CENTRE FOR APPLIED MACROECONOMIC ANALYSIS

The Australian National University

CAMA Working Paper Series January, 2010

THRESHOLD PRICING IN A NOISY WORLD

Timo Henckel
Centre for Applied Macroeconomic Analysis, ANU

Gordon D. Menzies
University of Technology Sydney
Centre for Applied Macroeconomic Analysis, ANU

Daniel J. Zizzo
University of East Anglia
Centre for Applied Macroeconomic Analysis, ANU

CAMA Working Paper 1/2010
http://cama.anu.edu.au

Threshold Pricing in a Noisy World*

Timo Henckel[§] Gordon D. Menzies[†,§]

Daniel J. Zizzo[‡,§]

§ Centre for Applied Macroeconomic Analysis, ANU

† University of Technology Sydney

‡ University of East Anglia

29 December 2009

Abstract

We propose that the formation of beliefs be treated as statistical hypothesis tests, and label such beliefs inferential expectations. If a belief is overturned due to sufficient contrarian evidence, we assume agents switch to the rational expectation. We build a state dependent Phillips curve, and show that adjustments to equilibria may be contaminated by signal censoring, where agents in possession of extreme information are the first to adjust to changed economic circumstances. This approach is able to replicate recent micro-level evidence on firms' pricing behavior and sheds light onto the dynamics of disaggregated prices.

JEL classification: E30; E50

Keywords: Price stickiness; expectations; monetary policy; inflation; rationality; macroeconomics

*We thank Austin Gerig, Bruce Preston, Julio Rotemberg and seminar participants at the 2009 CAMA Behavioral Macro Conference for useful comments.

"While I was aware a lot of these [subprime] practices were going on, I had no notion of how significant they had become until very late. I didn't really get it until very late in 2005 and 2006."
Alan Greenspan on CBS "60 minutes", 16 September 2007

1 Introduction

Recent decades have seen a theoretical backing away from Rational Expectations (RE).[1] Examples include: (1) near rationality (Akerlof and Yellen, 1985), (2) agents as econometricians (Sargent 1993), (3) parameter uncertainty and econometric learning (Evans and Honkapohja, 2001), (4) model uncertainty and robustness (Hansen and Sargent, 2001), (5) infrequent rational expectations, when inattentive agents consult experts (Carroll, 2003), (6) information processing constraints and 'rational inattention' (Sims, 2003), (7) utility-based beliefs, or 'optimal' expectations (Brunnermeier and Parker, 2005), and, the subject of this paper, (8) inferential expectations (Menzies and Zizzo, 2009).

The approach of inferential expectations is captured in Alan Greenspan's quote: evidence must build up to a 'significant' threshold before agents 'get it'. This seems reasonable for three reasons.

First, asset prices sometimes move far more sharply than the underlying fundamentals at the onset of a crisis, when beliefs are typically being revised. Figure 1 shows evidence from three decades.

Our explanation for the mismatch between the *gradual* erosion of fundamentals and the *rapid* asset price declines is that the fundamental factors are noticed by agents as they form expectations. But, as described in the quote, they take some time to 'get it' before dramatically downgrading their forecasts. At that instant, asset prices decline sharply.

Second, in Menzies and Zizzo (2005) undergraduates were shown the contents of two urns, each with different combinations of white and orange balls. One urn was then selected randomly, and subjects received signals about its contents by the means of random ball draws with replacement from the chosen urn. The prior probability of a particular urn being chosen was 0.5 at the start of the experiment, but should then have evolved under RE according to Bayes' rule, as *each* new ball was drawn. In fact, agents displayed belief conservatism. Even allowing for rounding, agents failed to switch their probability guess far more often - over 30 per cent of the time - than would be expected under RE. The evidence suggests that subjects often noticed the evidence without changing their mind (further, less direct experimental evidence for belief conservatism is discussed in the exchange rate experiment by Menzies and Zizzo, 2007).

[1] RE has equivocal support in the experimental literature. While its predictions are not rejected as null hypotheses in some contexts (see Dwyer et al., 1993), the most common outcome is that individuals do not hold RE (e.g., Schmalensee, 1976; Blomqvist, 1989; Beckman and Downs, 1997; Swenson, 1997). In addition, experimental research often finds either under-utilization or over-utilization of priors (Camerer, 1995).

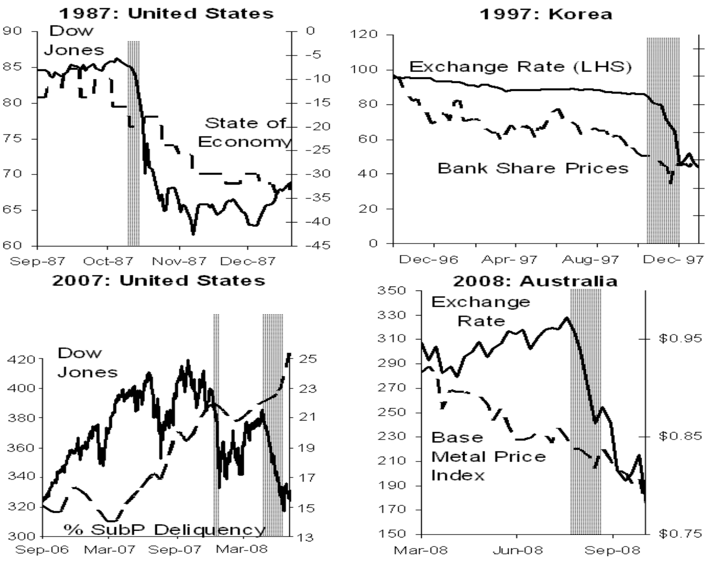

Figure 1: Sudden Asset Price Movements in Three Decades of Crises

Third, we appeal to the philosophy of science. While economists have become uneasy with RE in recent decades, a parallel debate has been carried on in the physical sciences on the status of hypothesis testing as a model of belief formation. Mayo (1996) and Mayo and Spanos (2006) have argued that there is a paradoxical reluctance of scientists to describe their own belief changes in terms of Neyman-Pearson testing. The key point here is that, if we, as social scientists, habitually use hypothesis testing to help form our own beliefs,[2] it would be natural to assume that our representative agents also use hypothesis tests.

A hypothesis test is a good model for belief formation of the kind described by Greenspan. In particular, until the rejection region is reached, *agents notice information without changing their minds*. We call our model 'inferential expectations' (IE) where 'inferential' refers to classical Neyman-Pearson inference.

We assume that, when a belief is overturned, agents switch to RE. Thus, RE is a special case of IE if agents are unconcerned about mistakenly changing their beliefs (the test size α equals unity). Seen in this way, α becomes a metric for rationality. When it is zero agents are completely unresponsive to evidence, and when it is unity they are completely untainted by belief conservatism. By linking RE to IE in this way, we ground our expectations theory ultimately in the structure of the model, which provides modeling discipline. Furthermore, if RE is defined in terms of the best unbiased estimator for a model variable or para-

[2] One author received criticism from a referee of a top-tier journal because significance was accorded to a test in the submitted paper with a p-value of 0.052.

meter – what we call feasible RE – then the model may be at the full information equilibrium in an expected value sense whilst exhibiting some volatility.[3]

In any IE model, there is a *cognitive target* about which *hypotheses* are tested using a *test statistic* based on a model variable, a *rejection region* and a *test size*.

This paper extends the theory of inferential expectations, outlined in Menzies and Zizzo (2009). In that paper the cognitive target was the money stock in a Dornbusch overshooting model, the hypotheses concerned the duration of that shock, the test statistic was the period of time the monetary expansion was observed and the rejection region was the time beyond which an observed expansion was believed to be permanent.

In this paper, the cognitive target is an economy wide real marginal cost shock, the hypotheses concern the size of this shock, the test statistic is a noisy signal and the rejection region comprises shocks that are 'too large' given beliefs about noise.

Treating beliefs as hypotheses tests in this context delivers two steady state features of micro level price movements suggested by the literature. First, a large fraction of price changes at the micro level are *negative* (Nakamura and Steinsson, 2008). Second, the absolute price changes are large (greater than 10 per cent) or small (less than 5 per cent), rather than medium (Klenow and Kryvtsov, 2008). Together, these imply price changes are somewhat 'tri-modal'. IE explains this by α per cent of agents mistakenly rejecting the null each period. If α is small this implies no change, or large (positive or negative) changes.

Furthermore, we build an IE Phillips curve and demonstrate that forming beliefs 'at the threshold' as opposed to 'at the margin' implies a distorted path on the adjustment toward equilibria. In particular, so called 'signal censoring' means that those agents in possession of extreme views (based on extreme draws of noise) are the first to 'get it'. The IE Phillips curve is state-dependent and offers an alternative explanation for price stickiness advanced in recent theoretical work, including Bonomo and Carvalho (2004), Caballero and Engel (2006, 2007), Caplin and Leahy (1997), Danziger (1999), Dotsey and King (2005), Dotsey et al. (1999), Golosov and Lucas (2007), John and Wolman (2008), Kiley (2000), Maćkowiak and Wiederholt (2007), Mankiw and Reis (2002, 2006, 2007), McAdam and Willman (2007), Reis (2006), and Woodford (2003, 2008a, 2008b).

