BANK FOR INTERNATIONAL SETTLEMENTS

BIS Working Papers

No 138

Public and private information in monetary policy models

by Jeffery D Amato* and Hyun Song Shin**

Monetary and Economic Department

September 2003

* Bank for International Settlements

** London School of Economics

Abstract

This paper examines the impact of public information in an economy where agents also have diverse private information. Since disclosures by central banks are an important source of public information, we are able to assess how the words of central bankers shape expectations, in addition to their actions. In an otherwise standard macro model, the disproportionate role of public information degrades the information value of economic outcomes, alters the welfare consequences of increased precision of public information and generates distinctive time series characteristics of some macro variables.

JEL Classification Numbers: D43, D82, D84, E31, E52
Keywords: imperfect information, monopolistic competition, targeting rule, Markov chain, Kalman filter

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The views expressed in them are those of their authors and not necessarily the views of the BIS.

Copies of publications are available from:

Bank for International Settlements
Press & Communications
CH-4002 Basel, Switzerland

E-mail: publications@bis.org

Fax: +41 61 280 9100 and +41 61 280 8100

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)

ISSN 1682-7678 (online)

Foreword

On 28-29 March 2003, the BIS held a conference on “Monetary stability, financial stability and the business cycle”. This event brought together central bankers, academics and market participants to exchange views on this issue (see the conference programme and list of participants in this document). This paper was presented at the conference. Also included in this publication are the comments by the discussants. The views expressed are those of the author(s) and not those of the BIS. The opening speech at the conference by the BIS General Manager and the prepared remarks of the four participants on the policy panel are being published in a single volume in the BIS Papers series.

Conference on
“Monetary stability,financial stability and the business cycle”
28-29 March 2003, Basel

Conference programme

Opening keynote remarks

Andrew Crockett (Bank for International Settlements)

Session I: The lessons from history

Chair: William White (Bank for International Settlements)

Paper 1: The price level, relative prices and economic stability: aspects of the interwar debate

Author: David Laidler (University of Western Ontario)

Discussants: Olivier Blanchard (Massachusetts Institute of Technology)
Nobuhiro Kiyotaki (London School of Economics)

Paper 2: The Great Depression as a credit boom gone wrong

Authors: Barry Eichengreen (University of California, Berkeley)
Kris Mitchener (Santa Clara University)

Discussants: Michael Bordo (Rutgers University)
Charles Goodhart (London School of Economics)

Session II: Monetary and financial frictions in business fluctuations

Chair: John Moore (London School of Economics)

Paper 3: Public and private information in monetary policy models

Authors: Jeffery Amato (Bank for International Settlements)
Hyun Song Shin (London School of Economics)

Discussants: Marvin Goodfriend (Federal Reserve Bank of Richmond)
Lars Svensson (Princeton University)

Paper 4: External constraints on monetary policy and the financial accelerator

Authors: Mark Gertler (New York University)
Simon Gilchrist (Boston University)
Fabio Natalucci (Board of Governors of the Federal Reserve System)

Discussants: Philippe Bacchetta (Study Center Gerzensee)
Philip Lowe (Reserve Bank of Australia)

Session III: Monetary policy challenges

Chair: Charles Freedman (Bank of Canada)

Paper 5: Asset prices, financial imbalances and monetary policy: are inflation targets enough?

Author: Charles Bean (Bank of England)

Discussants: Ignazio Visco (Bank of Italy)
Sushil Wadhwani (Wadhwani Asset Management LLP)

Paper 6: Financial strains and the zero lower bound: the Japanese experience

Author: Mitsuhiro Fukao (Keio University)

Discussants: Ignazio Angeloni (European Central Bank)
Jürgen von Hagen (University of Bonn)

Session IV: Achieving monetary and financial stability

Panel discussion

Chair: Andrew Crockett (Bank for International Settlements)

Panellists: Roger Ferguson (Board of Governors of the Federal Reserve System)
Otmar Issing (European Central Bank)
Michael Mussa (Institute for International Economics)
Yutaka Yamaguchi (formerly Bank of Japan)

Conference on
"Monetary stability,financial stability and the business cycle"
28-29 March 2003, Basel

Participants in the conference

Ignazio Angeloni	European Central Bank
Philippe Bacchetta	Study Center Gerzensee
Armando Baqueiro Cãrdenas	Bank of Mexico
Charles Bean	Bank of England
Olivier J Blanchard	Massachusetts Institute of Technology
Michael Bordo	Rutgers University
Barry Eichengreen	University of California, Berkeley
Charles Freedman	Bank of Canada
Mitsuhiro Fukao	Keio University
Simon Gilchrist	Boston University
Marvin Goodfriend	Federal Reserve Bank of Richmond
Charles Goodhart	London School of Economics
Otmar Issing	European Central Bank
Nigel Jenkinson	Bank of England
Thomas J Jordan	Swiss National Bank
Nobuhiro Kiyotaki	London School of Economics
David E Laidler	University of Western Ontario
Flemming Larsen	International Monetary Fund
Philip Lowe	Reserve Bank of Australia
Kris J Mitchener	Santa Clara University
John Moore	London School of Economics
Michael Mussa	Institute for International Economics
Fabio M Natalucci	Board of Governors of the Federal Reserve System
Peter Praet	National Bank of Belgium

Jan F Qvigstad	Central Bank of Norway
Hermann Remsperger	Deutsche Bundesbank
Hyun Song Shin	London School of Economics
Marc-Olivier Strauss-Kahn	Bank of France
Lars E O Svensson	Princeton University
Giovanni Toniolo	University of Rome Tor Vergata
José Viñals	Bank of Spain
Ignazio Visco	Bank of Italy
Jürgen von Hagen	University of Bonn
Sushil B Wadhwani	Wadhwani Asset Management LLP
Charles Wyplosz	Graduate Institute of International Studies
Yutaka Yamaguchi	formerly Bank of Japan
Gang Yi	The People's Bank of China
Andrew Crockett André Icard William White Renato Filosa Claudio Borio Gabriele Galati Jeffery Amato William English Andrew Filardo Ben Fung (Representative Office for Asia and the Pacific)	Bank for International Settlements

Contents

1 Introduction[1]

One of the often-cited virtues of a decentralised 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 policymakers 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 inflationary forces operating 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 pricing pressure. Such signs can be met by prompt central bank action to head off any inflationary forces through the use of monetary policy instruments. More generally, these actions can be codified in a more systematic framework for the setting of nominal interest rates, for instance as part of an 'inflation-forecast targeting' regime.

However, by the nature of its task, the central bank cannot confine its role merely to being a vigilant, but detached observer. Its monetary policy role implies that it must also engage in the active shaping and influencing of events (see Blinder, Goodhart, Hildebrand, Lipton and Wyplosz (2001)). For economic agents, who are all interested parties in the future course of action of the central bank, the signals conveyed by the central bank in its deeds and words have a material impact on how economic decisions are arrived at. For this reason, Svensson

[1]We thank Andy Filardo, Marvin Goodfriend, Nobu Kiyotaki, Stephen Morris and Lars Svensson for their comments on earlier drafts. Hyun Shin thanks the BIS for its hospitality during his visit in the summer of 2002, when this paper was prepared. The views are those of the authors and do not necessarily represent those of the BIS.

(2002) and Svensson and Woodford (2003) have advocated the announcement of the future path of the short-term policy interest rate as part of a central bank's overall policy of inflation-forecast targeting.

Monetary policy thus entails a dual role. As well as being a vigilant *observer* of outcomes, the central bank must also be able to *shape* the outcomes. In an economy with dispersed information, the central bank's actions and the information it releases constitute a shared benchmark in the information processing decisions of economic agents. In particular, the central bank's disclosures — or, in general, any type of credible public information — become a powerful focal point for the coordination of expectations among such agents.

Against this backdrop, this paper assesses the implications of public information in a small-scale monetary policy model in which agents have imperfect common knowledge on the state of the economy. We employ a model that is standard in most respects, but one that recognises the importance of decentralised information gathering and the resulting differential information in the economy. In particular, building on recent work by Woodford (2003a), our focus is on the pricing behaviour of monopolistically competitive firms with access to both private and public information.

Our analysis proceeds in two steps. Beginning with a series of simplified examples, we show how differentially informed firms follow pricing rules that suppress their own information, but instead put disporportionately large weight on commonly shared information; that is, firms suppress their private information on the underlying demand and cost conditions far more than is justified when the estimates of fundamentals are common knowledge. For reasonable values for the strength of strategic complementarity, the aggregate price suffers substantial information loss, and therefore ceases to be an informative signal of the underlying demand and cost conditions.

Following up on our partial equilibrium example, we then develop a general equilibrium model incorporating households and the central bank. A complete monetary policy model allows us to consider the dynamic implications of the presence of both public and private information under specific monetary policies. Our first objective is to solve for a rational expectations equilibrium with a finite dimensional state vector. In addition, we also wish to show whether equilibria exist under policies that follow simple rules, as explored in the recent monetary policy literature. This is followed by an investigation of the equilibrium properties of the model. First, we examine how changes in the degree of strategic complementarity and precision of public information affect the sample paths of the price level. Second, we investigate the dynamic responses of higher-order expectations to shocks in the underlying economic fundmentals, with particular emphasis on the role of public signals. Third, we trace out the impact of the relative precision of public and private signals on the volatility of macroeconomic aggregates.

In the next section, we provide a brief overview of related literature. Section 3 provides some conceptual background by means of simplified examples of pricing under differential information. Section 4 develops a complete macroeconomic model. Equilibrium is solved for in Section 5, while Section 6 explores some properties of the equilibrium. Section 7 concludes. An appendix contains further technical results.

2 Related literature

From a theoretical perspective, we have good grounds to conjecture that the 'climate of opinion' as embodied in the commonly shared information in an economy will play a disproportionate role in determining the outcome. A strand of the macroeconomics literature begun by Townsend (1983) and Phelps (1983), and recently developed and quantified by Woodford (2003a), examines the impact of decentralised information processing by individual agents in an environment where their interests are intertwined. Indeed, Phelps's paper is explicitly couched in terms of the importance of higher-order beliefs — that is, beliefs about the beliefs of others. For Woodford, the intertwining of interests arises from the strategic complementarities in the pricing decisions of firms. In setting prices, firms try to second-guess the pricing strategies of their potential competitors for market share. Even 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 fundamentals. The implication is that average prices suffer some impairment in serving as a barometer of the underlying cost and demand conditions.

These results are bolstered by recent theoretical studies into the impact of public and private information in a number of related contexts. They suggest that there is potential for the aggregate outcome to be overly sensitive to commonly shared information relative to reactions that are justified when all the available information is used in a socially efficient way. Morris and Shin (2002) note how increased precision of public information may impair social welfare in a game of second-guessing in the manner of Keynes's 'beauty contest' that has close formal similarities with the papers by Phelps and Woodford. Allen, Morris and Shin (2002) note that an asset's trading price may be a biased signal of its true value in a rational expectations equilibrium with uncertain supply, where the bias is toward the ex ante value of the asset.

A number of recent papers have revisited macroeconomic models with imperfect common knowledge by drawing on the recent modelling innovations for dealing with differential information. In independent work, Hellwig (2002) analyses the impact of public announcements in a semi-structural model with imperfect competition. He shows that public announcements allow quicker adjustment to fundamentals, but at the cost of greater noise. Adam (2003) considers optimal policy in a model with imperfect common knowledge, invoking results from the literature on information processing capacity. Bacchetta and van Wincoop (2002) explore the impact of public information in an asset pricing context. Pearlman and Sargent (2002) and Kasa (2000) extend the analysis of the models developed by Townsend (1983).

There has also been growing interest in examining more deeply the underlying rationale for imperfect common knowledge among agents. Is it possible that agents observe only noisy signals of aggregate fundamentals? If so, why do agents lack common knowledge? The latter question is easier to address, since it is presumed to be self-evident that agents have access to (at least partially) private information in the conduct of their own activities. One answer to the first question is that data on macroeconomic aggregates are subject to persistent measurement errors. Publicly available statistics rarely provide a completely accurate measure of the true underlying aggregates of economic interest. Bomfim (2001) has analysed the general equilibrium implications of measurement error in a common knowledge rational expectations setting. A second answer is that agents have limited information processing capabilities, along the lines of Sims (2002). The story is as follows. Consider dividing agents' activities into two parts: an information processing stage and a decision-making stage. Given the vast quantity of information at their disposal, both private and public in nature, it is conjectured that agents can only imperfectly filter this data into a set of statistics upon which to base decisions. But conditional upon their information sets, agents act optimally. A related argument is that a good deal of public information that agents pay attention to is imperfectly filtered by public sources, for example, newspaper reports or commentators on television.

The existence and likely use of both public and private information suggests that models with disparately-informed agents should take both types of signals into account. The strong likelihood that measurement errors in some key macroeconomic data series or that processing errors by agents persist indefinitely into the future suggests that the true state of the economy is never revealed. Combining these two features in a monetary policy model is a novel contribution of this paper.

One potential argument against the plausibility of the importance of higher-order beliefs in agents' behaviour is the degree of complexity involved in forming these beliefs (see, for instance, Svensson's (2003a) comments on Woodford (2003a)). If agents have only limited information processing capabilities, then how could they be expected to form expectations about others' expectations about others' expectations and so on? However, there is a clear distinction between the behaviours exhibited by agents and the informational constraints they face. Agents form and act upon higher-order beliefs because it is rational for them to do so. Invoking the well known billiard player analogy, agents act *as if* they have knowledge of the workings of the economy, which in our setting requires that they implictly second-guess others. By contrast, it is not clear how they can act *as if* they have perfect common knowledge of the economy's state. Indeed, a differential-information rational expectations economy places less stringent requirements upon agents than full information rational expectations models that are typical in the literature. The elegance of these latter models can be misleading regarding the enormous demands placed upon agents in both their behaviour, which we also impose, and information processing abilities, which we relax.

3 Conceptual background

Before developing our main arguments in a dynamic general equilibrium setting, let us introduce our conceptual building blocks by means of two simplified examples in a static context — for the discrete case and the Gaussian case. Our focus is on the equilibrium consequences of the pricing rule for firms that takes the form:

$$p_i = E_i p + \xi E_i x \tag{1}$$

where p_i is the (log) price set by firm i, p is the (log) average price across firms, x denotes the output gap (in real terms) — our "fundamental variable" — and ξ is a constant between 0 and 1. A rigorous derivation of (1) is presented in Section 4. The operator E_i denotes the conditional expectation with respect to firm i's information set. The pricing rule given by (1) arises in the classic treatment by Phelps (1983), and has been developed more recently by Woodford (2003a) for an economy with imperfectly competitive firms.

In a discussion that has subsequently proved to be influential, Phelps (1983) compared this pricing rule 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 (1983) also emphasised the importance of higher-order expectations — that of forecasting the forecasts of others. To see this, rewrite (1) in terms of the nominal output gap, defined as $q \equiv x + p$, yielding $p_i = (1 - \xi) E_i p + \xi E_i q$. Taking the average across firms,

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

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

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

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. Morris and Shin (2002) show how such a failure of the law of iterated expectations affects the welfare consequences of decision rules of this form, and note that increased precision of public information may be detrimental to welfare. The size of the parameter ξ proves to be crucial in determining the impact of differential information. In a monopolistically competitive model, ξ reflects, among other things, the degree of competition between firms. The more intense the competition — that is, the larger the elasticity of substitution between firms' goods — the smaller will be ξ, and hence the more important higher-order expectations in determining prices.

3.1 Discrete state space

Let us begin with the case when the underlying fundamental variable — the nominal output gap q — takes on finitely many possible values. In addition, all firms share common prior information and receive private signals of the fundamental during the period. More specifically, no firm observes q perfectly, but firm i observes an imperfect signal z_i of q, where z_i takes on finitely many possible values. Each firm observes the realisation of its own signal, but not the signals of other firms. Let us further suppose 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 define a *state* ω to be an ordered tuple:

$$\omega \equiv (q, z_1, z_2, \cdots, z_N)$$

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 q and the signals z_i. The prior is known to all firms, and represents the commonly shared assessment of the likelihood of various outcomes. However, once the firm observes its own signal z_i, it makes inferences on the economy based on the realisation of its own signal z_i. Thus, in this example of a static economy, all firms begin with common knowledge, but receive private signals before making decisions. However, this model can also be interpreted within the context of a *dynamic* economy, but one where all information is fully revealed at the end of each period. Seen from this perspective, the examples in this section are based on the extreme opposite assumption about information revelation compared to the macro model developed in later sections, where it is assumed that the true state is never revealed.

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 realisation of z_i is the same at ω and ω'. Some matrix notation is useful. Index the state space Ω by the set $\{1, 2, \cdots, |\Omega|\}$. In this section we adopt the convention of denoting a random variable $f : \Omega \to \mathbb{R}^{|\Omega|}$ as a *column vector* of length $|\Omega|$, while denoting any probability density over Ω as a *row vector* of the same dimension. Thus, 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 firm i's 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

$$\phi = \phi B_i \tag{4}$$

for all firms i. In other words, ϕ is a fixed point of the mapping defined by B_i. More specifically, note that B_i is a stochastic matrix — 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 (4) implies that the prior density ϕ is an *invariant distribution* over the states for this Markov chain. This formalisation of differential information environments in terms of Markov chains follows Shin and Williamson (1996) and Samet (1998).

