution, and from the fact that $0 < \xi < 1$, we have

$$\begin{aligned} q &= \xi \sum_{i=0}^{\infty} ((1 - \xi) B)^k Bz \\ &= \xi (I - (1 - \xi) B)^{-1} Bz \\ &= MBz \end{aligned} \tag{11}$$

where M is the matrix

$$M = \xi (I - (1 - \xi) B)^{-1}$$

Thus, equilibrium average price q is given by (11). Let us note some preliminary observations on the comparison between (11) and the perfect information benchmark, in which the cost function is common knowledge across firms. With perfect information, we have

$$q = z \tag{12}$$

There are two differences between (11) and (12). First, there is the effect of *fundamental uncertainty*. Since the firms do not know the true marginal cost, Bz is different from z. However, if the noise ε_i is very small, or when N is large (there are many classes of firms), $Bz \simeq z$.

However, there is a second, more important difference given by the appearance of the matrix M in the equilibrium of the differential information economy. The matrix M is a stochastic matrix (i.e. a matrix of non-negative entries whose rows sum to one) since each row of the matrix $((1-\xi)B)^k$ sums to $(1-\xi)^k$ so that the matrix $(I-(1-\xi)B)^{-1} = \sum_{i=0}^{\infty}((1-\xi)B)^k$ has rows which sum to $1+(1-\xi)+(1-\xi)^2+\cdots = 1/\xi$. Thus, the matrix $M = \xi(I-(1-\xi)B)^{-1}$ is a stochastic matrix.

The matrix M serves the role of "adding noise" (in the sense of Blackwell(1951)) to the random variable Bz. The effect of the noise is to smooth out the variability of prices across states. Thus, in going from (12) to (11) the average prices become less reliable signals of the underlying cost conditions of the economy. Since the noise matrix M is a convex combination of the higher order beliefs $\{B^k\}$, we must first understand what determines these higher order beliefs. In general, higher order expectations contain much less information than lower order expectations in the following precise sense. For any random variable f, denote by $\max f$ the highest realization of f, and define $\min f$ analogously as the smallest realization of f. Then for any stochastic matrices C and D and any random variable f,

$$\begin{aligned} \max CDf &\leq \max Df \\ \min CDf &\geq \min Df \end{aligned}$$

CD is a "smoother" or "noisier" version of D in the sense of Blackwell. So, the higher is the order of the iterated expectation, the more rounded are the peaks and troughs of the iterated expectation across states.

The importance of the parameter ξ is now apparent. The smaller is this parameter, the greater is the weighting received by the higher order beliefs in the noise matrix M, so that the prices are much less informative about the underlying cost conditions. In particular, the limiting case of B^k as $k \to \infty$ is an important benchmark, since this is the limit of the solution for equilibrium average price q as $\xi \to 0$. We turn to this question now.

3.3 Public Information Limit

The limiting case for higher order beliefs B^k as k becomes large has a special property. From (9), we know that

$$\phi = \phi B \tag{13}$$

so that the prior density ϕ is an invariant distribution for the Markov chain defined by the average belief matrix B. By post-mulitiplying both sides by B, we have

$$\phi = \phi B = \phi B^2$$

so that ϕ is an invariant density for B^2 also. By extension, we can see that ϕ is an invariant density for B^k for any kth order average belief operator. We also know from the elementary theory of Markov chains that under certain regularity conditions (which we will discuss below), the sequence $\{B^k\}_{k=1}^{\infty}$ converges to a matrix B^∞ whose rows are identical, and given by the unique stationary distribution over Ω. Since we know that the prior density ϕ is an invariant distribution, we can conclude that under the regularity conditions, all the rows of B^∞ are given by ϕ. That is

$$B^\infty = \begin{bmatrix} - & \phi & - \\ - & \phi & - \\ & \vdots & \\ - & \phi & - \end{bmatrix} \tag{14}$$

In other words, the limiting case of higher order beliefs B^k as k becomes large is so noisy that all information is lost, and the average beliefs converge to the prior density ϕ at every state. In particular, for any random variable f, successively higher order beliefs are so noisy that all all peaks and troughs into a constant function, where the constant is given by the prior expectation

$\bar{f}$ (i.e. the expectation of f with respect to the prior density ϕ). In other words,

$$B^k f \to \begin{bmatrix} \bar{f} \\ \bar{f} \\ \vdots \\ \bar{f} \end{bmatrix} \quad \text{as } k \to \infty \tag{15}$$

To introduce the regularity conditions that ensure this, and to delve further into the underlying structure of our results, let us denote the (j,k)th entry of B by

$$b(j,k)$$

This is the probability of one-step transition from state j to state k in this Markov chain. The condition that guarantees (14) is the following.

Condition 3 *For any two states j and k, there is a positive probability of making a transition from j to k in finite time.*

Condition 3 ensures that the matrix B corresponds to a Markov chain that is *irreducible*, *persistent* and *aperiodic*. It is irreducible since all states are accessible from all other states. For finite chains, this also means that all states are visited infinitely often, and hence persistent. Finally, the aperiodicity is trivial, since all diagonal entries of B are non-zero irrespective of condition 3. We can then prove lemma 4.

Lemma 4 *Suppose B satisfies condition 3. Then, the prior density ϕ is the unique stationary distribution, and $B^k \to B^\infty$, where B^∞ is the matrix whose rows are all identical and given by ϕ.*

Condition 3 has an interpretation in terms of the degree of information shared between the firms. It corresponds to the condition that

$$\bigcap_i \mathcal{I}_i = \emptyset \tag{16}$$

In other words, the intersection of the information sets across all firms is empty. There is no signal that figures in the information set of all the firms. This condition is satisfied in our case since the signals of firms' costs are highly correlated, but not exactly identical. Thus, (16) holds, so that condition 3 is satisfied. Another way to phrase this is to say that there is no non-trivial event that is common knowledge among the firms. The only event that is common knowledge is the trivial event Ω, which is the whole space itself.

If the intersection $\bigcap_i \mathcal{I}_i$ were non-empty, then some signals are observed by every firm, and the outcomes of these signals become *public* and hence common knowledge. The equilibrium pricing decision of firms can be analysed for this more general case in which firms have access to public information, as well as their private information.

In this more general case, the limiting results for the higher order average belief matrices B^k correspond to the beliefs conditional on *public signals*. In order to introduce these ideas, let us recall the notion of an information partition for a firm. Let firm i's information partition be defined by the equivalence relation $\sim_i$ where $\omega \sim_i \omega'$ if firm i cannot distinguish between states ω and ω'. Denote firm i's information partition by $\mathcal{P}_i$, and consider set of all information partitions $\{\mathcal{P}_i\}$ across firms. The *meet* of $\{\mathcal{P}_i\}$ is defined as the finest partition that is at least as coarse as all of the paritions in $\{\mathcal{P}_i\}$. The meet of $\{\mathcal{P}_i\}$ is thus the greatest lower bound of all the individual partitions in the lattice over partitions ordered by the relation "is finer than". The meet of $\{\mathcal{P}_i\}$ is denoted by

$$\bigwedge_i \mathcal{P}_i$$

The meet is the information partition that is generated by the public signals - i.e. those signals that are in the information set of every firm, and hence in the intersection

$$\bigcap_i \mathcal{I}_i$$

The meet has the following property[6].

