DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

MARTINGALE RESTRICTIONS ON EQUILIBRIUM PRICES OF ARROW-DEBREU SECURITIES UNDER RATIONAL EXPECTATIONS AND CONSISTENT BELIEFS

Peter Bossaerts

SOCIAL SCIENCE WORKING PAPER 958

May 1996

Martingale Restrictions
On Equilibrium Prices of Arrow-Debreu Securities
Under Rational Expectations And Consistent Beliefs

Peter Bossaerts

Abstract

Consider the Rational Expectations price history of an Arrow-Debreu security that matures in the money: $p_1, p_2, ..., p_T$. Past information can be used to predict the return $(p_{t+1} - p_t)/p_t$. Now consider a simple alternative performance measure: $(p_{t+1} - p_t)/p_{t+1}$. It differs from the return only in that the future price is used as basis. This variable cannot be forecasted from past information. The result obtains even if investors' beliefs are biased, i.e., prices are not set in a Rational Expectations Equilibrium (REE). It depends only on investors' using the rules of conditional probability to process information. More precisely, the result continues to hold in the Bayesian Equilibrium with Consistent Beliefs (CBE) introduced by Harsanyi [1967]. Many related results are proved in this paper and extensions to the pricing of equity subject to bankruptcy risk are discussed.

JEL Classification : C22, D84, G14.

Keywords: Arrow-Debreu Securities, Rational Expectations Equilibrium, Bayesian Equilibrium, Consistent Beliefs, Market Efficiency, Learning, Martingales, Reverse Time, Survivorship Bias.

Martingale Restrictions On Equilibrium Prices of Arrow-Debreu Securities Under Rational Expectations And Consistent Beliefs*

Peter Bossaerts

1 Introduction

This paper studies the time series behavior of security returns when a crucial cash flow determinant remains constant over time. This determinant has to be interpreted in a wide sense: it could be a company's bankruptcy status at a certain future date; whether a natural resources firm will obtain long-term monopoly rights for the exploitation of an oil field at some time in the future; whether an electronics firm will have its technology accepted as unique nationwide standard by the government. More abstractly, the cash flow determinant is best described as a *state variable* uniquely describing the path of a security's payouts in the future. This state variable could simply be a real number fixed at the outset, or an infinite-dimensional vector whose entries specify the realizations over time of a latent variable generating the cash flows.

Investors do not know the value of the state variable, but gather information as time passes by. Prices will be set in a Rational Expectations Equilibrium (REE). This means that investors use Bayes' rule to process information. If certain agents fail to observe pieces of information that others see, they learn from the equilibrium price. In accordance with standard definitions of REE, investors' inference will be assumed to be based on correct beliefs. This implies, among other things, that their predictions are unbiased in cross-section (Muth [1961], Radner [1967], Lucas [1972], Green [1973]).

The results of the paper, however, will not depend on Rational Expectations. Investors may have beliefs that deviate substantially from the laws that Nature uses to

*California Institute of Technology. Mailing address: m/c 228-77, California Institute of Technology, Pasadena, CA 91125; Phone (818) 395-4028; e-mail: pbs@rioja.caltech.edu. This is part of a paper that the author presented at the Université des Sciences Sociales, Toulouse, in April 1995. He is grateful for the many comments from the seminar participants. He also thanks Christian Gouriéroux for suggesting the analysis in reverse time. In addition, the paper benefited through discussions at seminars at Carnegie Mellon University, U.C. Riverside, University of Minnesota, and through comments from Oleg Bondarenko, Alan Kraus, P.C.B. Phillips and Richard Roll. Oleg Bondarenko provided able research assistance. The usual disclaimer applies.

determine the value of the state variable in a particular history. Their inference, however, must continue to be rational, i.e., it must be based on Bayes' rule. Investors' beliefs may even differ, but, these differences must actually be made inconsequential by assuming that they satisfy Harsanyi's consistent prior restriction (Harsanyi [1967]). This rules out investors' agreeing to disagree (Aumann [1976]) and the corresponding difficulty of having to deal with speculation (Harrison and Kreps [1978], Tirole [1982]). The extension of REE to arbitrary but consistent beliefs will be referred to as *Consistent Beliefs Equilibrium* (CBE).[1]

The canonical form of an asset whose payoff depends on a state variable that is fixed in a given time series is the *Arrow-Debreu security*. Consequently, the paper will focus on this rather abstract yet powerful security. As is standard in Arrow-Debreu securities analysis, the time horizon is taken to be finite. The discount rate can therefore be taken to be zero. The extension to an infinite horizon would force one to leave the elegant world of Arrow-Debreu securities. Also, in order to focus on asset return characteristics induced by the updating of beliefs about the unknown state variable, investors will be assumed to be risk neutral.

In this paper's world, securities analysis would be simple if one had a cross-section of independent realizations. One could exploit the unbiasedness property of the REE or readily estimate the initial common prior in the CBE. The paper starts from the premise that the analyst either has *only a single price series* at her disposal or is reluctant to embark on a cross-sectional analysis by *lack* of identifying information such as *cross-sectional independence or exchangeability*. In other words, the paper initially focuses on the type of analysis that is predominant in asset pricing, namely time series analysis. Nevertheless, the findings are also of importance for the analysis of cross-sections of independent price histories, because it allows one to readily distinguish rational learning from biases in beliefs without having to estimate priors. In other words, it provides a test of the validity of REE within the larger class of CBE. To underscore this point, results from a study of data from the Iowa experimental markets will briefly be discussed.

From the perspective of the marginal investor setting prices, i.e., using her information and beliefs, securities price changes, and, hence, returns, ought not to be predictable. Samuelson [1965] was one of the first to point out this *martingale property* in the context of a REE. It holds in a CBE as well, because it is based entirely on the rationality of the updating rule.

Samuelson's martingale result has often been interpreted as implying that returns in a time series ought to be unpredictable. In investments analysis, this property is usually referred to as *market efficiency*. The time series mean return should be zero on average (under the assumption of risk neutrality and zero discount rates), and projections of returns onto past information should reveal insignificant slope coefficients.

[1]See Biais and Bossaerts [1995] for an explicit comparison of REE, CBE and Bayesian Equilibrium under Inconsistent Beliefs (speculative equilibrium) in a financial markets setting.

It does not take much imagination to see that market efficiency will not follow in the world of this paper. The distribution that generates a particular return series differs from the one used by the marginal investor to set prices. Under the latter, returns cannot be forecasted. In contrast, predictability will generally be observed under the former. In particular, Arrow-Debreu securities that eventually pay off will feature price series that trend upward; those that mature out of the money may exhibit negative returns on average.

The import of the paper would be minor if its only contribution were to advance this obvious point.[2] Instead, the main purpose is to present martingale results for payoff-related quantities that obtain under the distribution underlying time series analysis. The martingale restrictions do not obtain for traditionally computed returns (except for one involving the investigation of series in reverse time). The paper will mainly use a modified return measure, where the price change is divided by the *future price* instead of the lagged price.

The martingale restrictions obtain under REE and CBE alike. The latter is attractive: REE have recently been criticized for being irrealistic because investors' beliefs must be unbiased from the outset. Temporary, and, certainly, permanent biases are not allowed. The former are likely to occur. But even the latter cannot be ruled out, as Kurz [1994] argues. The CBE retain only the most important feature of REE, namely *Rational Learning*. In this sense, it is comforting to know that this paper's martingale restrictions on returns and modified returns are robust to a generalization of REE to CBE.

The paper also discusses extensions to the pricing of equity subject to bankruptcy risk. The martingale restrictions on modified returns become weaker. Their behavior, however, continues to be sharply distinguishable from that of the traditional return measure. This allows the empiricist to test the rationality of equity pricing even if bankruptcy leads to otherwise difficult statistical problems, namely survivorship bias and sample truncation.

Empirical tests of asset pricing models have invariably been based on REE instead of CBE, perhaps because of a widespread fear that the results would otherwise depend all too delicately on the specification of priors (a criticism often raised against Bayesian analysis). This paper's contribution is to show that there are aspects of asset price data that are invariant across beliefs structures. These ought to form the focus of future empirical research in finance. Only *after* one fails to observe the martingale restrictions derived in this paper, can one safely question the rationality of pricing in financial markets. This contrasts with many empiricists' unfortunate tendency of immediately regarding any deviation from REE as evidence of irrational or boundedly rational behavior (e.g., De Bondt and Thaler [1985]).

[2] Discussions of the problem abound in the empirical finance literature, mostly in terms of an environment where the underlying fixed parameter determines eventual survival of the asset. The empiricist is warned for the consequences of observing complete price series only for assets that have survived. See Brown, Goetzmann and Ross [1995] for a discussion of how survivorship bias could explain several asset pricing anomalies.

The remainder is organized as follows. The next section formally describes the problem that the paper addresses. Section 3 presents the results under REE. A simple, finite-state example is discussed in Section 4 (some readers may find that they understand the issues more readily by first looking at this example). Section 5 proves that the same restrictions hold when extending the notion of equilibrium from REE to CBE. Section 6 provides illustrations. Section 7 discusses extensions to the price behavior of equity subject to bankruptcy risk. The relevance of the results of the paper for cross-sectional analysis is treated and illustrated with experimental market data in Section 8. Section 9 lists some open questions. Section 10 concludes.

2 Formal Problem Statement

Let time be indexed $t = 1, 2, 3, ..., T+1$. The possible states of the world are listed in a sample space Ω. We shall consider assets with a payoff V at $T+1$. V is a random variable defined on Ω. (At this point, we are not yet restricting our attention to Arrow-Debreu securities, so we remain aspecific about V.) The flow of information to the market will be represented by a filtration $\{\mathcal{F}_t\}_{t=1}^{T+1}$.[3] $V \in \mathcal{F}_{T+1}$. The probability measure P governs $(\Omega, \{\mathcal{F}_t\}_{t=1}^{T+1})$.

Let p_t denote the equilibrium price at t. $p_t \in \mathcal{F}_t$. Under REE, and subject to risk neutrality and zero discounting (restrictions that will be kept throughout the paper),

$$p_t = E[V|\mathcal{F}_t].$$

This implies the following fundamental result, originally derived in Samuelson [1965]. Let r_{t+1} denote the return over the period $t, t+1$:

$$r_{t+1} = \frac{p_{t+1} - p_t}{p_t}. \tag{1}$$

Lemma 1 *Under REE,*

$$E[r_{t+1}|\mathcal{F}_t] = 0.$$

(All proofs are collected in the Appendix.)

Lemma 1 has generally been interpreted as implying that there must not be any predictability in time series of asset returns, a property that is usually referred to as *market efficiency* (Fama [1970]). Historical averages of returns ought to be zero (except for sampling error); projections of a history of returns onto lagged information ought to produce insignificant coefficients.

[3]If individual agents' information differs, $\mathcal{F}_t$ is to be interpreted as the information reflected in the equilibrium price.

In many cases, however, absence of predictability does not follow from Lemma 1. One case is focused on here, namely where Ω can be factored into a parameter set Θ and a remainder $\tilde{\Omega}$:

$$\Omega = \Theta \times \tilde{\Omega}. \tag{2}$$

Elements of Θ are denoted θ; those of $\tilde{\Omega}$ are denoted $\tilde{\omega}$. Nature picks a $\theta \in \Theta$ at the outset. It determines the information flow $\{\mathcal{F}_t\}_{t=1}^{T}$ as well as the final payoff V (V is a nontrivial function of θ). Nevertheless, $\theta \not\in \mathcal{F}_t$, all $t \leq T$. In other words, signals are never fully revealing. As will be demonstrated in Section 4, however, the theoretical results continue to hold even if signals may fully reveal the state θ prior to $T+1$. This full-revelation case nevertheless does lead to serious theoretical-statistical problems. Fortunately, the practical relevance of those problems will often be limited, as the simulations in Section 6 and 7 illustrate.

The example that we are going to use throughout is that of an Arrow-Debreu security that pays one dollar if θ is revealed at T to be $\overline{\theta}$. In that case,

$$V(\theta, \tilde{\omega}) = 1_{\{\theta=\overline{\theta}\}}$$

($1_{\{\}}$ denotes the indicator function). In addition to $p_1, ..., p_t$, $\mathcal{F}_t$ could include a history of signals $s_1, ..., s_t$ (a signal is a function of θ and $\tilde{\omega}$). Since the signals reveal information about θ, the equilibrium price p_t ($= E[V|\mathcal{F}_t]$) will progressively reflect more about θ. The return r_{t+1} will reflect the updating between time t and $t+1$.

In this situation, it will still be true that

$$E[r_{t+1}|\mathcal{F}_t] = 0$$

(Lemma 1). The *actual* behavior of the signals, the equilibrium prices, and, hence, the return series $\{r_{t+1}\}_{t=1}^{T-1}$, however, will be determined by the value of θ picked by Nature at the outset. More formally: signals, prices and returns are draws from the *conditional* probability measure P_θ. Despite the fact that $E[r_{t+1}|\mathcal{F}_t] = 0$ (Lemma 1), $E[r_{t+1}|\mathcal{F}_t \vee \theta]$ differs from zero (except in trivial cases), and, hence,

$$\begin{aligned} & E[\frac{1}{T-1}\sum_{t=1}^{T-1} r_{t+1}|\theta] \\ = \quad & E[\frac{1}{T-1}\sum_{t=1}^{T-1}(r_{t+1} - E[r_{t+1}|\mathcal{F}_t \vee \theta])|\theta] \\ & + E[\frac{1}{T-1}\sum_{t=1}^{T-1} E[r_{t+1}|\mathcal{F}_t \vee \theta]|\theta]. \end{aligned}$$

The first term in the last expression is zero; the second term generally is not.[4]

In fact, it is not difficult to prove rigorously the following.

