AMBIGUITY, INFORMATION QUALITY AND ASSET PRICING*

Larry G. Epstein Martin Schneider

July 24, 2006

Abstract

When ambiguity-averse investors process news of uncertain quality, they act as if they take a worst-case assessment of quality. As a result, they react more strongly to bad news than to good news. They also dislike assets for which information quality is poor, especially when the underlying fundamentals are volatile. These effects induce ambiguity premia that depend on idiosyncratic risk in fundamentals as well as skewness in returns. Moreover, shocks to information quality can have persistent negative effects on prices even if fundamentals do not change.

1 INTRODUCTION

Financial market participants absorb a large amount of news, or signals, every day. Processing a signal involves quality judgments: news from a reliable source should lead to more portfolio rebalancing than news from an obscure source. Unfortunately, judging quality itself is sometimes difficult. For example, stock picks from an unknown newsletter without a track record might be very reliable or entirely useless – it is simply hard to tell. Of course, the situation is different when investors can draw on a lot of experience that helps them interpret signals. This is true especially for "tangible" information, such as earnings reports, that lends itself to quantitative analysis. By looking at past data, investors may become quite confident about how well earnings forecast returns.

*Epstein: Department of Economics, U. Rochester, Rochester, NY, 14627, lepn@troi.cc.rochester.edu; Schneider: Department of Economics, NYU, and Federal Reserve Bank of Minneapolis, ms1927@nyu.edu.
We are grateful to Robert Stambaugh (the editor) and an anonymous referee for very helpful comments. We also thank Monika Piazzesi, Pietro Veronesi, as well as seminar participants at Cornell, IIES (Stockholm), Minnesota, the Cowles Foundation Conference in honour of David Schmeidler, and the SAMSI workshop on model uncertainty for comments and suggestions. This paper reflects the views of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

This paper focuses on information processing when there is incomplete knowledge about signal quality. The main idea is that, when quality is difficult to judge, investors treat signals as *ambiguous.* They do not update beliefs in standard Bayesian fashion, but behave as if they have multiple likelihoods in mind when processing signals. To be concrete, suppose that θ is a parameter that an investor wants to learn. We assume that a signal s is related to the parameter by a *family of likelihoods*:

$$s = \theta + \epsilon, \qquad \epsilon \sim \mathcal{N}\left(0, \sigma_s^2\right), \qquad \sigma_s^2 \in \left[\underline{\sigma}_s^2, \overline{\sigma}_s^2\right]. \tag{1}$$

The Bayesian approach is concerned with the special case of a single likelihood, $\underline{\sigma}_s^2 = \overline{\sigma}_s^2$, and measures the quality of information via the signal precision $1/\sigma_s^2$. In our model, information quality is captured by the *range of precisions* $[1/\overline{\sigma}_s^2, 1/\underline{\sigma}_s^2]$. The quality of ambiguous signals therefore has two dimensions.

To model preferences (as opposed to merely beliefs), we use recursive multiple-priors utility, axiomatized in Epstein and Schneider [12]. The axioms describe behavior that is consistent with experimental evidence typified by the Ellsberg Paradox. They imply that an ambiguity averse agent behaves *as if* he maximizes, every period, expected utility under a worst-case belief that is chosen from a *set* of conditional probabilities. In existing studies, the set is typically motivated by agents' *a priori* lack of confidence in their information. In this paper, it is instead derived explicitly from information processing – its size thus depends on information quality. In particular, we present a thought experiment to show that ambiguity-averse behavior can be induced by poor information quality alone: an *a priori* lack of confidence is not needed.

Ambiguous information has two key effects. First, *after* ambiguous information has arrived, agents respond asymmetrically: bad news affect conditional actions – such as portfolio decisions – more than good news. This is because agents evaluate any action using the conditional probability that minimizes the utility of that action. If an ambiguous signal conveys good (bad) news, the worst case is that the signal is unreliable (very reliable). The second effect is that even *before* an ambiguous signal arrives, agents who anticipate the arrival of low quality information will dislike consumption plans for which this information may be relevant. This intuitive effect does not obtain in the Bayesian model, which precludes any effect of *future* information quality on *current* utility.[1]

To study the role of ambiguous information in financial markets, we consider a representative agent asset pricing model. The agent's information consists of (i) past prices and dividends and (ii) an additional, ambiguous, signal that is informative about future dividends. Our setup thus distinguishes between *tangible* information – here dividends – that lends itself to econometric analysis, and *intangible* information – such as news reports – that is hard to quantify, yet important for market participants' decisions. We assume that intangible information is ambiguous while tangible information is not. This approach generates several properties of asset prices that are hard to explain otherwise.

[1] Indeed, the law of iterated expectations implies that conditional expected utility is not affected by changes in the precision of future signals about consumption, as long as the distribution of future consumption itself does not change. In contrast, ambiguity-averse agents fear the discomfort caused by future ambiguous signals and their anticipation of low quality information directly lowers current utility.

In markets with ambiguous information, expected excess returns decrease with *future* information quality. Indeed, ambiguity averse investors require compensation for holding an asset simply because low quality information about that asset is expected to arrive. This result cannot obtain in a Bayesian model where changes in future information quality are irrelevant for current utility and hence for current asset prices. It implies that conclusions commonly drawn from event studies should be interpreted with caution. In particular, a negative abnormal return need not imply that the market views takes a dim view of fundamentals. Instead it might simply reflect the market's discomfort in the face of an upcoming period of hard-to-interpret, ambiguous information. To illustrate this effect, we provide a calibrated example that views September 11, 2001 as a shock that not only increased uncertainty, but also changed the nature of signals relevant for forecasting fundamentals. We show that shocks to information quality can have drawn out negative effects on stock prices even if fundamentals do not change.

Expected excess returns in our model increase with *idiosyncratic* volatility in fundamentals. It is natural that investors require more compensation for poor information quality when fundamentals are more volatile. Indeed, in markets where fundamentals do not move much to begin with, investors do not care whether information quality is good or bad, so that premia should be small even if information is highly ambiguous. In contrast, when fundamentals are volatile, information quality is more of a concern and the premium for low quality should be higher. What makes ambiguity premia different from risk premia is ambiguity averse investors' first-order concern with uncertainty: an asset that is perceived as more ambiguous is treated *as if* it has a lower mean payoff. This is why expected excess returns in our model depend on total (including idiosyncratic) volatility of fundamentals, and not on covariance with the market or marginal utility.

Ambiguous information also induces skewness in measured excess returns. Indeed, the asymmetric response to ambiguous information implies that investors behave as if they "overreact" to bad intangible signals. At the same time, they appear to "underreact" to bad tangible signals. This is because investors do not perceive ambiguity about (tangible) dividends per se, but only about (intangible) signals that are informative about future dividends. The arrival of tangible dividend information thus tends to "correct" previous reactions due to intangible signals. Overall, the skewness of returns depends on the relative importance of tangible and intangible information in a market: negative skewness should be observed for assets about which there is relatively more intangible information. This is consistent with the data: individual stocks that are "in the news" more, such as stocks of large firms, have negatively skewed returns, while stocks of small firms do not.

The volatility of prices and returns in our model can be much larger than the volatility of fundamentals. Volatility depends on how much the worst-case conditional expectation of fundamentals fluctuates. If the range of precisions contemplated by ambiguity averse agents is large, they will often attach more weight to a signal than agents who know the true precision. Intangible information can thus cause large price fluctuations. In addition, the relationship between information quality and the volatility of prices and returns is different from that with risky (or noisy) signals. West [27] has shown that, with higher precision of noisy signals, volatility of prices increases but the volatility of returns decreases. In our framework, information quality can also change if signals become less ambiguous. Changes in information quality – due to improvements in information technology, for example – can then affect the volatility of prices and returns in the same direction.

Our asset pricing results rely on the distinction between tangible and intangible information emphasized by Daniel and Titman [11]. Most equilibrium asset pricing models assume that all relevant information is tangible – prices depend only on past and present consumption or dividends. An exception is Veronesi [26], who has examined the effect of information quality on the equity premium in a Lucas asset pricing model that also features an intangible (but unambiguous) signal. His main result is that, with high risk aversion and a low intertemporal elasticity of substitution, there is no premium for low information quality in a Bayesian model. Another exception is the literature on overconfidence as a source of overreaction to signals and excess volatility (for example, Daniel et al [10]). In these models, an investor's perceived precision of an intangible (private) signal is higher than the "true" precision, which makes reactions to signals more aggressive than under rational expectations.[2]

A large literature has explored the effects of Bayesian learning on excess volatility and in-sample predictability of returns. Excess volatility arises from learning when agents' subjective variance of dividends is higher than the true variance, which induces stronger reactions to news.[3] Our model is different in that the subjective variance of dividends is equal to the *true* variance. Our model thus applies also when the distribution of (tangible) fundamentals is well understood, as long as ambiguous *intangible* signals are present. Finally, the mechanism generating skewness in our model differs from that in Veronesi [25], who shows, in a Bayesian model with risk averse agents, that prices respond more to bad news in good times and conversely. Our result does not rely on risk aversion and is therefore relevant also if uncertainty is idiosyncratic and investors are well diversified. Moreover, ambiguous signals entail an asymmetric response whether or not times are good.

The paper is organized as follows. Section 2 introduces our model of updating with ambiguous information. Section 3 discusses a simple representative agent model and derives its properties. Here we also contrast the Bayesian and ambiguity aversion approaches to thinking about information quality and asset pricing. Section 4 considers the calibrated model of 9/11 as an example of shocks to information quality. Proofs are collected in an appendix.

[2]Importantly, overconfidence and ambiguity aversion are not mutually exclusive. A model of overconfident, ambiguity averse agents would assume that agents are uncertain about precision, but the true precision lies close to (or even below) the lower bound of the range.

[3]See Timmermann [23, 24], Bossaerts [3], and Lewellen and Shanken [20] for models of nonstationary transitions and Brandt, Xeng, and Zhang [4], Veronesi [25] and Brennan and Xia [6] for models with persistent hidden state variables. A related literature has tried to explain post-event abnormal returns ("underreaction") through the gradual incorporation of information into prices (see Brav and Heaton [5] for an overview).

2 AMBIGUOUS INFORMATION

In this section, we first propose a thought experiment to illustrate how ambiguous information can lead to behavior that is both intuitive and inconsistent with the standard expected utility model. While the experiment is related to the static Ellsberg Paradox, it is explicitly dynamic and focuses on information processing. We then present a simple model of updating with multiple normal distributions, already partly described in the introduction, that is the key tool for our applications. Finally, we discuss the axiomatic underpinnings of our approach as well as its connection to the more general model of learning under ambiguity introduced in Epstein and Schneider [14].

2.1 A Thought Experiment

Consider two urns that have been filled with black and white balls as follows. First, a ball is placed in each urn according to the outcome of a fair coin toss. If the coin toss for an urn produces heads, the "coin ball" placed in that urn is black; it is white otherwise. The coin tosses are independent across urns. In addition to a coin ball, each urn contains n "non-coin balls", of which exactly $\frac{n}{2}$ are black and $\frac{n}{2}$ are white. For the first urn, it is known that $n = 4$: there are exactly two black and two white non-coin balls. Since the description of the experiment provides objective probabilities for the composition of this urn, we refer to it as the *risky* urn.

In contrast, the number of non-coin balls in the second urn is unknown – there could be either $n = 2$ (one white and one black) or $n = 6$ (three white and three black) non-coin balls. Since objective probabilities about its composition are not given, the second urn is called the *ambiguous* urn. The possibilities are illustrated in Figure 1. Consider now an agent who knows how the urns were filled, but does not know the outcome of the coin tosses. This agent is invited to bet on the color of the two coin balls. Any bet (on a ball of some color drawn from some urn) pays one dollar (or one util) if the ball has the desired color and zero otherwise.

