Taylor & Francis
Taylor & Francis Group

Rational Expectations and Ambiguity

Author(s): Thomas J. Sargent

Source: *Financial Analysts Journal*, March/April 2014, Vol. 70, No. 2 (March/April 2014), pp. 14-19

Published by: Taylor & Francis, Ltd.

Stable URL: https://www.jstor.org/stable/24586262

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

Taylor & Francis, Ltd. is collaborating with JSTOR to digitize, preserve and extend access to *Financial Analysts Journal*

JSTOR

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms

FAJ | Financial Analysts Journal
Volume 70 • Number 2
©2014 CFA Institute

PERSPECTIVES

Rational Expectations and Ambiguity

Thomas J. Sargent

On 20 May 2013 at the 66th CFA Institute Annual Conference in Singapore, Thomas J. Sargent discussed his research on rational expectations and macroeconomics. He examined the probability distribution function of potential future outcomes in terms of ambiguity aversion and discussed market equilibrium under ambiguity aversion. He also considered whether swings in the market represent changes in the degree of ambiguity and, if so, whether those states can be predicted.

The concept of rational expectations is widely used in academic finance and macroeconomics. For better or worse, in model shops at central banks, it is used almost exclusively.

Widely used in economics, finance, and portfolio management, expected utility models comprise two elements: a probability distribution of potential outcomes and a utility function that indicates what we think about a particular outcome. The *utility function* expresses preferences about risk—whether we are risk averse (and so require compensation for accepting additional risk), risk neutral (do not require compensation for additional risk, making the risk-free rate universally appropriate for discounting), or risk loving (require something less than the risk-free rate to take on additional risk).

When contemplating or constructing the probability distribution function describing potential future outcomes, we try to model our "ignorance" about the future completely. We are uncertain about what the future will be but pretend to be certain about the possible future outcomes and the relative frequency or probability of those possible future outcomes. Again, we do not know what the future holds, but we do know the probability distribution function of what can happen in the future. So, we do know something about the future and can make decisions about the "uncertain" future.

A natural question arises: Where does this probability distribution over future events come from? And if the answer is, "From the participants in the market," where did they find it?

Thomas J. Sargent is professor of economics at New York University.

Rational Expectations

The concept of rational expectations is a widely used assumption in game theory, macroeconomics, and finance. It is almost automatically used by people who build models of financial crises and who build models for central banks. It assumes that a single probability distribution is agreed upon by all the participants within a given model. It is accurate only in the sense that it is shared by the reality that the analyst is trying to model.

Because everyone is assumed to think about the future probability distribution in the same way, I like to call this model a form of "assumed communism." The rational expectations assumption of a commonly shared probability distribution function makes us all communists, in one sense.

The assumed single probability distribution function over future outcomes is a very powerful tool. It is exploited cleverly in Lars Peter Hansen's innovative applications of generalized method of moments estimation to asset-pricing and other macroeconomic models. It is used to formulate the so-called dynamic general equilibrium models used by central banks. Ben Bernanke, the former US Federal Reserve chairman, relied on it when he cited the Diamond–Dybvig model (1983) of bank runs to justify many of his actions during the financial crisis.[1] I myself have been using the rational expectations assumption most of my adult life.

What is the justification for this assumption, beyond the defense that a lot of smart people use it? One justification is *simplicity*. It is dauntingly difficult to describe or model one person's probability distribution function, and it would be dreadfully more difficult to figure out a different probability

 ©2014 CFA Institute

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms

distribution possessed by a collection of people living in a model. So, for the sake of simplicity, the concept of rational expectations assumes that everyone thinks about the future probability distribution in the same way.

Another justification comes from assuming that everybody living in a model is very experienced and that there are enough observations to apply the law of large numbers. If we have enough observations, there are results that show we will all eventually agree about what we know. In a sense, the law of large numbers means that we all end up learning the same thing. Many people, including me, have used this argument.

The late Milton Friedman (1953) justified a rational expectations assumption on the basis of survival in a competitive process that yields survival of the fittest. Those who have realistic beliefs about the future eventually extract all the wealth from those who do not have realistic beliefs and consequently dominate the market. This evolutionary argument explains how the market finally settles on a particular future probability distribution.

These justifications call for some caveats. Simplicity is a double-edged sword. Making a model tractable and "simple" automatically means ignoring some possibilities that can emerge when people have diverse beliefs. Sometimes we oversimplify, to our detriment.

