The material for this lecture is based on joint work with J.B. Heaton at Bartlit Beck Herman Palenchar & Scott
Motivation

- Competing theories of financial anomalies:
  - “Behavioral” theories of financial anomalies where agents have cognitive biases.
  - “Rational” theories of financial anomalies where agents have incomplete information about the “structure” of the economy.

- We seek:
  - A framework for comparing these theories.
  - An understanding of the ability to distinguish them empirically.
  - An understanding of their normative differences.
  - An understanding of their possible connections.
Plan of the lecture

1. Define rational expectations and present competing hypotheses: Behavioral versus Rational Structural Uncertainty models.
   - Illustrate both frameworks in simple models generating two financial anomalies: overreaction and underreaction.

2. Explore ability to distinguish the two models.
3. Explore normative differences between the models.
4. Explore connections between the two models.
5. Argue that we should integrate, rather than select, between the two theories.
6. Conclusions and work in progress.
The Rational Expectations Ideal

Friedman (1979, JME):

(A) *information exploitation*: “economic agents use efficiently whatever information is available.”

(B) *information availability*: “the information which is available to economic agents is sufficient to permit them to form expectations characterized by conditional subjective distribution outcomes that are equivalent to the conditional objective distribution of outcomes indicated by ‘the relevant economic theory’.”

• Sargent: two parts - 1) Individual behavior is optimal according to some perceived constraint. - 2) Consistency of perception across people. (consistency, at least implicitly, requires people to be forming beliefs about other’s decisions, about their decision processes, and even about their beliefs.

• Bray & Kreps: In equilibrium, there is a functional relationship b/w private signals and equilibrium price that establishes a statistical relationship b/w payoff relevant variables and equilibrium prices. If an agent knew both the statistical relationship and equilibrium price, the agent would use Bayes rule to calculate posterior assessment of payoff relevant variables and use this posterior belief to formulate demand. In a REE it is supposed that every agent knows the statistical relationship, and markets clear at equilibrium prices when information contained in equilibrium prices is taken into account by each agent.

• Huang and Litzenberger (p.263): (1) An equilibrium is a REE if individuals’ expectations are fulfilled ex-post. (2) A price system has two roles. First, determining an individual budget constraint as in a competitive equilibrium; Second, conveying information. An equilibrium where the price system plays both roles will be termed REE.
Behavioral Approach

Information exploitation

Information availability

- Investors cannot process available information correctly due to cognitive biases.
- Consistent with experimental results motivating behavioral finance. Cognitive biases arise in experiments despite the fact that experimental subjects can observe relevant data generating processes. Grether (1980).

Outside conventional rational expectations.
Structural Uncertainty Approach

Information exploitation

Information availability

- These rational investors optimally exploit all information available to them, but do not have all relevant information about the economy.


Merton (1987) presents a model of capital market equilibrium where a given investor is informed only about a subset of all securities, showing why, for example, the small-firm effect might arise.

Lewis (1989) studies dollar forecast errors during the 1980s subsequent to an increase in the U.S. money demand. She argues that these errors could have resulted from investors' prior beliefs that the change in the demand would not persist and the subsequent gradual learning about the true process generating fundamentals.

Timmerman (1993, QJE and 1996, RES) analyzes a representative investor who must learn dividend growth rates to price securities (i.e., they lack structural knowledge of this valuation-relevant parameter), and shows that least-squares learning can generate volatility and predictability.

Barsky and DeLong (1993) take up the problem of the rational investor who must estimate an unknown and possibly nonstationary dividend growth rate, showing that a simple learning process generates stock market volatility that is highly consistent with the data.

Morris (1996) presents a model where different Bayesian prior beliefs about an asset’s expected cash flows lead to the patterns of underperformance associated with initial public offerings.

