August 17, 2004

Understanding the Equity Risk Premium Puzzle

GEORGE M. CONSTANTINIDES*

ABSTRACT

The unconditional mean of the aggregate equity risk premium is almost six percent per year even after adjusting downwards the sample mean premium for unanticipated events in the latter part of the 20th century. In this essay I present my theoretical and empirical research on three classes of generalizations of the standard neoclassical model and discuss their contribution towards a better understanding of the equity risk premium: preferences exhibiting habit persistence; borrowing constraints over the households' life cycle that limit capital market participation and concentrate the stock market risk on the saving middle-aged households; and the recognition that idiosyncratic income shocks are persistent, uninsurable and concentrated in economic recessions.

*University of Chicago and NBER. I thank …..for their insightful comments and constructive criticism. Finally, I thank Vito Gala for his excellent research assistance and insightful comments. Naturally, I remain responsible for errors.

1. Introduction

The term "equity risk premium puzzle", coined by Mehra and Prescott (1985), originally referred to the inability of standard neoclassical economic theory to reconcile the historically large realized premium of stock market return over the risk free rate with its low covariability with aggregate consumption growth. By now, the term encompasses a number of empirical regularities in the prices of capital assets that are at odds with the predictions of standard economic theory.

The neoclassical rational-expectations economic model parsimoniously links the returns of all assets to the per capita consumption growth through the Euler equations of consumption.[1] According to the theory, the risk premia of financial assets are explained by their covariance with per capita consumption growth. However, per capita consumption growth covaries too little with the returns of most classes of financial assets and this creates a whole class of asset pricing puzzles: the aggregate equity return, the long-term bond return, and the returns of various subclasses of financial assets are too large, too variable, and too predictable. Attempts to leverage the low co-variability typically backfire, implying that the observed risk-free rate is too low and has too low variance. I discuss in some depth the aggregate equity puzzle because it exemplifies many of the problems that arise in attempting to explain the premium of any subclass of financial assets.

The covariance of the per capita consumption growth with the aggregate equity return is *positive*. The rational model explains why the aggregate equity premium is positive. However, the covariance is typically one order of magnitude lower than what is needed to explain the premium. Thus, the equity premium is a *quantitative* puzzle.[2]

---

[1] See Breeden (1979), Lucas (1978), Merton (1973), and Rubinstein (1976)

[2] Ferson and Constantinides (1991), Grossman and Shiller (1981), Hansen and Jagannathan (1991), Hansen and Singleton (1982), and many others tested and rejected the Euler equations of consumption. Mehra and Prescott (1985) calibrate an economy to match the process of consumption growth. They demonstrate that the unconditional mean annual premium of the aggregate equity return over the risk-free rate is, at most, 0.35 percent. This is too low, no matter how one estimates the unconditional mean equity premium. Weil (1989) stressed that the puzzle is a dual puzzle of the observed too high equity return and too low risk-free rate.

The equity premium puzzle is *robust*. One may address the problem by testing the Euler equations of consumption or by calibrating the economy. Either way, it is a puzzle. In calibrating an exchange economy, the model cannot generate the first and second unconditional moments of the equity returns. In testing and rejecting the Euler equations of consumption, one abstracts from the market clearing conditions. The rejections tell us that variations in the assumptions on the supply side of the economy do not resolve the puzzle.

The challenge is a *dual puzzle* of the equity premium that is too high and the risk-free rate that is too low relative to the predictions of the model. In calibrating an economy, the strategy of increasing the risk aversion coefficient in order to lever the effect of the problematic low covariance of consumption growth with equity returns increases the predicted risk-free rate and aggravates the risk-free-rate puzzle. In testing the Euler equations of consumption, the rejections are strongest when the risk-free rate is included in the set of test assets.

Several generalizations of essential features of the model have been proposed to mitigate its poor performance. They include alternative assumptions on preferences,[3] modified probability distributions to admit rare but disastrous market-wide events,[4] incomplete markets,[5] and market imperfections.[6] They also include a better understanding of data problems such as limited participation of consumers in the stock market,[7] temporal aggregation,[8] and the survival bias of the United States capital market.[9]

[3] For example, Abel (1990), Cagetti (2002), Bansal and Yaron (2004), Benartzi and Thaler (1995), Boldrin, Christiano, and Fisher (2001), Campbell and Cochrane (1999), Constantinides (1990), Epstein and Zin (1991), and Ferson and Constantinides (1991).

[4] The merits of this explanation are discussed in Mehra and Prescott (1988), and Rietz (1988).

[5] For example, Bewley (1982), Brav, Constantinides, and Geczy (2002), Constantinides and Duffie (1996), Heaton and Lucas (1996), Krebs (2002), Mankiw (1986), Mehra and Prescott (1985), and Storesletten, Telmer, and Yaron (2001, 2004).

