

# THE QUARTERLY JOURNAL OF ECONOMICS

---

Vol. CXVI

February 2001

Issue 1

---

## PROSPECT THEORY AND ASSET PRICES\*

NICHOLAS BARBERIS  
MING HUANG  
TANO SANTOS

We study asset prices in an economy where investors derive direct utility not only from consumption but also from fluctuations in the value of their financial wealth. They are loss averse over these fluctuations, and the degree of loss aversion depends on their prior investment performance. We find that our framework can help explain the high mean, excess volatility, and predictability of stock returns, as well as their low correlation with consumption growth. The design of our model is influenced by prospect theory and by experimental evidence on how prior outcomes affect risky choice.

### I. INTRODUCTION

For many years now, the standard framework for thinking about aggregate stock market behavior has been the consumption-based approach. As is well-known, this approach presents a number of difficulties. In its simplest form, it does not come close to capturing the stock market's high historical average return and volatility, nor the striking variation in expected stock returns over time.<sup>1</sup> Over the past decade researchers have used ever more sophisticated specifications for utility over consumption in an attempt to approximate the data more closely.<sup>2</sup> These efforts have

\* We are grateful to John Cochrane, George Constantinides, Kent Daniel, Darrell Duffie, Lars Hansen, Sendhil Mullainathan, Canice Prendergast, Andrei Shleifer, Kenneth Singleton, Richard Thaler, Stanley Zin, three anonymous referees, the editor Edward Glaeser, and participants in numerous workshops in the United States and Great Britain for helpful comments on earlier drafts.

1. See, for example, Hansen and Singleton [1983], Mehra and Prescott [1985], and Hansen and Jagannathan [1991].

2. Recent papers in this line of research include Abel [1990], Campbell and Cochrane [1999], Constantinides [1990], Epstein and Zin [1989, 1991], and Sundaresan [1989]. Another strand of the literature emphasizes market incompleteness due to uninsurable income shocks; see, for example, Heaton and Lucas

yielded some success. However, some basic features of stock returns, such as their low correlation with consumption growth, remain hard to understand.

In this paper we make the case for an alternative way of thinking about the aggregate stock market. Instead of trying to refine the consumption-based model further, we propose departing from it in a particular way. In the model we present below, the investor derives direct utility not only from consumption but also from changes in the value of his financial wealth. When deciding how much to invest in the stock market, he takes both types of utility into account: the objective function he maximizes includes an extra term reflecting a direct concern about financial wealth fluctuations. This contrasts with the traditional approach to asset pricing which holds that the only thing people take into account when choosing a portfolio is the future consumption utility that their wealth will bring.

Our specification of this additional source of utility captures two ideas we think are important for understanding investor behavior. First, our investor is much more sensitive to reductions in his financial wealth than to increases, a feature sometimes known as *loss aversion*. Second, how loss averse the investor is, depends on his prior investment performance. After prior gains, he becomes less loss averse: the prior gains will cushion any subsequent loss, making it more bearable. Conversely, after a prior loss, he becomes more loss averse: after being burned by the initial loss, he is more sensitive to additional setbacks.

By extending the traditional asset pricing framework in this way, we find that we are able to understand many of the hitherto perplexing features of aggregate data. In particular, starting from an underlying consumption growth process with *low* variance, our model generates stock returns with a high mean, high volatility, significant predictability, and low correlation with consumption growth, while maintaining a low and stable riskless interest rate.

In essence, our story is one of changing risk aversion. After a run-up in stock prices, our agent is less risk averse because those gains will cushion any subsequent loss. After a fall in stock prices, he becomes more wary of further losses and hence more risk averse. This variation in risk aversion allows returns in our model to be much more volatile than the underlying dividends: an unusually

---

[1996] and Constantinides and Duffie [1996]. Cochrane [1998] and Kocherlakota [1996] provide excellent surveys.

good dividend raises prices, but this price increase also makes the investor less risk averse, driving prices still higher. We also generate predictability in returns much like that observed in the data: following a significant rise in prices, the investor is less risk averse, and subsequent returns are therefore on average lower.

The model also produces a substantial equity premium: the high volatility of returns means that stocks often perform poorly, causing our loss-averse investor considerable discomfort. As a result, a large premium is required to convince him to hold stocks.

Our framework offers a distinct alternative to consumption-based models that attempt to understand the high mean, high volatility, and significant predictability of equity returns. Campbell and Cochrane [1999] explain these empirical features using an external habit level for consumption which generates time-varying risk aversion as current consumption moves closer to or farther from habit. Although our model is also based on changing risk aversion, we generate it by introducing loss aversion over financial wealth fluctuations and allowing the degree of loss aversion to be affected by prior investment performance.

The differences between our framework and a consumption-based approach like Campbell and Cochrane [1999] are highlighted by the distinct predictions of each. In the consumption-based model, a large component of stock return volatility comes from changes in risk aversion that are ultimately driven by consumption. It is therefore inevitable that stock returns and consumption are significantly correlated, although this is not the case in the data. In our framework, changes in risk aversion are driven by past stock market movements and hence ultimately by news about dividends. Since dividends are only weakly correlated with consumption, returns in our model are also only weakly correlated with consumption.

Our approach is also related to the literature on first-order risk aversion, as introduced using recursive utility by Epstein and Zin [1990] among others. So far, this literature has not allowed for time-varying risk aversion and is therefore unable to account for the high volatility of stock returns, although this could be incorporated without difficulty. A more basic difference is that most implementations of first-order risk aversion effectively make the investor loss averse over *total* wealth fluctuations as opposed to financial wealth fluctuations, as in this paper. This distinction is important because it underlies a number of our predictions, including the low correlation between consumption growth and stock returns.

At a more fundamental level, our framework differs from the consumption-based approach in the way it defines risk. In consumption-based models, assets are only risky to the extent that their returns covary with consumption growth. In our framework, the investor cares about fluctuations in financial wealth whether or not those fluctuations are correlated with consumption growth. Since we are measuring risk differently, it is not surprising that the level of risk aversion we need to explain the data is also affected. While we do assume a substantial level of risk aversion, it is not nearly as extreme as that required in many consumption-based approaches.

The design of our model draws on two long-standing ideas in the psychology literature. The idea that people care about changes in financial wealth and that they are loss averse over these changes is a central feature of the prospect theory of Kahneman and Tversky [1979]. Prospect theory is a descriptive model of decision making under risk, originally developed to help explain the numerous violations of the expected utility paradigm documented over the years.

The idea that prior outcomes may affect subsequent risk-taking behavior is supported by another strand of the psychology literature. Thaler and Johnson [1990], for example, find that when faced with sequential gambles, people are more willing to take risk if they made money on prior gambles, than if they lost. They interpret these findings as revealing that losses are less painful to people if they occur after prior gains, and more painful if they follow prior losses. The result that risk aversion goes down after prior gains, confirmed in other studies, has been labeled the “house money” effect, reflecting gamblers’ increased willingness to bet when ahead.

Our work is related to that of Benartzi and Thaler [1995], who examine single-period portfolio choice for an investor with prospect-type utility. They find that loss aversion makes investors reluctant to invest in stocks, even in the face of a sizable equity premium. This suggests that bringing prospect theory into a formal pricing model may help us understand the level of average returns. While our work confirms this, we find that loss aversion cannot *by itself* explain the equity premium; incorporating the effect of prior outcomes is a critical ingredient as well. To see this, we also examine a simpler model where prior outcomes are ignored and hence where the pain of a loss is the same, regardless of past history. The investor’s risk aversion is then constant over time, and stock prices lose an important source of volatility. With less volatile returns and hence less risk, we are no longer able to produce a substantial equity premium.

Another set of papers, including Barberis, Shleifer, and Vishny [1998] and Daniel, Hirshleifer, and Subrahmanyam [1998], explains some empirical features of asset returns by assuming that investors exhibit irrationality when making forecasts of quantities such as cash flows. Other papers, including Hong and Stein [1999], suppose that investors are only able to process subsets of available information. Here, we take a different approach. While we do modify the investor's preferences to reflect experimental evidence about the sources of utility, the investor remains rational and dynamically consistent throughout.<sup>3</sup>

In Section II we show how loss aversion over financial wealth fluctuations and the effect of prior outcomes can be introduced into an asset pricing framework. Section III discusses studies in the psychology literature that we draw on in specifying the model. Section IV characterizes equilibrium asset prices and presents intuition for the results. In Section V we investigate the model's ability to explain the aggregate data through a detailed numerical analysis. In Section VI we examine the importance of taking account of prior outcomes by analyzing a simpler model where they are ignored. Section VII concludes.

## II. INVESTOR PREFERENCES

Our starting point is the traditional consumption-based asset pricing model of Lucas [1978]. There is a continuum of identical infinitely lived agents in the economy, with a total mass of one, and two assets: a risk-free asset in zero net supply, paying a gross interest rate of  $R_{f,t}$  between time  $t$  and  $t + 1$ ; and one unit of a risky asset, paying a gross return of  $R_{t+1}$  between time  $t$  and  $t + 1$ . In the usual way, the risky asset—stock—is a claim to a stream of perishable output represented by the dividend sequence  $\{D_t\}$ , where dividend growth is given by

$$(1) \quad \log(D_{t+1}/D_t) = g_D + \sigma_D \epsilon_{t+1},$$

where  $\epsilon_{t+1} \sim \text{i.i.d. } N(0,1)$ .

Up to this point, our framework is entirely standard. We depart from the usual setup in the way we model investor preferences. In particular, our agents choose a consumption level  $C_t$  and an allocation to the risky asset  $S_t$  to maximize

3. See Shleifer [1999] for a recent treatment of irrationality in financial markets.

$$(2) \quad \mathbf{E} \left[ \sum_{t=0}^{\infty} \left( \rho^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_t \rho^{t+1} v(X_{t+1}, S_t, z_t) \right) \right].$$

The first term in this preference specification, utility over consumption  $C_t$ , is a standard feature of asset pricing models. Although our framework does not require it, we specialize to power utility, the benchmark case studied in the literature. The parameter  $\rho$  is the time discount factor, and  $\gamma > 0$  controls the curvature of utility over consumption.<sup>4</sup>

The second term represents utility from fluctuations in the value of financial wealth. The variable  $X_{t+1}$  is the gain or loss the agent experiences on his financial investments between time  $t$  and  $t + 1$ , a positive value indicating a gain and a negative value, a loss. The utility the investor receives from this gain or loss is  $v(X_{t+1}, S_t, z_t)$ . It is a function not only of the gain or loss  $X_{t+1}$  itself, but also of  $S_t$ , the value of the investor's risky asset holdings at time  $t$ , and a state variable  $z_t$  which measures the investor's gains or losses prior to time  $t$  as a fraction of  $S_t$ . By including  $S_t$  and  $z_t$  as arguments of  $v$ , we allow the investor's prior investment performance to affect the way subsequent losses are experienced, and hence his willingness to take risk. Finally,  $b_t$  is an exogenous scaling factor that we specify later.

The utility that comes from fluctuations in financial wealth can be interpreted in a number of different ways. We prefer to think of it as capturing feelings unrelated to consumption. After a big loss in the stock market, an investor may experience a sense of regret over his decision to invest in stocks; he may interpret his loss as a sign that he is a second-rate investor, thus dealing his ego a painful blow; and he may feel humiliation in front of friends and family when word leaks out.<sup>5</sup>

In summary, the preference specification in (2) recognizes

4. For  $\gamma = 1$ , we replace  $C_t^{1-\gamma}/(1-\gamma)$  with  $\log(C_t)$ .

5. One could potentially also interpret the second term in (2) as capturing utility over *anticipated* consumption: when an investor finds out that his wealth has gone up, he may get utility from savoring the thought of the additional future consumption that his greater wealth will bring. The difficulty with this interpretation is that it is really only an explanation of why people might get utility from fluctuations in *total* wealth. To motivate utility over financial wealth fluctuations, one would need to argue that investors track different components of their wealth separately and get utility from fluctuations in each one. It would then be natural to add to (2) a term reflecting a concern for fluctuations in the value of human capital, another major source of wealth. In fact, it turns out that doing so does not affect our results so long as the labor income process underlying human capital is exogenously specified.

that people may get direct utility from sources other than consumption, and also says that they anticipate these other sources of utility when making decisions today. This is a departure from traditional approaches, which hold that the only thing people think about when choosing a portfolio is the future consumption utility that their wealth will bring. While our preferences are nonstandard, this does not mean that they are irrational in any sense: it is not irrational for people to get utility from sources other than consumption, nor is it irrational for them to anticipate these feelings when making decisions.

Introducing utility over gains and losses in financial wealth raises a number of issues: (i) how does the investor measure his gains and losses  $X_{t+1}$ ? (ii) how does  $z_t$  track prior gains and losses? (iii) how does utility  $v$  depend on the gains and losses  $X_{t+1}$ ? and (iv) how does  $z_t$  change over time? Subsections A through D below tackle each of these questions in turn. Finally, subsection E discusses the scaling factor  $b_t$ .

#### A. Measuring Gains and Losses

The gains and losses in our model refer to changes in the value of the investor's *financial wealth*, even if this is only one component of his overall wealth. For simplicity, we go one step farther. Even though there are two financial assets, we suppose that the investor cares only about fluctuations in the value of the risky asset.<sup>6</sup>

Next, we need to specify the horizon over which gains and losses are measured. Put differently, how often does the agent seriously evaluate his investment performance? We follow the suggestion of Benartzi and Thaler [1995] that the most natural evaluation period is a year. As they point out, we file taxes once a year and receive our most comprehensive mutual fund reports once a year; moreover, institutional investors scrutinize their money managers' performance most carefully on an annual basis. Since this is an important assumption, we will investigate its impact on our results later in the paper.

Our investor therefore monitors fluctuations in the value of his stock portfolio from year to year and gets utility from those fluctuations. To fix ideas, suppose that  $S_t$ , the time  $t$  value of the

6. A simple justification for this is that since the return on the risk-free asset is known in advance, the investor does not get utility from changes in its value in the way that he does from changes in risky asset value. We also show later that for one reasonable way of measuring gains and losses, it makes no difference whether they are computed over total financial wealth or over the risky asset alone.

investor's holdings of the risky asset, is \$100. Imagine that by time  $t + 1$ , this value has gone up to  $S_t R_{t+1} = \$120$ . The exact way the investor measures this gain depends on the reference level with which \$120 is compared. One possible reference level is the status quo or initial value  $S_t = \$100$ . The gain would then be measured as \$20, or more generally as  $X_{t+1} = S_t R_{t+1} - S_t$ .