For any random variable $f : \Omega \to \mathbb{R}^{|\Omega|}$, 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 have $E_i f = B_i f$. 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 k-th order iterated average expectation of f. Then, the equilibrium pricing rule (1) can be expressed in matrix form as

$$p_i = \xi B_i q + (1 - \xi) B_i p$$

where p_i is now a column vector whose j-th element corresponds to firm i's price in state j, and with similar redefinitions for p and q respectively. Taking the average across firms,

$$p = \xi B q + (1 - \xi) B p \tag{5}$$

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

$$\begin{aligned} p &= \xi \sum_{i=0}^{\infty} ((1-\xi) B)^k Bq \\ &= \xi (I - (1-\xi) B)^{-1} Bq \qquad (6) \\ &= MBq \end{aligned}$$

where $M = \xi (I - (1-\xi) B)^{-1}$. Thus, equilibrium average price p is given by (6).

Let us note some comparisons between (6) and the case where all firms observe the same signal, and hence where the law of iterated expectations holds. When all firms observe the same signal, the k-fold iterated average expectation collapses to the single average expectation, and we have the pricing rule:

$$p = Bq \qquad (7)$$

The difference between (6) and (7) lies in the role played by matrix M. Note that M is a stochastic matrix 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 $1 + (1-\xi) + (1-\xi)^2 + \cdots = 1/\xi$. It serves the role of "adding noise" (in the sense of Blackwell) to the average expectation of the fundamentals q. The effect of the noise is to smooth out the variability of prices across states. Thus, in going from (7) to (6) the average price becomes a less reliable signal of the output gap.

The noise matrix M is a convex combination of the higher-order beliefs $\{B^k\}$, and higher-order expectations contain much less information than lower-order expectations in the following sense. For any random variable f, denote by $\max f$ the highest realisation of f, and define $\min f$ analogously as the smallest realisation 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" version of D; or, equivalently, CDf is a "noisier" version of Df. So, the higher the order of the iterated expectation, the more rounded the peaks and troughs of the iterated expectation across states. The importance of the parameter ξ is now apparent. The smaller this parameter, the greater the weighting received by the higher-order beliefs in the noise matrix M, so that the prices are much less informative about the underlying fundamentals.

The limiting case for higher-order beliefs B^k as k becomes large is especially noteworthy. From (4), we know that

$$\phi = \phi B \tag{8}$$

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

$$\phi = \phi B = \phi B^2 = \phi B^3 = \cdots$$

so that ϕ is an invariant density for B^k, for any kth order average belief operator. 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

$$B^\infty = \begin{bmatrix} - & \phi & - \\ - & \phi & - \\ & \vdots & \\ - & \phi & - \end{bmatrix} \tag{9}$$

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. For any random variable f, successively higher-order beliefs are so noisy that all peaks and troughs converge to a constant function, where the constant is given by the prior expectation $\bar{f}$ (ie the expectation of f with respect to the prior density ϕ):

$$B^k f \to \begin{bmatrix} \bar{f} \\ \bar{f} \\ \vdots \\ \bar{f} \end{bmatrix} \quad \text{as } k \to \infty \tag{10}$$

The condition that guarantees (9) is the following.

Condition 1 *For any two states j and k, there is a positive probability of making a transition from j to k in finite time.*

In our context, condition 1 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 1. We then have lemma 2, which mirrors Samet's (1998) analogous result for the iteration of individual beliefs.

Lemma 2 *Suppose B satisfies condition 1. 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 1 has an interpretation in terms of the degree of information shared between the firms. It corresponds to the condition that

$$\bigcap_i \mathcal{I}_i = \emptyset \tag{11}$$

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. 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.

When the intersection $\bigcap_i \mathcal{I}_i$ is non-empty, then this means that there are signals that are observed by every firm. Hence, the outcomes of signals in $\bigcap_i \mathcal{I}_i$ become common knowledge among all firms. One such example would be an announcement by a central bank. Information contained in $\bigcap_i \mathcal{I}_i$ is thus *public*. 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 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 the 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 partitions 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 — 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 whose proof is given in Shin and Williamson (1996).

Lemma 3 *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 3 extends condition 1. The idea is that the transition matrix of the Markov chain defined by the average belief matrix B can be expressed in block diagonal form:

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

where each sub-matrix A_j defines an irreducible Markov chain that corresponds to an element of the meet $\bigwedge_i \mathcal{P}_i$.[2] Furthermore, we have $\phi = \phi B^\infty$, so that for

[2] In the static examples considered here, there is a simple way to view the relation between the model with private signals only and the model with both private and public signals. Consider the prior ϕ over the state space Ω in an economy with private and public signals. This can be transformed into an equivalent economy with only private signals where the prior is given by $\tilde{\phi}$ and the state space is redefined to be $\tilde{\Omega}$. The new state space $\tilde{\Omega}$ is a subset of Ω, where $\Omega/\tilde{\Omega}$ is the set of states ruled out by the revelation of the public signal.

any random variable f, the limit of the higher-order expectation is the conditional expectation based on the public signals only. In other words, we have:

Theorem 4 *As* $k \to \infty$,

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

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

In the appendix, 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 4 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.

The argument so far has relied on a finite state space Ω, but it can be extended to more general discrete spaces. Such an extension would be important for embedding the pricing decisions in a dynamic economy. Let time be discrete, indexed by the non-negative integers. There is a countable set of economic variables $\{f_1, f_2, f_3, \cdots\}$ that reflect the fundamentals of the economy such as productivity, preferences and other exogenous shocks, together with all signals observed by any economic agent of these variables. Each economic variable f_k can take on a countable number of realisations, drawn from the set S_k. The *outcome space* is the product space $S \equiv \prod_k S_k$. The *outcome* of the economy at time t — given by a specified outcome for each of the economic variables f_s — is thus an element of S. Since each S_k is countable, so is the outcome space S.

The *state space* Ω is then defined to be the set of all sequences drawn from the set S. Thus, a typical state ω is given by the sequence

$$\omega = (s_0, s_1, s_2, \cdots)$$

where each s_t is an element of the outcome space S. Thus, a state ω specifies the outcome of all economic variables at every date, and so is a maximally specific description of the world over the past, present and future.

Let Ω be endowed with a prior probability measure ϕ. Each economic variable f_s then defines a stochastic process in the usual way in terms of the sequence

$$(f_{s,0}, f_{s,1}, f_{s,2}, \cdots)$$

where $f_{s,t}$ is the random variable that maps each state ω to the outcome of the economic variable f_s at time t. The information set of agent i at date t is a set of random variables whose outcomes are observed by firm i at date t. We denote by $\mathcal{I}_{i,t}$ the information set of firm i at date t. The information set $\mathcal{I}_{i,t}$ defines the information partition of agent i at date t over the state space Ω. This information partition is denoted by $\mathcal{P}_{i,t}$. The *meet* of the individual partitions at t is the finest partition of Ω that is at least as coarse as each of the partitions in $\{\mathcal{P}_{i,t}\}$. The meet at t is denoted by $\mathcal{P}_t$. It is the partition generated by the intersection of all information sets at date t, as in our earlier discussion. The meet $\mathcal{P}_t$ represents the set of events that are common knowledge at date t.

The analysis of pricing decisions by firms can then be generalised to this new setting. By construction, the state space Ω is countable. Almost all of the notation and apparatus developed above for the finite Ω can then be used in our new setting, except that we should be mindful of those rules for matrix manipulation that are not valid for infinite matrices. Kemeny, Snell and Knapp (1966) is a textbook reference for how infinite matrices can be used in the context of countable state spaces.

As before, any probability measure over Ω is denoted as a *row vector*, while a random variable f is denoted as a column vector. For each date t, the *average belief matrix* B_t is defined in the natural way. The s-th row of B_t is the probability measure over Ω that represents the mean across firms of their conditional beliefs over Ω at date t. Then, the average price at date t satisfies

$$p_t = \xi B_t q_t + (1 - \xi) B_t p_t \tag{12}$$

where p_t is the average price at t, and q_t is the date t version of the random variable q in the static case. By successive substitution, and from the fact that $0 < \xi < 1$, we can solve for p_t.

$$p_t = \xi \sum_{i=0}^{\infty} \left((1 - \xi) B_t \right)^k B q_t \tag{13}$$

For finite Ω, we wrote the sum $\sum_{i=0}^{\infty} \left((1 - \xi) B_t \right)^k$ as $(I - (1 - \xi) B)^{-1}$. However, for infinite matrices, the notion of an inverse is not well defined, and we cannot simplify (13) any further (see Kemeny, Snell and Knapp (1966, Chapter 1)). There is also a more substantial change to our results in this more general framework. Condition 1 is no longer sufficient for the convergence of higher-order beliefs to the public expectation (that is, the analogue of lemma 2 fails). The Markov chain associated with B_t must also be *recurrent* in the sense of every state being visited infinitely often by the Markov chain. With this additional strengthening, we can then appeal to the standard result for Markov chains on the convergence to stationary distributions (see Karlin and Taylor (1975, p 35)) to extend theorem 4 to our more general setting.

3.2 Gaussian case

Having established the intuition for the importance of higher-order beliefs, we can now show how they can be translated into a Gaussian setting. For reasons of space, we confine ourselves to the static case. Thus, let θ be a normally distributed random variable with mean μ and variance $1/\beta_0$ representing the fundamentals of the economy, and let agent i's information set $\mathcal{I}_i$ contain signals $\{x_1, x_2, \cdots, x_n\}$, where

$$x_i = \theta + \varepsilon_i$$

and ε_i is normal with mean 0 and variance $1/\beta_i$, and ε_i is independent of θ, as well as other noise terms ε_j. Appealing to the formula for conditional expectations for jointly normal random variables,[3] agent i's conditional expectation of θ is:

$$E_i(\theta) = \mu + V_{\theta x} V_{xx}^{-1} (x - \mu) \tag{14}$$

[3] See, for example, Searle (1971, p 47).

where $V_{\theta x}$ is the row vector of covariances between θ and $(x_1, \cdots, x_n)$, V_{xx} is the covariance matrix of $(x_1, \cdots, x_n)$, and $(x - \mu)$ is the column vector of deviations of each x_i from its mean μ. In our case, we have

$$V_{\theta x} = \tfrac{1}{\beta_0}\left[1, 1, \cdots, 1\right]$$

$$V_{xx} = \begin{bmatrix} \frac{1}{\beta_0} + \frac{1}{\beta_1} & \frac{1}{\beta_0} & \cdots & \frac{1}{\beta_0} \\ \frac{1}{\beta_0} & \frac{1}{\beta_0} + \frac{1}{\beta_2} & \cdots & \frac{1}{\beta_0} \\ \frac{1}{\beta_0} & \frac{1}{\beta_0} & \ddots & \vdots \\ \frac{1}{\beta_0} & \frac{1}{\beta_0} & \cdots & \frac{1}{\beta_0} + \frac{1}{\beta_n} \end{bmatrix}$$

Also, it can be verified by multiplication that the (i, j)-th entry of the inverse matrix V_{xx}^{-1} is given by

$$\begin{cases} \frac{-\beta_i \beta_j}{\sum_{k=0}^{n} \beta_k} & \text{if } i \neq j \\ \beta_i \left(1 - \frac{\beta_i}{\sum_{k=0}^{n} \beta_k}\right) & \text{if } i = j \end{cases}$$

Thus,

$$V_{\theta x} V_{xx}^{-1} = \frac{1}{\sum_{k=0}^{n} \beta_k} \left[\beta_1, \beta_2, \cdots, \beta_n\right] \tag{15}$$

so that (14) is given by:

$$E_i(\theta) = \frac{\beta_0 \mu + \sum_{k=1}^{n} \beta_k x_k}{\sum_{k=0}^{n} \beta_k} \tag{16}$$

In other words, agent i's conditional expectation of θ is a convex combination of the signals in his information set $\mathcal{I}_i$ and the prior mean μ, where the weights are given by the relative precision of each signal.

Now, let us consider the set of all random variables in the economy. Using superscript notation, let y^0 be a vector of all public signals about the fundamentals θ. This vector includes all signals in the intersection $\cap_i \mathcal{I}_i$. The prior mean of θ is a public signal, and so belongs to y^0. Let y^i be a vector of non-public signals in i's information set (ie signals in $\mathcal{I}_i \setminus \cap_j \mathcal{I}_j$). Let z be the stacked vector:

$$z \equiv \begin{bmatrix} y^0 \\ y^1 \\ \vdots \\ y^N \\ \theta \end{bmatrix}$$

where y^0 is a vector of length $n_1 + 1$, y^i $(i = 1, ..., N)$ are vectors of length n_2 and $n_1 + n_2 = n$. Suppose that z is jointly normally distributed with covariance matrix V. Individual i's information set $\mathcal{I}_i$ consists of signals in y^0 and y^i, where y^0 are the signals that are shared by everyone, while y^i consist of the remaining signals in $\mathcal{I}_i$. Let $E_i z$ be i's conditional expectation of z. From (16), and from the fact that the noise terms ε_i all have mean zero, there is a stochastic matrix A_i such that

$$E_i z = A_i z$$

The matrix A_i has entries that correspond to the weights in (15) and the weight on the prior mean μ. The average expectation $\bar{E}z$ is the arithmetic average $\frac{1}{N}\sum_{i=1}^{N} E_i z = \frac{1}{N}\sum_{i=1}^{N} A_i z$. We denote:

$$\bar{E}z = Az$$

where $A \equiv \frac{1}{N}\sum_{i=1}^{N} A_i$. Individual i's expectation about the average expectation is given by

$$E_i \bar{E}z = A_i A z$$

Hence, the average expectation of the average expectation is given by

$$\bar{E}\bar{E}z = \left(\frac{1}{N}\sum_{i=1}^{N} A_i\right) Az = A^2 z$$

In general, the kth order iterated average expectation of z is given by $A^k z$. Let us partition A so that

$$A = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix} \tag{17}$$

where I is the identity matrix whose order is the number of public signals. That is, I is the same dimension as y^0. The top right-hand cell of the partitioned matrix is the zero matrix, since the average expectation of y^0 is y^0 itself. In other words, the average expectation of y^0 places zero weight on any of the non-public signals. On the other hand, note that $R \neq 0$, provided that the public signals have some information value. Hence, Q is a matrix with norm strictly less than 1, so that $Q^k \to 0$ as $k \to \infty$.

higher-order average expectations then have the following property. First, as the order of expectation becomes higher, more and more weight is placed on the public signals, and less weight is placed on the non-public signals. This is so, since

$$A^k = \begin{bmatrix} I & 0 \\ \left(\sum_{i=0}^{k-1} Q^i\right) R & Q^k \end{bmatrix}$$

and $\left\{\left(\sum_{i=0}^{k-1} Q^i\right) R\right\}_{k=0}^{\infty}$ is a sequence whose norm is increasing in k, while $\{Q^k\}$ is a sequence whose norm is decreasing in k. In the limit where $k \to \infty$, we have

$$\begin{aligned} A^k &\to \begin{bmatrix} I & 0 \\ \left(\sum_{i=0}^{\infty} Q^i\right) R & 0 \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ (I-Q)^{-1} R & 0 \end{bmatrix} \end{aligned}$$

Thus, in the limit as $k \to \infty$, the higher-order average expectation places weight only on the *public signals*. The private signals receive zero weight. We therefore have the analogue of theorem 4, but this time for the Gaussian world.

A Markov chain interpretation can also be given, although the Markov chain in the Gaussian example is one over *signals*, rather than *states of the world*. Each random variable in z is associated with a state in a Markov chain, whose transition matrix is given by A. The fact that A can be partitioned as in (17) means that the public signals correspond to the *absorbing states* of the Markov chain — that is, once the system settles on such a state, it never emerges. The private signals and the fundamentals θ correspond to all the transient states in the chain. The long-run probability of being in such a state is zero. The weights on the public signals in the higher-order expectations matrix A^k thus give the probability of having been absorbed at date k. As k becomes large, the probability of being absorbed tends to 1.

4 A monetary policy model

We now consider the general equilibrium implications of the presence of both public and private information in monetary policy models. Our analysis is based on a model with standard behavioural assumptions on households and firms. All agents are rational, in the sense that they know the structure of the economy and make optimal decisions based on their information sets. The only departure we make from the benchmark full information rational expectations setting is the absence of common knowledge of the state of the economy among some agents. Specifically, as in the partial equilibrium example studied in the previous section, we assume that firms receive private and public noisy signals of current shocks. By contrast, households and the central bank are assumed to observe these shocks perfectly. This helps keep the focus on the pricing decisions, where the presence of strategic complementarities allows differential information to have important dynamic effects.

In this section we describe the behaviour and information sets of households, firms and the central bank, respectively. In Section 5 we characterise equilibrium, while in Section 6 we provide some simulation results illustrating the properties of the model.

4.1 Households

Households maximise their discounted expected utility of consumption subject to their budget constraint. One issue that must be addressed at the outset is the potential implications of having households possess private information. As mentioned above, we assume that households have full knowledge of the state. This allows households to mitigate idiosyncractic risk in incomes through insurance markets without greatly complicating our analysis. Households make identical consumption choices and we avoid having to keep track of the distribution of wealth. However, our assumption of perfect income insurance is only reasonable if we assume that households have perfect common knowledge without introducing complications regarding costly state verification. In addition, we would need to consider how rational expectations equilibria are established in asset markets under differential information. Incorporating asset market issues would take us too far astray, and divert attention from the main focus of our paper. Thus, both for the purpose of ensuring identical consumption decisions and for the purpose of avoiding asset market complications with differential information, we model households as having maximally-specific information sets with regard to all economic variables that have been realised to date.

