HOPE FOR THE BEST, PLAN FOR THE WORST

KENNETH KASA AND XIAOWEN LEI

ABSTRACT. This paper studies asset pricing when individuals struggle to strike a balance between doubt and hope. We argue this internal struggle is consistent with recent evidence from neuroscience. We operationalize it using the robust control and filtering approach of Hansen and Sargent (2008). Our key innovation is to assume that filtering is *optimistically* biased. Investment decisions, however, reflect doubts about model specification, and are pessimistically biased. We show that doubts about model specification, combined with optimistically distorted beliefs about dividend growth can not only explain low average price/dividend ratios and high average returns, but can also generate the sort of large procyclical swings in price/dividend ratios that are observed in the data. High and volatile returns occur despite the fact that investors have low (approximately logarithmic) risk aversion. The model's belief distortions are empirically plausible, with detection error probabilities over 10%.
JEL Classification Numbers: G12, D81

I am large, I contain multitudes. – WALT WHITMAN (*Song of Myself*, 1855)

1. INTRODUCTION

The 1980s and 90s were a dismal period for asset pricing research. The empirical flaws of standard models highlighted by Shiller (1981), Hansen and Singleton (1983) and Mehra and Prescott (1985) were remarkably resiliant to theoretical modifications. Fixing one problem invariably exacerbated other problems. Economists could not even explain the unconditional first moments of risky versus safe assets.

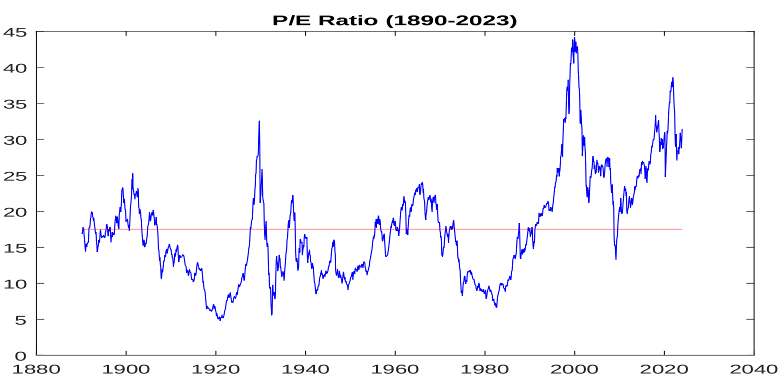
Since then the asset pricing literature has advanced significantly. The equity premium and risk free rate puzzles are no longer puzzles. If you want to know why mean returns on risky assets are so high while returns on safe assets are so low there are a menu of model options available. There are theories based on 'exotic preferences', e.g., the backward-looking habit persistence model of Campbell and Cochrane (1999) or the forward-looking long-run risk model of Bansal and Yaron (2004). There are models based on even more exotic preferences, e.g., Prospect Theory (Benartzi and Thaler (1995)). Alternatively, there are theories based on beliefs, which deviate from the Rational Expectations Hypothesis, e.g., Barillas, Hansen, and Sargent (2009) or Ju and Miao (2012). Other theories add frictions, leading to incomplete markets. Although initially this seemed like the obvious candidate, it proved to be surprisingly difficult to implement empirically. While Constantinides and Duffie (1996) showed that highly persistent idiosyncratic labor income shocks

---

Date: April, 2025.
We would like to thank Sev Hou for helpful comments.

could in principle explain the equity premium, data on labor income didn't seem to support this. However, recent work by Constantinides and Ghosh (2017) obtains promising results by allowing for 'rare disasters' in idiosyncratic labor income, as in Barro (2006).

Given these recent successes, the asset pricing literature has become more ambitious, and is now focused on a broader spectrum of data, e.g., the *conditional* moments of asset returns and their comovements with other observed data. Figure 1 in particular remains a challenge. It depicts (annualized) monthly data on the Price/Earnings ratio in the USA for the period 1890-2023 (obtained from Shiller's website).



The flat line is the unconditional mean, around 17.5. Its level reflects of the relatively high mean return on the stock market. Taking the reciprocal as an estimate of the dividend yield, and adding 2% as the mean growth rate of dividends yields a 7-8% (annualized, real) return. What is most striking about Figure 1, however, is the volatility. The PE rose above 30 during the 'roaring 20s' before it collapsed at the outset of the Great Depression. It rose even higher, to above 40, at the peak of the 'dotcom boom' of the 1990s. Conversely, it fell into the single digits during the bear markets of the early 1900s and 1970s. This volatility was the original inspiration for Shiller's (1981) classic paper, which showed that these fluctuations could not be explained by revisions of rational expectations of future dividend growth. Shiller argued that bubbles provide the most likely explanation. Other researchers focus on the possibility of a strongly countercyclical risk premium. No consensus has yet emerged on either possibility.

Our paper proposes a new possibility. It is not based on bubbles or risk aversion. Nor is it based on incomplete markets and idiosyncratic risk. Instead, it is based on beliefs. It is perhaps most closely related to Hansen and Sargent (2010). Hansen and Sargent were responding to criticisms of the Bansal-Yaron (2004) long-run risks model. For example, Beeler and Campbell (2012) noted that empirical evidence in support of the model's two key ingredients, i.e., stochastic growth and stochastic volatility, is rather weak. Hansen and Sargent (2010) introduce model uncertainty, and show that it is enough that agents *suspect* that growth rates *might* drift, even if in fact they do not. They are also able to generate apparently strong countercyclical fluctuations in risk prices by incorporating robust model averaging. Agents think that bad times will be persistent, whereas good times will be transitory. However, in Hansen and Sargent (2010) these fluctuations do not represent changes in the willingness of agents to bear risk. Instead, they represent changes in the price of model uncertainty.

We make one small, but important, change to Hansen and Sargent's Fragile Beliefs model. Instead of supposing that agents revise their beliefs in a defensively pessimistic manner, we assume beliefs are distorted *optimistically.* Operationally, this is accomplished by simply 'flipping the sign' of Hansen and Sargent's robust filtering parameter. Now agents think good times will persist, while bad times are transitory. This generates the sort of prolonged booms and sharp sudden crashes that are observed in the data (see Figure 1). To generate procyclical movements in the PE ratio, we suppose relative risk aversion is slightly less than one.[1] In some respects, our model resembles the 'intrinsic bubbles' model of Froot and Obstfeld (1991). However, in our model transversality conditions are obeyed and we do not confront the uncomfortable question of how agents' beliefs coordinate on a bubble. Even more important, we maintain Hansen and Sargent's assumption that agents doubt the specification of their dividend growth models, and so introduce *pessimistic* drift distortions into their models. This allows us to generate a low mean PE ratio (and hence high mean equity returns), while at the same time generating high volatility and procyclical movements in the PE ratio.

Anytime a model is based on distorted, 'nonrational' beliefs, it is important to demonstrate that such distortions are plausible, and do not leave too many $5 bills on the sidewalk. We respond to this challenge in two ways. First, following Hansen and Sargent (2008), we argue that a bit of defensive pessimism is *not* irrational. It is simply an acknowledgment of the limits of your own knowledge. As such, defensive pessimism could even be interpreted as a form of 'higher-order' rationality. The real question is why agents do not respond to model uncertainty in a Bayesian manner. For that, we refer readers to the detailed discussion in Hansen and Sargent (2008). Here we simply note that the founding father of Bayesian decision theory, Savage (1972), cautioned against the misapplication of 'small worlds' Bayesian methods to 'large worlds' problems (see, e.g., p. 16). We claim that the problem of investing in the stock market and placing value on new technologies is a large worlds problem. Even if one grants us our nonBayesian approach, we still confront the practical question of *quantifying* the degrees of optimism and pessimism. For that we again follow Hansen and Sargent (2008) by calibrating our belief parameters to 'detection error probabilities', so that agents do not base their forecasts and investment decisions on empirically implausible models of dividend growth.

Our second response is to argue that the distortions we employ are consistent with evidence from the psychology literature, and more recently, from the neuroscience literature. The novel aspect of our model is that agents are being *simultaneously* optimistic *and* pessimistic. They base dividend forecasts on optimistically biased estimates of the underlying unobserved growth state. At the same time, they base investment decisions on a pessimistically biased model of the process generating observed dividends. Isn't this schizophrenic, or at least bipolar?[2] The key to reconciling this apparent contradiction is to note that the problems of forming and revising beliefs about the future on the one

[1]If risk aversion exceeded one, then the higher discount rate associated with good news about future dividends would dominate the higher expected cash flows, and the PE ratio would be countercyclical.