[6] For example, Aiyagari and Gertler (1991), Bansal and Coleman (1996), Constantinides, Donaldson, and Mehra (2002), Daniel and Marshall (1997), Danthine, Donaldson, and Mehra (1992), He and Modest (1995), and Heaton and Lucas (1996).

[7] Attanasio, Banks and Tanner (2002), Brav, Constantinides, and Geczy (2002), Brav and Geczy (1995), Mankiw and Zeldes (1991), and Vissing-Jorgensen (2002).

[8] Gabaix and Laibson (2001), Heaton (1995), and Lynch (1996).

Many of these generalizations contribute in part toward our better understanding of the economic mechanism that determines the pricing of assets. I refer the reader to the introductory essay by Mehra and Prescott and several essays in this Handbook; the textbooks by Campbell, Lo, and McKinlay (1997) and Cochrane (2001); and the articles by Campbell (2001, 2003), Cochrane (1997), Cochrane and Hansen (1992), Constantinides (2002), Kocherlakota (1996), and Mehra and Prescott (2003).

My objective in this essay is to provide a better understanding of these issues when I generalize, one at a time, three of the key assumptions of the standard theory. First, I relax the assumption that preferences are time separable and introduce preferences that exhibit habit persistence. Second, I recognize that participation in the stock market by households is limited and that some aggregate consumption measures reflect the consumption of both participating and non participating households in the stock market. Lastly, I relax the assumption that the market is complete and address the implications of incomplete consumption insurance.

I begin with the standard assumptions of neoclassical economic theory, as adapted in finance. There are I consumers $i, i = 1,...,I$, each consuming c_t^i units of the consumption good in period t and having von Neumann-Morgenstern preferences

$$E_0\left[U_i\left(c_0^i,...,c_t^i,c_{t+1}^i,...\right)\right]. \tag{1.1}$$

There are J capital assets, $j = 1,...,J$, traded by the consumers in perfect markers and having returns $R_{j,t+1}$ between periods t and $t+1$. Utility maximization by the i^{th} consumer in trading the j^{th} asset at time t yields the following Euler equation of consumption between periods t and $t+1$:

$$E_t\left[-\frac{\partial U_i}{\partial c_t^i}+\frac{\partial U_i}{\partial c_{t+1}^i}R_{j,t+1}\right]=1. \tag{1.2}$$

---

[9] See Brown, Goetzmann, and Ross (1995). However, Jorion and Goetzmann (1999, Table 6) found that the average real capital gain rate of a United States equities index exceeds the average rate of a global equities index that includes both markets that have and have not survived by merely one percent per year.

In the standard theory, preferences are typically specialized to be time separable as

$$E_0\left[\sum_{t=0}^{\infty}\beta^t u_i\left(c_t^i\right)\right] \tag{1.3}$$

with constant subjective discount factor β. There is no empirical justification for this specialization of preferences. In Section 2, I relax this assumption, introduce preferences that exhibit *habit persistence* and explore its ramifications in understanding the observed equity premium. In any case, with time separable preferences as in equation (1.3), the Euler equation of consumption simplifies into the following equation:

$$E_t\left[\beta\frac{\partial u_i\left(c_{t+1}^i\right)/\partial c_{t+1}^i}{\partial u_i\left(c_t^i\right)/\partial c_t^i}R_{j,t+1}\right]=1. \tag{1.4}$$

In the standard theory, the model is typically further specialized by assuming that the market is complete. There is little, if any, empirical justification for this assumption either. In Section 4, I relax this assumption, considering instead an *incomplete market*, and explore its implications in addressing the observed equity premium. In any case, under the assumption of market completeness, the equilibrium in this heterogeneous-consumer economy is isomorphic in its pricing implications to the equilibrium in a representative-consumer economy. The strong assumption of market completeness is indirectly built into asset pricing models in finance and neoclassical macroeconomic models through the assertions that a representative consumer exists. For our purposes, market completeness implies that there exists a period utility function $u\left(C_t\right)$ of per capita consumption $C_t \equiv I^{-1}\sum_{i=1}^{I}c_t^i$, such that we may replace the I consumer-specific Euler equations (1.4) with the following Euler equation of consumption of the representative consumer:

$$E_t\left[\beta \frac{\partial u(C_{t+1})/\partial C_{t+1}}{\partial u(C_t)/\partial C_t} R_{j,t+1}\right] = 1\,. \tag{1.5}$$

It is often assumed that the period utility function is of the power form, $u(C_t) = (1-A)^{-1}(C_t)^{1-A}$, with constant relative risk aversion (RRA) coefficient *A*. With power utility, the Euler equation (1.5) further simplifies into the following:

$$E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-A} R_{j,t+1}\right] = 1\,. \tag{1.6}$$

An advantage of the assumption that the period utility function is of the power form is that per capita consumption appears in the Euler equation only as per capita consumption *growth,* C_{t+1}/C_t, which may be stationary even if the consumption *level* is non-stationary.