In Section 2 we set up some basic notation. In section 3 we show the relationship between signal extraction and IE. In section 4 we model information sets of agents and consider when an 'information deluge' does, or does not, bring the economy closer to the full information equilibrium. In Section 5 we set up the basic monopolistic competition pricing model. Sections 6 and 7 derive the steady state price change distribution, and an IE Phillips curve. Section 8 examines signal censoring 'at the threshold', and Section 9 discusses the adjustment of inflation to a 'newsworthy' shock, drawing on the previous sections. Section

[3] This is related to the distinction made in modeling between rational expectations and perfect foresight, where the former departs from the latter by a zero mean error.

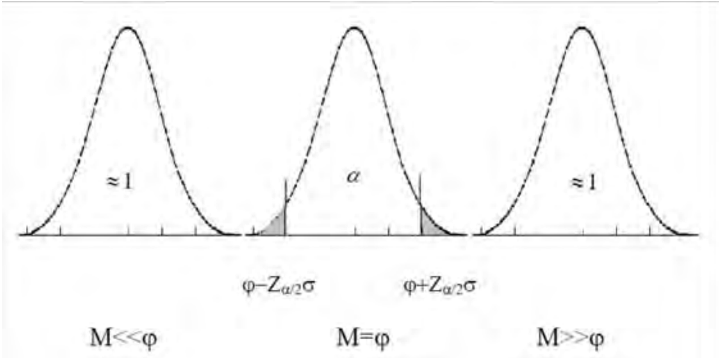

Figure 2:

10 concludes.

2 Notation and Basic Setup

We now introduce some notation and illustrate the possibility of sharp belief changes under IE. We do so in the simplest of informational environments, namely a single draw of a noisy signal.

In a standard way (Lucas, 1972), Nature draws an economically important parameter from a random variable $M \sim N\left(\phi, \sigma_M^2\right)$ and agents observe a noisy signal of this draw: $Y = M + v$, where $v \sim N\left(0, \sigma_v^2\right)$. Given the informational constraints, the feasible rational expectation of the draw M is the signal extraction solution:

$$E\left(M|Y\right) = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_M^2}\phi + \frac{\sigma_M^2}{\sigma_v^2 + \sigma_M^2}Y. \tag{1}$$

IE provides another guess of the cognitive target $E(M|Y)$. A reasonable H_0 belief is $E(M|Y) = \phi$. Conditioning on this, the hypotheses are:

$$\begin{aligned} H_0 &: M = \phi \\ H_1 &: M \neq \phi. \end{aligned} \tag{2}$$

The test statistic is based on Y, and α determines the rejection region as follows:

$$\text{Reject } H_0 \text{ if } \frac{|Y - \phi|}{\sigma_v} > Z_{\alpha/2}.$$

Figure 2 shows the distribution of Y for three very different draws of M and the probability that $H_0 : M = \phi$ is rejected, conditional on the M draw.

It is a matter of elementary definitions that the probability that H_0 is rejected equals the probability of a type I error α if H_0 is true, and, it is the probability complement of a type II error, namely $1 - \beta$ (also known as the power) if H_0 is false. Thus, if the truth is $M = \phi$, α per cent of the time agents find themselves in the shaded rejection regions. If, on the other hand, the true value of M is either very big or very small, the probability mass of $Y = M + v$ will be almost entirely to the right or left of the critical values. Equivalently, with these extreme draws of M the probability of a type II error is effectively zero, since it is extremely unlikely that H_0 will be believed. For values of M close to the critical values the probability that H_0 will be rejected will lie somewhere between 1 and α (not shown in Figure 2).

A key difference between signal extraction (1) and a hypothesis test (2) is the impact of an increment in Y on the beliefs held. With signal extraction:

$$\frac{dE\left(M|Y\right)}{dY} = \frac{\sigma_M^2}{\sigma_v^2 + \sigma_M^2} = \frac{1}{1+\delta}, \qquad \delta \equiv \frac{\sigma_v^2}{\sigma_M^2}. \tag{3}$$

Every increment in Y impacts upon the estimator, but discounted by δ. The greater the variance of the noise relative to the variance of M the more the influence of Y is discounted. The derivative is independent of the value of Y.

Under IE, increments in Y have no impact on beliefs up to the critical value because H_0 is maintained, viz. when the Greenspans of the world don't 'get it'.

$$\frac{dE\left(M|Y\right)}{dY} = 0 \qquad \text{if} \qquad \frac{|Y-\phi|}{\sigma_v} < Z_{\alpha/2} \tag{4}$$

At the critical value there is a jump to the RE belief, and $dE(M|Y)/dY = \infty$. For further increments in Y, within the rejection region, the derivative depends on what the RE belief is assumed to be.

If the variance of M is known, signal extraction is feasible and the derivative could be (3). However, we assume that hypothesis testing is a 'fast and frugal heuristic' (Gigerenzer et al., 1999) which avoids the need to form beliefs about M by conditioning on the H_0 value. We then define the RE belief to be Y, a parcel of information that must after all be available to agents in order to do the hypothesis test.[4]

Under this RE belief the hypothesis test in (2) is:

$$\begin{aligned} H_0 &: M = \phi \\ H_1 &: M \neq \phi \qquad \Longleftrightarrow \qquad M = Y \end{aligned} \tag{5}$$

and we have

$$\frac{dE\left(M|Y\right)}{dY} = \begin{cases} 0 & \text{if } \frac{|Y-\phi|}{\sigma_v} < Z_{\alpha/2} \\ \infty & \text{if } \frac{|Y-\phi|}{\sigma_v} = Z_{\alpha/2} \\ 1 & \text{if } \frac{|Y-\phi|}{\sigma_v} < Z_{\alpha/2} \end{cases} \tag{6}$$

[4] Giere (1988, cited in Mayo 1996 pg. 90) goes further than we need to. He suggests that even if the variance of M is known, signal extraction is unlikely to occur in practice: "Human beings are not naturally Bayesian information processors. And even considerable familiarity with probabilistic models seems not generally sufficient to overcome the natural judgment mechanisms, whatever they might be."

At the moment of rejection, the last increment in Y is the proverbial 'straw that broke the camel's back.' This is the feature of IE that can explain sharp asset price movements, such as those in Figure 1.

3 Signal Extraction and Expected IE Beliefs

Section 2 outlined the differences between the RE and IE solutions, but there is a relationship between the RE solution and the expected IE solution.

We define the Expected IE Belief (EIEB) as the expected value of the beliefs which will be held at the conclusion of the hypothesis test, where the expectation is taken over the distribution of M.

In principle, the EB requires consideration of the cases where H_1 is true and where H_1 is false. However, H_1 is always true in a probabilistic sense if M is continuous: the precise value of the mean under H_0 has measure (probability) zero.

The EB is therefore simply a convex combination of the beliefs implied by H_0 and H_1 where we may assume H_1 is true. Denoting the probability of a type II error as β we have

$$EIEB = \beta\,[H_0\text{-belief}] + (1-\beta)\,[H_1\text{-belief}] = \beta\phi + (1-\beta)\,Y, \tag{7}$$

where β depends on the rejection regions chosen in figure 2 as follows:

$$\beta = \Pr\left(\text{fail to reject } H_0\right) = \Pr\left(\frac{|Y-\phi|}{\sigma_v} \leq Z_{\alpha/2}\right).$$

When we allow for the variance of the M component of the variance of Y, σ_M^2, we can calculate this probability in terms of a χ^2-distribution, and the weights in (1) appear in the resultant expression

$$\begin{aligned}\beta &= \Pr\left(\frac{(Y-\phi)^2}{\sigma_v^2} \leq \chi_\alpha^2\right) \\ &= \Pr\left(\frac{(Y-\phi)^2}{\sigma_v^2+\sigma_M^2} \leq \frac{\sigma_v^2}{\sigma_v^2+\sigma_M^2}\chi_\alpha^2\right) \\ &= \Pr\left(\chi_1^2 \leq \frac{\sigma_v^2}{\sigma_v^2+\sigma_M^2}\chi_\alpha^2\right) \leq 1-\alpha.\end{aligned} \tag{8}$$

The intuition for (8) is provided by Figure 3. The larger σ_M^2 is relative to σ_v^2, the more the (bold) unconditional distribution stretches, and the smaller the probability mass in the acceptance region. β gets smaller and the weight in (7) shifts off ϕ onto Y, just as it does in (1).

The EIEB and the signal extraction solution are therefore both convex combinations of the mean of M (ϕ) and the signal Y. In both cases, the weight on ϕ is increasing and the weight on Y is decreasing in $\sigma_v^2/\left(\sigma_v^2+\sigma_M^2\right)$.