Lewellen and Shanken (1999) examine an economy populated by rational Bayesian investors who possess imperfect knowledge regarding valuation-relevant parameters. They show that as these investors attempt to update their beliefs about these parameters, the resulting asset prices will exhibit predictability, excess volatility, and deviations from the CAPM.
Models

- A single, one-period risky asset.
  - At the end of each period t pays 1 with probability \( \theta_t \) and 0 with probability \((1- \theta_t)\).
- The representative investor is risk neutral, does not know the value of \( \theta_t \), but estimates it according to estimators described below, the forms of which depend on whether the agent is irrational or rational.
- The key structural feature of the economy is the stability of \( \theta_t \). Call \( \theta_t \) “stable” if it is time invariant, that is, if \( \theta_t = \theta \ \forall \ t \). Call \( \theta_t \) “unstable” if it varies through time.
- For simplicity we assume that \( \theta \) has changed at most one time in the last \( n \) time periods (though perhaps not at all).
- Complete structural knowledge entails:
  - (A) knowledge as to whether \( \theta_t \) is stable or unstable; and
  - (B) if \( \theta_t \) is unstable, the location of the change-point \( r \in (1,\ldots,N) \).

Given his risk neutrality, the representative investor values the asset at its expected payoff, \( r \), so \( r \) is the valuation-relevant parameter.

\( r \) denotes the observation after which payoffs are generated by the new probability.

The key structural feature of the economy (about which the behavioral investor is informed, but the structural uncertainty investor is not) is the stability of \( \theta_t \). Call "stable" if it is time invariant, that is, if \( \theta_t = \theta \ \forall \ t \). Call \( \theta_t \) “unstable” if it varies through time.
Behavioral Model #1: “Conservatism” (Leads to Underreaction)

- The conservatism bias occurs when prior beliefs receive excessive weight and data are underweighted [Edwards (1968)].

- We model the conservative investor as estimating $\theta$ as follows:

$$\hat{\theta}_{Beh,C} = \left( \frac{n}{n+c} \right) \frac{D_n}{n} + \left( \frac{c}{n+c} \right) \frac{1}{2}$$

where $c > 2$.

- The estimator always puts higher than optimal weight on the prior belief given the above assumptions.


Example:

Assume that after time $t = 10$, the history of asset payoffs is 0, 0, 0, 0, 0, 1, 0, 1, 1, 1. The sample mean is then 0.4. Assuming that $\theta_t$ is stable, the Bayesian posterior mean for $\theta$ (and the price of the asset at time $t = 11$) should be:

$$\hat{\theta} = \frac{10 \cdot 4}{12} + \frac{2 \cdot 1}{12} = 0.416$$

However, the investor employing the “conservative version” with, for example, $c = 5$, sets:

$$\hat{\theta} = \frac{10 \cdot 4}{15} + \frac{2 \cdot 1}{15} = 0.434$$

which places too much weight on the prior ($1/2$) and too little on the sample mean ($4/10$), resulting in an estimate that is, in this example, too high.
Behavioral Model #2: “Representativeness” (Leads to Overreaction)

- Subjects seem to expect key population parameters to be "represented" in any recent sequence of generated data, this phenomenon is known as the “representativeness heuristic” [Kahneman and Tversky (1972)].
- We assume (arbitrarily) that the representativeness heuristic leads the investor to consider only the most recent half of the available data, ignoring prior beliefs and older data.

\[ \hat{\theta}_{\text{Beh,RH}} = \left( \frac{D_{n/2}}{n/2} \right) \]

Example:
Assume that after time \( t = 10 \), the history of asset payoff is: 0, 0, 0, 0, 0, 1, 0, 1, 1, 1. Assuming that \( \theta \) is stable, the irrational investor using the representativeness heuristic will estimate \( \theta \) (and set the price of the asset at time \( t = 11 \)) as:

\[ \hat{\theta}_{\text{Beh,RH}} = \frac{4}{5} = 0.8 \]
A Model of Rational Structural Uncertainty

- Our rational investors employ Bayesian methods.
- However, unlike irrational investors, rational investors do not know whether or not $\theta$ is stable. The rational investor’s estimator for $\theta$ must incorporate this ignorance, and this is the ultimate source of financial anomalies.
- Assume that a change can occur at most once in the sample.
- $\theta_A$ denotes the probability generating the data up until the change point $r$.
- If there was a change, $\theta_B$ denotes the probability generating the data after the change point.
- Need to integrate over all possible change points $r \in \{1, \ldots, N\}$ [Smith (1975)].
A Model of Structural Uncertainty (Cont’)

- Diffuse prior beliefs on the \( n \) possible change points, but informative prior on probability of some change.

- Diffuse prior beliefs on \( \theta_A \) and \( \theta_B \).