[2]Kasa (2012) showed how in certain cases these two biases could exactly offset each other and produce what looks like, to an outside observer, Rational Expectations.

hand, and taking actions based on those beliefs on the other, involve different cognitive processes, and take place (primarily but not exclusively) in different parts of the brain. As discussed below, neuroscientific evidence suggests that learning and belief revision takes place in the left hemisphere of the brain, and is optimistically biased, in the sense that positive signals elicit over reactions while negative signals elicit under-reaction. This bias is tempered, however, by the right hemisphere, which is regarded as the locus of a person's 'alert system', which produces cautionary, pessimistically biased responses to the information being transmitted from the left hemisphere. Of course, the Pop Science notion that the world is divided into left-brain people, who are analytical, and right-brain people, who are emotional, is a vast oversimplification. The fact is, unless afflicted by a stroke or some other brain injury, everyone is both a left and right-brained person, and the two halves work together in intricate ways to optimize an individual's response to an ever changing environment. Still, brain lateralization *is* real, and should not be too surprising to economists, who are attuned to the benefits of specialization and the gains from trade.

By relating stock market dynamics to nonrational belief dynamics our paper is consistent with recent work by Bordalo, Gennaioli, Porta, and Shleifer (2024). They directly measure expectations using survey data, and argue that it is important to distinguish between short-term and long-term growth forecasts. Bordalo et. al. (2024) show that long-term earnings growth forecasts overreact to good news, and that this overreation can explain many asset pricing puzzles, both in aggregate time series data and in firm level cross-sectional data. There are three key differences between our paper and theirs. First, and most importantly, we provide an explicit 'micro-foundation' for belief overreaction, based on recent advances in neuroscience. Second, we argue that a form of persistent pessimism is also important, which allows our model to also explain high average returns. Third, we validate the empirical magnitudes of our belief distortions using detection error probabilities, rather than using survey data. This allows us to consider a much longer sample period, which can be important when focusing on low frequency distortions. Despite these differences, we view our work as quite complementary to theirs.

The remainder of the paper is organized as follows. Section 2 reviews evidence from the psychology and neuroscience literatures on the co-existence of optimism and pessimism. As a benchmark, Section 3 computes asset prices when agents are Bayesians. Here we rediscover the equity premium and volatility puzzles. Asset prices are too high (and mean returns too low), and are nearly unresponsive to revisions in beliefs about future dividend growth. Section 4 presents our main results. We show how drift distortions in the dividend growth model and Kalman filtering equation can be used to operationalize and quantify the internal tensions between optimism and pessimism that agents struggle to balance. Now prices become much more responsive to belief revisions, and come close to matching the volatility of the observed Price/Earnings ratio displayed in Figure 1, although it falls short of matching the extreme spikes witnessed during the late 1920s and late 1990s. Section 5 shows that the belief distortions we introduce are empirically plausible, in the sense that detection error probabilities are not too low. Finally, Section 6 offers concluding remarks, and a technical appendix contains proofs and mathematical details.

2. Optimism & Pessimism: Evidence From Psychology and Neuroscience

The central premise of this paper is that individuals confront a tension between optimism and pessimism, and that this tension manifests itself differently in different decisions. We start by briefly reviewing evidence in support of this tension, coming first from the psychology literature, and then more recently from the neuroscience literature.

2.1. **Psychology.** We are certainly not the first persons to suggest that individuals balance optimism and pessimism. The homespun wisdom of our paper's title is so old that no one knows who first said it. We also would bet that references to it could be found in Shakespeare, if one cared to look. However, the first formal, scientific study that we are aware of is Kahneman and Lovallo (1993). They argue that the combination of 'bold forecasts' and 'timid actions' derives from a common underlying psychological bias, namely, 'narrow framing'. Narrow framing refers to the tendency to view problems as more unique than they really are. Individuals become excessively risk averse because they do not adequately account for the risk reducing benefits of diversification and the law of large numbers. In the context of forecasting, Kahneman and Lovallo refer to narrow framing as the 'inside view', which anchors predictions of exogenous future events on your own plans and scenarios, thus creating an illusion of excessive control and optimism.

More recently, Haselton and Nettle (2006) adopt an evolutionary perspective. They appeal to something called 'Error Management Theory', and argue that if the costs of false positives and false negatives have been asymmetric over the course of evolutionary history, then natural selection could produce a population of 'paranoid optimists'. The provocative example they use to motivate their analysis is the case of 'sexual inference', i.e., the tendency of men to overestimate a woman's sexual interest. If the cost of a missed opportunity to spread your genes exceeds the cost of a slap in the face, then men might develop an optimism bias when gauging their own chances.[3] Although this theory does not directly predict that individuals display optimism when forecasting, but pessimism when investing, it could be made to generate such behavior under plausible conditions. In particular, if the cost of being too optimistic when deciding on a portfolio allocation exceeds the cost of being too pessimistic, then agents might appear to be excessively risk averse. This is in fact precisely the theory behind 'loss aversion' and Prospect Theory. It could even arise with standard preferences in the presence of borrowing constraints or subsistence levels. Conversely, conditional on participation, agents might be optimistically biased when forecasting if the costs of pessimism about future returns exceed the costs of being too optimistic. This combination might occur in a 'Gold Rush' or 'Superstars' economy, which produces small probabilities of huge fortunes.

At a more general level, our model of Hoping For the Best While Planning for the Worst is consistent with a common thread running through much of the psychology literature, namely, the view that individuals are composites of competing, or duelling selves. Kahneman's (2011) well known book, *Thinking, Fast and Slow*, contains dozens of examples. This composite view of individuals has even become common within economics, examples being the hyperbolic discounting model of Laibson (1997) and Hansen and Sargent's

[3]If you've seen the movie *Dumb and Dumber* you cannot help but be reminded of Jim Carrey's reaction when told by a woman that his chances are one in a million. He breaks into a wide grin and responds, "so you're saying I have a chance!"

(2008) model of ‘evil agents’ and robust control. Interestingly, the notion that individuals are composites of competing interests is also consistent with evidence from the rapidly developing field of neuroscience. We turn to this evidence next.

2.2. **Neuroscience.** Evidence of brain lateralization emerged somewhat by chance. Epileptic seizures are abnormal electrical discharges that spread from one half of the brain to the other. Currently, there are effective medications that inhibit these discharges. This wasn't the case back in the 1960s, however. A neurologist named Roger Sperry proposed the idea of cutting the corpus callosum (a tract of nerve fibers separating the two halves of the brain) in an effort to contain and mitigate the electrical discharges causing epileptic seizures. The procedure turned out to be a success, and Sperry was awarded the Nobel Prize in 1981.

Over time, however, as these split-brain patients continued to be tested and monitored, some subtle changes began to be detected. For example, patients would often move their heads when processing visual information, or make symbolic hand movements. Using the fact that human vision is ‘cross-wired’, with the left hemisphere processing information from the right visual field (and vice versa), experiments were run where different kinds of information were shown to first one eye and then the other. It was discovered that when a picture of a face was shown to the right eye (and so processed by the left-side of the brain) it was unrecognized, but was easily recognized when shown to the left eye. Conversely, when a passage of text was shown to the left eye, the patient could not read it, but had no trouble when it was shown to the right eye. Although the idea that language and speech are processed in the left-side of the brain can be traced back to 1800s, these studies were the first solid evidence in support of it.[4]

For the purposes of our paper, the most interesting of these kinds of studies were performed by Wolford, Miller, and Gazzaniga (2000). They engaged subjects in a probability-guessing experiment. In these experiments subjects are shown random sequences of two outcomes, e.g., a red or green light appears on their computer screen. The probabilities of the two outcomes are different, e.g., red appears with probability 3/4 while green appears with probability 1/4. Subjects are rewarded based on the number of correct guesses. It has long been known that humans engage in ‘frequency matching’. In the above example, they guess red about 75% of the time. This is clearly suboptimal. To maximize your payoff, you should *always* guess red. Interestingly, nonhumans (e.g., rats) usually optimize. Inspired by the earlier research on split-brain patients, Wolford et. al. added a new twist to this classic experiment. They played the game separately on the left- and right-hemispheres (by displaying the events to only the right and left eyes, respectively). Interestingly, they found that the left hemisphere engages in frequency matching, whereas the right hemisphere optimizes.

In his popular book *Human: The Science Behind What Makes Your Brain Unique*, Gazzaniga (2008), who is a student of Sperry, summarizes the left/right brain research as follows:

[4]The idea first arose from autopsies on people with speech impediments. People with speech impediments were observed to have damage in the left-side of their brains.

> *This is consistent with the hypothesis that the left-hemisphere interpreter constructs theories to assimilate perceived information into a comprehensible whole. By going beyond simply observing events to asking why they happened, a brain can cope with such events more effectively if they happen again...Accuracy remains high in the right hemisphere, however, because it does not engage in these interpretive processes. The advantage of having such a dual system is obvious. The right hemisphere maintains an accurate record of events, leaving the left hemisphere free to elaborate and make inferences about the material presented.* (Gazzaniga (2008), p. 296)

What, you might ask, does this have to do with stock market investment? Investing involves *two* conceptually different activities. One must first formulate some belief about the future consequences of any given investment. This involves developing theories and hypotheses about how the future may unfold. This is a creative process, which is akin to engaging in frequency matching. We claim this problem is primarily processed in the left-hemisphere, and is optimistically biased. Second, an individual must decide *how much* of his scarce capital he should devote to the left-hemisphere's theories. We claim this investment allocation decision is allocated to the alert system in the right hemisphere, and is pessimistically biased. How do we know the left-hemisphere is optimistically biased while the right hemisphere is pessimistic? Much of the rapid progress made in neuroscience during the past 20 years derives from the development of new technologies that allow scientists to *directly* measure and isolate brain activity. Scientists no longer have to rely on autopsies, illnesses, or accidents. These recent studies provide strong evidence that indeed the left hemisphere is optimistic while the right hemisphere is pessimistic.