To be more specific, we will assume that at any date t, households' information sets are identical, and include the realisations of all current and past economic variables $\{f_1, f_2, \cdots\}$. Thus, at date t, all households have the information set

$$\mathcal{I}_t^* \equiv \cup_s \{f_{s,0}, f_{s,1}, \cdots, f_{s,t}\}$$

Households' conditional expectations operator at date t is given by

$$E_t(\cdot) \equiv E(\cdot|\mathcal{I}_t^*)$$

At date t, households know at least as much as any other agent in the economy, including Nature, who has chosen the latest realisations of the economic variables.

Each household z supplies labour services of one type, $H_t(z,i)$, for firm i, and seeks to maximise

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[u(C_t(z)) - v(H_t(z,i)) \right] \right\} \tag{18}$$

subject to the budget constraint

$$E_t[\delta_{t,t+1}\Xi_{t+1}] \leq \Xi_t + W_t(i)H_t(z,i) + \Phi_t - P_tC_t(z) \tag{19}$$

Within each period, the household derives utility, $u(\cdot)$, from consuming the Dixit-Stiglitz aggregate, $C_t(z)$, defined as

$$C_t(z) \equiv \left[\int_0^1 C_t(z,i)^{\frac{\epsilon-1}{\epsilon}} di \right]^{\frac{\epsilon}{\epsilon-1}} \tag{20}$$

where $C_t(z,i)$ is household z's consumption of product i and $\epsilon > 1$ is the elasticity of substitution between differentiated products. As ϵ increases, goods become ever closer substitutes (ie firms have *less* market power), and hence the degree of strategic complementarity increases. Supplying $H_t(z,i)$ hours reduces welfare, as indicated by the function $v(\cdot)$. We assume that labour markets are competitive and a equal number of households supply labour of type i.

Households can insure against idiosyncratic risk in incomes (as mentioned above) and therefore consume the identical amount given by C_t. In the budget constraint, P_t denotes the price index corresponding to the aggregate C_t defined as

$$P_t \equiv \left[\int_0^1 P_t(i)^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}} \tag{21}$$

where $P_t(i)$ is the price of product i; Ξ_t denotes the nominal value of the household's holdings of financial assets at the beginning of period t; $W_t(i)$ is the nominal hourly wage for supplying labour of type i; Φ_t is the household's share of firms' profits, which we assume are distributed lump-sum to households, and $\delta_{t,s}$ is a stochastic discount factor, pricing in period t assets whose payoffs are realised in period s. We assume there exists a riskless one-period nominal bond, the gross return on which is given by $R_t \equiv (E_t\delta_{t,t+1})^{-1}$. Finally, notice that we have not assumed that housholds can insure against idiosyncratic variation in labour supply, although, in equilibrium, households who supply labour to firm i will work the same amount, $H_t(i)$.

Given the overall level of consumption, households allocate their expenditures across goods according to

$$C_t(i) = \left[\frac{P_t(i)}{P_t}\right]^{-\epsilon} C_t \tag{22}$$

The first-order condition for determining the optimal level of consumption, given the allocation of consumption across goods expressed in (22), is $\Lambda_t = u_c(C_t)$, where Λ_t is the marginal utility of real income, and the standard Euler equation is given by

$$\Lambda_t/P_t = \beta R_t E_t[\Lambda_{t+1}/P_{t+1}] \tag{23}$$

A log-linear approximation of (23) around $\Lambda_t = \bar{\Lambda}$, $R_t = \bar{R}$ and $P_{t+1}/P_{t+1} = 1$ results in

$$\lambda_t = E_t\lambda_{t+1} + r_t - E_t\pi_{t+1} \tag{24}$$

where $\pi_{t+1} \equiv \log(P_{t+1}/P_t)$ is the inflation rate and lower case represents percent deviation of a variable from its steady state.

Market clearing requires that $C_t = Y_t - G_t$, where Y_t is the aggregate demand for output and G_t is an exogenous component of demand (eg exogenous government expenditures). Since $\Lambda_t = u_c(Y_t - G_t)$, λ_t can be expressed as

$$\lambda_t = -\sigma \left(y_t - g_t\right) \tag{25}$$

where $\sigma \equiv u_{cc}(\bar{C})\bar{C}/u_c(\bar{C})$ is the inverse of the intertemporal elasticity of substitution. Substituting out for λ_t in (24) yields a "forward-looking IS equation":

$$y_t - g_t = E_t\left(y_{t+1} - g_{t+1}\right) - \sigma^{-1}\left[r_t - E_t\pi_{t+1}\right] \tag{26}$$

It is convenient to write (26) in terms of the output gap, $x_t \equiv y_t - y_t^n$, where y_t^n is the "natural rate of output", the level of output that would be obtained in a full information rational expectations equilibrium. The resulting expression is

$$x_t = E_t x_{t+1} - \sigma^{-1}\left[r_t - E_t\pi_{t+1} - r_t^n\right] \tag{27}$$

where $r_t^n \equiv \sigma E_t\left[\left(y_{t+1}^n - g_{t+1}\right) - \left(y_t^n - g_t\right)\right]$ is the "natural rate of interest" (see Woodford (2003b)). It will turn out that r_t^n is a sufficient summary measure of all exogenous shocks in our model. As such, instead of specifying stochastic processes for the more fundamental shocks, we specify a process for r_t^n directly. In particular, r_t^n is assumed to follow a Markov process given by

$$r_t^n = \rho r_{t-1}^n + \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \tag{28}$$

Finally, the first-order condition for optimal labour supply is found by equating the marginal rate of substitution of consumption for leisure with the real wage

$$\frac{W_t(i)}{P_t} = \frac{v_h(H_t(i))}{\Lambda_t} \tag{29}$$

4.2 Firms

Consider first the optimal pricing decisions of firms, taking as given each firm's information set. Each firm i faces a Cobb-Douglas production technology with constant returns to scale

$$Y_t(i) = K_t(i)^\zeta (A_t H_t(i))^{1-\zeta} \tag{30}$$

where $K_t(i)$ is the capital input of firm i, A_t denotes a labour-augmenting technology shock and $0 < \zeta < 1$. For simplicity, we assume that the level of the capital stock is fixed and equal across firms (ie $K_t(i) = \bar{K}$). This assumption means that the demand for each good has the same form as (22), namely

$$Y_t(i) = \left[\frac{P_t(i)}{P_t}\right]^{-\epsilon} Y_t \tag{31}$$

The pricing decision by the firm is a static optimisation problem, where the first-order condition is given by

$$E_t^i\left[\frac{\partial \Pi_t(i)}{\partial P_t(i)}\right] = E_t^i\left[(1-\epsilon)\frac{Y_t(i)}{P_t} + \epsilon\frac{Y_t(i)}{P_t(i)}\frac{MC_t(i)}{P_t}\right] = 0 \tag{32}$$

where $\Pi_t(i)$ is firm i's real profit function and $MC_t(i)$ is its nominal marginal cost of producing an extra unit of output. Firms' conditional expectations operator at date t is given by

$$E_t^i(\cdot) \equiv E\left(\cdot|\mathcal{I}_t^i\right)$$

where $\mathcal{I}_t^i$ is the information set of firm i (see below).

Rearranging (32) yields

$$E_t^i\left[\frac{P_t(i)}{P_t} - \frac{\epsilon}{\epsilon-1}\frac{MC_t(i)}{P_t}\right] = 0 \tag{33}$$

Thus, the firm chooses its price such that its expected relative price is a constant markup over expected real marginal cost. In a situation of complete common knowledge, equation (33) reduces to the familiar condition that firms set their price equal to a fixed mark-up over marginal cost.

A log-linear approximation of (33) around $P_t(i)/P_t = 1$ and $S_t(i) \equiv MC_t(i)/P_t = (\epsilon-1)/\epsilon$ gives

$$E_t^i\left[\hat{p}_t(i) - s_t(i)\right] = 0 \tag{34}$$

where $\hat{p}_t(i) \equiv \log(P_t(i)/P_t)$.

Since real marginal cost is equal to the ratio of the real wage to the marginal product of labour, and in equilibrium the real wage must also equal the marginal rate of substitution, as given in (29), a log-linear approximation of real marginal cost can be expressed as

$$s_t(i) = \omega y_t(i) - (\nu + 1)a_t - \lambda_t \tag{35}$$

where $\nu \equiv v_{hh}(\bar{H})\bar{H}/v_h(\bar{H})$ is the inverse of the Frisch elasticity of labour supply and $\omega \equiv \left(\frac{\nu+\zeta}{1-\zeta}\right)$. Substituting (25) into (35) and rearranging gives

$$s_t(i) = (\omega + \sigma)(y_t - y_t^n) - \omega\epsilon\hat{p}_t(i)$$

where y_t^n, defined above as the natural rate of output, is given by

$$y_t^n \equiv \frac{1}{(\omega + \sigma)}\left[(\nu + 1)a_t + \sigma g_t\right] \tag{36}$$

We can now substitute the expression for marginal cost, given by (35), into the first-order condition for pricing, (34), to yield

$$p_t(i) = E_t^i p_t + \xi E_t^i x_t \tag{37}$$

where $\xi \equiv (\omega + \sigma)/(1 + \omega\epsilon)$. This equation is analogous to (1). Averaging (37) across firms gives

$$p_t = \bar{E}_t p_t + \xi \bar{E}_t x_t \tag{38}$$

where the average expectations operator, $\bar{E}_t(\bullet) \equiv \int_0^1 E_t^i(\bullet)\, di$, is the average expectation across firms.

Next, consider the information sets of firms. The underlying sources of aggregate disturbances are the demand shock g_t and the productivity shock a_t, which enter the model through the natural rate of interest r_t^n. To simplify matters, we assume that each firm observes one private and one public signal of r_t^n. Specifically, firm i's information set is given by

$$\mathcal{I}_t^i \equiv \{r_s^n(i), r_s^n(P)\}_{s=0}^t$$

where $r_t^n(i)$ and $r_t^n(P)$ are the private and public signals, respectively, of r_t^n. The conditional distribution of each signal, given r_t^n, is assumed to be normal with mean r_t^n and constant variance; namely,

$$r_t^n(i) = r_t^n + v_t(i), \qquad v_t(i) \overset{iid}{\sim} N(0, \sigma_v^2) \tag{39}$$

$$r_t^n(P) = r_t^n + \eta_t, \qquad \eta_t \overset{iid}{\sim} N(0, \sigma_\eta^2) \tag{40}$$

The innovations in (28) and (39)-(40) are assumed to be independent of each other at all leads and lags.

Other plausible assumptions on firms' information sets could also be incorporated into our framework. For example, one alternative approach would be to have firms obtain signals of endogenous variables directly, instead of the underlying fundamental shocks. For instance, firm i might observe a private signal of the price level such as $p_t^S(i) = p_t + e_t^p(i)$. We could also allow firms to observe all of the variables involved in their own production activities, such as their own output, hours hired and wages paid. In the current setup, if firms can observe their own output and hours employed when making pricing decisions, then they can infer without error the value of the technology shock A_t (or equivalently, a_t) from the production function (30). However, firms would still not be able to infer the exact value of g_t, and hence r_t^n.

4.3 Monetary policy

A large literature has developed recently examining the properties of different monetary policies. One approach taken has been to solve for optimal policy, where the central bank maximises a measure of expected discounted utility of the representative agent (see, eg, Rotemberg and Woodford (1997)).[4] An alternative approach is to specify the conduct of policy directly in terms of a (fixed) instrument rule. The type of instrument rule typically studied is an interest rate reaction function due to the fact that most central banks conduct monetary policy in practice by setting a target for a short-term nominal interest rate. Yet another approach, and the one followed in this paper, is to specify a targeting rule for the central bank. A targeting rule is a relation, analogous to a first-order condition, to be satisifed between some combination of the endogenous and exogenous variables in the model. Svensson (2003b) and Svensson and Woodford (2003) provide a general characterisation of targeting rules and describe their

[4] In other work (Amato and Shin (2003b)), we consider optimal monetary policy in a model similar to the one presented here.

merits.[5]

One advantage of employing a targeting rule is that it provides a transparent description of what monetary policy aims to achieve. In this paper, we consider targeting rules of the form

$$p_t + \lambda x_t = \delta r_t^n \tag{41}$$

Targeting rules expressed in terms of the price level, similar to (41), have been shown to have desirable welfare properties in sticky-price models. For instance, Svensson (1999) and Vestin (1999), among others, have demonstrated that when the central bank is unable to commit to its future actions, a price-level targeting rule performs better than an inflation-targeting rule even if society's welfare directly depends upon inflation but not the price level.

It should be noted, however, that (41) does not tell the central bank how to set the level of the short-term nominal rate on a period-by-period basis. This would require finding an instrument rule that is consistent with obtaining the relationship (41) in equilibrium subject to the behavioural equations (27) and (38). In fact, for a given model describing the behaviour of the private sector, there may be several interest rate rules consistent with the targeting rule (41). As an example, in the next section we will illustrate that an instrument rule of a common form can implement (41) in an equilibrium.

One important additional assumption we make is that the central bank has the same information set as households.[6] This means that policymakers observe, among other things, the current price level and output without error. The reason for assuming that the central bank observes the state perfectly is, once again, to keep our focus on the impact of differential information on firms' pricing behaviour and its macroeconomic consequences.

[5] Additional assumptions may also be required to characterise policy depending upon which approach is taken. For example, there are different notions of optimality that are linked to the treatment of the time-consistency problem (see, eg, Giannoni and Woodford (2002)).

[6] Recall that households' information sets are maximally specific with regard to all random variables realised to date.

5 General equilibrium

The complete model is given by the behavioural equations (27) and (38); the central bank's targeting rule (41); the process for the natural rate of interest (28); and the processes for the signals (39)-(40). In this section, we set up the model in state-space form, solve for the stochastic process followed by the state and then determine the equilibrium of the price level, output gap and the interest rate. In the next section, we illustrate some of the properties of the model.

The first step in solving the model is to describe the state space and determine the stochastic process followed by the state. In the present model, the state, denoted by X_t, is given by

$$X_t \equiv \begin{bmatrix} \theta_t \\ \psi_t \end{bmatrix} \tag{42}$$

where θ_t is a vector of exogenous variables and ψ_t is defined as

$$\psi_t \equiv \sum_{k=1}^{\infty} \xi_\lambda \left(1 - \xi_\lambda\right)^{k-1} \bar{E}_t^k \left(\theta_t\right) \tag{43}$$

where $\xi_\lambda \equiv \xi/\lambda$ and $\bar{E}_t^k(\bullet)$ is the k-th order average expectations operator. The exogenous state variables are

$$\theta_t \equiv \begin{bmatrix} r_t^n \\ \eta_t \end{bmatrix}$$

which follows a Markov process given by

$$\theta_t = B\theta_{t-1} + bu_t \tag{44}$$

where

$$B \equiv \begin{bmatrix} \rho & 0 \\ 0 & 0 \end{bmatrix}, b \equiv I_2$$

$$u_t \equiv \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}, \qquad u_t \overset{iid}{\sim} N(0, \Omega_u)$$

$$\Omega_u \equiv \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix}$$

I_n and 0_n denote the $n \times n$ identity and null matrices, respectively.[7]

Each firm observes the vector of variables

$$y_t^{sig}(i) \equiv \begin{bmatrix} r_t^n(i) \\ r_t^n(P) \end{bmatrix}$$

In terms of X_t, $y_t^{sig}(i)$ can be expressed as

$$y_t^{sig}(i) = ZX_t + zv_t(i) \tag{45}$$

where

$$Z \equiv \begin{bmatrix} Z_1 & 0_2 \end{bmatrix}, Z_1 \equiv \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, z \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and the process for $v_t(i)$ is given in (39).

Lemma 5 *Given equations (44) and (45), the state X_t, defined in (42), follows the Markov process given by*

$$X_t = MX_{t-1} + mu_t \tag{46}$$

where

$$M \equiv \begin{bmatrix} B & 0_2 \\ G & H \end{bmatrix}, m \equiv \begin{bmatrix} b \\ h \end{bmatrix},$$

and the matrices G, H and h are given in equations (60), (64) and (62), respectively.

Proof. See Appendix A.2. ■

It is now straightforward to find the equilibrium processes of p_t and x_t as a function of the state X_t. Substituting (41) into (38) yields

$$p_t = (1 - \xi_\lambda) \bar{E}_t p_t + \delta \xi_\lambda \bar{E}_t r_t^n$$

[7] We have started to recycle notation here. However, in the following, the appropriate reference object should be clear.