[4]One exception is when T is large and beliefs converge, which implies that $E[r_{t+1}|\mathcal{F}_t \vee \theta] \to E[r_{t+1}|\mathcal{F}_t] = 0$. The Cesàro sum in the second term will then converge in P to zero.

Lemma 2 *Let $\{\mathcal{F}_t\}_{t=1}^T$ be the information filtration generated by prices and (conditionally) independent signals $s_1, s_2, ..., s_T$. Under REE,*

$$E[r_{t+1}|\mathcal{F}_t \vee \overline{\theta}] > 0,$$

where r_{t+1} denotes the return over the period $t, t+1$ on an Arrow-Debreu security that pays one dollar when $\theta = \overline{\theta}$.

In words: the return on Arrow-Debreu securities that eventually pay off one dollar is positive on average.

The implications of the foregoing are straightforward: historical average returns on Arrow-Debreu securities will predominantly be found to be significantly different from zero. Likewise, projections of return histories on past information will often exhibit significant coefficients. Substantial evidence against "market efficiency" will be found. Yet price setting follows the rigour of REE; Lemma 1 is not violated.

We shall refer to the conditional measure P_θ as the *empirically relevant measure.* It is the one that governs the behavior of signals, prices and returns in a particular history. Hence, it is the one that the empiricist will be able to estimate. To put it differently: P_θ determines the behavior of the statistics that the empiricist computes from historical data.

The problem that this paper addresses can now be stated formally. In the above situation, time series analysis of asset returns will reveal violations of martingale restrictions. Are there payoff-related variables for which martingale behavior would obtain under the empirically relevant measure? A payoff-related variable is one that can easily be computed from prices and that measures somehow the payoffs on the asset. We shall use the notation x_{t+1} to denote any payoff-related variable measured using prices observed at time t and $t+1$. What we are asking is really: for which payoff-relevant variables can we state that:

$$E[x_{t+1}|\mathcal{F}_t \vee \theta] = 0? \tag{3}$$

Clearly, the traditional return measure, r_{t+1} (see Eqn. (1)), is not a candidate. If we can find an affirmative answer for a particular payoff-related variable, the time series analyst could focus on it and verify that it cannot be predicted, confirming REE.

3 Empirically Relevant Martingale Restrictions Under REE

We shall maintain the assumption throughout that the filtration $\{\mathcal{F}_t\}_{t=1}^T$ is generated by prices and a sequence of (conditionally) independent signals $s_1, ..., s_T$. Most of the results to be discussed continue to hold when the signals are dependent over time. For the time being, we should refrain from relaxing the independence assumption, in order not to distract. Later on, we shall come back to the issue of independence.

Given this assumption, Lemma 2 obtains. It states that returns tend to be positive. The question one should ask is whether there exists a *simple correction* to the calculation of returns that cancels the positive drift. Returns are traditionally computed by dividing an asset's payoff by its price at the *beginning* of the period (see Eqn. (1)). Would the drift be canceled if we were instead to divide the asset's payoff by its *end-of-period* price? The resulting payoff-relevant variable, to be denoted x_{t+1}^{+}, would be defined as follows:

$$x_{t+1}^{+} = \frac{p_{t+1} - p_t}{p_{t+1}}. \tag{4}$$

Because of Lemma 2, the end-of-period price tends to be higher than the beginning-of-period price. Using the former as basis to compute returns may offset the bias.

Surprisingly, this works.

Theorem 3 *Let $\{\mathcal{F}_t\}_{t=1}^{T}$ be the information filtration generated by the prices $p_1, ..., p_T$ as well as (conditionally) independent signals $s_1, ..., s_T$. Under REE,*

$$E[x_{t+1}^{+} | \mathcal{F}_t \vee \overline{\theta}] = 0,$$

where x_{t+1}^{+} is computed as in Eqn. (4), using prices of an Arrow-Debreu security that pays one dollar when $\theta = \overline{\theta}$.

In words: returns on the Arrow-Debreu securities that eventually pay off one dollar will be a martingale difference sequence, provided they are computed using the end-of-period price as basis.

The proof relies on arguments that have been used to show that likelihood ratio sequences are martingale processes (see, e.g., Doob [1953], II.7 and VII.9). Likelihood functions appear here because prices change as the result of an information accumulation process that can be written explicitly in terms of likelihood functions and a prior (Bayes' rule). Likelihood functions in particular play a crucial role in the determination of the properties of the equilibrium price processes. It turns out that the second determinant, the prior, is irrelevant, a property that we will exploit to show that Theorem 3 continues to hold when the scope of equilibrium is extended from REE to CBE (where agents' priors do not necessarily coincide with the law according to which Nature draws θ). The arguments of this extension are, however, subtle. Since the proof of Theorem 3 is really simple, we therefore plan to formally discuss it in the next section. This will provide the right background against which the subtleties of the extension to CBE will be best understood.

Theorem 3 restricts the behavior of a payoff-related variable for the Arrow-Debreu securities that will mature "in the money." These securities are often the *survivors*, because usually only they generate complete price records. One can then interpret Lemma 2 as providing an expression for the *survivorship bias*, whereas Theorem 3 suggests a *correction.* Many have pointed out the existence and nature of survivorship biases in securities price data (most recently, Brown, Goetzman and Ross [1995]), but no correction has been advanced yet that does not rely on a cross-section that either is not available or is contaminated by strong dependence. Therefore, Theorem 3 provides the first foundation of a formal analysis of securities price data under survivorship bias. We shall come back to this point in Section 7.

Can we say something about payoff-related variables for the other Arrow-Debreu securities? Can one obtain martingale results for Arrow-Debreu securities that eventually mature "out of the money"? Fortunately, the answer is affirmative. Moreover, it does not involve complicated payoff-related variables. In contrast, it relies on the traditional return measure (Eqn. (1)). What is novel, however, is that it requires an analysis of return series in *reverse time.* Returns will constitute a martingale difference sequence in reverse time only, which means that verification entails a projection of returns onto *future information.* Unfortunately, the result obtains only for binary parameter spaces.

Let $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Introduce an index n which increases backward in time. $n = 1, 2, ..., T$. It is related to the index forward in time, hence, we can write: $n(t)$. Set: $n(t) = T - t + 1$. A sequence of prices in reverse time starts with the value of the security at maturity (p_T) and ends with the price at $t = 1$. E.g., the fifth price in the reverse-time sequence is the $(T-4)$th price in the original sequence.

The return in period $(n-1, n)$ in reverse time, x_n^-, will be defined as the traditional return r_{t+1} with $t = T - n + 1$:

$$\begin{aligned} x_n^- &= \frac{p_{T-n+2} - p_{T-n+1}}{p_{T-n+1}} \\ &= r_{T-n+2}. \end{aligned} \tag{5}$$

Since we are analyzing price and return series in reverse time, we cannot use the original information filtration $\{\mathcal{F}_t\}_{t=1}^T$, which accumulates forward in time. The analysis of time series in reverse requires that one accumulate information backward in time. Let us introduce the filtration $\{\mathcal{F}_n^-\}_{n=1}^T$. We take $\mathcal{F}_n^-$ to be generated by the signals observed at $t = T, ..., T-n+2$ (when $n > 1$), and the prices observed at $t = T, ..., T-n+1$.[5] The probability measure governing $(\Omega, \{\mathcal{F}_n^-\}_{n=1}^T)$ continues to be P. As before, P_θ denotes the measure corresponding to conditioning on θ. It is the empirically relevant measure used to compute expectations, such as $E[x_n^-|\theta]$.

We obtain the following attractive theorem.

Theorem 4 *Let $\{\mathcal{F}_n^-\}_{n=1}^T$ be the information filtration generated by (conditionally) independent signals as well as market prices observed prior to n in reverse time. Let Θ be*

[5]The signal observed at $T - n + 1$ is *not* included in $\mathcal{F}_n^-$. If it were, the empiricist would be able to infer p_{T-n}, using Bayes' law. Yet, p_{T-n} is not observed till point $n+1$ in reverse time. See the proof of Theorem 4 for details.

binary, namely, $\Theta = \{\underline{\theta}, \overline{\theta}\}$. *Under REE,*

$$E[x_n^- | \mathcal{F}_{n-1}^- \vee \underline{\theta}] = 0,$$

where x_n^- *is the traditional return over the period* $n-1, n$*, computed as in Eqn. (5), using prices of an Arrow-Debreu security that pays one dollar when* $\theta = \overline{\theta}$*.*

Notice the simplicity of this result: when the parameter space is binary, returns on Arrow-Debreu securities that eventually mature "out of the money" are not predictable from *future information.* Forward in time, returns on these securities are predictable (although the analogue to Lemma 2 does not obtain - we discuss this point next). Backward in time, however, the empiricist cannot predict.

Using the law of iterated expectations, we can condition out $\mathcal{F}_{n-1}^-$ in Theorem 4 and we obtain the following.

Corollary 5 *Under the conditions of Theorem 4,*

$$E[r_t | \underline{\theta}] = 0,$$

where r_t *is the traditional return over the period* $(t-1, t)$*, computed as in Eqn. (1), using prices of an Arrow-Debreu security that pays one dollar when* $\theta = \overline{\theta}$*.*

In words: the unconditional expected return on the Arrow-Debreu security that matures "out of the money" is zero. This remarkable result implies that there is no a priori drift to the price of this security (when expressed as percentage price change). It contrasts with Lemma 2, which stated that the price of the Arrow-Debreu security that matures "in the money" does display a positive drift. In analogy to this Lemma, one would have expected the drift in the price of the Arrow-Debreu security that matures with a zero payoff to be negative. Such is not the case.

The binary parameter space may appear overly restrictive until it is realized that the payoff on many financial securities can be written as a function of a binary unknown state. The most important example is equity: at a fixed date in the future, the firm is either bankrupt or not; if it is, equity pays zero; if not, equity pays a (random) positive amount. The Arrow-Debreu security is simpler to analyze, because it pays a fixed amount (one dollar) in the second state. Hence, it is the ideal primitive security with which to study time series analysis under bankruptcy potential. In this sense, Theorem 4 provides the necessary foundation for the formal analysis of the history of prices of equity in bankrupt firms.

To complete the analysis of returns in REE, we consider the behavior of x_n^- in reverse time for the Arrow-Debreu security that matures "in the money." The following result is proved in the Appendix.

Theorem 6 *Let $\{\mathcal{F}_n^-\}_{n=1}^T$ be the information filtration generated by (conditionally) independent signals as well as market prices observed prior to n in reverse time. Let Θ be binary, namely, $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Under REE,*

$$E[x_n^- | \mathcal{F}_{n-1}^- \vee \overline{\theta}] > 0,$$

where x_n^- is the traditional return over the period $n-1, n$, computed as in Eqn. (5), using prices of an Arrow-Debreu security that pays one dollar when $\theta = \overline{\theta}$.

This result is complementary to that in Lemma 2. The latter predicted a positive average return on the Arrow-Debreu security whose payoff is nonzero at maturity. According to Theorem 6, the empiricist would forecast a positive return in reverse time as well. Unlike Lemma 2, however, it requires a binary parameter space.

4 A Simple, Finite–State Example

Let us briefly discuss an explicit example. Besides illustrating the main results, it facilitates the discussion of the extension to the case of a potentially fully revealing signal prior to $T+1$, as well as the theoretical-statistical problems it would create.

Let Θ be binary: $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Let $T = 3$. Consider the price p_t ($t = 1, 2, 3$) of a security that pays \$1 if $\theta = \overline{\theta}$, \$0 otherwise. At $t = 1$, agents are agnostic about the state of the world, i.e., $\mathcal{F}_1 = \{\phi, \Omega\}$. At $t = 2$, they receive a signal, s_2, as follows. If $\theta = \overline{\theta}$ then:

$$s_2 = \begin{cases} 1 & \text{with probability } 1/2, \\ 0 & \text{with probability } 1/2; \end{cases}$$

if $\theta = \underline{\theta}$ then:

$$s_2 = \begin{cases} 1 & \text{with probability } 1/3, \\ 0 & \text{with probability } 2/3. \end{cases}$$

This constitutes $\mathcal{F}_2$. Later on, we will discuss $\mathcal{F}_3$.

Setting $P\{\theta = \overline{\theta}\} = 1/4$, we obtain:

$$p_1 = \frac{1}{4}.$$

Letting $l_2(s_2|\theta)$ denote the conditional likelihood of s_2, Bayes' rule implies that if $s_2 = 1$:

$$\begin{aligned} p_2 &= \frac{l_2(s_2 = 1|\overline{\theta})p_1}{l_2(s_2 = 1|\overline{\theta})p_1 + l_2(s_2 = 1|\underline{\theta})(1 - p_1)} \\ &= \frac{\frac{1}{2}\frac{1}{4}}{\frac{1}{2}\frac{1}{4} + \frac{1}{3}\frac{3}{4}} \\ &= \frac{1}{3}. \end{aligned}$$

Analogously, if $s_2 = 0$:

$$\begin{aligned} p_2 &= \frac{l_2(s_2 = 0|\overline{\theta})p_1}{l_2(s_2 = 0|\overline{\theta})p_1 + l_2(s_2 = 0|\underline{\theta})(1 - p_1)} \\ &= \frac{\frac{1}{2}\frac{1}{4}}{\frac{1}{2}\frac{1}{4} + \frac{2}{3}\frac{3}{4}} \\ &= \frac{1}{5}. \end{aligned}$$

Consequently:

$$\begin{aligned} E[p_2|\mathcal{F}_1 \vee \overline{\theta}] &= \frac{1}{3}\frac{1}{2} + \frac{1}{5}\frac{1}{2} \\ &= \frac{4}{15} \\ &> \frac{1}{4} \\ &= p_1, \end{aligned}$$

i.e.,

$$E[r_2|\mathcal{F}_1 \vee \overline{\theta}] > 0. \tag{6}$$

This is the content of Lemma 2. In contrast,

$$\begin{aligned} E[\frac{1}{p_2}|\mathcal{F}_1 \vee \overline{\theta}] &= 3\frac{1}{2} + 5\frac{1}{2} \\ &= 4 \\ &= \frac{1}{p_1}, \end{aligned}$$

i.e.,

$$E[x_2^+|\mathcal{F}_1 \vee \overline{\theta}] = 0, \tag{7}$$

confirming Theorem 3.