Lemma 5 *If two states ω and ω' belong to the same element of the meet $\bigwedge_i \mathcal{P}_i$, then there is positive probability of making a transition from ω to ω' in finite time in the Markov chain associated with B.*

Lemma 5 gives a generalization of condition 3. The idea is that the Markov chain defined by the average belief matrix B can be expressed in block diagonal form:

$$B = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_J \end{bmatrix}$$

and where each sub-matrix A_j defines an irreducible Markov chain that corresponds to an element of the meet $\bigwedge_i \mathcal{P}_i$. Then, the higher-order belief limit is given by

$$B^\infty = \begin{bmatrix} A_1^\infty & & & \\ & A_2^\infty & & \\ & & \ddots & \\ & & & A_J^\infty \end{bmatrix}$$

Furthermore, we have

$$\phi = \phi B^\infty = \phi \begin{bmatrix} A_1^\infty & & & \\ & A_2^\infty & & \\ & & \ddots & \\ & & & A_J^\infty \end{bmatrix}$$

[6] See, for instance, Shin and Williamson (1996).

and so for any random variable f, the higher order expectation of f at each state has the following limiting property, in which the limit of the higher order expectation is the conditional expectation based on the public signals only. In other words, we have:

Theorem 6 *As $k \to \infty$,*

$$B^k f \to \begin{bmatrix} E\left(f \mid \cap_i \mathcal{I}_i\right)(\omega_1) \\ E\left(f \mid \cap_i \mathcal{I}_i\right)(\omega_2) \\ \vdots \\ E\left(f \mid \cap_i \mathcal{I}_i\right)(\omega_N) \end{bmatrix}$$

where $E\left(f \mid \cap_i \mathcal{I}_i\right)(\omega)$ is the conditional expectation of f at state ω based on public information only.

In appendix B, we provide an alternative proof of this result that uses the eigenvalues of the average belief matrix that bring out some additional features of the problem. Theorem 6 implies that for small values of ξ, the dominant influence in determining the average price level p is given by the set of *public signals.* For example, suppose the central bank announces a forecast for the price level, and this is a sufficient statistic for any public signals available to firms. Then the equilibrium average price p will largely reflect the central bank's forecast regardless of the underlying cost conditions in the economy.

3.4 Example

We illusrate our results through a simple example. Consider an economy with three firms. Suppose that marginal cost z can assume one of two values, z_H or z_L, with probabilities p_H and $1 - p_H$ respectively, where $z_L = z_H - \epsilon$. Each firm $i \in \{1, 2, 3\}$ observes a private signal of z, denoted z_i, given by

$$z_i = z + \epsilon_i \tag{17}$$

where the idiosyncratic noise term ϵ_i takes realizations ϵ, 0 or $-\epsilon$ with probabilities p_ϵ, $1 - 2p_\epsilon$ and p_ϵ, respectively. The noise terms are independent across firms.

In this economy there is a total of 54 states - 2 possible realizations for cost and 3^3 configurations of signals for each cost realization. A state is identified with the realization of the tuple $(z, \epsilon_1, \epsilon_2, \epsilon_3)$. These are illustrated in Table 1, along with the objective (prior) probability of the occurrence of each state.

Given the assumptions we have made, firms can observe four possible values of their private signals: $z_H + \epsilon$, z_H, z_L or $z_L - \epsilon$. When either the first or last of these is observed, firms know the value of marginal cost with certainty (z_H and z_L, respectively). However, since the signals of other firms are not observed, firms cannot identify the state even when either of these two signals is observed. For example, if firm 1 receives the signal $z_H + \epsilon$, then it knows that the index of the realised state (see Table 1) is between 1 and 9, but it cannot further distinguish amongst these states. Thus, the first nine rows of the posterior probability matrix for firm 1, B_1, has the (renormalised) prior for states 1 to 9 in the first nine columns, and zeros in the remaining columns. The rest of the posterior for firm 1 can be constructed by undertaking a similar analysis, as can the posteriors for the other two firms.

To make our example concrete, let us choose values for the parameters in the model. For the marginal cost parameters, our choices are: $z_H = 5$, $z_L = 3$ and $p_H = 0.5$. Since $\epsilon = z_H - z_L$, we have $\epsilon = 2$. Of greater relevance, as will be made clear below, are the choices of ξ and p_ϵ. As a baseline value, a reasonable choice for ξ is 0.15 (see Woodford (2003b) for a discussion of the size of this parameter). The baseline value for p_ϵ is 1/3, which means the distribution of the noise term gives equal weight to the three possible values.

Turning now to the features of this economy we wish to highlight, Figure 1 shows the set of feasible tuples for marginal cost and the price level in pooled and differential information versions of the model. These are given by the circles in the respective panels. Recall that in an economy with perfect information (i.e. no uncertainty), $q = z$. In both panels, this equilibrium is denoted by "$\times$". Finally, the square marks the tuple of the means of the two variables, which is identical in all three types of economies.

In the pooled information benchmark, the true marginal cost is fully revealed provided that at least one firm receives an extreme signal. For instance, if $z = z_L$, then it takes just one firm to receive the signal $z_L - \epsilon$ for the true marginal cost to become common knowledge, and we have the full information outcome. Conditional on $z = z_L$, this happens with probability $1 - \left(\frac{2}{3}\right)^3 = \frac{19}{27}$, and this outcome is marked by the circle that is superimposed on the "$\times$". With probability $\left(\frac{2}{3}\right)^3$, none of the firms receive the extreme signal $z_L - \epsilon$, and given our assumption of a uniform prior, all firms give equal weight to $z = z_L$ and $z = z_H$. In this case, the firms learn nothing, and we have $q = \frac{z_L + z_H}{2} = 4$. This is represented by the circle to the left of the square at the value 4. Clearly, as the number of firms increases, the probability of the full information outcome increases to 1.

Contrast this with the price distribution under the equilibrium with differential information, given by the right hand panel of Figure 1. Conditional on the true marginal cost, there are four possible values of the average price, corresponding to how many firms have received an extreme signal that reveals the true cost level. The uniform density assumption implies that this is all that matters - if an extreme value is not observed, then the signal is uninformative of the underlying true cost. The equilibrium outcomes in the differential information economy are more tightly bunched together near the mean but are farther away from the true value of the underlying fundamental (marginal cost). In particular, note that even in the case where all three firms have observed an extreme outcome (so that all firms *know* the true cost), they end up choosing a price that is very far from the price in the perfect information case. The reason here is that even though each firm knows the true cost, this true cost is *not common knowledge*. Each firm allows the possibility that one or both of the other firms do not know the true cost, and "shade" their decisions towards the mean. This illustrates the smoothing property of higher-order beliefs: as the order of belief increases, the posterior places relatively more weight on the prior (which results in prices being closer to their mean). Hence, our conclusion follows that prices may not be a good indicator of underlying cost pressures when information is dispersed across firms with strategic interactions.