- The marginal distribution for \( \theta \) at time \( t = n \) is therefore given by:

\[
\sum_{r=1}^{n} p_{\theta_A}(\theta_A|r) \cdot p_n(r) + p_{\theta_B}(\theta_B|r=n) \cdot p_n(n)
\]

- Interested in the posterior mean.

A uniform prior over the possible change-points, is in fact an "informative prior" that models a fairly strong belief in the potential instability of \( \theta \). Assigning identical probability to each possible change-point means that the no change point \( r = n \) receives prior probability \( 1/n \), while the total probability assigned to the event "some change" is \((n-1)/n \).

If we relax the uniform prior on the \( \theta_A \) and \( \theta_B \) and allow for an informative symmetric prior \( \text{B}(\alpha, \alpha) \) on both then the posterior for a change at point \( r \) is:

\[
\left( \begin{array}{c} r \\ D^r \end{array} \right) \cdot \left( \begin{array}{c} n-r \\ D^{n-r} \end{array} \right) \cdot \frac{1}{\binom{r+2\alpha-1}{D^r+a} \cdot \binom{n-r+2\alpha-1}{D^{n-r}+a}}
\]
Simulation Results: Overreaction

Figure 1 sets $n = 40$, and the true value of $\theta$ stable at 0.5. Based on 40 randomly generated observations, we calculate estimates under three models. “RE” is the “rational expectations” ideal in the stable environment, the Bayesian solution given in (3). This is the estimate that would be calculated by a Bayesian investor with complete structural knowledge, i.e., the knowledge that $\theta$ was stable. “SU” is the “structural uncertainty” estimator given in (9). The SU prices reflect fully Bayesian behavior, but the investor lacks the complete structural knowledge that would allow him to ignore the possibility that $\theta$ is unstable. “BehRH” is the behavioral “representativeness heuristic” estimator given in equation (1).

Viewing RE as a benchmark, two things are immediate from Figure 1. First, both SU and BehRH exhibit overreaction around RE. Prices (estimates) are “too extreme” versions of the RE ideal. Second, SU and BehRH exhibit virtually the same pattern of overreaction. Mechanically, this is no surprise. Figure 2 presents the weights placed on data points and prior beliefs after 25 observations. Both SU and BehRH share a heavy weighting of recent data compared to what is optimal in this environment (equal weighting of all past data). Estimates in the SU model are moderated two ways by comparison to BehRH, however. First, by positive weight on the prior. Second, by continuing to attach weight to older data. Both features capture the moderating influence of base rates that is lost by an investor employing the representativeness heuristic.

**Discuss the weights!**
Simulation Results: Underreaction

Figure 3 also sets $n = 40$, but now $\theta = 0.75$ generates all observations from 1,…,20, while $\theta = 0.25$ generates the remaining observations from 21,…,40. Based on 40 randomly generated observations, we again calculate estimates under three models. “RE” is again the “rational expectations” ideal, this time in the unstable environment. This means simply that the RE investor would use the Bayesian solution given in (3) on the data through time 20, and would then discard the data 1,…,20 and begin using (3) on the data 21,…,40. In this way, the Bayesian RE investor would not only act rationally but also incorporate his structural knowledge about the instability of. “SU” again is the “structural uncertainty” estimator given in (9). “BehC” is the behavioral “conservatism” estimator given in equation (4). Under our assumption that the behavioral investors do have structural knowledge (they know that there is instability in $\theta$ and they know the time location of the change from $\theta_A$ to $\theta_B$), the irrational investor applies his estimator to the data much as the RE investor, except, of course, that his estimator is not optimal given its conservatism. We arbitrarily set $c = 10$ in equation (4) (where $c$ is the parameter determining the degree of conservatism).