This is essentially our approach, but for one modification which we think is realistic: we take the reference level to be the status quo *scaled up by the risk-free rate*,  $S_t R_{f,t}$ . In our example, and with a risk-free rate of say 5 percent, this means a reference level of 105. An end-of-period risky asset value of 120 would then lead the investor to code a gain of 15, while a value of 100 would generate a loss of  $-5$ . In general terms, the investor will code a gain or loss of

$$(3) \quad X_{t+1} = S_t R_{t+1} - S_t R_{f,t}.$$

The idea here is that in an economy offering a riskless return of 5 percent, the investor is likely to be disappointed if his stock market investment returns only 4 percent. The riskless return may not be the investor's only point of comparison, although we suggest that it is a reasonable one. We will examine the sensitivity of our results to this choice later in the paper.<sup>7</sup>

### B. Tracking Prior Investment Outcomes

Now that we have explained how gains and losses are measured, we need to specify the utility they bring the investor. The simplest approach is to say that the utility of a gain or loss  $X_{t+1}$  is  $v(X_{t+1})$ , in other words, a function of the size of the gain or loss alone.

In our model, we allow the pain of a loss to depend not only on the size of the loss but also on investment performance *prior* to the loss. A loss that comes after substantial prior gains may be less painful than usual because it is cushioned by those earlier gains. Put differently, the investor may not care much about a stock market dip that follows substantial prior gains because he can still tell himself that he is "up, relative to a year ago," say.

Conversely, losses that come on the heels of substantial prior losses may be more painful than average for the investor. If he

7. Note that if the investor does use the risk-free rate as a reference level, it is irrelevant whether gains and losses are calculated over total financial wealth or over the risky asset alone: if  $B_t$  and  $S_t$  represent the investor's holdings of the risk-free asset and the risky asset, respectively, at time  $t$ , then  $(B_t R_{f,t} + S_t R_{t+1}) - (B_t + S_t) R_{f,t}$  is the same as  $S_t(R_{t+1} - R_{f,t})$ .



has been burned by a painful loss, he may be particularly sensitive to additional setbacks.

To capture the influence of prior outcomes, we introduce the concept of a *historical benchmark level*  $Z_t$  for the value of the risky asset.<sup>8</sup> We propose that when judging the recent performance of a stock, investors compare  $S_t$ , the value of their stock holdings today, with some value  $Z_t$  which represents a price that they remember the stock trading at in the past. Different investors will form this benchmark in different ways. For some investors, it may represent an average of recent stock prices. For others, it may be the specific stock price at salient moments in the past, such as the end of a year. Whichever way the benchmark level is formed, the difference  $S_t - Z_t$ , when positive, is the investor's personal measure of how much "he is up" on his investment at time  $t$  and conversely, when negative, how much "he is down."

Introducing  $Z_t$  is helpful in modeling the influence of prior outcomes on the way subsequent gains and losses are experienced. When  $S_t > Z_t$ , the investor has had prior gains, making subsequent losses less painful and lowering the investor's risk aversion. Conversely, when  $S_t < Z_t$ , the investor has endured prior losses. Subsequent losses are more painful, and the investor is more risk averse than usual.

Since  $S_t$  and  $Z_t$  summarize how the investor perceives his past performance, a simple way of capturing the effect of prior outcomes would be to write the utility of financial wealth fluctuations as  $v(X_{t+1}, S_t, Z_t)$ . For modeling purposes, we find it more convenient to write it as  $v(X_{t+1}, S_t, z_t)$ , where  $z_t = Z_t/S_t$ .

### C. Utility from Gains and Losses

In defining  $v(X_{t+1}, S_t, z_t)$ , we consider three separate cases:  $z_t = 1$ , where the investor has neither prior gains nor prior losses on his investments;  $z_t < 1$ , the case of prior gains; and  $z_t > 1$ , the case of prior losses.

We start with the case of  $z_t = 1$ . We want to model the idea that investors are much more sensitive to reductions in financial

8. We use the term *benchmark level* to distinguish  $Z_t$  from the *reference level*  $S_t R_{f,t}$ . The reference level determines the size of the gain or loss. The benchmark level  $Z_t$  determines the magnitude of the utility received from that gain or loss, in a way that we soon make precise. While we are careful to stick to this terminology, some readers may find it helpful to think of  $Z_t$  as a secondary reference level that also affects the investor's decisions.

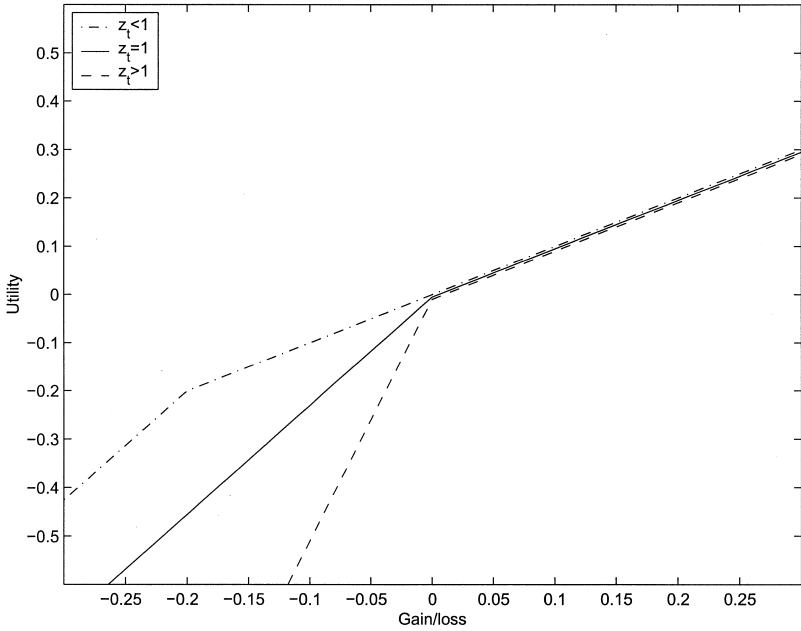


FIGURE I  
Utility of Gains and Losses

The dash-dot line represents the case where the investor has prior gains, the dashed line the case of prior losses, and the solid line the case where he has neither prior gains nor losses.

wealth than to increases, a feature sometimes known as *loss aversion*. We capture this by defining

$$(4) \quad v(X_{t+1}, S_t, 1) = \begin{cases} X_{t+1} & \text{for } X_{t+1} \geq 0 \\ \lambda X_{t+1} & \text{for } X_{t+1} < 0, \end{cases}$$

with  $\lambda > 1$ . This is a piecewise linear function, shown as the solid line in Figure I. It is kinked at the origin, where the gain equals zero.

We now turn to  $z_t < 1$ , where the investor has accumulated prior gains in the stock market. The dash-dot line in Figure I shows the form of  $v(X_{t+1}, S_t, z_t)$  in this case. It differs from  $v(X_{t+1}, S_t, 1)$  in the way losses are penalized. Small losses are not penalized very heavily, but once the loss exceeds a certain amount, it starts being penalized at a more severe rate. The intuition is that if the investor has built up a cushion of prior

gains, these gains may soften the blow of small subsequent losses, although they may not be enough to protect him against larger losses.

To understand how we formalize this intuition, an example may be helpful. Suppose that the current stock value is  $S_t = \$100$ , but that the investor has recently accumulated some gains on his investments. A reasonable historical benchmark level is  $Z_t = \$90$ , since the stock must have gone up in value recently. As discussed above, we can think of  $\$90$  as the value of the stock one year ago, which the investor still remembers. The difference  $S_t - Z_t = \$10$  represents the cushion, or reserve of prior gains that the investor has built up. Suppose finally that the risk-free rate is zero.

Imagine that over the next year, the value of the stock falls from  $S_t = \$100$  down to  $S_t R_{t+1} = \$80$ . In the case of  $z_t = 1$ , where the investor has no prior gains or losses, equations (3) and (4) show that we measure the pain of this loss as

$$(80 - 100)(\lambda) = -40$$

for a  $\lambda$  of 2.

When the investor has some prior gains, this calculation probably overstates actual discomfort. We propose a more realistic measure of the pain caused: since the first  $\$10$  drop, from  $S_t = \$100$  down to  $Z_t = \$90$ , is completely cushioned by the  $\$10$  reserve of prior gains, we penalize it at a rate of only 1, rather than  $\lambda$ . The second part of the loss, from  $Z_t = \$90$  down to  $S_t R_{t+1} = \$80$  will be *more* painful since all prior gains have already been depleted, and we penalize it at the higher rate of  $\lambda$ . Using a  $\lambda$  of 2 again, the overall disutility of the  $\$20$  loss is

$$\begin{aligned} (90 - 100)(1) + (80 - 90)(\lambda) \\ = (90 - 100)(1) + (80 - 90)(2) = -30, \end{aligned}$$

or in general terms

$$\begin{aligned} (Z_t - S_t)(1) + (S_t R_{t+1} - Z_t)(\lambda) \\ = S_t(z_t - 1)(1) + S_t(R_{t+1} - z_t)(\lambda). \end{aligned}$$

Note that if the loss is small enough to be completely cushioned by the prior gain—in other words, if  $S_t R_{t+1} > Z_t$ , or equivalently,  $R_{t+1} > z_t$ —there is no need to break the loss up into

two parts. Rather, the entire loss of  $S_t R_{t+1} - S_t$  is penalized at the gentler rate of 1.

In summary, then, we give  $v(X_{t+1}, S_t, z_t)$  the following form for the case of prior gains, or  $z_t \leq 1$ :

$$(5) \quad v(X_{t+1}, S_t, z_t) = \begin{cases} S_t R_{t+1} - S_t \\ S_t(z_t - 1) + \lambda S_t(R_{t+1} - z_t) \end{cases} \quad \text{for} \quad \begin{cases} R_{t+1} \geq z_t \\ R_{t+1} < z_t. \end{cases}$$

For the more relevant case of a nonzero riskless rate  $R_{f,t}$ , we scale both the reference level  $S_t$  and the benchmark level  $Z_t$  up by the risk-free rate, so that<sup>9</sup>

$$(6) \quad v(X_{t+1}, S_t, z_t) = \begin{cases} S_t R_{t+1} - S_t R_{f,t} \\ S_t(z_t R_{f,t} - R_{f,t}) + \lambda S_t(R_{t+1} - z_t R_{f,t}) \end{cases} \quad \text{for} \quad \begin{cases} R_{t+1} \geq z_t R_{f,t} \\ R_{t+1} < z_t R_{f,t}. \end{cases}$$

Finally, we turn to  $z_t > 1$ , where the investor has recently experienced losses on his investments. The form of  $v(X_{t+1}, S_t, z_t)$  in this case is shown as the dashed line in Figure I. It differs from  $v(X_{t+1}, S_t, 1)$  in that losses are penalized more heavily, capturing the idea that losses that come on the heels of other losses are more painful than usual. More formally,

$$(7) \quad v(X_{t+1}, S_t, z_t) = \begin{cases} X_{t+1} \\ \lambda(z_t) X_{t+1} \end{cases} \quad \text{for} \quad \begin{cases} X_{t+1} \geq 0 \\ X_{t+1} < 0, \end{cases}$$

where  $\lambda(z_t) > \lambda$ . Note that the penalty  $\lambda(z_t)$  is a function of the size of prior losses, measured by  $z_t$ . In the interest of simplicity, we set

$$(8) \quad \lambda(z_t) = \lambda + k(z_t - 1),$$

where  $k > 0$ . The larger the prior loss, or equivalently, the larger  $z_t$  is, the more painful subsequent losses will be.

We illustrate this with another example. Suppose that the current stock value is  $S_t = \$100$ , and that the investor has recently experienced losses. A reasonable historical benchmark level is then  $Z_t = \$110$ , higher than  $\$100$  since the stock has been falling. By definition,  $z_t = 1.1$ . Suppose for now that  $\lambda = 2$ ,  $k = 3$ , and that the risk-free rate is zero.

9. Although the formula for  $v$  depends on  $R_{t+1}$  as well, we do not make the return an explicit argument of  $v$  since it can be backed out of  $X_{t+1}$  and  $S_t$ .

Imagine that over the next year, the value of the stock falls from  $S_t = \$100$  down to  $S_t R_{t+1} = \$90$ . In the case of  $z_t = 1$ , where the investor has no prior gains or losses, equations (3) and (4) show that we measure the pain of this loss as

$$(90 - 100)(\lambda) = (90 - 100)(2) = -20.$$

In our example, though, there has been a prior loss and  $z_t = 1.1$ . This means that the pain will now be

$$(90 - 100)(\lambda + 3(0.1)) = (90 - 100)(2 + 3(0.1)) = -23,$$

capturing the idea that losses are more painful after prior losses.

#### *D. Dynamics of the Benchmark Level*

To complete our description of the model, we need to discuss how the investor's cushion of prior gains changes over time. In formal terms, we have to specify how  $z_t$  moves over time, or equivalently how the historical benchmark level  $Z_t$  reacts to changes in stock value  $S_t$ . There are two ways that the value of the investor's stock holdings can change. First, it can change at time  $t$  because of an action taken by the investor: he may take out the dividend and consume it, or he may buy or sell some shares. For this type of change, we assume that  $Z_t$  changes in proportion to  $S_t$ , so that  $z_t$  remains constant. For example, suppose that the initial value of the investor's stock holdings is  $S_t = \$100$  and that  $Z_t = \$80$ , implying that he has accumulated \$20 of prior gains. If he sells \$10 of stock for consumption purposes, bringing  $S_t$  down to \$90, we assume that  $Z_t$  falls to \$72, so that  $z_t$  remains constant at 0.8. In other words, when the investor sells stock for consumption, we assume that he uses up some of his prior gains.

The assumption that the investor's actions do not affect the evolution of  $z_t$  is reasonable for transactions of moderate size, or more precisely, for moderate deviations from a strategy in which the investor holds a fixed number of shares and consumes the dividend each period. However, larger deviations—a complete exit from the stock market, for example—might plausibly affect the way  $z_t$  evolves. In supposing that they do not, we make a strong assumption, but one that is very helpful in keeping our analysis tractable. We discuss the economic interpretation of this assumption further in Section IV when we compute equilibrium prices.

The second way stock value can change is simply through its return between time  $t$  and time  $t + 1$ . In this case, the only

requirement we impose on  $Z_t$  is that it respond *sluggishly* to changes in the value of the risky asset. By this we mean that when the stock price moves up by a lot, the benchmark level also moves up, but by less. Conversely, if the stock price falls sharply, the benchmark level does not adjust downwards by as much.

Sluggishness turns out to be a very intuitive requirement to impose. To see this, recall that the difference  $S_t - Z_t$  is the investor's measure of his reserve of prior gains. How should this quantity change as a result of a change in the level of the stock market? If the return on the stock market is particularly good, investors should feel as though they have increased their reserve of prior gains. Mathematically, this means that the benchmark level  $Z_t$  should move up less than the stock price itself, so that the cushion at time  $t + 1$ , namely  $S_{t+1} - Z_{t+1}$ , be larger than the cushion at time  $t$ ,  $S_t - Z_t$ . Conversely, if the return on market is particularly poor, the investor should feel as though his reserves of prior gains are depleted. For this to happen,  $Z_t$  must fall less than  $S_t$ .