A priori, before any draw is observed, one should be indifferent between bets on the coin ball from either urn - all these bets amount to betting on a fair coin. Suppose now that one draw from each urn is observed and that both balls drawn are black. For the risky urn, it is straightforward to calculate the conditional probability of a black coin ball. Let n denote the number of non-coin balls. Since the unconditional probability of a black coin ball is equal to that of a black draw (both are equal to $\frac{1}{2}$), we have

$$\Pr(\text{coin ball black}|\text{black draw}) = \Pr(\text{black draw}|\text{coin ball black}) = \frac{n/2+1}{n+1},$$

and with $n = 4$ for the risky urn, the result is $\frac{3}{5}$.

The draw from the ambiguous urn is also informative about the coin ball, but there is a difference between the information provided about the two urns. In particular, it is intuitive that one would prefer to bet on a black coin ball in the risky urn rather than in

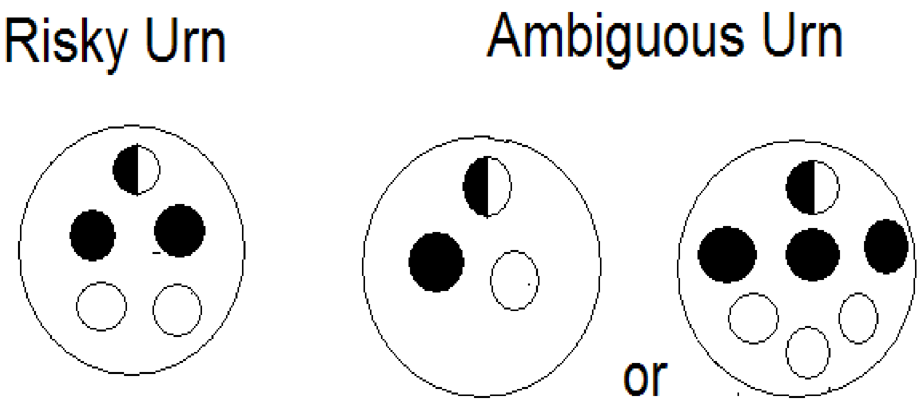

Figure 1: Risky and ambiguous urns for the experiment. The coin balls are drawn as half black. The ambiguous urn contains *either* $n = 2$ *or* $n = 6$ non-coin balls.

the ambiguous urn. The reasoning here could be something like "if I see a black ball from the risky urn, I know that the probability of the coin ball being black is exactly $\frac{3}{5}$. On the other hand, I'm not sure how to interpret the draw of a black ball from the ambiguous urn. It would be a strong indicator of a black coin ball if $n = 2$, but it could also be a much weaker indicator, since there might be $n = 6$ non-coin balls. Thus the posterior probability of the coin ball being black could be anywhere between $\frac{6/2+1}{6+1} = \frac{4}{7} \approx .57$ and $\frac{2/2+1}{2+1} = \frac{2}{3}$. So I'd rather bet on the risky urn." By similar reasoning, it is intuitive that one would prefer to bet on a white coin ball in the risky urn rather than in the ambiguous urn. One might say "I know that the probability of the coin ball being white is exactly $\frac{2}{5}$. However, the posterior probability of the coin ball being white could be anywhere between $\frac{1}{3}$ and $\frac{3}{7} \approx .43$. Again I'd rather bet on the risky urn."

Could a Bayesian agent exhibit these choices? In principle, it is possible to construct a subjective probability belief about the composition of the ambiguous urn to rationalize the choices. However, any such belief must imply that the number of non-coin balls in the ambiguous urn depend on the color of the coin ball, contradicting the description of the experiment. To see this, assume independence and let p denote the subjective probability that $n = 2$. The posterior probability of a black coin ball given a black draw is

$$\tfrac{2}{3}p + \tfrac{4}{7}(1 - p).$$

Strict preference for a bet on a black coin ball in the risky urn requires that this posterior probability be greater than $\frac{3}{5}$ and thus reveals that $p > \frac{3}{10}$. At the same time, strict preference for a bet on a white coin ball in the risky urn reveals that $p < \frac{3}{10}$, a contradiction. While this limitation of the Bayesian model is similar to that exhibited in the Ellsberg Paradox, a key difference is that the Ellsberg Paradox arises in a static context, while here ambiguity is only relevant *ex post,* after the signal has been observed.

Information Quality and Multiple Likelihoods

The preference to bet on the risky urn is intuitive because the ambiguous signal – the draw from the ambiguous urn – appears to be of lower quality than the noisy signal – the draw from the risky urn. A perception of low information quality arises because the distribution of the ambiguous signal is not objectively given. As a result, the standard Bayesian measure of information quality, precision, is not sufficient to adequately compare the two signals. The precision of the noisy signal is parametrized by the number of non-coin balls n: when there are few non-coin balls that add noise, precision is high. We have shown that a single number for precision (or, more generally, a single prior over n) cannot rationalize the intuitive choices. Instead, behavior is *as if* one is using different precisions depending on the bet that is evaluated.

Indeed, in the case of bets on a black coin ball, the choice is made as if the ambiguous signal is less precise than the noisy one, so that the available evidence of a black draw is a weaker indicator of a black coin ball. In other words, when the new evidence – the drawn black ball – is "good news" for the bet to be evaluated, the signal is viewed as relatively imprecise. In contrast, in the case of bets on white, the choice is made as if the ambiguous signal is more precise than the noisy one, so that the black draw is a stronger indicator of a black coin ball. Now the new evidence is "bad news" for the bet to be evaluated and is viewed as relatively precise. The intuitive choices can thus be traced to an asymmetric response to ambiguous news. In our model, this is captured by combining worst-case evaluation as in Gilboa-Schmeidler [16] with the description of an ambiguous signal by multiple likelihoods.

More formally, we can think of the decision-maker as trying to *learn* the colors of the two coin balls. His *prior* is the same for both urns and places probability $\frac{1}{2}$ on black. The draw from the risky urn is a noisy signal of the color of the coin ball. Its (objectively known) distribution is that black is drawn with probability $\frac{3}{5}$ if the coin ball is black, and $\frac{2}{5}$ if the coin ball is white. However, for the ambiguous urn, the signal distribution is unknown. If $n = 2$ or 6 is the unknown number of non-coin balls, then black is drawn with probability $\frac{n/2+1}{n+1}$ if the coin ball is black and $\frac{n/2}{n+1}$ if it is white. Consider now updating about the ambiguous urn conditional on observing a black draw. Bayes' Rule applied in turn to the two possibilities for n gives rise to the posterior probabilities for a black coin ball of $\frac{4}{7}$ and $\frac{2}{3}$ respectively, which leads to the range of posterior probabilities $\left[\frac{4}{7}, \frac{2}{3}\right]$.[4] If bets on the ambiguous urn are again evaluated under worst-case probabilities, then the expected payoff on a bet on a black coin ball in the ambiguous urn is $\frac{4}{7}$, strictly less than $\frac{3}{5}$, the payoff from the corresponding bet on the risky urn. At the same time, the expected payoff on a bet on a black coin ball in the ambiguous urn is $\frac{1}{3}$, strictly less than the risky urn payoff of $\frac{2}{5}$.

Normal Distributions

To write down tractable models with ambiguous signals, it is convenient to use normal distributions. The following example features a normal ambiguous signal that inherits all the key features of the ambiguous urn from the above thought experiment. This example

[4] Because the agent maximizes expected utility under the worst-case probability, his behavior is identical if he uses the entire interval of posterior probabilities or if he uses only its endpoints.

is at the heart of our asset pricing applications below. Let θ denote a parameter that the agent wants to learn about. This might be some aspect of future asset payoffs. Assume that the agent has a unique normal prior over θ, that is $\theta \sim \mathcal{N}(m, \sigma_\theta^2)$ – there is no ambiguity ex ante. Assume further that an ambiguous signal s is described by the set of likelihoods (1) from the introduction. For comparison with the thought experiment, the parameter θ here is analogous to the color of the coin ball, while the variance σ_s^2 of the shock ε plays the same role as the number of non-coin balls in the ambiguous urn.

To update the prior, apply Bayes' rule to all the likelihoods to obtain a family of posteriors:

$$\theta \sim \mathcal{N}\left(m + \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_s^2}(s - m), \frac{\sigma_s^2 \sigma_\theta^2}{\sigma_\theta^2 + \sigma_s^2}\right), \qquad \sigma_s^2 \in \left[\underline{\sigma}_s^2, \overline{\sigma}_s^2\right]. \tag{2}$$

Even though there is a unique prior over θ, updating leads to a nondegenerate set of posteriors – the signal induces ambiguity about the parameter. Suppose further that in each period, choice is determined by maximization of expected utility under the worst-case belief chosen from the family of posteriors. Now it is easy to see that, after a signal has arrived, the agent responds asymmetrically. For example, when evaluating a bet, or asset, that depends positively on θ, he will use a posterior that has a low mean. Therefore, if the news about θ is good ($s > m$), he will act as if the signal is imprecise (σ_s^2 high), while if the news is bad ($s < m$), he will view the signal as reliable (σ_s^2 low). As a result, bad news affect conditional actions more than good news.

2.2 A Model of Learning under Ambiguity

Recursive multiple-priors utility extends the Gilboa-Schmeidler [16] model to an intertemporal setting. Suppose that S is a finite *period* state space. One element $s_t \in S$ is observed every period. At time t, the decision-maker's information consists of the history $s^t = (s_1, ..., s_t)$. Consumption plans are sequences $c = (c_t)$, where each c_t depends on the history s^t. Given a history, preferences over future consumption are represented by a conditional utility function U_t, defined recursively by

$$U_t(c; s^t) = \min_{p_t \in \mathcal{P}_t(s^t)} E^{p_t}\left[u(c_t) + \beta U_{t+1}(c; s^t, s_{t+1})\right], \tag{3}$$

where β and u satisfy the usual properties. The set $\mathcal{P}_t(s^t)$ of probability measures on S captures conditional beliefs about the next observation s_{t+1}. Thus beliefs are determined by the whole *process* of conditional one-step-ahead belief sets $\{\mathcal{P}_t(s^t)\}$.

Epstein and Schneider [14] propose a particular functional form for $\{\mathcal{P}_t(s^t)\}$ in order to capture learning from a sequence of conditionally independent signals. Let Θ denote a *parameter space* that represents features of the data that the decision maker tries to learn. Denote by $\mathcal{M}_0$ a set of probability measures on Θ that represents initial beliefs about the parameters, perhaps based on prior information. Taking $\mathcal{M}_0$ to be a set allows the decision-maker to view this initial information as ambiguous. For most of the effects emphasized in this paper, we do not require ambiguous prior information, and hence assume $\mathcal{M}_0 = \{\mu_0\}$. However, we will compare the effects of ambiguous prior information and ambiguous signals in Section 3 below.

The distribution of the signal s_t conditional on a parameter value θ is described by a *set* of likelihoods $\mathcal{L}$. Every parameter value $\theta \in \Theta$ is thus associated with a set of probability measures $\mathcal{L}(\cdot \mid \theta)$. The size of this set reflects the decision maker's (lack of) confidence in what an ambiguous signal means, given that the parameter is equal to θ. Signals are unambiguous only if there is a single likelihood, that is $\mathcal{L} = \{\ell\}$. Otherwise, the decision-maker feels unsure about how parameters are reflected in data. The set of normal likelihoods described in (1) is a tractable example of this that will be important below.