The important feature illustrated in Figure 3 is the effect of the change from $\theta = 0.75$ to $\theta = 0.25$. The RE estimator moves quickly to both the initial level and the new level after the change. Viewing RE as the benchmark, it is clear that both SU and BehC underreact to structural changes. The reasons are related but slightly different. The behavioral investor is attaching far too much weight to his prior beliefs and failing to extract enough information about the new environment. This despite the fact that the behavioral investor “knows” the right data to use. Consider observation 25, the fifth observation generated by new parameter. The optimal weight on the prior is, from (3), 0.29 (=2/7). The weight placed on the prior by the conservative investor is 0.67 (=10/15). At this point, the SU investor is not necessarily overweighting prior beliefs (his weight is about 20%) but this is not a general feature of the model. The SU investor’s prior weighting will in general be too high or too low, depending on the occurrence and location of actual parameter change in the data. More importantly for the SU investor, his inability to know exactly where the change occurred leaves him using substantial amounts of data from the old ($\theta_A$) environment. His estimates tend to drift rather slowly, as more and more data from the new environment enters his estimates.
Distinguishing the Competing Theories

Mathematical similarity

- Evident in simulations; proven in Appendix 3 that rational structural uncertainty model embodies the two essential features of the behavioral models — heavy weighting of prior opinion and heavy weighting of recent data.

Predictive similarity

- Predictions from simple behavioral and structural uncertainty models are equally consistent with available empirical evidence.

Use of data: A) \( \theta \) is stable: Investors placing heavy weight on prior beliefs may be acting irrationally and displaying conservatism, but they also may be placing more weight on prior beliefs in the (rational) belief that the underlying parameters are unstable, rendering old data irrelevant and thus lowering the available sample size. With a lower sample size, heavy weight on prior beliefs is optimal. Alternatively, investors placing heavy weight on recent data may be acting irrationally and displaying the representativeness heuristic, but they also may be placing more weight on recent data in the (rational) belief that the underlying parameters are unstable, rendering the older data less relevant to their estimates. B) \( \theta \) is unstable: then the SU investor does not necessarily overweight his prior (that depends on where the change occurs) but he does use too much of the old data and in that sense he is acting “conservatively”.

Note that increasing the prior weight on the no-change will, in case (A) lead to behavior which looks more like an informed rational Bayesian. In case B, obviously, it will strengthen the underreaction result.

1. Results on overreaction. Most of the evidence suggests that investors sometimes (but not always) attach too much weight to recently good or bad performance and over- or undervalue stocks accordingly. That the evidence is consistent with the major behavioral hypothesis that investors fall prey to the representativeness heuristic. But the nature of these tests also makes them consistent with the structural uncertainty model presented above. The portfolio formation strategies in overreaction studies tend to sort on proxies for stability. Thus, they may tend to select many stocks priced by investors whose concern with potential instability was (ex post) too severe. Consider the superiority of value stock investment strategies over growth stock investment strategies documented by Lakonishok, Shleifer, and Vishny (1994). Value stocks that outperform growth stocks are consistent poor performers, in terms of earnings, cash flow, and sales growth, both before and after portfolio formation. Growth stocks are consistent good performers, in terms of earnings, cash flow, and sales growth, both before and after portfolio formation. This portfolio formation strategy (and the similar returns-based approach of DeBondt and Thaler (1985)) may identify value stocks that are priced too low due to the overweighting of especially bad recent performance, and growth stocks that are priced too high due to overweighting of especially good recent performance.

2. Underreaction is consistent with conservatism, as investors overweight prior beliefs and fail to take proper account of the new evidence. But underreaction is also consistent with the presence of investors who are unsure whether the change in fact occurred, and continued to use too much data from an earlier environment. And, perhaps not surprisingly, the portfolio formation strategies in underreaction studies have sorted on good proxies for instability. Sample of such stocks may therefore include stocks priced by investors holding a mistaken belief in stability. Consider the superiority of momentum strategies documented by Chan, Jegadeesh, and Lakonishok (1996). They sort firms based on standardized unexpected earnings, extreme recent returns, and changes in analysts’ forecasts. On each measure, winners continue to be winners in the immediate future, while losers continue to be losers. It is at least plausible that such shocks reflect fundamental change in some underlying valuation-relevant parameter. The authors find little evidence of price reversals. The drift to new price levels is permanent, consistent with the existence of an actual change in a valuation-relevant parameter that investors recognized only slowly.
Normative Differences Between the Competing Theories

- Cost of Capital [Stein (1996)].


- Capital market regulatory policy [but can regulators do better?].

Positive:
Important questions in finance are likely to look very different in behavioral and structural uncertainty approaches [see Kuhn (1996)]. The very notion of risk aversion and the risk-return trade-off that plays such a crucial role in traditional finance will almost surely be reinvented in a behavioral theory, perhaps with a much greater role for risk seeking behavior in speculative markets and loss aversion [see O’Dean (1998)]. This, in turn, will call into question the traditional focus on diversification and hedging and could lead to a radically different theory of finance than we currently know.