A simple way of modeling the sluggishness of the benchmark level  $Z_t$  is to write the dynamics of  $z_t$  as

$$(9) \quad z_{t+1} = z_t \frac{\bar{R}}{R_{t+1}},$$

where  $\bar{R}$  is a fixed parameter. This equation then says that if the return on the risky asset is particularly good, so that  $R_{t+1} > \bar{R}$ , the state variable  $z = Z/S$  falls in value. This is consistent with the benchmark level  $Z_t$  behaving sluggishly, rising less than the stock price itself. Conversely, if the return is poor and  $R_{t+1} < \bar{R}$ , then  $z$  goes up. This is consistent with the benchmark level falling less than the stock price.<sup>10</sup>

$\bar{R}$  is not a free parameter in our model, but is determined endogenously by imposing the reasonable requirement that in equilibrium, the median value of  $z_t$  be equal to one. In other words, half the time the investor has prior gains, and the rest of the time he has prior losses. It turns out that  $\bar{R}$  is typically of similar magnitude to the average stock return.

We can generalize (9) slightly to allow for varying degrees of

10. The benchmark level dynamics in (9) are one simple way of capturing sluggishness. More generally, we can assume dynamics of the form  $z_{t+1} = g(z_t, R_{t+1})$ , where  $g(z_t, R_{t+1})$  is strictly increasing in  $z_t$  and strictly decreasing in  $R_{t+1}$ .

sluggishness in the dynamics of the historical benchmark level. One way to do this is to write

$$(10) \quad z_{t+1} = \eta \left( z_t \frac{\bar{R}}{R_{t+1}} \right) + (1 - \eta)(1).$$

When  $\eta = 1$ , this reduces to (9), which represents a sluggish benchmark level. When  $\eta = 0$ , it reduces to  $z_{t+1} = 1$ , which means that the benchmark level  $Z_t$  tracks the stock value  $S_t$  one-for-one throughout—a very fast moving benchmark level.

The parameter  $\eta$  can be given an interpretation in terms of the investor’s memory: it measures how far back the investor’s mind stretches when recalling past gains and losses. When  $\eta$  is near zero, the benchmark level  $Z_t$  is always close to the value of the stock  $S_t$ : prior gains and losses are quickly swallowed up and are not allowed to affect the investor for long. In effect, the investor has a short-term memory, recalling only the most recent prior outcomes. When  $\eta$  is closer to one, though, the benchmark level moves sluggishly, allowing past gains and losses to linger and affect the investor for a long time; in other words, the investor has a long memory.<sup>11</sup>

*E. The Scaling Term  $b_t$*

We scale the prospect theory term in the utility function to ensure that quantities like the price-dividend ratio and risky asset risk premium remain stationary even as aggregate wealth increases over time. Without a scaling factor, this will not be the case because the second term of the objective function will come to dominate the first as aggregate wealth grows. One reasonable specification of the scaling term is

$$(11) \quad b_t = b_0 \bar{C}_t^{-\gamma},$$

where  $\bar{C}_t$  is the *aggregate* per capita consumption at time  $t$ , and hence exogenous to the investor. By using an exogenous variable, we ensure that  $b_t$  simply acts as a neutral scaling factor, without affecting the economic intuition of the previous paragraphs.

The parameter  $b_0$  is a nonnegative constant that allows us to control the overall importance of utility from gains and losses in

11. A simple mathematical argument can be used to show that the “half-life” of the investor’s memory is equal to  $-0.693/\log \eta$ . In other words, after this amount of time, the investor has lost half of his memory. When  $\eta = 0.9$ , this quantity is 6.6 years, and when  $\eta = 0.8$ , it equals 3.1 years.

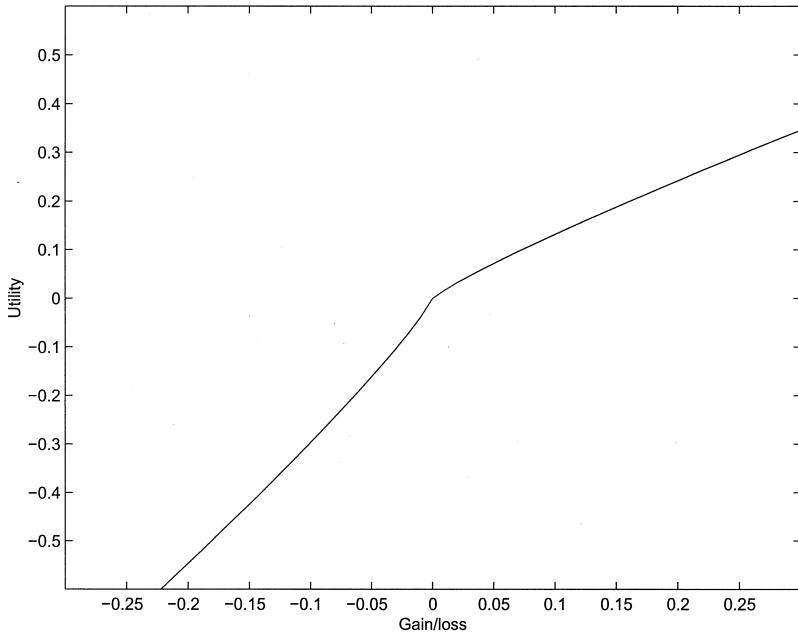


FIGURE II  
Kahneman-Tversky Value Function

financial wealth relative to utility from consumption. Setting  $b_0 = 0$  reduces our framework to the much studied consumption-based model with power utility.

### III. EVIDENCE FROM PSYCHOLOGY

The design of our model is influenced by some long-standing ideas from psychology. The idea that people care about changes in financial wealth and that they are loss averse over these changes is a central feature of the prospect theory of Kahneman and Tversky [1979]. Prospect theory is a descriptive model of decision making under risk that was originally developed to help explain the numerous violations of the expected utility paradigm documented over the years.

While our model is influenced by the work of Kahneman and Tversky [1979], we do not attempt an exhaustive implementation of all aspects of prospect theory. Figure II shows Kahneman and



Tversky's utility function over gains and losses:<sup>12</sup>

$$(12) \quad w(X) = \begin{cases} X^{0.88} & \text{for } X \geq 0 \\ -2.25(-X)^{0.88} & \text{for } X < 0. \end{cases}$$

It is similar to  $v(X_{t+1}, S_t, 1)$ —the solid line in Figure I—but is also mildly concave over gains and convex over losses. This curvature is most relevant when choosing between prospects that involve *only* gains or between prospects that involve *only* losses.<sup>13</sup> For gambles that can lead to both gains and losses—such as the one-year investment in stocks that our agent is evaluating—loss aversion at the kink is far more important than the degree of curvature away from the kink. For simplicity then, we make  $v$  linear over both gains and losses.

In our framework the “prospective utility” the investor receives from gains and losses is computed by taking the expected value of  $v$ , in other words by weighting the value of gains and losses by their probabilities. As a way of understanding Allais-type violations of the expected utility paradigm, Kahneman and Tversky [1979] suggest weighting the value of gains and losses not with the probabilities themselves but with a nonlinear transformation of those probabilities. Again, for simplicity, we abstract from this feature of prospect theory, and have no reason to believe that our results are qualitatively affected by this simplification.<sup>14</sup>

The idea that prior outcomes may affect willingness to take risk is also supported by recent studies in psychology. Thaler and Johnson [1990] study risk-taking behavior in an experimental setting using a large sample of Cornell undergraduate and MBA

12. This functional form is proposed in Tversky and Kahneman [1992] and is based on experimental findings.

13. Indeed, it is by offering subjects gambles over *only* losses or *only* gains that Kahneman and Tversky [1979] deduce the shape of the value function. They propose concavity over gains because subjects prefer a gamble that offers \$2000 w.p.  $\frac{1}{4}$ , \$4000 w.p.  $\frac{1}{4}$ , and \$0 w.p.  $\frac{1}{2}$  to the mean-preserving spread offering \$6000 w.p.  $\frac{1}{4}$  and \$0 otherwise. Preferences switch when the signs are flipped, suggesting convexity over losses.

14. Two other assumptions inherent in our framework are worth noting. First, we assume that investors rationally forecast the loss aversion that they will feel in the future. Loewenstein, O'Donoghue, and Rabin [1999] suggest that reality is more complex in that people have trouble forecasting how they will feel about future events. A more complete model would take this evidence into account, although we do not attempt it here. Second, we assume that investors rationally forecast future changes in the reference level  $S_t R_{f,t}$ . It is sometimes claimed that such rationality is inconsistent with loss aversion: if people know they are eventually going to reset their reference level after a stock market drop, why are they so averse to the loss in the first place? Levy and Wiener [1997] provide one answer, noting that changing the reference level forces the investor to confront and accept the loss, and this is the part that is painful.

students. They offer subjects a sequence of gambles and find that outcomes in earlier gambles affect subsequent behavior: following a gain, people appear to be more risk seeking than usual, taking on bets they would not normally accept. This result has become known as the “house money” effect, because it is reminiscent of the expression “playing with the house money” used to describe gamblers’ increased willingness to bet when ahead. Thaler and Johnson argue that these results suggest that losses are less painful after prior gains, perhaps because those gains cushion the subsequent setback.

Thaler and Johnson also find that after a loss, subjects display considerable reluctance to accept risky bets. They interpret this as showing that losses are more painful than usual following prior losses.<sup>15</sup>

The stakes used in Thaler and Johnson [1990] are small—the dollar amounts are typically in double digits. Interestingly, Gertner [1993] obtains similar results in a study involving much larger stakes. He studies the risk-taking behavior of participants in the television game show “Card Sharks,” where contestants place bets on whether a card to be drawn at random from a deck will be higher or lower than a card currently showing. He finds that the amount bet is a strongly increasing function of the contestant’s winnings up to that point in the show. Once again, this is evidence of more aggressive risk-taking behavior following substantial gains.

The evidence we have presented suggests that in the context of a sequence of gains and losses, people are less risk averse following prior gains and more risk averse after prior losses. This may initially appear puzzling to readers familiar with Kahneman and Tversky’s original value function, which is concave in the region of gains and convex in the region of losses. In particular, the convexity over losses is occasionally interpreted to mean that “after a loss, people are risk seeking,” contrary to Thaler and Johnson’s evidence. Hidden in this interpretation, though, is a critical assumption, namely that people *integrate* or “merge” the outcomes of successive gambles. Suppose that you have just suffered a loss of \$1000, and are contemplating a gamble equally

15. It is tempting to explain these results using a utility function where risk aversion decreases with wealth. However, any utility function with sufficient curvature to produce lower risk aversion after a \$20 gain, regardless of initial wealth level, inevitably makes counterfactual predictions about attitudes to large-scale gambles.

likely to win you \$200 as to lose you \$200. Integration of outcomes means that you make the decision about whether to take the gamble by comparing the average of  $w(-1200)$  and  $w(-800)$  with  $w(-1000)$ , where  $w$  is defined in equation (12). Of course, under this assumption, convexity in the region of losses leads to risk seeking after a loss.

The idea that people integrate the outcomes of sequential gambles, while appealingly simple, is only a hypothesis. Tversky and Kahneman [1981] themselves note that prospect theory was originally developed only for elementary, one-shot gambles and that any application to a dynamic context must await further evidence on how people think about sequences of gains and losses. A number of papers, including Thaler and Johnson [1990], have taken up this challenge, conducting experiments on whether people integrate sequential outcomes, segregate them, or do something else again. That these experiments have uncovered *increased* risk aversion after prior losses does not contradict prospect theory; it simply rejects the hypothesis that people integrate sequential gambles.<sup>16</sup>

Thaler and Johnson [1990] do find some situations where prior losses lead to risk-seeking behavior. These are situations where after the prior loss, the subject can take a gamble that offers a good chance of breaking even and only limited downside. In conjunction with the other evidence, this suggests that while losses after prior losses are very painful, gains that enable people with prior losses to break even are especially sweet. We have not been able to introduce this break-even effect into our framework in a tractable way. It is worth noting, though, that outside of the special situations uncovered by Thaler and Johnson, it is increased risk aversion that appears to be the norm after prior losses.

#### IV. EQUILIBRIUM PRICES

We now derive equilibrium asset prices in an economy populated by investors with preferences of the type described in Sec-

16. There is another sense in which integration of sequential outcomes is an implausible way to implement prospect theory in a multiperiod context. If investors did integrate many years of stock market gains and losses, they would essentially be valuing absolute levels of wealth, and not the *changes* in wealth that are so important to prospect theory. We thank J. B. Heaton for this observation.

tion II. It may be helpful to summarize those preferences here. Each investor chooses consumption  $C_t$  and an allocation to the risky asset  $S_t$  to maximize

$$(13) \quad \mathbf{E} \left[ \sum_{t=0}^{\infty} \left( \rho^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_0 \bar{C}_t^{-\gamma} \rho^{t+1} v(X_{t+1}, S_t, z_t) \right) \right],$$

subject to the standard budget constraint, where

$$(14) \quad X_{t+1} = S_t R_{t+1} - S_t R_{f,t},$$

and where for  $z_t \leq 1$ ,

$$(15) \quad v(X_{t+1}, S_t, z_t) = \begin{cases} S_t R_{t+1} - S_t R_{f,t} \\ S_t (z_t R_{f,t} - R_{f,t}) + \lambda S_t (R_{t+1} - z_t R_{f,t}) \end{cases} \quad \text{for } \begin{cases} R_{t+1} \geq z_t R_{f,t} \\ R_{t+1} < z_t R_{f,t} \end{cases}$$

and for  $z_t > 1$ ,

$$(16) \quad v(X_{t+1}, S_t, z_t) = \begin{cases} S_t R_{t+1} - S_t R_{f,t} \\ \lambda (z_t) (S_t R_{t+1} - S_t R_{f,t}) \end{cases} \quad \text{for } \begin{cases} R_{t+1} \geq R_{f,t} \\ R_{t+1} < R_{f,t} \end{cases}$$

with

$$(17) \quad \lambda(z_t) = \lambda + k(z_t - 1).$$

Equations (15) and (16) are pictured in Figure I. Finally, the dynamics of the state variable  $z_t$  are given by

$$(18) \quad z_{t+1} = \eta \left( z_t \frac{\bar{R}}{R_{t+1}} \right) + (1 - \eta)(1).$$

We calculate the price  $P_t$  of a dividend claim—in other words, the stock price—in two different economies. The first economy, which we call “Economy I,” is the one analyzed by Lucas [1978]. It equates consumption and dividends so that stocks are modeled as a claim to the future consumption stream.