Let us now investigate the effect of potential full revelation by studying the following information flow at $t = 3$, conditional on $s_2 = 1$ (i.e., $p_2 = 1/3$). If $\theta = \overline{\theta}$ then:

$$s_3 = 1 \text{ with probability } 1;$$

if $\theta = \underline{\theta}$ then:

$$s_3 = \begin{cases} 1 & \text{with probability } 1/2, \\ 0 & \text{with probability } 1/2. \end{cases}$$

This constitutes the information in $\mathcal{F}_3$ when $s_2 = 1$. Letting $l_3(s_3|\theta)$ denote the conditional likelihood of s_3, and appealing to Bayes' law, we obtain that, if $s_3 = 1$,

$$\begin{aligned} p_3 &= \frac{l_3(s_3 = 1|\overline{\theta})p_2}{l_3(s_3 = 1|\overline{\theta})p_2 + l_3(s_3 = 1|\underline{\theta})(1 - p_2)} \\ &= \frac{1\frac{1}{3}}{1\frac{1}{3} + \frac{1}{2}\frac{2}{3}} \\ &= \frac{1}{2}. \end{aligned}$$

Analogously, if $s_3 = 0$:

$$\begin{aligned} p_3 &= \frac{l_3(s_3 = 0|\overline{\theta})p_2}{l_3(s_3 = 0|\overline{\theta})p_2 + l_3(s_3 = 0|\underline{\theta})(1 - p_2)} \\ &= \frac{0}{0 + \frac{1}{2}\frac{2}{3}} \\ &= 0. \end{aligned}$$

Consequently:

$$\begin{aligned} E[p_3|\{s_2 = 1\} \vee \overline{\theta}] &= \frac{1}{2}1 + 0.0 \\ &= \frac{1}{2} \\ &> \frac{1}{3} \\ &= p_2, \end{aligned}$$

i.e.,

$$E[r_3|\{s_2 = 1\} \vee \overline{\theta}] > 0. \tag{8}$$

Again, this is the content of Lemma 2. Now, however, we have:

$$\begin{aligned} E[\frac{1}{p_3}|\{s_2 = 1\} \vee \overline{\theta}] &= 2.1 + \frac{1}{0}0 \\ &= ? \end{aligned}$$

i.e., $E[x_3^+|\{s_2 = 1\} \vee \overline{\theta}]$ is indeterminate. The full revelation of the state when $s_3 = 0$ complicates matters.

To obtain the result of Theorem 3, one can use the following limit argument. Consider a sequence of economies, indexed $n = 1, 2, 3, ...$ In the nth economy,

$$P\{s_3 = 1|\{s_2 = 1\} \vee \overline{\theta}\} = 1 - \frac{1}{n}.$$

As $n \to \infty$, we obtain the economy described above. In the nth economy, if $s_3 = 1$:

$$\begin{aligned} p_3 &= \frac{(1 - \frac{1}{n})\frac{1}{3}}{(1 - \frac{1}{n})\frac{1}{3} + \frac{1}{2}\frac{2}{3}} \\ &= \frac{1 - \frac{1}{n}}{2 - \frac{1}{n}}, \end{aligned}$$

and if $s_3 = 0$:

$$\begin{aligned} p_3 &= \frac{\frac{1}{n}\frac{1}{3}}{\frac{1}{n}\frac{1}{3} + \frac{1}{2}\frac{2}{3}} \\ &= \frac{\frac{1}{n}}{1 + \frac{1}{n}}. \end{aligned}$$

Hence,

$$
\begin{aligned}
&\lim_{n\to\infty} E[\frac{1}{p_3}|\{s_2 = 1\} \vee \overline{\theta}] \\
&= \lim_{n\to\infty} \left(\frac{2-\frac{1}{n}}{1-\frac{1}{n}}(1-\frac{1}{n}) + \frac{1+\frac{1}{n}}{\frac{1}{n}}\frac{1}{n} \right) \\
&= \lim_{n\to\infty} \left(2 - \frac{1}{n} + 1 + \frac{1}{n} \right) \\
&= 3.
\end{aligned}
$$

Consequently,

$$\lim_{n\to\infty} E[x_3^+|\{s_2 = 1\} \vee \overline{\theta}] = 0, \tag{9}$$

extending the validity of Theorem 3.

While Theorem 3 is not necessarily invalidated when signals are potentially fully revealing, such a case does create difficult *statistical problems*. Heuristically, this can be explained as follows. When $\theta = \overline{\theta}$, one never observes price paths with $p_3 = 0$, which means that x_3^+ will always be finite. As a matter of fact, with unit probability, one observes that $x_3^+ = 1 - (1/3)/(1/2) = 1/3 > 0$.

Mathematically, the statistical problems are created by the nonexistence of moments of order higher than one. Take, for instance, the second moment:

$$
\begin{aligned}
&\lim_{n\to\infty} E[\left(\frac{1}{p_3}\right)^2 |\{s_2 = 1\} \vee \overline{\theta}] \\
&= \lim_{n\to\infty} \left(\left(\frac{2-\frac{1}{n}}{1-\frac{1}{n}}\right)^2 (1-\frac{1}{n}) + \left(\frac{1+\frac{1}{n}}{\frac{1}{n}}\right)^2 \frac{1}{n} \right) \\
&= \lim_{n\to\infty} \left(\frac{(2-\frac{1}{n})^2}{1-\frac{1}{n}} + \frac{(1+\frac{1}{n})^2}{\frac{1}{n}} \right).
\end{aligned}
$$

The latter limit does not exist.

In fact, the above is a nice example of how *unbiasedness* is a weak statistical property. While the (time series) sample average is unbiased, it will generate almost always (i.e., with unit probability) a value that is higher than the theoretical mean. In other words, the sample average is *inconsistent*. This is no surprise given the nonexistence of higher moments.

In the above example, the problem introduced by full revelation before or at T is severe. This is because the probability of full revelation is high. In cases where the probability of full revelation at any point in time is low, statistics such as the sample average are much better estimates of their population counterparts. The illustration in Section 6 is an example.

5 Extension To Consistent Beliefs Equilibria (CBE)

The results in Section 3 have all been derived under the assumption that securities prices are set in a REE. To see why they continue to be valid when we extend the scope of equilibrium to include those where agents have more general (but still consistent) beliefs, we need to analyze the updating rule (Bayes' rule) in a formal way and use the result to express the various conditional expectations in terms of (Lebesgue-Stieltjes) integrals.

Continuing with the assumption of Lemma 2, we assume that agents receive a sequence of (conditionally) independent signals $s_1, s_2, ..., s_T$. Priors about θ are updated with these signals. The updating is rational, i.e., Bayes' rule is used. The equilibrium price reflects Rational Expectations, which means that the prior is *derived from the actual probability measure that generates the outcomes*, P. Factor P into an unconditional μ defined on Θ and a conditional P_θ defined on $\tilde{\Omega}$. μ represents the law Nature uses to draw θ from Θ. Under Rational Expectations, the prior of the agents is given by μ.

Let $l_t(s_t|\theta)$ denote the density of s_t corresponding to P_θ. Agents' beliefs about θ can be written as a sequence of probability measures over Θ, $\{\lambda_t\}_{t=0}^T$, whose elements are recursively defined by:

$$\lambda_0(d\theta) = \mu(d\theta), \tag{10}$$

and

$$\lambda_t(d\theta) = \frac{l_t(s_t|\theta)\lambda_{t-1}(d\theta)}{\int_\Theta l_t(s_t|\theta)\lambda_{t-1}(d\theta)} \tag{11}$$

when $t = 1, ..., T$.

Hence, the time-$(t+1)$ price of the Arrow-Debreu security that pays one dollar if $\theta = \overline{\theta}$, equals:

$$\begin{aligned} p_{t+1} &= E[V|\mathcal{F}_{t+1}] \\ &= E[1_{\{\theta=\overline{\theta}\}}|\mathcal{F}_{t+1}] \\ &= \lambda_{t+1}(d\overline{\theta}) \\ &= \int_\Theta 1_{\{\theta=\overline{\theta}\}}\lambda_{t+1}(d\theta) \\ &= \int_\Theta 1_{\{\theta=\overline{\theta}\}}\frac{l_{t+1}(s_{t+1}|\theta)}{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}\lambda_t(d\theta) \\ &= \frac{l_{t+1}(s_{t+1}|\overline{\theta})\lambda_t(d\overline{\theta})}{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}. \end{aligned} \tag{12}$$

Now that we have expressed prices in terms of integrals, we are ready to prove Theorem 3. As mentioned in the previous section, the proof is simple enough to be included here. It enables us, however, to illustrate the apparent irrelevance of the prior, at the same time pointing out some subtle arguments needed to extend the results of Section 3 to CBE.

Consider p_t/p_{t+1}:

$$\begin{aligned} \frac{p_t}{p_{t+1}} &= \frac{\lambda_t(d\overline{\theta})}{\lambda_t(d\overline{\theta})} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} \\ &= \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})}. \end{aligned} \tag{13}$$

For the empiricist, the relevant expectations condition on $\mathcal{F}_t$ *and* $\overline{\theta}$. These conditional expectations can be computed by integrating with respect to the conditional density $l_{t+1}(s_{t+1}|\overline{\theta})$. If we apply this to (13), $l_{t+1}(s_{t+1}|\overline{\theta})$ will cancel against the denominator of (13). We are left with the density of s_{t+1} conditional on $s_1, ..., s_t$, which integrates out to 1.

To show this formally, let S_{t+1} denote the sampling space of s_{t+1}.[6] Let $P_{s_1,\ldots,s_t}$ denote the probability of s_{t+1} conditional on the history $s_1, ..., s_t$. We have:

$$\begin{aligned} E[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \vee \overline{\theta}] &= \int_{S_{t+1}} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} P_{\overline{\theta}}(ds_{t+1}) && (14) \\ &= \int_{S_{t+1}} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} l_{t+1}(s_{t+1}|\overline{\theta}) ds_{t+1} && (15) \\ &= \int_{S_{t+1}} \int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta) ds_{t+1} \\ &= \int_{S_{t+1}} P_{s_1,\ldots,s_t}(ds_{t+1}) && (16) \\ &= 1. \end{aligned}$$

Noting that

$$x^+_{t+1} = \frac{p_{t+1} - p_t}{p_{t+1}} = 1 - \frac{p_t}{p_{t+1}},$$

Theorem 3 obtains immediately.

This proof reveals an important fact: the nature of the priors is irrelevant. In the above, $P_{s_1,\ldots,s_t}$ is obtained from P. Any other $P^*_{s_1,\ldots,s_t}$ would do, as long as it emerges from a measure P^* on Ω which can be factored into an arbitrary prior λ_0 on Θ and P_θ on $\tilde{\Omega}$ (the same P_θ that emerges from a factorization of P into μ and P_θ). In all cases, the crucial equality

$$\int_{S_{t+1}} P^*_{s_1,\ldots,s_t}(ds_{t+1}) = 1 \tag{17}$$

would hold (see Eqn. (16)).

Whereas the priors are not crucial, the updating rule in Eqn. (11) is. Without it, the cancelations in (13) and (15) would not have obtained. Consequently, *Rational Learning is crucial; Rational Expectations is not.* Theorem 3 continues to hold when

[6] We shall implicitly assume that S_{t+1} is some real (sub)space. $l_t(s_{t+1}|\theta)$ is taken to be the density of P_θ with respect to the Lebesgue measure. It is trivial to extend the analysis to more general cases.

agents have priors that deviate from the "correct" one. The resulting equilibrium will not be REE anymore, but will share with it the rationality of updating. This is essentially the Bayesian Equilibrium with Consistent Beliefs developed in the game-theoretic literature by Harsanyi [1967]. We refer to it as the Consistent Beliefs Equilibrium (CBE).

Actually, we are not done yet. There is a subtle complication which needs to be addressed before the arguments of the proof of Theorem 3 continue to be valid under CBE. It concerns the nature of λ_0 (subsequent beliefs, i.e., λ_t for $t > 0$, can be obtained from λ_0 and the signals, so need not be considered separately). In the world of REE, λ_0 is fixed, being obtained from factorization of the actual probability measure P. In particular, $\lambda_0 = \mu$. In the world of CBE, however, λ_0 will generally be a random variable for the empiricist. Indeed, because priors are arbitrary, it would be unrealistic to assume that she knows agents' priors in any given history.