The most common of these new technologies is Functional Magnetic Resonance Imaging, or fMRI. It is popular because it is relatively inexpensive and is the least invasive. It exploits the correlation that exists in the brain between neural activity and blood flow. However, because it does not directly measure neural activity, it is relatively imprecise, and must be coupled with statistical methods that extract signal from noise.[5] More accurate technologies directly measure neural activity. Two examples are Transcranial Direct Current Stimulation (tDCS) and Transcranial Magnetic Stimulation (TMS). These methods temporarily stimulate (or inhibit) neural activity by applying electrical currents or magnetic pulses to selective regions of the brain. Clearly, this is more invasive. Sharot, Kanei, Marston, Korn, Rees, and Dolan (2012) employed TMS to study whether belief updating could be influenced by disrupting either the left or right hemispheres. Participants were asked to estimate the likelihood of various adverse life events (e.g., Alzheimers, crime, etc.). After each trial participants were then given new information stating the likelihood of these events for individuals matching their characteristics (e.g., age, race, gender, etc.). Finally, participants were given the opportunity to revise their estimates based on the new information. They found that participants with a disrupted left hemisphere became relatively more responsive to 'bad news' (population probability $>$ prior), whereas participants with disrupted right hemispheres became relatively more responsive to 'good news' (population probability $<$ prior). Fecteau, Knoch, Sultani, Boggio, and

[5]fMRI studies are heavily used in the recent neuroeconomics literature. See Fehr and Rangel (2011) for a survey.

Pascual-Leone (2007) employed a tDCS study to examine risk-taking, the second stage of our investment problem. Participants were asked to play a gambling game, consisting of 100 trials of guessing the color of the box hiding a winning ticket. Each trial had six boxes of two colors, with the relative proportions differing from trial to trial. Payoffs were structured so that riskier bets yielded higher returns. Treated patients received tDCS stimulation to the right hemisphere. They found that the average bets of treated patients were significantly less risky, which supports the notion that the right hemisphere is in charge of regulating risk taking behavior.[6] These are just two examples of the many neuroscientific studies that support our maintained assumption that belief formation in the left hemisphere is optimistically biased, while risk taking behavior is regulated in the right hemisphere and is pessimistically biased. Hecht (2013) provides a comprehensive survey.

Finally, before moving on to the model, we should note that not everyone is convinced that these neuroscientific advances call for a rethinking of economic theory. In a widely cited critique of neuroeconomics, Gul and Pesendorfer (2008) state the following:

> *Since expected utility theory makes predictions only about choice behavior, its validity can be assessed only through choice evidence. If economic evidence leads us to the conclusion that expected utility theory is appropriate in a particular set of applications, then the inability to match this theory to the physiology of the brain might be considered puzzling. But this puzzle is a concern for neuroscientists, not economists....* (p. 25)
>
> *Our central argument is simple: neuroscience evidence cannot refute economic models because the latter make no assumptions or draw no conclusions about the physiology of the brain. Conversely, brain science cannot revolutionize economics because it has no vehicle for addressing the concerns of the latter. Economics and psychology differ in the questions they ask. Therefore, abstractions that are useful for one discipline will typically not be useful for the other. The concepts of a preference, a choice function, demand function, GDP, utility, etc. have proven to be useful abstractions in economics. The fact that they are less useful for the analysis of the brain does not mean that they are bad abstractions in economics.* (p. 26)

We've included this extended quotation because we agree with it. Like the behaviorists before them, early advocates of neuroeconomics were arguably too devoted to tearing down economics, rather than building it up. Our goal is to show that insights from neuroscience can indeed improve economic models, even if you're like Gul and Pesendorfer and only care about observed choices. If an alternative theory can explain asset prices equally well, we should prefer the theory that is more consistent with the 'microfoundations' of human behavior.

3. A Bayesian Benchmark

We start by studying an economy inhabited by individuals who face no tension between doubt and hope. They are Bayesians. This will provide a useful benchmark.

[6] It should be noted that the motivation for their study was to develop potential treatments for other types of risky behavior, e.g, taking drugs, not stock market investment!

3.1. **Environment.** Consider a frictionless, complete markets economy, consisting of identical infinitely lived agents. Their only choice in life is to decide how much to invest in a 'tree' yielding a flow of nondurable 'fruit', which provides consumption utility.[7] The optimal choice is characterized by an Euler equation. Exploiting the law-of-iterated-expectations and imposing a transversality condition, we can iterate forward the Euler equation and obtain the following present value asset pricing equation:

$$P_t U'(C_t) = E_t \int_t^{\infty} e^{-\delta(s-t)} U'(C_s) D_s ds \tag{3.1}$$

where δ is the rate of time preference. Notice that the marginal utility of consumption plays the role of a 'stochastic discount factor'. When times are bad and consumption is low, marginal utility is high, which raises the value of the reduced flow of dividends. This trade-off becomes clear if we assume utility takes the constant relative risk aversion form, $U(C) = \frac{1}{1-\gamma} C^{1-\gamma}$, where γ is the coefficient of relative risk aversion. Imposing the equilibrium condition, $C_t = D_t$, we get

$$P_t D_t^{-\gamma} = E_t \int_t^{\infty} e^{-\delta(s-t)} D_s^{1-\gamma} ds \tag{3.2}$$

This expresses the equilibrium price, P_t, as an expectation of the exogenous flow of future dividends. Notice that the impact of future dividends on the current price is ambiguous. If $\gamma > 1$ then the discount rate effect dominates the cash flow effect, and asset prices decline in response to good news about future dividends. Conversely, if $\gamma < 1$ then prices rise. This tends to create problems when taking this model to the data. Returns are high on average, which calls for a high γ. However, prices are also procyclical. They are high when times are good, and low in bad times. In principle, this could be consistent with $\gamma > 1$ if there is sufficient mean reversion in consumption and dividends, so that current good times create expectations of future bad times. However, the data suggest that mean reversion is quite weak. In the next section we avoid this dilemma by following Barillas, Hansen, and Sargent (2009), and introduce a new parameter that captures model uncertainty. This allows us to keep $\gamma < 1$ in order to generate procyclicality, while at the same time generating high average returns via a model uncertainty premium. For now, however, in this Bayesian world, we are stuck. We shall see that when $\gamma < 1$ the equilibrium Price/Dividend ratio is far too high (and mean returns far too low).

To evaluate the expectation in eq. (3.2) agents must take a stand on the process generating dividends. For now we assume agents have no doubts about this process. Dividends are known to follow the process

$$\frac{dD_t}{D_t} = \mu_t dt + \sigma dB_t \tag{3.3}$$

where B_t is a Brownian motion. Notice that the mean growth rate, μ_t, is stochastic. Not only that, it is unobserved. However, it is known to be governed by the following mean-reverting Ornstein-Uhlenbeck process

$$d\mu_t = \rho(\bar{\mu} - \mu_t)dt + \sigma_\mu dB_t^\mu \tag{3.4}$$

[7]Zero net supply bonds are redundant in this economy. However, we can still compute a market-clearing (shadow) interest rate.

where $\bar{\mu}$ is the long-run average growth rate and $\rho > 0$ is a mean reversion parameter. For simplicity, we assume the Brownian motion driving the growth rate, B_t^{μ}, is uncorrelated with the Brownian motion driving dividends. So far, this looks a lot like the Bansal-Yaron (2004) long-run risks model. Remember, however, that preferences are time separable here, so that the long-run risk created by eqs. (3.3)-(3.4) is not priced in equilibrium.[8]

3.2. **Learning.** Since the growth rate is unobserved, agents must formulate an estimate, $\hat{\mu}_t$, and sequentially revise it as new information arrives. Given knowledge of the process (3.4), they can do this by employing the Kalman filter.