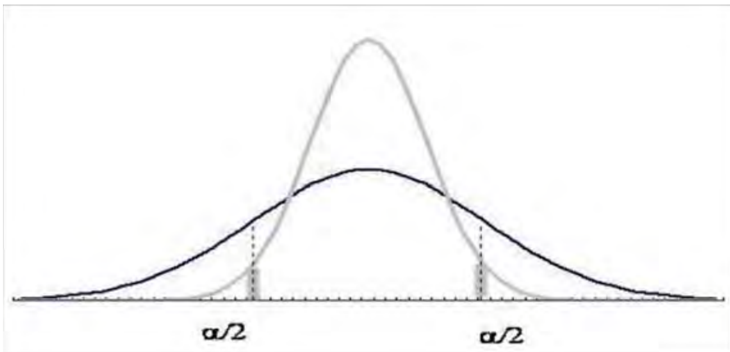

Figure 3:

The importance of this is that any environment in which signal extraction occurs over multiple agents, such as in Lucas (1972), can be recast into an IE framework, with the same qualitative interpretation. This is guaranteed by the appearance of $\sigma_v^2/\left(\sigma_v^2+\sigma_M^2\right)$ in both (1) and (8). The concept of an expected IE belief (EB) then just becomes the aggregation of beliefs over a large number of agents, prior to a draw on M or v.

4 Information Deluges

During times when the economy is newsworthy, such as crises or a change in policy direction, it is common for the quantity of information to increase rapidly. We refer to this as an information deluge. We need to make some analytic points in anticipation of the pricing model.

In what follows, we assume that agents form beliefs about a draw on M, denoted m, by accessing n different information sources each of which is a noisy signal. The set of all $m+v_j$ must be combined somehow to arrive at an estimate of m.

We allow the v_j to be dependent, and we let the common correlation across all shocks be ρ. The estimation of m is a GLS problem.

$$v=\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \qquad E\left(vv'\right)=\Omega=\sigma_v^2\begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \cdots & \cdots & 1 \end{bmatrix}.$$

$$
\begin{aligned}
y &= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} m + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \ell m + v \\
&\Longrightarrow \quad m_{GLS} = \left(\ell'\Omega^{-1}\ell\right)^{-1} \ell'\Omega^{-1} y,
\end{aligned}
$$

where

$$
\Omega^{-1} = \frac{1}{\sigma_v^2 \left[(1-\rho)(1-\rho+n\rho)\right]} \begin{bmatrix} 1+\rho(n-2) & -\rho & \cdots & -\rho \\ -\rho & 1+\rho(n-2) & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ -\rho & \cdots & \cdots & 1+\rho(n-2) \end{bmatrix}
$$

Now for any symmetric matrix A, the expression $\ell' A \ell$ just sums all the elements in the matrix, so

$$
\begin{aligned}
\ell'\Omega^{-1}\ell &= \frac{1}{\sigma_v^2 \left[(1-\rho)(1-\rho+n\rho)\right]} \left[n(1+\rho(n-2)) - \left(n^2-n\right)\rho\right] \\
&= \frac{n}{\sigma_v^2 (1-\rho+n\rho)}.
\end{aligned}
$$

$$
\Longrightarrow \left(\ell'\Omega^{-1}\ell\right)^{-1} = \frac{\sigma_v^2 (1-\rho+n\rho)}{n}.
$$

The expression for the GLS estimator is therefore the sample mean,

$$
\begin{aligned}
m_{GLS} &= \left(\ell'\Omega^{-1}\ell\right)^{-1} \ell'\Omega^{-1} y \\
&= \frac{\sigma_v^2 (1-\rho+n\rho)}{n} \frac{1}{\sigma_v^2 \left[(1-\rho)(1-\rho+n\rho)\right]} \\
&\quad \times \ell' \begin{bmatrix} 1+\rho(n-2) & -\rho & \cdots & -\rho \\ -\rho & 1+\rho(n-2) & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ -\rho & \cdots & \cdots & 1+\rho(n-2) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\
&= \frac{1}{n(1-\rho)} \left[1+\rho(n-2)-(n-1)\rho\right] \sum_{j=1}^{n} y_j = \overline{y}.
\end{aligned}
$$

Hence, in the analysis of sections 2 and 3, $Y = M + v$ could be replaced with $Y = M + \overline{v}$ if agents receive multiple noisy signals which they optimally combine with GLS.

However, although the estimator is the same as the standard case, the variance depends on the degree of correlation between the signals:

$$
\begin{aligned}
V(m_{GLS}) &= E(m - m_{GLS})(m - m_{GLS})' \\
&= E\left(\ell'\Omega^{-1}\ell\right)^{-1} \ell'\Omega^{-1} v v' \Omega^{-1}\ell \left(\ell'\Omega^{-1}\ell\right)^{-1} \\
&= \left(\ell'\Omega^{-1}\ell\right)^{-1} = \frac{\sigma_v^2 (1+\rho(n-1))}{n}
\end{aligned} \tag{9}
$$

Clearly, when $\rho = 0$ we obtain the textbook variance for a random sample, namely σ_v^2/n. The analysis of sections 2 and 3 holds in full, except that v becomes $\overline{v}$ and σ_v^2 becomes σ_v^2/n. An information deluge (where n approaches infinity) results in the elimination of all noise, unmasking the signal. The IE solution (and, for that matter, the signal extraction solution) converges to Y which is now a perfect measure of m.

However, it is more realistic to consider what happens when the correlation between parcels of information is not zero.

As ρ approaches unity (9) converges to σ_v^2, regardless of the size of n. This is identical to the variance of a single signal. The intuition is that if all information is perfectly correlated there is no point collecting more than one piece of essentially the same commodity.

To consider the asymptotic variance, note $(n-1)/n$ is close to unity for large n:

$$V\left(m_{GLS}\right) = \frac{\sigma_v^2\left(1+\rho\left(n-1\right)\right)}{n} \approx \sigma_v^2\left(\frac{1}{n}+\rho\right). \tag{9'}$$

Clearly, (9') does not converge to zero when ρ is non-zero. The limit as n approaches infinity is $\rho\sigma_v^2$. Thus, there is a floor on the variance of the feasible GLS estimator and any agent who rejects the null cannot expect to have a fully accurate belief about the true signal. Feasible RE cannot coincide with full information beliefs about the economy, no matter how much information is examined.

The economic intuition of this is that one can go so far by looking at extra stories on a Reuters screens. Market commentary may be influenced by common perceptions which lead to a correlation of the information.

Considering (9) and (9') together reveals that agents may be *worse* off in an information deluge if the correlation between (more abundant) signals increases. This would be economically important if a lot of market commentary at the time of newsworthy events is repetitive and uninformative. How much does the correlation have to change by to undo the advantage of an extra signal? Totally differentiate (9) and set it equal to zero.

$$\begin{aligned}\frac{dV\left(m_{GLS}\right)}{dn} &= \frac{\partial V}{\partial\rho}d\rho + \frac{\partial V}{\partial n}dn = \sigma_v^2\frac{\rho-1}{n^2}dn \equiv 0 \\ &\Longrightarrow \frac{\Delta\rho}{\Delta n} \approx \frac{1-\rho}{n\left(n-1\right)}.\end{aligned} \tag{10}$$

Setting $\Delta n = 1$, if ρ and n are large, any extra correlation between parcels of information need not be very large to undo the positive effects of one more signal information. For example, if ρ is 0.5 and n is 5, $\Delta\rho$ is 0.025. That is, if the correlation between parcels of information grew by more than 0.025 it would undo the variance benefits of one more signal.

Although IE grounds the model by jumping to the RE solution when H_0 is rejected, this section has shown that feasible RE need not be very close to the full information beliefs for individual agents if the signals collected by an individual agent are highly correlated. This result is robust to an information deluge that one might expect as the economy becomes 'newsworthy', provided that some correlation between news remains.

5 Pricing Setup

We assume a single producer of a final good in a centralized economy—the 'aggregator' demands a continuum of highly substitutable intermediate goods y_{it} on an index i over $[0,1]$ and combines them using CES production,[5]

$$y_t = \left[\int_0^1 y_{it}^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}, \qquad \theta > 1. \tag{11}$$

The aggregator solves the following problem:

$$\max \left\{ p_t y_t - \int_0^1 p_{it} y_{it} di \right\} \Longrightarrow y_{it} = y_t \left(\frac{p_{it}}{p_t} \right)^{-\theta} \equiv y_t \rho_{it}^{-\theta}, \tag{12}$$

where ρ_{it} is the relative price of the input. Combining (12) and (11) and cancelling y_t gives the price index

$$p_t^{1-\theta} = \int_0^1 p_{it}^{1-\theta} di. \tag{13}$$

We assume each intermediate firm has a production function such that real marginal cost is constant across all firms at time t, namely $rmc_t = rmc_{ss}(1+e_t)$ where ss refers to the steady state. Firms average n i.i.d. noisy signals of rmc_t, to obtain a statistic on rmc_t as in section 4. The idiosyncratic perception for firm i is $\widehat{rmc}_{it}$. In terms of the earlier notation, the draw from M is e_t and it has zero mean. We have:

$$\widehat{rmc}_{it} = \frac{\sum_{j=1}^{n} rmc_{ss} \left(1 + e_t + v_{ijt}\right)}{n} = rmc_{ss} \left(1 + e_t + \overline{v}_{it}\right).$$

We now assume that (9') is valid and that ρ is not too large.[6]

Firms conduct the following hypothesis test:

$$\begin{array}{lcl} H_0 : rmc = rmc_{ss} & \Longleftrightarrow & e_t = 0 \\ H_1 : rmc = \widehat{rmc}_{it} & \Longleftrightarrow & e_t \neq 0 \end{array}$$

If H_0 is believed, firms opt for the steady state version of real marginal cost, rmc_{ss}. If H_0 is rejected, firms use $\widehat{rmc}_{it}$.