Beliefs about every signal in the sequence $\{s_t\}$ are described by the same set $\mathcal{L}$. Moreover, for a given parameter value $\theta \in \Theta$, the signals are known to be independent over time. However, the decision-maker is not confident that the data are actually identically distributed over time. In contrast, he believes that any sequence of likelihoods $\ell^t = (\ell_1, .., \ell_t) \in \mathcal{L}^t$ could have generated a given sample s^t and any likelihood in $\mathcal{L}$ might underlie the next observation. The set $\mathcal{L}$ represents factors that the agent perceives as being relevant but which he understands only poorly - they can vary across time in a way that he does not understand beyond the limitation imposed by $\mathcal{L}$. Accordingly, he has decided that he will not try to (or is not able to) learn about these factors. In contrast, because θ is fixed over time, he can try to learn the true θ.

Conditional independence implies that the sample s^t affects beliefs about future signals (such as s_{t+1}) only to the extent that it affects beliefs about the parameter. We can therefore construct beliefs $\{\mathcal{P}_t(s^t)\}$ in two steps. First, we define a set of posterior beliefs over the parameter. For any history s^t, prior $\mu_0 \in \mathcal{M}_0$ and sequence of likelihoods $\ell^t \in \mathcal{L}^t$, let $\mu_t\left(\cdot\, ; s^t.\mu_0, \ell^t\right)$ denote the posterior obtained by updating μ_0 by Bayes Rule if the sequence of likelihoods is known to be ℓ^t. Updating can be described recursively by

$$d\mu_t\left(\cdot\, ; s^t, \mu_0, \ell^t\right) = \frac{\ell_t(s_t \mid \cdot)}{\int_\Theta \ell_t(s_t \mid \theta')\, d\mu_{t-1}(\theta'; s^{t-1}, \mu_0, \ell^{t-1})} d\mu_{t-1}(\cdot\, ; s^{t-1}, \mu_0, \ell^{t-1}).$$

The set of posteriors $\mathcal{M}_t\left(s^t\right)$ now contains all posteriors that can be derived by varying over all μ_0 and ℓ^t:

$$\mathcal{M}_t(s^t) = \left\{\mu_t\left(s^t; \mu_0, \ell^t\right) : \ \mu_0 \in \mathcal{M}_0,\, \ell^t \in \mathcal{L}^t\right\}. \tag{4}$$

Second, we obtain one-step-ahead beliefs by integrating out the parameter. This is analogous to the Bayesian case. Indeed, if there were a single posterior μ_t and likelihood ℓ, the one-step-ahead belief after history s^t would be $p_t\left(\cdot|s^t\right) = \int_\Theta \ell(\cdot \mid \theta)\, d\mu_t(\theta|s^t)$. With multiple posteriors and likelihoods, we define

$$\mathcal{P}_t(s^t) = \left\{p_t\left(\cdot\right) = \int_\Theta \ell_{t+1}(\cdot \mid \theta)\, d\mu_t(\theta) : \mu_t \in \mathcal{M}_t(s^t),\, \ell_{t+1} \in \mathcal{L}\right\}. \tag{5}$$

This is the process of one-step-ahead beliefs that enters the specification of recursive multiple priors preferences (3). The Bayesian model of learning from conditionally i.i.d.

signals obtains as the special case of (5) when both the prior and likelihood sets have only a single element.

3 TREE PRICING

In this section, we derive two key properties of asset pricing with ambiguous news: market participants respond more strongly to bad news than to good news, and returns must compensate market participants for enduring periods of ambiguous news. We derive these properties first in a simple three-period setting. In this context, we also compare the properties of information quality in our model to those of Bayesian models. We then move to an infinite horizon setting, where we derive a number of implications for observed moments.

3.1 An Asset Market with Ambiguous News

There are three dates, labelled 0, 1 and 2. We focus on news about one particular asset (asset A). There are $\frac{1}{n}$ shares of this asset outstanding, where each share is a claim to a dividend

$$d = m + \varepsilon^a + \varepsilon^i. \tag{6}$$

Here m is the mean dividend, ε^a is an aggregate shock and ε^i is an idiosyncratic shock that affects only asset A. In what follows, all shocks are mutually independent and normally distributed with mean zero. We summarize the payoff on all other assets by a dividend $\tilde{d} = \tilde{m} + \varepsilon^a + \tilde{\varepsilon}^i$, where $\tilde{m}$ is the mean dividend and $\tilde{\varepsilon}^i$ is a shock. There are $\frac{n-1}{n}$ shares outstanding of other assets and each pays $\tilde{d}$. The market portfolio is therefore a claim to $\frac{1}{n}d + \frac{n-1}{n}\tilde{d}$.

In the special case $n = 1$, asset A is itself the market. For n large, it can be interpreted as stock in a single company. Under the latter interpretation, one would typically assume that the payoff on the other assets $\tilde{d}$ is itself be a sum of stock payoffs for other companies. One concrete example is the symmetric case of n stocks that each promise a dividend of the form (6), with the aggregate shock ε^a identical and the idiosyncratic shocks ε^i independent across companies. We use this symmetric example below to illustrate the relationship of our results to the law of large numbers. However, the precise nature of $\tilde{d}$ is irrelevant for most of our results.

News

Dividends are revealed at date 2. The arrival of news about asset A at date 1 is represented by the signal

$$s = \alpha\varepsilon^a + \varepsilon^i + \varepsilon^s. \tag{7}$$

Here the number $\alpha \geq 0$ measures how specific the signal is to the particular asset on which we focus. For example, suppose n is large, and hence that d represents future dividends of a single company. If $\alpha = 1$, then the signal s is simply a noisy estimate of future cash flow d. As such, it partly reflects future aggregate economic conditions ε^a. In contrast, if $\alpha = 0$, then the news is 100% company-specific: while it helps to forecast company cash flow d, the signal is not useful for forecasting the payoff on other assets (that is, $\tilde{d}$). Examples of company-specific news include changes in management or merger announcements.

We assume that the signal is ambiguous: the variance of the shock ε^s is known only to lie in some range, $\sigma_s^2 \in [\underline{\sigma}_s^2, \overline{\sigma}_s^2]$. This captures the agent's lack of confidence in the signal's precision. This setup is very similar to the normal distributions example in the previous section. The one difference is that the parameter $\theta = (\varepsilon^a + \varepsilon^i, \varepsilon^a)'$ that agents try to infer from the signal s is now two-dimensional. Apart from that, there is again a single normal prior for θ and a set of normal likelihoods for s parametrized by σ_s^2. The set of one-step-ahead beliefs about s at date 0 consists of normals with mean zero and variance $\alpha^2\sigma_a^2 + \sigma_i^2 + \sigma_s^2$, for $\sigma_s^2 \in [\underline{\sigma}_s^2, \overline{\sigma}_s^2]$. The set of posteriors about θ at date 1 is calculated using standard rules for updating normal random variables. For fixed σ_s^2, let γ denote the regression coefficient

$$\gamma\left(\sigma_s^2\right) = \frac{cov\left(s, \varepsilon^a + \varepsilon^i\right)}{var\left(s\right)} = \frac{\alpha\sigma_a^2 + \sigma_i^2}{\alpha^2\sigma_a^2 + \sigma_i^2 + \sigma_s^2}.$$

Given s, the posterior density of $\theta = (\varepsilon^a + \varepsilon^i, \varepsilon^a)'$ is also normal. In particular, the sum $\varepsilon^a + \varepsilon^i$ is normal with mean $\gamma\left(\sigma_s^2\right) s$ and variance $(1 - \alpha\gamma\left(\sigma_s^2\right))\sigma_a^2 + \left(1 - \gamma\left(\sigma_s^2\right)\right)\sigma_i^2$, while its covariance with ε^a is $\left(1 - \alpha\gamma\left(\sigma_s^2\right)\right)\sigma_a^2$. These conditional moments will be used below. As σ_s^2 ranges over $[\underline{\sigma}_s^2, \overline{\sigma}_s^2]$, the coefficient $\gamma\left(\sigma_s^2\right)$ also varies, tracing out a family of posteriors. In other words, the ambiguous news s introduces ambiguity into beliefs about fundamentals.

Measuring Information Quality

To compare information quality across situations, it is common to measure the information content of a signal *relative* to the volatility of the parameter. For fixed σ_s^2, the coefficient $\gamma(\sigma_s^2)$ provides such a measure since it determines the *fraction* of prior variance in θ that is resolved by the signal. Under ambiguity, $\underline{\gamma} = \gamma\left(\bar{\sigma}_s^2\right)$ and $\bar{\gamma} = \gamma(\underline{\sigma}_s^2)$ provide lower and upper bounds on (relative) information content, respectively. In the Bayesian case, $\bar{\gamma} = \underline{\gamma}$, and agents know precisely how much information the signal contains. More generally, the greater is $\bar{\gamma} - \underline{\gamma}$, the less confident they feel about the true information content. This is the new dimension of information quality introduced by ambiguous signals. At the same time, $\underline{\gamma}$ continues to measure known information content - if $\underline{\gamma}$ increases, everybody knows that the signal has become more reliable.

In the present asset market example, the signal s captures the sum of all intangible information that market participants obtain during a particular trading period, such as a day. The range $\bar{\gamma} - \underline{\gamma}$ describes their confidence in that information. It may differ across markets or time due to differences in information production. For example, stocks that do well often become "hot news", that is, popular news coverage increases. Such coverage will typically not increase the potential for truly valuable news: $\bar{\gamma}$ remains nearly constant. However, the "typical day's news" s will now be affected more by trumped up, irrelevant news items that cannot be easily distinguished from relevant ones: $\underline{\gamma}$ falls. As a second example, suppose a foreign stock is newly listed on the New York Stock Exchange. This will entice more U.S. analysts to research this particular stock, because trading costs for their American clients have fallen. Again, the competence of the information providers is uncertain, especially since the stock is foreign. It again becomes harder to know how reliable is the typical day's news. However, since most of the new coverage is by experts, one would now expect $\bar{\gamma}$ to increase, while $\underline{\gamma}$ remains nearly constant.

3.2 Asymmetric Response and Price Discount

We assume that there is a representative agent who does not discount the future and cares only about consumption at date 2. He has recursive multiple-priors utility with beliefs as described above. We begin with a Bayesian benchmark, where the agent maximizes expected utility and beliefs are as above with $\underline{\gamma} = \bar{\gamma}$. We also allow for risk aversion: let period utility be given by $u(c) = -e^{-\rho c}$, where ρ is the coefficient of absolute risk aversion.

Bayesian Benchmark

It is straightforward to calculate the price of asset A at dates 0 and 1:

$$\begin{aligned} q_0 &= m - \rho cov\left(d, \frac{1}{n}d + \frac{n-1}{n}\tilde{d}\right) = m - \rho\left(\sigma_a^2 + \frac{1}{n}\sigma_i^2\right); \\ q_1(s) &= m + \underline{\gamma}s - \rho\left(\left(1 - \alpha\underline{\gamma}\right)\sigma_a^2 + \frac{1}{n}\left(1-\underline{\gamma}\right)\sigma_i^2\right). \end{aligned} \tag{8}$$

At both dates, price equals the expected present value minus a risk premium that depends on risk aversion and covariance with the market. At date 0, the expected present value is simply the prior mean dividend m. At date 1, it is the posterior mean dividend $m + \underline{\gamma}s$: it now depends on the value of the signal s provided that the signal is informative $\left(\underline{\gamma} > 0\right)$. The risk premium depends only on time (and not on s) – it is smaller at date 1 as the signal resolves some uncertainty. At either date, it consists of two parts, one driven by the variance of the common shock ε^a, and one equal to the variance of the idiosyncratic shock multiplied by $\frac{1}{n}$, the market share of the asset. As n becomes large, idiosyncratic risk is diversified away and does not matter for prices.