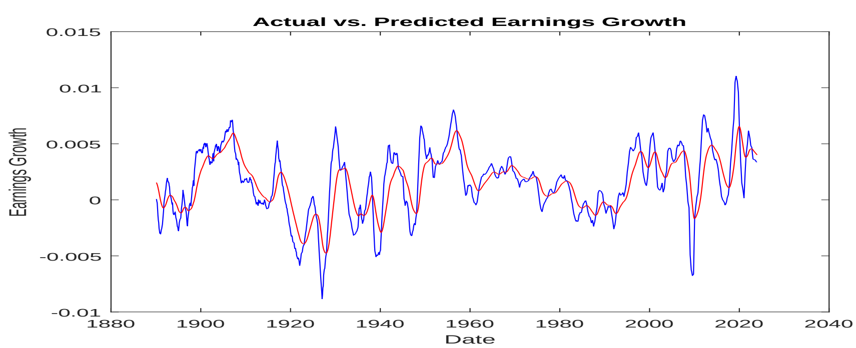
$$d\hat{\mu}_t = \rho(\bar{\mu} - \hat{\mu}_t)dt + \frac{Q_t}{\sigma} d\hat{B}_t \tag{3.5}$$

$$dQ_t = \left(\sigma_\mu^2 - 2\rho Q_t - \frac{Q_t^2}{\sigma^2}\right) dt \tag{3.6}$$

Eq. (3.5) is the sequence of conditional means and eq. (3.6) is the sequence of conditional variances, Q_t. Note that the conditional variance evolves deterministically. The innovation process, $d\hat{B}_t$ is defined as follows

$$d\hat{B}_t \equiv \frac{1}{\sigma}\left(\frac{dD_t}{D_t} - \hat{\mu}_t dt\right)$$

From the perspective of the agent, it is an i.i.d. process. It is simply the (standardized) error when forecasting the current dividend growth rate. Figure 2 plots $\hat{\mu}_t$ along with the sequence of actual earnings growth rates



The parameter values that were used are contained in Table 1. Note that the preference parameters, (γ, δ), are not needed to implement the Kalman filter. The remaining four parameters $(\sigma, \sigma_\mu, \rho, \bar{\mu})$ were calibrated from observed earnings growth.[9]

[8]Epstein, Farhi, and Strzalecki (2014) criticize the Bansal-Yaron model on introspection grounds. They argue that the implied premium on the early resolution of uncertainty is implausible.

[9]Following Shiller, we use earnings rather than dividends, since changes in tax laws have reduced the incentive for firms to pay dividends, which produces a trend in the Price/Dividend ratio.

TABLE 1
Benchmark Parameter Values (Annualized)

γ	δ	σ	σ_μ	ρ	$\bar{\mu}$
0.955	.0156	.0163	.00087	.0096	.0185

The mean reversion parameter, ρ, was obtained by fitting an ARMA(1,1) process to the first difference in log earnings, which is the discretized version of eqs. (3.3)-(3.4). The (annualized) coefficient on lagged earnings growth is 0.99 (highly significant), which then suggests a value of $1 - .99 \approx .01$ for ρ. This turns out to be an important parameter in determining the dynamics of the PE ratio.

Figure 2 shows that the predicted growth rate (red line) tracks the actual growth rate reasonably well on average, but not surprisingly, it is less volatile. The question now becomes – can this process of belief revision explain the observed fluctuations in the PE ratio seen in Figure 1?

3.3. **Equilibrium.** To evaluate the expected present value in (3.2) we shall employ the tools of dynamic programming. Define the right-hand side of (3.2) by the function $V(D, \hat{\mu})$. To simplify, we assume that the learning process has already converged to its steady state, so that Q_t is constant, and hence, not a state variable. It is given by the unique positive root of the following quadratic

$$\bar{Q}^2 + 2\rho\sigma^2\bar{Q} - \sigma_\mu^2\sigma^2 = 0$$

When expressed in terms of the agent's own information set, the state variable D is governed by the process

$$\frac{dD_t}{D_t} = \hat{\mu}_t dt + \sigma d\hat{B}_t \tag{3.7}$$

while the state variable $\hat{\mu}$ is governed by the Kalman filter in eq. (3.5). The Hamilton-Jacobi-Bellman (HJB) equation associated with this problem is then given by the following partial differential equation (PDE)

$$\delta V(D,\hat{\mu}) = D^{1-\gamma} + \hat{\mu} D V_D + \frac{1}{2}\sigma^2 D^2 V_{DD} + \rho(\bar{\mu} - \hat{\mu})V_{\hat{\mu}} + \frac{1}{2}\left(\frac{\bar{Q}}{\sigma}\right)^2 V_{\hat{\mu}\hat{\mu}} + \bar{Q} D V_{D\hat{\mu}} \tag{3.8}$$

where subscripts denote partial derivatives. To solve this PDE we can employ a standard guess-and-verify/separation of variables strategy. Conjecture that the solution takes the following form

$$V(D,\hat{\mu}) = F(\hat{\mu}) D^{1-\gamma} \tag{3.9}$$

One can readily verify that a common $D^{1-\gamma}$ term appears in each term of eq. (3.8), and so can be cancelled out. Note that given the form of the solution in (3.9), and the fact that the left-side of (3.2) contains the term $D^{-\gamma}$, the equilibrium Price/Dividend ratio is then simply given by the function $F(\hat{\mu})$. It is determined by the solution of the following linear, nonhomogeneous, 2nd-order ODE:

$$\delta F(x) = 1 + (1-\gamma)\left(x - \frac{1}{2}\sigma^2\gamma\right)F(x) + \rho(\bar{\mu} - x)F'(x) + \frac{1}{2}(\bar{Q}/\sigma)^2 F''(x) + (1-\gamma)\bar{Q}F'(x) \tag{3.10}$$

where for notational convenience in what follows, we substitute x in place of $\hat{\mu}$. To solve this equation we employ an old trick, called 'variation of parameters'. Suppose we can find a solution, $F^h(x)$, to the homogeneous part of (3.10). We then conjecture that the solution to the nonhomogeneous equation takes the form $F(x) = v(x)F^h(x)$ and derive a *first-order* ODE for $v'(x)$, exploiting the fact that $F^h(x)$ solves the homogeneous equation. This equation can then be solved by simple integration.

After a change of variables, the homogeneous part of (3.10) takes the form of Kummer's equation, one of the classical ODEs of applied mathematics. Its series solution is called the 'confluent hypergeometric function'. We summarize this result in the following:

Lemma 3.1. *The solution of the homogeneous part of the ODE (3.10), $F^h(x)$, is given by*

$$F^h(x) = e^{ax} M\left(\frac{2(\delta + .5\sigma^2(1-\gamma)(\gamma - c)\rho^2 + (g(1-\gamma))^2}{4\rho^3}; \frac{1}{2}; \frac{[(c-x)\rho^2 + (1-\gamma)g^2]^2}{(\rho g)^2}\right) \tag{3.11}$$

where $a \equiv (1-\gamma)/\rho$, $c \equiv \bar{\mu} + \bar{Q}(1-\gamma)/\rho$, $g = \bar{Q}/\sigma$, and where M is the confluent hypergeometric (or Kummer) function (see Abramowitz and Stegun (1972, chpt. 13)).

All of our results below could be obtained, and would continue to apply, if we used this confluent hypergeometric solution. However, to make the mathematics cleaner and more transparent, we impose the following parameter restriction that converts it to a nice exponential function.

Lemma 3.2. *If the parameters satisfy the restriction*

$$\delta = \rho + (1-\gamma)\left[\bar{\mu} - \frac{1}{2}\gamma\sigma^2 + \frac{(1-\gamma)}{2\rho}\left(2\bar{Q} + g^2/\rho\right)\right] \tag{3.12}$$

then the homogeneous solution takes the exponential form

$$F^h(x) = K \exp\left\{-ax + \frac{\rho}{g^2}(c-x)^2\right\} \tag{3.13}$$

where the constant $K = \exp\left\{[(1-\gamma)[2c + (1-\gamma)g^2]/\rho\right\}$

The proof follows from the fact $M(a; a; z) = e^z$ for any argument a. The restriction in eq. (3.12) then ensures that the first argument of the Kummer function in eq. (3.11) is equal to $1/2$, which is the fixed value of the second argument. Expanding the third argument and combining with the exponential function multiplying M in eq. (3.11) then delivers eq. (3.13). This restricition links together the rate of time preference, the long-run growth rate, the rate of mean reversion, and volatility. All else equal, if $\gamma < 1$ then a higher long-run growth rate must be offset by a higher rate of time preference or less mean reversion.

We can now state the main result of this section:

Proposition 3.3. *If the parameters satisfy the restriction in eq. (3.12) then the equilibrium Price/Dividend ratio with Bayesian learning is given by*

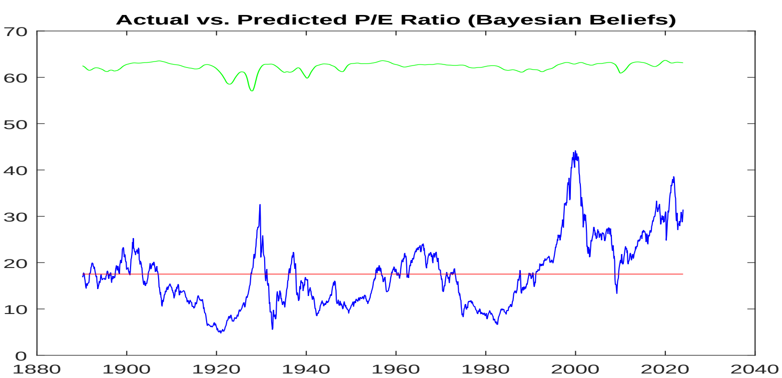
$$F(x) = \frac{1}{\delta + \frac{1}{2}\sigma^2\gamma(1-\gamma)}\left\{1 - (1-\gamma)^2 - (1-\gamma)e^{-ax+b(c-x)^2}\left[K_1 \operatorname{erf}\left(\frac{a+2b(c-x)}{2\sqrt{b}}\right)\right]\right\} \tag{3.14}$$

where $a \equiv (1-\gamma)/\rho$, $b \equiv \rho(\sigma/\bar{Q})^2$, $c \equiv \bar{\mu} + \bar{Q}(1-\gamma)/\rho$, *and* K_1 *is a positive constant that is a complicated function of all the model's parameters. The function* erf(u) *denotes the 'error function', which is defined as follows*

$$\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}}\int_0^u e^{-s^2}ds$$

The proof requires some algebra and calculus manipulations, which are relegated to Appendix A. As a reality check, note that $F(x) = 1/\delta$ when $\gamma = 1$, as it should. When $\gamma = 1$ and preferences are logarithmic, the discount rate effect exactly offsets the cash flow effect, and the price dividend ratio becomes constant. Also note that when $\gamma < 1$ uncertainty about growth, $\bar{Q}$, effectively increases the expected long-run growth rate, and hence the mean PE ratio. This is the Jensen's inequality effect highlighted by Pastor and Veronesi (2003).