[5] In this section ρ_{it} is not the same as ρ defined earlier. The un-subscripted ρ remains a correlation in what follows.

[6] This whole section could be written with independent observations, in which case the average of the v's would have variance σ_v^2/n. The generalization above relies on $(n-1)/n$ being close to unity (for (9') to be valid) and on the mutual correlation being small enough that the average of the v's converges to a Gaussian distribution.

The hypothesis test rejection region is:

$$\text{Reject } H_0 \text{ if } \left| \frac{\widehat{rmc}_{it} - rmc_{ss}}{rmc_{ss}\sigma\sqrt{\frac{1}{n}+\rho}} \right| > Z_{\alpha/2} \iff \left| \frac{e_t + \overline{v}_{it}}{\sigma\sqrt{\frac{1}{n}+\rho}} \right| > Z_{\alpha/2}$$

Firms need beliefs about real marginal costs to maximize profits. Firms choose p_{it} to maximize current profits.[7] Call mc_{it} nominal marginal cost. Firm i's real profits are given by

$$\Pi_i^{\text{real}} = \frac{\{p_{it} - mc_{it}\}\, y_{it}}{p_t}. \tag{14}$$

Let the relative price between the chosen and aggregate price in period t be ρ_{it} and real marginal costs be rmc. The aggregator's demand for y_{it} will be given by (12). Substituting this in for y_{it} in (14) gives:

$$\begin{aligned} \Pi_i^{\text{real}} &= (\rho_{it} - rmc_{it})\, y_t \rho_{it}^{-\theta} \\ &= \left(\rho_{it}^{1-\theta} - \rho_{it}^{-\theta} rmc_{it}\right) y_t. \end{aligned} \tag{15}$$

Differentiation w.r.t. ρ_{it} we obtain:[8]

$$\frac{d\Pi_i^{\text{real}}}{d\rho_{it}} = \left((1-\theta)\,\rho_{it}^{-\theta} + \theta\rho_{it}^{-(1+\theta)} rmc_{it}\right) y_t = 0,$$

which implies

$$\rho_{it} = \frac{\theta}{\theta - 1} rmc_{it}.$$

Or, equivalently,

$$p_{it} = \frac{\theta}{\theta - 1} rmc_{it} p_t. \tag{17}$$

We can now be specific about how agents calculate (17) under inferential expectations. They need beliefs about rmc_{it} and p_t. Under H_0 agents set rmc_{it} equal to its expected value rmc_{ss}. Furthermore, they set $p_t = p_{t-1}$ for simplicity in (17). Under H_1 agents work out the full model RE solution for p_t in (17) and use $\widehat{rmc}_{it}$ for rmc_{it}.

6 Steady State Analysis

Before building a Phillips curve, we will show that the H_0 and H_1 beliefs in (17) can deliver a 'tri-modal' distribution of price changes, as suggested in the

[7] Unlike the Calvo model, all firms are free to optimize. It is true that IE agents could predict non-optimization in the future if they do not fall in a rejection region, in which case a Calvo structure emerges, but we are more comfortable assuming our agents would 'wait and see'.

[8] The second order condition holds, and monopolistically competitive firms ignore their effect on y_t.

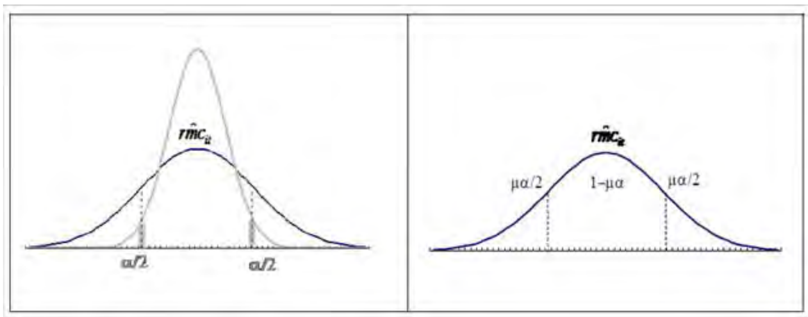

Figure 4:

literature. IE explains this by at least α per cent of agents mistakenly rejecting the null each period. If α is small this can imply no change, or large (positive or negative) changes.

We begin by assuming a stochastic steady state where inflation is zero.[9] Therefore $p_t = p_{t-1}$ and the only difference between H_0 and H_1 beliefs are then beliefs about rmc.

More subtly, we assume a constant fraction of agents reject the null for the consecutive periods where we calculate the distribution of price changes. From (8), we know the fraction of agents believing the null is no greater than $1 - \alpha$, and so we can write the fraction of agents *rejecting* the null as $\mu\alpha$ where $\mu \geq 1$. Finally, we have assumed that θ is large, so $\theta/(\theta - 1)$ is close to one. That being so, we have

$$\begin{aligned} \rho_{it} &\approx rmc_{ss} \quad \text{under } H_0, \\ \rho_{it} &\approx \widehat{rmc}_{it} \quad \text{under } H_1, \end{aligned}$$

which allows a direct graphical mapping of the distribution of $\widehat{rmc}_{it}$ onto the distribution of relative prices and relative price changes.

In figure 4 firm i receives a signal $\widehat{rmc}_{it}$. Under H_0 the sampling distribution of rmc is the taller *pdf* in the left panel, and it rejects H_0 in the tails. However, the actual probability of rejection is larger, because H_0 is false. The p-value of the shaded critical values, namely $\mu\alpha$ is shown on the right.

The distribution of (relative) prices can be derived easily from the above and is illustrated in Figure 5. A proportion $\mu\alpha$ fall in the two tails, and their estimate is actually $\widehat{rmc}_{it}$. So the distribution of prices has exactly the same *pdf* in that region. However, under H_0 the agents believe $\rho_t = rmc_{ss}$, so probability mass $1 - \mu\alpha$ is a discrete random variable at the single point rmc_{ss}.

[9] We will make use of a non-stochastic steady state to build the Phillips curve.

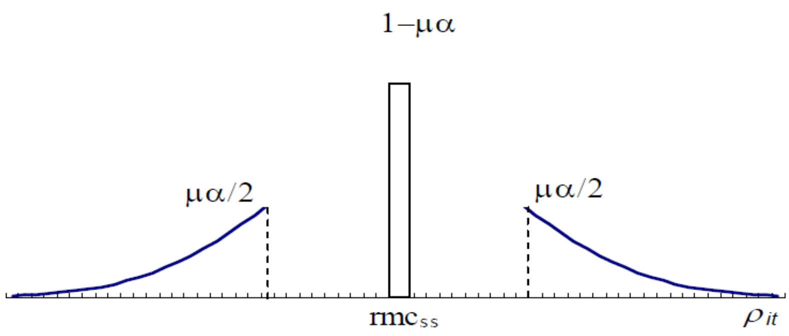

Figure 5:

We can now derive the distribution of price *changes* assuming that each probability mass in the above distribution is independently allocated across prices in the next period, in the proportions above. That is, the proportion of agents $\mu\alpha/2$ in, say, the left tail is divided up next period across the left, middle and right with proportions $(\mu\alpha/2)^2$, $(\mu\alpha/2)(1-\mu\alpha)$, $(\mu\alpha/2)(1-\mu\alpha)$, respectively. In the following table, the position of the firm in period t is read off the rows, and the position of the agent in period $t+1$ is read off the columns.

Table 1

Transition Matrix for Price Changes

	Left Tail_{t+1}	Middle_{t+1}	Right Tail_{t+1}
Left Tail_t	$(\mu\alpha)^2/4$	$(\mu\alpha/2)(1-\mu\alpha)$	$(\mu\alpha)^2/4$
Middle_t	$(1-\mu\alpha)(\mu\alpha/2)$	$(1-\mu\alpha)^2$	$(1-\mu\alpha)(\mu\alpha/2)$
Right Tail_t	$(\mu\alpha)^2/4$	$(\mu\alpha/2)(1-\mu\alpha)$	$(\mu\alpha)^2/4$

We denote an unchanged position between t and $t+1$ as '0', a shift up one region (from the left tail to the middle or from the middle to the right tail) as '+' and a shift from the left tail to the right tail as '++'. Using the same convention for '−' and '−−' we obtain a matrix of qualitative changes, summarized in table 2.

Table 2

Qualitative Transition Matrix for Price Changes

	Left Tail_{t+1}	Middle_{t+1}	Right Tail_{t+1}
Left Tail_t	0	+	++
Middle_t	−	0	+
Right Tail_t	−−	−	0

If α is small, '++' and '−−' are second-order small, and crossing from one tail to another becomes nigh impossible. Taking the limit as $\alpha \to 0$ for the proportions in Table 1 we use Table 2 to group a small-α distribution of *changes* between areas.