Solving this expression by repeated substitution (as in Section 2), we get

$$\begin{aligned} p_t &= \delta \sum_{k=1}^{\infty} \xi_\lambda \left(1-\xi_\lambda\right)^{k-1} \bar{E}_t^k r_t^n &(47)\\ &= \delta e_3' X_t &(48)\end{aligned}$$

where e_i is the 4x1 unit vector with 1 in the i-th position. Substituting (48) into (41) implies that

$$x_t = \frac{\delta}{\lambda}(e_1 - e_3)' X_t \quad (49)$$

We can also determine the process followed by r_t as a function of the state X_t. Using the solutions for p_t and x_t in (48)-(49) and the stochastic process for X_t given by (46), the solution for r_t can be found by rearranging (27) and making the appropriate substitutions. This gives

$$\begin{aligned} r_t &= \sigma E_t \left(x_{t+1} - x_t\right) + E_t \left(p_{t+1} - p_t\right) + r_t^n \\ &= \left([M' - I] \left[\frac{\sigma\delta}{\lambda} \left(e_1 - e_3\right) + \delta e_3 \right] + e_1 \right)' X_t &(50)\end{aligned}$$

While equation (50) describes how the interest rate should respond to the state X_t, it is not necessarily a description of how policy should be *implemented.* In other words, (50) does not have to be the instrument rule followed by the central bank in determining the appropriate level of its policy rate target on a period-by-period basis. In fact, a policy of setting interest rates directly according to (50) may have some undesirable consequences. For instance, in the special case of full information in models of the type considered here, it is well known that rules that specify the interest rate to be a function solely of exogenous variables lead to indeterminancy of equilibrium (eg Woodford (2003b)).

For now, it is informative to show that the targeting rule (41), and the resulting equilibrium characterised by (48)-(50), can be implemented by a simple instrument rule. Taylor's (1993) rule (and its generalisations) is a well-known example. Here we show that a rule where the short-term nominal interest rate responds only to the price level and output gap is consistent with the equilibrium relation (41). Specifically, we consider an instrument rule of the form

$$r_t = \alpha_p p_t + \alpha_x x_t \quad (51)$$

The main difference between (51) and the Taylor rule is the inclusion of the price level instead of the inflation rate.[8]

Lemma 6 *The targeting rule (41) and the resulting equilibrium processes for* p_t*,* x_t *and* r_t *given in (48)-(50) can be implemented by an instrument rule of the form (51).*

Proof. See Appendix A.3. ■

6 Model properties

Here we examine several features of the model presented above. Before proceeding, we must choose values for the parameters. These are given in Table 1. Our choices for the preference and technology parameters fall within the range of values typically used in the literature. The parameters governing the process of r_t^n can be rationalised on the basis of estimates provided in Rotemberg and Woodford (1997) (see Woodford (1999) for further discussion). The variances of the noise terms in the signals have been chosen somewhat arbitrarily because there is not much evidence to draw upon in these cases. In the baseline, as well as the alternatives considered below, the variance of the noise terms (0.2% each) has been chosen to be much smaller than the variance of the fundamental r_t^n (set equal to 1%). Introspection would suggest that measurement and filtering errors are typically smaller in magnitude than variability in the fundamentals of the economy; whether this is true in actual economies, however, remains to be determined. Finally, regarding monetary policy, we set both λ and δ equal to one. This implies that the central bank aims for the nominal output gap, defined as $p_t + x_t$, to fluctuate one-for-one with the natural rate of interest. This is similar to nominal GDP targeting, except that account is taken of fluctuations in the natural rate of output.

Before proceeding, however, it is worth noting that perfect stabilisation of the price level and the output gap is actually feasible in the current version of

[8] In addition, the coefficients α_p and α_x will be determined as a function of the model's structural parameters and the parameters of the targeting rule (41).

our model. This can be seen by setting $\delta = 0$ in the targeting rule (41), and hence the solutions for p_t, x_t and r_t in (48)-(50). If we also assume that the natural rate of output is the efficient level of output (ie resulting from a subsidy to firms to eliminate the distortion due to monopolistic competition and thereby raise steady-state output), then perfect stabilisation would correspond to the first-best equilibrium. While this is an interesting property of the model, we view it as not being very relevant for the purposes of understanding how monetary policy can work in actual economies. The reason is that complete stabilisation can only be achieved under our assumption that the central bank perfectly observes current and past values of the state. In the more realistic setting where the central bank also obtains only noisy signals of fundamentals, this equilibrium is no longer feasible. The virtue of the current analysis is its relative simplicity in demonstrating the basic properties of a differential information economy.

6.1 Changing weights on higher-order beliefs

Recall that one of the key parameters of the model is ξ, which, being the numerator of ξ_λ, determines in part the relative weight attached to higher-order expectations in the pricing relation (47). Among other things, ξ depends inversely upon the elasticity of substitution, ϵ. Thus, an increase in ϵ, which increases the coordination motive among firms and produces a smaller steady-state markup, gives a more prominent role to higher-order beliefs by lowering ξ.[9] One feature of the macro model we wish to highlight is the implication of changing ξ on the sample paths of the output gap and the price level. We do this by altering the value of ϵ, since it enters the model only through ξ.

The results of one such experiment are shown in Figure 1. Each panel of the figure plots one sample realisation (time series) of the price level against the output gap using the same randomly drawn sample of shocks. The cases in the panels are distinguished by their treatment of ϵ and the relative precision of the public signal, defined as $1/\sigma_\eta^2$. The data in the left-hand side panels have

[9] As already noted by Woodford (2003a), such changes are more critical in the current setting than in standard sticky-price models, where an increase in competition lowers the elasticity of inflation to the output gap, but no more.

been generated under a steady-state markup of 25%, whereas the right-hand side panels correspond to a markup of 5%. In addition, the top panels report cases with high-precision public signals ($1/\sigma_\eta^2 = 10\%$), whereas the lower panels are based on low-precision public signals ($1/\sigma_\eta^2 = 5\%$). The plots suggest that, conditional on the output gap, an increase in competition (lower markup) or a decline in the precision of the public signal spreads out prices. This is most evident in the lower right panel, where prices depend relatively more on higher-order expectations (due to lower ξ), which in turn are adversely affected by noisier information (less precise public signals).

These scatter plots intimate the potential degradation of the information value of price as a signal of the output gap. For economies that have relatively noisy public signals and a high degree of competition, prices convey poor quality information about the underlying output gap.

6.2 Impulse responses of higher-order beliefs

One way to illustrate the dynamic impact of differential information is to plot the impulse responses of higher-order beliefs of the fundamentals. In particular, recalling that the aggregate price level is given by the infinite weighted sum of k-th order average expectations of r_t^n, we wish to examine the evolution of random variables such as $\bar{E}_t^k(\theta_t)$. To compute the impulse responses of $\bar{E}_t^k(\theta_t)$ to innovations in θ_t, we first must determine its law of motion. Define

$$\Psi_t^{(k)} \equiv \begin{bmatrix} \bar{E}_t^k(\theta_t) \\ \bar{E}_t^{k-1}(\theta_t) \\ \vdots \\ \bar{E}_t(\theta_t) \\ \theta_t \end{bmatrix} \tag{52}$$

The following lemma gives the stochastic process followed by $\Psi_t^{(k)}$.

Lemma 7 *The (k+1)-dimensional vector of sequential higher-order beliefs $\Psi_t^{(k)}$, defined in (52), follows the Markov process given by*

$$\Psi_t^{(k)} = B_{(k)}\Psi_{t-1}^{(k)} + b_{(k)}u_t$$

where $B_{(k)}$ and $b_{(k)}$ are given in (90) and (91), respectively.

Proof. See Appendix A.4. ■

Figure 2 shows the responses of the first eight orders of average expectations of r_t^n with respect to a cumulative one-percent deviation in r_t^n from zero (recall that all variables are expressed as deviations from steady state). The solid line shows the path followed by r_t^n itself. The other lines show the responses of the first-order (solid with circles) through eighth-order (solid with asterisks) average expectations. It is evident that higher-order expectations respond more sluggishly to the shock, with virtually no initial response in expectations as low as order four (solid with square). The discrepancy between $\bar{E}_t^k(r_t^n)$ and r_t^n is also monotonically increasing in k in each period after the shock.

Similar to Figure 2, Figure 3 shows the responses of $\bar{E}_t^k(r_t^n)$ to an innovation in the noise of the public signal (ie η_t). For clarity, expectations for $k = 1, 2, 4, 8$ are only plotted. In the period of the shock, only the response of the first-order average expectation is much different than zero. Thus, even though a larger weight is given to the public signal as the order of expectation increases, this is more than outweighed by the dampening effect of the *presence* of public information on higher-order expectations. In addition, notice that there is a delay in the peak response in expectations of order higher than one, with the delay increasing in k.

Lastly, Figure 4 compares the responses of higher-order expectations to a shock in r_t^n in the current model with public information (solid lines with symbols, as in Figure 2) to the responses in an analogous model without public signals (dashed lines with symbols). Again, the plain solid line is the path of r_t^n. For low orders ($k = 1$ or $k = 2$), the dynamic response in expectations in the presence of public signals is always closer to r_t^n than in the model without public signals. Note that in this experiment the public signal always equals the true value of r_t^n (ie the noise term in the public signal is assumed to be zero at all times). Thus, this figure demonstrates the beneficial effect of public information in aligning low-order average expectations closer to the fundamental. However, the relative initial reponse of expectations of a higher order ($k = 4$ or $k = 8$) is the opposite. The larger weight agents place on the public signal in these cases is not sufficient to counterbalance the relatively more sluggish adjustment of expectations overall in the presence of public information. Nonetheless, the response of $\bar{E}_t^k(r_t^n)$ converges to r_t^n more quickly when there is public information (see also Hellwig (2002)). This effect is largely due to the higher persistence imparted to r_t^n compared to the noise in the public signal, η_t.

6.3 Volatility and the quality of public information

We next demonstrate that more precise public information does not necessarily lead to lower volatility among endogenous variables. This result is evident in Figure 5. This figure plots values of the variances of the endogenous variables as a function of the precision of the public signal. In each panel, the solid line is the case when firms' private signals have relatively high precision ($1/\sigma_v^2 = 10\%$), whereas the dashed line is the case when these signals have relatively low precision ($1/\sigma_v^2 = 2\%$). The figure demonstrates that increases in the precision of the public signal can result in a higher variance of the price level (and inflation). In particular, the lowest values for these variances are achieved under the *least* precise public signal. The fact that similar effects are evident in both cases (solid and dashed lines) suggests that these results are robust across a wide range of values for the precision of the private signal.

Figure 5 illustrates one key effect of public information. From the results in Section 3, recall that more precise public signals get a higher weight in both individuals' and average k-fold expectations. A higher weight on a common (public) signal necessarily means that individuals' expectations are distributed more closely together around the public signal. However, this can lead to greater volatility in the aggregate if the public signal is not very precise relative to private information. Since higher-order beliefs play a direct role only in firms' pricing decisions, it is perhaps not surprising that these effects largely pertain to price level and, by extension, inflation outcomes; note that the change in the variance of the output gap and interest rate is small, both relatively and absolutely. These results are reflective of the finding by Morris and Shin (2002), extended here to a dynamic macroeconomic setting, that more precise public information does not necessarily lead to better welfare outcomes. Importantly, this is not predicated on inefficiencies that arise due to poor information available to the central bank. On the contrary, the central bank operates with *full* information on the state of the economy in our model.

7 Conclusions

An economy with diverse private information has features that are not always well captured in representative individual models where all agents share the same information. The most distinctive of these features is the relatively greater impact of common, shared information at the expense of private information. The source of the greater impact of public information lies in the strategic complementarity of the price setting behaviour of firms, and 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 contribution has been to demonstrate how the theoretical results can be embedded in a standard macroeconomic model that is rich enough to engage in questions of significance for policy purposes. Moreover, our discussion of the conceptual background in Section 3 has been motivated by the need to unravel the main mechanisms at work. By developing the argument by means of a series of simple examples, our intention has been to convey the main intuitions, and so show that the results do not rely in sensitive ways on specific functional forms or distributional assumptions.

In illustrating the basic effects of the presence of both public and private information in a complete macroeconomic model, we have made several simplifying assumptions, such as the fact that consumers and the central bank are fully informed. At the cost of some additional complexity, we can extend our model to contexts where agents observe noisy signals of the endogenous variables directly and the central bank has less than perfect information as well (see Amato and Shin (2003a)). Nevertheless, the results in this paper reveal that the impact of public information in differential information economies is large, and shifts in the precision of public signals can have significant effects on observable variables that enter into calculations of welfare.

A Proofs

A.1 Alternative proof of theorem 4

An alternative proof of theorem 4 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 generalises. Suppose there are N agents. Since p_{ij} is the average conditional probability of state j at state i, we have

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

where $p_k\left(j|i\right)$ is the k-th agent'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, $S\left(i,j\right) = S\left(j,i\right)$. Denote by $P_k\left(i\right)$ the ex ante probability of the cell of individual k's partition that contains state i. Then,

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

where $P\left(i,j\right)$ 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

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

so that $P\left(i,j\right) = P\left(j,i\right)$. 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 diagonalisable 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

$$B = \begin{bmatrix} \frac{p_1}{P(1,1)} & \frac{\sqrt{p_1p_2}}{P(1,2)} & \cdots & \frac{\sqrt{p_1p_n}}{P(1,n)} \\ \frac{\sqrt{p_2p_1}}{P(2,1)} & \frac{p_2}{P(2,2)} & & \frac{\sqrt{p_2p_n}}{P(2,n)} \\ \vdots & & \ddots & \\ \frac{\sqrt{p_np_1}}{P(n,1)} & \frac{\sqrt{p_np_2}}{P(n,2)} & & \frac{p_n}{P(n,n)} \end{bmatrix}$$

It can be verified that $B = D^{-1}AD$. Since A is a symmetric matrix, it is diagonalisable 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 diagonalisable, 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 (7), we can write the matrix of eigenvectors C as follows.

$$C = \left[\begin{array}{c|c|c|c|c} 1 & \vdots & \vdots & & \vdots \\ 1 & c_2 & c_3 & & c_n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & & & & \end{array}\right]$$

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

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 8 *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,

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

where $E(.)$ is the expectations operator with respect to public information only (ie 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

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

since $\lambda_j < 1$ for $j \geq 2$. Thus, theorem 4 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.

A.2 Proof of lemma 5

Recall that X_t is defined as

$$X_t \equiv \begin{bmatrix} \theta_t \\ \psi_t \end{bmatrix} \tag{53}$$

where θ_t is a vector of variables that are exogenous with respect to p_t, y_t and r_t, and ψ_t is defined as

$$\psi_t \equiv \sum_{k=1}^{\infty} \xi_\lambda (1 - \xi_\lambda)^{k-1} \bar{E}_t^k (\theta_t) \tag{54}$$

θ_t is governed by the process

$$\theta_t = B\theta_{t-1} + bu_t \tag{55}$$

for known matrices B and b and where $u_t \sim N(0, \Omega_u)$ is a vector of *iid* random variables.

The state-space model is completed by specifying the observation equation. Let $y_t^{sig}(i)$ be the $n_y x1$ vector of variables observed by firm i at date t. The observation equation is

$$y_t^{sig}(i) = ZX_t + zv_t(i)$$

for known matrices $Z \equiv [Z_1 \quad 0_{n_y xn}]$ and z, where 0_{kxl} is the null matrix of dimension kxl, and $v_t(i) \sim N(0, \sigma_v^2)$ is independently and identically distributed across time and firms. These assumptions, and the law of large numbers, imply that $\int_0^1 v_t(i)di = 0$.

Our method follows the steps of, but also generalises, the proof in Woodford (2003a). For now assume (to be confirmed later) that the state, X_t, is given by the process

$$X_t = MX_{t-1} + mu_t \tag{56}$$

where

$$M \equiv \begin{bmatrix} B & 0_2 \\ G & H \end{bmatrix}, m \equiv \begin{bmatrix} b \\ h \end{bmatrix}$$

and the matrices G, H and h are yet to be determined. When there is no ambiguity, the subscript will be omitted from I_n and 0_n.

Now consider the firm's problem of estimating the state, X_t, using the Kalman filter. Given the assumptions made so far, the Kalman filter produces minimum mean squared error estimates of the state for the log-linearised version of the model. Assume that a time-invariant filter exists that is also independent of i, with the Kalman gain denoted by K. Let $X_{t|s}(i) \equiv E_s^i X_t$. Combining the prediction and updating equations from the Kalman filter for firm i gives

$$X_{t|t}(i) = MX_{t-1|t-1}(i) + K\left(y_t^{sig}(i) - ZMX_{t-1|t-1}(i)\right) \tag{57}$$

Averaging across i and rearranging gives

$$\begin{aligned} X_{t|t} &= (I - KZ)\, MX_{t-1|t-1} + KZX_t \\ &= (I - KZ)\, MX_{t-1|t-1} + KZMX_{t-1} + KZmu_t \end{aligned}$$

Defining $\Xi \equiv [\xi_\lambda I \quad (1-\xi_\lambda)I]$ and $\hat{K} \equiv \Xi K$, first notice that $\psi_t = \Xi X_{t|t}$, and thus $(1-\xi_\lambda)\psi_{t-1|t-1} = \psi_{t-1} - \xi_\lambda \theta_{t-1|t-1}$. This implies

$$\psi_t = (\Xi - \hat{K}Z)MX_{t-1|t-1} + \hat{K}ZMX_{t-1} + \hat{K}Zmu_t \tag{58}$$

and

$$X_{t-1|t-1} = \varphi_1 \psi_{t-1} + \varphi_2 \theta_{t-1|t-1} \tag{59}$$

where $\varphi_1 \equiv [0 \quad \frac{1}{1-\xi_\lambda}I]'$ and $\varphi_2 \equiv [I \quad -\frac{\xi_\lambda}{1-\xi_\lambda}I]'$. Substituting (59) into (58) and expanding gives

$$\begin{aligned}\psi_t &= \hat{K}Z_1B\theta_{t-1} + \frac{1}{(1-\xi_\lambda)}\hat{\Xi}_2\psi_{t-1} \\ &+ \left[\hat{\Xi}_1 - \frac{\xi_\lambda}{(1-\xi_\lambda)}\hat{\Xi}_2\right]\theta_{t-1|t-1} + \hat{K}Z_1bu_t\end{aligned}$$

where $\hat{\Xi}_1 \equiv \left(\xi_\lambda I - \hat{K}Z_1\right)B + (1-\xi_\lambda)\,G$ and $\hat{\Xi}_2 \equiv (1-\xi_\lambda)\,H$.