Figure 3 plots eq. (3.14) (green line) using the parameters in Table 1 and the sequence of state estimates, $\hat{\mu}_t$, given by the Kalman filter in eq. (3.5) (red line in Figure 2).



Evidently, as many others have discovered before us, we can see that Bayesian learning in a Complete Markets/Representative Agent economy really doesn't get you very far. The mean PE ratio is far too high (corresponding to a low mean return), and it is nearly flat. Of course, if we amplify γ enough we can make the green line approach the red, and at the same time make it more volatile. Unfortunately, this would also produce countercyclical PE fluctuations, which is strongly counterfactual.

A ray of hope in the Bayesian learning asset pricing literature was offered by Weitzman (2007). Weitzman shows that in models featuring stochastic latent variables, like ours, posteriors about asset returns are *very* sensitive to priors about the parameters of the latent variable process. He provides examples where model predicted risky returns would be too high! In the next section we exploit this sensitivity to explain the volatlity and comovement puzzles. Our contribution relative to Weitzman (2007) is to provide a theory of why agents might want to perturb their prior in a way that better fits the data.

4. Asset Pricing When Agents Have A Human Brain

Although this section contains the paper's main result, most of the heavy lifting was done in the previous section. A key advantage of the Hansen-Sargent approach to ambiguity and model uncertainty is that it requires relatively minor modifications of 'business as usual'. We just need to add drift distortions to the dividend process (3.7) and the Kalman filter in (3.5). Both distortions are constrained by a relative entropy cost. As in Barillas, Hansen, and Sargent (2009), the distortion to the dividend process is negative, and reflects doubts about model specification. As discussed in Section 2, this pessimism is being driven by the alert system of right hemisphere. The distortion to the Kalman filter is motivated by Hansen and Sargent (2007). However, here there is an important difference. Hansen and Sargent (2007) introduce a pessimistic drift distortion to the Kalman filter, reflecting an agent's doubts about his own prior. As discussed in Section 2, here we do precisely the opposite. Our belief distortion is *optimistic*, reflecting the left hemisphere's role in formulating optimistically biased plans and expectations.

The interaction between the left and right hemispheres of the brain can be formalized via the following dynamic zero sum game version of the expected present value of dividends

$$V(D,\hat{\mu}) = \min_{h_1}\max_{h_2} \tilde{E}_t \int_t^\infty e^{-\delta(s-t)}\left\{D_s^{1-\gamma} + \frac{1}{2}(1-\gamma)\left(\varepsilon_1^{-1}h_{1s}^2 - \varepsilon_2^{-1}h_{2s}^2\right)\right\}ds \tag{4.15}$$

where the quadratic cost functions in the h_{it} sequences represent relative entropy penalties. The free parameters $(\varepsilon_1, \varepsilon_2)$ determine the level of pessimism and optimism. They are scaled by $(1-\gamma)^{-1}$ to put them in units of utility. Notice that if they are 0, then the entropy costs become infinite, and the optimal choices are $h_{1t} = h_{2t} = 0$. In this case, the model collapses to the Bayesian world of Section 3.

The expectations operator is taken with respect to the following drift distorted dividend process

$$dD_t = (\hat{\mu}_t + \sigma h_{1t})D_t dt + \sigma D_t d\tilde{B}_t \tag{4.16}$$

and the following drift distorted sequence of state estimates

$$d\hat{\mu}_t = [(\rho(\bar{\mu} - \hat{\mu}_t) + (\bar{Q}/\sigma)h_{2t}]dt + \frac{\bar{Q}}{\sigma}d\hat{B}_t \tag{4.17}$$

The idea here is that the cautious right hemisphere controls the model distortion sequence, h_{1t}, while the bold and imaginative left hemisphere controls the belief distortion, h_{2t}.

Performing the optimization and subbing the optimal h_i functions into HJB equation yields the following modified HJB equation

$$\begin{aligned}\delta V(D,\hat{\mu}) \quad = \quad & D^{1-\gamma} + \hat{\mu}DV_D + \frac{1}{2}\sigma^2 D^2 V_{DD} + \rho(\bar{\mu}-\hat{\mu})V_{\hat{\mu}} + \frac{1}{2}\left(\frac{\bar{Q}}{\sigma}\right)^2 V_{\hat{\mu}\hat{\mu}} + \bar{Q}DV_{D\hat{\mu}} \\ & -\frac{1}{2}\frac{\varepsilon_1}{1-\gamma}\sigma^2 D^2 (V_D)^2 + \frac{1}{2}\frac{\varepsilon_2}{1-\gamma}(\bar{Q}/\sigma)^2 (V_{\hat{\mu}})^2 \end{aligned} \tag{4.18}$$

The first few terms are identical to the Bayesian HJB equation in (3.8). Notice, however, that the final two terms make the equation *nonlinear*. Nonlinear PDEs are a no man's land. They can only be solved numerically, or approximately via perturbation expansions. Another problem is that this nonlinearity will in general make the solution nonhomothetic. Depending on whether $\gamma \lesseqgtr 1$, optimism and pessimism will either wax or wane as the economy grows.

To deal with these problems we employ a trick first proposed by Maenhout (2004). This involves scaling the cost parameters $(\varepsilon_1, \varepsilon_2)$. Notice that if ε_1 is scaled by V, so that $\varepsilon_1 = \bar{\varepsilon}_1/V$, where $\bar{\varepsilon}_1$ is constant, then the term involving $(V_D)^2$ becomes homothetic, and a conjectured solution of the form $V \sim D^{1-\gamma}$ will be consistent. Behaviorally, homotheticity is being induced by making the cost of pessimism increase as the economy grows (assuming $\gamma < 1$). The filtering cost parameter, ε_2, requires a slightly different scaling. Suppose as before we conjecture a separable solution of the form $V = D^{1-\gamma}F(\hat{\mu})$. Then notice that if we employ the scaling $\varepsilon_2 = (1-\gamma)\bar{\varepsilon}_2/(D^{1-\gamma}F')$, then equation becomes linear and homothetic.

Given this scaling we can once again posit a solution of the form $V = D^{1-\gamma}F$, and after cancelling out $D^{1-\gamma}$, obtain the following *linear* ODE for the PE ratio

$$\begin{aligned}\delta F(x) \quad &= \quad 1 + (1-\gamma)\left(x - \frac{1}{2}\sigma^2\gamma\right)F(x) + \rho(\bar{\mu} - x)F'(x) + \frac{1}{2}(\bar{Q}/\sigma)^2 F''(x) + (1-\gamma)\bar{Q}F'(x) \\ &\quad -\frac{1}{2}\bar{\varepsilon}_1\sigma^2(1-\gamma)F(x) + \frac{1}{2}\bar{\varepsilon}_2(\bar{Q}/\sigma)^2 F'(x) \qquad (4.19)\end{aligned}$$

Mathematically, this is the exact same equation as before. All that has changed are the coefficient functions. We can solve it using variation of parameters just like before. Of course, the parameter restriction delivering an exponential solution is now slightly different. The parameters must now satisfy the weaker restriction

$$\delta + \frac{1}{2}\bar{\varepsilon}_1\sigma^2(1-\gamma) = \rho + (1-\gamma)\left[\bar{\mu} + \frac{1}{2}\bar{\varepsilon}_2(\bar{Q}/\sigma)^2/\rho - \frac{1}{2}\gamma\sigma^2 + \frac{(1-\gamma)}{2\rho}\left(2\bar{Q} + g^2/\rho\right)\right] \quad (4.20)$$

Notice that when $\gamma < 1$, greater optimism ($\bar{\varepsilon}_2$) must be accompanied by greater pessimism ($\bar{\varepsilon}_1$). This correlation can be avoided, however, if we are willing to adjust one of the other parameters. As in Barillas, Hansen, and Sargent (2009), model uncertainty ($\bar{\varepsilon}_1$) can be interpreted as a form of risk aversion. The hopeful optimism of the left hemisphere ($\bar{\varepsilon}_2$) effectively increases the long-run growth rate of dividends, $\bar{\mu}$.