Ambiguous Signals

We now calculate prices when the signal is ambiguous. For simplicity, we assume that the agent is risk neutral; of course, he is still averse to ambiguity.[5] As discussed in Section 2, with recursive multiple-priors utility, actions are evaluated under the worst-case conditional probability. We also know that the representative agent must hold all

[5] This approach allows us to derive transparent closed form solutions for key moments of prices and returns. In the numerical example considered below, risk aversion is again introduced.

assets in equilibrium. It follows that the worst-case conditional probability minimizes conditional mean dividends. Therefore, the price of asset A at date 1 is

$$q_1(s) = \min_{\sigma_s^2 \in [\underline{\sigma}_s^2, \bar{\sigma}_s^2]} E[d|s] = \begin{cases} m + \underline{\gamma} s & \text{if } s \geq 0 \\ m + \bar{\gamma} s & \text{if } s < 0. \end{cases} \quad (9)$$

A crucial property of ambiguous news is that the worst-case likelihood used to interpret a signal depends on the value of the signal itself. Here the agent interprets bad news ($s < 0$) as very informative, whereas good news are viewed as imprecise. The price function $q_1(s)$ is thus a straight line with a kink at zero, the cutoff point that determines what "bad news" means. If the agent is not ambiguity averse $(\bar{\gamma} = \underline{\gamma})$, the price function is the same as that for a Bayesian agent who is not risk averse $(\rho = 0)$.

At date 0, the agent knows that an ambiguous signal will arrive at date 1. His one-step-ahead conditional beliefs about the signal s are normal with mean zero and variance $\alpha^2\sigma_a^2 + \sigma_i^2 + \sigma_s^2$, where σ_s^2 is unknown. Again, the worst-case probability is used to evaluate portfolios. Since the date 1 price is concave in the signal s, the date zero conditional mean return is minimized by selecting the highest possible variance $\bar{\sigma}_s^2$. We thus have

$$\begin{aligned} q_0 &= \min_{\sigma_s^2 \in [\underline{\sigma}_s^2, \bar{\sigma}_s^2]} E[q_1] \\ &= \min_{\sigma_s^2 \in [\underline{\sigma}_s^2, \bar{\sigma}_s^2]} E\left[m + \underline{\gamma} s + (\bar{\gamma} - \underline{\gamma}) \min\{s, 0\}\right] \\ &= m - (\bar{\gamma} - \underline{\gamma}) \frac{1}{\sqrt{2\pi\underline{\gamma}}} \sqrt{\alpha\sigma_a^2 + \sigma_i^2} \end{aligned} \quad (10)$$

The date zero price thus exhibits a discount, or ambiguity premium. This premium is directly related to the extent of ambiguity, as measured by $\bar{\gamma} - \underline{\gamma}$. It is also increasing in the volatility of fundamentals, including the volatility σ_i^2 of idiosyncratic risk. Without ambiguity aversion, we obtain risk neutral pricing ($q_0 = m$), exactly as in the case of no risk aversion ($\rho = 0$) in (8).

Comparison of (10) and (8) reveals two key differences between risk premia and premia induced by ambiguous information. The first is the role of *idiosyncratic shocks* for the price of small assets. Ambiguous company-specific news not only induces a premium, but the size of this premium depends on total (including idiosyncratic) risk. In the Bayesian case, whether company-specific news is of low quality barely matters even ex post. Indeed, for $\sigma_a^2 = 0$ and n large, the average price at date 1 equals the price at date 0, and both are equal to the unconditional mean dividend. Second, under ambiguity, prices depend on the *prospect* of low information quality. It is intuitive that if it becomes known today that information about asset A will be more difficult to interpret *in the future*, this makes asset A less attractive, and hence cheaper, already today. This is exactly what happens when the signal is ambiguous. In contrast, a change of information quality in the Bayesian model does not have this effect. While the prospect of lower information quality in the future produces a larger discount ex post after the news has arrived (q_1 is increasing in $\underline{\gamma}$), the ex ante price q_0 is independent of $\underline{\gamma}$.

Both properties can be traced to one behavioral feature: for ambiguity averse investors, uncertainty about the distribution of future payoffs is a first-order concern. While the Bayesian model assumes that agents treat all model uncertainty as risk, the multiple-priors model accommodates intuitive behavior by assuming that agents act *as if* they adjust the mean of uncertain assets (or bets). For example, in the thought experiment of Section 2, a preference for betting on the risky urn derives from the fact that agents evaluate bets on the ambiguous urn using a lower posterior mean. The same effect is at work here. To elaborate, consider first the impact of idiosyncratic shocks. If uncertainty about mean earnings changes because of company-specific news, then Bayesians treat this as a change in risk. There will be only a second-order effect on the Bayesian valuation of a company as long as the covariance with the market remains the same. In contrast, ambiguity averse investors act as if mean earnings themselves have changed. This is a first-order effect, even if the company is small.

Second, suppose that Bayesian market participants are told at date 0 that hard-to-interpret news will arrive at date 1. They believe that, at date 1, everybody will simply form subjective probabilities about the meaning of the signal at date 1 and average different scenarios to arrive at a forecast for dividends. As long as the volatility of fundamentals does not change, total risk is the same and there is no need for prices to change. In contrast, ambiguity averse market participants know that they will not be confident enough to assign subjective probabilities to different interpretations of the signal at date 1. Instead they will demand a discount once they have seen the signal. As a result, prices reflect this discount even at date 0. The prospect of ambiguous news is thus enough to cause a drop in prices.

Idiosyncratic uncertainty and the law of large numbers

To gain more intuition about the role of idiosyncratic volatility, consider the symmetric case with n assets, indexed by i, each with payoff (6). Assume further that there is no aggregate risk ($\sigma_a^2 = 0$) and that an independent signal of the type (7) arrives about each asset at date 1. From (8), the date 0 value of any vector of portfolio holdings α with $\Sigma_i \alpha_i = 1$ can be written as $m - \rho cov(\alpha' \varepsilon^i, \frac{1}{n} \Sigma_j \varepsilon^j)$.

A key feature of rational asset pricing is that assets are not evaluated in isolation. Uncertainty is only reflected in prices to the extent it actually lowers investor utility. Uncertainty may be due to either risk or ambiguity. When uncertainty consists of risk, it lowers utility by increasing the volatility of consumption, or, under additional assumptions that are satisfied here, the volatility of the market portfolio. With purely idiosyncratic risks, the law of large numbers implies that the variance of the market portfolio tends to zero as n becomes large. As a result, the value of any portfolio converges to the mean m, linearly in n. In other words, as the market portfolio becomes riskless, the uncertainty of any particular portfolio does not affect utility much at the margin, resulting in a small premium.

Ambiguity averse investors also do not evaluate assets in isolation. As in the Bayesian case, uncertainty is reflected in prices only to the extent it lowers utility. However, uncertainty is now captured by the range of probabilities that describe beliefs, and it affects utility by making the worst case probability less favorable to the investor. In particular, in the case of risk neutrality considered here, uncertainty lowers the worst case mean. In contrast to the Bayesian case, the market portfolio does not become less uncertain as the number of assets increases.[6] Intuitively, the presence of uncertainty makes ambiguity averse investors act as if the mean is lower. Moreover, in a setting with independent and identical sources of uncertainty, behavior towards each source is naturally the same. Therefore, investors act as if the mean on each source – here each individual asset – is lower. Summing up, they act as if the mean of the market portfolio is lower, regardless of the number of assets.[7]

The above argument shows that, in a setting of iid ambiguity, the market portfolio does not become less uncertain with the number of assets. In addition, the marginal change in utility from investing in portfolio α does not shrink with n. If a marginal dollar is spent on portfolio α, what matters is the marginal change in the worst case mean, which does not depend on the number of assets. If ambiguity is induced by ambiguous information, (10) shows that the change in the worst case mean scales with the volatility of fundamentals. In particular, the value of the portfolio α depends on the weighted idiosyncratic volatility of the assets; it can now be written as $m - \frac{\bar{\gamma}-\underline{\gamma}}{\sqrt{2\pi\underline{\gamma}}} \sum_i \alpha_i \sigma_i$.

Ambiguous Signals vs. Ambiguous Prior Information

It is interesting to compare premia induced by ambiguous signals with premia due to ambiguous prior information. To extend the model to ambiguous prior information, assume that $m = 0$, but that the mean of the parameter vector $\theta = (\varepsilon^i, \varepsilon^i + \varepsilon^a)$ is perceived to lie in the set $[\underline{m}_i, \overline{m}_i] \times [\underline{m}_i + \underline{m}_a, \bar{m}_i + \overline{m}_a]$.[8] This assumption defines a set of priors over θ. Applying the normal updating formula (2) to every possible prior, as suggested by the general formula (4), we obtain a set of posteriors. The family of marginals on the relevant parameter $\varepsilon^i + \varepsilon^a = d$ is

$$d \sim \mathcal{N}\left(m_i + m_a, \gamma(s - \alpha m_a - m_i), (1-\gamma)\,\sigma^2\right); \quad (m_i, m_a) \in [\underline{m}_i, \overline{m}_i] \times [\underline{m}_a, \overline{m}_a],\ \gamma \in [\underline{\gamma}, \overline{\gamma}].$$

The pricing equation (9) changes only in that there is now a joint minimization over the means (m_i, m_a) and the signal variance σ_s^2. Naturally, investors at date 1 evaluate asset A using the worst case prior means $(\underline{m}_i, \underline{m}_a)$. This behavior is anticipated by investors at date 0, so that the lowest prior mean also enters the price formula (10).

[6] This is not to say that a model with ambiguity aversion does not allow for any benefits from diversification. For example, if the agent were risk averse in addition to ambiguity averse, the effect of idiosyncratic risk through risk aversion would diminish as n grows. The point is that here we isolate ambiguity about the mean, which is not diversified away.

[7] More formally, versions of the law of large numbers for iid ambiguous random variables show that (*i*) sample averages must (almost surely) lie in an interval bounded by the highest and lowest possible mean, and (*ii*) these bounds are tight in the sense that convergence to a narrower interval does not occur (see Marinacci [21] or Epstein and Schneider [13]).

[8] Since the aggregate and idiosyncratic components ε^a and ε^i are independent, it is natural to allow for all possible combinations of their means m_a and m_i.

Equilibrium prices are thus

$$
\begin{aligned}
q_1(s) &= \underline{m}_i + \underline{m}_a + \underline{\gamma}\left(s - \alpha \underline{m}_a - \underline{m}_i\right) + \left(\bar{\gamma} - \underline{\gamma}\right) \min\left\{s - \alpha \underline{m}_a - \underline{m}_i, 0\right\}, \\
q_0 &= \underline{m}_i + \underline{m}_a - \left(\bar{\gamma} - \underline{\gamma}\right) \frac{1}{\sqrt{2\pi \underline{\gamma}}} \sqrt{\alpha \sigma_a^2 + \sigma_i^2}.
\end{aligned}
$$

Like the anticipation of future ambiguous signals, ambiguous prior information also gives rise to a price discount at date zero. For example, when comparing two securities with similar expected payoffs, the one about which there is more prior uncertainty would typically be modeled by a wider interval of prior means. It would thus have a lower worst case mean and a lower price. However, the discount induced by ambiguous prior information does not scale with volatility: an explicit link between ambiguity premia and the volatility of fundamentals is unique to the case of ambiguous information.

As before, the price at date 1 reflects investors' asymmetric response to news. However, the presence of ambiguous prior information may change the meaning of "good news": if the mean of the signal is itself ambiguous, the signal is now treated as unreliable if it is higher than the worst case prior mean. In this case, ambiguous signals draw out the effect of any ambiguous prior information on prices. Indeed, for a given "true" data generating process, an investor with a wider interval of prior means will be more likely to receive "good news", which he will not weigh heavily, so that learning is slower. Of course, interaction between ambiguous priors and signals requires that the mean of the signal is itself ambiguous. If this is not the case – for example, if $\alpha = 0$ and $\underline{m}_i = \overline{m}_i$ – then prior ambiguity simply lowers the price by a constant amount at both dates.[9]

3.3 Asset Price Properties

To compare the predictions of the model to data, we embed the above three-period model of news release into an infinite-horizon asset pricing model. Specifically, we chain together a sequence of short learning episodes of the sort modeled above. Agents observe just *one* intangible signal about the next innovation in dividends before that innovation is revealed and the next learning episode starts.