What is the effect of changing the weights on the orders of beliefs in the expectations of firms? Evidence of this is presented in Figure 2. The left panel of the figure repeats the differential information case from Figure 1. In the right panel of Figure 2 the weights on higher-order beliefs have been increased by changing the value of ξ from 0.15 to 0.05. Recall from equation (11) that the weights on higher orders are a decreasing function of ξ, i.e. for smaller values of ξ, the weight on beliefs of order k declines more slowly as k becomes large. Again, the importance of the smoothing property of higher-order beliefs is evident; prices in the economy with greater competition are distributed even more closely together near the mean.

It was argued above that many industrial economies may have witnessed the degree of competition having intensified in the 1990s. In the context of our model, this would have the effect of lowering ξ because one factor that determines ξ is the elasticity of substitution across goods, as measured by θ – more intense competition is embodied by a higher value for θ and hence a smaller value for ξ. If competition has indeed intensified, then the distorted picture presented by aggregate price outcomes may have gotten worse in recent years.

Lastly, one versatile feature of our model is that, apart from the requirement of a finite state space, our equilibrium solution has been obtained independently of any distributional assumptions on the common prior or signals. But, of course, specific choices for these distributions – by varying p_H or p_ϵ, or by specifying the distribution of ϵ_i to be asymmetric – will impact upon the actual realisations of the price level. The final feature of the model we illustrate brings to light the effect of different distributional assumptions.

We highlight the importance of the relative precision of the private signals on the dispersion of the price level. In general, we would expect that more precise private signals (relative to commonly held information, i.e. the prior) should result in more efficient outcomes; that is, prices should converge to marginal cost, the perfect information benchmark. In order to assess this, notice that the variance of the noise term in the private signal is given by $2\epsilon p_\epsilon$, and is therefore an increasing function of p_ϵ. We therefore reduce p_ϵ, from 1/3 to 1/12; the results are shown in Figure 3. When we reduce p_ϵ in this way, the uniform posterior density is lost, and we must allow a greater number of possible combinations of signals. Conditional on the true cost level, there are 10 possible equilibrium average prices[7].

Some interesting effects are evident by reducing p_ϵ to 1/12. We can see

[7] Either all firms observe the true cost, or only two of the firms observe the true cost (which can happen in two ways), or only one firm observes the true cost (which can happen in three ways), or none of the firms observe the true cost (which can happen in four ways). Thus, there are $1 + 2 + 3 + 4 = 10$ possible outcomes for average price.

that the range of feasible prices is now larger, both in the direction towards the true value of the state (marginal cost) and away from it. The price level is furthest away from the true cost level when all the firms have noise that takes them further from the true cost level. This happens with small probability, but when it does happen, the firms take their signals at face value, and the average price is far away from the true cost. Of course, in ex ante terms, these extreme outcomes would be very unlikely. Thus, the price level will sometimes be further away from marginal cost compared to when the signals are noisier, but more often they will be closer to the fundamental.

4 Implications for Monetary Policy

In expanding upon previous work, we have developed a model of pricing behaviour that has potentially strong implications for the information value of prices. In a nutshell, our main result is that goods prices, particularly the aggregate price, may provide a distorted picture of underlying activity in the economy. In this section we discuss some possible ramifications of our results for the role played by aggregate goods price measures in the conduct of monetary policy. It should be noted that while our analysis has focused on the behaviour of the price level, the policy implications drawn here are likely to be equally valid for measures of goods price inflation as well. In view of the greater emphasis in current policy practice on inflation stabilisation, rather than price level stabilisation, we will generally refer to inflation[8]. Our discussion will focus on two issues: price level or inflation targets as a goal of monetary policy and inflation as an indicator variable in the conduct of policy. We do not, however, seek to provide new empirical evidence to test

[8] A formal analysis of inflation would require us to extend our model to a truly dynamic setting. See Woodford (2003a), Amato and Shin (2003) or Hellwig (2002) for dynamic models with imperfect common knowledge.

the conjectures put forward here.

Inflation targeting has now become one of the main paradigms of monetary policy. Fry et al. (2000), for example, document the rise of inflation targeting in advanced industrial and emerging market economies alike. In many countries, the adoption of inflation targeting was a solution to problems with previous monetary regimes.[9] In part, the move towards a regime centered around an inflation target was a natural progression for policy makers who had been primarily focused on inflation stabilisation in any case. Inflation targeting has also generated considerable interest in academic circles (see, e.g., Svensson (1999) and Svensson and Woodford (2003), among many others). Ultimately, inflation targeting became credible because of the widespread view that high and variable inflation rates reduce welfare and that the best contribution of monetary policy to improving welfare is to control inflation[10]. Price level targeting is not currently practiced in any countries, at least not explicitly. But with the attainment of low and stable inflation, some authors have called for central banks to move one step further and aim at price targets.

Apart from recognition of the direct costs of high and variable inflation, the appeal of inflation targeting has been rooted in the view that inflation movements generally reflect underlying economic imbalances. The basis for this is some version of the Phillips curve, which is the standard model of inflation determination. In the Phillips curve, misalignments between demand and supply translate into fluctuations in inflation. Other shocks might also

[9] Despite progress during the 1980s in controlling inflation, several central banks were forced to abandon their monetary regime in the early 1990s. As an alternative, and to solidify the gains in bringing down inflation, many chose to adopt inflation targeting.

[10] Even under this perspective of monetary policy, there may be instances when deviating from primary attention on inflation would be a preferable strategy; see Mussa (2003) for a recent discussion of this point.

buffet inflation (e.g. inefficient supply shocks), making it more difficult to identify the true source of price changes; but it is generally the case that inflation is seen to be a sufficient statistic of imbalances in real activity.

In this regard, the desirability of either inflation or price level targeting might need to be reevaluated in the light of the issues addressed in our paper. We have argued that the price level may be a poor reflection of underlying economic activity. For instance, in Figure 1, price outcomes are much closer to the mean than to marginal cost in each state of the world. Where competitive pressures are even more intense, such as in Figure 2, the price level is distributed even more tightly around the mean. Using as a metric the squared or absolute distance between the price level and its mean (assumed to be the policy target), one would judge monetary policy to have been a success in the differential information economy. But this conclusion would be erroneous if the true objective of monetary policy is to keep prices at their efficient level in each state of the world, namely, equal to marginal cost.

Whether or not a central bank pursues price level or inflation stabilisation as a goal of monetary policy, changes in aggregate prices can serve a separate role as an indicator of imbalances between aggregate demand and supply. For instance, inflation can enter as a separate term in a policy rule (e.g. the Taylor rule). In various versions of New Keynesian dynamic general equilibrium models, interest rate rules that give a dominant role to inflation have been shown to be near optimal (see, e.g., Woodford (2003b) and the references therein). But the same arguments above apply equally well to the efficacy of inflation as an indicator variable. Just as the price level or inflation may be a poor metric of efficient allocations, it may also be a poor guide of the true state of underlying fundamentals. A central bank that moves its policy interest rate mainly in response to inflation fluctuations could end up displaying too little activism. The smoothing effect of higher-order beliefs on prices could induce a false sense of efficiency in goods markets.