We can now state our paper's main result

Proposition 4.1. *If the parameters satisfy the restriction in eq. (4.20) then the equilibrium Price/Dividend ratio when agents hope for the best, but plan for the worst is given by*

$$F(x) = \frac{1}{\delta + \frac{1}{2}\sigma^2(1-\gamma)(\gamma + \bar{\varepsilon}_1)}\left\{1 - (1-\gamma)^2 - (1-\gamma)e^{-ax+b(c-x)^2}\left[K_1 \operatorname{erf}\left(\frac{a + 2b(c-x)}{2\sqrt{b}}\right)\right]\right\} \quad (4.21)$$

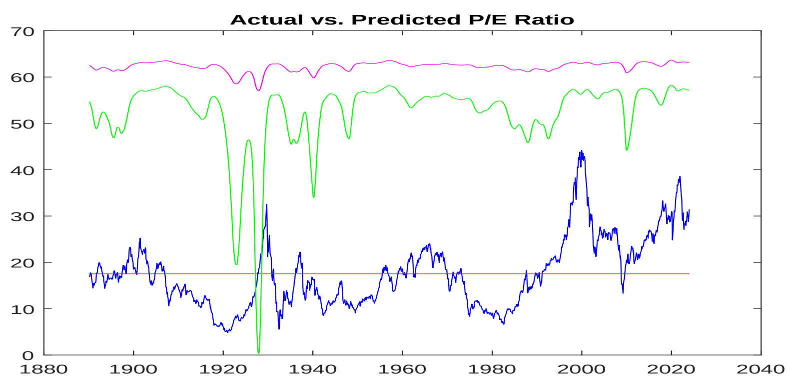
where $a \equiv (1-\gamma)/\rho$, $b \equiv \rho(\sigma/\bar{Q})^2$, $c \equiv \bar{\mu} + .5\bar{\varepsilon}_2(\bar{Q}/\sigma)^2/\rho + \bar{Q}(1-\gamma)/\rho$, *and* K_1 *is a positive constant that is a complicated function of all the model's parameters. The function* $\operatorname{erf}(u)$

denotes the 'error function', which is defined as follows

$$\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-s^2} ds$$

Comparing this to the Bayesian PE ratio in eq. (3.14) highlights the distinct roles being played by the right and left hemispheres. The pessimism of the right hemisphere effectively pulls down the mean PE ratio, thus increasing mean returns. The fact that it delivers just an unconditional mean adjustment derives from our homothetic scaling, which eliminates state dependence. This is not the case, however, with the optimism of the left hemisphere. By increasing the parameter c in the exponential and error functions, it makes beliefs more responsive to news. The parameter $a = (1-\gamma)/\rho$ is also important, since it interacts with c. It will be large when shocks are persistent (ρ is small), and when agents respond more vigorously to shocks (as γ shrinks toward zero).

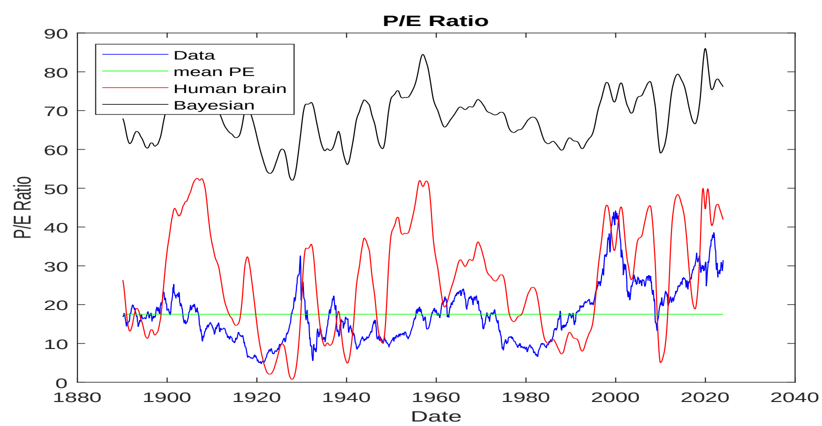
Figure 4 plots the PE ratio in an economy where agents hope for the best, but plan for the worst. The parameters are the same as before. The new parameters, $(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$ were set to $\bar{\varepsilon}_1 = 206$ and $\bar{\varepsilon}_2 = 255$. By themselves, these values are meaningless, since they are not unit free. What matters is the magnitude of the implied distortions. The pessimism parameter $\bar{\varepsilon}_1$ implies a 12 basis point reduction in the unconditional growth rate of dividends. Conversely, the optimism parameter $\bar{\varepsilon}_2$ implies a 52 basis point positive distortion to $\bar{\mu}$.



Evidently, the model improves along two dimensions. First, the mean PE is now closer to the data. However, a 12 basis point pessimistic distortion to the dividend process results in only a modest decrease in the mean PE ratio. Predicted mean returns only increase by about 50-60 basis points. Of course, we could increase this effect by increasing $\bar{\varepsilon}_1$, but as we shall see below, this would produce an excessively low detection error probability given the way our analytic solution links together $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$. Second, notice that the volatility is substantially increased. The challenge now is to maintain this high volatility while at the same time decreasing the mean PE ratio.

4.1. **Numerical Solution.** The previous analytic solution is convenient for developing intuition for how the parameters influence the moments of the predicted PE ratio. However, it does suffer from a couple of disadvantages. First, scaling the cost parameters to deliver a linear PDE imples that the distortions are *constant*. This can still increase volatility (as seen above), but it seems clear that allowing the distortions to be state-dependent will make it easier for the model to match the data. Moreover, state dependent optimism and pessimism are consistent with the survey evidence presented in La Porta et. al. (2024). (See below for further discussion). Second, the above analytic solution links together $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$. If $\gamma < 1$ then greater optimism must be accompanied by greater pessimism. Effectively, there is only *one* free parameter. Clearly, allowing the left- and right-hemispheres to work independently will enhance the model's ability to fit the data.

Figure 5 presents the Bayesian and Human Brain solutions for the PE ratio from unscaled/nonlinear PDE in equation (4.18). [Blah, Blah....Xiaowen, can you add a few sentences describing the algorithm, e.g., the pert. approx., differencing, boundary conditions].[10] Crucially, now we adjust $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$ independently so as to match both the mean and volatility of the observed PE ratio.



Evidently, the fit is much improved. The mean PE ratio is now 26.2, only slightly above the actual value of 17.5. The standard deviation is 13.9, which is actually higher than the observed value of 7.72. However, if one corrects for differences in the mean by computing the coefficient of variation, the the volatilities match up pretty well. The predicted value is .53, while the actual is .42. Perhaps most impressive is the model's ability to match the rise and fall of the PE ratio during the Roaring 20s/Great Depression and Dotcom Boom periods. Table 2 summarizes the results

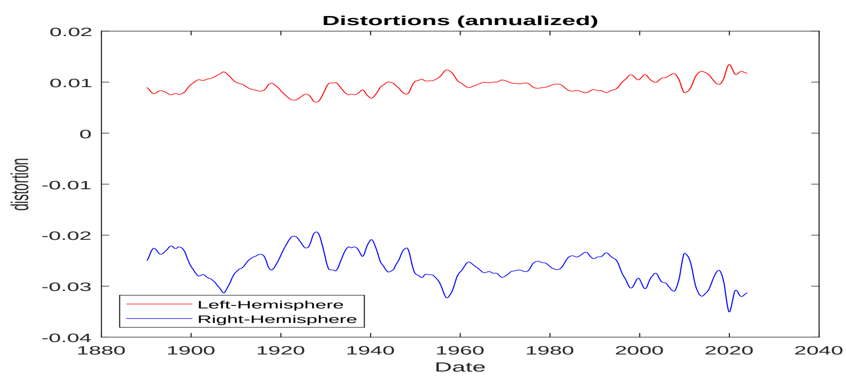
[10]The Matlab code is available upon request and posted on the author's webpage.

TABLE 2
Actual vs. Predicted Moments

	Mean	st. dev.	cvar	corr$(PE, \hat{PE})$
Data	17.5	7.42	.424	–
Scaled	52.3	7.72	.148	.263
Numerical	26.2	13.9	.530	.379
Bayes	68.8	6.94	.101	.402

Notes: (1) Bayes results are for the numerical solution.

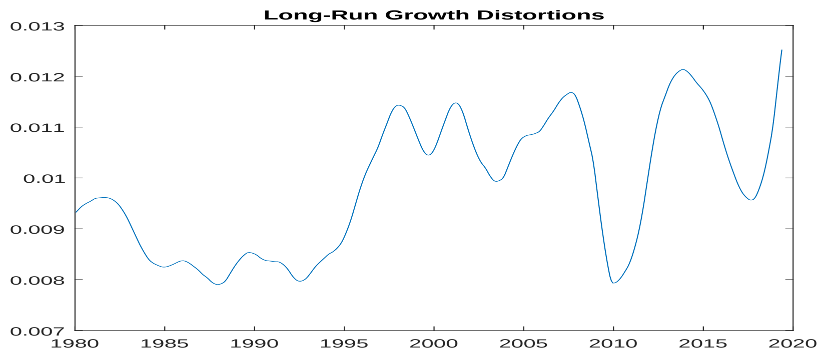
Both the dividend model and filtering distortions now fluctuate over time. Figure 6 presents the results. Both distortions are annualized, and the optimistic filtering distortion has been scaled to be units of long-run mean growth (by dividing by the mean reversion parameter ρ).