Normative:
Barsky and DeLong (1993) and Lakonishok, Shleifer, and Vishny (1994). The former study of stock market fluctuations acknowledges that prices move more than future dividends, but there is no hint that the authors think they could have done better than the market in dealing with the uncertainty about dividend growth. The latter study of the value-growth anomaly reads as a virtual how-to manual in exploiting investor irrationality. Indeed, the authors have put their money where their mouths were: LSV Asset Management now has billions of dollars under management investing along the lines suggested by the paper.
Connections Between the Competing Theories

The “Limits of arbitrage” literature, Shleifer and Vishny (1997):
The survival of irrationality-induced anomalies, is linked to rational structural uncertainty on the part of capital providers.

- Structural uncertainty as a necessary condition for limits on arbitrage activity.

Shleifer and Vishny (1997): “Both arbitrageurs and their investors are fully rational…We assume that investors [who provide arbitrageurs with funds] have no information about the structure of the model determining asset prices…Implicitly we are assuming that the underlying structural model is sufficiently nonstationary and high dimensional that investors are unable to infer the underlying structure of the model from past returns data…”

This is where I start hinting about merging the two.

For example, the Shleifer and Vishny (1997) model could have assumed instead that the arbitrageur raised money from irrational investors whose beliefs are correlated with the noise traders against whom the arbitrageurs speculate. But such a theory would have been less convincing. Assuming that all investors are irrational but the capital-poor arbitrageur is somehow less convincing than assuming that there are some rational investors with capital who nevertheless fail to exploit every instance of mispricing. But if irrationality-induced anomalies survive because these rational actors with money do not bet fully against them, there must be a reason. Rational structural uncertainty provides such a reason, and Shleifer and Vishny (1997) rest their model upon it.

First, if there are rational arbitrageurs with capital, then rational structural uncertainty is virtually a necessary condition for the presence of an irrationality-induced anomaly. Behavioral theories that posit the survival of irrationality-induced anomalies should state explicitly the sources of rational structural uncertainty that will limit arbitrage. Second, further studies of arbitrage will likely benefit from an explicit acknowledgment of the tension between investor irrationality and rational structural uncertainty. On the one hand, more investor irrationality creates more opportunity for arbitrage profits. On the other hand, investor irrationality and structural uncertainty may be highly correlated. The optimal allocation of arbitrage resources requires a trade-off, and anecdotal evidence suggests that arbitrage resources are allocated roughly accordingly [see Shleifer and Vishny (1997)].
A Case for Integrating the Two Theories using Structural Uncertainty

- We have shown the similarity in the use of prior beliefs, old data, and recent data.

- The structural uncertainty approach,

  2) Provides a plausible *economic context* to derive various psychological behaviors (Conservatism, Overreaction and… Overconfidence).
A Case for Integrating the Two Theories using Structural Uncertainty

And,

3) Gives the flexibility to “hard-wire” certain beliefs. These beliefs are, after all, still mistaken.
4) Consistent with lab results which find that agents act less than fully rationally [Winkler and Murphy (1973), Gigerenzer (1991), Cosmides and Tooby (1994, 1996)].

Gigerenzer (1991): “One possibility is that the mind generalizes the specific content to a broader mental model that uses implicit domain-dependent knowledge about these structural assumptions. If so, then the content of problems is of central importance for understanding judgement—it embodies implicit knowledge about the structure of an environment.”

Shifts the emphasis from the labeling of behavior as “Rational” or “Irrational”. Agents differ in their views on the economic structure.
Structural Uncertainty and Overconfidence

- Individuals tend to overestimate their own ability. They overweight the precision of their favored hypothesis [Griffin and Tversky (1992)].

- Structural uncertainty provides an *economic* context within which overconfidence arises.
  - Overconfidence and overreaction will tend to coincide in stable environments.