The law of large numbers does not tell us how many observations we need in order to "learn" or to agree about what we have "learned." And learning can take a very long time. I view appealing to the law of large numbers as a convenient "bluff" in that we are essentially saying that eventually we *will* learn but do not know *when* we will learn. Further, some infrequent events that may matter a great deal are ignored by the law of large numbers.

In considering Friedman's evolutionary argument, we must assume that markets are complete and that they allow participants to bet on the discrepancies between their beliefs. Friedman's evolutionary process toward a situation in which the most accurate beliefs about the future dominate the market is attenuated if there are too few markets or overregulated markets. Such circumstances may reverse the process by allowing unrealistic expectations of the future to prosper and not be competed away. Recently, some university professors have used this insight to argue for more regulation of the markets because they think it unfair that markets transfer resources from people with stupid beliefs to people with smarter beliefs about outcomes. This argument is not one I would make, but it is alive and maybe even influential.

Bayesians and Waiting on the Law of Large Numbers

When I talk about rational expectations in the sense that all of us have the same beliefs about the future, I am talking as an academic or as someone building a model for a central bank. But if we are decision makers and know that we will have to wait awhile for the law of large numbers to become informative, we will want to think differently about the source of our probability distribution over the future outcomes—that is, think outside the box provided by the law of large numbers.

One approach—the use of Bayesian statistics—is considered by some to be the most brilliant of all solutions and by others a refuge for scoundrels. Bayesians completely summarize what they do not know in terms of a subjective probability distribution function over outcomes. They create this distribution through introspection and head scratching (and maybe some beer). The subjective probability distribution function is considered just as valid as an objective probability distribution function based on the law of large numbers.

What is the source of that subjective probability distribution? The famous University of Minnesota economist and Nobel laureate Leo Hurwicz used to say that it is rude to ask a Bayesian where this distribution comes from; it is a personal matter, and from the point of view of Bayesian theory, one distribution is as good as another.

The upshot is that a Bayesian statistician has a unique joint probability distribution over all possible outcomes. For a Bayesian, the statistical theory of learning is trivial: Hand your initial subjective joint probability distribution to your staff and tell them to update conditional distributions as data flow in. Bayesians call this process applying the Bayes rule to generate conditional probabilities. During the learning process, the underlying unique initial subjective joint distribution does not change, but the conditional probability distributions change as they are updated with data over time.

The Bayesian approach is an essential part of expected utility theory. The process can be thought of as describing what one does while waiting for the law of large numbers to become applicable. Eventually, the data will update the conditional probability to make beliefs converge to something unique. Unfortunately, the data may be unable to tell us everything about the future—say, will the

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms

euro survive? If the euro does not survive, what will replace it? The future is so complicated and history dependent that it seems unrealistic to expect data to reveal everything we would like to know in order to make good decisions.

We could use an imaginative approach initiated by Fischer Black and Bob Litterman (1992), who asked, Given our utility function, does our subjective probability distribution of the future (also called our "prior") imply decisions that seem sensible?

To understand what Black and Litterman did, let us assume that portfolios are based on mean–variance analysis. Thus, we assume that our utility function is quadratic such that a security's mean return and return variance are the only statistical attributes that matter to the investor. Put differently, the investor is concerned with only risk and return (sound familiar?). When Black and Litterman (1992) performed their risk–return optimization for portfolio weights based on good least-squares estimates of means and variances of returns, what emerged were portfolio weights that were completely wacky, with extreme values for short and long positions in various securities.

The Black–Litterman response was to act as if they believed that the estimated covariances between security returns were correct but that the estimated mean returns of each security were not to be trusted. (There are deep statistical reasons that recommend those assumptions.) Then, Black and Litterman did some very clever things. Looking at the market portfolio, they backed out what a representative investor's subjective mean returns would have to be to rationalize that investor's being content to hold the market portfolio. They then computed the discrepancy between those market-implied subjective mean returns and the sample mean return. That discrepancy is the core of the Black–Litterman model. The reason I praise Black and Litterman's approach in this piece is that it is one of the first and most serious attempts to address the possibility of an important discrepancy between what investors believe and act on and what is actually out there in the real world. Their work naturally makes you think about multiple probability distributions and the differences between them.

Ambiguity

To look at this matter another way, let us consider something important called the Ellsberg Paradox (Ellsberg 1961). Richard Nixon hated Daniel Ellsberg because he released the Pentagon Papers. However, some Bayesian people hated him before Nixon did, and the following exercise will explain why. To examine the Ellsberg Paradox, we will imagine the following mental experiment, designed to distinguish "risk" from "uncertainty."