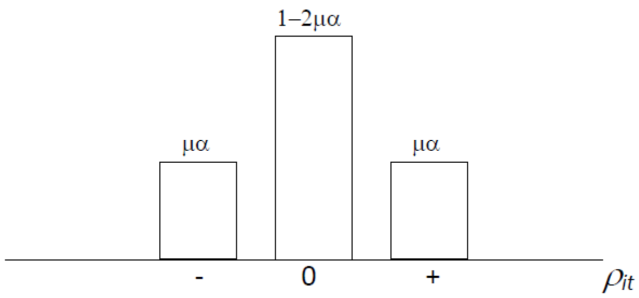

Figure 6:

Figure 6 gives area-to-area price changes related to Figure 5. That is, the proportion in '+' move from the left tail to the centre, or, from the centre to the right tail. The probability mass in the zero position is composed partly by agents who move literally zero between periods (they stay at the single point rmc_{ss} in the central column in figure 5) and those who stay in the same tail.

The key insight here is that if $\mu\alpha$ is small, because there is not much signal noise (μ is approximately unity) and agents exhibit belief conservatism (α is small), then agents who stay in the same tail in Figure 5 will not move very much in terms of actual price changes.[10] Thus, for a small $\mu\alpha$, there will be some agents with literally zero change (the ones in the central column in Figure 5), some with small changes (the ones who stay in the same tail in Figure 5), and some with large changes (since $\mu\alpha$ is small moving '+' or '−' involves moving a significant distance in Figure 5). Naturally, as $\mu\alpha$ grows these regularities dissolve, because the tails become so large that 'staying in the same tail' no longer implies small price changes.

Thus the model explains a tri-modal distribution under the assumption that agents are belief conservative and that underlying shocks to real marginal costs are not too large. This is in keeping with Nakamura and Steinsson (2008) and Klenow and Kryvtsov (2008).[11] IE also helps explain why sometimes these results might *not* be seen. For example, during times of structural change or rapid inflation, agents may lower their degree of belief conservatism, and $\mu\alpha$ may rise.

[10] As α approaches zero, the probability mass in an infinitesimally small neighbourhood of $\pm Z_{\alpha/2}\sigma_v$ converges to 1/2.

[11] Other papers studying micro-level pricing behavior include Bils and Klenow (2004), Boivin et al. (2009), Levy et al. (1997), Midrigan (2008), Rumler and Vilmunen (2005), Wolman (2007), and Zbaracki et al. (2004).

7 An IE Phillips Curve

Returning to the model (11) – (17) we condition on a particular shock e_t and derive an economy-wide Phillips curve. We break up the integral (13) into those firms that believe H_0 and those that believe H_1. Denote the proportion that reject H_0 as $\mu_t\alpha$, where the parameter μ_t is now conditioned on a particular economy-wide draw of e_t. (In section 6, it was not conditioned on a draw of e_t). We divide up firms on $[0,1]$ with $[0,1-\mu_t\alpha]$ indexing firms which believe H_0, and $[1-\mu_t\alpha,1]$ indexing those which believe H_1. From (13) the integral is

$$\begin{aligned} p_t^{1-\theta} &= \int_0^{1-\mu_t\alpha}\left(p_{it}^{H_0}\right)^{1-\theta}di+\int_{1-\mu_t\alpha}^{1}\left(p_{it}^{H_1}\right)^{1-\theta}di \\ &= \underbrace{(1-\mu_t\alpha)\int_0^1\left[\frac{\theta}{\theta-1}rmc_{ss}p_{t-1}\right]^{1-\theta}di}_{H_0\text{ beliefs}}+\underbrace{\int_{1-\mu_t\alpha}^{1}\left[\frac{\theta}{\theta-1}\widehat{rmc}_{it}p_t\right]^{1-\theta}di}_{H_1\text{ beliefs}} \\ &= (1-\mu_t\alpha)\left[\frac{\theta}{\theta-1}rmc_{ss}p_{t-1}\right]^{1-\theta}+\int_{1-\mu_t\alpha}^{1}\left[\frac{\theta}{\theta-1}rmc_{ss}\left(1+e_t+\overline{v}_{it}\right)p_t\right]^{1-\theta}di \\ &= (1-\mu_t\alpha)\left[\frac{\theta}{\theta-1}rmc_{ss}p_{t-1}\right]^{1-\theta}+\left[\frac{\theta}{\theta-1}rmc_{ss}p_t\right]^{1-\theta}\int_{1-\mu_t\alpha}^{1}\left(1+e_t+\overline{v}_{it}\right)^{1-\theta}di. \end{aligned}$$

We assume the average of the noise is small, so we can use the following approximation:

$$\begin{aligned} \int_{1-\mu_t\alpha}^{1}\left(1+e_t+\overline{v}_{it}\right)^{1-\theta}di &\approx \int_{1-\mu_t\alpha}^{1}\left(1+(1-\theta)\left(e_t+\overline{v}_{it}\right)\right)di \\ &= \mu_t\alpha\left(1+(1-\theta)e_t+(1-\theta)\int_0^1\overline{v}_{it}di\right). \end{aligned}$$

The last line uses the properties of Lebesgue integrals.[12]

We assume the integral of the noise is small and we therefore have

$$\begin{aligned} \int_{1-\mu_t\alpha}^{1}\left(1+e_t+\overline{v}_{it}\right)^{1-\theta}di &\approx \mu_t\alpha\left(1+e_t+\int_0^1\overline{v}_{it}di\right)^{1-\theta} \\ &\equiv \mu_t\alpha\left(1+e_t+\int\right)^{1-\theta}. \end{aligned}$$

[12] Consider $\int_{1-a}^{1}x(i)\,di$ where i is randomly assigned to all the $x(i)$'s. The Lebesgue integral is the supremum of all simple integrable function summations of the form $\sum_{j=1}^{k}X_j m(X_j)$ where $\cup_{j=1}^{k}m(X_j)=a$ and $0\leq X_j\leq x_j$. (See, for example, Capinski and Kopp, 2004, page 77). We re-write the summation as $a\sum_{j=1}^{k}X_j\frac{m(X_j)}{a}$ and $\frac{m(X_j)}{a}=M(X_j)$ where $\cup_{j=1}^{k}M(X_j)=1$. Therefore, $\int_{1-a}^{1}x(i)\,di=a\int_0^1 x(i)\,di$ and $M(\cdot)$ summing to unity implies that any Lebesgue integral summed over the index $[0,1]$ is an expected value, where the random variable is uniform with support $[0,1]$. We use this in section 8.

where the integral $\int$ of the noise in the range where agents reject H_0 will be discussed further momentarily. For now, we leave it as a symbol and derive the IE Phillips curve.

$$p_t^{1-\theta} = (1-\mu_t\alpha)\left[\frac{\theta}{\theta-1}rmc_{ss}p_{t-1}\right]^{1-\theta} + \mu_t\alpha\left[\frac{\theta}{\theta-1}rmc_{ss}p_t\right]^{1-\theta}\left(1+e_t+\int\right)^{1-\theta}.$$

Divide both sides by p_{t-1},

$$(1+\pi_t)^{1-\theta} = (1-\mu_t\alpha)\left[\frac{\theta}{\theta-1}rmc_{ss}\right]^{1-\theta} + \mu_t\alpha\left[\frac{\theta}{\theta-1}rmc_{ss}\left(1+\pi_t\right)\right]^{1-\theta}\left(1+e_t+\int\right)^{1-\theta}.$$

In the (non-stochastic) steady state $\pi = 0$ and all underlying shocks are set to zero which yields

$$\begin{aligned} 1 &= (1-\alpha)\left[\frac{\theta}{\theta-1}rmc_{ss}\right]^{1-\theta} + \alpha\left[\frac{\theta}{\theta-1}rmc_{ss}\right]^{1-\theta} \\ &\implies \frac{\theta}{\theta-1}rmc_{ss} = 1. \end{aligned}$$

This substitution is made, and the model is log linearized to deliver

$$\begin{aligned} \left[1+(1-\theta)\pi_t\right] &= (1-\mu_t\alpha) + \mu_t\alpha\left[1+(1-\theta)\pi_t\right]\left(1+(1-\theta)\left(e_t+\int\right)\right) \quad (18) \\ (1-\theta)\pi_t &= -\mu_t\alpha + \mu_t\alpha\left(1+(1-\theta)\left(\pi_t+e_t+\int\right)\right) \\ \pi_t &= \frac{\mu_t\alpha}{1-\mu_t\alpha}\left(e_t+\int\right). \end{aligned}$$

In (18) both $\mu_t\alpha$ and $\int$ are determined by e_t. The dependence of $\mu_t\alpha$ is obvious from figure 2: a large shock pushes most agents in the rejection region, and the dependence of $\int$ will be presently.