If X_t is governed by (56), then it must be the case that

$$G = \hat{K}Z_1B \tag{60}$$

$$H = \frac{1}{1-\xi_\lambda}\hat{\Xi}_2 \tag{61}$$

$$h = \hat{K}Z_1b \tag{62}$$

$$\hat{\Xi}_1 = \frac{\xi_\lambda}{1-\xi_\lambda}\hat{\Xi}_2 \tag{63}$$

The solutions for G and h are given directly by (60) and (62), respectively. By the definition of $\hat{\Xi}_2$, it can be seen that (61) is satisfied. Finally, the solution for H is obtained by substituting the result for G into (63):

$$H = \left(I - \hat{K}Z_1\right)B \tag{64}$$

The last step is to determine the value of K, or equivalently, $\hat{K}$. Under the above assumptions, we have (see Harvey (1989))

$$\hat{K} = \Xi\Sigma Z'F^{-1} \tag{65}$$

where

$$\Sigma \equiv var\left(X_t - X_{t|t-1}(i)\right) = MVM' + m\Omega_u m' \tag{66}$$

$$V \equiv var\left(X_t - X_{t|t}(i)\right) = \Sigma - \Sigma Z'F^{-1}Z\Sigma \tag{67}$$

$$F \equiv var\left(y_t^{sig}(i) - ZX_{t|t-1}(i)\right) = Z\Sigma Z' + \sigma_v^2 zz' \tag{68}$$

Substituting (67)-(68) into (66), we obtain a Riccati equation:

$$\Sigma = M\left(\Sigma - \Sigma Z'\left(Z\Sigma Z' + \sigma_v^2 zz'\right)^{-1} Z\Sigma\right) M' + m\Omega_u m' \tag{69}$$

It is possible to solve (69) explicitly for Σ. In fact, if we partition Σ as

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{21} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

it can be seen from (65) and the definition of Z that we need only determine Σ_{11} and Σ_{21} to obtain the solution for $\hat{K}$. As it turns out, Σ_{11} and Σ_{21} can be solved for recursively without having to solve for Σ_{22} as well.

We begin by isolating the upper-left block of equations in (66):

$$\Sigma_{11} = BV_{11}B' + \Omega_u \tag{70}$$

where

$$V_{11} \equiv \Sigma_{11} - \Sigma_{11}Z_1'\left(Z_1\Sigma_{11}Z_1' + \sigma_v^2 zz'\right)^{-1} Z_1\Sigma_{11}$$

Notice that (70) is a set of three equations that involves only the elements of Σ_{11}. Let σ_{ij} denote the (i,j)-th element of Σ_{11}. Thus, by the definition of B, we have

$$\begin{bmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \rho^2 \begin{bmatrix} v_{11} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \tag{71}$$

where v_{11} is the $(1,1)$ element of the matrix V_{11}. It is immediate from (71) that

$$\sigma_{21} = 0$$

$$\sigma_{22} = \sigma_\eta^2$$

$$\begin{aligned} \sigma_{11} &= \rho^2 v_{11} + \sigma_\varepsilon^2 \\ &= \rho^2\left(\sigma_{11} - \frac{\left(\sigma_\eta^2 + \sigma_v^2\right)\sigma_{11}^2}{\left(\sigma_\eta^2 + \sigma_v^2\right)\sigma_{11} + \sigma_\eta^2\sigma_v^2}\right) + \sigma_\varepsilon^2 \end{aligned} \tag{72}$$

Rewriting (72), we get

$$\left(\sigma_\eta^2 + \sigma_v^2\right)\sigma_{11}^2 - \left(\sigma_\varepsilon^2\left[\sigma_\eta^2 + \sigma_v^2\right] - \left[1 - \rho^2\right]\sigma_\eta^2\sigma_v^2\right)\sigma_{11} - \sigma_\varepsilon^2\sigma_\eta^2\sigma_v^2 = 0$$

which is a quadratic equation in σ_{11} that has two real roots, one positive and one negative. Since σ_{11} is a variance, its solution must be the positive root, which is given by

$$\sigma_{11} = \frac{\sigma_\varepsilon^2}{2} - \frac{1-\rho^2}{2}\frac{\sigma_\eta^2\sigma_v^2}{\left[\sigma_\eta^2+\sigma_v^2\right]} + \sqrt{\left(\frac{\sigma_\varepsilon^2}{2} - \frac{1-\rho^2}{2}\frac{\sigma_\eta^2\sigma_v^2}{\left[\sigma_\eta^2+\sigma_v^2\right]}\right)^2 + \sigma_\varepsilon^2\frac{\sigma_\eta^2\sigma_v^2}{\left[\sigma_\eta^2+\sigma_v^2\right]}}$$

The second step is to solve for Σ_{21}. From the lower-left block of equations in (66), we see that Σ_{21} depends only upon the elements of Σ_{11} (and other known parameters):

$$\Sigma_{21} = \left(G\Sigma_{11} + H\Sigma_{21}\right)\left(I - Z_1'F^{-1}Z_1\Sigma_{11}\right)B' + h\Omega_u \tag{73}$$

Let s_{ij} denote the (i,j)-th element of Σ_{21}. Again, by the definition of B, we have

$$\begin{aligned}\begin{bmatrix} s_{12} \\ s_{22}\end{bmatrix} &= h\Omega_u\bar{e}_2 \\ &= \sigma_\eta^2\begin{bmatrix}\hat{K}_{12} \\ \hat{K}_{22}\end{bmatrix}\end{aligned}$$

where $\hat{K}_{ij}$ is the (i,j)-th element of $\hat{K}$. Since F does not depend on Σ_{21}, it is evident from (65) that there is a linear relationship between $\hat{K}$ and Σ_{21}. In particular, $\hat{K}_{12}$ is a linear function of only s_{11} and s_{12}; similarly, $\hat{K}_{22}$ is a linear function of only s_{21} and s_{22}. We can therefore obtain expressions for s_{12} and s_{22} as linear functions of s_{11} and s_{21}, respectively; namely,

$$\begin{bmatrix} s_{12} \\ s_{22}\end{bmatrix} = \frac{\xi_\lambda\sigma_\eta^2}{1-\kappa_2}\Sigma_{11}Z_1'F^{-1}\bar{e}_2 + \frac{\kappa_1}{1-\kappa_2}\begin{bmatrix} s_{11} \\ s_{21}\end{bmatrix} \tag{74}$$

where

$$\kappa_1 \equiv \left(1-\xi_\lambda\right)\sigma_\eta^2\bar{e}'F^{-1}\bar{e}_2, \kappa_2 \equiv \left(1-\xi_\lambda\right)\sigma_\eta^2\bar{e}_2'F^{-1}\bar{e}_2$$

$\bar{e} \equiv [1 \quad 1]'$ and $\bar{e}_i$ is the 2x1 unit vector with 1 in the i-th position.

It remains to solve for s_{11} and s_{21}. Expanding (73), it turns out that the upper-left equation involves only s_{11} and $\hat{k}_1 \equiv \hat{K}_{11} + \hat{K}_{12}$. Noting again (65), it can be seen that $\hat{k}_1$ is a linear function of s_{11} and s_{12}:

$$\hat{k}_1 = \xi_\lambda\bar{e}_1'\Sigma_{11}Z_1'F^{-1}\bar{e} + \left(1-\xi_\lambda\right)\bar{e}'F^{-1}\bar{e}s_{11} + \left(1-\xi_\lambda\right)\bar{e}'F^{-1}\bar{e}_2s_{12} \tag{75}$$

By (74), we can substitute out for s_{12} in (75) to obtain

$$\hat{k}_1 = \chi_1 + \chi_0 s_{11} \tag{76}$$

where

$$\chi_1 \equiv \xi_\lambda \bar{e}_1' \Sigma_{11} Z_1' F^{-1} \left[\bar{e} + \frac{\kappa_1}{1-\kappa_2} \bar{e}_2 \right], \chi_0 \equiv (1-\xi_\lambda) \bar{e}' F^{-1} \left[\bar{e} + \frac{\kappa_1}{1-\kappa_2} \bar{e}_2 \right]$$

Thus, the upper-left equation of (73) can be written as a quadratic equation in either $\hat{k}_1$ or s_{11}, which does not depend, in particular, upon s_{21}. In terms of $\hat{k}_1$, this equation is

$$\varpi_2 \hat{k}_1^2 + \varpi_1 \hat{k}_1 + \varpi_0 = 0$$

where

$$\begin{aligned}
\varpi_0 &\equiv \sigma_{11} \left(\frac{\chi_1}{\chi_0} \left[\bar{e}' F^{-1} \bar{e} + \frac{\kappa_1}{1-\kappa_2} \bar{e}' F^{-1} \bar{e}_2 \right] - \frac{\xi_\lambda \sigma_\eta^2}{1-\kappa_2} \left[\bar{e}' F^{-1} \bar{e}_2 \right] \left[\bar{e}_1' \Sigma_{11} Z_1' F^{-1} \bar{e}_2 \right] \right) \\
&\quad - \frac{\chi_1}{\chi_0} \left(1 - \frac{1}{\rho^2} \right) \\
\varpi_1 &\equiv -\frac{1}{\rho^2} \left(\frac{1}{\chi_0} - \sigma_\varepsilon^2 \right) + \sigma_{11} \left(1 + \frac{\xi_\lambda \sigma_\eta^2}{1-\kappa_2} \left[\bar{e}' F^{-1} \bar{e}_2 \right] \left[\bar{e}_1' \Sigma_{11} Z_1' F^{-1} \bar{e}_2 \right] \right) + \frac{1+\chi_1}{\chi_0} \\
&\quad - \left[\sigma_{11}^2 + \frac{\sigma_{11}}{\chi_0} (1+\chi_1) \right] \bar{e}' F^{-1} \bar{e} - \frac{1}{\chi_0} \frac{\kappa_1}{1-\kappa_2} \left[\sigma_{11} (1+\chi_1) \right] \bar{e}' F^{-1} \bar{e}_2 \\
\varpi_2 &\equiv \frac{1}{\chi_0} \left[\sigma_{11} \left(\bar{e}' F^{-1} \bar{e} + \frac{\kappa_1}{1-\kappa_2} \bar{e}' F^{-1} \bar{e}_2 \right) - 1 \right]
\end{aligned}$$

It is difficult to simplify the expressions for ϖ_0, ϖ_1 and ϖ_2 much further. The roots of $\hat{k}_1$ can be determined numerically for given values of the parameters. Given a solution for $\hat{k}_1$, we can then find the value of s_{11} using (76).

Under the range of values for the parameters in the simulations in Section 6, $\hat{k}_1$ has two real roots, one positive and one negative. The fact that $\hat{k}_1$ is a linear combination of Kalman gains does not, by itself, rule out either of these roots. However, a restriction can be placed upon the chosen root if we wish X_t to be stationary — which is desirable since we have assumed that θ_t is stationary. Recalling the solutions for M and m, we have

$$[r_t^n - \psi_{1t}] = \rho \left(1 - \hat{k}_1 \right) [r_t^n - \psi_{1t}] + \left(1 - \hat{k}_1 \right) \varepsilon_t + \hat{K}_{12} \eta_t$$

where ψ_{1t} is the first element of ψ_t. Since r_t^n itself is assumed to be stationary, $r_t^n - \psi_{1t}$ is stationary if and only if $\left|\rho\left(1-\hat{k}_1\right)\right| < 1$. If we assume that $0 < \rho < 1$, this condition simplifies to

$$1 - \frac{1}{\rho} < \hat{k}_1 < 1 + \frac{1}{\rho}$$

For the parameter values considered, only the positive root falls within this range, therefore, this is the one that is selected.

Finally, analogous to $\hat{k}_1$, $\hat{k}_2 \equiv \hat{K}_{21} + \hat{K}_{22}$ is a linear function of s_{21}:

$$\begin{aligned} \hat{k}_2 &= \xi_\lambda \bar{e}_2' \Sigma_{11} Z_1' F^{-1} \bar{e} + (1-\xi_\lambda) \bar{e}' F^{-1} \bar{e} s_{21} + (1-\xi_\lambda) \bar{e}' F^{-1} \bar{e}_2 s_{22} &(77)\\ &= \chi_2 + \chi_0 s_{21} &(78)\end{aligned}$$

where

$$\chi_2 \equiv \xi_\lambda \bar{e}_2' \Sigma_{11} Z_1' F^{-1} \left[\bar{e} + \frac{\kappa_1}{1-\kappa_2} \bar{e}_2\right]$$

Thus, the lower-left equation of (73) is *linear* in s_{21} as a function of s_{11}, s_{12} and other known parameters. The solution is

$$s_{21} = \frac{\vartheta \chi_2}{1 - \vartheta \chi_0}$$

where

$$\vartheta \equiv \sigma_\varepsilon^2 + \rho^2 \left([\sigma_{11} - s_{11}]\left[1 - \sigma_{11} \bar{e}' F^{-1} \bar{e}\right] + \sigma_{11} s_{12} \bar{e}' F^{-1} \bar{e}_2\right)$$

A.3 Proof of lemma 6

Substituting (51) into (27), we get

$$x_t = \mu_1 E_t x_{t+1} - \mu_1 \sigma^{-1} \left[(\alpha_p + 1) p_t - E_t p_{t+1} - r_t^n\right] \tag{79}$$

If we assume, for now, that (51) can implement the targeting rule (41), (48) can be used as an equilibrium solution for the price level in terms of the state X_t. Substituting for p_t in (79), solving forward, and computing expectations of X_t

from (46), we obtain

$$\begin{aligned} x_t &= \mu_1 E_t x_{t+1} - \mu_1 \sigma^{-1} \phi' X_t \\ &= -\mu_1 \sigma^{-1} \phi' \sum_{i=0}^{\infty} \mu_1^i E_t X_{t+i} \\ &= -\mu_1 \sigma^{-1} \phi' \sum_{i=0}^{\infty} (\mu_1 M)^i X_t \\ &= -\mu_1 \sigma^{-1} \phi' (I - \mu_1 M)^{-1} X_t \end{aligned} \tag{80}$$

where $0 < \mu_1 \equiv \left(\sigma^{-1} + 1\right)^{-1} < 1$ and

$$\phi \equiv \delta \left[(\alpha_p + 1) I - M' \right] e_3 - e_1 \tag{81}$$

and assuming that $N \equiv (I - \mu_1 M)^{-1}$ is nonsingular.

If the instrument rule (51) is to be consistent with the targeting rule (41), it must be the case that the equilibrium processes for x_t given in (49) and (80) are consistent with each other. This requires

$$\frac{\delta}{\lambda}(e_1 - e_3) = -\mu_1 \sigma^{-1} N' \phi \tag{82}$$

Thus, it remains to be shown whether (82) holds for some value of α_p. First, notice that because M is block lower diagonal, N is also block lower diagonal:

$$N = \begin{bmatrix} N_{11} & 0 \\ N_{21} & N_{22} \end{bmatrix}$$

Partition ϕ accordingly as

$$\phi = \begin{bmatrix} -\left(\delta G' + I\right) \bar{e}_1 \\ \left[(\alpha_p + 1) I - H' \right] \bar{e}_1 \end{bmatrix}$$

Expanding the right-hand side of (82), the first two equalities require

$$\frac{\delta}{\lambda} \bar{e}_1 = -\mu_1 \sigma^{-1} \left(N_{21}' \left[(\alpha_p + 1) I - H' \right] - N_{11}' \left[\delta G' + I \right] \right) \bar{e}_1 \tag{83}$$

whereas the last two require

$$\frac{\delta}{\lambda} \bar{e}_1 = \mu_1 \sigma^{-1} N_{22}' \left((\alpha_p + 1) I - H' \right) \bar{e}_1 \tag{84}$$

Equating (83) and (84), we have

$$N_{22}' \left((\alpha_p + 1) I - H' \right) \bar{e}_1 = -\left(N_{21}' \left[(\alpha_p + 1) I - H' \right] - N_{11}' \left[\delta G' + I \right] \right) \bar{e}_1 \tag{85}$$

which is a system of two equations in one unknown, α_p. Rearranging (85) gives

$$(\alpha_p + 1)\, C_1' \bar{e}_1 = C_2' \bar{e}_1 \tag{86}$$

where

$$C_1 \equiv N_{21} + N_{22}, \quad C_2 \equiv H\,[N_{21} + N_{22}] + [\delta G + I]\, N_{11}$$

By the definitions of B and M, N_{11} is diagonal. Noting (60) and (64), it can be seen that $G\bar{e}_2 = H\bar{e}_2 = 0_2$, which implies that N_{22} is lower diagonal. Taken together, these results imply that the two equalities in (86) are satisfied if $C_{1,11} \neq 0$, where C_{ij} is the *(i,j)*-th element of matrix C, since in (86) both sides of the second equality are zero and a solution for α_p can be obtained from the first equality and is given by:

$$\alpha_p = \frac{C_{2,11}}{C_{1,11}} - 1 \tag{87}$$

A.4 Proof of lemma 7

Recall that $\Psi_t^{(k)}$ is defined as

$$\Psi_t^{(k)} \equiv \begin{bmatrix} \bar{E}_t^k(\theta_t) \\ \bar{E}_t^{k-1}(\theta_t) \\ \vdots \\ \bar{E}_t(\theta_t) \\ \theta_t \end{bmatrix} \tag{88}$$

Proceeding in a similar way as in the proof of Lemma 5, we begin by conjecturing the form of a state-space model in terms of $\Psi_t^{(k)}$ and the observable vector $y_t^{sig}(i)$. We then determine the stochastic process of $\Psi_t^{(k)}$ by solving each firm's optimal filtering problem and averaging across firms. Accordingly, for now assume (to be confirmed later) that the state $\Psi_t^{(k)}$ follows the Markov process

$$\Psi_t^{(k)} = B_{(k)} \Psi_{t-1}^{(k)} + b_{(k)} u_t \tag{89}$$

where

$$B_{(k)} \equiv \begin{bmatrix} B_{k,k} & B_{k,k-1} & \cdots & B_{k,1} & B_{k,0} \\ 0_n & B_{k-1,k-1} & \cdots & B_{k-1,1} & B_{k-1,0} \\ 0_n & 0_n & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & B_{1,1} & B_{1,0} \\ 0_n & 0_n & \cdots & 0_n & B_{0,0} \end{bmatrix} \tag{90}$$

$$b_{(k)} \equiv \begin{bmatrix} b_k \\ b_{k-1} \\ \vdots \\ b_1 \\ b_0 \end{bmatrix} \tag{91}$$

$$B_{0,0} \equiv B, b_0 \equiv b \tag{92}$$

The state-space model is completed by specifying the observation equation. This is given by

$$y_t^{sig}(i) = Z_\Psi \Psi_t^{(k)} + z v_t(i)$$

where

$$Z_\Psi \equiv [0_{n_y x n k} \quad Z_1]$$

We wish to determine the matrices $B_{i,j}$ and b_i in terms of known parameters of the model.