A related argument about the potential problems with using consumer price inflation as an indicator is what has been called "the paradox of credibility" (see, e.g., Borio and Lowe (2002)). The idea is that as central banks have become more credible with the achievement of low and stable inflation, inflation itself may have become a less reliable signal of other imbalances. By policy being more credible, inflation expectations are anchored more closely around target inflation rates, and are less prone to move with changing cost conditions. Firms are rational in our model and the probability distribution of marginal cost has been taken as given; thus we have implicitly assumed that monetary policy is credible. But this simply means that, if anything, our analysis is more applicable today with high credibility of monetary policy. Our results, therefore, support the view that prices may be a poor indicator of marginal costs is such an environment.

5 Concluding Remarks

Our main task in this paper has been to show that an economy with diverse private information has features that are not captured in representative individual models where all firms share the same information. The most distinctive of these features is the crowding out of private information by the commonly shared information. The impact of public information is greater for those economies where price competition is more fierce.

The observation that public signals have a disproportionately large impact in games with coordination elements is not new, but our main focus has been to present the arguments in a framework that is sufficiently simple so as to isolate the key effects at work. By basing our analysis on a static framework with a finite state space, we have been able to employ elementary methods from Markov chain theory. The results we obtain show how increased competitive pressures magnify the effects of strategic complementarity on the choice of prices across firms, and thereby increase the weight placed on higher order average beliefs in the optimal choice of firms. Higher order beliefs are less informative of the underlying fundamentals, and hence the equilibrium prices become more detached from the underlying cost conditions. The simulation results reported by Woodford (2003a) and others rely on these key features. By showing the origin of these features in a simple theoretical model and numerical example, we hope our results can throw light on precisely how differential information affects prices. This, and our preliminary discussion of policy implications, will hopefully lead to further avenues of research in this area.

A Derivation of Pricing Rule

We give an illustrative derivation of the pricing rule (1) in a partial equilibrium setting. A continuum of firms indexed by the unit interval $[0,1]$ choose prices for differentiated goods, where firm i faces the demand curve

$$y_i = \left(\frac{p_i}{P}\right)^{-\theta} Y \tag{18}$$

where p_i is the price charged by firm i, P is aggregate price level, given by the Dixit-Stiglitz price index $[\int_0^1 p_i^{1-\theta} di]^{\frac{1}{1-\theta}}$, and Y is the aggregate output given by $[\int_0^1 y_i^{\frac{\theta-1}{\theta}} di]^{\frac{\theta}{1-\theta}}$, and the parameter $\theta > 1$ is the elasticity of substitution between goods produced by different firms. The firm has some pricing power, but the higher is θ, the more sensitive is the demand for a firm's output with respect to the relative prices across firms. In the absence of uncertainty, the first order condition for profit maxmization is given by

$$\frac{p_i}{P} = \frac{\theta}{\theta - 1} c\left(y_i, Y\right) \tag{19}$$

where $c\left(y_i, Y\right)$ is the real marginal cost for firm i when his own output is y_i and the aggregate output is Y. Thus, firms charge a mark-up over marginal cost, where the mark-up is decreasing in θ. For illustration, let us suppose that the marginal cost for firm i takes the form:

$$c\left(y_i, Y\right) = A y_i^{\alpha} w\left(Y\right) \tag{20}$$

where α and A are positive constants, and $w\left(Y\right)$ is the real wage that depends on the aggregate output. Substituting out y_i from (19) using (18), we can write the first-order condition as

$$\frac{p_i}{P} = \left[\tfrac{\theta}{\theta-1} c\left(Y, Y\right)\right]^{\xi} \tag{21}$$

where the parameter ξ is given by

$$\xi = \frac{1}{1 + \alpha\theta} \tag{22}$$

From (21) and introducing uncertainty, we have

$$p_i = E_i\left(P\left[\frac{\theta}{\theta-1}c\left(Y,Y\right)\right]^{\xi}\right)$$

Taking logs,

$$\log p_i = \log E_i\left(P\left[\frac{\theta}{\theta-1}c\left(Y,Y\right)\right]^{\xi}\right) \tag{23}$$

To obtain (1) from (21) (up to an additive constant), we take two approximations. The first approximation is

$$\log E_i\left(P\cdot c\left(Y,Y\right)^{\xi}\right) \simeq E_i \log\left(P\cdot c\left(Y,Y\right)^{\xi}\right) + C \tag{24}$$

where C is some constant. This relationship would be exact in some special cases, such as when conditional on firm i's information set, P and $c\left(Y,Y\right)$ have a joint lognormal distribution. In such a case,

$$\begin{aligned} E_i\left(P\cdot c\left(Y,Y\right)^{\xi}\right) &= E_i\left(\exp\left[\log P + \xi\log c\left(Y,Y\right)\right]\right) \\ &= E_i\left(\exp\left[X\right]\right) \\ &= \exp\left[E_i\left(X\right) + \frac{1}{2}var_i\left(X\right)\right] \end{aligned}$$

where $X \equiv \log P + \xi\log c\left(Y,Y\right)$ is normally distributed conditional on firm i's information set. Thus, we have

$$\begin{aligned} \log E_i\left(P\cdot c\left(Y,Y\right)^{\xi}\right) &= E_i\left(X\right) + \frac{1}{2}var_i\left(X\right) \\ &= E_i \log\left(P\cdot c\left(Y,Y\right)^{\xi}\right) + C \end{aligned}$$

if we assume that $C = var_i\left(X\right)/2$ is identical across i. This will hold, for example, when the firms are symmetric *ex ante*, so that the signals of $\log\left(P\cdot c\left(Y,Y\right)^{\xi}\right)$ that each firm observes are (conditionally) normally distributed, with mean equal to $\log\left(P\cdot c\left(Y,Y\right)^{\xi}\right)$ and constant variance equal across firms.

Substituting (24) into (23), and omitting constants, we have

$$\log p_i = E_i \log P + \xi E_i \log c\,(Y, Y) \tag{25}$$

The second approximation we take is $q \simeq \log P$. Making the substitution into (25) leads to (1) directly, where $\chi \equiv \log c\,(Y, Y)$. Although we already have a linear relation between q_i, $E_i \log P$ and $E_i \chi$ in (25), many of the results in the paper require us to express (1) in terms of q. Specifically, this approximation is required when recursively substituting out for q to obtain (3), after averaging (2) across firms. Avoiding the approximation $q \simeq \log P$ would significantly increase the complexity of the calculation of higher-order beliefs; in particular, analytic solutions would no longer be attainable.

It is worth noting, however, that the closely related literature on staggered price setting (e.g. Calvo contracts) has also relied on a log-linear approximation of the aggregate price index in the derivation of the New Keynesian Phillips Curve, as shown in Sbordone (*JME*, 2002, p. 270), Woodford (*Interest and Prices*, 2003) and many other recent papers. While the details are slightly different, the principle is the same: a log-linearisation of the price index is needed to derive a tractable expression of aggregate price dynamics.