A proper understanding of (18) requires an economic interpretation of e_t. Recalling that $rmc_t = rmc_{ss}(1+e_t)$ it is apparent that e_t is the proportional deviation of real marginal cost from its steady state. Equation (18) therefore shows how cost shocks impact upon inflation.[13] Since agents do not care about aggregate output, we may scale (14) by y_t and rewrite their profit maximization problem as a marginal equivalence between revenue and cost scaled by output. Let $y = y_{it}/y_t$. Then

$$\Pi_i^* = \frac{\{p_{it}-mc_{it}\}\,y_{it}}{p_t y_t} = \rho_{it}\frac{y_{it}}{y_t} - rmc_{it}\frac{y_{it}}{y_t} = (\rho_{it}-rmc_{it})\,y = R(y) - C(y).$$

Use (12) to write $R(y)$ as a function of y, and to find marginal revenue, which is given by

$$\begin{aligned} R(y) &= \rho_{it}y = y^{-1/\theta}y = y^{\frac{\theta-1}{\theta}} \\ MR(y) &= \frac{dR(y)}{dy} = \frac{\theta-1}{\theta}y^{-1/\theta}. \end{aligned}$$

[13] It is standard to map deviations of rmc onto an output gap, though we do not do so here.

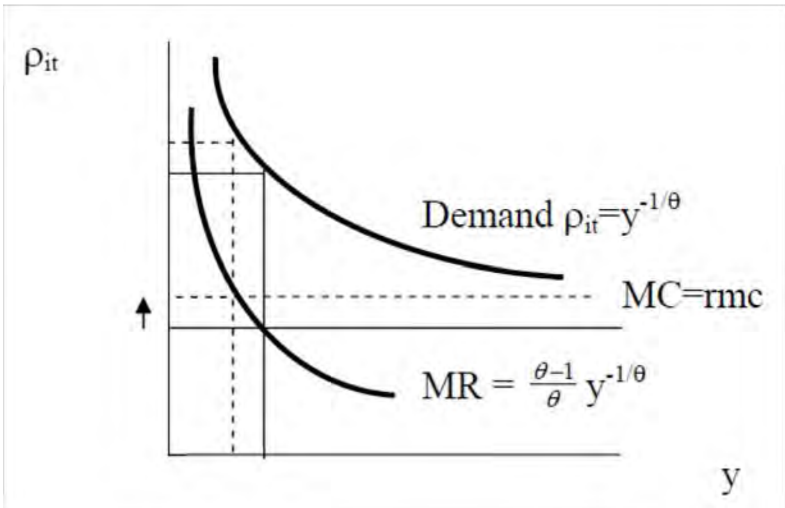

Figure 7:

Figure 7 shows price setting for a monopolistically competitive firm:

The axes shown are prices and quantities, relative to the aggregation of all other firms. Firms equate MR and MC taking into account the feasible relative price from the demand curve, which is derived from the aggregators demand for the intermediate inputs (12).

As shown by the dashed segments, if marginal cost rises the last unit of output becomes extra-marginal. The optimal response is to cut output relative to other firms and thereby raise prices relative to other firms. As the optimizing firms (those who reject H_0) do this, it feeds into the aggregate price level, and hence creates inflation.

According to (18) the extent of the increase in inflation is increasing in the share of firms who reject H_0 and optimize—that is, $\mu_t\alpha$. Since we are conditioning on a particular shock, figure 2 is the appropriate diagram. Clearly, large shocks will push a large probability mass into the rejection region. In the limit, as $\mu_t\alpha$ approaches unity, a shock to real marginal cost will, according to (18), have a very large impact on inflation.

The intuition is that increasing real marginal costs make firms want to cut back output *relative to other firms* and thereby raise prices *relative to the aggregate price index of all firms* ((13) and figure 7). Since an adjustment of *relative* prices is required, it becomes problematic if too many firms optimize. If virtually all firms increase prices in (13) then the aggregate rises by nearly the same amount, and further rises are required to take account of the very small measure of firms which do not change prices in (13). There is a runaway cost-push inflation spiral as most firms try to increase prices relative to everyone else.

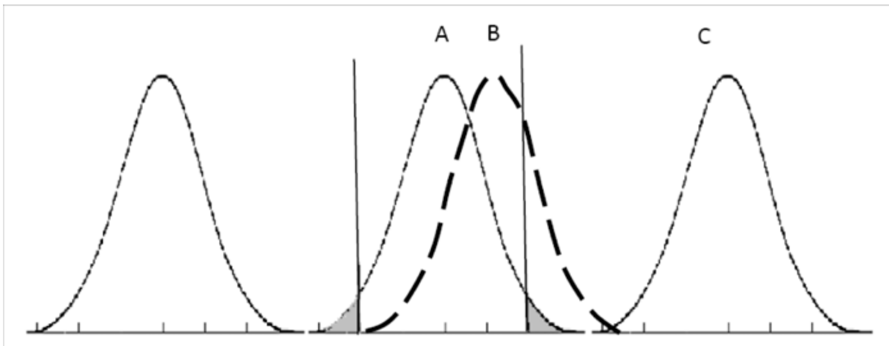

Figure 8:

Another feature of (18) is that if shocks to e_t are small, then $\mu_t\alpha$ will be low and stable. In these circumstances, (18) looks very much like a Calvo (1983) setup, where the Calvo Fairy's arrival frequency is dictated by the belief conservatism of firms parameterized by α. In a noisy world, agents must decide how seriously to take the occasional perverse signal. If they do not take it seriously, the proportion of firms changing prices each period will be low.

We emphasize a feature of (18) and IE more generally, which is that *agents sometimes make mistakes based on misreading information.* This seems an uncontroversial proposition to us, but it is surprisingly difficult to get optimizing models to deliver it.

8 Threshold Pricing: Signal Censoring

We now consider the final term in (18):

$$\pi_t = \frac{\mu_t\alpha}{1-\mu_t\alpha}\left(e_t + \int_0^1 \overline{v}_{it}di\right) \tag{18}$$

The integral is indexed over $[0,1]$ and it represents the probability weighted sum of all the $\overline{v}_{it}$'s observed by firms *who reject the null.* It is the average noise term, where a conditional distribution has been created by re weighting the measure by the probability that the null is rejected ($\mu_t\alpha$) (cf. footnote 16).

It turns out that this function is not analytic, but some qualitative results are possible. We proceed graphically. Consider figure 8 which is an expanded version of figure 2.

For a draw of e_t that is zero (A), the probability mass of rejecters is the shaded tails, $\mu_t = 1$ and so $\mu_t\alpha = \alpha$. The distribution of $\overline{v}$ is just the rejection tails of a normal scaled up so that each tail has probability one-half. Its expected value is clearly zero.

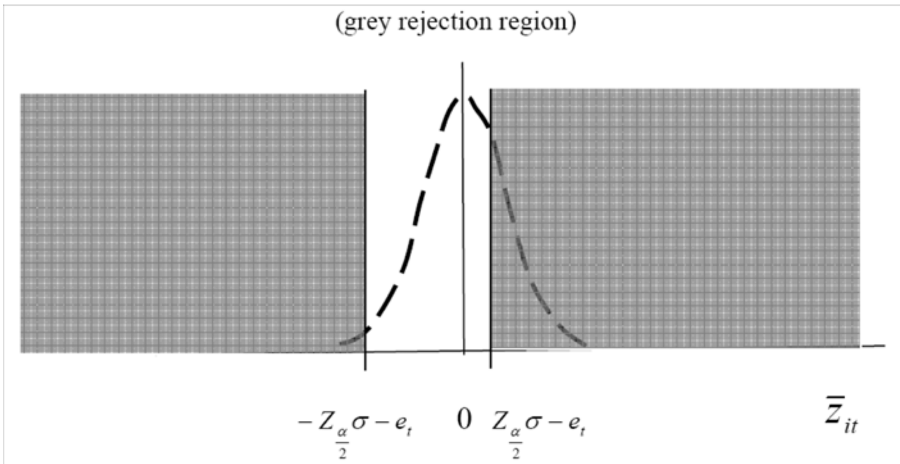

Figure 9:

For a small positive draw of e_t (B), the probability mass in the lower rejection region becomes very small, but the probability mass to the the upper rejection region increases rapidly. In Figure 8 $\mu_t\alpha$ looks like it is around 0.2. Therefore, the $\mu_t\alpha$-scaled *pdf* of $\overline{v}$ will consist of a large right tail and a small left tail both multiplied by 5 (1/0.2).

Finally, for a big enough shock to e_t (C) virtually all agents reject and $\mu_t\alpha$ is around 1. The distribution of $\overline{v}$ just becomes a normal (scaled up slightly because $\mu_t\alpha$ is not quite unity).

Economically, when the proportion of negative noise for optimizers differs from the proportion of positive noise, we say that there is noise adverse-selection. What is of concerning about this is that the *average* noise for those who reject H_0 departs from zero.