Due to its simplicity, the first economy is the one typically studied in the literature. However, we also calculate stock prices in a more realistic economy—“Economy II”—where consumption and dividends are modeled as separate processes. We can then allow the volatility of consumption growth and of dividend growth to be very different, as they indeed are in the data. We can think of the difference between consumption and dividends as arising

from the fact that investors have other sources of income besides dividends. Equivalently, they have other forms of wealth, such as human capital, beyond their financial assets.

In our model, changes in risk aversion are caused by changes in the level of the stock market. In this respect, our approach differs from consumption-based habit formation models, where changes in risk aversion are due to changes in the level of consumption. While these are different ideas, it is not easy to illustrate their distinct implications in an economy like Economy I, where consumption and the stock market are driven by a single shock, and are hence perfectly conditionally correlated. This is why we emphasize Economy II: since consumption and dividends do not have to be equal in equilibrium, we can model them as separate processes, driven by shocks that are only imperfectly correlated. The contrast between our approach and the consumption-based framework then becomes much clearer.

In both economies, we construct a one-factor Markov equilibrium in which the risk-free rate is constant and the Markov state variable  $z_t$  determines the distribution of future stock returns. Specifically, we assume that the price-dividend ratio of the stock is a function of the state variable  $z_t$ :

$$(19) \quad f_t \equiv P_t/D_t = f(z_t),$$

and then show for each economy in turn that there is indeed an equilibrium satisfying this assumption. Given the one-factor assumption, the distribution of stock returns  $R_{t+1}$  is determined by  $z_t$  and the function  $f(\cdot)$  using

$$(20) \quad R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{1 + P_{t+1}/D_{t+1}}{P_t/D_t} \frac{D_{t+1}}{D_t} \\ = \frac{1 + f(z_{t+1})}{f(z_t)} \frac{D_{t+1}}{D_t}.$$

### A. Stock Prices in Economy I

In the first economy we consider, consumption and dividends are modeled as identical processes. We write the process for aggregate consumption  $\bar{C}_t$  as

$$(21) \quad \log(\bar{C}_{t+1}/\bar{C}_t) = \log(D_{t+1}/D_t) = g_C + \sigma_C \epsilon_{t+1},$$

where  $\epsilon_t \sim \text{i.i.d. } N(0,1)$ . Note from equation (1) that the mean  $g_D$  and volatility  $\sigma_D$  of dividend growth are constrained to equal  $g_C$

and  $\sigma_C$ , respectively. Together with the one-factor Markov assumption, this means that the stock return is given by

$$(22) \quad R_{t+1} = \frac{1 + f(z_{t+1})}{f(z_t)} e^{g_C + \sigma_C \epsilon_{t+1}}.$$

Intuitively, the value of the risky asset can change because of news about consumption  $\epsilon_{t+1}$  or because the price-dividend ratio  $f$  changes. Changes in  $f$  are driven by changes in  $z_t$ , which measures past gains and losses: past gains make the investor less risk averse, raising  $f$ , while past losses make him more risk averse, lowering  $f$ .

In equilibrium, and under rational expectations about stock returns and aggregate consumption levels, the agents in our economy must find it optimal to consume the dividend stream and to hold the market supply of zero units of the risk-free asset and one unit of stock at all times.<sup>17</sup> Proposition 1 characterizes the equilibrium.<sup>18</sup>

**PROPOSITION 1.** For the preferences given in (13)–(18), there exists an equilibrium in which the gross risk-free interest rate is constant at

$$(23) \quad R_f = \rho^{-1} e^{\gamma g_C - \gamma^2 \sigma_C^2 / 2},$$

and the stock's price-dividend ratio  $f(\cdot)$ , as a function of the state variable  $z_t$ , satisfies for all  $z_t$ :

$$(24) \quad 1 = \rho \mathbf{E}_t \left[ \frac{1 + f(z_{t+1})}{f(z_t)} e^{(1-\gamma)(g_C + \sigma_C \epsilon_{t+1})} \right] \\ + b_0 \rho \mathbf{E}_t \left[ \hat{v} \left( \frac{1 + f(z_{t+1})}{f(z_t)} e^{g_C + \sigma_C \epsilon_{t+1}}, z_t \right) \right],$$

where for  $z_t \leq 1$ ,

$$(25) \quad \hat{v}(R_{t+1}, z_t) \\ = \begin{cases} R_{t+1} - R_{f,t} \\ (z_t R_{f,t} - R_{f,t}) + \lambda(R_{t+1} - z_t R_{f,t}) \end{cases} \quad \text{for } \begin{cases} R_{t+1} \geq z_t R_{f,t} \\ R_{t+1} < z_t R_{f,t}, \end{cases}$$

17. We need to impose rational expectations about aggregate consumption because the agent's utility includes aggregate consumption as a scaling term.

18. We assume that  $\log \rho + (1 - \gamma)g_C + 0.5(1 - \gamma)^2 \sigma_C^2 < 0$  so that the equilibrium is well behaved at  $t = \infty$ .

and for  $z_t > 1$ ,

$$(26) \quad \hat{v}(R_{t+1}, z_t) = \begin{cases} R_{t+1} - R_{f,t} & \text{for } R_{t+1} \geq R_{f,t} \\ \lambda(z_t)(R_{t+1} - R_{f,t}) & \text{for } R_{t+1} < R_{f,t}. \end{cases}$$

We prove this formally in the Appendix. At a less formal level, our results follow directly from the agent's Euler equations for optimality at equilibrium, derived using standard perturbation arguments:

$$(27) \quad 1 = \rho R_f \mathbf{E}_t[(\bar{C}_{t+1}/\bar{C}_t)^{-\gamma}],$$

$$(28) \quad 1 = \rho \mathbf{E}_t[R_{t+1}(\bar{C}_{t+1}/\bar{C}_t)^{-\gamma}] + b_0 \rho \mathbf{E}_t[\hat{v}(R_{t+1}, z_t)].$$

Readers may find it helpful to compare these equations with those derived from standard asset pricing models with power utility over consumption. The Euler equation for the risk-free rate is the usual one: consuming a little less today and investing the savings in the risk-free rate does not change the investor's exposure to losses on the risky asset. The first term in the Euler equation for the risky asset is also the familiar one first obtained by Mehra and Prescott [1985]. However, there is now an additional term. Consuming less today and investing the proceeds in the risky asset exposes the investor to the risk of greater losses. Just how dangerous this is, is determined by the state variable  $z_t$ .

In constructing the equilibrium in Proposition 1, we follow the assumption laid out in subsection II.D, namely that buying or selling on the part of the investor does not affect the evolution of the state variable  $z_t$ . Equivalently, the investor believes that his actions will have no impact on the future evolution of  $z_t$ . As we argued earlier, this is a reasonable assumption for many actions the investor might take, but is less so in the case of a complete exit from the stock market. In essence, our assumption means that the investor does not consider using his cushion of prior gains in a strategic fashion, perhaps by waiting for the cushion to become large, exiting from the stock market so as to preserve the cushion and then reentering after a market crash when expected returns are high.<sup>19</sup>

19. Allowing the investor to consider such strategies does not change the qualitative nature of our results. It may affect our quantitative results depending on how one specifies the evolution of the investor's cushion of prior gains after he leaves the stock market.

### B. Stock Prices in Economy II

In Economy II, consumption and dividends follow distinct processes. This allows us to model the stock for what it really is, namely a claim to the dividend stream, rather than as a claim to consumption. Formally, we assume

$$(29) \quad \log(\bar{C}_{t+1}/\bar{C}_t) = g_C + \sigma_C \eta_{t+1},$$

and

$$(30) \quad \log(D_{t+1}/D_t) = g_D + \sigma_D \epsilon_{t+1},$$

where

$$(31) \quad \begin{pmatrix} \eta_t \\ \epsilon_t \end{pmatrix} \sim \text{i.i.d. } N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix}\right).$$

This assumption, which makes  $\log \bar{C}_t/D_t$  a random walk, allows us to construct a one-factor Markov equilibrium in which the risk-free interest rate is constant and the price-dividend ratio of the stock is a function of the state variable  $z_t$ .<sup>20</sup> The stock return can then be written as

$$(32) \quad R_{t+1} = \frac{1 + f(z_{t+1})}{f(z_t)} e^{g_D + \sigma_D \epsilon_{t+1}}.$$

Given that the consumption and dividend processes are different, we need to complete the model specification by assuming that each agent also receives a stream of nonfinancial income  $\{Y_t\}$ —labor income, say. We assume that  $\{Y_t\}$  and  $\{D_t\}$  form a joint Markov process whose distribution gives  $\bar{C}_t \equiv D_t + Y_t$  and  $D_t$  the distributions in (29)–(31).

We construct the equilibrium through the Euler equations of optimality (27) and (28). The risk-free rate is again constant and given by (27). The one-factor Markov structures of stock prices in (19) and (32) satisfy the Euler equation (28). The next proposition characterizes this equilibrium. The Appendix gives more detailed calculations and proves that the Euler equations indeed characterize optimality.<sup>21</sup>

20. Another approach would model  $\bar{C}_t$  and  $D_t$  as cointegrated processes, but we would then need at least one more factor to characterize equilibrium prices.

21. We assume that  $\log \rho - \gamma g_C + g_D + 0.5(\gamma^2 \sigma_C^2 - 2\gamma \omega \sigma_C \sigma_D + \sigma_D^2) < 0$  so that the equilibrium is well behaved at  $t = \infty$ .



PROPOSITION 2. In Economy II, the risk-free rate is constant at

$$(33) \quad R_f = \rho^{-1} e^{\gamma g_C - \gamma^2 \sigma_C^2 / 2},$$

and the stock's price-dividend ratio  $f(\cdot)$  is given by

$$(34) \quad 1 = \rho e^{g_D - \gamma g_C + \gamma^2 \sigma_C^2 (1 - \omega^2) / 2} \mathbf{E}_t \left[ \frac{1 + f(z_{t+1})}{f(z_t)} e^{(\sigma_D - \gamma \omega \sigma_C) \epsilon_{t+1}} \right] \\ + b_0 \rho \mathbf{E}_t \left[ \hat{v} \left( \frac{1 + f(z_{t+1})}{f(z_t)} e^{g_D + \sigma_D \epsilon_{t+1}}, z_t \right) \right],$$

where  $\hat{v}$  is defined in Proposition 1.

### C. Model Intuition

In Section V we solve for the price-dividend ratio numerically and use simulated data to show that our model provides a way of understanding a number of puzzling empirical features of aggregate stock returns. In particular, our model is consistent with a low volatility of consumption growth on the one hand, and a high mean and volatility of stock returns on the other, while maintaining a low and stable risk-free rate. Moreover, it generates long horizon predictability in stock returns similar to that observed in empirical studies and predicts a low correlation between consumption growth and stock returns.

It may be helpful to outline the intuition behind these results before moving to the simulations. Return volatility is a good place to start: how can our model generate returns that are more volatile than the underlying dividends? Suppose that there is a positive dividend innovation this period. This will generate a high stock return, increasing the investor's reserve of prior gains. This makes him less risk averse, since future losses will be cushioned by the prior gains, which are now larger than before. He therefore discounts the future dividend stream at a lower rate, giving stock prices an extra jolt upward. A similar story holds for a negative dividend innovation. It generates a low stock return, depleting prior gains or increasing prior losses. The investor is more risk averse than before, and the increase in risk aversion pushes prices still lower. The effect of all this is to make returns substantially more volatile than dividend growth.

The same mechanism also produces long horizon predictability. Put simply, since the investor's risk aversion varies over time depending on his investment performance, expected returns on

the risky asset also vary. To understand this in more detail, suppose once again that there is a positive shock to dividends. This generates a high stock return, which lowers the investor's risk aversion and pushes the stock price still higher, leading to a higher price-dividend ratio. Since the investor is less risk averse, subsequent stock returns will be lower on average. Price-dividend ratios are therefore inversely related to future returns, in exactly the way that has been documented by numerous studies, including Campbell and Shiller [1988] and Fama and French [1988b].

If changing loss aversion can indeed generate volatile stock prices, then we may also be able to generate a substantial equity premium. On average, the investor is loss averse, and fears the frequent drops in the stock market. He may therefore charge a high premium in return for holding the risky asset. Earlier research provides hope that this will be the case: Benartzi and Thaler [1995] analyze one-period portfolio choice for loss-averse investors—a partial equilibrium analysis where the stock market's high historical mean and volatility are exogenous—and find that these investors are unwilling to invest much of their wealth in stocks, even in the face of the large historical premium. This suggests that loss aversion may be a useful ingredient for equilibrium models trying to understand the equity premium.

Finally, our framework also generates stock returns that are only weakly correlated with consumption, as in the data.<sup>22</sup> To understand this, note that in our model, stock returns are made up of two components: one due to news about dividends, and the other to a change in risk aversion caused by movements in the stock market. Both components are ultimately driven by shocks to dividends, and so in our model, the correlation between returns and consumption is very similar to the correlation between dividends and consumption—a low number. This result distinguishes our approach from consumption-based habit formation models of the stock market such as Campbell and Cochrane [1999]. In those models, changes in risk aversion are caused by changes in consumption levels. This makes it inevitable that returns will be significantly correlated with consumption shocks, in contrast to what we find in the data.

Another well-known difficulty with consumption-based models is that attempts to make them match features of the stock

22. This is a feature that is unique to Economy II, which allows for a meaningful distinction between consumption and dividends.

market often lead to counterfactual predictions for the risk-free rate. For example, these models typically explain the equity premium with a high curvature  $\gamma$  of utility over consumption. However, this high  $\gamma$  also leads to a strong desire to smooth consumption intertemporally, generating high interest rates. Furthermore, the habit formation feature that many consumption-based models use to explain stock market volatility can also make interest rates counterfactually volatile.<sup>23</sup>

In our framework we use loss aversion over financial wealth fluctuations rather than a high curvature of utility over consumption to explain the equity premium. We do not therefore generate a counterfactually high interest rate. Moreover, since changes in risk aversion are driven by past stock market performance rather than by consumption, we can maintain a stable, indeed constant interest rate.

One feature that our model does share with consumption-based models like that of Campbell and Cochrane [1999] is contrarian expectations on the part of investors. Since stock prices in these models are high when investors are less risk averse, these are also times when investors require—and expect—lower returns than on average.

Durell [1999] has examined investor expectations about future stock market behavior and found evidence of extrapolative, rather than contrarian expectations. In other words, some investors appear to expect higher than average returns precisely at market peaks. Shiller [1999] presents results of a survey of investor expectations over the course of the U. S. bull market of the last few years. He finds no evidence of extrapolative expectations; but neither does he find evidence of contrarian expectations. It is not clear whether the samples used by Durell and Shiller are representative of the investing population, but they do suggest that the story in this paper—or indeed the consumption-based story—may not be a complete description of the facts.

