Hope For The Best, Plan For The Worst

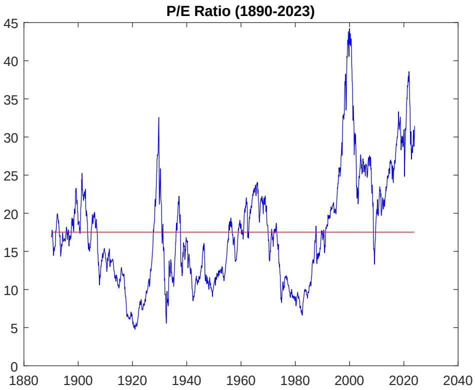
Kenneth Kasa[1]   Xiaowen Lei[2]

[1]Department of Economics
Simon Fraser University

[2]Department of Economics
University of Guelph

March 11, 2026

Some Challenging Data



An Idea That Doesn't Work

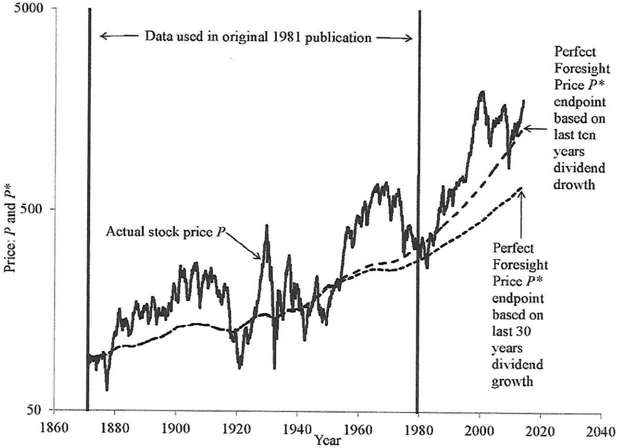


Figure 1. Real Standard & Poor's Composite Stock Price Index along with Present values with constant discount rate of subsequent real dividends accruing to the index 1871–1913. The two present values differ in their assumption about dividend growth after 2013.

Our Paper

An Experiment

- Subjects shown random sequences of 2 outcomes, R or G.
- $Pr(R) = .75 > Pr(G) = .25$
- Reward based on number of correct guesses.
- Humans 'frequency match'. (Guess R about 75% of time). Other animals optimize. (Always guess R).
- Miller et. al. (2000) played this game separately with the Left and Right Hemispheres of the brain.
- The LH frequency matches. The RH optimizes.

A Quote

This is consistent with the hypothesis that the left-hemisphere interpreter constructs theories to assimilate perceived information into a comprehensive 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.

– Michael Gazzaniga (2008). In *Human: The Science Behind What Makes Your Brain Unique*

Main Idea

- Investment involves *two* cognitive processes.
- 1.) Formulate beliefs about the future consequences of investment.
- This involves developing hypotheses about how the future may unfold.
- This is a creative process, which is akin to engaging in frequency matching.
- We assume this problem is processed in the LH, and is optimistically biased.
- 2.) Decide how much scarce capital to allocate to the LH's theories.
- We assume this decision is processed by the alert system of the RH, and is pessimistically biased.

A Bayesian Benchmark

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

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

$$\frac{dDt}{D_t} = \mu_t dt + \sigma dB_t$$

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

Parameter Values

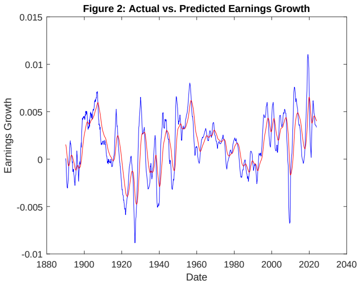
TABLE 1
Benchmark Parameter Values (Annualized)