At distribution A there is no signal censoring because the rejecters are acting on the basis of an equal amount of positive and negative noise. However, for increasing values of et firms that optimize possess extreme positive noise (see B). Beyond a certain point, when virtually the entire probability mass is in the rejection region (say C), there is very little signal censoring because, again, rejecters are just as likely to have received positive or negative noise. For a given shock et the average noise is found by finding an expected value conditional on the draw of e_t. In Figure 9 we fix and show the distribution conditional on draws of e_t rather than sliding the distribution along the *rmc* axis by e_t, as in figure 8. The conditional distribution is the *pdf* scaled up outside of the interval $\left[-Z_{\alpha/2}\sigma - e_t, Z_{\alpha/2}\sigma - e_t\right]$, the scaled *pdf* in the grey rejection regions.[14]

[14] In this section σ is whatever the standard deviation of the noise is. If there is correlation, $(9')$ applies.

We now find the expected value of the noise, conditional on being in the grey regions. This is $\int_0^1 \overline{v}_{it}di$ and $\mu_t\alpha$ is our existing terminology for the probability of being in the grey region. A preliminary result required is the integral of $xf(x)$, where x is a normal random variable, over the range $[a, b]$:

$$\int_a^b x\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}}dx.$$

Let $y = x^2/2\sigma^2$ and rewrite the above as

$$\int_{\frac{a^2}{2\sigma^2}}^{\frac{b^2}{2\sigma^2}} \frac{\sigma}{\sqrt{2\pi}}e^{-y}dx = \frac{\sigma}{\sqrt{2\pi}}\left(e^{-\frac{a^2}{2\sigma^2}} - e^{\frac{-b^2}{2\sigma^2}}\right)$$

From Figure 9, the expected noise is

$$\int_0^1 \overline{v}_{it}di = \frac{1}{\mu_t\alpha}\left(\int_{-\infty}^{-Z_{\alpha/2}\sigma-e_t} x\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}}dx + \int_{Z_{\alpha/2}\sigma-e_t}^{\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}}dx\right).$$

Using the above result, this simplifies to

$$\int_0^1 \overline{v}_{it}di = \frac{1}{\mu_t\alpha}\frac{\sigma}{\sqrt{2\pi}}\left(e^{-\frac{\left(Z_{\alpha/2}\sigma-e_t\right)^2}{2\sigma^2}} - e^{\frac{\left(Z_{\alpha/2}\sigma+e_t\right)^2}{2\sigma^2}}\right)$$

$$\mu_t\alpha = 1 + cdf\left(Z_{\alpha/2} - \frac{e}{\sigma}\right) - cdf\left(-\left(Z_{\alpha/2} + \frac{e}{\sigma}\right)\right)$$

where $cdf(\cdot)$ is the cumulative density function for a $N(0,1)$ and $Z_{\alpha/2} = cdf^1(1-\alpha/2)$. The $1/\mu_t\alpha$ term is not analytic. So the function is plotted for $\sigma = 0.01$ and $\alpha = 0.05$ in figure 10. A similar shape occurs for other parameter values.

Figure 10 shows the expected noise term for firms for given shocks to real marginal costs. Signal censoring is not a problem for very small shocks (in the immediate neighbourhood of zero) or very large ones (more than 3 standard deviations, e.g. ± 0.03 in figure 10). In the latter case, virtually the whole distribution of *rmc* in figure 2 lies in the rejection region and therefore positive and negative noise is almost equally prevalent.

A technically reassuring point from calibrations based on (19) is that the size of the integral is fairly small, justifying the derivation of (18).

The economic point of this section and figure 10 is that signal censoring will magnify the effects of marginal cost shocks on inflation. Firms in possession of extreme signals (driven by large noise terms – either positive or negative) reject before other firms. Thus, the integral term in (18) is positively related to e_t and reinforces the effect of e_t on π_t.

This point is made in reference to our model of inflation, but it is likely to apply to other IE models 'at the threshold'.

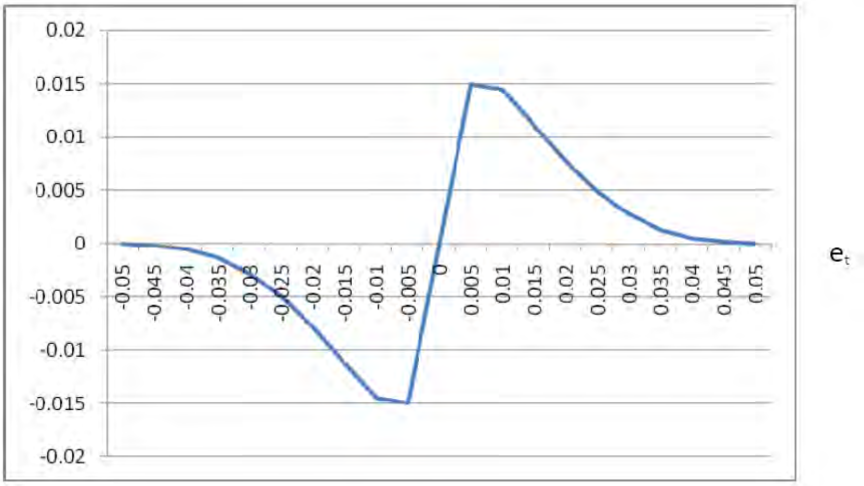

Figure 10:

9 A Newsworthy Shock

All the analytic apparatus has now been assembled to trace through the full effects of a shock to real marginal cost. To make use of figure 10, we assume that the agents each receive 20 noisy signals, with noise standard deviation of 0.02 and correlation of 0.2. From (9'), agents average signals as dictated by GLS to arrive at a noise standard deviation σ of $0.02(1/20 + 0.2)^{0.5} = 0.01$, as in figure 10. We shall assume the shock to real marginal cost, e_t, is equal to 2 per cent, or 0.02.

When the shock hits, the distribution of the noisy signal shifts to the right (figure 2) and a fraction of agents reject the null. For these parameters, the fraction is 0.5160 (from the expression under (19), where σ is understood to be the standard deviation inclusive of correlation as in (9')). The fraction which rejects H_0 makes up a 'multiplier' in front of the square brackets in (18).

$$\pi_t = \pi_t = \frac{\mu_t \alpha}{1 - \mu_t \alpha} \left(e_t + \int_0^1 \overline{v}_{it} di \right)$$

The multiplier here is $0.516/(1 - 0.516) = 1.0660$, or, approximately unity.

As described at the end of section 7, the H_1 firms who optimize reduce their extra marginal production to raise their prices relative to the aggregate price. With just under half the firms holding onto H_0, the optimizers can quite easily achieve the increase in relative price without much inflation; there is inertia in the aggregate price index (13) because the H_0-agents—half the total—leave their prices unchanged. As a result of these effects, inflation rises to just over 2 per cent (0.02×1.066).

However, signal censoring adds around 0.8 of a percentage point onto headline inflation. (See the last term in (18)). From figure 10, a positive shock of 0.02 leads to an average noise term of 0.0077. As seen in both figures 2 and 9, the positive shock means that H_1-agents are more influenced by positive noise 'at the threshold' than by negative noise. Given the multiplier of 1.0660, this extra effect transmits into inflation virtually one-for-one.

What now happens if there is an information deluge as a result of the shock? There are many degrees of freedom in this setup, so there are many possible outcomes. We will consider two scenarios.

Suppose that there are now 1000 pieces of information, so that $1/n$ effectively vanishes, and that their correlation drops to 0.14. With these parameters, the noise standard deviation drops to $0.02(0.14)^{0.5} = 0.0075$. Agents realize this and draw in their rejection regions closer to $e_t = 0$.

For a shock of 0.2, many agents—76 per cent—optimize. The equivalent of figure 10 drawn with noise standard deviation of 0.0075 leads to somewhat less signal censoring—the expected noise drops from half a percentage point to approximately 0.3 of a percentage point. However, the bad news is that with the higher proportion of firms optimizing the multiplier $\mu_t\alpha/(1-\mu_t\alpha)$ rises to over 3.1. A cost-push inflation (described at the end of section 7) ensues and inflation rises to over 6 per cent.

Now let us consider the case where the deluge is extremely uninformative. If the correlation between the signals rises to unity, the variance of the noise becomes 0.02 (regardless of how many observations there are – see (9)). The proportion rejecting H_0 drops to 0.17, and the multiplier becomes approximately 0.2.

The signal censoring becomes very important here. As seen in the figure 11 (figure 10 with variance of 0.02), nearly 3 percentage points are added to the term in square brackets in (18). When the 5 per cent (3 from signal censoring plus 2 from e_t) in the square brackets pass through the multiplier of 0.2, the net outcome is an inflation rate of 1 per cent.

This section has shown that signal censoring could potentially add to inflation. It would be interesting to explore any links between signal censoring and the 'missing inflation rates' between moderate inflation and hyperinflation.

We leave to future research the question of the impact on inflation if price signals become extremely noisy. Could the optimizers then provide an impetus towards hyperinflation due to signal censoring?

10 Conclusion

In this paper we have extended the theory of inferential expectations, outlined in Menzies and Zizzo (2009) by building an IE model of pricing.