#### *D. A Note on Aggregation*

The equilibrium pricing equations in subsections IV.A and IV.B are derived under the assumption that the investors in our

23. Campbell and Cochrane's paper [1999] is perhaps the only consumption-based model that avoids problems with the risk-free rate. A clever choice of functional form for the habit level over consumption enables them to use precautionary saving to counterbalance the strong desire to smooth consumption intertemporally.

economy are completely homogeneous. This is certainly a strong assumption. Investors may be heterogeneous along numerous dimensions, which raises the question of whether the intuition of our model still goes through once investor heterogeneity is recognized. For any particular form of heterogeneity, we need to check that loss aversion remains in the aggregate and, moreover, that aggregate loss aversion still varies with prior stock market movements. If these two elements are still present, our model should still be able to generate a high premium, volatility, and predictability.

One form of heterogeneity does aggregate satisfactorily: this is the case where investors have different wealth levels, but identical wealth to income ratios. We can model this by having several cohorts of investors, each cohort containing a continuum of equally wealthy investors. Since wealth is not a nontrivial state variable, all our results go through.<sup>24</sup>

There is reason to hope that our intuition will also survive other forms of heterogeneity. For example, it is possible that investors differ in the extent of their prior gains or losses, perhaps because they entered the stock market at different times. In other words, investors may have different  $z_t$ 's.

Note that even if  $z_t$  varies across investors, each individual investor is still more sensitive to losses than to gains, and there is no reason to believe that this will be lost in the aggregate. Hence there is no reason to think that the equity premium will be much reduced in the presence of this kind of heterogeneity. Furthermore, if the stock market experiences a sustained rise, this will increase prior gains for most investors, making them less risk averse. Therefore, it is very reasonable to think that risk aversion will also fall in the aggregate. If aggregate loss aversion still varies over time, our model should still be able to generate substantial volatility and predictability.

## V. NUMERICAL RESULTS AND FURTHER DISCUSSION

In this section we present price-dividend ratios  $f(z_t)$  that solve equations (24) and (34). We then create a long time series of simulated data and use it to compute various moments of asset

24. The only subtlety is that since aggregate consumption  $\bar{C}_t$  enters preferences, we need to assume that people use the average consumption of a reference group of people with identical wealth to set  $\bar{C}_t$ , rather than average consumption in the economy as a whole.

TABLE I  
PARAMETER VALUES FOR ECONOMY I

Parameter	
$g_C$	1.84%
$\sigma_C$	3.79%
$\gamma$	1.0
$\rho$	0.98
$\lambda$	2.25
$k$	(range)
$b_0$	(range)
$\eta$	0.9

returns which can be compared with historical numbers. We do this for both economies described in Section IV: Economy I, where stocks are modeled as a claim to the consumption stream; and the more realistic Economy II where stocks are a claim to dividends, which are no longer the same as consumption.

#### A. Parameter Values

Table I summarizes our choice of parameter values for Economy I. For  $g_C$  and  $\sigma_C$ , the mean and standard deviation of log consumption growth, we follow Cecchetti, Lam, and Mark [1990] who obtain  $g_C = 1.84$  percent and  $\sigma_C = 3.79$  percent from a time series of annual data from 1889 to 1985. These numbers are very similar to those used by Mehra and Prescott [1985] and Constantinides [1990].

The investor's preference parameters are  $\gamma$ ,  $\rho$ ,  $\lambda$ ,  $k$ , and  $b_0$ . We choose the curvature  $\gamma$  of utility over consumption and the time discount factor  $\rho$  so as to produce a sensibly low value for the risk-free rate. Given the values of  $g_C$  and  $\sigma_C$ , equation (23) shows that  $\gamma = 1.0$  and  $\rho = 0.98$  bring the risk-free interest rate close to  $R_f - 1 = 3.86$  percent.

The value of  $\lambda$  determines how keenly losses are felt relative to gains in the case where the investor has no prior gains or losses. This is the case that is most frequently studied in the experimental literature: Tversky and Kahneman [1992] estimate  $\lambda = 2.25$  by offering subjects isolated gambles, and we use this value.

The parameter  $k$  determines how much more painful losses are when they come on the heels of other losses. It is an important

determinant of the investor's *average* degree of loss aversion over time. In the results that we present, we pick  $k$  in two different ways. Our first approach is to choose  $k$  so as to make the investor's average loss aversion close to 2.25, where average loss aversion is computed in a way that we make precise in the Appendix. After prior gains, the investor does not fear losses very much, so his effective loss aversion is less than 2.25; after prior losses, he is all the more sensitive to additional losses, so his effective loss aversion is higher than 2.25, to a degree governed by  $k$ . We find that choosing  $k = 3$  keeps average loss aversion close to 2.25. To understand what a  $k$  of 3 means, suppose that the state variable  $z_t$  is initially equal to 1, and that the stock market then experiences a sharp fall of 10 percent. From equation (10) with  $\eta = 1$ , this means that  $z_t$  increases by approximately 0.1, to 1.1. From (8), any additional losses will now be penalized at  $2.25 + 3(0.1) = 2.55$ , a slightly more severe penalty.

Our second approach to picking  $k$  is to go to the data for guidance: we simply look for values of  $k$  that bring the predicted equity premium close to its empirical value.

The parameter  $b_0$  determines the relative importance of the prospect utility term in the investor's preferences. We do not have strong priors about what constitutes a reasonable value for  $b_0$  and so present results for a range of values.<sup>25</sup>

The two final parameters,  $\eta$  and  $\bar{R}$ , arise in the definition of the state variable dynamics.  $\bar{R}$  is not a parameter we have any control over: it is completely determined by the other parameters *and* the requirement that the equilibrium median value of  $z_t$  be equal to one. The variable  $\eta$  controls the persistence of  $z_t$  and hence also the persistence of the price-dividend ratio. We find that an  $\eta$  of 0.9 brings the autocorrelation of the price-dividend ratio that we generate close to its empirical value.

### B. Methodology

Before presenting our results, we briefly describe the way they were obtained. The identical technique is used for both Economy I and II, so we describe it only for the case of Economy I. The difficulty in solving equation (24) comes from the fact that

25. One way to think about  $b_0$  is to compare the disutility of losing a dollar in the stock market with the disutility of having to consume a dollar less. When computed at equilibrium, the ratio of these two quantities equals  $b_0\rho\lambda$ . By plugging numbers into this expression, we can see how  $b_0$  controls the relative importance of consumption utility and nonconsumption utility.

$z_{t+1}$  is a function of both  $\epsilon_{t+1}$  and  $f(\cdot)$ . In economic terms, our state variable is endogenous: it tracks prior gains and losses, which depend on past returns, themselves endogenous. Equation (24) is therefore self-referential and needs to be solved in conjunction with

$$(35) \quad z_{t+1} = \eta \left( z_t \frac{\bar{R}}{R_{t+1}} \right) + (1 - \eta)(1),$$

and

$$(36) \quad R_{t+1} = \frac{1 + f(z_{t+1})}{f(z_t)} e^{g_C + \sigma_C \epsilon_{t+1}}.$$

We use the following technique. We start out by guessing a solution to (24),  $f^{(0)}$  say. We then construct a function  $h^{(0)}$  so that  $z_{t+1} = h^{(0)}(z_t, \epsilon_{t+1})$  solves equations (35) and (36) for this  $f = f^{(0)}$ . The function  $h^{(0)}$  determines the distribution of  $z_{t+1}$  conditional on  $z_t$ .

Given the function  $h^{(0)}$ , we get a new candidate solution  $f^{(1)}$  through the following recursion:

$$(37) \quad 1 = \rho \mathbf{E}_t \left[ \frac{1 + f^{(i)}(z_{t+1})}{f^{(i+1)}(z_t)} e^{(1-\gamma)(g_C + \sigma_C \epsilon_{t+1})} \right] \\ + b_0 \rho \mathbf{E}_t \left[ \hat{v} \left( \frac{1 + f^{(i)}(z_{t+1})}{f^{(i+1)}(z_t)} e^{g_C + \sigma_C \epsilon_{t+1}}, z_t \right) \right], \quad \forall z_t.$$

With  $f^{(1)}$  in hand, we can calculate a new  $h = h^{(1)}$  that solves equations (35) and (36) for  $f = f^{(1)}$ . This  $h^{(1)}$  gives us a new candidate  $f = f^{(2)}$  from (37). We continue this process until convergence occurs:  $f^{(i)} \rightarrow f$ , and  $h^{(i)} \rightarrow h$ .

### C. Stock Prices in Economy I

Figure III presents price-dividend ratios  $f(z_t)$  that solve equation (24) for three different values of  $b_0$ : 0.7, 2, and 100, and with  $k$  fixed at 3. Note that  $f(z_t)$  is a decreasing function of  $z_t$  in all cases. The intuition for this is straightforward: a low value of  $z_t$  means that recent returns on the asset have been high, giving the investor a reserve of prior gains. These gains cushion subsequent losses, making the investor less risk averse. He therefore discounts future dividends at a lower rate, raising the price-dividend ratio. Conversely, a high value of  $z_t$  means that the

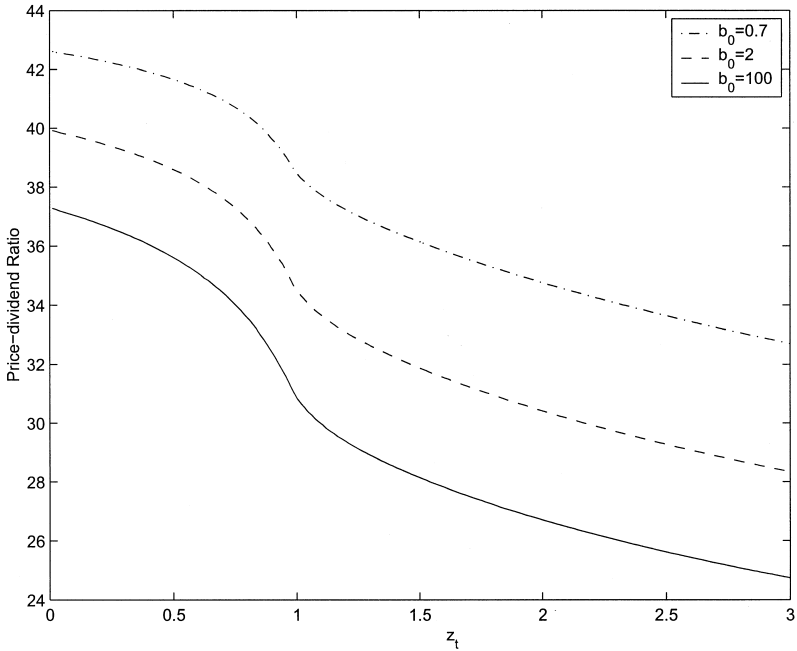


FIGURE III  
Price-Dividend Ratios in Economy I

The price-dividend ratios are plotted against  $z_t$ , which measures prior gains and losses: a low  $z_t$  indicates prior gains. The parameter  $b_0$  controls how much the investor cares about financial wealth fluctuations. We fix the parameter  $k$  at 3, bringing average loss aversion close to 2.25 in all cases.

investor has recently experienced a spate of painful losses; he is now especially sensitive to further losses which makes him more risk averse and leads to lower price-dividend ratios.

Figure III by itself does not tell us the range of price-dividend ratios we are likely to see in equilibrium. For that, we need to know the equilibrium distribution of the state variable  $z_t$ . Figure IV shows this distribution for one case that we will consider in more detail later:  $b_0 = 2$ ,  $k = 3$ . To obtain it, we draw a long time series  $\{\epsilon_t\}_{t=1}^{50,000}$  of 50,000 independent draws from the standard normal distribution and starting with  $z_0 = 1$ , use the function  $z_{t+1} = h(z_t, \epsilon_{t+1})$  described in subsection V.B to generate a time series for  $z_t$ . Note from the graph that the average  $z_t$  is close to one, and this is no accident. The value of  $\bar{R}$  in equation (10) is chosen precisely to make the median value of  $z_t$  as close to one as possible.



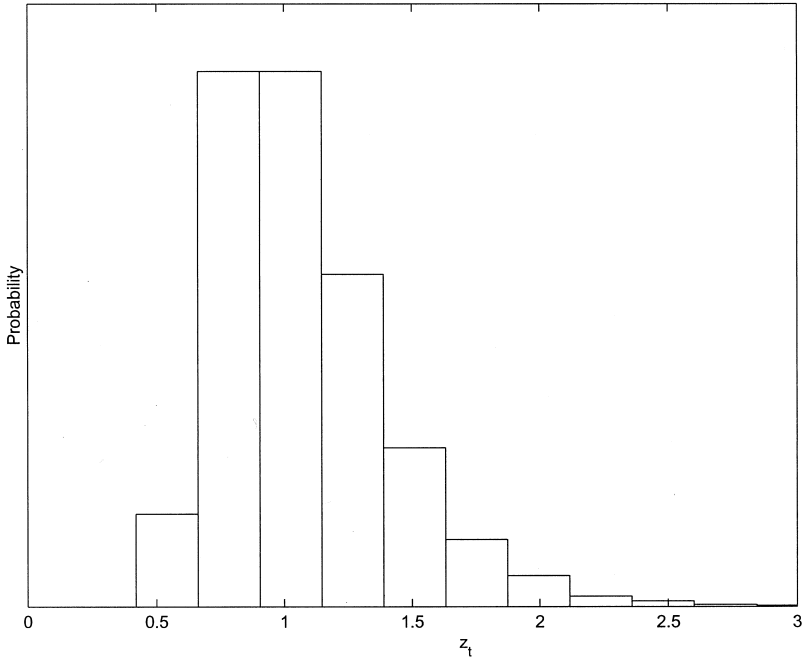


FIGURE IV  
 Distribution of the State Variable  $z_t$   
 The distribution is based on Economy I with  $b_0 = 2$  and  $k = 3$ .

As we generate the time series for  $z_t$  period by period, we also compute the returns along the way using equation (20). We now present sample moments computed from these simulated returns. The time series is long enough that sample moments should serve as good approximations to population moments.

Table II presents the important moments of stock returns for different values of  $b_0$  and  $k$ . In the top panel, we vary  $b_0$  and set  $k$  to 3, which keeps average loss aversion over time close to 2.25. At one extreme we have  $b_0 = 0$ , the classic case considered by Mehra and Prescott [1985]. As we push  $b_0$  up, the asset return moments eventually reach a limit that is well approximated with a  $b_0$  of 100. The table also reports the investor’s average loss aversion, calculated in the way described in the Appendix.