- Applications.
Conclusions

1. The two competing theories relax, in two different ways the rational expectations assumptions.
   - Behavioral theories relax complete rationality while structural uncertainty theories relax complete structural knowledge.
2. Similarity of both theories, mathematically and predictively.
   - All theories allocate weights to (1) older data; (2) newer data; and prior beliefs similarly.
4. Connections between the two theories.
5. Structural uncertainty as a mean to integrate the two theories.
   - Application to modeling overconfidence.
6. Extensions.
Notes:

RE is embraced since it restricts the number of possible outcomes. Hence, we can make sharp predictions! Furthermore, do not need to deal with belief formation.

• Lucas’s critique: “The general hypothesis that economic agents are Bayesian decision makers has, in many applications, little empirical content: without some way of inferring what an agent’s subjective view of the future is, this hypothesis is of no help in understanding his behavior. Even psychotic behavior can be (and today, is) understood as “rational” given a sufficiently abnormal view of relevant probabilities. (1977:15)

Bray & Kreps (1987) Their paper focuses on the learning aspect of the statistical relationship. They consider two types of models: Irrational learning models (ILM) and rational learning models (RLM).

• ILM: Assume some learning process and check if converge to a \textit{stationary} and \textit{correct} model (Bray (1982)). The assumption is that agents “stay” with a model for a long time before deciding whether to change to a new model or not. In Blume and Easley (1982), agents believe that one of a set of stationary models is the right one. Every new data point they reassess the posterior on the models - but b/c they are changing their beliefs - they actual model is non-stationary so none of their models is actually correct. In any case, convergence to RE depends on how close are the initial prior beliefs to the true model (local stability). They find that RE is locally stable but that other non-rational equilibria are locally stable as well. Moreover, RE is not globally stable as well and agents may learn the wrong model. In Foster and Frierman (1990) the authors analyze the Blume and Easley economy and find conditions for global stability.

In Bray and Savin (1984), agents know the correct statistical model relating prices and payoff relevant parameters but they are uncertain about some of the parameters. They use OLS to estimate the parameters which is optimal \textit{only if the environment is indeed stationary}. But the environment is non-stationary since when their estimate changes, so does the equilibrium relationship. In these previous papers convergence \textit{may} occur if the effect of learning on equilibrium is small. Otherwise, cycles and divergence may occur, and convergence may occur to the wrong stationary model.

• RLM: In such a model agents’ uncertainty about the stat’ relationship stems from their lack of knowledge regarding some parameters in the economy. The stat’ relationship b/w pay-off relevant parameters and prices emerges when their learning and optimization is specified. This relationship is usually non-stationary. A RLM is one in which at every date \( t \) agents know this correct relationship up to the value of the unknown parameters. -- Starting with the way in which agents learn, we generate actual equil’ relationship, and then close the system by insisting that the learning process is in accord with the actual equilibrium relationship. The question of how agents learn about the equil’ relationship is moot. They know what the relationship is-- what we have is a model of rational learning of the (unknown) parameters within a REE. Bray & Kreps show that if you maintain RLM then beliefs will converge w.p.1 and beliefs will not converge to the wrong model. They cannot show though that the economy will settle into some stationary REE. [Frydman (1982) studies a RLM. To get convergence he needs that all firms know the “average opinion” in every period which in in his setup does not hold, hence he argues that convergence is unlikely to occur.]

• DeCanio (1979) makes the point that convergence to REE depends on the assumed learning process, the structural and stochastic parameters of the market. Makes the point that, in reality, agent may not converge to RE even if theoretically one exists and that it is an empirical issue.

----------

• Huang and Litzenberger When markets are \textit{complete} and individuals have differential information before trading, it is necessary that individuals have RE for an equilibrium to exist. LR define RE by the requirement that agent agree on the mapping b/w signals and the price functional (which is a function of the signals, state). In complete markets and asymmetrically informed individuals a competitive equilibrium does not exist. When markets are \textit{incomplete} then a competitive equilibrium may exist but individuals ignore the information in prices.

In any REE with complete markets, the price system must symmetrize the difference in information among individuals. In such an equilibrium the price system is fully revealing. Moreover, a fully revealing REE with complete markets always exists under standard conditions. Thus observationally, there is no distinction b/w a REE and a competitive equilibrium in an artificial economy that is identical in every aspect except that individuals are endowed with common prior information held in the REE. Remember that the fully revealing RE is problematic (GS) as no trader has an incentive to gather costly information hence the introduction of some sort of noise.

• When markets are incomplete then a REE may not exist.