B Alternative Proof of Theorem 6

An alternative proof of theorem 6 can be given in terms of the eigenvalues and eigenvectors of the average belief matrix. Let there be n states in Ω, and denote by p_{ij} the (i, j)-th entry of B. For the moment, we will assume that $p_{ij} > 0$ for all i, j. We'll return to comment on how the result genernalizes. Suppose there are N firms. Since p_{ij} is the average conditional probability of state j at state i, we have

$$p_{ij} = \frac{1}{N}\left(p_1\,(j|i) + p_2\,(j|i) + \cdots + p_n\,(j|i)\right)$$

where $p_k(j|i)$ is the k-th firm's conditional probability of state j at state i. Let $S(i,j)$ be the subset of individuals for whom states i and j belong to the same element of their information partition. Clearly, $S(i,j) = S(j,i)$. Denote by $P_k(i)$ the ex ante probability of the cell of individual k's partition that contains state i. Then,

$$p_{ij} = \frac{1}{N} \sum_{k \in S(i,j)} \frac{p_j}{P_k(i)} = \frac{p_j}{P(i,j)}$$

where $P(i,j)$ is defined as $\frac{1}{P(i,j)} \equiv \frac{1}{N} \sum_{k \in S(i,j)} \frac{1}{P_k(i)}$. Note that

$$\sum_{k \in S(i,j)} \frac{1}{P_k(i)} = \sum_{k \in S(j,i)} \frac{1}{P_k(i)} = \sum_{k \in S(j,i)} \frac{1}{P_k(j)}$$

so that $P(i,j) = P(j,i)$. Thus, the matrix B can be written as

$$B \equiv \begin{bmatrix} \frac{p_1}{P(1,1)} & \frac{p_2}{P(1,2)} & \cdots & \frac{p_n}{P(1,n)} \\ \frac{p_1}{P(2,1)} & \frac{p_2}{P(2,2)} & \cdots & \frac{p_n}{P(2,n)} \\ \vdots & \vdots & \ddots & \\ \frac{p_1}{P(n,1)} & \frac{p_2}{P(n,2)} & \cdots & \frac{p_n}{P(n,n)} \end{bmatrix}$$

where p_i is the ex ante probability of state i. We can show that B is diagonalizable and has real-valued eigenvalues. To see this, define two matrices D and A. D is the diagonal matrix defined as:

$$D = \begin{bmatrix} \sqrt{p_1} & & & \\ & \sqrt{p_2} & & \\ & & \ddots & \\ & & & \sqrt{p_n} \end{bmatrix}$$

A is a symmetric matrix defined as

$$A = \begin{bmatrix} \frac{p_1}{P(1,1)} & \frac{\sqrt{p_1 p_2}}{P(1,2)} & \cdots & \frac{\sqrt{p_1 p_n}}{P(1,n)} \\ \frac{\sqrt{p_2 p_1}}{P(2,1)} & \frac{p_2}{P(2,2)} & & \frac{\sqrt{p_2 p_n}}{P(2,n)} \\ \vdots & & \ddots & \\ \frac{\sqrt{p_n p_1}}{P(n,1)} & \frac{\sqrt{p_n p_2}}{P(n,2)} & & \frac{p_n}{P(n,n)} \end{bmatrix}$$

It can be verified that $B = D^{-1}AD$. Since A is a symmetric matrix, it is diagonalizable and has real-valued eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, and there is an orthogonal matrix E whose columns are the eigenvectors of A. In other words, $A = E\Lambda E'$ where

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

and where E' is the transpose of E. Thus,

$$B = D^{-1}AD = D^{-1}E\Lambda E'D = C\Lambda C^{-1}$$

where $C = D^{-1}E$. Thus, B is diagonalizable, has real valued eigenvalues, and whose eigenvectors are given by the columns of C. The matrix C of eigenvectors can be derived as follows. Since the rows of B sum to one, we know that the vector

$$u = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

satisfies $u = Bu$. Thus, u is the eigenvector that corresponds to the eigenvalue 1, which is the largest eigenvalue of B. From this, we have

$$u = Bu = D^{-1}ADu$$

so that $Du = ADu$. In other words, Du is the eigenvector corresponding to the eigenvalue 1 in A. Du is the column vector

$$\begin{bmatrix} \sqrt{p_1} \\ \vdots \\ \sqrt{p_n} \end{bmatrix}$$

Thus, the orthogonal matrix E of eigenvectors of B has the form:

$$E = \left[\begin{array}{c|c} \sqrt{p_1} & \cdots \\ \sqrt{p_2} & \cdots \\ \vdots & \vdots \\ \sqrt{p_n} & \cdots \end{array} \right]$$

and

$$E^{-1} = E' = \left[\begin{array}{cccc} \sqrt{p_1} & \sqrt{p_2} & \cdots & \sqrt{p_n} \\ \hline \vdots & \vdots & & \vdots \end{array} \right]$$

From this, and from (12), we can write the matrix of eigenvectors C as follows.

$$\begin{aligned} C &= \left[\begin{array}{c|c|c|c|c} 1 & \vdots & \vdots & & \vdots \\ 1 & c_2 & c_3 & & c_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & & & & \end{array} \right] \\ C^{-1} &= \left[\begin{array}{ccccc} p_1 & p_2 & p_3 & \cdots & p_n \\ \hline p_1 c_{21} & p_2 c_{22} & p_3 c_{23} & \cdots & p_n c_{2n} \\ \vdots & \vdots & & & \vdots \\ p_1 c_{n1} & p_2 c_{n2} & p_3 c_{n3} & \cdots & p_n c_{nn} \end{array} \right] \end{aligned}$$

where c_k is the kth eigenvector of B, and where c_{kj} is the jth entry of c_k. Bringing all the elements together, we have:

Lemma 7 *The matrix B of average conditional beliefs satisfies*

$$B = \left[\begin{array}{c|c|c|c} 1 & \vdots & & \vdots \\ 1 & c_2 & & c_n \\ \vdots & \vdots & \cdots & \vdots \\ 1 & & & \end{array} \right] \left[\begin{array}{cccc} 1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{array} \right] \left[\begin{array}{cccc} p_1 & p_2 & \cdots & p_n \\ \hline p_1 c_{21} & p_2 c_{22} & \cdots & p_n c_{2n} \\ \vdots & \vdots & & \vdots \\ p_1 c_{n1} & p_2 c_{n2} & \cdots & p_n c_{nn} \end{array} \right]$$