In empirical work and calibration, the per capita consumption growth in the Euler equation (1.6) is often taken to be the aggregate consumption growth, where the aggregate consumption is taken from the National Income and Product Accounts (NIPA). However, the aggregate consumption reported by NIPA is the sum total over all households, irrespective of whether these households are marginal in the equities market or not. In Section 3, I relax this assumption, recognizing instead the *limited stock market participation* by households and explore its ramifications in addressing the observed equity premium. I pursue these issues in the context of a lifecycle model also in which young consumers are constrained from borrowing against their future labor income to invest in the stock market.

In any case, by ignoring the limited capital market participation, the Euler equations (1.6) on the market return, $R_{M,t}$, and on the risk free return, $R_{F,t}$, may be combined and simplified either under some distributional assumptions (for instance, joint log-normality of the consumption growth rate and returns) or approximations (for example, Taylor-series expansion or log-linearization) to obtain the following result:

$$E\left[R_{M,t} - R_{F,t} \right] = A \operatorname{cov}\left(\frac{C_{t+1}}{C_t}, R_{M,t+1} \right). \tag{1.7}$$

The unconditional mean equity premium is equal to the product of the RRA coefficient of the representative consumer and the covariance of per capita consumption growth with the market return.

Empirically, the covariance of the per capita consumption growth with the market return is low and cannot generate the mean equity premium with a reasonable value of the RRA coefficient. If one assumes a sufficiently high value of the RRA coefficient in order to generate a realistic mean premium, the Euler equation of consumption with respect to the risk free rate now implies that the risk free rate is too low. Thus the puzzle is a joint puzzle that the observed mean equity premium is too high and the risk free rate is too low. These are but two of the implications of the Euler equations (1.5) that are at odds with the data either in empirical work or in calibration. In the following sections of this essay, I sequentially examine preferences that exhibit habit persistence, recognize the limited stock market participation by households, and relax the assumption of market completeness and explore their ramifications in better understanding the equity premium, the risk free rate and the Euler equations of consumption.

2. Habit persistence

In this section, I explore the contribution of habit persistence in resolving the equity premium puzzle and related anomalies. Habit persistence has a long tradition in economic theory, dating back to Marshall (1920) and Duesenberry (1949). It is the property of preferences that an increase in consumption increases the marginal utility of consumption at adjacent dates relative to the marginal utility of consumption at distant ones.

I model *internal habit*, x_t, as an exponentially weighted sum of the representative consumer's own past consumption, c_t, as

$$x_t = e^{-at} x_0 + \int_0^t e^{a(s-t)} c_s ds \, . \tag{2.1}$$

with parameters $a \, and \, x_0$. The half-life of internal habit is $a^{-1} \ln 2$.

Pollak (1970) introduced a model of *external* habit persistence in which the consumer does not take into account the effect of her own current consumption on her future preferences but rather takes into account the effect of other consumers' consumption on her future preferences. I model *external habit*, X_t, as an exponentially weighted sum of the per capita past consumption, C_t, of *identical* consumers as

$$X_t = e^{-a't} X_0 + \int_0^t e^{a'(s-t)} C_s ds \tag{2.2}$$

with parameters $a' \, and \, X_0$. The key assumption is that there are a large number of identical consumers so that a consumer chooses her optimal consumption taking the path of per capita consumption as exogenous. Naturally, in equilibrium, $c_t = C_t$.

I specialize the von Neumann-Morgenstern preferences stated in equation (1.1) as follows in a continuous time framework:

$$\int_0^\infty e^{-\rho t}\left(1-A\right)^{-1}\left(f\left(c_t,x_t,X_t\right)\right)^{1-A}dt \tag{2.3}$$

where $\rho \equiv -\ln\beta$ is the subjective discount rate. The concavity parameter A has the interpretation of the RRA coefficient in the time separable model, $f\left(c_t,x_t,X_t\right)=c_t$, but not so in the habit model. The function $f=f\left(c_t,x_t,X_t\right)$ is nonlinear in consumption, internal habit and external habit. More generally, I may define $f\left(\cdot\right)$ to be a function of internal and external habits at several different frequencies.

Building on earlier work by Ryder and Heal (1973) and Sundaresan (1989), I demonstrated in Constantinides (1990) that internal habit persistence can, in principle, reconcile the high mean equity premium with the low variance of consumption growth and with the low covariance of consumption growth with equity returns. Habit persistence lowers the intertemporal elasticity of substitution in consumption, given the risk aversion. The mean equity premium is equal to the covariance of consumption growth with equity returns, divided by this elasticity. Therefore, given the risk aversion, habit persistence lowers the elasticity and raises the mean equity premium.

I begin with the special case where the habit is *internal* and *linear*, as studied in Constantinides (1990):

$$f\left(c_t,x_t,X_t\right)=c_t-bx_t\,. \tag{2.4}$$

where b is a parameter. I embed these preferences in a representative-consumer production economy and discuss their implications. There exists only one production good that is also the consumption good. This good may be consumed or invested without adjustment costs in two constant-returns-to-scale technologies. I do not explicitly model the labor input to production because my focus is on the Euler equations of consumption and their implications on the equity premium puzzle.