We maintain the assumption of risk neutrality, but now fix an exogenous riskless interest rate r and a discount factor $\beta = \frac{1}{1+r}$ for the agent. In addition, we omit the distinction between systematic and idiosyncratic shocks, since agents' reaction to ambiguous signals is similar in the two cases. The level of dividends on some asset is given by a mean-reverting process,

$$
d_t = \kappa \bar{d} + (1 - \kappa) d_{t-1} + u_t, \tag{11}
$$

[9] In this context, it is not essential that ε^a is an aggregate shock and that ε^i is an idiosyncratic shock. We could also imagine ε^a to be an idiosyncratic component of the dividend that is a priori ambiguous, but that does not affect the signal.

where u_t is a shock and $\kappa \in (0,1)$. The parameter κ measures the speed with which dividends adjust back to their mean $\bar{d}$.[10]

Every period, agents observe an ambiguous signal about next period's shock:

$$s_t = u_{t+1} + \varepsilon_t^s,$$

where the variance of ε_t^s is $\sigma_{s,t}^2 \in [\underline{\sigma}_s^2, \overline{\sigma}_s^2]$. The relevant state of the world for the agent is (s_t, d_t). The components s_t and d_t are conditionally independent, because s_t provides information only about u_{t+1}, which in turn is independent of d_t. Beliefs about s_{t+1} are normal with mean zero and (unknown) variance $\sigma_u^2 + \sigma_{s,t}^2$. Beliefs about d_{t+1} are given by (11) and by the set of posteriors about u_{t+1} given s_t described in the previous section.

Our goal is to derive asset pricing properties that would be observed by an econometrician who studies the above asset market. We thus assume that there is a *true* variance of noise $\sigma_s^{*2} \in [\underline{\sigma}_s^2, \overline{\sigma}_s^2]$. It is also useful to define $\gamma^* = \gamma(\sigma_s^{*2})$, a measure of the true information content of the news that arrives in a typical trading period. In addition, we assume that the true distribution of the fundamentals u coincides with the subjective beliefs of agents. The latter assumption distinguishes the present model from existing approaches to asset pricing under ambiguity. Indeed, existing models are driven by ambiguity about fundamentals. The degree of ambiguity is then often motivated by how hard it is to measure fundamentals. The present setup illustrates that ambiguity can matter even if the true process of dividends is *known* by both the econometrician and market participants. The point is that market participants typically have access to ambiguous information, other than past dividends, that is not observed by the econometrician.

Let q_t denote the stock price. In equilibrium, the price at t must be the worst-case conditional expectation of the price plus dividend in period $t+1$:

$$q_t = \min_{\left(\sigma_{s,\,t}^2, \sigma_{s,t+1}^2\right) \in \left[\underline{\sigma}_s^2, \overline{\sigma}_s^2\right]^2} \beta E_t\left[q_{t+1} + d_{t+1}\right]. \tag{12}$$

We focus on stationary equilibria. The price is given by

$$q_t = \frac{\bar{d}}{r} + \frac{1-\kappa}{r+\kappa}\left(d_t - \bar{d}\right) + \frac{1}{r+\kappa}\gamma_t s_t - \left(\bar{\gamma} - \underline{\gamma}\right)\frac{\sigma_u}{r\sqrt{2\pi\underline{\gamma}}}, \tag{13}$$

where γ_t is a random variable that is equal to $\bar{\gamma}$ if $s_t < 0$ and equal to $\underline{\gamma}$ otherwise.[11]

The first two terms reflect the present discounted value of dividends – without intangible news, prices are determined only by the interest rate and the current dividend level. The third term captures the response to the current ambiguous signal. As in (9), this

[10]Under these assumptions, dividends are stationary in levels, which is not realistic. However, it is straightforward to extend the model to allow for growth. Let observed dividends be given by $\hat{d}_t = g^t(\bar{d} + d_t)$ and $d_t = (1-\kappa)\,d_{t-1} + u_t$, where $g-1$ is the average growth rate, $g-1 < r$. The observed stock price in the growing economy is then $\widehat{q}_t = g^t q_t$. The analysis below applies to the detrended stock price q_t if β is replaced by βg.

[11]Conjecture a time invariant price function of the type $q_t = \bar{Q} + Q_{\hat{d}}\hat{d}_t + Q_s\gamma_t s_t$. Inserting the guess into (12) and matching undetermined coefficients delivers (13).

response is asymmetric: the distribution of γ_t implies that bad news are incorporated into prices more strongly. In addition, the strength of the reaction now depends on the persistence of dividends: if κ is smaller, then the effect of news on prices is stronger since the information matters more for payoffs beyond just the next period. The fourth term captures anticipation of *future* ambiguous news – it is the present discounted value of the premium in (10). As before, it may contain compensation for asset-specific shocks.

We are interested also in the behavior of excess returns. In the present setup, it is convenient to focus on *per share* excess returns, defined as

$$\begin{aligned} R_{t+1} &= q_{t+1} + d_{t+1} - (1+r)\, q_t \\ &= \frac{1}{r+\kappa}\gamma_{t+1} s_{t+1} + \frac{1+r}{r+\kappa}\left(u_{t+1} - \gamma_t s_t\right) + \left(\bar{\gamma} - \underline{\gamma}\right) \frac{\sigma_u}{\sqrt{2\pi\underline{\gamma}}} \end{aligned} \tag{14}$$

It is helpful to decompose the excess return realized into three separate parts. The first is the response to the current ambiguous signal, $R_{t+1}^{(1)} := \frac{1}{r+\kappa}\gamma_{t+1}s_{t+1}$. It reflects the response of the current price to intangible news about future dividends. The second component is the response to the realization of dividends, $R_{t+1}^{(2)} := \frac{1+r}{r+\kappa}\left(u_{t+1} - \gamma_t s_t\right)$. It is proportional to the difference between the current innovation to dividends u_{t+1} and the response to last period's signal about that innovation, $\gamma_t s_t$. The two stochastic components of excess returns $R_{t+1}^{(1)}$ and $R_{t+1}^{(2)}$ are independent of each other. The last, constant term in (14) represents a constant premium that investors obtain for enduring low information quality in the future.

If there is no intangible information, all γ's are zero and excess returns are $\frac{1+r}{r+\kappa}u_{t+1}$; thus they depend only on the shock to fundamentals and are always expected to be zero. If the signals are unambiguous $\left(\bar{\gamma} = \underline{\gamma} = \gamma^*\right)$, $R_{t+1}^{(1)}$ is proportional to investors' forecast in $t+1$ of the next dividend innovation for $t+2$, while $R_{t+1}^{(2)}$ is the forecast error that investors realize in $t+1$. The signals now affect the volatility of returns. However, expected excess returns are still zero since the representative agent is risk neutral.

Finally, consider the case of ambiguous signals. Although the representative agent does not think in terms of probabilities, he behaves *as if* he holds beliefs that give rise to forecasts $\gamma_{t+1}s_{t+1}$ and forecast errors $u_{t+1} - \gamma_t s_t$. In particular, to an observing econometrician it will appear as if the agent overreacts to bad intangible news, by making a particularly pessimistic forecast $\underline{\gamma} s_{t+1}$. In contrast, it will appear as if the forecast error realized when bad dividends are realized is "too small", since it is computed relative to an earlier forecast that was already pessimistic. In this sense, the ambiguity averse agent overreacts to bad intangible news, but underreacts to bad tangible news. By the same logic, he appears to underreact to good tangible news, but overreact to good intangible news.

Equity Premium and Idiosyncratic Risk

The mean excess return under the true probability is

$$E^*\left[R_{t+1}\right] = \left(\bar{\gamma} - \underline{\gamma}\right) \frac{1}{\sqrt{2\pi\underline{\gamma}}} \left(1 + \frac{r}{r+\kappa}\sqrt{\underline{\gamma}/\gamma^*}\right) \sigma_u.$$

The presence of ambiguous news induces an ambiguity premium. It is well-known that such a premium can arise as a result of ambiguity in fundamentals. What is new here is that it is driven by ambiguity about the quality of news. Since news is in turn driven by fundamentals, this introduces a direct link between the *volatility* of fundamentals and the ambiguity premium: $E^* [R_{t+1}]$ is increasing in σ_u. Ambiguous signals thus provide a reason why idiosyncratic risk is priced. Indeed, as discussed above, it does not matter for the order of magnitude of the effect whether the news is company-specific or not.

Recent literature has documented a link between idiosyncratic risk and excess returns in the cross section (for example, Lehmann [18] and Malkiel and Xu [19]). The explanation typically put forward is that agents do not fully diversify. There is indeed evidence that some agents, such as individual households, hold undiversified portfolios.[12] However, a Bayesian model will produce equilibrium premia for idiosyncratic risk only if there is *no* agent who holds a well-diversified portfolio. Any well-diversified mutual fund, for example, will bid up prices until the discount on idiosyncratic risk is zero. In contrast, the present model features one, well-diversified, investor. As long as this investor views company-specific news as ambiguous, he will want to be compensated for it. Since institutional investors with low transaction costs still have to process news, this delivers a more robust story for why idiosyncratic risk can matter.

Additional implications of the ambiguous news model could be used to explore further the trade-off between idiosyncratic shocks and expected returns. In particular, the premium changes to different degrees depending on the way in which information quality is increased. Other things equal, the premium is increasing in both $\bar{\gamma}$ and $\underline{\gamma}$, but the derivative with respect to $\bar{\gamma}$ is always larger. This implies that, at the margin, an increase in coverage by potential experts (higher $\bar{\gamma}$) induces a larger increase in the premium than an increase in popular news coverage $\left(\text{lower } \underline{\gamma}\right)$. The intuition is that potential high quality news moves prices more, and hence induces more uncertainty per unit of volatility of fundamentals σ_u.

Excess Volatility

A classic question in finance is why stock prices are so much more volatile than measures of the expected present value of dividends. We now reconsider the link between "excess volatility" and information quality. The variance of the stock price is

$$var\left(q_t\right) = \sigma_u^2 \left(\frac{1-\kappa}{r+\kappa}\right)^2 \left(\frac{1}{\kappa\left(2-\kappa\right)} + \frac{1}{2\gamma^*}\left(\bar{\gamma}^2 + \underline{\gamma}^2 - \frac{1}{\pi}\left(\bar{\gamma} - \underline{\gamma}\right)^2\right)\right).$$

Price volatility is proportional to the volatility of the shock u. When there is no intangible information $\left(\underline{\gamma} = \bar{\gamma} = 0\right)$, the second term in the big bracket is zero, and the volatility of prices is equal to that of the present value of dividends .

The big bracket reflects the propagation of shocks through news. In the Bayesian case ($\underline{\gamma} = \bar{\gamma} = \gamma^*$), it is equal to $\frac{1}{\kappa(2-\kappa)} + \gamma^*$: given persistence κ, price volatility is

[12] In fact, ambiguity aversion may be responsible for such underdiversification. See Epstein and Schneider [14] for an analysis of portfolio choice with recursive multiple-priors.

increasing in information quality and bounded above. In particular, with persistent dividends (κ small), changes in information quality typically have only a small effect on price volatility, which is dominated by the volatility of the present value of dividends. With ambiguous signals, at the benchmark $\bar{\gamma} = \underline{\gamma} = \gamma^*$, $var(q_t)$ is increasing in both $\bar{\gamma}$ and $\underline{\gamma}$. In other words, more news coverage by potential experts (higher $\bar{\gamma}$) increases volatility, while more popular coverage (lower $\underline{\gamma}$) tends to reduce it. Also, the possibility of ambiguous news removes the upper bound on price volatility imposed by the Bayesian model.[13] Ambiguous signals can thus contribute to excess volatility of prices.