Interestingly, notice that the two distortions are negatively correlated. When optimism about growth increases, the agent becomes more pessimistic about his model.

4.2. **Comparison to Bordalo et. al. (2024).** Below we report evidence that these distortions are empirically plausible. Following Hansen and Sargent (2008), we do this by computing detection error probabilites. However, as emphasized by Bordalo et. al. (2024), one could instead adopt a less model dependent stategy by simply *asking* people what they expect growth to be over various horizons. In this section we show that our results are quite consistent with this survey evidence.

Figure 7 plots the optimistic distortion to $\bar{\mu}$ from Figure 6, corresponding to the shorter sample reported in Bordalo et. al. (2025), i.e., 1980-2020. It should be compared to Figure 1 in their paper. Notice that our theory-implied long-run growth forecast distortion exhibits a very similar pattern. Long-run growth forecasts decline by about 20 basis points during the 1980s, then rise rapidly, by about 30-40 basis points during the Dotcom boom period of the 1990s. These optimistic long-run growth forecasts are then reversed during the Financial Crisis of 2009-10.



5. Detection Error Probabilities

The key mechanism in our model is belief distortions. Investors are defensively pessimistic when constructing models of dividend growth, but conditional on their models, they 'hope for the best' when formulating and revising forecasts about future dividend growth. We argued above that evidence from neuroscience supports this tension between optimism and pessimism. Unfortunately, the neuroscience literature does not provide much guidance on the *magnitude* of these distortions. We have introduced two new free parameters (the relative entropy costs on pessimism and optimism), so it seems rather obvious we could tune these parameters to explain both high and volatile returns. How do we know that the resulting parameter values are empirically plausible? Is there any external evidence we can consider that would discipline our calibration?

Here we follow Hansen and Sargent (2008) and quantify plausibility using 'detection error probabilities'. Agents are only allowed to consider models that could have plausibly generated the observed data. Specifically, we view agents as statisticians, who attempt to discriminate among models using likelihood ratio statistics. When likelihood ratio statistics are large, detection error probabilities are small, and models are easy to distinguish. Detection error probabilities will be small when models are very different, or when there is a lot of data. Classical inference is based on adopting a null hypothesis, and fixing the probability of falsely rejecting this hypothesis. Detection error probabilities treat the null and alternative symmetrically, and average between Type I and Type II errors. In particular,

$$DEP = \frac{1}{2}\text{Prob}(H_A|H_0) + \frac{1}{2}\text{Prob}(H_0|H_A)$$

Hence, a DEP is analogous to a p-value. Our results would therefore be implausible if the DEP is too small. Small DEP's imply agents are hedging against models that could easily be rejected using observed data.

In general, entropy constrained distortions are state dependent, and so the error probabilities must be computed using monte carlo simulation (Hansen and Sargent (2008, Chpt. 9). However, we have purposely rigged things (by appropriately scaling the entropy costs by the value function) so that the resulting distortions are *constant*. Hence, the null and alternative models for log dividend growth are both linear Gaussian ARMA(1,1) models, which only differ in their intercepts. This allows us to compute detection errror probabilities using a simple z-score.[11]

By quasi-differencing the Kalman filter for $\hat{\mu}_t$, (4.17), and then substituting into the discretized version of the dividend growh model in (4.16) we obtain

$$[1-(1-\rho)L]\Delta\log(D_{t+1}) = \rho\bar{\mu} + (\Delta_\mu - \rho\Delta_d) + \sigma\left[\frac{\bar{Q}}{\sigma^2} + 1 - (1-\rho)L)\right]d\hat{B}_{t+1} \quad (5.22)$$

where $\Delta_\mu > 0$ is the (optimistic) distortion to the Kalman filter and $\Delta_d > 0$ is the (pessimistic) distortion to the dividend growth model, both expressed as (annualized) percentage points. As noted in the previous section, our benchmark parameters imply $\Delta_\mu = .0064$ and $\Delta_d = .0113$. There are a couple of important points to notice here. First, notice the two distortions *offset* each other. This makes it harder to detect them, based solely on observed data.[12] Second, notice that the dividend model distortion is scaled by ρ, which is close to zero, given the observed weak mean reversion in dividend growth. This means that detection error probabilities are driven primarily by the optimistic filtering distortion.

Since the detection errors only depend on the intercept, the Type I and Type II error probabilities are the same, and we obtain

$$DEP = \Phi\left(\frac{-\sqrt{T}|\Delta_\mu - \rho\Delta_d|}{\hat{\sigma}}\right)$$

where Φ denotes the standard normal CDF and $\hat{\sigma}$ is the steady state standard deviation of the error term in (5.22). It is given by

$$\hat{\sigma} = \sigma\sqrt{(1+\bar{Q}/\sigma)^2 + (1-\rho)^2}$$

[11]We take a couple short cuts. First, we assume Q_t has converged to its steady state, and so is constant. Second, we ignore the MA(1) structure of the model's error term. Since this tends to bias toward rejection, it should bias our detection error probabilities downwards, which reinforces our results.

[12]As noted earlier, Kasa (2012) considered the case where the two distortions exactly offset each other.

If we set $T = 130$, which is roughly equal to the (annualized) length of our data set, we find $DEP = .101$. Thus, even after 130 years our left-brain/right-brain investor would not be able to resolve their optimism/pessimism tension with much statistical confidence.[13]

6. Conclusion

This paper has shown that when agents are pessimistic about their models but optimistic about their beliefs, stock returns will be high on average and Price/Earnings ratios will exhibit large procyclical fluctuations. This is true even though markets are complete and agents have modest (nearly logarithmic) risk aversion. The key innovation in our paper is to link this apparently schizophrenic combination of optimism and pessimism to recent work in neuroscience, which shows that beliefs are processed primarily by the optimistically-biased left-hemisphere of the brain, while decisions are primarily processed by the pessimistically biased right-hemisphere.

An important task for future work is to examine the robustness of these results to alternative model specifications and parameter values. Preliminary analysis suggests that model-implied price volatility is quite sensitive to mean reversion in earnings (as parameterized by ρ). In the Introduction we noted that our paper is similar to Bordalo, Gennaioli, Porta, and Shleifer (2024). One advantage of their work is to show that belief overreaction also explains cross-sectional asset pricing puzzles. Since earnings persistence differs across firms, it would be interesting to see whether our model can also shed light on firm-level stock return data.

[13] In reality, agents are finite-lived and seem to place more weight on their own experiences (Malmendier and Nagel (2011)). In this case, detection error probabilities would of course be even larger, and would persist (within the economy) indefinitely.

Appendix A. Proof of Proposition 3.3

Since the ODE we want to solve is linear, the general solution is the sum of a particular solution and two linearly independent solutions of the homogeneous equation. Moreover, given Lemma 3.2, we have at hand a solution of the homogeneous equation. Hence, we can apply a standard 'reduction-of-order' and 'variation of parameters' method to constuct the general solution. Details can be found in Boyce and DiPrima (1977, p. 126).

Lemma A.1. *Consider the following non-homogeneous linear 2nd-order ODE*

$$y''(x) + p(x)y'(x) + q(x)y(x) = g(x) \tag{A.23}$$

and suppose $y_1(x)$ is a known solution to the homogeneous equation. Then the general solution is given by

$$y(x) = c_1 y_1(x) + c_2 y_1(x)\int^x \frac{ds}{y_1^2(s)h(s)} + y_1(x)\int^x \frac{1}{y_1^2(s)h(s)}\left[\int^s y_1(u)h(u)g(u)du\right]ds \tag{A.24}$$

where (c_1, c_2) are arbitrary constants of integration, and where $h(s) = \exp(\int^s p(u)du)$.