Let f be a random variable, expressed as a column vector conformable with B. Then,

$$C^{-1}f = \begin{bmatrix} p_1 & p_2 & \cdots & p_n \\ \hline p_1 c_{21} & p_2 c_{22} & \cdots & p_n c_{2n} \\ \vdots & \vdots & & \vdots \\ p_1 c_{n1} & p_2 c_{n2} & \cdots & p_n c_{nn} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} E(f) \\ E(c_2 f) \\ \vdots \\ E(c_n f) \end{bmatrix}$$

where $E(.)$ is the expectations operator with respect to public information only (i.e. with respect to the ex ante probabilities $p_1, p_2, \cdots, p_n$). $E(c_k f)$ denotes the expectation of the state by state product of c_k and f. Since $B^k = C\Lambda^k C^{-1}$, we can write

$$\begin{aligned} B^k f &= C\Lambda^k C^{-1} f \\ &= \begin{bmatrix} 1 & c_{21} & c_{31} & & c_{n1} \\ 1 & c_{22} & c_{32} & & c_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & c_{2n} & c_{3n} & & c_{nn} \end{bmatrix} \begin{bmatrix} E(f) \\ \lambda_2^k E(c_2 f) \\ \vdots \\ \lambda_n^k E(c_n f) \end{bmatrix} \\ &= \begin{bmatrix} E(f) + \sum_{j=2}^n \lambda_j^k c_{j1} E(c_k f) \\ E(f) + \sum_{j=2}^n \lambda_j^k c_{j2} E(c_k f) \\ \vdots \\ E(f) + \sum_{j=2}^n \lambda_j^k c_{jn} E(c_k f) \end{bmatrix} \to \begin{bmatrix} E(f) \\ E(f) \\ \vdots \\ E(f) \end{bmatrix} \quad \text{as } k \to \infty \end{aligned}$$

since $\lambda_j < 1$ for $j \geq 2$. Thus, theorem 6 holds when matrix B has positive entries for all i and j. When B has zero entries, we know that there is some t such that the power matrix B^t has entries that are all strictly positive. This is due to the ergodicity of the Markov chain. When the meet of the individual partitions is non-trivial, then there are as many unit eigenvalues as there are elements in the meet. So, the above analysis would apply to each element of the meet.

References

[1] Adam, Klaus (2003): "Optimal monetary policy with imperfect common knowledge", unpublished paper, Frankfurt University.

[2] Allen, Franklin, Stephen Morris and Hyun Song Shin (2002): "Beauty contests, bubbles and iterated expectations in asset markets", http://nuff.ox.ac.uk/users/shin/working.htm

[3] Amato, Jeffery D and Hyun Song Shin (2003): "Monetary Policy in a Differential Information Economy", unpublished paper, London School of Economics.

[4] Andersen, Palle and William Wascher (2001) "Understanding the Recent Behaviour of Inflation: an Empirical Study of Wage and Price Developments in Eight Countries" BIS papers, no. 3, Bank for International Settlements, `www.bis.org`

[5] Bacchetta, Philippe and Eric van Wincoop (2002): "Can information heterogeneity explain the exchange rate determination puzzle?", unpublished paper, Studienzentrum Gerzensee.

[6] Bacchetta, Philippe and Eric van Wincoop (2004) "Higher Order Expectations in Asset Pricing", unpublished paper, Studienzentrum Gerzensee.

[7] Ball, L. and D. Romer (1990) "Real Rigidities and the Non-Neutrality of Money" *Review of Economic Studies*, 57, 183-203.

[8] Blackwell, D. (1951) "The Comparison of Experiments" *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93-102, University of California Press.

[9] Blinder, Alan, Charles Goodhart, Philipp Hildebrand, David Lipton and Charles Wyplosz (2001): *How Do Central Banks Talk?* Geneva Report on the World Economy 3, Centre for Economic Policy Research, London.

[10] Bomfim, Antulio N (2001): "Measurement error in general equilibrium: the aggregate effects of noisy economic indicators", *Journal of Monetary Economics*, 48, pp 585-603.

[11] Borio, Claudio and Philip Lowe (2002) "Asset Prices, Financial and Monetary Stability: Exploring the Nexus" BIS working paper, no. 114, Bank for International Settlements, `www.bis.org`

[12] Bowman, D. (2003) "Market Power and Inflation" Board of Governers of the Federal Reserve System, International Finance Discussion Paper 783.

[13] Fry, Maxwell, Deanne Julius, Lavan Mahadeva, Sandra Roger and Gabriel Sterne (2000) "The Use of Explicit Targets for Monetary Policy: Practical Experiences in 93 Economies in the 1990s", in Lavan Mahadeva and Gabriel Sterne (eds), *Monetary Policy Frameworks in a Global Context*, Routledge, London, UK, pp 29-56.

[14] Gali, J. and M. Gertler (1999) "Inflation Dynamics: a Structural Econometric Analysis" *Journal of Monetary Economics*, 44, 195-222.

[15] Giannoni, Marc and Michael Woodford (2002): "Optimal interest-rate rules: I. General theory", unpublished paper, Princeton University.

[16] Hellwig, Christian (2002): "Public announcements, adjustment delays and the business cycle", unpublished paper, UCLA.

[17] Kasa, Kenneth (2000): "Forecasting the forecasts of others in the frequency domain", *Review of Economic Dynamics*, 3, pp 726-756.

[18] Keynes, John Maynard (1936): *The General Theory of Employment, Interest, and Money*, MacMillan, London.

[19] Morris, Stephen and Hyun Song Shin (2002): "Social value of public information", *American Economic Review*, 52, pp 1521-1534.

[20] Mussa, Michael (2003) "Remarks on Achieving Monetary and Financial Stability" BIS papers, no. 18, Bank for International Settlements, `www.bis.org`

[21] OECD (2002) "Product Market Competition and Economic Performance" working paper ECO/CPE/WP1 (2002) 11.

[22] Pearlman, Joseph G and Thomas J Sargent (2002): "Knowing the forecasts of others", unpublished paper, NYU.

[23] Phelps, Edmund S (1983): "The trouble with 'rational expectations' and the problem of inflation stabilization" in R Frydman and E S Phelps (eds) *Individual Forecasting and Aggregate Outcomes*, Cambridge University Press, New York, pp 31-41.

[24] Rogoff, K. (2003) "Globalization and Global Disinflation" Paper for the Federal Reserve Bank of Kansas conference in Jackson Hole, 2003.

[25] Rotemberg, Julio and Michael Woodford (1997): "An optimisation-based econometric framework for the evaluation of monetary policy", in Ben S Bernanke and Julio Rotemberg (eds), *NBER Macroeconomics Annual*, MIT Press, Cambridge, MA, pp 297-346.

[26] Samet, Dov (1998): "Iterated expectations and common priors," *Games and Economic Behavior*, 24, pp 131-141.

[27] Sargent, Thomas (1991) "Equilibrium with Signal Extraction from Endogenous Variables" *Journal of Economic Dynamics and Control*, 15, pp 245-273.

[28] Sbordone, Argia M (2002): "Prices and unit labour costs: a new test of price stickiness", *Journal of Monetary Economics*, 49, pp 265-292.

[29] Shin, Hyun Song and Timothy Williamson (1996): "How much common belief is necessary for a convention?" *Games and Economic Behavior*, 13, pp 252-268.

[30] Sims, Christopher A (2002): "Implications of rational inattention", unpublished paper, Princeton University.

[31] Svensson, Lars (1999): "Price-level targeting vs inflation targeting: a free lunch?", *Journal of Money, Credit and Banking*, 31, pp 277-295.

[32] Svensson, Lars (2002): "Monetary policy and real stabilization", paper presented at the Federal Reserve Bank of Kansas City's Jackson Hole symposium (available at www.princeton.edu/~svensson).

[33] Svensson, Lars (2003a): "Comment on Woodford", in P Aghion, R Frydman, J Stiglitz and M Woodford (eds) *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S Phelps*, Princeton University Press, Princeton, pp 59-63.