Note that as we raise  $b_0$  while keeping  $k$  fixed, the equity premium goes up. There are two forces at work here. As  $b_0$  gets larger, prior outcomes affect the investor more, causing his risk

TABLE II  
ASSET RETURNS IN ECONOMY I

	$b_0 = 0$	$b_0 = 0.7$ $k = 3$	$b_0 = 2$ $k = 3$	$b_0 = 100$ $k = 3$	Empirical value
Log risk-free rate	3.79	3.79	3.79	3.79	0.58
Log excess stock return					
Mean	0.07	0.63	0.88	1.26	6.03
Standard deviation	3.79	4.77	5.17	5.62	20.02
Sharpe ratio	0.02	0.13	0.17	0.22	0.3
Average loss aversion		2.25	2.25	2.25	
		$b_0 = 0.7$ $k = 150$	$b_0 = 2$ $k = 100$	$b_0 = 100$ $k = 50$	
Log risk-free rate		3.79	3.79	3.79	
Log excess stock return					
Mean		3.50	3.66	3.28	
Standard deviation		10.43	10.22	9.35	
Sharpe ratio		0.34	0.36	0.35	
Average loss aversion		10.7	7.5	4.4	

Moments of asset returns are expressed as annual percentages. Empirical values are based on Treasury Bill and NYSE data from 1926–1995. The parameter  $b_0$  controls how much the investor cares about financial wealth fluctuations, while  $k$  controls the increase in loss aversion after a prior loss.

aversion to vary more, and hence generating more volatile stock returns. Moreover, as  $b_0$  grows, loss aversion becomes a more important feature of the investor's preferences, pushing up the Sharpe ratio. The higher volatility and higher Sharpe ratio combine to raise the equity premium.

Although the results in Table II are encouraging from a qualitative standpoint, the magnitudes are not impressive. Changes in risk aversion do give returns a volatility higher than the 3.79 percent volatility assumed for dividend growth, but this effect is not nearly large enough to match historical volatility. Since the investor does not observe any particularly large market crashes, he does not charge a particularly high equity premium either. The highest equity premium we can possibly generate for this value of  $k$  is 1.28 percent.

The bottom panel of Table II shows that we can come closer to matching the historical equity premium by increasing the value of  $k$ . Since the investor is now extremely loss averse in some states of the world, average loss aversion also climbs steeply. Note, however, that return volatility here is still far too low.

TABLE III  
PARAMETER VALUES FOR ECONOMY II

Parameter	Economy II	Economy I
$g_C$	1.84%	1.84%
$g_D$	1.84%	—
$\sigma_C$	3.79%	3.79%
$\sigma_D$	12.0%	—
$\omega$	0.15	—
$\gamma$	1.0	1.0
$\rho$	0.98	0.98
$\lambda$	2.25	2.25
$k$	(range)	(range)
$b_0$	(range)	(range)
$\eta$	0.9	0.9

The parameter values used for Economy I and shown in Table I are repeated here for ease of comparison.

An unrealistic feature of Economy I is that consumption and dividends are constrained to follow the same process. This means that we are modeling stocks as a claim to a very smooth consumption stream, rather than as what they really are, namely a claim to a far more volatile dividend stream. We therefore turn to Economy II which allows us to relax this constraint.

#### *D. Stock Prices in Economy II*

We now calculate stock prices in a more general economy where consumption and dividends are modeled as distinct processes. Table III presents our choice of parameters in this economy. To make comparison easier, we also show the parameters used for Economy I alongside.

Allowing consumption and dividends to follow different processes introduces three new parameters:  $g_D$ , the mean dividend growth rate;  $\sigma_D$ , the volatility of dividend growth; and  $\omega$ , the correlation of shocks to dividend growth and consumption growth. For simplicity, we set  $g_D = g_C = 0.0184$ . Using NYSE data from 1926–1995 from CRSP, we find  $\sigma_D = 0.12$ . Campbell [2000] estimates  $\omega$  in a time series of U. S. data spanning the past century, and based on his results, we set  $\omega = 0.15$ . As the table shows, we keep all other parameters fixed at the values discussed in subsection V.A.

Figure V presents price-dividend ratios that solve equation

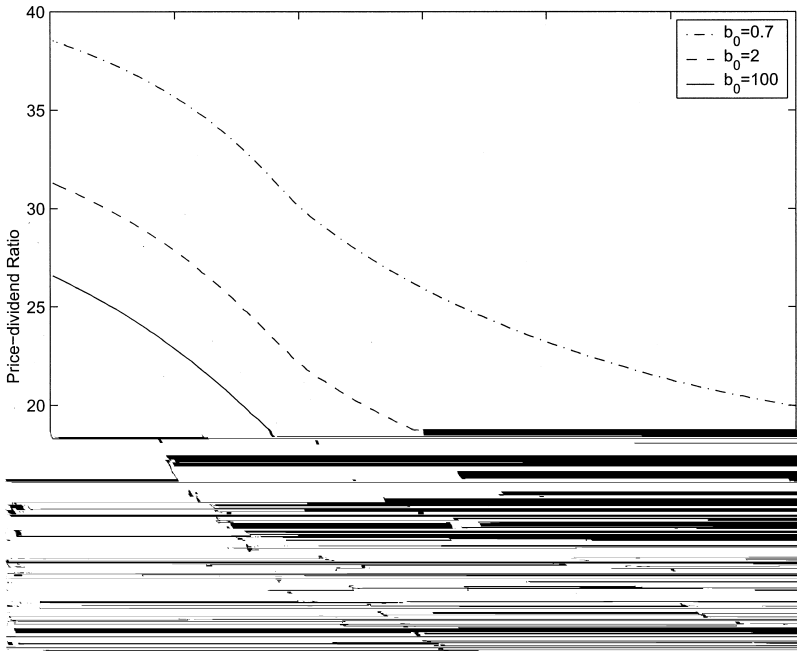


FIGURE V  
Price-Dividend Ratios in Economy II

The price-dividend ratios are plotted against  $z_t$ , which measures prior gains and losses: a low  $z_t$  indicates prior gains. The parameter  $b_0$  controls how much the investor cares about financial wealth fluctuations. We fix the parameter  $k$  at 3, bringing average loss aversion close to 2.25 in all cases.

(34) for three different values of  $b_0$ : 0.7, 2, and 100, with  $k$  fixed at 3. Unconditional moments of stock returns are shown in Table IV for various values of  $b_0$  and  $k$ . The top panel keeps  $k$  fixed at 3, bringing the average degree of loss aversion close to 2.25. A quick comparison with Table II shows that separating consumption and dividends improves the results significantly. The volatility of returns is now much higher than what we obtained in Economy I because dividend growth volatility is now 12 percent rather than just 3.79 percent. The equity premium is also much higher: the investor now sees far more severe downturns in the stock market and charges a much larger premium as compensation.

Modeling consumption and dividends separately also leads to higher return volatility in consumption-based models such as that of Campbell and Cochrane [1999]. However, the fact that the

TABLE IV  
ASSET PRICES AND RETURNS IN ECONOMY II

	$b_0 = 0$	$b_0 = 0.7$ $k = 3$	$b_0 = 2$ $k = 3$	$b_0 = 100$ $k = 3$	Empirical value
Log risk-free rate	3.79	3.79	3.79	3.79	0.58
Log excess stock return					
Mean	-0.65	1.3	2.62	3.68	6.03
Standard deviation	12.0	17.39	20.87	20.47	20.02
Sharpe ratio	-0.05	0.07	0.13	0.18	0.3
Correlation w/consumption growth	0.15	0.15	0.15	0.15	0.1
Price-dividend ratio					
Mean	76.6	29.8	22.1	17.5	25.5
Standard deviation	0	2.9	2.7	2.4	7.1
Average loss aversion		2.25	2.25	2.25	

	$b_0 = 0.7$ $k = 20$	$b_0 = 2$ $k = 10$	$b_0 = 100$ $k = 8$
Log risk-free rate	3.79	3.79	3.79
Log excess stock return			
Mean	5.17	5.02	5.88
Standard deviation	25.85	23.84	24.04
Sharpe ratio	0.2	0.21	0.24
Correlation w/consumption growth	0.15	0.15	0.15
Price-dividend ratio			
Mean	14.0	14.6	12.7
Standard deviation	2.6	2.5	2.2
Average loss aversion	5.8	3.5	3.2

Moments of asset returns are expressed as annual percentages. Empirical values are based on Treasury Bill and NYSE data from 1926–1995. The parameter  $b_0$  controls how much the investor cares about financial wealth fluctuations, while  $k$  controls the increase in loss aversion after a prior loss.

equity premium goes up is unique to our model: as Campbell and Cochrane demonstrate, separating consumption and dividends has *no effect* on the equity premium in a consumption-based model. The reason is that in a consumption-based model, the canonical measure of a stock's risk is its covariance with consumption growth. By separating consumption and dividends, the volatility of stock returns goes up, but the correlation of stock returns with consumption goes down, since dividends are poorly correlated with consumption. Overall, the covariance with consumption remains unaffected, and the equity premium remains as hard to explain as before.

The investor in our model cares not only about consumption but also about fluctuations in the value of his investments. An increase in dividend volatility makes stocks more volatile, scaring the investor into charging a higher equity premium. It is true that stocks are now less correlated with consumption, but this does not matter in our model, since the investor cares about fluctuations in the stock market per se, not merely about how those fluctuations covary with consumption growth.

Table IV confirms that stock returns are only weakly correlated with consumption growth in our model. As discussed earlier, this prediction is unique to our framework: since consumption-based models ascribe a significant fraction of price volatility to changes in consumption, stock returns are inevitably highly correlated with consumption growth.

We also report the mean and standard deviation of the simulated price-dividend ratio. It is striking that while we *are* successful at matching the volatility of returns, we significantly underpredict the volatility of the price-dividend ratio. To understand how this can be, it is helpful to consider the following approximate relationship, derived by Campbell, Lo, and MacKinlay [1997] with the help of a log-linear approximation<sup>26</sup>:

$$r_{t+1} \approx A + \log f_{t+1} - \log f_t + \sigma_D \epsilon_{t+1}.$$

In words, up to a constant  $A$ , the log return  $r_{t+1}$  is approximately equal to the change in the log price-dividend ratio plus the innovation to dividend growth  $\sigma_D \epsilon_{t+1}$ . This gives

$$\text{var}(r_{t+1}) \approx \text{var}\left(\log \frac{f_{t+1}}{f_t}\right) + \sigma_D^2 + 2 \text{cov}\left(\log \frac{f_{t+1}}{f_t}, \sigma_D \epsilon_{t+1}\right).$$

Our simulated price-dividend ratio is not sufficiently volatile, making  $\text{var}(\log(f_{t+1}/f_t))$  too low in our model relative to what it is in the data. The reason we are still able to match the volatility of returns  $\text{var}(r_{t+1})$  is that  $\text{cov}(\log(f_{t+1}/f_t), \sigma_D \epsilon_{t+1})$  is too *high* in our model relative to what it is in the data: in our framework, changes in price-dividend ratios are perfectly conditionally correlated with dividend shocks.

Since our model relies on only one factor to generate movement in the price-dividend ratio, it is not surprising that we

26. More specifically, this follows from equation 7.1.19 in Chapter 7 of the book.

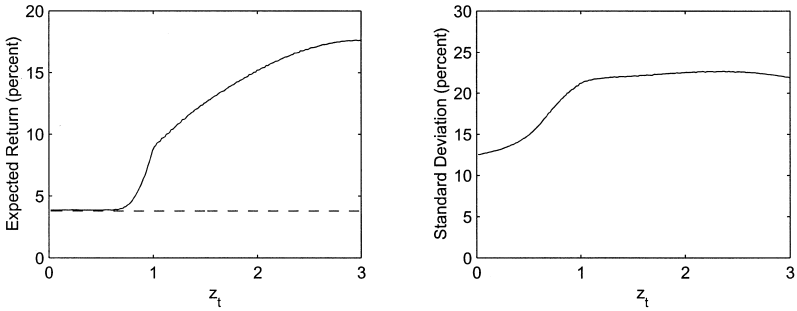


FIGURE VI  
 Conditional Moments of Stock Returns

The conditional stock return mean and volatility are plotted against  $z_t$ , which measures prior gains and losses: a low  $z_t$  indicates prior gains. The results are for Economy II with  $b_0 = 2$  and  $k = 3$ . The dashed line in the left-hand panel indicates the level of the risk-free rate.

underpredict its volatility. In a more realistic model, other factors will also affect the price-dividend ratio: consumption relative to habit is one possible factor suggested by the habit formation literature. Adding such factors will increase the volatility of the price-dividend ratio and decrease its correlation with dividend shocks, improving the model’s fit with the data. It is worth emphasizing, though, that a *purely* consumption-based approach will not fare well, since it will predict that changes in price-dividend ratios are perfectly conditionally correlated with consumption shocks, and that returns are highly correlated with consumption. A model that merges habit formation over consumption with a direct concern about financial wealth fluctuations may be much more fruitful.

The bottom panel in Table IV shows that we can significantly improve our results by increasing  $k$ , which increases investors’ loss aversion after prior losses. In Economy I, we had to make the investor extraordinarily loss averse in some states of the world to even come close to matching the equity premium. Interestingly, the increases in  $k$  we need now are much more modest. For  $b_0 = 2$ , for example, a  $k$  of 10 is enough to give a premium of 5.02 percent and a volatility of 23.84 percent; note also that this corresponds to an average loss aversion of only 3.5; this is not a small level of risk aversion, but neither is it extreme.

Figure VI presents some additional results of interest. The left panel plots the conditional expected stock return as a function

TABLE V  
AUTOCORRELATIONS OF LOG RETURNS AND PRICE-DIVIDEND RATIOS IN ECONOMY II

	$b_0 = 2$ $k = 3$	$b_0 = 2$ $k = 10$	Empirical value
$\text{corr}(r_t, r_{t-j})$			
$j = 1$	-0.07	-0.12	0.07
$j = 2$	-0.05	-0.09	-0.17
$j = 3$	-0.04	-0.06	-0.05
$j = 4$	-0.04	-0.04	-0.11
$j = 5$	-0.02	-0.03	-0.04
$\text{corr}((P/D)_t, (P/D)_{t-j})$			
$j = 1$	0.81	0.72	0.70
$j = 2$	0.66	0.52	0.50
$j = 3$	0.53	0.38	0.45
$j = 4$	0.43	0.28	0.43
$j = 5$	0.35	0.20	0.40

Empirical values are based on NYSE data from 1926–1995.

of  $z_t$ , obtained by numerically integrating the return equation (32) over the conditional distribution of  $z_{t+1}$  given by  $z_{t+1} = h(z_t, \epsilon_{t+1})$ . The value of  $b_0$  here is 2 and  $k$  is 3. The conditional expected return is an increasing function of the state variable. Low values of  $z_t$  mean that the investor has accumulated prior gains that will cushion future losses. He is therefore less risk averse, leading to a lower expected return in equilibrium. The dashed line shows the level of the constant risk-free rate for comparison.