In a room are two urns: Urn A and Urn B. In Urn A are 10 black balls and 10 white balls. If we choose Urn A, we choose a color. Then nature randomly draws a ball from the urn. If the color we chose matches the color of the drawn ball, we receive \$10 million; otherwise, we receive nothing. The probability of winning by using Urn A is 50% because the number of black balls equals the number of white balls. So, the expected payoff associated with "playing" Urn A is \$5 million.

In Urn B are 20 balls that are black or white, but we do not know how many balls are black and how many are white. If we choose Urn B, the game is much the same as with Urn A. First, we choose a color, and then nature draws a ball. If the randomly drawn ball is the color we chose, we receive \$10 million; otherwise, we receive nothing. What is the expected payoff of choosing Urn B? We simply do not know, which was Ellsberg's way of expressing the idea that for Urn B there is *uncertainty* and for Urn A there is only *risk*.

The question now is, Which urn should we choose, A or B? Bayesians should choose Urn B because from a utility perspective, any subjective distribution of the proportion of white balls to black balls is equivalent to (if we assume 10 black and 10 white balls) or even better than (for any subjective distributions other than fifty-fifty) choosing Urn A. Think about it: Any prior that has a majority of balls of a particular color in Urn B favors choosing Urn B, then that color. For any subjective probability for the two colors, the expected payoff associated with choosing Urn B is at least \$5 million. For example, if the prior is 12 black balls and 8 white balls, then the expected payoff for choosing black is \$6 million (= 12/20 × \$10 million + 8/20 × \$0) and not \$5 million, the expected payoff for Urn A (= 10/20 × \$10 million + 10/20 × \$0). Consequently, by choosing Urn B, a Bayesian has, at worst, the same odds of winning as with Urn A and potentially better odds of winning, depending on the Bayesian's subjective probability for the balls in Urn B.

Ellsberg induced a group of famous economists and statisticians to participate in this experiment. Surprisingly, at least from a Bayesian perspective, almost all of them chose Urn A. Ellsberg concluded that those people were not acting like Bayesians. And that is why Bayesians, such as my good friend Chris Sims, do not like Ellsberg—or at least his experiment.

Why do people seem to prefer Urn A? Smart followers of Ellsberg hit upon an explanation that expresses people's aversion to "uncertainty" or

 ©2014 CFA Institute

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms

"statistical model ambiguity": When we do not know the distribution of future events and are unwilling to make up a subjective distribution of the future, there is ambiguity about the future and we cannot simply maximize an expected utility function to make a decision. What should we do? Some smart people, starting with Abraham Wald and Leo Hurwicz, said that we should embrace "min-max" behavior. At first, this approach may sound very paranoid when I explain it, but please wait before rejecting it too soon.

The reasoning behind using a min-max decision rule in the urn example is that when we choose a color, nature will choose a distribution from which to draw a ball that minimizes our expected payoff. At least, we act as if nature will always work against our color selection to "minimize" us. So, we seek to maximize our expected utility while assuming that nature will minimize our opportunity to win the game. Such min-max behavior induces us to choose Urn B. We do not behave like Bayesians.

To me, min-max expected utility behavior is not about being paranoid. Rather, it is a device to bound expected utility with respect to a set of possible probability models.

The preceding discussion of the Ellsberg Paradox sets the stage for the following brief discussion of a promising approach to coping with model uncertainty that comes from engineering and applied mathematics. It is called *robust control theory*.

The End of Communism

Now, let us not be communists in the limited sense of having only one probability distribution function over future events that everyone believes. Instead, we could perhaps agree that there is a set of potential models of the future (i.e., a set of different probability distribution functions over future outcomes). Because we have different interests (i.e., utility functions), we might make decisions on the basis of min-max decision theory; that is, the *min* part of our min-max decision problems would tell us to "choose" different probability distribution functions from the agreed-upon set of models against which to maximize our different utility functions. Such behavior would generate what looks like, *ex post*, belief heterogeneity.

How should we measure model ambiguity? Although we have a set of potential probability distribution functions of the future, we are unwilling to combine these distributions into a single probability distribution, as a Bayesian would. Suppose that we perform some econometric or statistical analysis to develop a model of returns and risks. After taking our best cut, suppose that, like Black and Litterman (1992), we ask ourselves whether we completely trust and believe in our model. If someone should ask us why we do not, we would just say that we know there are other probability distributions that fit the data as well or almost as well but we do not have enough data to prove convincingly that one distribution is better than another. The way my friend Lars Hansen and I proceeded in this situation was to measure the statistical proximity of different distributions (Hansen and Sargent 2008). To measure statistical discrepancy between two distributions, we used the statisticians' measure called *entropy*.