Therefore, the sampling space ought to be extended to include the space of all probability measures over Θ, to be denoted $\Pi(\Theta)$. The resulting sampling space Ω equals $\Theta \times \tilde{\Omega} \times \Pi(\Theta)$. The information filtration can be as before. Let Q denote the extension of P to the new Ω (and the filtration generated by the signals and prices). Because agents' beliefs at t, λ_t, are not in $\mathcal{F}_t$, they ought to be integrated out in the the right-hand-side of (14) using Q.

Factor Q into a probability measure μ on Θ and another one, Q_θ, on $\tilde{\Omega} \times \Pi(\Theta)$. As before, the signals are mutually (conditionally) independent. We now add to this that they be perceived to be (conditionally) independent of the prior λ_0. Altogether, this implies that $Q_{\overline{\theta}}(ds_{t+1}d\lambda_0) = Q_{\overline{\theta}}(ds_{t+1})Q_{\overline{\theta}}(d\lambda_0)$. Using this, we can adapt Eqns. (14) – (16) to the case of CBE.

$$\begin{aligned}
E[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \vee \overline{\theta}] &= \int_{\Pi(\Theta)} \int_{S_{t+1}} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} Q_{\overline{\theta}}(ds_{t+1}) Q_{\overline{\theta}}(d\lambda_0) \\
&= \int_{\Pi(\Theta)} \int_{S_{t+1}} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} l_{t+1}(s_{t+1}|\overline{\theta}) ds_{t+1} Q_{\overline{\theta}}(d\lambda_0) \\
&= \int_{\Pi(\Theta)} \int_{S_{t+1}} \int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta) ds_{t+1} Q_{\overline{\theta}}(d\lambda_0) \\
&= \int_{\Pi(\Theta)} \int_{S_{t+1}} P^*_{s_1,\ldots,s_t}(ds_{t+1}) Q_{\overline{\theta}}(d\lambda_0) \\
&= \int_{\Pi(\Theta)} Q_{\overline{\theta}}(d\lambda_0) \\
&= 1.
\end{aligned}$$

Noting again that $x^+_{t+1} = 1 - p_t/p_{t+1}$, we have shown the following.

Theorem 7 *Let $\{\mathcal{F}_t\}_{t=1}^T$ be the information filtration generated by the prices $p_1, ..., p_T$ as well as (conditionally) independent signals $s_1, ..., s_T$. Under CBE,*

$$E[x^+_{t+1}|\mathcal{F}_t \vee \overline{\theta}] = 0,$$

where x^+_{t+1} is computed as in Eqn. (4), using prices of an Arrow-Debreu security that pays one dollar when $\theta = \overline{\theta}$.

The importance of this result must be underscored. It implies a simple test of the rationality of learning in financial markets: verify whether x^+ for "surviving" Arrow-Debreu securities forms a martingale difference sequence. Rational Expectations are not required. Market prices may reflect biased priors. Information, however, must be absorbed into prices in a rational way, i.e., conform with the rules of conditional probability (Bayes' rule).

The remaining propositions of the previous section continue to hold under CBE as well. We collect them in the following theorem.

Theorem 8 *Theorems 4 and 6, as well as Corollary 5, continue to hold if REE is relaxed to CBE.*

What is accomplished? Several payoff-related variables continue to exhibit martingale difference characteristics even in the world of Arrow-Debreu securities and even if prices are not set according to Rational Expectations, but in accordance with arbitrary, potentially biased, beliefs. The crucial determinant of the martingale results is not beliefs, but rationality of learning.

6 Illustration

We now turn to a simple simulation exercise. It is meant to illustrate Theorems 3 and 7. In contrast with the theoretical model of the previous section, the signals will not be (conditionally) independent. As mentioned there, Theorem 3 is unaffected by a relaxation of this assumption. The analysis of returns in reverse time (e.g., Theorem 4), however, is invalidated. We will discuss why, and use the simulations to illustrate the violations.

The example is inspired by the analysis in Brown, Goetzmann and Ross [1995]. It deviates from the preceding theoretical analysis in that the signals may perfectly reveal the true state before maturity (remember that we have been assuming that $\theta \not\in \mathcal{F}_t$, all $t \leq T$). As argued in Section 4, the results continue to hold for a potentially perfectly revealing signal. The theoretical properties of statistics constructed on the basis of the simulations, however, become hard to analyze. Nevertheless, the example is included here because of its prominence in the literature. Further simulations revealed that the conclusions remain qualitatively the same if we had chosen a non-perfectly revealing case.

The Model: The signals s_t are taken to be the increments of a Brownian motion with volatility σ. The initial value of the Brownian motion is known and equal to W_0. The signals are really pieces of the sample path of a Brownian motion (functions of time).

Most often, we need only the value at the end. So, s_t will usually refer to this endpoint value. The value of the Brownian motion at t, W_t, can then be obtained as:

$$W_t = W_0 + \sum_{\tau=1}^{t} s_\tau.$$

We define a barrier B, with $W_0 > B$. B is known. Θ is binary, with $\theta = \overline{\theta}$ for those sample paths of the Brownian motion such that $W_t > B$ for all $t \in [0, T^*]$, where $T^* > T$. Otherwise, $\theta = \underline{\theta}$.

We consider the prices of the Arrow-Debreu security that pays \$1 if $\theta = \overline{\theta}$. For $W_t > B$, standard analysis of first-passage-time of Brownian motions reveals that:

$$\begin{aligned} p_t &= E[1_{\{\theta=\overline{\theta}\}}|\mathcal{F}_t] \\ &= 2N(\frac{W_t - B}{\sigma\sqrt{T^* - t}}) - 1, \end{aligned} \tag{18}$$

where N denotes the standard normal distribution function (see, e.g., Karatzas and Shreve [1987], p. 80).

Figure 1: We set $B = 0$ and consider the price paths for which $\theta = \overline{\theta}$. $\sigma = 0.40$, $W_0 = 1$, $T = 200$ and $T^* = 1000$. Figure 1 displays a typical sample path. The upward drift in prices is not very apparent in this figure, but becomes clear when aggregating the results from a large number of simulations. This is done in the subsequent figures.

Figure 2: Figure 2 displays the empirical density of the (time series) average return,

$$\frac{1}{199}\sum_{t=1}^{199} r_{t+1},$$

across 200 simulations. Plotted are: (i) the histogram, (ii) a kernel estimate of the density using a normal kernel, (ii) the normal density evaluated at the sample mean and standard deviation of the outcomes. The upward bias is striking. As a matter of fact, the mean is several standard errors above zero. The mean equals 0.0194; its standard error 0.0006.

Figure 3: In contrast, Figure 3 displays the empirical density of the (time series) average x^+, i.e.,

$$\frac{1}{199}\sum_{t=1}^{199} x_{t+1}^+.$$

According to Theorem 3, this payoff-related variable should be unbiased. Figure 3 confirms that it is. Again the histogram, a kernel estimate of the density and the normal density evaluated at the sample mean and standard deviation are plotted. The mean is within one standard error from zero: it equals 0.0005, whereas its standard error equals 0.0006. (The skewness is reflected in the fact that the median is well above the mean: it equals 0.0034.)

Figures 4 and 5: Theorem 3 not only claims that the average x^+ would be unbiased, but also that x^+ cannot be predicted from past information. To verify this, we project x^+_{t+1} onto the beginning-of-period price level, p_t. This is a time series projection, i.e., the data run from $t = 1$ to $t = T - 1$ ($T = 50$). (There is a reason why this particular projection was chosen; it will be discussed later.) Figures 4 and 5 display the empirical densities (histogram, kernel estimate and normal density) for the OLS intercept and slope coefficient, respectively.

The figures illustrate Theorem 3: both the intercept and slope coefficient have densities that are concentrated around zero. Since x^+_{t+1} is regressed on an endogenous variable (p_t), neither the OLS intercept estimator nor the OLS slope estimator will be unbiased. Indeed, in both cases, the mean of the estimates is more than two standard errors away from zero.[7] Nevertheless, the densities have substantial mass both below and above zero. The intercept estimates are anywhere between -0.3253 and 0.1483; the slope estimates are between -0.6428 and 0.6665.

Figures 6 and 7: Contrast this with the empirical densities of the intercept and slope coefficient in a time series OLS projection of the return, r_{t+1}, onto the beginning-of-period price, p_t. Figures 6 and 7 display the result. The estimates of the intercept are *always positive*: they are between 0.0329 and 0.3598. Those of the slope are *always negative*: between -1.4995 and -0.501. The explanation for the negative slope coefficient is simple: if initial signals convey that the security is going to expire worthless, the price quickly drops to a very low level; since the security eventually does mature in the money, prices are very likely to increase subsequently; conversely, if positive signals reach the market early on, prices gradually increase to a high level; subsequent signals most often only confirm this price runup, hence, returns are more likely to be low. In other words: returns are negatively correlated with the price level.

The choice of projection can now be justified. It has been documented that the returns on common stock are indeed inversely related to the price level. Among the first to notice this were Keim and Stambaugh [1986]. It reappears, disguised, in the success of contrarian strategies (e.g., Jegadeesh [1990]), and, in the long run, in the (albeit controversial) literature on long-run mean reversion in stock prices (De Bondt and Thaler [1985]). There has been some theoretical work trying to explain this in a dynamic, stationary asset pricing framework (Bossaerts and Green [1989], Berk [1995]). Recently, it has been suggested that bankruptcy may explain some of the phenomenon (Brown, Goetzmann and Ross [1995]). The results in Figures 6 and 7 support this view. As mentioned before, Arrow-Debreu securities reflect the essence of bankruptcy: there is a state ($\underline{\theta}$) where the security pays zero – this is the bankruptcy state; in the other state ($\overline{\theta}$), the security pays one dollar – this is the no-bankruptcy state.

Figure 8: Throughout this example, signals are increments of a Brownian motion. Hence, they are independent. Unfortunately, they are not independent *conditional on* θ.

[7] When increasing T to 500 and T^* to 5000, the biases disappeared: the mean estimates of the intercept and slope became insignificant.

Given $\overline{\theta}$, for instance, the distribution of s_t depends on the accumulation of past signals, i.e., W_{t-1}.[8] The theorems in the previous sections all assume conditional independence. It was pointed out that the results concerning forward time series analysis continue to hold if the signals are dependent, but that the reverse-time martingale difference behavior would not obtain.

It is easy to see why. Section 5 revealed that the theorems rely on a cancelation of two density functions: (i) the density used by the market to evaluate the likelihood of observing a signal conditional on a given θ, and (ii) the density used by the outsider to evaluate the likelihood of the same signal. Forward in time, the two densities always coincide if conditioned on the same θ. In reverse time, however, the two densities may not coincide even if conditioned on the same θ. If the signals are dependent over time, the market will use lagged signals as additional conditioning information to assess the likelihood of the next signal. This means that the density it uses will depend not only on θ, but also on lagged signals. When analyzing return data in reverse time, however, the empiricist cannot condition on these signals, because they do not occur prior (in reverse time) to the return. In fact, the empiricist evaluates the likelihood of a signal by incorporating signals that occur in the *future* when dated forward in time. Hence, the density she uses differs from that used by the market, and the cancelations that were so crucial to the theorems will not obtain. There is only one exception: the signals are (conditionally) independent. In that case, neither the market nor the empiricist conditions on lagged (forward or backward in time) signals. Hence, the likelihoods of signal outcomes coincide and the cancelations can be carried out.

There is another source of dependence in the simulations.[9] Forward in time, all price paths start at the conditional probability of a nonzero payoff given $W_0 = 1$. One thereby introduces a selection bias in reverse time: after 200 periods, all price paths revert back to the same level. Hence, returns will not be independent.

To illustrate how lack of independence invalidates the theorems involving return analysis in reverse time, Figure 8 plots the densities (histogram, kernel estimate and normal curve) of the returns on the Arrow-Debreu security that pays one dollar in the alternative state, i.e., $\underline{\theta}$. Returns are computed conditional on the true state being $\overline{\theta}$. According to Corollary 5[10], the time series average return should be zero. Figure 8 illustrates that this is not the case. In fact, its mean is several standard errors below zero: it equals -0.0032; its standard error is 0.0001.

REE vs. CBE: In the simulations used to produce the figures, Nature (the computer) *always* set $\theta = \overline{\theta}$. In other words, Nature did not obtain the signals from the

[8]See Brown, Goetzmann and Ross [1995] for an extensive analysis.

[9]Oleg Bondarenko quite rightly pointed this out.

[10]Corollary 5 is stated in terms of the Arrow-Debreu security that pays one dollar if $\theta = \overline{\theta}$ and the returns are observed when the true state is $\underline{\theta}$. Its validity obviously is unaffected by a reversal of the roles of $\underline{\theta}$ and $\overline{\theta}$.

solutions to the following stochastic differential equation:

$$dW_t = \sigma dZ_t, \tag{19}$$

where dZ_t denote the instantaneous increments of a Wiener process. Instead, the signals were generated by:

$$dW_t = \frac{2\sigma n(\frac{W_t}{\sigma\sqrt{T^*-t}})}{\sqrt{T^*-t}(2N(\frac{W_t}{\sigma\sqrt{T^*-t}})-1)}dt + \sigma dZ_t. \tag{20}$$

(n denotes the standard normal density.) The solutions to the stochastic differential equation in (20) provide sample paths that do not cross the barrier $B = 0$ before T^*.[11]

Investors did use Eqn. (19) to update their beliefs about θ. It differs from the process truely used by Nature, Eqn. (20). Consequently, *investors did not make decisions on the basis of Rational Expectations.* Nevertheless, their *learning was rational.* The simulations therefore illustrate Theorems 7 and 8 instead of the Theorems of Section 3. They confirm that the results of this paper do not depend on REE. Instead, all what is required is the Rational Learning of a CBE.