Recent empirical work on firm level volatility has mostly focused on volatility of *returns*. In our model,

$$var(R_t) = \sigma_u^2 \left(\frac{1+r}{r+\kappa}\right)^2 \left\{1 - \bar{\gamma} - \underline{\gamma} + \frac{1+\beta^2}{2\gamma^*}\left(\bar{\gamma}^2 + \underline{\gamma}^2 - \frac{1}{\pi}\left(\bar{\gamma} - \underline{\gamma}\right)^2\right)\right\}$$

At the point $\bar{\gamma} = \underline{\gamma} = \gamma^*$, the derivatives of $var(R_t)$ with respect to both $\bar{\gamma}$ and $\underline{\gamma}$ are again positive. Changes in information quality due to changes in the ambiguity of signals thus affect the volatility of prices and returns in the same way. This is in sharp contrast to the Bayesian case, where price and return volatilities move in opposite directions.[14] As explained by West [27], the latter result obtains because higher information quality simply means that more information about future cash flows is released earlier, when these future cash flows are still being discounted at a higher rate.

Campbell et al. [8] document an upward trend in individual stock return volatility. One question is whether this development can be connected to improvements in information technology. Campbell et al. interpret such improvements as an increase in γ^*, and dismiss the explanation, because return volatility should decrease, rather than increase. In our framework, this argument only applies if the effect of the new technology is immediately fully known. If agents have also become less certain about how much improvement there is, the outcome is no longer obvious. In particular, while higher γ^* lowers return volatility, higher $\bar{\gamma}$ increases it. Increased uncertainty about the potential of information technology is thus consistent with higher volatility of returns.

Skewness

Since ambiguity averse market participants respond asymmetrically to news, the model tends to produce skewed distributions for prices and returns, even though both dividends and noise have symmetric (normal) distributions. Skewness of a random variable x is usually defined by $\mu_3(x) / (\sigma(x))^3$, where μ_3 is the centered third moment and σ is the

[13] This does not mean that price volatility in the ambiguous news model is arbitrary. It just says that the model cannot be rejected using relative volatility of prices and dividends. In a quantitative application, $\bar{\gamma}$, $\underline{\gamma}$ and γ^* can be identified from the full distribution of prices. See Section 4 for an example.

[14] Formally, in the Bayesian case, $var(R_t) = \frac{\sigma_u^2}{(1-\beta\kappa)^2}\left(1 - \gamma^*(1-\beta^2)\right)$, so that an increase in γ^* makes price more volatile but returns less so.

standard deviation. For returns, the third moment is

$$\mu_3 \left(R_{t+1} \right) = \left(\frac{1+r}{r+\kappa} \right)^3 \left\{ - \left(1 - \beta^3 \right) \mu_3 \left(\gamma_t s_t \right) - \frac{\sigma_u^3}{\sqrt{2\pi\gamma^{*3}}} \left(\bar{\gamma} - \underline{\gamma} \right) \left(\bar{\gamma} + \underline{\gamma} - \gamma^* \right) \right\}.$$

The appendix shows that $\mu_3 (\gamma_t s_t)$ is negative and proportional to σ_u^3, where the proportionality factor depends on $\bar{\gamma}$, $\underline{\gamma}$ and γ^*. Since $\gamma^* \leq \bar{\gamma}$, the second term is negative. It follows that returns are negatively skewed if the discount factor β is high enough. Moreover, since the standard deviation of returns is proportional to σ_u, skewness is independent of the volatility of fundamentals and depends only on relative information quality. In the Bayesian case, skewness is zero. It can also be verified that for small enough discount factors, a small increase in ambiguity, starting from a Bayesian benchmark, actually leads to positive skewness of returns.

The intuition follows from the above decomposition of returns into a response to the current ambiguous signal $R_{t+1}^{(1)}$ and a response to the current dividend shock $R_{t+1}^{(2)}$. The signal s_{t+1} will be weighted more heavily the less favorable it is, so that $R_{t+1}^{(1)}$ is negatively skewed. In contrast, the dividend shock u_{t+1} will often offset strong negative responses to bad signals in the previous period, while it will amplify weak positive responses to past signals. As a result, $R_{t+1}^{(2)}$ is positively skewed. The overall skewness of returns thus depends on how much of their variability is contributed by the response to intangible news $R_{t+1}^{(1)}$ vs. the reaction to tangible news $R_{t+1}^{(2)}$.

Since the signal s_{t+1} is about future cash flows (beyond $t+1$), the contribution of the negatively skewed component $R_{t+1}^{(1)}$ increases with the discount factor. Moreover, if the signal provides any information, the "forecast error" $u_{t+1} - \gamma_t s_t$ will be less positively skewed than the "forecast" $\gamma_{t+1} s_{t+1}$ is negatively skewed. Without discounting ($\beta = 1$), we thus always obtain negative skewness of returns: summing up $R_{t+1}^{(1)}$ and $R_{t+1}^{(2)}$ does not simply cancel skewness in successive signals, since the past signal is correlated with the current dividend. In contrast, an ambiguous signal that provides little information (and thus generates little volatility) and is also heavily discounted cannot contribute enough negative skewness to offset positive skewness due to the (undiscounted) forecast error that enters $R_{t+1}^{(2)}$. In this case, positive skewness obtains.

The model thus predicts that negative skewness of returns should prevail at high frequencies, which here corresponds to a lower discount rate or higher β. This fact has been widely documented in the empirical literature. It is closely connected to the finding that volatility increases in times of low returns, which has been found in both aggregate (index) and individual stock returns.[15] The model also predicts that negative skewness should be more pronounced for stocks for which there is more intangible information. This feature helps explain existing evidence on skewness in the cross-section of stocks.

In particular, Harvey and Siddique [17] show that negative skewness tends to be more pronounced for stocks with larger market capitalization. Our model suggests that this is because large cap stocks are "in the news" more, so that traders in those stocks have

[15] See, for example, Bekaert and Wu [2] for references.

to digest more ambiguous, intangible information. In contrast, for small stocks that are not followed widely by the media, so that information mostly arrives in the form of the occasional tangible earnings report, one would not expect a lot of negative skewness, or perhaps even positive skewness.

It is important for this last point that our model does *not* predict that skewness and volatility are related in the cross section: large stocks can be more negatively skewed, but still have lower volatility and hence lower ambiguity premia, since their fundamentals are more stable (lower σ_u^2). It is also important that skewness is driven by the ambiguity of signals, but need not be affected by ambiguous *prior* information. One might expect that investors perceive small stocks to be more ambiguous ex ante. Incorporating prior ambiguity about the mean dividend u_t, say, following the example at the end of the previous subsection, would increase the ambiguity premium and could thus help to explain the size effect. However, it would not change the centered third moment and skewness, provided that the mean of the signal remains unambiguous. Thus higher moments would still depend only on the perception of news.

4 SHOCKS TO INFORMATION QUALITY

The previous section compared markets with different information quality. We now focus on changes to information quality in a given market. In particular, asset markets often witness shocks that not only increase uncertainty about fundamentals, but also force markets to deal with unfamiliar news sources. Our leading example for such a shock is September 11, 2001. On the one hand, the terrorist attack increased uncertainty about economic growth. On the other hand, news about terrorism and foreign policy – that were arguably less important for U.S. growth and stock returns earlier – suddenly became important for market participants to follow. It thus became more difficult for market participants to assess how much weight to place on any given piece of news. For familiar news, such as Fed announcements and macro statistics releases, market participants had – through years of experience – developed a feel for the relevance of any given piece of news. Such experience was lacking for the news that became relevant after 9/11.

We view 9/11 as the beginning of a learning process where market participants were trying to infer the possibility of structural change to the U.S. economy from unfamiliar signals. Figure 2 plots the price-dividend ratio for the S&P 500 index for 19 trading days, starting 9/17, including the pre-attack value, 9/10, as day 0. The stock market was closed in the week after the attack; trading resumed on Monday, 9/ 17. The large drop on that day was followed by another week of losses, before the market began to rebound. At the end of our window – Friday, October 5 – the market had climbed, for the first time, to the pre-attack level. It subsequently remained between 68 and 73 for another three weeks (not shown). With hindsight, we know that structural change did not occur. The goal is thus to explain what moved the market during the learning process, conditional on that knowledge.

We argue that the key implications of ambiguous news - asymmetric response and price discount - help to explain the observed price pattern. In particular, a Bayesian model with known signal quality has problems explaining the initial slide in prices. Roughly, if signal precision is high, the arrival of enough bad news to explain the first week is highly unlikely. If signal precision is low, bad news will not be incorporated into prices in the first place. In our model, where signal precision is unknown, bad news are taken especially seriously and hence a much less extreme sequence of signals suffices to account for prices in the first week. In sum, ambiguous information can help to rationalize the delayed negative response observed after a shock to information quality. The exercise also illustrates how belief parameters can be identified from the distribution of prices.

Setup

There is an infinitely-lived representative agent. A single Lucas tree yields dividends $Y_t = \exp\left(\Sigma_{j=1}^t \Delta y_j\right) Y_0$, with Y_0 given. According to the true data generating process, the growth rate of dividends is $\Delta y_t \sim i.i.\mathcal{N}\left(\theta^{hi}, \sigma^2\right)$ for all t. The agent knows that the mean growth rate is θ^{hi} from time 0 up to some given time $T+1$. However, he believes that with probability $1-\mu$, the mean growth rate drops permanently to θ^{lo} after $T+1$. Information about growth beyond $T+1$ is provided, at each date $t \leq T$, by a signal s_t that takes the values 1 or 0. Signals are serially independent and also independent of dividends before $T+1$; they satisfy $\Pr\left(s_t = 1\right) = \pi$. At time $T+1$, the long run mean growth rate is revealed.

The information structure captures the following scenario. First, there was no actual permanent structural change caused by the attack.[16] Second, agents were initially unsure if there would be such a change. Third, news reports were initially much more informative about the possibility of structural change than were dividend or consumption data. Of course, to the extent that dividend data were available, they may have provided some information. But initially, they are likely to have largely reflected decisions taken before the attack occurred, becoming more informative only with time. Our model captures this shift in relative informativeness in a stark way. We divide the time after the attack into two phases, a learning phase ($t \leq T$) where dividends are entirely uninformative about structural change, and a "new steady state" phase ($t > T$) where structural change actually materializes in dividends. In our calibration below, T corresponds to 26 days.[17] Finally, imposing a fixed T at which θ is revealed is not a strong restriction if beliefs are already close to the true θ at time T. We show a plot of our posterior means below.

The agent believes that signals are informative about growth, but views them as

[16] The model can nevertheless accommodate drops in dividends in September and stock price movements that reflect these drops. All that is required is that such movements come from the same distribution as movements before September 11.

[17] The model could be extended to relax this strict division into phases. One might want to assume that both news reports and dividends are informative about structural change at all times. However, in such a setup, one would still like to let the informativeness of news reports decrease over time relative to that of dividends. It is plausible that the main effects of our setup would carry over to this more general environment.

ambiguous. This feature is modeled via a set of likelihoods ℓ, where

$$\ell\left(s_t = 1|\theta^{hi}\right) = \ell\left(s_t = 0|\theta^{lo}\right) = \lambda \in [\underline{\lambda}, \overline{\lambda}], \tag{15}$$

with $\underline{\lambda} > \frac{1}{2}$. Beliefs about signals up to time T are represented by the parameter space $\Theta = \left\{\theta^{hi}, \theta^{lo}\right\}$, the single prior given by μ and the set of likelihoods $\mathcal{L}$ defined by (15). The special case $\underline{\lambda} = \overline{\lambda}$ is a Bayesian model. To ease notation, assume that signals continue to arrive after T, but that for $t > T$, $\ell\left(s_t = 1|\theta^{hi}\right) = 1 = \ell\left(s_t = 0|\theta^{lo}\right)$.