Given the probability distribution functions f and $\hat{f}$, we can calculate the expected value of the log likelihood ratio under one of the distributions, which is called *relative entropy*:

$$ent = \hat{E}\log\left(\frac{\hat{f}}{f}\right) = E\left(\frac{\hat{f}}{f}\right)\log\left(\frac{\hat{f}}{f}\right).$$

Relative entropy (*ent*) is a measure of how close two models are statistically, and it governs the rate at which we can distinguish between the two distributions as datasets grow.

So, Lars and I surrounded the probability distribution function that we got from our econometric work with an "entropy ball" containing other possible probability distributions that were very close in terms of fitting the data. There are deep reasons to think that Black and Litterman (1992) were using an entropy ball when they said that they trusted the covariances but not the means in creating their model.

Now let us consider a simple investment decision for a risk-neutral investor. Two securities are available: a risk-free security with a return of r_0 and a risky security with a mean return of μ and an associated variance of σ^2, both of which we know with certainty. We can borrow or save at the risk-free rate, and we can be short or long one share of the risky security. Under the Dodd–Frank Wall Street Reform and Consumer Protection Act, speculation is limited to one share.

What should the investment decision be? If $\mu > r_0$, then we should borrow money to purchase one share of the risky security. If $\mu < r_0$, then we should short one share of the risky security and invest the proceeds at the risk-free rate. If $\mu = r_0$, then we are indifferent as to being long or short the risky security.

Now let us change the scenario. Suppose that we do not trust the distribution of the risky return. To keep it simple, let us assume there is a set of possible

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms

means or an interval in which the mean resides. Although there was no bid–ask spread before, there will be now because of min-max behavior. It works like this: If we go long, nature will work against us and generate the lowest possible mean. If we go short, nature will work against us and generate the highest possible mean. From the Dow–Werlang model (1992) with ambiguity, we obtain

$$\min_{v=\pm\sqrt{2\eta}} \max_{\alpha\in(-1,1)} (\mu+\sigma v-r_0)\alpha$$

and

$$\alpha=\begin{pmatrix} 1 \text{ if } \mu-r_0>0 \text{ and } \mu-r_0-\sigma\sqrt{2\eta}>0 \\ -1 \text{ if } \mu-r_0<0 \text{ and } \mu-r_0-\sigma\sqrt{2\eta}<0 \\ 0 \text{ otherwise} \end{pmatrix},$$

where η measures entropy. There is a region of inactivity whose size depends on the number of other possible models that can exist around the approximating model as measured by entropy. This calculation demonstrates how ambiguity measured as entropy can freeze a market over a certain interval or cause the bid–ask spread to grow.

Asset Pricing and Model Uncertainty

The key equation in asset pricing is

$$E_t\left(m_{t+1}R_{j,t+1}\right)=1,$$

which has a stochastic discount factor, m_{t+1}, and a risky return, $R_{j,t+1}$. The discount factor is a random variable that encodes the conditional covariance between the discount factor and the return that sets the expectation to 1 (see Cochrane 2005). All formulas about risk premiums come from this equation. Because the covariance is so integral to the formula, a natural question is, On what probability distribution function is this covariance based? John Cochrane's book is based on one particular probability distribution function, which makes his book a communist book (my making that statement would really annoy John).

Lars Hansen and I attacked the problem by manipulating the stochastic discount factor (Hansen and Sargent 2008). By staring at the stochastic discount factor, one can figure out how risks are priced. Under the standard communist interpretation (i.e., one probability distribution function exists for everyone), all risk premiums are generated by people's attitude toward risk. If we rework that theory with ambiguity present, we find that people may have been misinterpreting the prices of risk when using the communist interpretation. These "prices" are not about risk but, rather, reflect ambiguity or model uncertainty (uncertainty about the correct probability distribution function); people hate being confronted with a situation in which they do not know the probability distribution function of future events.

Conclusion

I credit Lars Hansen with this summing-up "punch line": A small amount of model uncertainty can substitute for a large amount of risk in terms of explaining return premiums.

Lars and I wrote a paper on fragile beliefs that, in a sense, addresses "market stability" under model ambiguity. In our model, beliefs change rapidly when small amounts of additional data are introduced (Hansen and Sargent 2010). This dynamic might help explain volatility.