Since this is a representative-consumer economy, I interpret the optimal consumption and investment paths as the equilibrium paths and the shadow prices of assets as the equilibrium prices. The return on the risk free technology is constant over time and this determines the risk free rate as this technology is always active, under the assumed parameter values. The risky technology has constant returns to scale and i.i.d. returns with given mean and standard deviation. This technology also is always active and determines the mean and standard deviation of the return on the stock market.

I do not model the consumers' labor income. Therefore consumption equals dividends plus interest, net of new investment in production. One way to rationalize the assumption that consumers do not receive labor income is to view the capital invested in production as including human capital, the capitalized value of future labor income. Then the return on the risky technology is interpreted as the return on both the stock market and human capital.

There is a single state variable in this economy, the habit-to-consumption ratio. In Constantinides (1990), I solved this model and determined the dynamics of the wealth, consumption, and RRA coefficient as functions of this ratio. In calibrating the model, I set the technology parameters such that the annual risk free rate is 1% and the annual market return has mean 7% and standard deviation 16.5%, implying mean annual equity premium of 6%. The goal is to set the remaining parameters $\rho, A, a, \text{ and } b$ so that annual consumption growth rate has mean 1.8% and standard deviation 3.6% while the RRA coefficient is kept at a plausibly low value. The order of magnitude of the subjective discount factor is not a controversial parameter, so I set $\rho = 0.037$ per year.

I begin with time separable preferences by setting $b = 0$. If I want the model to imply that 100% (or 50%) of the wealth is invested in the stock market, I must set the RRA coefficient at 2.2 (or, 4.4). Then the model implies that the annual consumption growth has mean 3% (or, 1%) and standard deviation 16.5% (or, 8.2%). The time separable model fails: in order to generate the moments of the market return and the risk free rate, it requires much higher variability of consumption growth and covariability with the market return than what is observed in the data.

With internal habit, several parameter combinations attain the calibration goal, as displayed in Constantinides (1990, Table 1). For example, with parameters $A = 2.2, a = 0.5, \text{ and } b = 0.405$, the RRA coefficient has unconditional mean 2.88, and the annual consumption growth has unconditional mean 1.8% and standard deviation 3.6%. Furthermore, on average, 79% of the wealth is invested in the stock market.

The intuition behind the model's success is that the RRA coefficient is an increasing function of the habit-to-consumption ratio. First, consider time separable preferences. The standard explanation for a positive equity premium is that in an economic recession equity returns are low and this event typically occurs when consumption is low and, therefore, the marginal utility of consumption is high. However, this effect is not sufficiently strong to generate the order of magnitude of the observed premium. Second, consider preferences exhibiting habit persistence. In an economic recession, the decrease in consumption increases the habit-to-consumption ratio, increases the RRA aversion coefficient, and raises the marginal utility of consumption even further than in a model with time separable preferences. It turns out that the combination of the two effects is sufficiently strong to generate the order of magnitude of the observed premium without an unreasonably high RRA coefficient.

The encouraging calibration results are reinforced by the empirical evidence. Ferson and Constantinides (1991) tested the Euler equations of consumption in the special case of the internal habit model in which the habit depends only on the *first* lag of own consumption. They reported that the habit model performs better than the time-separable model and that the habit persistence parameter is economically and statistically significant. Cecchetti, Lam, and Mark (1994), Cochrane and Hansen (1992), Gallant, Hansen and Jagannathan (1991), Hansen and Tauchen (1990), and Heaton (1995) found that the internal habit model brings the moments of the marginal rate of substitution closer to the feasible Hansen and Jagannathan (1991) region than the time separable model does.

I proceed to explore the implications of replacing the internal habit of equation (2.1) by the external habit of equation (2.2) in the model of Constantinides (1990). I embed these preferences in the above representative-consumer production economy and determine the dynamics of the wealth, consumption, and the RRA coefficient as functions of the habit-to-consumption ratio. For a set of parameter values that attain the calibration goal when habit is internal, $A = 2.2, a = 0.5, \textit{and } b = 0.405$, the RRA coefficient is 3.35, and the annual consumption growth has mean -1% and standard deviation 2.6% at the mode of the habit-to-consumption ratio. With parameter values $A = 2.2, a = 0.5, \textit{and } b = 0.25$, the RRA coefficient is 3.35, and the annual consumption growth has mean 1.37% and standard deviation 8.4% at the mode of the habit-to consumption ratio. The standard deviation of the annual consumption growth is three times its observed value. In general, the model with linear external habit requires that the standard deviation of the annual consumption growth be much higher than the value observed in the data. Part of the problem may be the stylized assumptions of the production economy in which these preferences are embedded. Ferson and Harvey (1992) tested the linear external habit model and reported positive results. However, Li (2001) reported that in the linear external habit model (and in the non-linear version, to be described next), the habit-to-consumption ratio has limited success in explaining the time series of the premia of stock and bond portfolios.