[34] Svensson, Lars (2003b): "What is wrong with Taylor rules? Using judgment in monetary policy through targeting rules", forthcoming, *Journal of Economic Literature.*

[35] Svensson, Lars and Michael Woodford (2003): "Implementing optimal policy through inflation-forecast targeting", unpublished paper, Princeton University (available at www.princeton.edu/~svensson).

[36] Taylor, John (1993): "Discretion versus policy rules in practice", *Carnegie-Rochester Conference Series on Public Policy*, 39, pp 195-214.

[37] Townsend, Robert M (1983): "Forecasting the forecasts of others", *Journal of Political Economy*, 91, pp 546-588.

[38] Ui, Takashi (2003): "A note on the Lucas model: iterated expectations and the neutrality of money", unpublished paper, Yokohama National University.

[39] Vestin, David (1999): "Price-level targeting versus inflation targeting in a forward-looking model", unpublished paper, Stockholm University.

[40] Woodford, Michael (1999): "Optimal monetary policy inertia", *NBER Working Paper No 7261.*

[41] Woodford, Michael (2003a): "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 Honor of Edmund S Phelps*, Princeton University Press, Princeton, pp 25-58.

[42] Woodford, Michael (2003b): *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, Princeton.

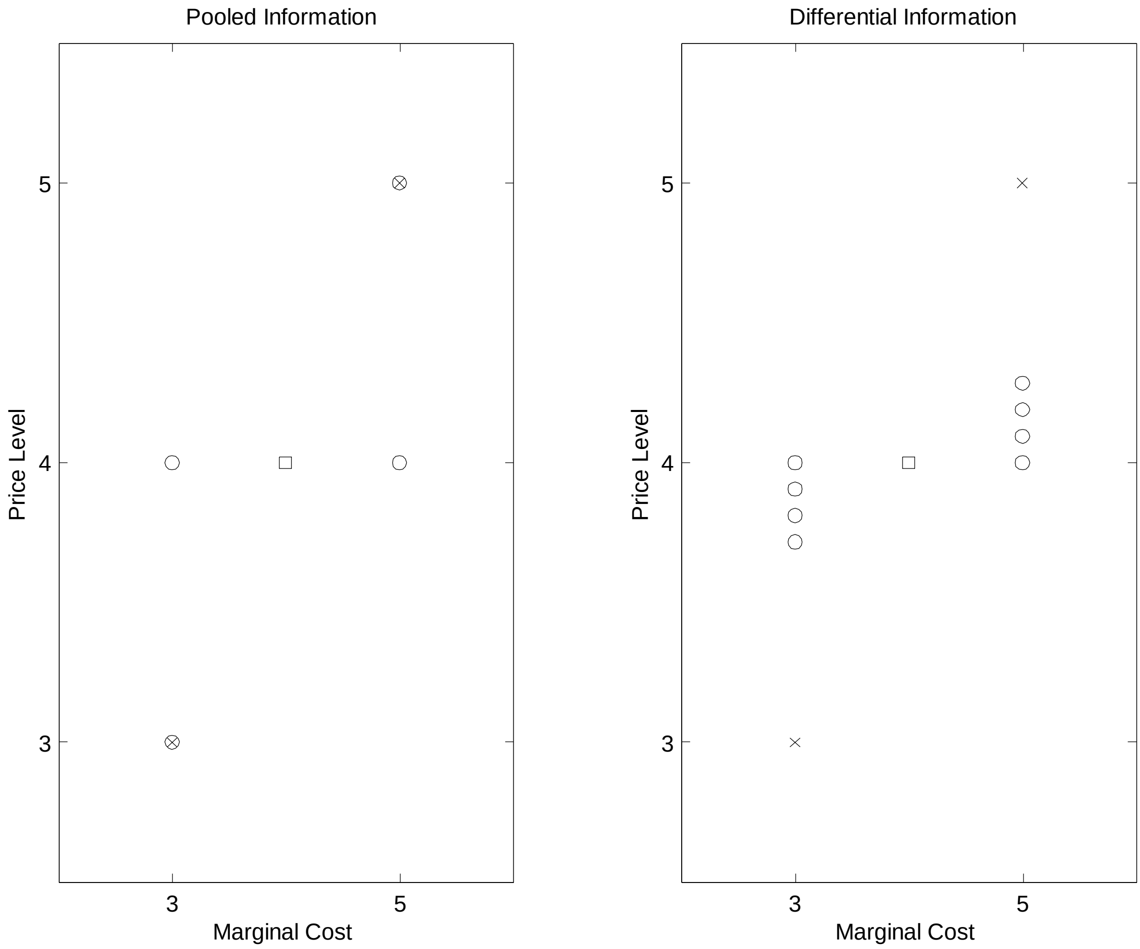
Table 1

Example of a State Space for an Economy with Differential Information

State	z	ϵ_1	ϵ_2	ϵ_3	Probability
1	z_H	ϵ	ϵ	ϵ	$p_H p_\epsilon^3$
2	z_H	ϵ	ϵ	0	$p_H p_\epsilon^2(1-2p_\epsilon)$
3	z_H	ϵ	ϵ	$-\epsilon$	$p_H p_\epsilon^3$
4	z_H	ϵ	0	ϵ	$p_H p_\epsilon^2(1-2p_\epsilon)$
5	z_H	ϵ	0	0	$p_H p_\epsilon(1-2p_\epsilon)^2$
6	z_H	ϵ	0	$-\epsilon$	$p_H p_\epsilon^2(1-2p_\epsilon)$
7	z_H	ϵ	$-\epsilon$	ϵ	$p_H p_\epsilon^3$
8	z_H	ϵ	$-\epsilon$	0	$p_H p_\epsilon^2(1-2p_\epsilon)$
9	z_H	ϵ	$-\epsilon$	$-\epsilon$	$p_H p_\epsilon^3$
10	z_H	0	ϵ	ϵ	$p_H p_\epsilon^2(1-2p_\epsilon)$
		...			
18	z_H	0	$-\epsilon$	$-\epsilon$	$p_H p_\epsilon^2(1-2p_\epsilon)$
19	z_H	$-\epsilon$	ϵ	ϵ	$p_H p_\epsilon^3$
		...			
27	z_H	$-\epsilon$	$-\epsilon$	$-\epsilon$	$p_H p_\epsilon^3$
28	z_L	ϵ	ϵ	ϵ	$p_L p_\epsilon^3$
		...			
54	z_L	$-\epsilon$	$-\epsilon$	$-\epsilon$	$p_L p_\epsilon^3$