Unlike Froot and Obstfeld (1991), we are not interested in intrinsic bubbles, which correspond to solutions of the homogeneous equation. So we set $c_1 = c_2 = 0$, and confine attention to the third term in (A.24). By inspection, and from basic economic reasoning, it is clear that when $\gamma = 1$ the solution to (3.10) is just a constant. Hence, let's conjecture that $F(x) = A + (1-\gamma)^2 G(x)$. Subbing this guess into (3.10) we find

$$A = \frac{1}{\delta + \frac{1}{2}\sigma^2\gamma(1-\gamma)}$$

and that $G(x)$ solves the ODE

$$(1-\gamma)\delta G(x) = (1-\gamma)Ax + (1-\gamma)\left[(1-\gamma)x - \frac{1}{2}\sigma^2\gamma\right]G(x) + (1-\gamma)[\rho(\bar{\mu}-x) + (1-\gamma)\bar{Q}]G'(x) + \frac{1}{2}(1-\gamma)(\bar{Q}/\sigma)^2 G''(x)$$

Note that this can be put into the standard form in (A.23), with $g(x) = -2(1-\gamma)(\sigma/\bar{Q})^2 Ax$ and $p(x) = 2b(c-x)$, where b and c are the constants defined in Proposition 3.3. Hence, $h(x) = \exp\left[-b(c-x)^2\right]$. Given the restriction in (3.12), the solution of the homogeneous equation is given by (3.13).[14] Note that the constant of integration in (3.13) cancels out of the third term of (A.24), so we ignore it in what follows. We can now evaluate the inner integral in (A.24) as follows

$$\int^s y_1(u)h(u)g(u)du = -2(1-\gamma)A(\sigma/\bar{Q})^2\int^s ue^{-au}du = 2(1-\gamma)A(\sigma/\bar{Q})^2\left[\frac{1}{a^2}e^{-as}(1+as)\right]$$

where from Proposition 3.3 $a \equiv (1-\gamma)/\rho$. Next, noting that

$$\frac{1}{y_1^2(s)h(s)} = e^{2as - b(c-x)^2}$$

we get the following integral to evaluate

$$\int^x \frac{1}{y_1^2(s)h(s)}\left[\int^s y_1(u)h(u)g(u)du\right]ds = \frac{2A\sigma^2\rho^2}{\bar{Q}^2(1-\gamma)}\int^x (1+as)e^{as-b(c-s)^2}ds$$

This is not an easy integral to evaluate by hand. However, it is easily handled by Wolfram-Alpha or Matlab's symbolic math toolbox. This gives us

$$\int^x (1+as)e^{as-b(c-s)^2}ds = -\left[\frac{ae^{ax-b(c-x)^2}}{2b} + \frac{\sqrt{\pi}(a^2+2b(1+ac))e^{a^2/4b+ac}\,\mathrm{erf}\left(\frac{a+2b(c-x)}{2\sqrt{b}}\right)}{4b^{3/2}}\right] \tag{A.25}$$

To obtain $G(x)$, we just need to multiply the right-side of (A.25) by $y_1(x)$, and reintroduce the constant. Note that the first term on the right-side of (A.25) is just $(a/2b)y_1^{-1}(x)$, so the first term of $G(x)$ just evaluates to -1. Finally, remembering that $F(x) = A + (1-\gamma)^2 G(x)$, we find that constant term in $F(x)$

[14]Note that the common factor $(1-\gamma)$ can be cancelled from both sides without changing the solution of the homogeneous equation. However, to retain the appropriate limit as $\gamma \to 1$, we do not want to cancel $(1-\gamma)$ from the the non-homogeneous term. Technically, this is a singular perturbation problem.

simplifies to $A[1-(1-\gamma)^2] = A$ when $\gamma = 1$ (as it should). The second component of the solution given by (3.14) then follows easily, noting that

$$K_1 \equiv \frac{\rho\sqrt{\pi}[a^2 + 2b(1+ac)]e^{a^2/4b+ac}}{2\sqrt{b}}$$

Appendix B. Proof of Proposition 4.1

The proof here is nearly identical to the above proof of Proposition 3.3. Given our assumptions about the scaling of the relative entropy cost parameters, the only difference arises in the coefficients of the ODE we need to solve. Comparing (3.10) and (4.19), we observe that the term $\frac{1}{2}\bar{\varepsilon}_2(\bar{Q}/\sigma)^2$ gets added the coefficient on $F'(x)$, while $\frac{1}{2}\bar{\varepsilon}_1\sigma^2\gamma(1-\gamma)^2$ is added to the coefficient of $F(x)$. Adding a constant to the coefficient of $F(x)$ just changes A to

$$A = \frac{1}{\delta + \frac{1}{2}\sigma^2(1-\gamma)(\gamma + (1-\gamma)\bar{\varepsilon}_1)}$$

which is the same as given in Proposition 4.1. Adding a constant to the coefficient of $F'(x)$ just changes the (perceived) long-run mean of x to

$$c = \bar{\mu} + .5\bar{\varepsilon}_2(\bar{Q}/\sigma)^2/\rho + \bar{Q}(1-\gamma)/\rho$$

which is the same value given in Proposition 4.1.

References

Abramowitz, M., and I. A. Stegun (1972): *Handbook of Mathematical Functions*. Dover Publications.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Barillas, F., L. P. Hansen, and T. J. Sargent (2009): "Doubts or Variability?," *Journal of Economic Theory*, 144, 2388–418.

Barro, R. J. (2006): "Rare Disasters and Asset Markets in the 20th Century," *Quarterly Journal of Economics*, 121, 823–66.

Beeler, J., and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–82.

Benartzi, S., and R. H. Thaler (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110, 73–92.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2024): "Belief Overreaction and Stock Market Puzzles," *Journal of Political Economy*, 132, 1450–84.

Boyce, W. E., and R. C. DiPrima (1977): *Elementary Differential Equations and Boundary Value Problems*. John Wiley & Sons, third edn.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–51.

Constantinides, G. M., and D. Duffie (1996): "Asset Pricing With Heterogeneous Consumers," *Journal of Political Economy*, 104, 219–40.

Constantinides, G. M., and A. Ghosh (2017): "Asset Pricing With Countercyclical Household Consumption Risk," *Journal of Finance*, 72, 415–59.

Epstein, L. G., E. Farhi, and T. Strzalecki (2014): "How Much Would You Pay to Resolve Long-Run Risk?," *American Economic Review*, 104, 2680–97.

Fecteau, S., D. Knoch, N. Sultani, P. Boggio, and A. Pascual-Leone (2007): "Diminishing Risk-Taking Behavior by Modulating Activity in the Prefrontal Cortex: A Direct Current Stimulation Study," *Journal of Neuroscience*, 14, 12500–05.

Fehr, E., and A. Rangel (2011): "Neuroeconomic Foundations of Economic Choice – Recent Advances," *Journal of Economic Perspectives*, 25, 3–30.

Froot, K. A., and M. Obstfeld (1991): "Intrinsic Bubbles: The Case of Stock Prices," *American Economic Review*, 81, 1189–1214.

Gazzaniga, M. S. (2008): *Human: The Science Behind What Makes Your Brain Unique*. HarperCollins.

Gul, F., and W. Pesendorfer (2008): " The Case for Mindless Economics," in *The Foundations of Positive and Normative Economics: A Hand Book*, ed. by A. Caplin, and A. Schotter. Oxford University Press.

Hansen, L. P., and T. J. Sargent (2007): "Recursive Robust Estimation and Control without Commitment," *Journal of Economic Theory*, 136, 1–27.

——— (2008): *Robustness*. Princeton University Press.

——— (2010): "Fragile Beliefs and the Price of Uncertainty," *Quantitative Economics*, 1, 129–62.

Hansen, L. P., and K. J. Singleton (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, 249–65.

Haselton, M. G., and D. Nettle (2006): "The Paranoid Optimist: An Integrative Evolutionary Model of Cognitive Biases," *Personality and Social Psychology Reveiw*, 10, 47–66.

Hecht, D. (2013): "The Neural Basis of Optimism and Pessimism," *Experimental Neurobiology*, 22, 173–99.

Ju, N., and J. Miao (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80, 559–91.

Kahneman, D. (2011): *Thinking, Fast and Slow*. Doubleday.

Kahneman, D., and D. Lovallo (1993): "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Takiing," *Management Science*, 39, 17–31.

Kasa, K. (2012): "A Behavioral Defense of Rational Expectations," Working Paper 12-05, Dept. of Economics, Simon Fraser University.

Maenhout, P. J. (2004): "Robust Portfolio Rules and Asset Pricing," *Review of Financial Studies*, 17, 951–83.

Malmendier, U., and S. Nagel (2011): "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *Quarterly Journal of Economics*, 126, 373–416.

Mehra, R., and E. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–61.

Pastor, L., and P. Veronesi (2003): "Stock Valuation and Learning About Profitability," *Journal of Finance*, 58, 1749–89.

Savage, L. J. (1972): *The Foundations of Statistics*. Dover Publications, second revised edn.

Sharot, T., R. Kanei, D. Marston, C. Korn, G. Rees, and R. Dolan (2012): "Selectively Altering Belief Formation in the Human Brain," *Proceedings of the National Academy of Sciences*, 109, 17058–62.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71, 421–36.

Weitzman, M. L. (2007): "Subjective Expectations and Asset-Return Puzzles," *American Economic Review*, 97, 1102–30.

Wolford, G., M. B. Miller, and M. Gazzaniga (2000): "The Left Hemishere's Role in Hypothesis Formation," *Journal of Neuroscience*, 20, 1–4.

Department of Economics, Simon Fraser University, 8888 University Dr., Burnaby BC V5A 1S6

Email address: kkasa@sfu.ca

URL: https://www.sfu.ca/~kkasa/

Department of Economics, University of Guelph, 50 Stone Rd E, Guelph, ON Canada. N1G 2W1

Email address: leix@uoguelph.ca

URL: http://www.uoguelph.ca/lang/people/xiaowen-lei