A common feature of the internal and external linear habit preference specifications is that the equilibrium consumption growth is auto-correlated. The auto-correlation is due to the constancy of the risk free rate, forced by the risk free technology with constant rate of return. This can be demonstrated in an exchange economy in which the endowment growth process (and, therefore, the equilibrium consumption growth process) is assumed to be i.i.d.. In this case, the risk free rate is endogenous and turns out to be excessively variable.

Leaving aside the empirical issue as to whether consumption growth exhibits slight positive auto-correlation or not, one may introduce a model of *nonlinear* habit that produces richer dynamics for the consumption growth, the risk free rate and the stock market return, and, in particular, implies constant risk free rate in an exchange economy with i.i.d. consumption growth. Abel (1990) and Campbell and Cochrane (1999) introduced preference models with nonlinear habit. Specifically, Campbell and Cochrane (1999) studied a model with nonlinear external habit. Rather than directly model the function $f = f(c_t, X_t)$, they reverse-engineered the process of the habit-to-consumption ratio, in a way that the interest rate is either constant or exhibits low variability, the long-horizon returns are predictable, and the behavior of equity prices along the business cycle is realistic. Calibrated to the actual history of aggregate consumption, the model hits the aggregate price-dividend ratio in a number of periods but misses it in the 1950s and 1990s. The model has in common with the linear internal habit model the property that the risk aversion is countercyclical. The model relies on a high and highly volatile RRA coefficient that has the value of eighty in the steady state and much higher still in economic recessions.

Turning to the empirical evidence, Li (2001) reported that in both the linear and non-linear versions of the external habit models, the habit-to-consumption ratio has limited success in explaining the time series of the premia of stock and bond portfolios. Menzly, Santos, and Veronesi (2004) developed a variant of the above nonlinear external habit model and reported that it helps explain the cross-section of asset returns. Chen and Ludvigson (2004) treated the functional form of nonlinear habit as an unknown and estimated it along with the other model parameters from the data in both the internal habit and external habit versions of the model. They concluded that habit is nonlinear, is internal and explains the cross-section of stock returns better than most extant factor models do.

A promising direction for future research is to endogenize the currently *ad hoc* specification of the nonlinear habit. Another direction is to address the predictability of asset returns and their behavior along the business cycle in a model that benefits from the flexibility of the nonlinear specification of habit but keeps risk aversion and its variability plausibly low.

Habit persistence is already gaining ground as an ingredient of economic models addressing a diverse set of economic problems beyond asset pricing, including the consumption-saving behavior and the home-equity puzzle. Habit persistence is a sensible property of preferences. It is also a property that allows for the separate specification of the RRA coefficient and the intertemporal elasticity of substitution within the class of Neumann-Morgenstern preferences.

3. Limited stock market participation

The limited participation of households in the stock market sheds light on the equity premium puzzle by addressing the low correlation of the equity return with the growth rate of per capita consumption. Whereas we understood all along that many households whose consumption is counted in the NIPA-reported aggregate capital consumption do not hold equities, it took a paper by Mankiw and Zeldes (1991) to point out that the emperor has no clothes.[10] Even though 52 percent of the United States adult population held stock either directly or indirectly in 1998, compared to 36 percent in 1989, stockholdings remain extremely concentrated in the hands of the wealthiest few. Furthermore, wealthy entrepreneurs may be infra marginal in the stock market if their wealth is tied up in private equity.

In Section 1, I pointed out that the unconditional mean equity premium is equal to the product of the RRA coefficient of the representative consumer and the covariance of the per capita consumption growth with the market return, as in equation (1.6). I now explicitly distinguish between the NIPA-reported aggregate consumption, C_t^{NIPA}, and the aggregate consumption, $I \times C_t \equiv \sum_{i=1}^{I} c_t^i$, by the subset of households who are marginal in the equities market, identified as the stockholders. Under the plausible assumption that the consumption of households infra marginal in the stock market is uncorrelated with the market return, equation (1.6) yields the following approximate expression for the mean equity premium:

$$E\left[R_{M,t} - R_{F,t} \right] = \frac{A}{\alpha} \operatorname{cov}\left(\frac{C_{t+1}^{NIPA}}{C_t^{NIPA}}, R_{M,t+1} \right), \tag{3.1}$$

[10] Since then, several papers have studied the savings and portfolio composition of households, stratified by income, wealth, age, education, and nationality. See Blume and Zeldes (1993), Haliassos and Bertaut (1995), Heaton and Lucas (1999, 2000), Poterba (2001), and the collected essays in Guiso, Haliassos, and Jappelli (2001).

where α denotes the average fraction of NIPA consumption attributed to stockholders. By recognizing the limited stock market participation of households, I may coax the theory to predict a higher value of the mean equity premium without increasing the RRA coefficient and without predicting an implausibly high value for the risk free rate.