As before, assume that a time-invariant filter exists that is also independent of i, with the Kalman gain denoted by $K_\Psi \equiv \left[K_{k+1}' K_k' \cdots K_1'\right]'$. Let $\Psi_{t|s}^{(k)}(i) \equiv E_s^i \Psi_t^{(k)}$. The updating equation from the Kalman filter for firm i is

$$\Psi_{t|t}^{(k)}(i) = B_{(k)} \Psi_{t-1|t-1}^{(k)}(i) + K_\Psi \left(y_t^{sig}(i) - Z_\Psi B_{(k)} \Psi_{t-1|t-1}^{(k)}(i) \right)$$

Averaging across i and rearranging gives

$$\begin{aligned} \Psi_{t|t}^{(k)} &= B_{(k)} \Psi_{t-1|t-1}^{(k)} + K_\Psi Z_\Psi \left(\Psi_t^{(k)} - B_{(k)} \Psi_{t-1|t-1}^{(k)} \right) \\ &= B_{(k)} \Psi_{t-1|t-1}^{(k)} + K_\Psi Z_1 \left(\theta_t - B \theta_{t-1|t-1} \right) \end{aligned} \tag{93}$$

The first n equations of the system (93) can be written as

$$\begin{aligned} \bar{E}_t^{k+1}(\theta_t) &= \sum_{i=0}^{k} B_{k,i} \bar{E}_{t-1}^{i+1}(\theta_t) + K_{k+1} Z_1 \left(\theta_t - B \theta_{t-1|t-1} \right) \\ &= \sum_{i=1}^{k} B_{k,i} \bar{E}_{t-1}^{i+1}(\theta_t) + \left(B_{k,0} - K_{k+1} Z_1 B \right) \bar{E}_{t-1}(\theta_t) \\ &\quad + K_{k+1} Z_1 B \theta_{t-1} + K_{k+1} Z_1 b u_t \end{aligned} \tag{94}$$

Yet, the conjectured law of motion for $\bar{E}_t^{k+1}(\theta_t)$ implied by (89) is

$$\bar{E}_t^{k+1}(\theta_t) = \sum_{i=0}^{k+1} B_{k+1,i}\bar{E}_{t-1}^{i}(\theta_t) + b_{k+1}u_t \tag{95}$$

Thus, the law of motion for $\bar{E}_t^{k}(\theta_t)$ can be obtained by first matching coefficients in (94) and (95), to get

$$\begin{aligned}
B_{k+1,0} &= K_{k+1}Z_1B \\
B_{k+1,1} &= B_{k,0} - K_{k+1}Z_1B \\
B_{k+1,i} &= B_{k,i-1}, \quad i = 2, 3, \dots, k+1 \\
b_{k+1} &= K_{k+1}Z_1b
\end{aligned}$$

and, in turn, noting that these equalities imply

$$B_{k,0} = K_kZ_1B \tag{96}$$

$$B_{k,i} = (K_{k-i} - K_{k+1-i})\, Z_1B, \quad 1 \le i < k \tag{97}$$

$$B_{k,k} = \left(Z_1^{-1} - K_1\right) Z_1B \tag{98}$$

$$b_k = K_kZ_1b \tag{99}$$

These arguments also apply to lower-order expectations to obtain analogous expressions for $B_{i,j} \quad (i = 1, 2, \dots, k-1; j = 0, 1, \dots, k-1; i \ge j)$ and $b_i \quad (i = 1, 2, \dots, k-1)$.

The elements of K_Ψ remain to be determined. As before, the assumption that a time-invariant filter exists means

$$K_\Psi = \Sigma_\Psi Z_\Psi' F_\Psi^{-1} \tag{100}$$

where

$$\Sigma_\Psi \equiv var\left(\Psi_t^{(k)} - \Psi_{t|t-1}^{(k)}(i)\right) = B_{(k)}V_\Psi B_{(k)}' + b_{(k)}\Omega_u b_{(k)}' \tag{101}$$

$$V_\Psi \equiv var\left(\Psi_t^{(k)} - \Psi_{t|t}^{(k)}(i)\right) = \Sigma_\Psi - \Sigma_\Psi Z_\Psi' F_\Psi^{-1} Z_\Psi \Sigma_\Psi \tag{102}$$

$$F_\Psi \equiv var\left(y_t^{sig}(i) - Z_\Psi \Psi_{t|t-1}^{(k)}(i)\right) = Z_\Psi \Sigma_\Psi Z_\Psi' + \sigma_v^2 zz' \tag{103}$$

By the definition of Z_Ψ, (100) implies that

$$K_{k+1} = \Sigma_{k,0} Z_1' F^{-1} \tag{104}$$

where

$$\Sigma_{k,0} \equiv cov(\left[\bar{E}_t^k \theta_t - E_{t-1}^i \left(\bar{E}_t^k \theta_t\right)\right], \left[\theta_t - E_{t-1}^i \left(\theta_t\right)\right])$$

and $\Sigma_{0,0} \equiv \Sigma_{11}$. Substituting (102) and (103) into (101), we obtain a Riccati equation for Σ_Ψ:

$$\Sigma_\Psi = B_{(k)} \left(\Sigma_\Psi - \Sigma_\Psi Z_\Psi' \left(Z_\Psi \Sigma_\Psi Z_\Psi' + \sigma_v^2 zz'\right)^{-1} Z_\Psi \Sigma_\Psi\right) B_{(k)}' + b_{(k)} \Omega_u b_{(k)}' \quad (105)$$

This last equation can be simplified and partitioned to yield an expression for $\Sigma_{k,0}$:

$$\Sigma_{k,0} = \sum_{i=0}^{k} B_{k,i} \Sigma_{i,0} \left(B - \Sigma_{11} Z_1' F^{-1} Z_1\right)' + b_k \Omega_u b' \quad (106)$$

Notice that (106) cannot be directly recursively solved for $\Sigma_{k,0}$ (given Σ_{11}) because the matrices $\{B_{k,i}\}$ and $\{b_k\}$ themselves are functions of $\{K_k\}$, which, in turn, are functions of $\{\Sigma_{k,0}\}$. Instead, we can invert (104) to get an expression for $\Sigma_{k,0}$ in terms of K_{k+1} (and Σ_{11}) and substitute this and (96)-(99) into (106) to obtain a recursive set of equations for K_k in terms of known parameters. The resulting set of equations has the form:

$$K_{k+1} = D_1 K_{k+1} D_2 + D_3 \quad (107)$$

where

$$\begin{aligned}
D_1 &\equiv Z_1' F^{-1} \left(B - \Sigma_{11} Z_1' F^{-1} Z_1 B\right) \\
D_2 &\equiv F(Z_1')^{-1} A' - Z_1 \Sigma_{11} \\
D_3 &\equiv D_3 \left(K_k, K_{k-1}, \dots, K_1\right) \\
&\equiv Z_1' F^{-1} \left[K_k Z_1 B \Sigma_{11} + \sum_{i=1}^{k-1} \left(K_{k-i} - K_{k+1-i}\right) Z_1 B K_{i+1} F(Z_1')^{-1}\right] \cdot \\
&\quad \left[B - \Sigma_{11} Z_1' F^{-1} Z_1\right]' + K_k Z_1 b \Omega_u b'
\end{aligned}$$

Both D_1 and D_2 are functions of known parameters. Applying the $vec(\cdot)$ operator to (107), and rearranging, the unique solution of the elements of K_{k+1} ($k \geq 1$) can be found recursively from

$$vec(K_{k+1}) = \left[I_{n \cdot n_y} - (D_2' \otimes D_1)\right]^{-1} vec(D_3 \left(K_k, K_{k-1}, \dots, K_1\right))$$

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", unpublished paper, London School of Economics (available at http://nuff.ox.ac.uk/users/shin/working.htm).

[3] Amato, Jeffery D and Hyun Song Shin (2003a): "Equilibrium determinacy in differential information economies", notes, London School of Economics.

[4] Amato, Jeffery D and Hyun Song Shin (2003b): "Optimal monetary policy in the presence of public and private information", notes, London School of Economics.

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

[6] 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.

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

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

[9] Harvey, Andrew (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

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

[11] Karlin, Samuel and H Taylor (1975): *A First Course in Stochastic Processes*, second edition, Academic Press, New York.

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

[13] Kemeny, J G, J L Snell and A W Knapp (1966): *Denumerable Markov Chains*, van Nostrand, New York.

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

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

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

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

[18] 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.

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

[20] Searle, S R (1971): *Linear Models*, Wiley, New York.

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

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

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

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

[25] 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.

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

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

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

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

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

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

[32] 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.

[33] Woodford, Michael (2003b): *Interest and Prices: Foundations of a Theory of Monetary Policy*, forthcoming, Princeton University Press, Princeton (draft version available at http://www.princeton.edu/˜woodford/).

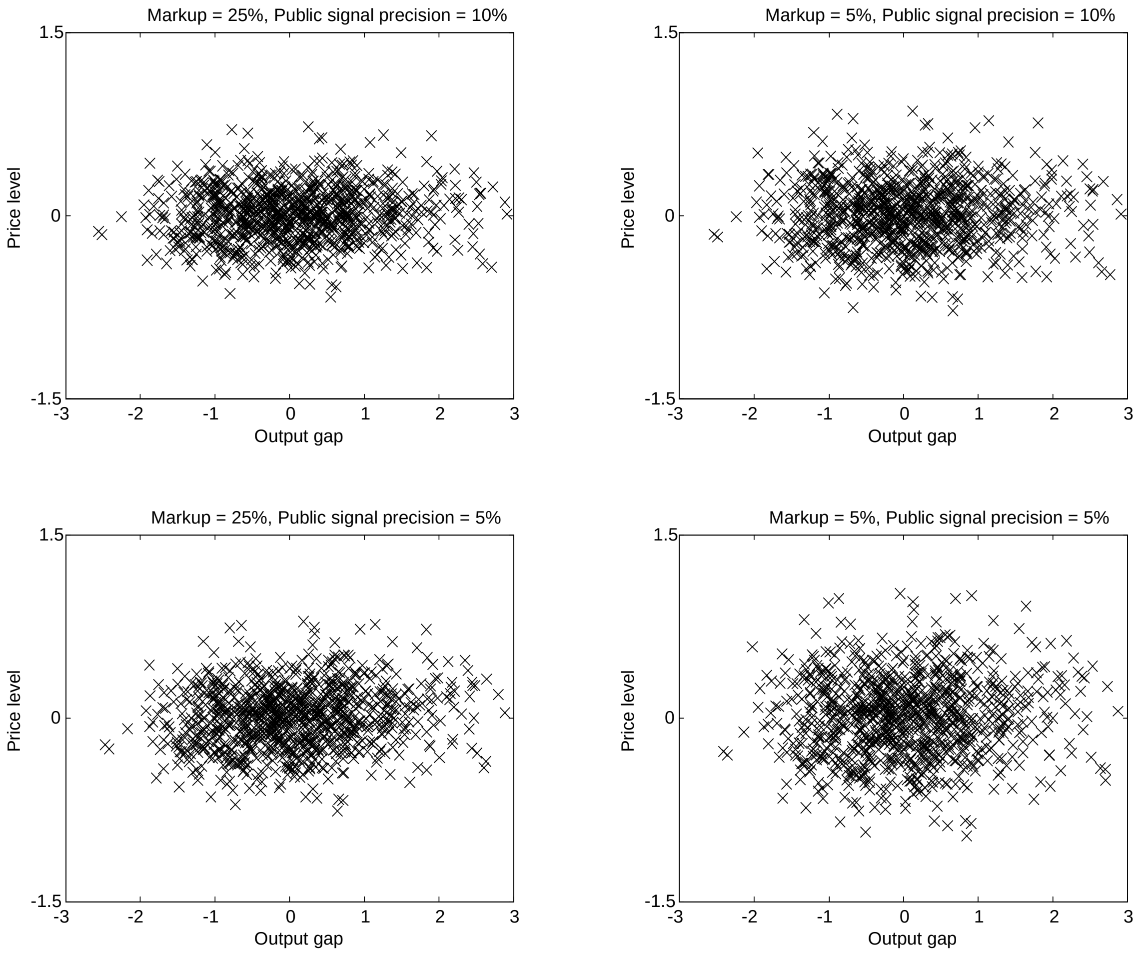
Table 1

Baseline calibrated parameters

Preferences and technology	
σ	2
ζ	0.3
ν	2
ϵ	11
Markup	10%
Natural rate of interest	
ρ	0.8
σ_ε^2	$(1-\rho^2)$%/quarter
$\text{var}(r_t^n)$	1%/quarter
Signals	
σ_v^2	0.2%/quarter
σ_η^2	0.2%/quarter
Monetary policy	
λ	1
δ	1

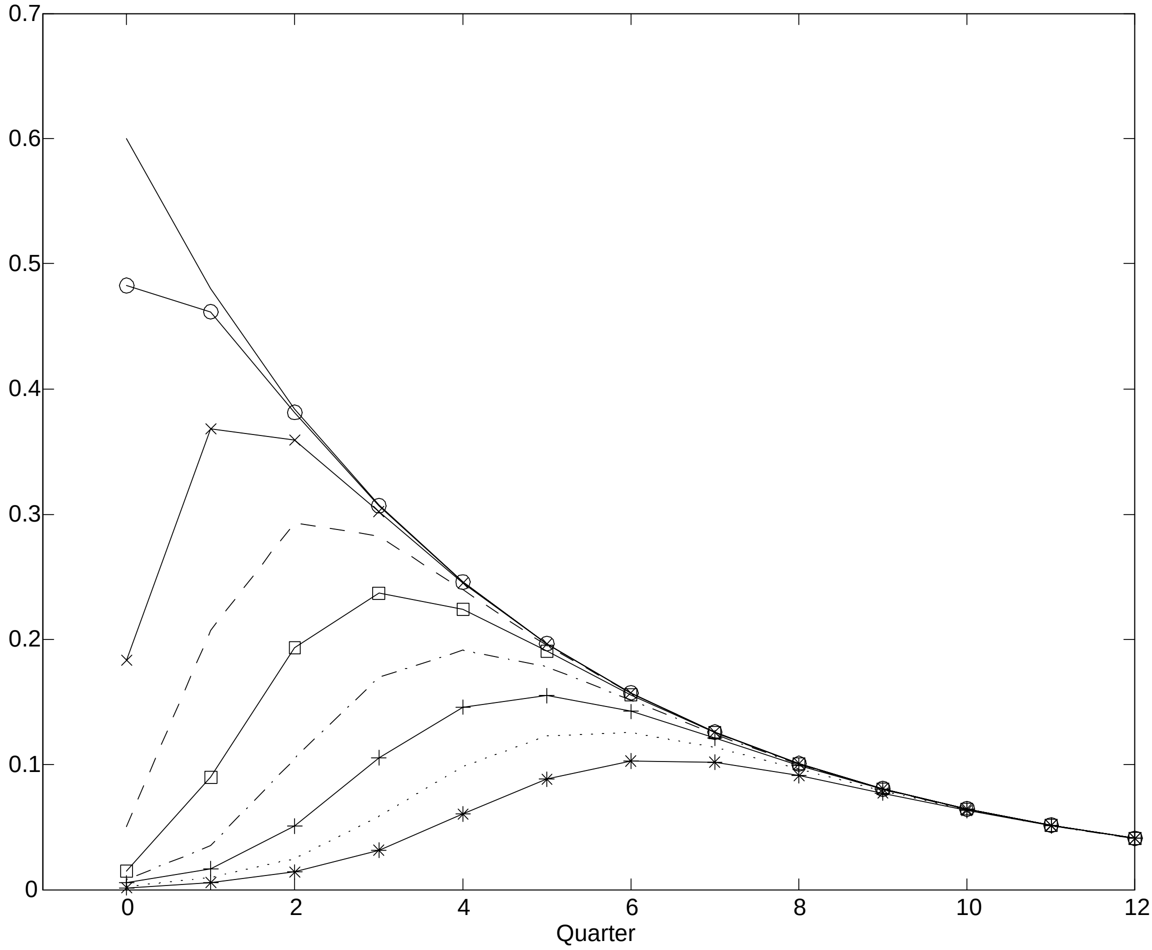
Figure 1

Effects of changing the markup and precision of public signals:
sample realisations of the output gap and price level