In terms of the notation of Section 2, the state space is $S = \{0, 1\} \times \mathbb{R}$. Since Y is independent of s, the one-step-ahead beliefs $\mathcal{P}_t\left(s^t, Y^t\right)$ for $t \leq T$ are given by the appropriate product of one-step-ahead beliefs about s_{t+1} and the conditional probability law for Y_{t+1}. Preferences over consumption streams are then defined recursively by

$$V_t\left(c; s^t, Y^t\right) = \min_{p_t \in \mathcal{P}_t(s^t, Y^t)} \left(c_t^{1-\gamma} + \beta E^{p_t}\left[\left(V_{t+1}\left(c; s^{t+1}, Y^{t+1}\right)\right)^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}, \tag{16}$$

where β and γ are the discount factor and the coefficient of relative risk aversion, respectively. Since only the signals are ambiguous, the minimization in (16) may be viewed as a choice over sequences $\lambda^{t+1} = (\lambda_1, ..., \lambda_{t+1})$ of precisions.

Connection between Truth and Beliefs

Discipline on beliefs is imposed in two ways. First, as above, assume that the true precision λ lies in $[\underline{\lambda}, \overline{\lambda}]$. This condition ensures that an agent's view of the world is not contradicted by the data. Suppose the agent looks back at the history of signals after he is told the true parameter at time T. If he is Bayesian ($\underline{\lambda} = \overline{\lambda}$), the distribution of the signals at the true parameter value is the same as the true distribution of the signals. In this sense, the agent has interpreted the signals correctly.[18] More generally, an ambiguity averse agent contemplates many 'theories' of how the signal history has been generated, each corresponding to a different sequence of precisions. One might thus be concerned that theories that do not satisfy $\lambda_t = \lambda$ infinitely often are contradicted by the data. However, this is not the case if $\lambda \in [\underline{\lambda}, \overline{\lambda}]$: there exists a large family of signal processes with *time varying precision* $\lambda_t \in [\underline{\lambda}, \overline{\lambda}]$ that cannot be distinguished from the true distribution on the basis of *any* finite sample.[19] While some of these processes will appear less likely than others in the short run, any of them is compatible with a sample that looks i.i.d. with precision λ. An agent who believes in the whole range $[\underline{\lambda}, \overline{\lambda}]$ need not, with hindsight, feel that he interpreted the signals incorrectly.

The second way in which discipline is imposed on beliefs is through the restriction that agents would learn the true state θ^{hi} even if it were not revealed at $T + 1$. This

[18] For example, he has not been "overconfident", interpreting every signal as more precise than it actually was.

[19] To construct such precision sequences, pick any ω such that $\omega\underline{\lambda} + (1 - \omega)\overline{\lambda} = \lambda$. Let $\tilde{\lambda}_t$ be an i.i.d. process valued in $\left\{\underline{\lambda}, \overline{\lambda}\right\}$ with $\Pr\left(\tilde{\lambda}_t = \underline{\lambda}\right) = \omega$. For almost every realization (λ_t) of $\left(\tilde{\lambda}_t\right)$, the empirical distribution of the nonstationary signal process with precision sequence (λ_t) converges to the true distribution of the signals. See Nielsen [22] for a formal proof.

precludes an excessively pessimistic interpretation of news. A sufficient condition is that the posterior probability of θ^{hi}, $\mu_t\left(s^t, \lambda^t\right)$, converges to 1 if the truth equals the *lower bound* of the precision range:

$$\lim_{t\to\infty} \min_{\lambda_1^t} \ \mu_t\left(s^t, \lambda^t\right) = 1, \ \text{ a.s. for } s_t \text{ i.i.d. with } \Pr\left(s_t = 1\right) = \underline{\lambda}. \tag{17}$$

If $\overline{\lambda}$ were too large for given $\underline{\lambda}$, agents could interpret negative signals as very precise and never be convinced that the true state has occurred if the fraction of good signals is $\underline{\lambda}$. Thus the condition bounds $\overline{\lambda}$ for a given $\underline{\lambda}$.

Supporting Measure and Asset Prices

Following [15], equilibrium asset prices can be read off standard Euler equations once a (one-step-ahead) "supporting measure" that achieves the minimum in (16) has been determined. Suppose that the intertemporal elasticity of substitution is greater than one. It is then easy to show that continuation utility is always higher after good news ($s = 1$) than after bad news ($s = 0$). Thus the sequence of precisions $\left(\lambda^{*t+1}\right)$ that determines the supporting measure at time t and history s^t is chosen to *minimize* the probability of a high signal in $t+1$. For the past signals s^t, this requires maximizing the precision of bad news ($\lambda_j^* = \overline{\lambda}$ if $s_j = 0$) and minimizing the precision of good news ($\lambda_j^* = \underline{\lambda}$ if $s_j = 1$). For the future signal s_{t+1}, it requires maximizing (minimizing) the precision λ_{t+1}^* whenever news are more likely to be bad (good) next period, that is, whenever the posterior probability of θ^{hi} is smaller (larger) than $\frac{1}{2}$.

Let q_t denote the price of the Lucas tree. Since signals and dividends are independent for $t \leq T$, the price-dividend ratio $v_t = q_t/Y_t$ depends only on the sequence of signals. It satisfies the difference equation

$$v_t\left(s^t\right) = \hat{\beta} E_t^*\left[\left(1 + v_{t+1}\left(s^{t+1}\right)\right)\right]$$

where E_t^* denotes expectation with respect to the (one-step-ahead) supporting measure and where the new discount factor $\hat{\beta} = \beta e^{(1-\gamma)\theta^{hi} - \frac{1}{2}(1-\gamma)^2\sigma^2}$ is adjusted for dividend risk, where γ here is the coefficient of relative risk aversion. Once θ has been revealed at date $T+1$, the price dividend ratio settles at a constant value.

Calibration

We set the discount rate to 4% p.a. and the coefficient of relative risk aversion to one half. The average growth rate of dividends is fixed to match the price-dividend ratio, yielding a number of 5.2% p.a. This is clearly larger than the historical average, which reflects the high p/d ratio. The volatility of consumption is set at the historical value of 2% p.a. reported by Campbell [7] for postwar data. Finally, we assume that the potential permanent shock corresponds to a drop in consumption growth of .5 % p.a. In steady state, this would correspond to a price-dividend ratio of 61.

Having fixed these parameters, we infer, for every learning model, the sequence of signals that must have generated our price-dividend ratio sample if the model is correct. If the signals had a continuous distribution, this map would be exact. Here we assume

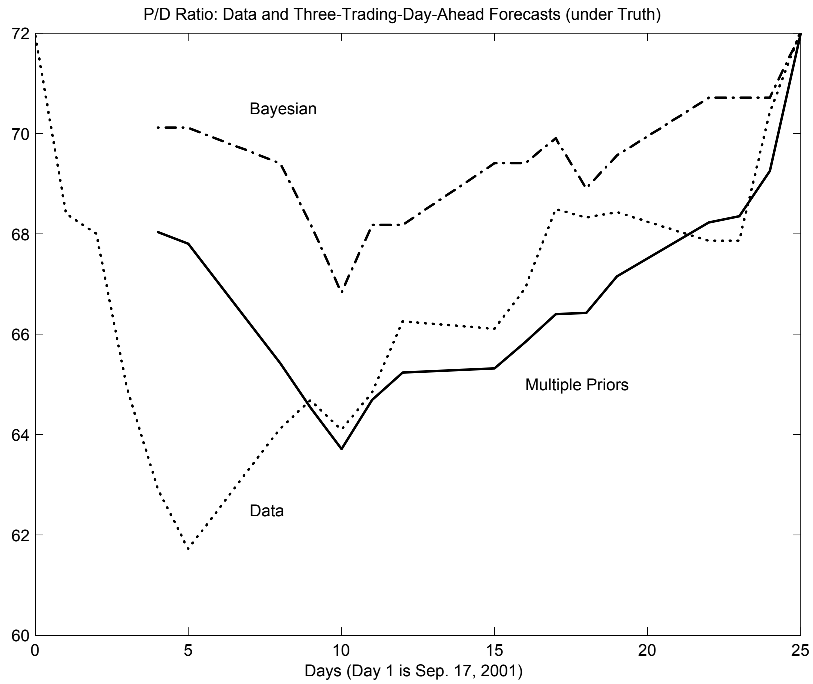

Figure 2: Data and In-Sample Forecasts for 9/11 Calibration.

that agents observe 20 signals per day. We then compute the model-implied price path that best matches the data. While the price distribution is still discrete, it is sufficiently fine to produce sensible results. A model is discarded if its 'pricing errors' are larger than .5 at any point in time. Finally, we compute the likelihood conditional on the first observation for each model, using the distribution of the fitted price paths. This is a useful criterion for comparing models, since the first observation is basically explained by the choice of the prior.

Numerical Results

To select a Bayesian model, we search over priors μ and precision parameters λ to maximize the likelihood. This yields an interior solution for both parameters. For example, for precision, the intuition is as follows: the path of posteriors is completely determined by the path of p/d ratios. Thus performance differences across Bayesian models depend on how likely the path of posteriors is under the truth. If precision is very large, then it is highly unlikely that there could have been enough bad news to explain the initial price decrease. In contrast, if precision is very small, then signals are so noisy that posteriors do not move much in response to any given news. Highly unlikely 'clusters' of first bad and then good news would be required to explain the price path. This trade-off gives rise to an interior solution for precision.

To select a multiple-priors model, we need to specify both the true precision and the range of precisions the agent thinks possible. To sharpen the contrast with Bayesian models, we focus on models where ambiguity is large; we set $\overline{\lambda}$ slightly (.001) below the upper bound associated with the requirement (17). We also assume that the truth corresponds to $\underline{\lambda}$.[20] With these two conventions, we search over $\underline{\lambda}$ to find our favorite multiple-priors model. The favorite multiple-priors model begins with a much higher prior probability, and the precision range for $\lambda \in [.58, .608]$ is higher than the precision for the best Bayesian model, $\lambda = .56$. The multiple-priors model (log likelihood = -33.29) outperforms the Bayesian model (log likelihood = -36.82). Inspection of the one-step-ahead conditional likelihoods reveals the main source of the difference: the multiple-priors model is better able to explain the downturn in the week of September 17. The models do about the same during the recovery. Figure 2 plots, together with the data, three-trading-day-ahead in-sample forecasts. This shows that the Bayesian model predicts a much faster recovery than the multiple-priors model throughout the sample.

The result shows how the effects discussed in the previous section operate in a setting with many signals. The two models represent two very different accounts of market movements in September 2001. According to the Bayesian story, all price movements reflect changes in beliefs about future growth. In particular, the initial drop in prices arose because market participants expected a permanent drop in consumption of .2%. During the first week, bad news increased the expected drop to almost .5%. In contrast, the ambiguity story says that agents begin with a prior opinion that basically nothing has changed. However, they know that the next few weeks will be one of increased confusion and uncertainty. Anticipation of this lowers their willingness to pay for stocks. In particular, they know that future bad news will be interpreted (by future ambiguity averse market participants) as very precise, whereas future good news will be interpreted as noisy. This makes it more likely that the market will drop further in the short run than for the Bayesian model.