Alternatively, one may avoid the use of NIPA consumption data and obtain alternative measures of the consumption by stockholders. Mankiw and Zeldes (1991) calculated the per capita food consumption of a subset of households, designated as asset holders according to a criterion of asset holdings above some threshold. They found that the implied RRA coefficient decreases as the threshold is raised. Brav and Geczy (1995) confirmed their result by using the non-durables and services per capita consumption, reconstructed from the Consumer Expenditure Survey (CEX) database. Attanasio, Banks, and Tanner (2002), Brav, Constantinides, and Geczy (2002), and Vissing-Jorgensen (2002) found some evidence that per capita consumption growth can explain the equity premium with a relatively high value of the RRA coefficient, once we account for limited stock market participation. However, Brav, Constantinides, and Geczy (2002) pointed out that the statistical evidence is weak and the results are sensitive to the experimental design.

Limited stock market participation is a fact of life and empirical tests of the Euler equations of consumption should account for it. However, my interpretation of these empirical results is that recognition of limited stock market participation alone is insufficient to explain asset returns.

Borrowing constraints provide an endogenous partial explanation for the limited participation of young consumers in the stock market. Constantinides, Donaldson, and Mehra (2002) constructed an overlapping-generations exchange economy in which consumers live for three periods. In the first period, a period of human capital acquisition, the consumer receives a relatively low endowment income. In the second period, the consumer is employed and receives wage income subject to large uncertainty. In the third period, the consumer retires and consumes the assets accumulated in the second period. The key feature is that the bulk of the *future* income of the young consumers is derived from their wages forthcoming in their middle age, while the *future* income of the middle-aged consumers is derived primarily from their savings in equity and bonds.

The young would like to invest in equity, given the observed large equity premium. However, they are unwilling to decrease their current consumption in order to save by investing in equity, because the bulk of their lifetime income is derived from their wages forthcoming in their middle age. They would like to borrow, but the borrowing constraint prevents them from doing so. Human capital alone does not collateralize major loans in modern economies for reasons of moral hazard and adverse selection. The model explains why many consumers do not participate in the stock market in the early phase of their life cycle.

The future income of the middle-aged consumers is derived from their current savings in equity and bonds. Therefore, the risk of holding equity and bonds is concentrated in the hands of the middle-aged saving consumers. This concentration of risk generates the high equity premium and the demand for bonds, in addition to the demand for equity, by the middle-aged.[11] The model recognizes and addresses simultaneously, at least in part, the equity premium, the limited participation in the stock market, and the demand for bonds.

Constantinides et al (2002) calibrated the model to match the following eight targets: the average share of income going to labor; the average share of income going to the labor of the young; the average share of income going to interest on government debt; the coefficient of variation of the 20-year wage income of the middle aged; the coefficient of variation of the 20-year aggregate income; the 20-year autocorrelation of the labor income; the 20-year autocorrelation of the aggregate income; and the 20-year cross-correlation of the labor income and the aggregate income. Since the length of one period in this model is twenty years, for all securities (equity, bond or consol), the annualized mean return is defined as the mean of $(20)^{-1}\ln$(20-year holding period return); and the annualized standard deviation of the return is defined as the standard deviation of

[11] See also the discussion in the related papers by Bertaut and Haliasos (1997), Bodie, Merton, and Samuelson (1992), Jagannathan and Kocherlakota (1996) and Storesletten, Telmer, and Yaron (2001).

$(20)^{-0.5}$ ln(20-year holding period return). Below, I reproduce the first panel of Table 1 in Constantinides et.al. (2002).[12]

	Borrowing-constrained	Borrowing-unconstrained
Mean equity return	8.4	10.2
Std of equity return	23.0	42.0
Mean bond return	5.1	9.0
Std of bond return	15.4	27.6
Mean premium over bond	3.4	1.1
Std of premium over bond	18.4	31.6
Mean consol return	3.7	9.9
Std of consol return	19.1	27.6
Mean premium over consol	4.7	0.3
Std of premium over consol	10.5	5.2
Corr of wages and dividend	-0.43	-0.43
Corr of wages and premium	-0.02	0.00

The borrowing constraint decreases the mean return of the 20-year or consol bond by about a factor of two. This observation is robust to the calibration of the correlation and auto-correlation of the labor income of the middle-aged with the aggregate income. The borrowing constraint goes a long way, albeit not all the way, towards resolving the risk free rate puzzle. If the young are able to borrow, they do so and purchase equity; the borrowing activity of the young raises the bond return, thereby exacerbating the risk-free rate puzzle.