7 Equity

One wonders how the martingale characterizations would apply to assets other than Arrow-Debreu securities. Of course, any asset can be considered to be a portfolio of such securities, but, since the martingale restrictions concern only specific Arrow-Debreu securities, they will generally not obtain for arbitrary portfolios of these primitive securities.

It is straightforward, however, to extend the results to assets whose payoff conditional on the fixed parameter (θ) is independent of the signals. As before, let V denote the security's (random) payoff at maturity (T) and $g(V|\overline{\theta})$ its density conditional on $\theta = \overline{\theta}$. Assume that the security's payoff is zero for all other values of θ. It now suffices to substitute $\int Vg(V|\overline{\theta})dV 1_{\{\theta=\overline{\theta}\}}$ for $1_{\{\theta=\overline{\theta}\}}$ and the proofs continue to hold.

As far as the analysis *forward* in time is concerned, it is even possible to obtain clean results for the case where V depends on the sequence of signals as well. This would mean that the conditional density, g, is a function of past signals, in addition to $\overline{\theta}$. The following is proved in the Appendix.

Theorem 9 *Let $\{\mathcal{F}_t\}_{t=1}^T$ be the information filtration generated by the prices $p_1, ..., p_T$ as well as (conditionally) independent signals $s_1, ..., s_T$. Under CBE,*

$$E[x_{t+1}^+|\mathcal{F}_t \vee \overline{\theta}] \le 0,$$

[11]See, e.g., Brown, Goetzmann and Ross [1995]. Actually, the simulations were based on Euler discretization of (19), whereby $(t, t+1)$ was split in ten subintervals of equal length; subsequently, the sample paths for which the (discretized) Brownian motion crossed the barrier were discarded.

where x^+_{t+1} is computed as in Eqn. (4), using prices of a security that pays a random amount V when $\theta = \overline{\theta}$ and zero otherwise; V and the signals need not be independent.

(The theorem is stated in terms of CBE; hence, it holds *a fortiori* for REE as well.) In words: the asset's payoff, normalized by the *end-of-period* price, must not have a positive drift.

Theorem 9 covers an important case: equity. The eventual payoff on equity ("liquidation value") depends on whether the firm has gone bankrupt or not.[12] Absent bankruptcy, the eventual payoff is correlated with the signals that the market receives about the likelihood of bankruptcy. Brown, Goetzmann and Ross [1995] argue that bankruptcy could explain a host of asset pricing anomalies. Many would object to this claim. When samples are truncated because of survival (we observe only the return histories of non-bankrupt firms), it is impossible to forcefully argue in favor of Brown, Goetzmann and Ross' conjecture, even under their (implicit) assumption of REE. We shall discuss how our statistics can shed light on this debate. To better understand the contribution, however, it is necessary to first study what happens if bankruptcy is indeed irrelevant.

The presence of bankruptcy is reflected in the conditioning on θ. If bankruptcy is irrelevant, the empiricist's conditioning on θ is ill advised. Investors do not use the signals to update their beliefs about whether equity will have a positive payoff; signals are only there to indicate the size of the likely payoff. Under REE, we have, of course, that

$$E[r_{t+1}|\mathcal{F}_t \vee \overline{\theta}] = 0. \tag{21}$$

(The formal conditioning on $\theta = \overline{\theta}$ is continued, despite its irrelevance.) What about the modified return, x^+_{t+1}? A simple application of Jensen's inequality reveals:

$$\begin{aligned} & E[x_{t+1}|\mathcal{F}_t \vee \overline{\theta}] \\ = \;& E[\frac{p_{t+1} - p_t}{p_t}\frac{p_t}{p_{t+1}}|\mathcal{F}_t \vee \overline{\theta}] \\ = \;& 1 - E[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \vee \overline{\theta}] \\ \leq \;& 1 - \frac{1}{E[\frac{p_{t+1}}{p_t}|\mathcal{F}_t \vee \overline{\theta}]} \\ = \;& 0. \end{aligned}$$

In short,

$$E[x_{t+1}|\mathcal{F}_t \vee \overline{\theta}] \leq 0, \tag{22}$$

as in Theorem 9. In other words, the theorem continues to hold even if the empiricist falsely assumes that conditioning on θ is relevant, i.e., that there is a selection bias.

This fact allows one to investigate the importance of bankrupcty in the pricing of equity under the maintained assumption of REE. If Eqn. (22) holds, equity prices are set

[12] As before, we assume that bankruptcy leads to total loss of value. This assumption is standard in capital structure theory. For empirical evidence, see Clark and Weinstein [1983].

rationally, whether there is anything (bankruptcy) to be learned or not. If, in addition, Eqn. (21) holds, learning is indeed irrelevant, and (22) obtains as a mere consequence of Jensen's inequality. Alternatively, if Eqn. (21) fails, (22) obtains precisely because agents are learning rationally about bankruptcy.

Consequently, the tests in this paper provide a means with which to identify the relevance of bankruptcy in a REE context. It uses rationality of updating in a REE to generate the identifying information. Absent this *economic* restriction, there is no way to distinguish the two cases. For every *statistical* model of the data which accounts for potential bankruptcy, there is another one where bankruptcy does not occur and which is observationally equivalent on strictly positive price paths (i.e., the return histories for the non-bankrupt firms obey the same law). An example is the Brownian motion in (19) subject to an absorbing barrier at zero. For strictly positive price histories, it is observationally equivalent to the model in (20), where the price never hits zero.[13]

Nevertheless, the contribution of the present paper goes beyond this: it can shed light on the nature of price setting *even if* there is a cross-section of price histories for both securities that mature in the money and those that expire out of the money. We will come back to this point in the next section.

In the meanwhile, let us look at Figure 9, which illustrates Theorem 9. It displays estimated densities (histogram, kernel estimate, normal curve) for the time series average x^+ for a case that is identical to that of the previous section, with the exception that the security pays W_{T^*} at maturity (instead of \$1) when $\theta = \overline{\theta}$.[14] The mean of the density (0.0007) is again within one standard error (0.0008) of zero. As a matter of fact, there is not even a noticeable negative bias (Theorem 9 allows the time series average x^+ to be negative).[15] In contrast, the time series average of the traditional return (not reported) was positive in *every* price path.

Other extensions of the martingale results of this paper are worth exploring. At this point, it suffices to add that no reverse-time equivalent to Theorem 9 obtains, because the empiricist conditions on future signals to evaluate the likelihood of past returns, whereas the market conditions on past signals to set prices. This is the same reason why reverse-time martingale results fail to obtain when the signals are correlated (see Section 6).

[13]It may be worth pointing out here that the survivorship bias in itself does not inhibit *estimation*. It is well known, for instance, how to estimate the parameters of a Markov chain with absorbing barrier (the absorbing barrier being the equivalent of the survival determinant). See, e.g., Basawa and Rao [1980], p.68.

[14]The parameters remained as in Section 5. It can be shown that $p_t = W_t$ (instead of Eqn. (18).

[15]The upper bound (zero) on the time series average x^+ implied by Theorem 9 is tight if the period-to-period variability in the expectation of V is low. Details can be found in the proof of the theorem in the Appendix.

8 Relevance For Cross-Sectional Analysis

The importance of the results in this paper for the analysis of single time series requires little discussion. It may not be readily transparent, however, that the same results can be a valuable tool in the analysis of cross-sections of return histories which are not subject to any selection bias.

It was shown in Section 5 that martingale difference behavior continued to obtain in a more general environment than that of REE, namely, CBE. Intuitively, it was simple to see why this was: the restrictions followed only because of the rationality of learning (the use of Bayes' law), and not because of the unbiasedness of initial beliefs in a REE.

Because the results are not affected by biases in initial beliefs, they can be used to distinguish REE from CBE where priors are biased. Under REE, expected returns should not be predictable in any cross-section that is not subject to a systematic selection bias. In that case, the average return observed by the empiricist is the one computed in Lemma 1 (the absence of sample selection bias is reflected in the fact that θ is integrated out). Absent REE, the cross-sectional average of the historical average return will be nonzero. If, however, REE fails to hold just because initial beliefs were biased, but agents where otherwise rational (correctly applied the rules of conditional probability), the other theorems can still be appealed to. For instance, the price histories for assets that matured in the money will still satisfy Theorem 3: the cross-sectional average of the historical mean modified return will be zero.

To illustrate this, consider a sampling of Arrow-Debreu price histories from the Iowa Experimental Markets (IEM): those corresponding to Arrow-Debreu securities that pay $1 when one of Apple, IBM, Microsoft and the S&P500 has the highest return over a one-month period, and those corresponding to whether the stock price of Microsoft went up or down in the same period. Take six months worth of daily closing prices, spanning the period October 1995–March 1996. In total, one obtains a cross-section of 4×6 plus 2×6 time series, of which twelve are for Arrow-Debreu securities that matured in-the-money.

The mean time series average return for the winning Arrow-Debreu securities is 0.1104. With a standard error of 0.0423, this mean is significantly positive, confirming Lemma 2. The mean time series average modified return (x^+) for the winning Arrow-Debreu securities equals -0.0171, which is insignificant in view of the standard error of 0.0183. Hence, Theorem 3 is verified, confirming the rationality of learning. Across all histories of winning *and loosing* markets, however, the mean time series average return was 0.0488, which is significantly positive because of a standard error of 0.0249.[16] Consequently, REE is rejected.

[16] Actually, the estimate of the standard error is biased upward: it ignores the strong negative correlation between return histories of complementary Arrow-Debreu securities such as the two securities deriving their value from the price change of Apple stock (one of them necessarily ends up out of the money; the other will expire in the money).

9 Open Questions

The main contribution of this paper lies in the martingale restrictions on payoff-related variables for Arrow-Debreu securities and their robustness to extensions of REE to equilibria with less controversial beliefs structures. The results rely almost exclusively on the rationality of belief updating, i.e., on the assumption that investors follow the rules of conditional probability in processing payoff-relevant information. Some remarks are in order.

Remark 1: It should be emphasized here that the robustness of the martingale results to initial beliefs is unique in the learning literature. The controversy surrounding Bayesian statistical analysis, for instance, can mainly be attributed to the sensitivity of the inference to the specification of priors.[17] In contrast, we have here a set of restrictions on security price data which are robust to investors' initial beliefs.

Remark 2: We have been assuming risk neutrality and zero discount rates. The latter can easily be relaxed. But risk aversion could now also be reintroduced. Any deviation from martingale behavior can then be analyzed as *risk premia*, in analogy with the literature on asset pricing under REE in a static, or a stationary, dynamic environment. This means that the results of this paper provide a starting point to build tests of standard asset pricing effects in an environment where there is learning, and where the learning may start from arbitrary beliefs.

Remark 3: About statistical inference. In the illustrations of Sections 6 and 7, the behavior of the time-series average payoff-related variables (such as returns) was investigated, as well as that of the parameter estimates in time series OLS projections onto lagged information. The martingale difference characterization of some of these variables would suggest that one could appeal to a martingale Central Limit Theorem (CLT) and use z-statistics and the standard normal distribution for testing in a given time series.

Martingale CLTs, however, do not readily obtain. For one thing, convergence of beliefs may make the asymptotic standard deviation of the (scaled) estimated parameters zero. In the illustrations of the previous section, one could avoid this by increasing the difference between T (maturity of the security) and T^* (the state is determined by whether W_t crosses the boundary B before $T^* > T$) as $T \to \infty$. Even so, the asymptotic standard deviation may be sample-dependent. In principle, this is not a problem. Martingale CLTs could be appealed to which allow the asymptotic standard deviation to vary across sample paths (see, e.g., Hall and Heyde [1980], p. 58).

Unfortunately, extensive simulations along the lines of those reported in the previous section revealed that either CLTs failed to obtain or the moderately long time series length was insufficient for CLT behavior to kick in. The reason behind the non-normal large-sample behavior of the z-statistics seems to be the dependence between the parameter's

[17] See, e.g., the arguments about Bayesian analysis of nonstationary time series: Sims and Uhlig [1991], Phillips [1990].

(scaled) sample value and its estimated standard deviation. Whether this is a small-sample result or an asymptotic property of the learning environment we are studying is not clear. In any event, suffice it to say that general martingale CLTs do require the (scaled) parameter estimate to be mean-independent of the asymptotic standard deviation.

This can easily be illustrated. Figure 10 displays density estimates (histogram, kernel estimate, normal curve) for the z-statistics computed from the time series average x^+ for an Arrow-Debreu security that paid one dollar if $\theta = \overline{\theta}$ (the true state was $\overline{\theta}$). The deviation from the standard normal density is not spectacular. Most pronounced, however, is the bias: the mean equals 0.4181 instead of zero. Less apparent is the lower standard deviation: 0.8794 instead of 1. Figure 11 plots the (scaled) sample average against the estimate of its asymptotic standard deviation. The scaled average return equals:

$$\frac{1}{\sqrt{T-1}} \sum_{t=1}^{T-1} x_{t+1}^{+},$$

whereas the estimate of the asymptotic standard deviation was obtained as the square root of:

$$\frac{1}{T-1} \sum_{t=1}^{T-1} (x_{t+1}^{+})^2.$$

A significant negative (linear) relationship appears, which caused the positive bias (relative to the standard normal) in the z-statistic shown in Figure 10.

The negative relationship is intriguing. A linear, negative relation between scaled average returns and estimated asymptotic standard deviation has been observed elsewhere. Bossaerts [1995b] documents this for experimental financial markets and points out that the finding invalidates the conclusions regarding market efficiency that one usually draws from the z-statistics.