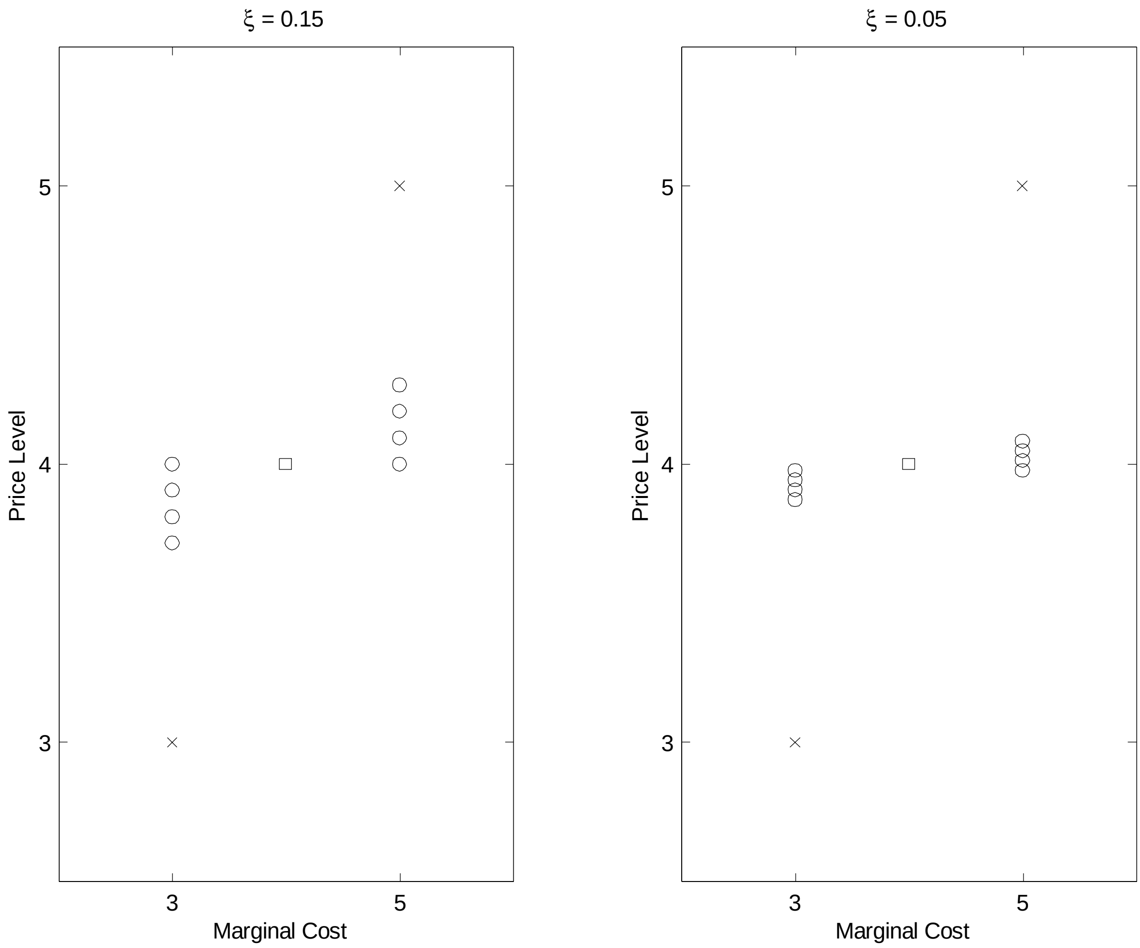
Figure 1

The Price Level Under Pooled and Differential Information Structures



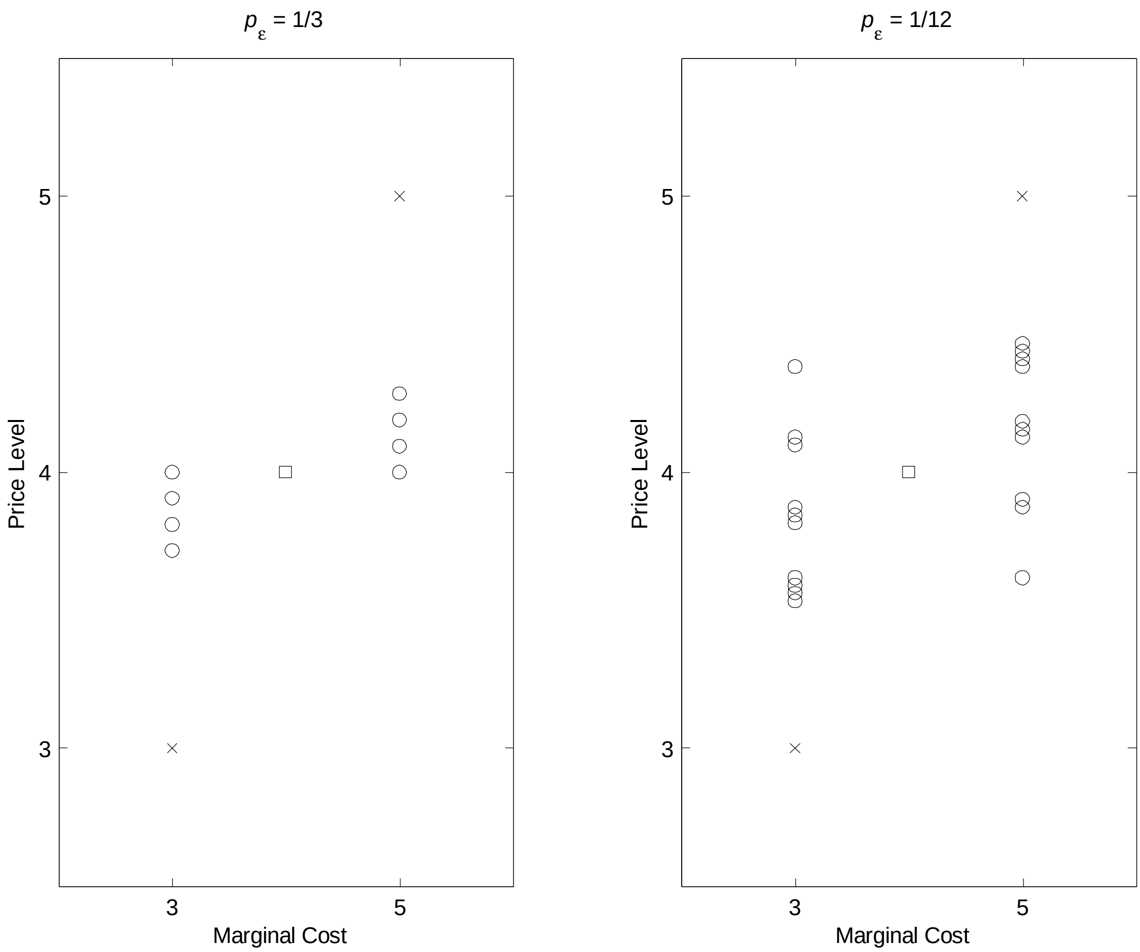
Notes: The circles represent the set of feasible values for the price level against marginal cost in pooled information (left panel) and differential information (right panel) economies. The markers denoted by an $\times$ give the tuples for price and marginal cost in a perfect information economy, while the square is the tuple of means in all the economies.

Figure 2
Impact on the Price Level of Changing Weights on Higher-Order Beliefs



Notes: The circles represent the set of feasible values for the price level against marginal cost in a differential information economy under different assumptions for ξ. The markers denoted by an $\times$ give the tuples for price and marginal cost in a perfect information economy, while the square is the tuple of means in both types of economies.

Figure 3
Impact on the Price Level of the Relative Precision of Private Signals



Notes: The circles represent the set of feasible values for the price level against marginal cost in a differential information economy under different assumptions for p_ϵ. The markers denoted by an $\times$ give the tuples for price and marginal cost in a perfect information economy, while the square is the tuple of means in both types of economies.