Second, the mean equity premium over the 20-year or consol bond is about 4%. This is satisfactory given that long term bond returns typically command a premium over the short term risk free rate. This premium drastically decreases when the borrowing constraint is relaxed. If the young are able to borrow, the increase in the bond return induces the middle-aged to shift their portfolio holdings in favor of the bond; the increase in the demand for equity by the young and the decrease in the demand for equity by the middle-aged work in opposite directions; on balance, the effect is to increase the return on both equity and the bond while simultaneously shrinking the equity premium.

[12] The RRA coefficient is set at 6; the coefficient of variation of the (20-year) aggregate income is set at 0.20; the coefficient of variation of the (20-year) wages is set at 0.25; the auto-correlation of aggregate income, the auto-correlation of wages, and the correlation of aggregate income and wages are all set at 0.1. The consol bond is in positive net supply and the one-period (20-year) bond is in zero net supply.

Third, the correlation of the labor income of the middle-aged and the equity premium over the 20-year bond is much smaller in absolute value than the exogenously-imposed correlation of the labor income of the middle-aged and the dividend. Thus, equity is attractive to the young because of the large mean equity premium *and* the low correlation of the premium with the wage income of the middle-aged, thereby corroborating another important dimension of the model. In equilibrium, it turns out that the correlation of the wage income of the middle-aged and the equity return is low.[13] The young consumers would like to invest in equity because equity return has low correlation with their future consumption, if their future consumption is derived from their future wage income. However, the borrowing constraint prevents them from purchasing equity on margin. Furthermore, since the young consumers are relatively poor and have an incentive to smooth their intertemporal consumption, they are unwilling to decrease their current consumption in order to save by investing in equity. Therefore, the young choose not participate in the equity market.

Finally, the borrowing constraint results in standard deviations of the annualized, 20-year equity and bond returns which are lower than in the unconstrained case and which are *comparable* to the sample moments.

The model also serves as a useful laboratory to address a range of related economic issues. Campbell et.al. (2001), and Constantinides, Donaldson, and Mehra (2004a) addressed the cost of Social Security reform. Storesletten, Telmer, and Yaron (2001, 2004) explored the interaction of life cycle effects and the uninsurable wage income shocks and found that the interaction plays an important role in explaining asset returns. Heaton and Lucas (1999) explored whether changes in market participation patterns account for the recent rise in stock prices and find that they do not.

The low covariance of the growth rate of aggregate consumption with equity returns is a major stumbling block in explaining the mean aggregate equity premium and the cross-section of the asset returns, in the context of a representative-consumer economy with time separable preferences. Mankiw and Shapiro (1986) found that the

[13] The low correlation of the wage income of the middle-aged and the equity return is a property of the equilibrium and obtains for a wide range of values of the assumed correlation of the wage income of the middle-aged and the dividend.

market beta often explains asset returns better than the consumption beta does. Over the years, a number of different economic models have been proposed that effectively increase the covariance of equity returns with the growth rate of aggregate consumption, by proxying the growth rate of aggregate consumption with the aggregate stock market return in the Euler equations of consumption.[14]

I present an old folks' tale, introduced in Constantinides, Donaldson, and Mehra (2004a,b), that accomplishes this goal without introducing Epstein-Zin (1991) preferences or preferences defined directly over wealth. Old folks who are rich enough to be non-trivial investors in the capital markets care about their wealth just as much as younger folks do, even though the state of their health and their medical expenses account for their consumption patterns better than fluctuations of their wealth do. This simple observation takes us a long way toward understanding why the stock market return does a better job than the growth of aggregate consumption does in explaining asset returns.

In the context of an overlapping-generations economy, the major investors in the market are the middle-aged households at the saving phase of their life cycle. These households save with the objective to maximize the utility of their "consumption" in their middle and old age. The insight here is that "consumption" of the old consists of two components, direct consumption, c_D; and the "joy of giving," c_B, in the form of *inter vivos* gifts and *post mortem* bequests. Since the old households' direct consumption is constrained by the state of their health, the correlation between the direct consumption of the old and the stock market return is *low*, a prediction that is born out empirically. Therefore, the balance of the old households' wealth, c_B, is *a fortiori* highly correlated with the stock market return. In terms of a utility function of consumption at the old age,

[14] Friend and Blume (1975) explained the mean equity premium with low RRA coefficient by assuming a single-period economy in which the end-of-period consumption inevitably equals the end-of-period wealth. Epstein and Zin (1991) introduced a recursive preference structure that emphasizes the timing of the resolution of uncertainty. Even though the preferences are defined over consumption alone, the stock market return enters directly in the Euler equations of consumption. Bakshi and Chen (1996) introduced a set of preferences defined over consumption and wealth—the spirit of capitalism—that also have the effect of introducing the stock market return in the Euler equations of consumption.

$u(c_D)+v(c_B)$, that is separable over direct consumption and bequests, the model predicts an Euler equation of consumption with marginal utility at the old age given by $v'(c_B)$ and not by $u'(c_D)$, where c_B is proxied by the stock market value.