If the results are an indication of the failure of standard CLTs, the surprisingly regular, linear relation between the scaled sample average and its asymptotic standard deviation, seems to suggest that a simple mixture-of-normals invariance result (nonergodic martingale CLT) may hold. We leave this question for future research.[18,19]

To put things into perspective, however, the deviations from standard normality in Figure 10 should be compared to the dramatic failure of the z-statistic computed from

[18] Examples of analytical asymptotics under learning can already be found in Bossaerts [1995a]. In two particular cases where beliefs ultimately converged, this article derived the asymptotic distribution of the z-statistic computed from average returns (among other things). In its second case, the asymptotic value of the scaled average return was negatively correlated with its asymptotic standard deviation, generating substantial deviations from asymptotic normality. The negative correlation reappears here (see Figure 11).

[19] Phillips [1990] discusses the bimodality of the asymptotic distribution of statistics where the numerator and denominator are correlated. See his p. 56. His paper is particularly relevant for the example here, because of the presence of infinite second moments – see Section 4.

the time series average return (as opposed to x^+). Of course, we do not expect a CLT to hold for this statistic, because returns do not constitute a martingale difference sequence (Lemma 2). Figure 12 displays estimates of the density of the z-statistic. Not only is the nonnormality (skewness) clear, the mean is now as high as 1.9657 and the standard deviation only 0.3522.

10 Conclusion

This paper proposes modifications to the traditional return measure that make it a martingale difference sequence (or a reverse-time martingale difference sequence) when a key determinant of the payoff of an asset remains constant over time and investors gradually learn about this determinant. The results are not only relevant to time series analysis subject to selection bias (e.g., survivorship bias). Since the restrictions continue to hold if investors' initial beliefs are biased, i.e., prices are not set in a REE, but a CBE, they can be used to test REE against CBE in cross-sections that are free from any selection bias.

Hitherto, tests of market efficiency and asset pricing have exclusively been based on REE. The appeal of REE as an equilibrium concept should be attributed to its emphasizing learning (from signals and prices). CBE shares this emphasis, without requiring priors to be unbiased at the outset. This paper demonstrates that allowing priors to be biased does not necessarily lead to a situation where "everything is possible." Robust restrictions still obtain solely as a result of rationality of learning.

The results of this paper therefore should shed new light on the debate about the sources of predictability in asset return data. Within REE, all predictability reflects time-varying risk premia. In CBE, predictability partly reveals investors' learning payoff determinants. The martingale difference results obtained here should allow empiricists to disentangle the two sources of predictability.

There are other means by which to disentangle Bayesian learning and risk premia, based on differences between in-sample and out-of-sample predictability. Bossaerts and Hillion [1996] prove propositions that provide precise discrimination. Using a cross-section of stock index return histories from fifteen countries, they fail to reject that all predictability that one can generate using linear prediction models is caused by learning.

References

Aumann, R.J., 1976, "Agreeing to Disagree," *Annals of Statistics* 4, 1236-39.

Basawa, I.V. and P. Rao, 1980, *Statistical Inference for Stochastic Processes.* New York: Academic Press.

Berk, J., 1995, "A Critique of Size-Related Anomalies," *Review of Financial Studies* 8, 275-286.

Biais, B. and P. Bossaerts, 1995, "Asset Prices and Trading Volume in a Beauty Contest," Caltech working paper.

Bossaerts, P., 1995a, "The Econometrics of Learning in Financial Markets," *Econometric Theory* 11, 151-189.

Bossaerts, P., 1995b, "Rational Price Discovery in Experimental and Field Data," Caltech working paper.

Bossaerts, P. and R. Green, 1989, "A General Equilibrium Model of Changing Risk Premia: Theory and Tests," *Review of Financial Studies* 2, 467-493.

Bossaerts, P. and P. Hillion, 1996, "Implementing Statistical Criteria To Select Return Forecasting Models: What Do We Learn?" Caltech working paper.

Brown, S., W. Goetzmann and S. Ross, 1995, "Survival," *Journal of Finance* 50, 853-874.

Clark, T. and M. Weinstein, 1983, "The Behavior of the Common Stock of Bankrupt Firms," *Journal of Finance* 38, 489-504.

De Bondt, W., and R.H. Thaler, 1985, "Does the Stock Market Overreact?" *Journal of Finance* 40, 793-805.

Doob, J.L. (1952): *Stochastic Processes.* New York: Wiley.

Fama, E., 1970, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance* 25, 383-417.

Green, J., 1973, "Information, Efficiency and Equilibrium," Harvard Institute of Economic Research, Discussion Paper 284.

Hall, P. and C.C. Heyde, 1980, *Martingale Limit Theory and Its Application.* New York: Academic Press.

Harrison, J.M. and D.M. Kreps, 1978, "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics* 92, 323-36.

Harsanyi, J.C., 1967, "Games with Incomplete Information Played by 'Bayesian' Players, I: The Basic Model," *Management Science* 14, 159-82.

Jegadeesh, N., 1990, "Evidence of Predictable Behavior of Security Returns," *Journal of Finance* 45, 881-898.

Karatzas, I. and S. Shreve, 1987, *Brownian Motion and Stochastic Calculus.* New York: Springer Verlag.

Keim, D. and R. Stambaugh, 1986, "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics* 17, 357-391.

Kurz, M., 1994, "On the Structure and Diversity of Rational Beliefs," *Economic Theory* 4, 1-24.

Lucas, R., 1972, "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-124.

Muth, J.F., 1961, "Rational Expectations and the Theory of Price Movements," *Econometrica* 29, 315-35.

Phillips, P.C.B., 1990, "Time Series Regression With A Unit Root And Infinite-Variance Errors," *Econometric Theory* 6, 44-62.

Phillips, P.C.B., 1990, "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends," Cowles Foundation Discussion Paper.

Radner, R., 1967, "Equilibre des marchés à terme et au comptant en cas d' incertitude," *Cahiers du Séminaire d' Econométrie*, 35-52.

Samuelson, P. A., 1965, "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review* 6, 41–50.

Tirole, J., 1982, "On the Possibility of Speculation under Rational Expectations," *Econometrica* 50, 1163-81.

Sims, C. and H. Uhlig, 1991, "Understanding Unit Rooters: A Helicopter Tour," *Econometrica* 59, 1591-1599.

Appendix

Proof of Lemma 1

$$
\begin{aligned}
& E[r_{t+1}|\mathcal{F}_t] \\
&= E[\frac{p_{t+1}-p_t}{p_t}|\mathcal{F}_t] \\
&= \frac{1}{p_t}E[p_{t+1}|\mathcal{F}_t] - 1 \\
&= \frac{1}{p_t}E[E[V|\mathcal{F}_{t+1}]|\mathcal{F}_t] - 1 \\
&= \frac{1}{p_t}E[V|\mathcal{F}_t] - 1 \\
&= 0.
\end{aligned}
$$

Proof of Lemma 2

From Eqn. (12),

$$\frac{p_{t+1}}{p_t} = \frac{\lambda_t(d\overline{\theta})}{\lambda_t(d\overline{\theta})} \frac{l_{t+1}(s_{t+1}|\overline{\theta})}{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}.$$

Hence,

$$\begin{aligned}
&E[\frac{p_{t+1}}{p_t}|\mathcal{F}_t \vee \overline{\theta}] \\
&= \int_{S_{t+1}} \frac{l_{t+1}(s_{t+1}|\overline{\theta})}{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)} P_{\overline{\theta}}(ds_{t+1}) \\
&= \int_{S_{t+1}} \frac{1}{\frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})}} l_{t+1}(s_{t+1}|\overline{\theta}) ds_{t+1} \\
&> \frac{1}{\int_{S_{t+1}} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} l_{t+1}(s_{t+1}|\overline{\theta}) ds_{t+1}} \\
&= \frac{1}{\int_{S_{t+1}} \int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta) ds_{t+1}}.
\end{aligned}$$

From Eqns. (15) and (16), it follows that the last expression equals 1. The result immediately obtains:

$$\begin{aligned}
E[r_{t+1}|\mathcal{F}_t \vee \overline{\theta}] &= E[\frac{p_{t+1}}{p_t}|\mathcal{F}_t \vee \overline{\theta}] - 1 \\
&> 0.
\end{aligned}$$

Proof of Theorem 3

See Section 5 (Eqns. (10) to (16)).

Proof of Theorem 4

Define:

$$\tilde{s}_n = s_{T-n+1},$$
$$\tilde{S}_n = S_{T-n+1},$$
$$\tilde{\lambda}_n = \tilde{\lambda}_{T-n+1}.$$

First find the evolution of beliefs in reverse time. Since

$$\tilde{\lambda}_{n-1}(d\overline{\theta}) = \frac{l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})\tilde{\lambda}_n(d\overline{\theta})}{l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})\tilde{\lambda}_n(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_n(d\underline{\theta})},$$

we can solve for $\tilde{\lambda}_n$ as a function of $\tilde{\lambda}_{n-1}$:

$$\tilde{\lambda}_n(d\overline{\theta}) = \frac{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta})}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})(1-\tilde{\lambda}_{n-1}(d\overline{\theta}))}.$$

Consequently,

$$p_{T-n+2} = \tilde{\lambda}_{n-1}(d\overline{\theta}))$$

and

$$\begin{aligned} p_{T-n+1} &= \tilde{\lambda}_n(d\overline{\theta})) \\ &= \frac{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta})}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})(1-\tilde{\lambda}_{n-1}(d\overline{\theta}))}, \end{aligned}$$

i.e.,

$$\begin{aligned} &\frac{p_{T-n+2}}{p_{T-n+1}} \\ &= \frac{\tilde{\lambda}_{n-1}(d\overline{\theta})}{\tilde{\lambda}_{n-1}(d\overline{\theta})} \frac{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})(1-\tilde{\lambda}_{n-1}(d\overline{\theta}))}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})} \\ &= \frac{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})(1-\tilde{\lambda}_{n-1}(d\overline{\theta}))}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})}. \end{aligned}$$

As mentioned in Section 3, $\tilde{s}_{n-1}$ is not in the empiricist's information set at time $(n-1)$, $\mathcal{F}^-_{n-1}$; it does not become available until time n. This is important in the evaluation of the following expectation.

$$\begin{aligned} &E[\frac{p_{T-n+2}}{p_{T-n+1}}|\mathcal{F}^-_{n-1} \vee \underline{\theta}] \\ &= \int_{\tilde{S}_{n-1}} \frac{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})(1-\tilde{\lambda}_{n-1}(d\overline{\theta}))}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})} l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})d\tilde{s}_{n-1} \\ &= \tilde{\lambda}_{n-1}(d\overline{\theta}) \int_{\tilde{S}_{n-1}} l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})d\tilde{s}_{n-1} \\ &\quad + (1-\tilde{\lambda}_{n-1}(d\overline{\theta})) \int_{\tilde{S}_{n-1}} l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})d\tilde{s}_{n-1} \\ &= 1. \end{aligned}$$

The results obtains from:

$$\begin{aligned} E[x^-_n|\mathcal{F}^-_{n-1} \vee \underline{\theta}] &= E[\frac{p_{T-n+2}}{p_{T-n+1}}|\mathcal{F}^-_{n-1} \vee \underline{\theta}] - 1 \\ &= 0. \end{aligned}$$

Proof of Corollary 5

This follows from Theorem 4 by the law of iterated expectations.

Proof of Theorem 6

From the proof of Theorem 4,

$$\begin{aligned}
&E[\frac{p_{T-n+2}}{p_{T-n+1}}|\mathcal{F}_{n-1}^{-} \vee \overline{\theta}] \\
&= \tilde{\lambda}_{n-1}(d\overline{\theta}) \int_{\tilde{S}_{n-1}} l_{n-1}(\tilde{s}_{n-1}|\overline{\theta}) d\tilde{s}_{n-1} \\
&\quad + (1-\tilde{\lambda}_{n-1}(d\overline{\theta})) \int_{\tilde{S}_{n-1}} \frac{l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})} l_{n-1}(\tilde{s}_{n-1}|\overline{\theta}) d\tilde{s}_{n-1} \\
&> \tilde{\lambda}_{n-1}(d\overline{\theta}) \\
&\quad + (1-\tilde{\lambda}_{n-1}(d\overline{\theta})) \frac{1}{\int_{\tilde{S}_{n-1}} l_{n-1}(\tilde{s}_{n-1}|\underline{\theta}) d\tilde{s}_{n-1}} \\
&= 1.
\end{aligned}$$

The result follows from $x_n^- = p_{T-n+2}/p_{T-n+1} - 1$.

Proof of Theorem 7

See Section 5, Eqn. (18).

Proof of Theorem 8

We shall prove the extension of Theorem 4 to CBE; the proofs of the extension of the other results work analogously. Q denotes the measure that the empiricist uses to evaluate the uncertainty about θ, signals, and, most importantly, $\tilde{\lambda}_1$. We get (for $n > 1$):

$$\begin{aligned}
&E[\frac{p_{T-n+2}}{p_{T-n+1}}|\mathcal{F}_{n-1}^{-} \vee \underline{\theta}] \\
&= \int_{\Pi(\Theta)} \int_{\tilde{S}_{n-1}} \frac{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})\tilde{\lambda}_{n-1}(d\overline{\theta}) + l_{n-1}(\tilde{s}_{n-1}|\overline{\theta})(1-\tilde{\lambda}_{n-1}(d\overline{\theta}))}{l_{n-1}(\tilde{s}_{n-1}|\underline{\theta})} \\
&\quad l_{n-1}(\tilde{s}_{n-1}|\underline{\theta}) d\tilde{s}_{n-1} Q_{\underline{\theta}}(d\tilde{\lambda}_1) \\
&= \int_{\Pi(\Theta)} Q_{\underline{\theta}}(d\tilde{\lambda}_1) \\
&= 1.
\end{aligned}$$

The result follows from $x_n^- = p_{T-n+2}/p_{T-n+1} - 1$.