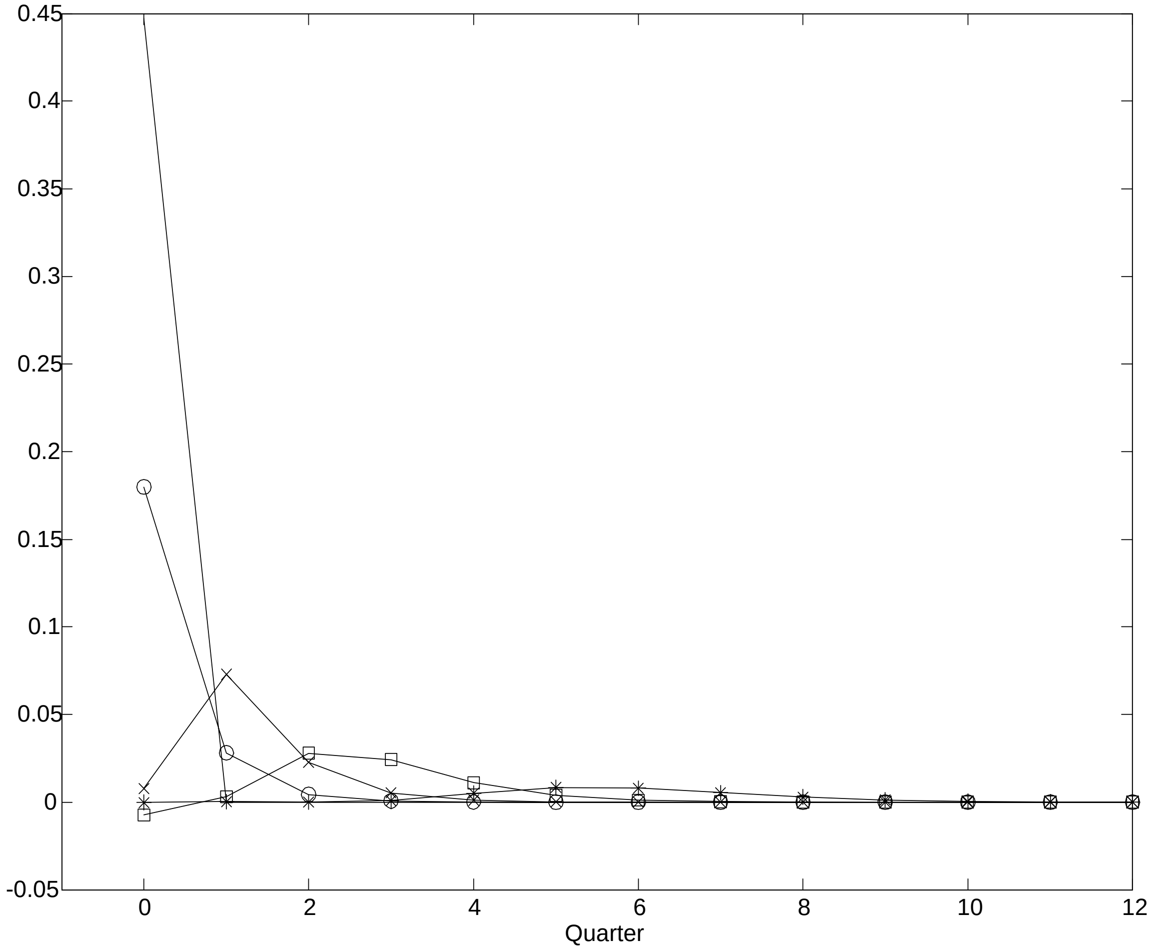
Notes: Each panel plots one sample realisation of the price level against the output gap. The same sample of randomly drawn shocks is used in each panel when simulating the time paths of the endogenous variables. Data is constructed for 1100 periods, but the first 100 observations are dropped to minimise the influence of initial values. The price level and output gap are in percentages.

Figure 2
Impulse responses of higher-order expectations of natural rate of interest: shock to natural rate of interest



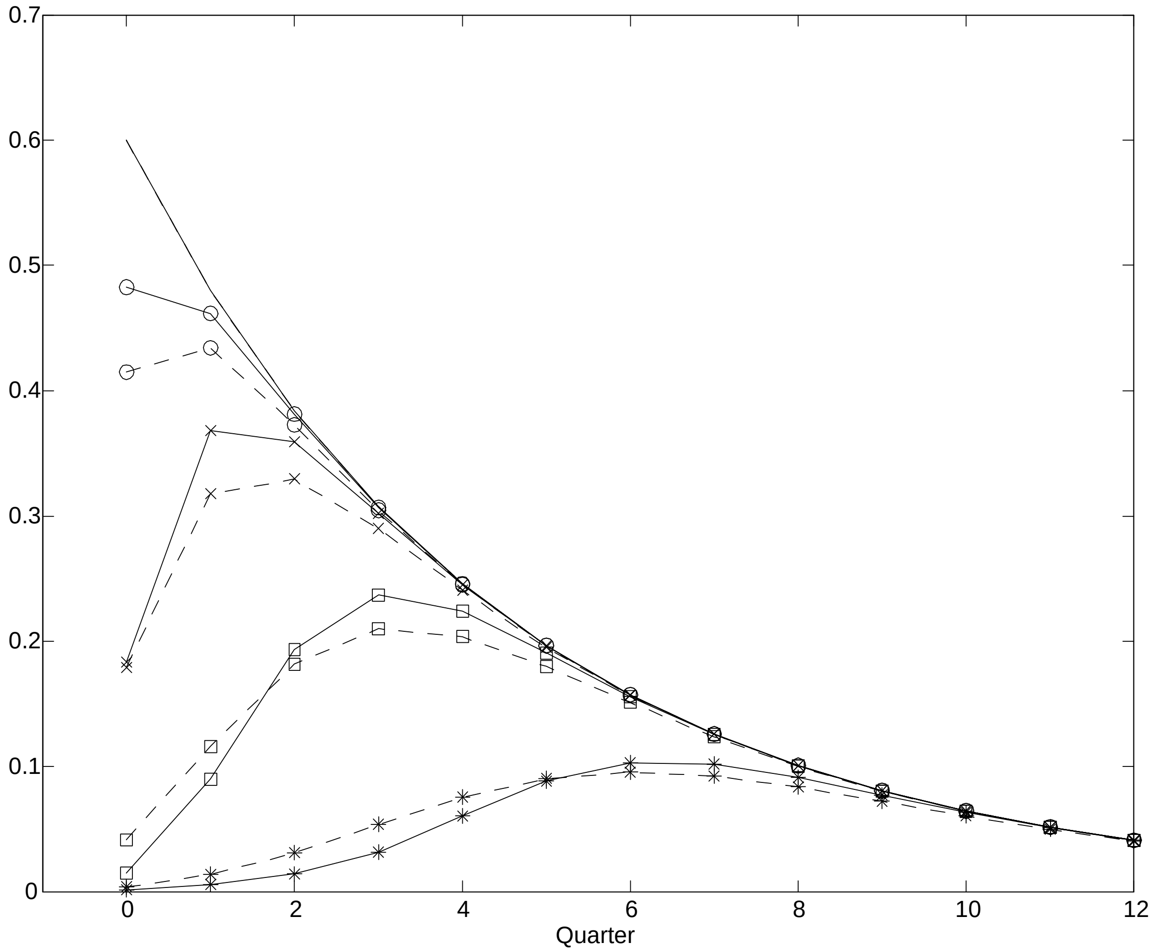
Notes: The figure shows the impulse responses of higher-order expectations of the natural rate of interest (in percentages) with respect to a one-standard deviation innovation in the natural rate of interest. The solid line is the path followed by the natural rate of interest, while the other lines correspond to successively higher orders k of expectations, from $k = 1$ (o) to $k = 8$ (*).

Figure 3
Impulse responses of higher-order expectations of natural rate of interest:
shock to public signal



Notes: The figure shows the impulse responses of higher-order expectations of the natural rate of interest (in percentages) with respect to a one-standard deviation innovation in the shock to the public signal. The solid line is the path followed by the public signal shock, while the other lines correspond to higher-order expectations: $k = 1$ (o), $k = 2$ (x), $k = 4$ (□) and $k = 8$ (*).

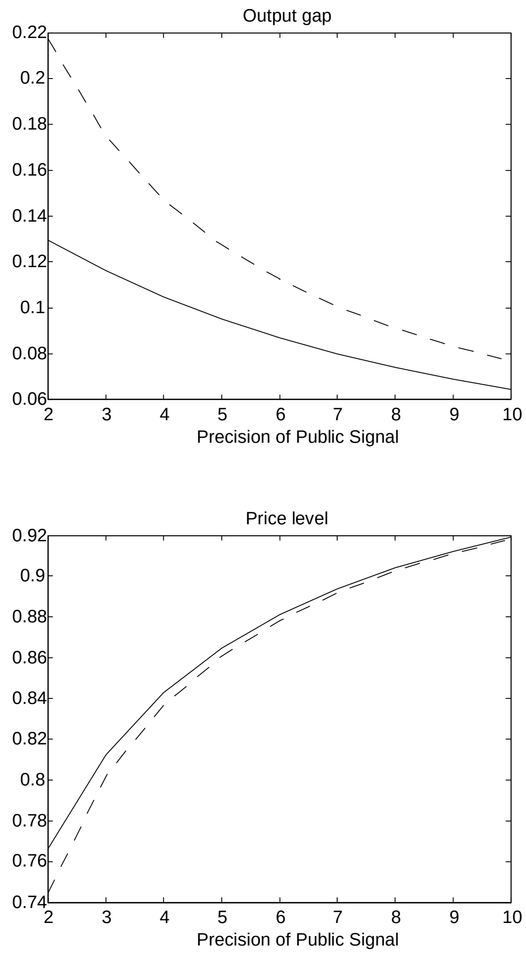
Figure 4
Impulse responses of higher-order expectations of natural rate of interest:
effects of public information



Notes: The figure shows the impulse responses of higher-order expectations of the natural rate of interest (in percentages) with respect to a one-standard deviation innovation in the natural rate of interest. The solid line is the path followed by the natural rate of interest. The solid lines with symbols represent the case when there are both public and private signals present, while the dashed lines are the case of private signals only. The lines distinguished by symbols, whether solid or dashed, correspond to different degrees of higher-order expectations: $k = 1$ (o), $k = 2$ (x), $k = 4$ (□) and $k = 8$ (*).

Figure 5

Variances of endogenous variables with respect to precision of public signal



Notes: The figure plots the variances of endogenous variables with respect to the precision of the innovation in the public signal. The precision of the private signal is set equal to 10 percent (solid line) or 2 percent (dashed line). Inflation and the interest rate are expressed in annualised percentages, while the price level, output gap and precision of signal innovations are in percentages.

Discussion of "Public and private information in monetary policy models", by Jeffery D Amato and Hyun Song Shin

Marvin Goodfriend[1]

Broadly speaking, modern central banks aim to facilitate the functioning of a market economy with minimal direct interference in the decision-making of households and firms. For the most part, central bankers used to think that the best way to do so was to operate in secret and out of the limelight.[2] More recently, central bankers have come to appreciate the importance of transparency in connection with the emphasis on price stability. Today's central bankers recognise that building credibility for the commitment to price stability is the best way to maximise the power of monetary policy to stabilise the macroeconomy over the business cycle. Advocates of inflation targeting emphasise that transparency rather than secrecy regarding the procedures and objectives of monetary policy is the best way to build that credibility.[3]

The paper by Amato and Shin recognises that central banks must actively shape and influence events to facilitate the functioning of the macroeconomy. However, they take their analysis in a different direction. They point out a fundamental tension in central banking. A central bank needs to react to data in order to manage interest rate policy. Yet it distorts the very data from which it seeks guidance in the process of influencing behaviour through public pronouncements of its objectives and its views on the fundamental state of the economy.

In other words, Amato and Shin show that a central bank's assessment of the underlying state of the economy may be reflected in aggregate behaviour in a way that causes the data generated by agents to obscure the fundamentals. There is a problem even if a central bank's views of the aggregate state are accurate. Of course, the problem is worse when the central bank's views are wrong. Either way, public announcements by the central bank become a powerful focal point for coordination of private agents.

The point is very interesting and potentially important. Let me give two examples of how I think Amato and Shin's tension has worked in practice. First, consider the low inflation period in the 1950s to the mid-1960s in the United States. At the time, the Fed's denial of any persistent inflationary potential may have succeeded in holding down actual and expected inflation somewhat. Households and firms probably held back on wage and price pressures for a while because they were confident that the low inflation equilibrium would be maintained. As a result, the economy operated at a higher level of real economic activity for a while than was ultimately sustainable.

In terms of Amato and Shin, the Fed's insistence that inflation was not a threat probably distorted behaviour in a way that seemed to confirm the belief that trend inflation would remain low. This may have been one reason why monetary policy was insufficiently pre-emptive in the early stages of the Great Inflation.

A similar dynamic may have been operating after the Fed restored credibility for low inflation in the late 1990s. Here again, in retrospect, the economy can be seen to have operated for a few years at levels that were not sustainable. One reason may have been that the Fed's credibility for low inflation made wage and price setters confident that the low inflation equilibrium would be maintained. In this case, the unsustainability did not precipitate an outbreak of inflation. It resulted instead in extreme asset price fluctuations, excessive consumption growth, and an unsustainable investment boom. The distorted behaviour helped to delay a tightening of monetary policy that might have avoided much of

---

1 Senior Vice President and Policy Advisor, Federal Reserve Bank of Richmond. The views expressed are those of the author and not necessarily those of the Federal Reserve Bank of Richmond, the Federal Reserve System or the BIS.

2 See, for instance, Goodfriend (1986).

3 See, for instance, Bernanke and Mishkin (1997) and Svensson (1999).

the cyclical instability. Although, clearly other factors such as the terrorist attack on the World Trade Center contributed to the 2001 recession.

The paper explores the fundamental tension identified by Amato and Shin in a monopolistically competitive macro model in which the nominal pricing decisions of individual firms depend on their own demand and cost conditions, and on the strategic complementarity among the pricing decisions of firms in the aggregate. Monopolistically competitive firms must guess the pricing decisions of their potential competitors in order to choose a nominal price that achieves their desired relative price. This, in turn, means that each firm must take into account its beliefs about the beliefs of other firms about pricing, and so on in a potentially infinite recursion.

The main point of the paper is that such concerns lead firms to set prices that are potentially far less sensitive to their best direct (idiosyncratic) estimates of the underlying fundamentals. In the limit, Amato and Shin show that firms will set prices conditional only on information known in common about the underlying fundamentals, whether that common information is accurate or not.

In the formal model used to study this issue in the paper, firms receive both private and public noisy signals of two current shocks, an aggregate demand and a productivity shock. Households and the central bank are assumed to observe the two underlying shocks perfectly. The labour market is perfectly competitive. And nominal prices and wages are perfectly flexible.

It is useful to note that the macro model utilised by Amato and Shin may be characterised as a monopolistically competitive real business cycle model with perfectly flexible prices and wages. Equivalently, for the full information case, the equilibrium resembles the flexible price, perfectly competitive labour market version of Blanchard and Kyotaki's monopolistically competitive macro model.[4] In this case, firms choose their product price to maintain the constant profit-maximising markup at all times. The size of the markup then acts like a tax that governs how far equilibrium aggregate employment falls below the level that would be attained in a perfectly competitive macro model.

With this insight in hand, we can consider what optimal monetary policy might look like in Amato and Shin's model where the central bank and households are fully informed but the firms are not. If firms are imperfectly informed, then they can no longer be counted upon to set prices to maintain the constant profit maximising markup. Presumably, the all-knowing central bank could see any pricing errors in this flexible price environment and offset them with monetary policy. In other words, optimal monetary policy would stabilise the markup to reproduce the full information, flexible price outcome.[5]

However, when Amato and Shin analyse monetary policy in their model, they use a kind of ad hoc Taylor rule. They should also discuss the extent to which the complications for macroeconomic stabilisation due to the information problems they highlight could be overcome if the central bank were allowed to behave optimally with imperfect information. I will return to this point shortly.