γ	δ	σ	σ_μ	ρ	$\bar{\mu}$
0.955	.0116	.0163	.00087	.0096	.0185

Kalman Filter

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

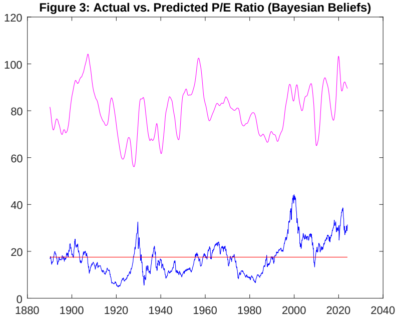
$$dQ_t = \left(\sigma_\mu^2 - 2\rho Q_t - \frac{Q_t^2}{\sigma^2}\right) dt$$



EQUILIBRIUM P/E RATIO

$$F^b(x) = e^{\lambda x}\left\{\sum_{n=0}^{\infty} a_n (A_0 - g\rho x)^n\right\}$$

where $\lambda = (1-\gamma)/\rho$, $g = 2(\sigma/\bar{Q})^2$, $A_0 = [\rho\bar{\mu} + \bar{Q}(1-\gamma)]$



History

- Knight (1921): There is a difference between risk and uncertainty.
- Keynes (1936): Financial markets are volatile because they price uncertainty, not risk.
- Savage (1954): Under certain axioms, the distinction between risk and uncertainty is behaviorally irrelevant.
- Ellsberg (1961): Experiments cast doubt on the Savage axioms.
- "It takes a model to beat a model". SEU dominates in economics until the 1990s (Gilboa & Schmeidler (1989)). Coincidentally, engineers develop Robust Control methods at around the same time.
- Hansen & Sargent brought these two literatures together.

THE ELLSBERG PARADOX

- Consider 2 urns, each known to contain 100 Red & Black balls.
  1. Urn 1: Known to have 50 Red/50 Black balls.
  2. Urn 2: No info given about Red vs. Black balls.

- Now suppose you can bet on the color of the ball drawn from one of the urns. You get $100 if you guess correctly, $0 otherwise.

- The SEU axioms predict that people should (weakly) prefer Urn 2. According to SEU, people have unique subjective prior probabilities about the proportion of R vs. B in Urn 2. This prior must assign at least a probability of 0.5 to either R or B, so Urn 2 should be (weakly) preferred.

- In practice, the vast majority of subjects prefer Urn 1. The rationale? It is better to wager with 'known probabilities'.

- Gilboa & Schmeidler (1989) relax the SEU axioms to incorporate a preference for known probabilities. According to SEU, if an individual is indifferent between two lotteries, he should be indifferent to a mixture between them (with known probabilities). The axioms of GS allow agents to (weakly) prefer the mixture.

Robust Control

- In Robust Control, agents do not adhere to the Savage axioms. They adhere to alternative axioms developed by Klibanoff, Marinacci, & Mukerji (2005) and Strzalecki (2011).
    1. They cannot form a unique prior.
    2. They do not reduce compound lotteries.
    3. They distinguish between model uncertainty and parameter uncertainty.
    4. They care about early resolution of uncertainty.
- All these behavioral modifications are controlled by a *single* free parameter, θ. This parameter is calibrated using detection error probabilities. As $\theta \longrightarrow \infty$, Rational Expectations/Bayesian results are recovered.
- Robust Control features two models, not one. One is the agent's benchmark model. It is exogenously specified. The other is his worst-case model.
- Since the worst-case model depends on his own policy, the agent views himself as being immersed in a dynamic zero-sum game. A hypothetical 'evil agent' picks models to minimize his payoff.
- The agent is prudent, not paranoid. To avoid undue pessimism, the evil agent must pay a relative entropy cost.
- It is typically assumed the agent's doubts are all in his head. His benchmark model is correct. He just doesn't know it.