Proof of Theorem 9

To understand the proof, first note that we needed to define V only for $\theta = \overline{\theta}$ (because the security pays zero in all other states). We can extend V arbitrarily for other θ.

We shall take this extension to be such that V has the same (conditional) distribution, independent of θ. This implies, in particular:

$$\int_R V g(V|s_1,...,s_t,\theta)dV = \int_R V g(V|s_1,...,s_t)dV. \tag{23}$$

Now consider p_{t+1}:

$$\begin{aligned}
&p_{t+1}\\
&= \int_\Theta (\int_R V g(V|s_1,...,s_t,s_{t+1})dV) 1_{\{\theta=\overline{\theta}\}} \lambda_{t+1}(d\theta)\\
&= (\int_R V g(V|s_1,...,s_t,s_{t+1})dV)\lambda_{t+1}(d\overline{\theta})\\
&= \int_R V g(V|s_1,...,s_t,s_{t+1})dV \frac{l_{t+1}(s_{t+1}|\overline{\theta})}{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}\lambda_t(d\overline{\theta}).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{p_t}{p_{t+1}}\\
&= \frac{\int_R V g(V|s_1,...,s_t)dV}{\int_R V g(V|s_1,...,s_t,s_{t+1})dV} \frac{\lambda_t(d\overline{\theta})}{\lambda_t(d\overline{\theta})} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})}\\
&= \frac{\int_R V g(V|s_1,...,s_t)dV}{\int_R V g(V|s_1,...,s_t,s_{t+1})dV} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})}.
\end{aligned}$$

We shall need an implication of the law of iterated expectations (and Eqn. (23)):

$$\begin{aligned}
&\int_{S_{t+1}} (\int_R V g(V|s_1,...,s_t,s_{t+1})dV) l_{t+1}(s_{t+1}|\theta)ds_{t+1}\\
&= \int_{S_{t+1}} (\int_R V g(V|s_1,...,s_t,s_{t+1},\theta)dV) l_{t+1}(s_{t+1}|\theta)ds_{t+1}\\
&= E[E[V|\mathcal{F}_{t+1} \vee \theta]|\mathcal{F}_t \vee \theta]\\
&= E[V|\mathcal{F}_t \vee \theta]\\
&= \int_R V g(V|s_1,...,s_t,\theta)dV\\
&= \int_R V g(V|s_1,...,s_t)dV.
\end{aligned}$$

This is used to prove the following.

$$\begin{aligned}
&E[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \vee \overline{\theta}]\\
&= \int_{\Pi(\Theta)} \int_{S_{t+1}} \frac{\int_R V g(V|s_1,...,s_t)dV}{\int_R V g(V|s_1,...,s_t,s_{t+1})dV} \frac{\int_\Theta l_{t+1}(s_{t+1}|\theta)\lambda_t(d\theta)}{l_{t+1}(s_{t+1}|\overline{\theta})} l_{t+1}(s_{t+1}|\overline{\theta})ds_{t+1} Q_{\overline{\theta}}(d\lambda_0)\\
&= \int_{\Pi(\Theta)} \int_\Theta \int_{S_{t+1}} \frac{\int_R V g(V|s_1,...,s_t)dV}{\int_R V g(V|s_1,...,s_t,s_{t+1})dV} l_{t+1}(s_{t+1}|\theta)ds_{t+1}\lambda_t(d\theta)Q_{\overline{\theta}}(d\lambda_0)\\
&\geq \int_{\Pi(\Theta)} \int_\Theta \frac{\int_R V g(V|s_1,...,s_t)dV}{\int_{S_{t+1}}(\int_R V g(V|s_1,...,s_t,s_{t+1})dV) l_{t+1}(s_{t+1}|\theta)ds_{t+1}} \lambda_t(d\theta)Q_{\overline{\theta}}(d\lambda_0)\\
&= 1.
\end{aligned}$$

(The weak inequality in this derivation cannot be made strict, because V may be independent of s_{t+1}, in which case:

$$\int_R Vg(V|s_1, ..., s_t, s_{t+1})dV = \int_R Vg(V|s_1, ..., s_t)dV,$$

which would imply: $E[\frac{p_t}{p_{t+1}}|\mathcal{F}_t \vee \overline{\theta}] = 1$. Also, if the expectation about V changes little over time, i.e.,

$$\int_R Vg(V|s_1, ..., s_t, s_{t+1})dV \approx \int_R Vg(V|s_1, ..., s_t)dV,$$

the inequality will provide a tight bound. This is illustrated in Figure 9.)

The result obtains because $x_{t+1}^{+} = 1 - p_t/p_{t+1}$.

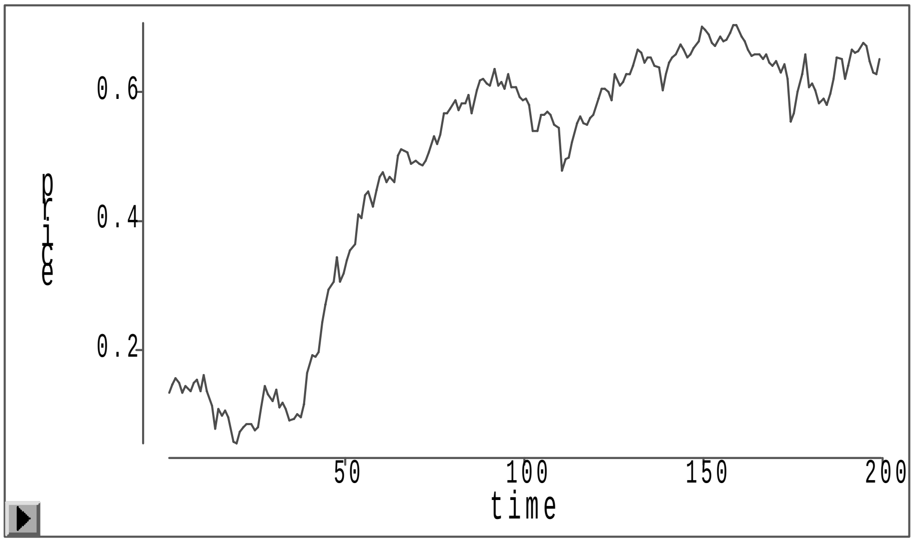

Figure 1: Typical price path of an Arrow-Debreu security that eventually matures "in-the-money."

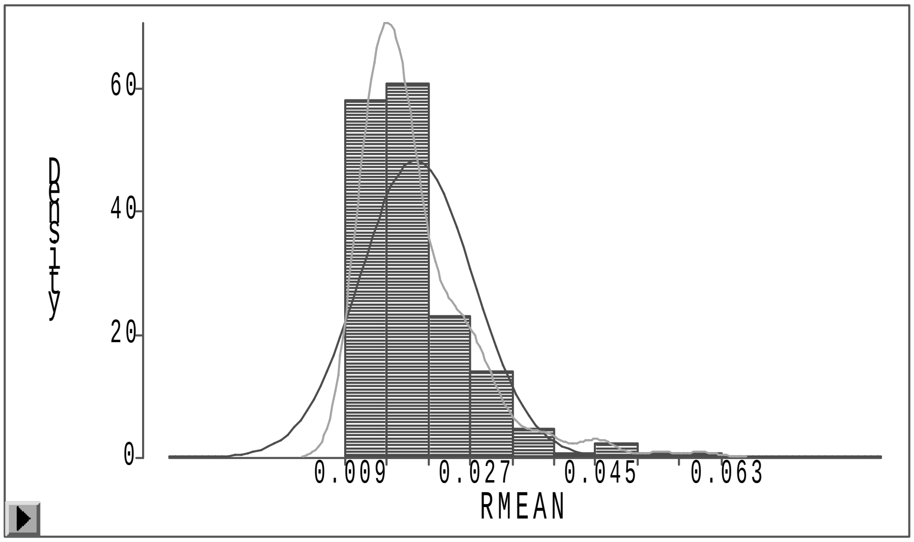

Figure 2: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the time series average return of an Arrow-Debreu security that matures "in-the-money" (RMEAN). Results based on 200 independent time series of length 200 each.

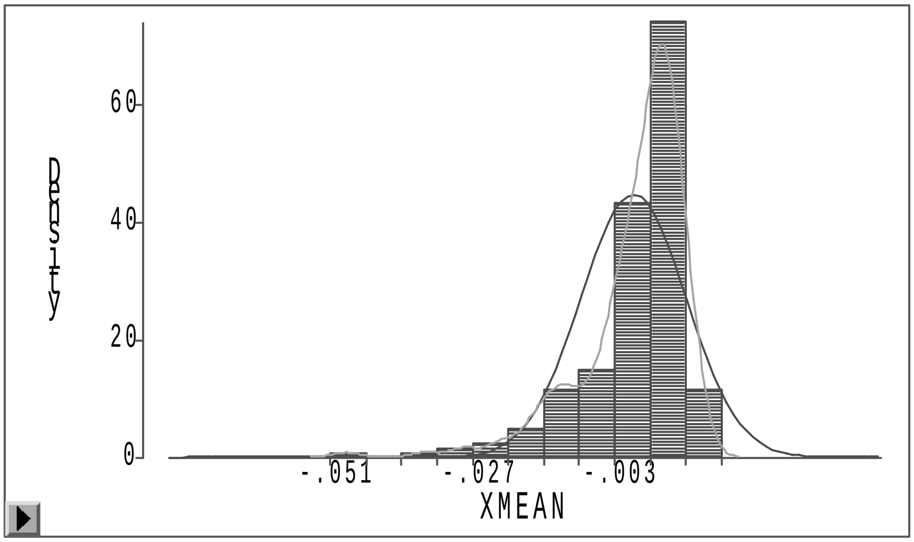

Figure 3: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the time series average payoff of an Arrow-Debreu security that matures "in-the-money" (XMEAN). The payoff is normalized by the *end-of-period* price. Results based on 200 independent time series of length 200 each.

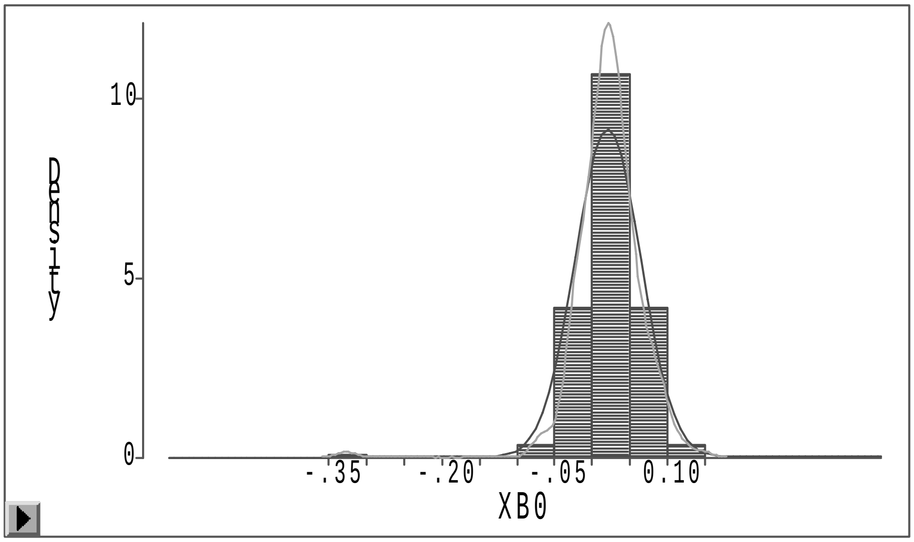

Figure 4: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the intercept estimate (XB0), in a time series OLS regression of the payoff on an Arrow-Debreu security that matures "in-the-money" onto its beginning-of-period price. The payoff is normalized by the *end-of-period* price. Results based on 200 independent time series of length 200 each.

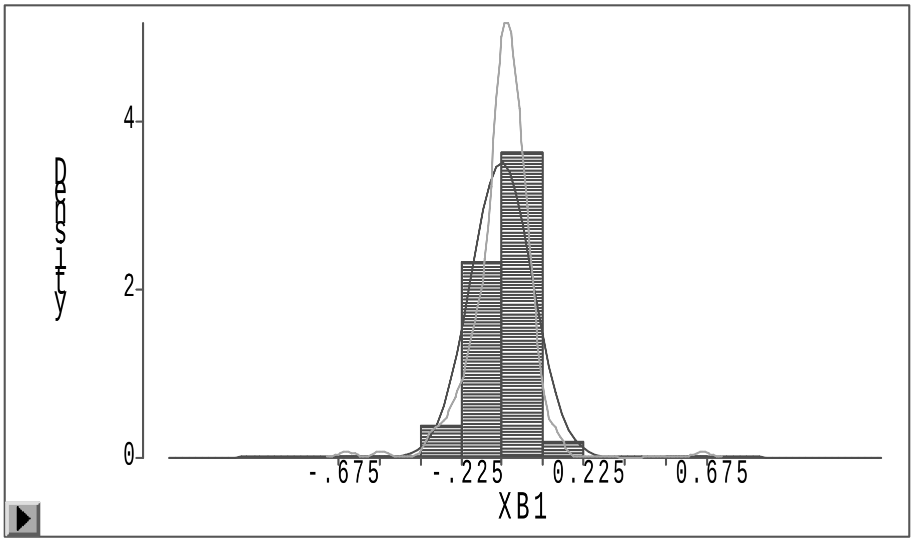

Figure 5: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the slope estimate (XB1), in a time series OLS regression of the payoff on an Arrow-Debreu security that matures "in-the-money" onto its beginning-of-period price. The payoff is normalized by the *end-of-period* price. Results based on 200 independent time series of length 200 each.