We have shown that belief conservatism, parameterized by a low α, can deliver a steady state 'tri-modal' distribution of price changes (Nakamura and Steinsson, 2008, and, Klenow and Kryvtsov, 2008). IE explains this by α per

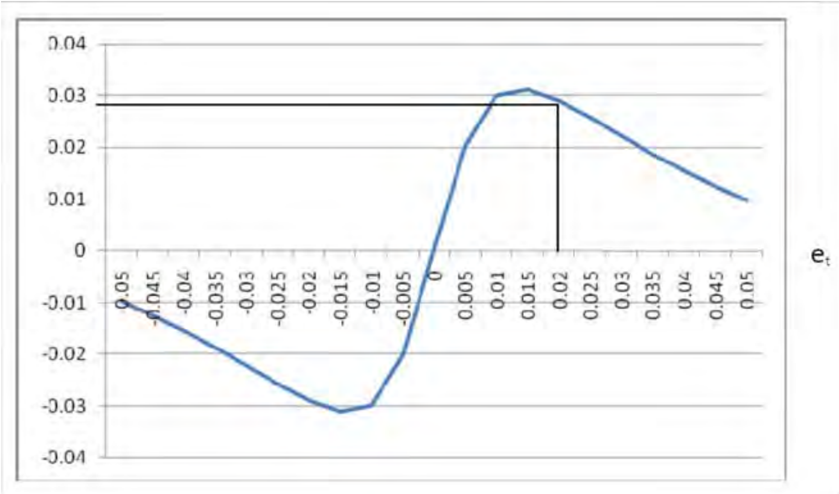

Figure 11:

cent of agents mistakenly rejecting the null each period. If α is small this implies no change, or large (positive or negative) changes.

Out of the steady state, our IE Phillips curve shows that signal censoring can lead to a distorted adjustment paths because the agents in possession of extreme views (based on extreme draws of noise) are the first to change beliefs and 'get it' when an underlying shock hits the economy.

We recognize that our analysis has been essentially static. It has been important to establish the link between signal extraction and expected IE beliefs, but future research will consider the build up of information in more dynamic models, and the potential to estimate IE parameters, such as the degree of belief conservatism.

Some of our results have been general, rather than specific. The relationship between signal extraction and IE bodes well for the latter's future modelling usefulness, as does the analysis of informative and uninformative 'information deluges'.

References

[1] Akerlof, George A. and Janet L. Yellen (1985), "Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?" *American Economic Review*, Vol. 75, 708-720.

[2] Bils, Mark, and Peter J. Klenow (2004). "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, Vol. 112, No. 5, 947-985.

[3] Boivin, Jean, Marc P. Giannoni, and Ilian Mihov (2009). "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data," *American Economic Review*, Vol. 99, No. 1, 350-384.

[4] Bonomo, Marco, and Carlos Carvalho (2004). "Endogenous Time-Dependent Rules and Inflation Inertia," *Journal of Money, Credit and Banking*, Vol. 36, No. 6, 1015-41.

[5] Brunnermeier, Markus K., and Jonathan A. Parker (2005), "Optimal Expectations," *American Economic Review*, Vol. 95, 1092-1118.

[6] Caballero, Ricardo J., and Eduardo Engel (2006). "Price Stickiness in Ss Models: Basic Properties," *mimeo*, MIT.

[7] Caballero, Ricardo J., and Eduardo Engel (2007). "Price Stickiness in Ss Models: New Interpretations of Old Results," *Journal of Monetary Economics*, Vol. 54 (supp. 1), 100-121.

[8] Calvo, Guillermo A. (1983). "Staggered Prices In a Utility-Maximizing Framework," *Journal of Monetary Economics*, Vol. 12, No. 3, 383-398.

[9] Caplin, Andrew, and John Leahy (1997). "Aggregation and Optimization with State-Dependent Pricing," *Econometrica*, Vol. 65, 601-625.

[10] Carroll, Christopher D. (2003), "Macroeconomic Expectations of Households and Professional Forecasters," *Quarterly Journal of Economics*, 118, 269-298.

[11] Danziger, Leif (1999). "A Dynamic Economy with Costly Price Adjustments," *American Economic Review*, Vol. 89, 878-901.

[12] Dotsey, Michael, and Robert G. King (2005). "Implications of State-Dependent Pricing for Dynamic Macroeconomic Models," *Journal of Monetary Economics*, Vol. 52, 213-242.

[13] Dotsey, Michael, Robert G. King, and Alexander L. Wolman (1999). "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics*, Vol. 114, No. 2, 655-90.

[14] Evans, George W., and Seppo Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.

[15] Gertler, Mark, and John Leahy (2008). "A Phillips Curve with an Ss Foundation," *Journal of Political Economy*, Vol. 116, No. 3, 533-572.

[16] Gigerenzer, Gerd, P. M. Todd, and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart.* Oxford: Oxford University Press.

[17] Golosov, Mikhail, and Robert E. Lucas Jr. (2007). "Menu Costs and Phillips Curves," *Journal of Political Economy*, Vol. 115, No. 2, 171-199.

[18] Hansen, Lars. P., and Thomas J. Sargent (2001), "Robust Control and Model Uncertainty," *American Economic Review Papers and Proceedings*, Vol. 91, 60-66.

[19] John, A. Andrew, and Alexander L. Wolman (2008). "Steady-state equilibrium with state-dependent pricing," *Journal of Monetary Economics*, Vol. 55, 383-405.

[20] Kiley, Michael T. (2000). "Endogenous Price Stickiness and Business Cycle Persistence," *Journal of Money, Credit, and Banking*, Vol. 32, No. 1, 28-53.

[21] Klenow, Peter J., and Oleksiy Kryvtsov (2008). "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" *Quarterly Journal of Economics*, Vol. 123, No. 3, 863-904.

[22] Levy, Daniel, Mark Bergen, Shantanu Dutta, and Robert Venable (1997). "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," *Quarterly Journal of Economics*, Vol. 112(3), 791-825.

[23] Lucas, Robert E. Jr. (1972). "Expectations and the Neutrality of Money," *Journal of Economic Theory*, Vol. 4, 103-124.

[24] Maćkowiak, Bartosz, and Mirko Wiederholt (2007). "Optimal Sticky Prices under Rational Inattendance," *mimeo*, Northwestern University, March.

[25] Mankiw, N. Gregory, and Ricardo Reis (2002). "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, Vol. 117, No. 471, 1295-1328.

[26] Mankiw, N. Gregory, and Ricardo Reis (2006). "Pervasive Stickiness," *American Economic Review*, Vol. 96, No. 2, 164-169.

[27] Mankiw, N. Gregory, and Ricardo Reis (2007). "Sticky Information in General Equilibrium," *Journal of the European Economic Assocation*, Vol. 5, Nos. 2-3, 603-613.

[28] Mayo, Deborah G., (1996). *Error and the Growth of Experimental Knowledge.*

[29] Mayo, Deborah G., and Aris Spanos (2006). "Severe Testing as a Basic Concept in a Neyman-Pearson Philosophy of Induction," *The British Journal for the Philosophy of Science*, 1-35.

[30] McAdam, Peter, and Alpo Willman (2007). "State-Dependency and Firm-Level Optimization - A Contribution to Calvo Price Setting," ECB Working Paper No. 806, Frankfurt.

[31] Menzies, Gordon D. and Daniel J. Zizzo (2005), "Inferential Expectations", Australian National University Centre for Applied Macroeconomic Analysis Discussion Paper No. 12, June.

[32] Menzies, Gordon D., and Daniel J. Zizzo (2007), "Exchange Rate Markets and Conservative Inferential Expectations", Australian National University Centre for Applied Macroeconomic Analysis Discussion Paper No. 2, January.

[33] Menzies, Gordon D., and Daniel J. Zizzo (2009), "Inferential Expectations," *The B.E. Journal of Macroeconomics: 9(1) (Advances),* http://www.bepress.com/bejm/vol9/iss1/art42.

[34] Midrigan, Virgiliu (2008). "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *mimeo*, New York University.

[35] Nakamura, Eki, and Jón Steinsson (2008). "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, Vol. 123, No. 4, 1415-1464.

[36] Reis, Ricardo (2006). "Inattentive Producers," *Review of Economic Studies*, Vol. 73, No. ??, 793-821.

[37] Rumler, F., and J. Vilmunen (2005). "Price setting in the euro area: some stylized facts from micro consumer data," ECB Working Paper #524.

[38] Sargent, Thomas J. (1993). *Bounded Rationality in Macroeconomics*. New York: Oxford University Press.

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

[40] Wolman, Alexander L. (2007). "The Frequency and Costs of Individual Price Adjustment," *Managerial and Decision Economics*, Vol. 28, 531-552.

[41] Woodford, Michael (2003). "Imperfect Common Knowledge and the Effects of Monetary Policy," in: Aghion, Philippe, Roman Frydman, Joseph Stiglitz, and Michael Woodford (eds.). *Knowledge, Information, and Expectations in Modern Macroeconomics*. Princeton: Princeton University Press.

[42] Woodford, Michael (2008a). "Inattention as a Source of Randomized Discrete Adjustment," *mimeo*, Columbia University, April.

[43] Woodford, Michael (2008b). "Information-Constrained State-Dependent Pricing," NBER Working Paper No. 14620, December.

[44] Zbaracki, Mark, Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen (2004). "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *Review of Economics and Statistics*, Vol. 86, 514-533.