For representative agent asset pricing models with multiple-priors utility, there is always an observationally equivalent Bayesian model that yields the same equilibrium price. This begs the question why one should not consider this Bayesian model directly.[21] Here, the reason is that this Bayesian model cannot be motivated by the same plausible *a priori* view of the environment as our ambiguity aversion model. We want to capture a scenario where signals are generated by a memoryless mechanism, and where precision does not depend on the state of the world: learning in good times is not expected to occur at a different speed than in bad times. An ambiguity aversion model with these features outperforms a Bayesian model with these features. Some other Bayesian model which does not have these features is not of interest. In addition, such a model would yield misleading comparative static predictions. The observationally equivalent model is much like a 'reduced form' which is not invariant to changes in the environment.

[20] Strictly speaking, this polar case is not permitted by the restriction that the truth lie in the interior of the precision range. However, there is always an admissible model arbitrarily close to the model we compute.

[21] This model would be an expected utility model with pessimistic beliefs, similar to that in Abel [1].

5 Appendix

The key step in calculating moments of prices and returns in Section 3.3 is to find moments of $\gamma_t s_t$, where γ_t denotes the random variable that is equal to $\bar{\gamma}$ when $s_t \leq 0$, and equal to $\underline{\gamma}$ otherwise. We summarize the properties of $\gamma_t s_t$ here. All moments are calculated under the true signal distribution. Since $E[s_t|\gamma_t = \bar{\gamma}] = E_t[s_t|s_t \leq 0])$, the calculations make heavy use of formulas for moments of truncated normal distributions, $E[s_t|s_t \geq 0] = \sigma_s\sqrt{\frac{2}{\pi}}$, $E[s_t^2|s_t \geq 0] = \sigma_s^2$ and $E[s_t^3|s_t \geq 0] = 2\sqrt{\frac{2}{\pi}}\sigma_s^3$. The mean and variance of $\gamma_t s_t$ are given by

$$\begin{aligned}
E[\gamma_t s_t] &= E[\gamma_t E[s_t|\gamma_t]] \\
&= \frac{1}{2}\bar{\gamma}\left(-\sqrt{\frac{2}{\pi}}\sigma\right) + \frac{1}{2}\underline{\gamma}\left(\sqrt{\frac{2}{\pi}}\sigma\right) \\
&= -\left(\bar{\gamma} - \underline{\gamma}\right)\frac{\sigma_s}{\sqrt{2\pi}}, \quad \text{and}
\end{aligned}$$

$$\begin{aligned}
var(\gamma_t s_t) &= E\left[\gamma_t^2 s_t^2\right] - E[\gamma_t s_t]^2 = E\left[\gamma_t^2 E\left[s_t^2|\gamma_t\right]\right] - E[\gamma_t E[s_t|\gamma_t]]^2 \\
&= E\left[\gamma_t^2 var(s_t|\gamma_t)\right] + var(\gamma_t E[s_t|\gamma_t]) \\
&= \frac{1}{2}\left(\bar{\gamma}^2 + \underline{\gamma}^2\right)\sigma_s^2\left(1 - \frac{2}{\pi}\right) + \frac{1}{4}\left(\sqrt{\frac{2}{\pi}}\sigma_s\left(\bar{\gamma} + \underline{\gamma}\right)\right)^2 \\
&= \frac{1}{2}\sigma_s^2\left(\bar{\gamma}^2 + \underline{\gamma}^2 - \frac{1}{\pi}\left(\bar{\gamma} - \underline{\gamma}\right)^2\right).
\end{aligned}$$

To determine the variance of returns, we also need the term

$$\begin{aligned}
cov(\gamma_t s_t, u_{t+1}) &= E[\gamma_t s_t u_{t+1}] = E[\gamma_t s_t E[u_{t+1}|s_t, \gamma_t]] \\
&= \gamma^* E\left[\gamma_t s_t^2\right] = \gamma^* E\left[\gamma_t E\left[s_t^2|\gamma_t\right]\right] \\
&= \gamma^* \frac{\bar{\gamma} + \underline{\gamma}}{2}\sigma_s^2 = \frac{\bar{\gamma} + \underline{\gamma}}{2}\sigma_u^2.
\end{aligned}$$

The third centered moment of $\gamma_t s_t$ is

$$\mu_3(\gamma_t s_t) = E\left[\gamma_t^3 s_t^3\right] - E[\gamma_t s_t]^3 - 3E[\gamma_t s_t]\, var(\gamma_t s_t).$$

The only as yet unknown term is

$$E\left[\gamma_t^3 s_t^3\right] = E\left[\gamma_t^3 E\left[s_t^3|\gamma_t\right]\right] = -\left(\bar{\gamma}^3 - \underline{\gamma}^3\right)\sqrt{\frac{2}{\pi}}\sigma_s^3,$$

where the second equality follows from the third moment of the truncated normal distribution. Some algebra then delivers

$$\mu_3(\gamma_t s_t) = -\frac{\sigma_s^3}{\sqrt{2\pi}}\left(\frac{\left(\bar{\gamma} - \underline{\gamma}\right)^3}{\pi} + \frac{3}{2}\bar{\gamma}\underline{\gamma}\left(\bar{\gamma} - \underline{\gamma}\right) + \frac{1}{2}\left(\bar{\gamma}^3 - \underline{\gamma}^3\right)\right) < 0.$$

Since $\sigma_s^2 = \sigma_u^2/\gamma^*$, this expression is indeed proportional to σ_u^3, where the factor of proportionality depends only on $\bar{\gamma}$, $\underline{\gamma}$ and γ^*.

Finally, consider the third centered moment of returns. Since $u_{t+1} - \gamma_t s_t$ and $\gamma_{t+1} s_{t+1}$ are independent, we just need to compute

$$\begin{aligned}\mu_3 \left(u_{t+1} - \gamma_t s_t\right) &= \mu_3 \left(u_{t+1}\right) - \mu_3 \left(\gamma_t s_t\right) \\ &\quad +3E\left[u_{t+1} \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)^2\right] - 3E\left[u_{t+1}^2 \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)\right].\end{aligned}$$

The first term is zero since since u_{t+1} is normal. In the third and fourth terms, use conditional normality of u_{t+1} and $E\left[s_t\right] = 0$ to get

$$\begin{aligned}E\left[E\left[u_{t+1}|s_t, \gamma_t\right] \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)^2\right] &= E\left[\left(\gamma^* s_t\right) \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)^2\right] \\ &= \gamma^* E\left[s_t \left(\gamma_t^2 s_t^2 - 2\gamma_t s_t E\left[\gamma_t s_t\right]\right)\right], \\ E\left[E\left[u_{t+1}^2|s_t, \gamma_t\right] \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)\right] &= E\left[\left(\gamma^{*2} s_t^2 + \left(1 - \gamma^*\right) \sigma_u^2\right) \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)\right] \\ &= \gamma^{*2} E\left[s_t^2 \left(\gamma_t s_t - E\left[\gamma_t s_t\right]\right)\right].\end{aligned}$$

Putting terms together, we have,

$$\begin{aligned}\mu_3 \left(R_{t+1}\right) \left(1 - \beta\kappa\right)^3 &= \left(\beta^3 - 1\right) \mu_3 \left(\gamma_t s_t\right) \\ &\quad +3\gamma^* \left(E\left[\gamma_t \left(\gamma_t - \gamma^*\right) s_t^3\right] - E\left[\gamma_t s_t\right] \left(2E\left[\gamma_t s_t^2\right] - \gamma^* E\left[s_t^2\right]\right)\right) \\ &= \left(\beta^3 - 1\right) \mu_3 \left(\gamma_t s_t\right) - \frac{\sigma_s^3}{\sqrt{2\pi}} \gamma^* \left(\bar{\gamma} - \underline{\gamma}\right) \left(\bar{\gamma} + \underline{\gamma} - \gamma^*\right).\end{aligned}$$

Since $\sigma_s^2 = \sigma_u^2/\gamma^*$, we obtain the expression in the text.

References

[1] A. Abel, An exploration of the effects of pessimism and doubt on asset returns, *J. Econ. Dynam. Control* 26 (2002), 7-8, 1075-1092.

[2] G. Bekaert and G. Wu, Asymmetric volatility and risk in equity markets, *Rev. Finan. Studies* 13 (2002), 1-42.

[3] P. Bossaerts, Learning-induced securities price volatility, Caltech, 2001.

[4] M. Brandt, Q. Zeng and L. Zhang, Equilibrium stock return dynamics under alternative rules of learning about hidden states, *J. Econ. Dynamics and Control* 28, 2004, 1925-1954.

[5] A. Brav and J. Heaton, Competing theories of financial anomalies, *Rev. Finan. Studies* 15 (2002), 575-606.

[6] M. J. Brennan and Y. Xia, Stock price volatility and the equity premium, *J. Monet. Econ.* 47 (2001), 249-283.

[7] J. Campbell, Asset Prices, consumption and the business cycle, in *Handbook of Macroeconomics,* J. Taylor and M. Woodford, eds., Amsterdam: North Holland, 1999.

[8] J. Campbell, M. Lettau, B. Malkiel and Y. Xu, Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk, *J. Finance* 56 (2001), 1-43.5

[9] Z. Chen and L.G. Epstein, Ambiguity, risk and asset returns in continuous time, *Econometrica* 70 (2002), 1403-43.

[10] K. Daniel, D. Hirshleifer, and A. Subrahmanyam, Investor psychology and security market under- and overreactions, *J. Finance* 53 (1998), 1839-1886.

[11] K. Daniel and S. Titman, Market reactions to tangible and intangible information, *J. Finance*, forthcoming.

[12] L.G. Epstein and M. Schneider, Recursive multiple-priors, *J. Econ. Theory* 113 (2003), 1-31.

[13] L.G. Epstein and M. Schneider, IID: independently and indistinguishably distributed, *J. Econ. Theory* 113 (2003), 32-50.

[14] L.G. Epstein and M. Schneider, Learning under ambiguity, mimeo, 2005.

[15] L.G. Epstein and T. Wang, Intertemporal asset pricing under Knightian uncertainty, *Econometrica* 62 (1994), 283-322.

[16] I. Gilboa and D. Schmeidler, Maxmin expected utility with nonunique prior, *J. Math. Econ.* 18 (1989), 141-153.

[17] C. Harvey and A. Siddique, Conditional skewness in asset pricing tests, *J. Finance* 55 (2000), 1263-1295.

[18] B. N. Lehmann, Residual risk revisited, *J. Econometrics,* 45 (1990), 71-97.

[19] B.G. Malkiel and Y. Xu, Idiosyncratic risk and security returns, UT Dallas 2002.

[20] J. Lewellen and J. Shanken, Learning, asset pricing tests and market efficiency, *J. Finance* 58 (2002), 1113-1145.

[21] M. Marinacci, Limit laws for non-additive probabilities, and their frequentist interpretation, *J. Econ. Theory* 84 (1999), 145-195.

[22] C.K. Nielsen, Rational belief structures and rational belief equilibria, *Economic Theory* 8 (1996) 3, 399-422.

[23] A. Timmermann, How learning in financial markets generates excess volatility and predictability in stock returns, *Quart. J. Economics* 108 (1993), 1135-1145.

[24] A. Timmermann, Excess volatility and predictability of stock returns in autoregressive dividend models with learning, *Rev. Ec. Stud.* (1996), 523-557.

[25] P. Veronesi, Stock market overreaction to bad news in good times: a rational expectations equilibrium model, *Rev. Finan. Stud.* 12 (1999), 976-1007.

[26] P. Veronesi, How does information quality affect stock returns?, *J. Finance* 55 (2000), 807-837.

[27] K. West, Dividend innovations and stock price volatility, *Econometrica* 56 (1988), 37-61.