Another important question is whether swings in the market represent changes in the degree of ambiguity and, if so, whether those states can be predicted. In our 2010 paper on fragility, Lars and I defined fragility in an environment where ambiguity is ever present, which means that informational changes create large swings in market participants' beliefs. Let us imagine a pessimist who operates under the min-max premise and views good news as temporary and bad news as permanent. This logic, which is interior to the problem, causes these big swings in beliefs, particularly in how bad news is evaluated relative to good news. That is the "spirit" of our idea. I do not think our model is ready for us to use in trying to run a country because we still have to think through a number of issues.

This article qualifies for 0.5 CE credit.

Notes

1. Recently succeeded by Janet Yellen as Fed chair, Bernanke has said that we live in a world of "unusual uncertainty." How has he adapted his models to account for that? When we consider what he did as Fed chair, Bernanke is a kind of empirical macroeconomist who studies how well rules work under normal circumstances. In other words, he studies monetary policy rules that have been applied in the past to do certain things. What he did over the past four or five years is unprecedented. Because there are no past data concerning the effects of quantitative easing, Bernanke was like the person in the Ellsberg Paradox who chooses Urn B (discussed later in the article). Although Bernanke is very smart and well versed in theory, he was in a situation in which the data did not reveal the probability distribution function of the future that he sought. So, he had to make it up or make an educated guess as to what the distribution was. Bernanke was dealing with a much more complicated version of my little model of ambiguity because his models had a lot more moving parts.

 ©2014 CFA Institute

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms

He was operating one of the biggest hedge funds—no disrespect intended—in the United States and possibly the world. He was like our Dow–Werlang (1992) person but with very large long and short positions in some securities.

References

Black, Fischer, and Robert Litterman. 1992. "Global Portfolio Optimization." *Financial Analysts Journal*, vol. 48, no. 5 (September/October):28–43.

Cochrane, John H. 2005. *Asset Pricing*. Revised edition. Princeton, NJ: Princeton University Press.

Diamond, Douglas W., and Philip H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, vol. 91, no. 3 (June):401–419.

Dow, James, and Sergio Ribeiro da Costa Werlang. 1992. "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio." *Econometrica*, vol. 60, no. 1 (January):197–204.

Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics*, vol. 75, no. 4 (November):643–669.

Friedman, Milton. 1953. "The Case for Flexible Exchange Rates." In *Essays in Positive Economics*. Chicago: University of Chicago Press.

Hansen, Lars Peter, and Thomas J. Sargent. 2008. *Robustness*. Princeton, NJ: Princeton University Press.

———. 2010. "Fragile Beliefs and the Price of Uncertainty." *Quantitative Economics*, vol. 1, no. 1 (August):129–162.

CFA INSTITUTE BOARD OF GOVERNORS 2013–2014

Chair
Charles J. Yang, CFA
T&D Asset Management
Tokyo, Japan

Vice Chair
Aaron Low, CFA
Lumen Advisors
Singapore

CFA Institute President and CEO
John D. Rogers, CFA
CFA Institute
Charlottesville, Virginia

Saeed M. Al-Hajeri, CFA
Abu Dhabi Investment Authority
Abu Dhabi, United Arab Emirates

Giuseppe Ballocchi, CFA
Le Grand-Saconnex,
Switzerland

Heather Brilliant, CFA
Morningstar, Inc.
Chicago, Illinois

Beth Hamilton-Keen, CFA
Mawer Investment Management Ltd.
Calgary, Alberta, Canada

Robert Jenkins, FSIP
London Business School
London, United Kingdom

James G. Jones, CFA
Sterling Investment Advisors, LLC
Bolivar, Missouri

Attila Koksal, CFA
Standard Unlu A.S.
Istanbul, Turkey

Mark J. Lazberger, CFA
Colonial First State Global Asset Management
Sydney, Australia

Frederic P. Lebel, CFA
HFS Hedge Fund Selection S.A.
Founex, Switzerland

Colin W. McLean, FSIP
SVM Asset Management Ltd.
Edinburgh, United Kingdom

Alan M. Meder, CFA
Duff & Phelps Investment Management Co.
Chicago, Illinois

Matthew H. Scanlan, CFA
RS Investments
San Francisco, California

Jane Shao, CFA
Lumiere Pavilions Limited
Beijing, China

Sunil Singhania, CFA
Reliance Mutual Fund
Mumbai, India

Roger Urwin
Towers Watson Limited
Surrey, United Kingdom

This content downloaded from
149.10.125.20 on Tue, 25 Jan 2022 16:10:25 UTC
All use subject to https://about.jstor.org/terms