The right panel in Figure VI graphs the conditional volatility of returns as a function of the state variable. Since much of the return volatility in our model is generated by changing risk aversion, the conditional volatility in any state depends on how sensitive the investor's risk aversion in that state is to dividend shocks. Empirically, volatility has been found to be higher after market crashes than booms, which in our context would mean an upward sloping conditional volatility curve. For much of the range of the state variable, this is exactly what we find. However, this result is sensitive to how we make the degree of loss aversion depend on prior outcomes, so we do not attach too much weight to it.

Table V presents autocorrelations of log returns and of the price-dividend ratio. As expected, our model produces negatively autocorrelated returns at all lags: high prices lower risk aversion



TABLE VI  
RETURN PREDICTABILITY REGRESSIONS IN ECONOMY II

$\beta_j, R^2(j)$	$b_0 = 2$ $k = 3$	$b_0 = 2$ $k = 10$	Empirical value
$\beta_1$	4.6	4.4	4.2
$\beta_2$	8.3	7.5	8.7
$\beta_3$	11.6	9.7	12.1
$\beta_4$	13.7	11.5	15.9
$R^2(1)$	2%	6%	7%
$R^2(2)$	4%	10%	16%
$R^2(3)$	5%	12%	22%
$R^2(4)$	6%	14%	30%

Coefficients and  $R^2$  in regressions of  $j$ -year cumulative log returns on the lagged dividend-price ratio,  $r_{t+1} + r_{t+2} + \dots + r_{t+j} = \alpha_j + \beta_j(D_t/P_t) + \epsilon_{j,t}$ . Empirical values are based on annual NYSE data from 1926–1995.

and lead to lower returns on average. These negative autocorrelations imply long horizon mean reversion of the kind documented by Poterba and Summers [1988] and Fama and French [1988a]. Moreover, the price-dividend ratio is highly autocorrelated in our model, closely matching its actual behavior.

Since the investor’s risk aversion changes over time, expected returns also vary, and hence returns are predictable. To demonstrate this, we use our simulated data to run regressions of cumulative log returns over a  $j$ -year horizon on the lagged dividend-price ratio for  $j = 1, 2, 3,$  and  $4$ :

$$(38) \quad r_{t+1} + r_{t+2} + \dots + r_{t+j} = \alpha_j + \beta_j(D_t/P_t) + \epsilon_{j,t}.$$

where  $r_t$  is the log return. Table VI presents the slope coefficients  $\beta_j$  and  $R^2(j)$  obtained from our simulated data alongside the empirical values. Note that our simulated results capture the main features of the empirical findings, including an  $R^2$  that increases with the return horizon.

*E. Sensitivity Analysis*

We now analyze the sensitivity of our results to various parameters of interest. For each parameter that we vary, the other parameters are kept fixed at the values given in Table III.

Table VII shows the effect of varying  $k$ , which governs how much loss aversion goes up after prior losses. Raising  $k$  has a large effect on the equity premium since it raises average loss

TABLE VII  
SENSITIVITY OF ASSET RETURNS TO  $k$

	$k = 3$	$k = 5$	$k = 10$	Empirical value
Log excess stock return				
Mean	2.62	3.15	5.02	6.03
Standard deviation	20.87	20.93	23.84	20.02
Sharpe ratio	0.13	0.15	0.21	0.3
Average loss aversion	2.25	2.6	3.5	

The parameter  $k$  controls how much loss aversion increases after a prior loss. Moments of asset returns are expressed as annual percentages. The results are for Economy II with  $b_0 = 2$ ; other parameters are fixed at the values in Table III.

aversion. It also raises volatility somewhat because a higher  $k$  means more rapid *changes* in risk aversion.

Throughout our analysis we have fixed  $\lambda$  at 2.25 since many independent experimental studies have estimated it at around this level. Table VIII shows how the results change as we vary  $\lambda$ . An increase in  $\lambda$  raises the equity premium because average loss aversion is now higher.

Table IX examines the effect of varying  $\eta$ . This parameter determines how far into the future the investor will be affected by a substantial gain or loss this year. The primary effect of  $\eta$  is on volatility: if  $\eta$  is high, a loss this year will make the investor more loss averse for many years to come, leading him to discount cash flows at a higher rate many years into the future, and causing a more dramatic price drop today. More directly,  $\eta$  affects the

TABLE VIII  
SENSITIVITY OF ASSET RETURNS TO  $\lambda$

	$\lambda = 1.5$	$\lambda = 2.25$	$\lambda = 3$	Empirical value
Log excess stock return				
Mean	3.8	5.02	5.6	6.03
Standard deviation	25.68	23.84	23.21	20.02
Sharpe ratio	0.15	0.21	0.24	0.3
Average loss aversion	3.2	3.5	3.9	

The parameter  $\lambda$  controls the investor's loss aversion. Moments of asset returns are expressed as annual percentages. The results are for Economy II with  $b_0 = 2$  and  $k = 10$ ; other parameters are fixed at the values in Table III.

TABLE IX  
SENSITIVITY OF ASSET RETURNS TO  $\eta$

	$\eta = 1$	$\eta = 0.9$	$\eta = 0.8$	Empirical value
Log excess stock return				
Mean	7.68	5.02	3.91	6.03
Standard deviation	34.54	23.84	19.12	20.02
Sharpe ratio	0.22	0.21	0.2	0.3
Autocorrelation of price-dividend ratio	0.81	0.72	0.65	0.7
Average loss aversion	4.5	3.5	3.0	

The parameter  $\eta$  governs how long-lasting the effects of prior gains and losses are. Moments of asset returns are expressed as annual percentages. The results are for Economy II with  $b_0 = 2$  and  $k = 10$ ; other parameters are fixed at the values in Table III.

persistence of the state variable and hence the autocorrelation of the price-dividend ratio.

So far in our analysis, we have taken the length of the period over which the investor evaluates gains and losses to be one year. Table X analyzes the effect of changing this. It shows that the length of the evaluation period mainly affects the equity premium: if the investor evaluates more frequently, he is more likely to experience losses, and since he is averse to losses, he will charge a higher premium.

As described in Section II, we suppose that the investor uses the risk-free rate as a reference level when calculating gains and losses. Table XI studies the sensitivity of our results to that

TABLE X  
SENSITIVITY OF ASSET RETURNS TO THE EVALUATION PERIOD

	6 months	1 year	2 years	Empirical value
Log excess stock return				
Mean	7.63	5.02	2.85	6.03
Standard deviation	27.78	23.84	20.15	20.02
Sharpe ratio	0.27	0.21	0.14	0.3
Average loss aversion	3.5	3.5	3.6	

The evaluation period is the length of time over which the investor measures his gains and losses. Moments of asset returns are expressed as annual percentages. The results are for Economy II with  $b_0 = 2$  and  $k = 10$ ; other parameters are fixed at the values in Table III.

TABLE XI  
SENSITIVITY OF ASSET RETURNS TO THE REFERENCE LEVEL

	0%	-1%	-2%	Empirical value
Log excess stock return				
Mean	5.02	4.11	3.43	6.03
Standard deviation	23.84	24.25	25.20	20.02
Sharpe ratio	0.21	0.17	0.14	0.3
Average loss aversion	3.5	3.5	3.7	

A "0%" reference level means that the investor compares the stock return with the risk-free rate when measuring gains and losses. "-1%" means a reference level one percent lower than the risk-free rate. Moments of asset returns are expressed as annual percentages. The results are for Economy II with  $b_0 = 2$  and  $k = 10$ ; other parameters are fixed at the values in Table III.

assumption. Given the parameter values in Table III, the risk-free rate is  $R_f - 1 = 3.86$  percent. The table shows the effect of using a reference level that is one or two percentage points lower than this. Note that the equity premium falls as we use a lower reference level: a lower reference point means that the investor is less likely to code stock market movements as losses, and hence less inclined to charge a high premium for holding stocks.

## VI. THE IMPORTANCE OF PRIOR OUTCOMES

The model in Section II makes use of both loss aversion and the effect of prior outcomes to match asset prices. The reader may wonder whether both these ingredients are truly necessary. After all, Benartzi and Thaler [1995] show that a loss-averse investor is very reluctant to allocate much of his portfolio to stocks even when faced with the large historical equity premium. This suggests that perhaps an equilibrium model with loss aversion *alone* would be enough to understand the data. Is it really necessary to incorporate the effect of prior outcomes?

We answer this by examining the predictions of a version of the model of Section II which ignores the effect of prior outcomes. In particular, we make the utility  $v$  of a gain or loss a function of the gain or loss  $X_{t+1}$  *alone*, and remove the dependence on the state variable  $z_t$ . In this model, the degree of loss aversion is the same in all circumstances, regardless of the investor's prior investment performance.

More formally, the investor chooses consumption  $C_t$  and an allocation to the risky asset  $S_t$  to maximize

$$(39) \quad \mathbf{E} \left[ \sum_{t=0}^{\infty} \left( \rho^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_t \rho^{t+1} v(X_{t+1}) \right) \right],$$

subject to the standard budget constraint, where

$$(40) \quad X_{t+1} = S_t R_{t+1} - S_t R_{f,t}$$

and

$$(41) \quad v(X_{t+1}) = \begin{cases} X_{t+1} & \text{for } X_{t+1} \geq 0 \\ \lambda X_{t+1} & \text{for } X_{t+1} < 0. \end{cases}$$

The next proposition presents the equations that govern equilibrium prices in an economy like Economy II where consumption and dividends are separated out. We show that there is an equilibrium in which the risk-free rate and the stock's price-dividend ratio are both constant and stock returns are i.i.d.

PROPOSITION 3. For the preferences given by (39)–(41), there exists an equilibrium in which the gross risk-free interest rate is constant at

$$(42) \quad R_f = \rho^{-1} e^{\gamma g_C - \gamma^2 \sigma_C^2 / 2},$$

and the stock's price-dividend ratio,  $f_t$ , is constant at  $f$  and given by

$$(43) \quad 1 = \rho e^{g_D - \gamma g_C + \gamma^2 \sigma_C^2 (1 - \omega^2) / 2} \frac{1+f}{f} \mathbf{E}_t [e^{(\sigma_D - \gamma \omega \sigma_C) \epsilon_{t+1}}] + b_0 \rho \mathbf{E}_t \left[ \hat{v} \left( \frac{1+f}{f} e^{g_D + \sigma_D \epsilon_{t+1}} \right) \right],$$

where

$$(44) \quad \hat{v}(R_{t+1}) = \begin{cases} R_{t+1} - R_{f,t} & \text{for } R_{t+1} \geq R_{f,t} \\ \lambda(R_{t+1} - R_{f,t}) & \text{for } R_{t+1} < R_{f,t}. \end{cases}$$

That returns are i.i.d. is a direct consequence of the fact that the price-dividend ratio is constant. To see this, note that the stock return is related to the stock's price-dividend ratio, denoted by  $f_t \equiv P_t/D_t$ , as follows:

$$(45) \quad R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{1 + P_{t+1}/D_{t+1}}{P_t/D_t} \frac{D_{t+1}}{D_t} = \frac{1 + f_{t+1}}{f_t} \frac{D_{t+1}}{D_t}.$$

TABLE XII  
PARAMETER VALUES FOR A MODEL WHERE PRIOR OUTCOMES HAVE NO EFFECT

Parameter	Section VI	Subsection V.D
$g_C$	1.84%	1.84%
$g_D$	1.84%	1.84%
$\sigma_C$	3.79%	3.79%
$\sigma_D$	12.0%	12.0%
$\gamma$	1.0	1.0
$\rho$	0.98	0.98
$\lambda$	2.25	2.25
$k$	—	(range)
$b_0$	(range)	(range)
$\eta$	—	0.9

Parameter values used for Economy II in subsection V.D are presented alongside for comparison.

Given the assumption that the dividend growth is i.i.d. (see (1)), a constant price-dividend ratio  $f_t = f$  implies that stock returns are i.i.d.

Table XII summarizes our choices of parameter values. Many of the parameters of the earlier model now no longer play a role. The parameters that remain are  $g_C$ ,  $g_D$ ,  $\sigma_C$ ,  $\sigma_D$ ,  $\gamma$ ,  $\rho$ ,  $\lambda$ , and  $b_0$ , and we assign them exactly the same values as those chosen earlier. For comparison, we list the earlier parameters alongside.

Figure VII presents the implied values of the price-dividend ratio and equity premium for different values of  $b_0$  and Table XIII

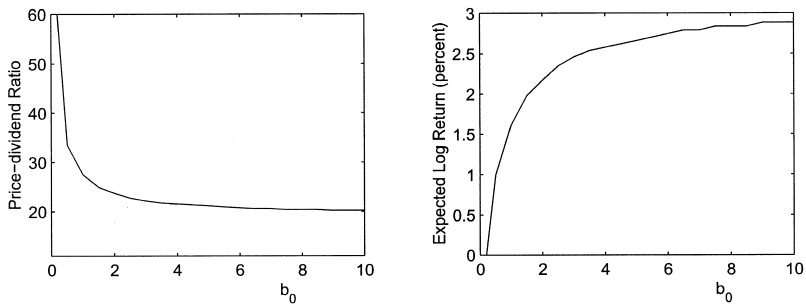


FIGURE VII

Stock Prices and Returns When Prior Outcomes Have No Effect

The price-dividend ratio and the equity premium are plotted against  $b_0$ , which controls how much the investor cares about financial wealth fluctuations.

TABLE XIII  
ASSET RETURNS WHEN PRIOR OUTCOMES HAVE NO EFFECT

	$b_0 = 0$	$b_0 = 0.7$	$b_0 = 2$	$b_0 = 100$	Empirical value
Log risk-free rate	3.79	3.79	3.79	3.79	0.58
Log excess stock return					
Mean	-0.65	1.30	2.17	2.88	6.03
Standard deviation	12.0	12.0	12.0	12.0	20.02
Sharpe ratio	-0.05	0.11	0.18	0.24	0.3

Moments of asset returns are expressed as annual percentages. Empirical values are based on Treasury Bill and NYSE data from 1926–1995. The parameter  $b_0$  controls how much the investor cares about financial wealth fluctuations.

reports unconditional moments of returns. By comparing the results in Table XIII with those in the bottom panel of Table IV, we can isolate the impact of prior outcomes.