Asset Pricing When Agents Have Human Brains

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

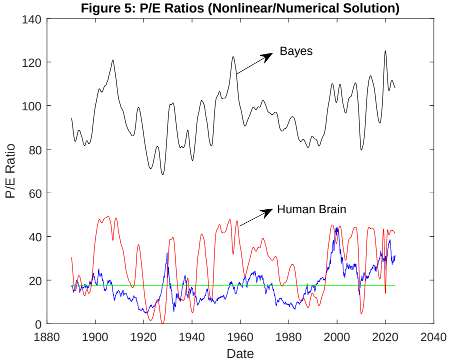
$$dD_t = (\hat{\mu}_t + \sigma h_{1t}) D_t dt + \sigma D_t d\tilde{B}_t$$

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

Some Progress

$$F(x) = F^b(x) + \varepsilon_1 F^p(x) + \varepsilon_2 F^o(x)$$

where $\varepsilon_1 = .084$ and $\varepsilon_2 = .302$

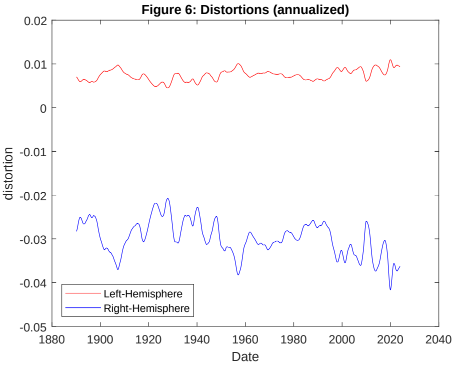


Actual vs. Predicted

TABLE 2
Actual vs. Predicted Moments

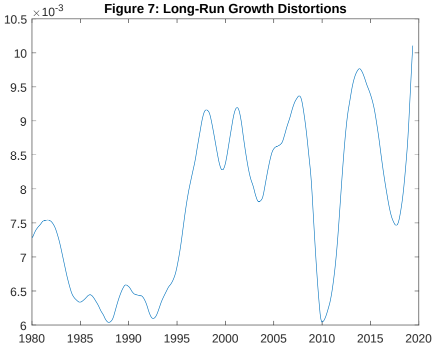
	Mean	st. dev.	cvar	corr($PE, \hat{PE}$)
Data	17.5	7.42	.424	–
Bayes	95.1	11.6	.122	.402
Scaled	43.7	5.47	.125	.279
Numerical	27.5	13.4	.489	.345

Optimism & Pessimism



LINK TO BORDALO ET. AL. (JPE 2024)

Our paper provides a microfoundation for the overreaction documented by Bordalo et. al. (JPE 2024). (See their Fig 1).



Detection Error Probabilities

- Neuroscience provides little guidance on the *magnitude* of optimism and pessimism.
- We assume agents only consider empirically plausible models.
- Det. Error Probs.

$$\frac{1}{2}P(H_A|H_0) + \frac{1}{2}P(H_0|H_A)$$

- Like a p-value, except null and alternative are symmetric.
- If DEP is small, then distorted model could be easily rejected based on historical data.

- With scaled entropy costs, optimism and pessimism only distort the intercepts.
- In terms of observable data the (discretized) distorted model becomes the ARMA(1,1) process

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

  where $\Delta_\mu = .0055$ and $\Delta_d = .0071$.
- The DEP then becomes

$$\begin{aligned} DEP &= \Phi\left(\frac{-\sqrt{T}|\Delta_\mu - \rho\Delta_d|}{\hat{\sigma}}\right) \\ &= .097 \qquad (\text{when } T = 130) \end{aligned}$$

- DEP would be even larger if agents are finite-lived and learn from their own experiences.

- Without scaled entropy costs, distortions become state dependent, and explicit expressions for the DEP are unavailable.
- However, we can develop the following asymptotic (large deviations) approximation

$$DEP \approx \frac{1}{2}\exp\left\{-TE(h_2^2)\right\} = .197$$

- For simplicity, we ignore $h_1(D, \hat{\mu})$. (It is scaled by $\rho \approx 0$). Since h_1 offsets h_2, omitting it reinforces our results.

Conclusion

- When agents are pessimistic about their models but optimistic about their beliefs, stock returns will be high on average and PE ratios will exhibit large procyclical fluctuations. Even with complete markets and low risk aversion.
- Our key innovation is to link this apparent 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.