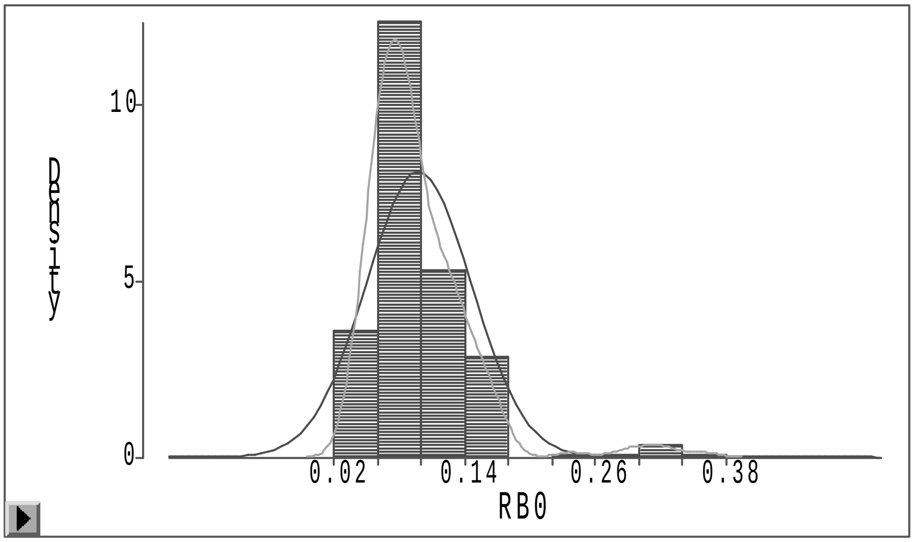

Figure 6: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the intercept estimate (RB0), in a time series OLS regression of the return on an Arrow-Debreu security that matures "in-the-money" onto its beginning-of-period price. Results based on 200 independent time series of length 200 each.

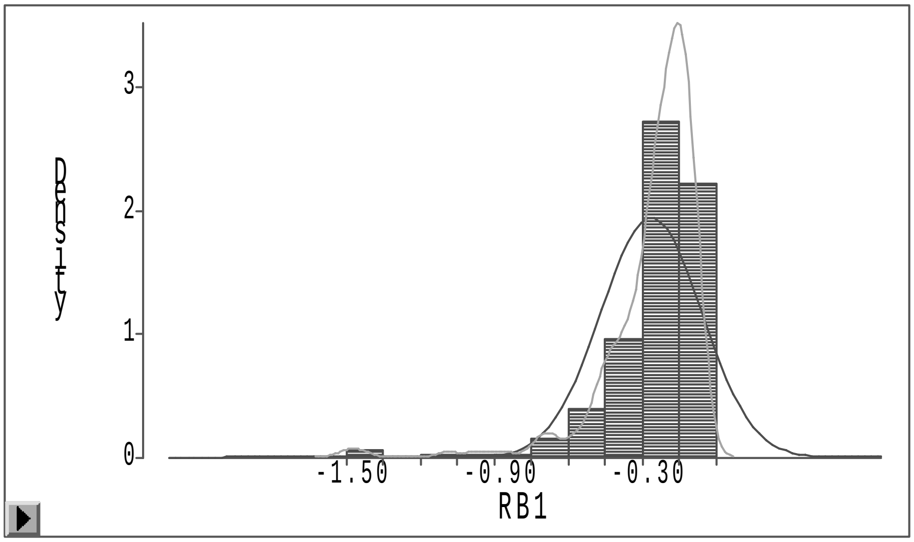

Figure 7: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the slope estimate (RB1), in a time series OLS regression of the return on an Arrow-Debreu security that matures "in-the-money" onto its beginning-of-period price. Results based on 200 independent time series of length 200 each.

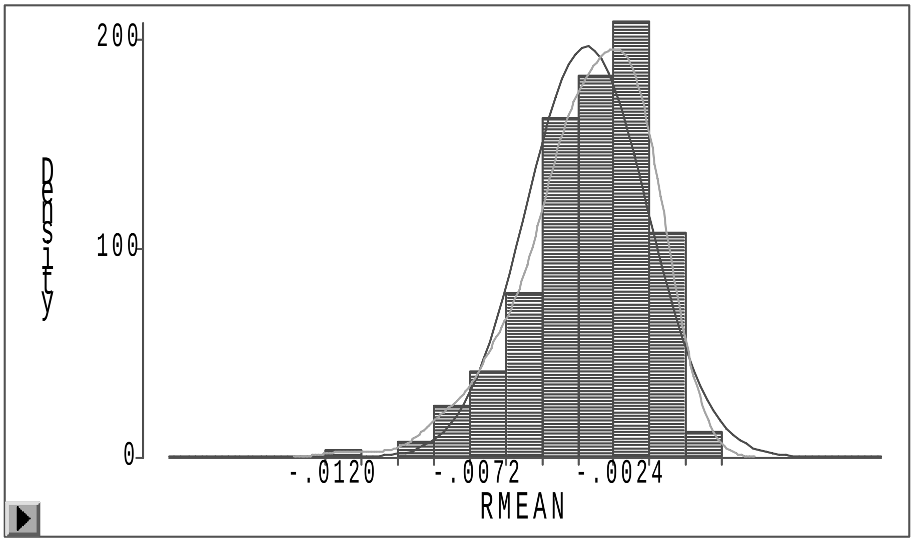

Figure 8: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the time series average return of an Arrow-Debreu security that matures "out-of-the-money" (RMEAN). Results based on 200 independent time series of length 200 each.

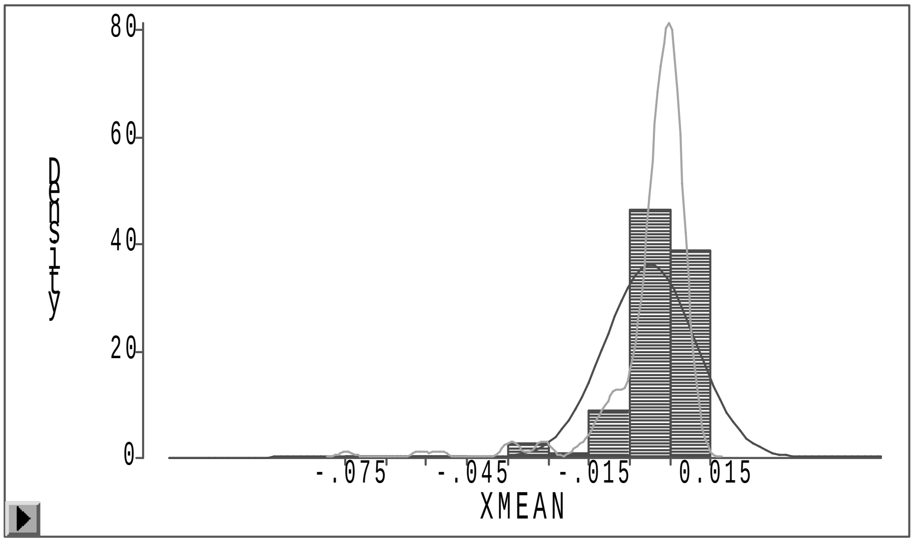

Figure 9: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the time series average payoff of a *random-payout* Arrow-Debreu security that matures "in-the-money" (XMEAN). The payoff is normalized by the *end-of-period* price. Results based on 200 independent time series of length 200 each.

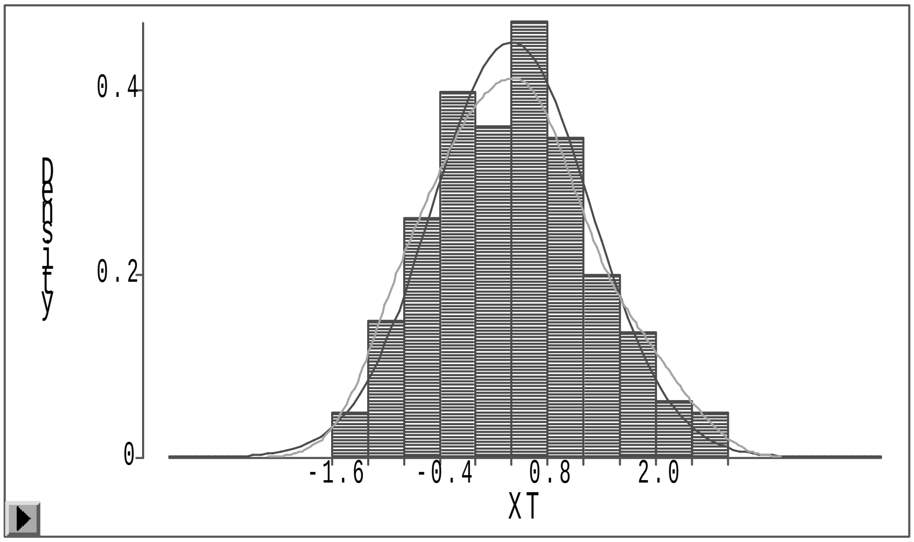

Figure 10: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the z-statistic computed from the time series average payoff of an Arrow-Debreu security that matures "in-the-money" (XT). The payoff is normalized by the *end-of-period* price. Results based on 200 independent time series of length 200 each.

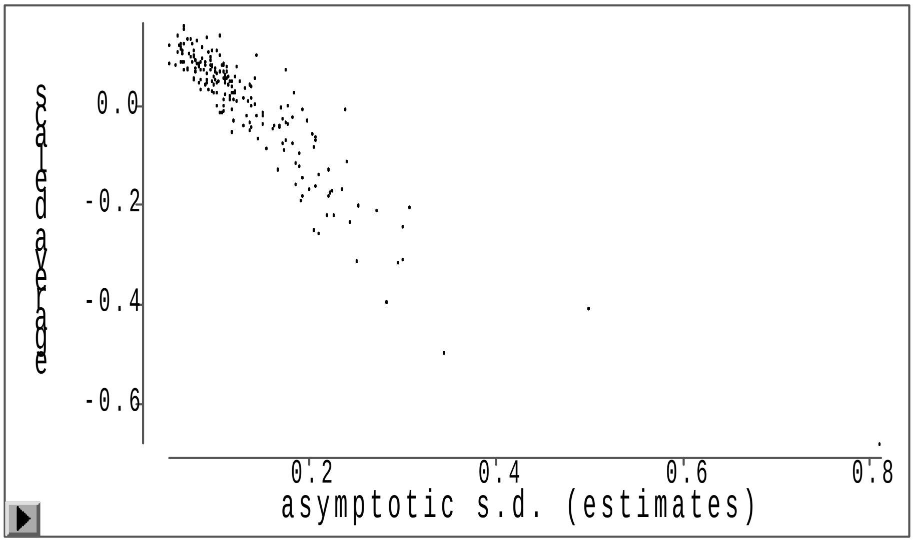

Figure 11: Plot of the scaled time series average payoff of an Arrow-Debreu security that matures "in-the-money," against the estimate of its asymptotic standard deviation (s.d.). The payoff is normalized by the *end-of-period* price. Results based on 200 independent time series of length 200 each.

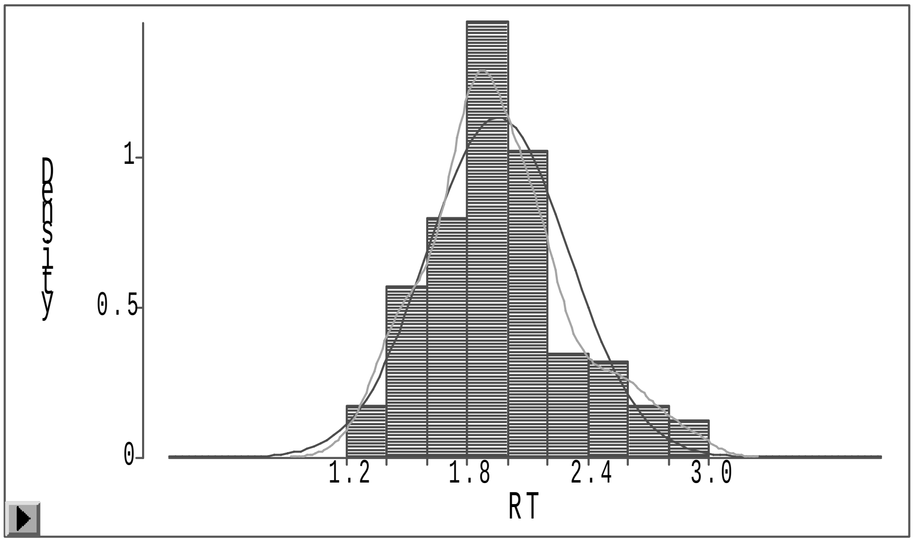

Figure 12: Histogram, kernel density estimate and Normal curve evaluated at the sample mean and sample standard deviation, of the z-statistic computed from the time series average return of an Arrow-Debreu security that matures "in-the-money" (RT) Results based on 200 independent time series of length 200 each.