The main surprising finding of the analysis of the flexible price macro model is that when "firms observe fairly precise private signals of aggregate shocks, the mere presence of the public signal, interpreted as a signal with precision greater than zero, actually makes inflation more volatile".[6] This is seen in the top right-hand panel of Figure 2. Note, however, in the bottom left-hand panel that a more precise public signal always reduces the volatility of real output.

Amato and Shin suggest that these results mean that more precise public information does not necessarily lead to better welfare outcomes. They emphasise that this finding cannot be attributed to poor information available to the central bank, since it is given full information in the analysis.

One can question their findings on two counts. First, smoothing the volatility of output is not necessarily the right metric for assessing welfare. According to the discussion above, maintaining markup constancy is a better one, since it produces the outcomes of a real business cycle model with perfectly flexible wages and prices. With markup constancy, employment could be relatively stable even if output were highly volatile due to volatile productivity shocks.

---

4 See Blanchard and Kiyotaki (1987).

5 See Goodfriend (2002b).

6 See page 32 of the paper.

Second, although the central bank is given full information, it is forced to operate with a suboptimal Taylor rule. More precise public information might improve welfare if the central bank were allowed to pursue monetary policy optimally.

One can question whether the conditions necessary to create the possibility that public information can be harmful are likely to be met. By its very nature, a private signal is not likely to be very informative about aggregates because it will reflect a relatively small part of the economy. On the other hand, private information gets aggregated over the entire economy. To make their argument more persuasive, Amato and Shin should explain in an intuitive quantitative way whether the precision of private information necessary for public information to be harmful is likely to be attained in practice.

Assuming that the above condition is met, what would the authors have us believe about the government's information policy? In any case, the government would need to continue to collect aggregate data to make monetary and fiscal policy. Leaks would be inevitable if such data were not made public officially. Moreover, private groups would continue to collect and disseminate data on various sectors of the economy. Surveys would continue to be collected to improve the understanding of fundamentals. Financial market prices would continue to provide valuable public information. All these sources of information would provide noisy public signals of the underlying aggregate state of the economy. Such common information would be valued and utilised by individual firms and households even if its use has negative consequences for social welfare as Amato and Shin suggest.

Actually, given that the government cannot suppress noisy common signals of the aggregate state of the economy, Amato and Shin's findings suggest that the government might actually increase social welfare by improving the precision of the public signals! This conclusion seems possible if it is infeasible to suppress common signals, and the economy inevitably operates to the right of the maximum of the curve in the left-hand panel of Figure 2.

Thus, the second best information policy might actually be just the opposite of the infeasible first best. If it were impossible to eliminate the common public information, then the next best thing to do might be to devote more resources to reducing the measurement error and improving the coverage and accuracy in our national statistics. In other words, the findings of Amato and Shin could be interpreted as favouring either greater opacity or greater transparency depending on the circumstances.

Amato and Shin compare the data generating implications of their flexible price, imperfect information model to those of a sticky price model without information imperfections. An interesting and important extension of their work would be to analyse a single model incorporating sticky prices and the kind of information processing that they emphasise in this paper. It would seem that the information imperfections that Amato and Shin are concerned about might matter less in a sticky price model, especially if strict inflation targeting were close to the optimum policy.

Amato and Shin conjecture that a central bank could do harm by targeting inflation too strictly in the short run. The reason is that doing so degrades the private information in inflation as a signal of the output gap.[7] This is an important point, and it should be taken seriously. Clearly, the credibility that the Fed acquired for price stability since the mid-1990s has helped to stabilise the inflation rate in spite of fairly large swings in the output gap since then. In fact, that is to be expected of successful explicit or implicit inflation targeting.[8]

In some ways, interest rate policy decisions are more difficult to make in the absence of cyclically volatile inflation. On the other hand, waiting for inflation to trigger interest rate policy actions created destabilising go/stop monetary policy when it was tried in the 1960s and 1970s.[9] The go/stop experience suggests that the benefits to firmly anchoring inflation and inflation expectations over the business cycle are well worth the loss of inflation as a guide for interest rate policy actions.[10]

[7] Borio and Lowe (2002) call this the "paradox of credibility".

[8] See Goodfriend (2003).

[9] See Goodfriend (1997).

[10] See Goodfriend (2003).

It will not be easy for central banks to learn to utilise signals other than inflation to guide interest rate policy. For instance, the Fed's policy problems in the late 1990s make that clear.[11] But with time, central banks will no doubt improve their ability to manage interest rate policy in a world of price stability.

References

Bernanke, B and R Mishkin (1997): "Inflation targeting: a new framework for monetary policy?", *Journal of Economic Perspectives*, 11, pp 97-116.

Blanchard, O and N Kyotaki (1987): "Monopolistic competition and the effects of aggregate demand", *American Economic Review*, 77, pp 647-66.

Borio, C and P Lowe (2002): "Asset prices, financial and monetary stability: exploring the nexus", *BIS Working Paper*, no 114, Basel, July.

Goodfriend, M (1986): "Monetary mystique: secrecy and central banking", *Journal of Monetary Economics*, 17, pp 63-92.

——— (1997): "Monetary policy comes of age: a 20th century odyssey", Federal Reserve Bank of Richmond, *Economic Quarterly*, 83, pp 1-22.

——— (2002): "The phases of US monetary policy: 1987 to 2001", Federal Reserve Bank of Richmond, *Economic Quarterly*, 88, pp 1-17.

——— (2002b): "Monetary policy in the new neoclassical synthesis: a primer", *International Finance*, 5, pp 165-91.

——— (2003): *Inflation targeting in the US?*, Paper prepared for the NBER Conference on Inflation Targeting, Miami, Florida, January. Forthcoming in the conference volume.

Svensson, L (1999): "Inflation targeting as a monetary policy rule", *Journal of Monetary Economics*, 43, pp 607-54.

[11] See Goodfriend (2002a).

Discussion of "Public and private information in monetary policy models", by Jeffery D Amato and Hyun Song Shin[1]

Lars E O Svensson[2]

Jeff Amato and Hyun Shin have produced a very fine paper (Amato and Shin (2003)). It is a pleasure to discuss it. The main message is that central bank information may have bad consequences. It could degrade the information value of private signals, and it could increase the volatility of inflation. This makes the paper something of an anti-transparency paper, a somewhat rare thing in this age of central banking transparency. However, I do not believe that the anti-transparency flavour stands up to scrutiny. Indeed, I will argue that the paper's main result can rather be interpreted as a pro-transparency one.

The paper discusses difficult issues with the help of a very elegant and powerful framework, modelling differential information with the help of Markov chains and related matrix algebra. First, the authors provide a simple static example of their analysis. Then they provide a more elaborate intertemporal model of a New Keynesian model of a monetary economy.

In the simple example, a typical firm i ($i = 1, 2, ..., N$) sets the (log) price p_i of its product according to

$$p_i = \mathrm{E}^i p + \xi \mathrm{E}^i (y - \bar{y}), \tag{1}$$

where E^i denotes the firm's expectation or estimate conditional on its private information; $p \equiv \frac{1}{N}\sum_{i=1}^{N} p_i$ denotes the aggregate (log) price level; ξ ($0 < \xi < 1$) is a parameter; and $y - \bar{y}$ denotes the output gap, the difference between (log) output, y, and (log) potential output, $\bar{y}$. This pricing equation can be rewritten as

$$p_i \equiv (1 - \xi)\mathrm{E}^i p + \xi \mathrm{E}^i (p + y - \bar{y}),$$

where $p + y - \bar{y}$ can be interpreted as (log) nominal GDP adjusted for potential output. By taking the average of this equation, we get

$$p = (1 - \xi)\bar{\mathrm{E}} p + \xi \bar{\mathrm{E}} (p + y - \bar{y}),$$

[1] These comments borrow a few points from my comments on Woodford (2003) in Svensson (2003a). I thank Kathleen Hurley for editorial and secretarial assistance. The views expressed are those of the author and not of the BIS.

[2] Princeton University.

where $\bar{\mathrm{E}}[\cdot] \equiv \bar{\mathrm{E}}^1[\cdot] = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}^i[\cdot]$ denotes the average (first-order) expectations operator. Since $0 < \xi < 1$, by recursive substitution of the term $\bar{\mathrm{E}}p$, we can write the average price equation as

$$p = \xi \sum_{k=1}^{\infty} (1-\xi)^{k-1} \bar{\mathrm{E}}^k (p + y - \bar{y}), \tag{2}$$

where $\bar{\mathrm{E}}^k$ denotes kth-order average expectations defined as

$$\bar{\mathrm{E}}^k[\cdot] \equiv \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}^i \left[\bar{\mathrm{E}}^{k-1}[\cdot]\right] \qquad (k \geq 2).$$

Equation (2) shows that the average price level depends on an infinite sum of higher-order expectations of nominal GDP adjusted for the output gap, with the weight on higher-order expectations being larger, the smaller the parameter ξ. The smaller the parameter ξ, the stronger the strategic complimentarity of the individual firms' pricing decisions.

The paper shows that, if there is public information, higher-order expectations converge to public expectations,

$$\bar{\mathrm{E}}^k[\cdot] \rightarrow \bar{\mathrm{E}}[\cdot|\text{Public information}] \qquad (k \rightarrow \infty).$$

The paper then shows that the outcome depends on the relative precision of private and public information. When private precision is good, introducing bad public information may increase the volatility of inflation in the New Keynesian model. If the precision of public information improves, however, the volatility of inflation falls, as seen in Figures 2 and 3 of the paper.

Indeed, I believe that it is this latter result that makes the paper a pro-transparency paper. In the real world, there is already considerable public information, for instance, data and forecasts published by various government agencies and private forecasters. Since there is already public information, the results of the paper indicate that central banks should provide as good *additional* public information as possible, to improve the precision of the public information. Looked at this way, the results of this paper become pro-transparency rather than anti-transparency.

The parameter ξ is crucial for the relative importance of public information (recall that a lower ξ implies more weight on higher-order expectations). The paper shows how ξ is determined in a rather complex way in the New Keynesian model. However, ξ could also depend on monetary policy. This can be illustrated in the simple example above. Suppose that monetary policy results in a targeting rule of the form

$$p + \lambda(y - \bar{y}) = q, \tag{3}$$

where $\lambda \geq 0$ is a parameter related to the monetary policy regime and q is some exogenous error term. The case $\lambda = 0$ could be interpreted as strict price level targeting, $\lambda > 0$ could be interpreted as flexible price-level targeting, $\lambda = 1$ could be interpreted as a kind of nominal GDP targeting (where nominal GDP is adjusted for potential output), and $\lambda \to \infty$ could be interpreted as strict output gap targeting.

We can use equation (3) to eliminate the output gap in equation (1). This results in the new pricing equation

$$p_i = (1 - \tilde{\xi})\mathrm{E}^i p + \tilde{\xi} q,$$

where the new parameter, $\tilde{\xi}$, is given by

$$\tilde{\xi} \equiv \frac{\xi}{\lambda}.$$

If $\lambda > \xi$, we have $\tilde{\xi} < 1$, and we can still do the recursive substitution leading to equation (2), where $\tilde{\xi}$ replaces ξ and the higher-order expectations refer to q rather than $p + y - \bar{y}$. Thus, λ affects the size of $\tilde{\xi}$ for given ξ, and thereby the relative weight on higher-order expectations. However, if $\lambda < \xi$, we have the $\tilde{\xi} > 1$, and the recursive substitution no longer makes sense. Indeed, firms' individual price setting decisions are then no longer strategic complements but strategic substitutes.

What order k of firms' expectations are sensible? How rational and sophisticated are the firms? In principle, one could find out via the surveys of inflation expectations that many central banks undertake these days. One could ask questions of the following form to individual firms: (1) What do you think the average price level is? (2) What do you think other firms think the average price level is? (3) What do you think other firms think other firms think the average price level is?; and so forth. These questions are obviously constructed such that averaging the responses to the kth question gives the kth-order average expectations. It would be very interesting to see whether firms could give sensible answers to higher-order questions. I would certainly have to think a while myself before answering such questions, and I am not sure how many high-order questions I would have an answer to.

One possibility is that agents would display bounded rationality and simplify the formation of higher-order expectations. Two alternatives immediately present themselves. One is that higher-order expectations beyond some fixed order K are set equal to the Kth-order expectations, $\bar{\mathrm{E}}^k[\cdot] = \bar{\mathrm{E}}^K[\cdot]$ for $k > K$. Another is that higher-order expectations beyond some fixed order K are set equal to a constant expectations operator, for instance, the expectations conditional on the public information, $\bar{\mathrm{E}}^k[\cdot] = \bar{\mathrm{E}}[\cdot|\text{Public information}]$ for $k > K$. Clearly, these two alternatives have very different consequences. The first would reduce the weight on public information; the second would increase that weight. It is not clear that one case is more plausible than the other.

These comments indicate that there remain quite a few interesting issues for future research, and I very much hope the authors will address them in their future research.

Finally, let me voice a complaint on the conference version of this otherwise so fine paper. The authors present a model in which they model firms' pricing and households' consumption not as following ad hoc rules of behaviour but those of rational and goal-directed agents; ie, by specifying objectives and constraints and then deriving optimal first-order conditions that describe private sector behaviour with a structural relation. But when the authors model monetary policy, they don't follow the same healthy principles of analysis. Instead, they model the central bank as following an ad hoc reaction function, an instrument rule, either a Taylor rule or a so-called forecast-based instrument rule. There is no reason to believe that such an ad hoc reaction function would be structural. As I have argued elsewhere, for instance, in Svensson (2003b), good central banks are at least as goal-directed and rational as the average household and firm (and they certainly employ more PhDs). Therefore, it makes a lot of sense to model good monetary policy as optimising, by using optimal targeting rules instead of ad hoc instrument rules. Indeed, Charles Bean's paper at this conference, Bean (2003), shows very pedagogically how this can be done and how helpful such an approach is in sorting out some common confusion about the role of asset prices regarding objectives and responses in monetary policy.

As stated above, the previous comment applies to the conference version of the paper. The post-conference version of June 2003 has abandoned the ad hoc reaction function and instead models inflation targeting as implementing a targeting rule similar to the form (3). Needless to say, I welcome this change. The optimal policy is further examined in another paper of the authors.

References

Amato, Jeffery D, and Hyun Song Shin (2003): "Public and private information in monetary policy models," working paper, BIS, Basel.

Bean, Charles (2003): "Asset prices, financial imbalances and monetary policy: are inflation targets enough?", working paper, Bank of England.

Svensson, Lars E O (2003a): "Comment on Michael Woodford, 'Imperfect common knowledge and the effects of monetary policy'," in Philippe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford, eds, *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S Phelps*, Princeton University Press, 2003, pp 59–63.

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

Woodford, Michael (2003): "Imperfect common knowledge and the effects of monetary policy," in Philippe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford, eds, *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S Phelps*, Princeton University Press, 2003, pp 25–58.

Recent BIS Working Papers

No	Title	Author
137 September 2003	The Great Depression as a credit boom gone wrong	Barry Eichengreen and Kris Mitchener
136 September 2003	The price level, relative prices and economic stability: aspects of the interwar debate	David Laidler
135 September 2003	Currency crises and the informational role of interest rates	Nikola A Tarashev
134 September 2003	The cost of barriers to entry: evidence from the market for corporate euro bond underwriting	João A C Santos and Kostas Tsatsaronis
133 September 2003	How good is the BankScope database? A cross-validation exercise with correction factors for market concentration measures	Kaushik Bhattacharya
132 July 2003	Developing country economic structure and the pricing of syndicated credits	Yener Altunbaş and Blaise Gadanecz
131 March 2003	Optimal supervisory policies and depositor-preference laws	Henri Pagès and João A C Santos
130 February 2003	Living with flexible exchange rates: issues and recent experience in inflation targeting emerging market economies	Corrinne Ho and Robert N McCauley
129 February 2003	Are credit ratings procyclical?	Jeffery D Amato and Craig H Furfine
128 February 2003	Towards a macroprudential framework for financial supervision and regulation?	Claudio Borio
127 January 2003	A tale of two perspectives: old or new challenges for monetary policy?	Claudio Borio, William English and Andrew Filardo
126 January 2003	A survey of cyclical effects in credit risk measurement models	Linda Allen and Anthony Saunders
125 January 2003	The institutional memory hypothesis and the procyclicality of bank lending behaviour	Allen N Berger and Gregory F Udell
124 January 2003	Credit constraints, financial liberalisation and twin crises	Haibin Zhu
123 January 2003	Communication and monetary policy	Jeffery D Amato, Stephen Morris and Hyun Song Shin
122 January 2003	Positive feedback trading under stress: Evidence from the US Treasury securities market	Benjamin H Cohen and Hyun Song Shin
121 November 2002	Implications of habit formation for optimal monetary policy	Jeffery D Amato and Thomas Laubach
120 October 2002	Changes in market functioning and central bank policy: an overview of the issues	Marvin J Barth III, Eli M Remolona and Philip D Wooldridge
119 September 2002	A VAR analysis of the effects of monetary policy in East Asia	Ben S C Fung
118 September 2002	Should banks be diversified? Evidence from individual bank loan portfolios	Viral V Acharya, Iftekhar Hasan and Anthony Saunders
117 September 2002	Internal rating, the business cycle and capital requirements: some evidence from an emerging market economy	Miguel A Segoviano and Philip Lowe
116 September 2002	Credit risk measurement and procyclicality	Philip Lowe