Our results suggest that a model which relies on loss aversion alone cannot provide a complete description of aggregate stock market behavior. The fundamental weakness of such a model comes in explaining volatility: the standard deviations of returns in Table XIII are much lower than both the empirical value and the values in Table IV. The reason for this failure is straightforward. Since

$$(46) \quad R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{1 + f_{t+1}}{f_t} \frac{D_{t+1}}{D_t} = \frac{1 + f}{f} e^{gD + \sigma_D \epsilon_{t+1}},$$

the volatility of log returns in this model is equal to the volatility of log dividend growth, namely 12 percent. This problem is not unique to the model of this section, but is shared by all consumption-based models with constant discount rates.<sup>27</sup>

The unrealistically low stock return volatility generated by the model also hampers its ability to explain the equity premium: the average stock returns in Table XIII are considerably lower than the empirical value. Intuitively, even though our investor is loss averse, the stock market fluctuations are not large enough to scare him into demanding a high equity premium as compensation. More generally, our results show that loss aversion can be helpful in understanding the equity premium, but only in con-

27. This category includes papers that use first-order risk-aversion, such as Epstein and Zin [1990]. See Campbell [2000] for a discussion of this issue.

junction with a mechanism that makes stock returns more volatile than underlying cash flows. In our model, this mechanism is a changing degree of loss aversion, but it could also be based on changing perceptions of risk, or excessive extrapolation of cash flows into the future.

## VII. CONCLUSION

In this paper we study asset prices in an economy where investors derive direct utility not only from consumption but also from fluctuations in the value of their financial wealth. They are loss averse over these fluctuations and how loss averse they are, depends on their prior investment performance.

Our main finding is that in such an economy, asset prices exhibit phenomena very similar to what has been observed in historical data. In particular, stock returns have a high mean, are excessively volatile, and are significantly predictable in the time series. They are also only weakly correlated with consumption growth.

The analysis in this paper raises a number of questions for further investigation. For the sake of simplicity, we have studied an economy containing a single risky asset. In a more realistic economy with many risky assets, it is not immediately clear what investors are loss averse about: do they feel loss averse over changes in the value of individual securities that they own, or only over portfolio fluctuations? In essence, this is a question about how people summarize and think about their investment performance, or about how they do their “mental accounting,” in the language of psychology. How mental accounting affects asset prices is as yet an unexplored topic.

Another question that may warrant further study asks to what extent our preferences can explain not only financial data but also experimental evidence on attitudes to risky gambles. In order to explain the high equity premium, consumption-based models typically assume a high curvature of utility over consumption. As is well-known, this high curvature is consistent with attitudes to small-scale gambles, but unfortunately also predicts that people will reject extremely attractive large-scale gambles.<sup>28</sup> In our model, we do not use a high curvature of utility of con-

28. See, for example, Epstein and Zin [1990] and Kandel and Stambaugh [1991].



sumption, and the loss aversion we assume is not far from that estimated from experimental evidence. This offers the intriguing possibility that preferences of the kind studied here may be able to reconcile attitudes to both large- and small-scale gambles with the empirical facts about stock returns.

APPENDIX

*Proof of Proposition 1.* Proposition 1 is a special case of Proposition 2 with  $Y_t = 0$  and  $C_t = D_t$  for all  $t$ .

*Proof of Proposition 2.* We first show that if the risk-free interest rate is given by (33) and stock returns are determined by (18), (19), (32), and (34), then the strategy of consuming  $C_t = \bar{C}_t = D_t + Y_t$  and holding the market supply of financial securities indeed satisfies the Euler equations of optimality (27) and (28). We then show that these Euler equations are necessary and sufficient conditions for optimality.

Given (29), the interest rate  $R_f$  in (33) satisfies the Euler equation (27). For Euler equation (28), note that  $z_{t+1}$  is determined only by  $z_t$  and  $\epsilon_{t+1}$  through (18) and (32), so we have

$$\begin{aligned} \mathbf{E}_t \left[ R_{t+1} \left( \frac{\bar{C}_{t+1}}{\bar{C}_t} \right)^{-\gamma} \right] &= \mathbf{E}_t \left[ \frac{1 + f(z_{t+1})}{f(z_t)} e^{gD + \sigma_D \epsilon_{t+1}} e^{-\gamma(gC + \sigma_C \eta_{t+1})} \right] \\ &= e^{gD - \gamma gC} \mathbf{E}_t \left[ \mathbf{E}_t [ e^{-\gamma \sigma_C \eta_{t+1}} | \epsilon_{t+1} ] e^{\sigma_D \epsilon_{t+1}} \frac{1 + f(z_{t+1})}{f(z_t)} \right] \\ &= e^{gD - \gamma gC + \gamma^2 \sigma_c^2 (1 - \omega^2) / 2} \mathbf{E}_t \left[ \frac{1 + f(z_{t+1})}{f(z_t)} e^{(\sigma_D - \gamma \omega \sigma_C) \epsilon_{t+1}} \right]. \end{aligned}$$

Applying this, we find that the strategy of consuming  $\bar{C}_t$  and holding the market securities indeed satisfies Euler equation (28).

Euler equations are necessary conditions for optimality. To prove that they are sufficient conditions as well, we apply a method used by Duffie and Skiadas [1994] and Constantinides and Duffie [1996].

To simplify notation, let  $u_t(C_t) = \rho^t C_t^{1-\gamma} / (1 - \gamma)$  and  $\bar{b}_t = \rho^{t+1} b_t$ . Assume that the strategy  $(C^*, S^*)$  satisfies the Euler equations,

$$(47) \quad u'_t(C_t^*) = R_f \mathbf{E}_t [ u'_{t+1}(C_{t+1}^*) ],$$

$$(48) \quad u'_t(C_t^*) = \mathbf{E}_t[R_{t+1}u'_{t+1}(C_{t+1}^*)] + \bar{b}_t \mathbf{E}_t[\hat{v}(R_{t+1}, z_t)].$$

Consider any alternative strategy  $(C^* + \delta C, S^* + \delta S)$  that satisfies the budget constraint

$$(49) \quad \delta W_{t+1} = (\delta W_t - \delta C_t)R_f + \delta S_t(R_{t+1} - R_f).$$

The increase in expected utility from using the alternative strategy is

$$(50) \quad \mathbf{E} \left[ \sum_{t=0}^{\infty} [u_t(C_t^* + \delta C_t) - u_t(C_t^*) + \bar{b}_t \delta S_t \hat{v}(R_{t+1}, z_t)] \right]$$

$$(51) \quad \leq \Delta(\delta C, \delta S) \equiv \mathbf{E} \left[ \sum_{t=0}^{\infty} [u'_t(C_t^*) \delta C_t + \bar{b}_t \delta S_t \hat{v}(R_{t+1}, z_t)] \right],$$

where we have made use of the concavity of  $u_t(\cdot)$  and the linearity of the prospect utility term with respect to  $S_t$ . It is therefore enough to show that  $\Delta(\delta C, \delta S) = 0$  under budget constraint (49).

Multiplying  $u'_{t+1}(C_{t+1}^*)$  with (49) and applying Euler equations (47) and (48), we have

$$(52) \quad \mathbf{E}_t[u'_t(C_t^*) \delta C_t + \bar{b}_t \delta S_t \hat{v}(R_{t+1}, z_t)] \\ = u'_t(C_t^*) \delta W_t - \mathbf{E}_t[u'_{t+1}(C_{t+1}^*) \delta W_{t+1}].$$

Summing up (52) for all  $t$  and taking expectations, we have

$$\Delta(\delta C, \delta S) = u_0(C_0^*) \delta W_0 - \lim_{T \rightarrow \infty} \mathbf{E}[u'_T(C_T^*) \delta W_T].$$

The budget constraint implies that  $\delta W_0 = 0$ . By requiring feasible strategies to use bounded units of financial securities, and with a unit of the risk-free security priced at one, we can show that the limiting term also goes to zero if our model parameters satisfy  $\log \rho - \gamma g_C + g_D + 0.5(\gamma^2 \sigma_C^2 - 2\gamma \omega \sigma_C \sigma_D + \sigma_D^2) < 0$ , a condition which we already noted in footnote 21.<sup>29</sup> So  $\Delta(\delta C, \delta S) = 0$  for any feasible alternative to  $(C^*, S^*)$ .

We have thus shown that any other budget feasible strategy cannot increase utility. The Euler equations are therefore necessary and sufficient conditions of optimality.

29. The proof of this last step is available upon request.

*Proof of Proposition 3.* This is a special case of Proposition 1 with  $\eta = 0$  and  $z_t = 1$  for all  $t$ , which is in turn a special case of Proposition 2, proved above.

*Computation of Average Loss Aversion.* Many of the tables present the investor's average loss aversion over time. There is only one difficulty in computing this quantity: when  $z_t < 1$ —in other words, when the investor has prior gains—part of any subsequent loss is penalized at a rate of one, and part of it is penalized at a rate of 2.25. Therefore, it is not obvious what single number should be used to describe the investor's loss aversion in this case. When  $z_t \geq 1$ , there is no such problem because losses are penalized at a single rate: 2.25 for  $z_t = 1$ , and higher than 2.25 for  $z_t > 1$ .

We adopt the following technique for dealing with the case of  $z_t \leq 1$ : we take the loss aversion component of the investor's utility, namely

$$v(X_{t+1}, S_t, z_t) = \begin{cases} S_t R_{t+1} - S_t R_{f,t} \\ S_t(z_t R_{f,t} - R_{f,t}) + \lambda S_t(R_{t+1} - z_t R_{f,t}) \end{cases} \quad \text{for } \begin{cases} R_{t+1} \geq z_t R_{f,t} \\ R_{t+1} < z_t R_{f,t}, \end{cases}$$

and for  $S_t = 1$  and a risk-free rate of  $R_{f,t} = 3.86$  percent, we compute  $E(v)$ , the expected loss aversion when the excess stock return is distributed as

$$\log R_{t+1} - \log R_f \sim N(0.06, 0.2^2),$$

a good approximation to the historical distribution of the excess stock return. We then compute the quantity  $\bar{\lambda}$  for which an investor with utility function

$$\bar{v}(X_{t+1}, S_t) = \begin{cases} S_t(R_{t+1} - R_{f,t}) \\ \bar{\lambda} S_t(R_{t+1} - R_{f,t}) \end{cases} \quad \text{for } \begin{cases} R_{t+1} \geq R_{f,t} \\ R_{t+1} < R_{f,t} \end{cases}$$

would have exactly the same expected loss aversion for this distribution of the excess stock return, again with  $S_t = 1$ . This  $\bar{\lambda}$  is our measure of the investor's effective loss aversion for any particular  $z_t < 1$ .

## REFERENCES

- Abel, Andrew, "Asset Prices under Habit Formation and Catching up with the Joneses," *American Economic Review*, LXXX (1990), 38–42.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, "A Model of Investor Sentiment," *Journal of Financial Economics*, XLIX (1998), 307–343.
- Benartzi, Shlomo, and Richard Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, CX (1995), 73–92.
- Campbell, John Y., "Asset Prices, Consumption and the Business Cycle," in *Handbook of Macroeconomics*, John Taylor and Michael Woodford, eds. (Amsterdam: North-Holland, 2000).
- Campbell, John Y., and John H. Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, CVII (1999), 205–251.
- Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, *The Econometrics of Financial Markets* (Princeton, NJ: Princeton University Press, 1997).
- Campbell, John Y., and Robert Shiller, "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, XLIII (1988), 661–676.
- Cecchetti, Stephen, Pok-sang Lam, and Nelson Mark, "Mean Reversion in Equilibrium Asset Prices," *American Economic Review*, LXXX (1990), 398–418.
- Cochrane, John H., "Where Is the Market Going: Uncertain Facts and Novel Theories," Federal Reserve Bank of Chicago Discussion Paper, 1998.
- Constantinides, George, "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, XCVIII (1990), 531–552.
- Constantinides, George, and Darrell Duffie, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, CIV (1996), 219–240.
- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, "Investor Psychology and Security Market Under- and Over-reactions," *Journal of Finance*, LIII (1998), 1839–1885.
- Duffie, Darrell, and Costis Skiadas, "Continuous-Time Security Pricing: A Utility Gradient Approach," *Journal of Mathematical Economics*, XXIII (1994), 107–131.
- Durell, Alan, "Consumer Confidence and Stock Market Returns," Working Paper, Harvard University, 1999.
- Epstein, Larry G., and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Asset Returns: A Theoretical Framework," *Econometrica*, LVII (1989), 937–968.
- Epstein, Larry G., and Stanley E. Zin, "First-Order Risk Aversion and the Equity Premium Puzzle," *Journal of Monetary Economics*, XXVI (1990), 387–407.
- Epstein, Larry G., and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption Growth and Asset Returns II: An Empirical Analysis," *Journal of Political Economy*, XCIX (1991), 263–286.
- Fama, Eugene, and Kenneth French, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, XCVI (1988a), 246–273.
- Fama, Eugene, and Kenneth French, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, XXII (1988b), 3–25.
- Gertner, Robert, "Game Shows and Economic Behavior: Risk Taking on 'Card Sharks'," *Quarterly Journal of Economics*, CLI (1993), 507–521.
- Hansen, Lars P., and Ravi Jagannathan, "Restrictions on Intertemporal Marginal Rates of Substitutions Implied by Asset Returns," *Journal of Political Economy*, XCIX (1991), 225–262.
- Hansen, Lars P., and Kenneth Singleton, "Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, XCI (1983), 249–265.
- Heaton, John, and Deborah Lucas, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, CIV (1996), 668–712.
- Hong, Harrison, and Jeremy Stein, "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets," *Journal of Finance*, LIV (1999), 2143–2184.
- Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, XLVII (1979), 263–291.

- Kandel, Shmuel, and Robert Stambaugh, "Asset Returns and Intertemporal Preferences," *Journal of Monetary Economics*, XXVII (1991), 39–71.
- Kocherlakota, Narayana, "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, XXXIV (1996), 42–71.
- Levy, Haim, and Zvi Wiener, "Prospect Theory and Utility Theory: Temporary and Permanent Attitudes toward Risk," Working Paper, Hebrew University of Jerusalem, 1997.
- Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," Working Paper, Carnegie Mellon University, 1999.
- Lucas, Robert, "Asset Prices in an Exchange Economy," *Econometrica*, XLVI (1978), 1419–1446.
- Mehra, Rajnish, and Edward Prescott, "The Equity Premium Puzzle," *Journal of Monetary Economics*, XV (1985), 145–161.
- Poterba, James, and Lawrence Summers, "Mean-Reversion in Stock Returns: Evidence and Implications," *Journal of Financial Economics*, XXII (1988), 27–60.
- Shiller, Robert, "Measuring Bubble Expectations and Investor Confidence," Working Paper, Yale University, 1999.
- Shleifer, Andrei, *Inefficient Markets: An Introduction to Behavioral Finance* (Oxford, UK: Oxford University Press, 1999).
- Sundareshan, Suresh, "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies*, II (1989), 73–88.
- Thaler, Richard H., and Eric J. Johnson, "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice," *Management Science*, XXXVI (1990), 643–660.
- Tversky, Amos, and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, CCXI (1981), 453–458.
- Tversky, Amos, and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, V (1992), 297–323.