This model remains to be tested. Nevertheless, it reinforces the general point that per capita consumption measures neither the total consumption of the marginal investor in the stock market nor that part of the marginal investor's consumption that is unconstrained by health and medical considerations.

4. Incomplete markets and idiosyncratic income shocks

In economic recessions, investors are exposed to the double hazard of stock market losses and job loss. Investment in equities not only fails to hedge the risk of job loss but also accentuates its implications. Investors require a hefty equity premium in order to be induced to hold equities. This is the argument that I formalize below and address the predictability of asset returns and their unconditional moments.

The observed correlation of per capita consumption growth with stock returns is low. Over the years, I have grown skeptical of how meaningful an economic construct *aggregate* (as opposed to *disaggregate*) consumption is, and how hard we should push aggregate or per capita consumption to explain returns. At a theoretical level, aggregate consumption is a meaningful economic construct if the market is complete or effectively so.[15]

In a complete market, heterogeneous households are able to equalize, state by state, their marginal rate of substitution. The equilibrium in a heterogeneous-household, full-information economy is isomorphic in its pricing implications to the equilibrium in a representative-household, full-information economy, if households have von Neumann-Morgenstern preferences.[16] The strong assumption of market completeness is indirectly built into asset pricing models in finance and neoclassical macroeconomic models through the assumption of the existence of a representative household.

Bewley (1982), Mankiw (1986), and Mehra and Prescott (1985) suggested the potential of enriching the asset-pricing implications of the representative-household paradigm, by relaxing the assumption of complete markets.[17] Constantinides and Duffie (1996) found that incomplete markets substantially enrich the implications of the

---

[15] The market is effectively complete when all households have preferences that imply one-fund or two-fund separation.

[16] See Constantinides (1982), Mehra and Prescott (1985, an unpublished earlier draft), and Negishi (1960).

[17] There is an extensive literature on the hypothesis of complete consumption insurance. See Altonji, Hayashi and Kotlikoff (1992), Attanasio and Davis (1997), Cochrane (1991), and Mace (1991).

representative-household model. Their main result is a proposition demonstrating, by construction, the existence of household income processes, consistent with given aggregate income and dividend processes, such that equilibrium equity and bond price processes match the given equity and bond price processes.

The starting point in the Constantinides and Duffie (1996) theory, are the $I \, x \, J$ Euler equations (1.4) of individual consumption by the I consumers for the J assets, specialized for preferences that imply constant relative risk aversion:

$$E_t\left[\beta g_{i,t+1}^{-A} R_{j,t+1}\right] = 1, \quad i = 1,...,I; \quad j = 1,...,J\,. \tag{4.1}$$

The consumption growth of the i^{th} consumer is $g_{i,t+1} \equiv c_{t+1}^i / c_t^i$. The point of departure from standard theory is the recognition that market incompleteness rules out the step of replacing the Euler equations (4.1) by the Euler equation (1.6) of the representative consumer. In principle, I may directly test the $I \, x \, J$ system of Euler equations (4.1) of household consumption. In the US, the best available disaggregated consumption data are in the department of Labor Statistics' Consumer Expenditure Survey (CEX) of quarterly consumption of selected households (not individual consumers).

A stochastic discount factor (SDF), or pricing kernel, is defined as any random variable m_{t+1} with the following property:

$$E_t\left[m_{t+1} R_{j,t+1}\right] = 1, \quad j = 1,...,J\,. \tag{4.2}$$

Therefore, equation (4.1) states that each consumer's marginal rate of substitution, $\beta g_{i,t+1}^{-A}$, is a valid SDF. Since household consumption data are reported with substantial error, it is difficult to directly test the hypothesis that each household's marginal rate of substitution is a valid SDF.

I sum the $I \, x \, J$ Euler equations (4.1) across households and obtain the expression

$$E_t\left[\beta\left\{I^{-1}\sum_{i=1}^{I} g_{i,t+1}^{-A}\right\}R_{j,t+1}\right]=1, \quad j=1,...,J \,. \tag{4.3}$$

Equation (4.3) states that the weighted sum of the consumers' marginal rate of substitution is a valid SDF also:

$$m_{t+1}=\beta\left\{I^{-1}\sum_{i=1}^{I} g_{i,t+1}^{-A}\right\}. \tag{4.4}$$

Whereas I may directly test this hypothesis, the equally weighted sum of the households' marginal rate of substitution is still susceptible to observation error because each term in the sum is raised to a high power, if the risk aversion coefficient is high.

I expand equation (4.4) as a Taylor series up to cubic terms and obtain the following approximation for the SDF:

$$m_{t+1}=\beta\, g_{t+1}^{-A}\left\{1+\frac{A\,(A+1)}{2}I^{-1}\sum_{i=1}^{I}\left(\frac{g_{i,t+1}}{g_{t+1}}-1\right)^2-\frac{A\,(A+1)\,(A+2)}{6}I^{-1}\sum_{i=1}^{I}\left(\frac{g_{i,t+1}}{g_{t+1}}-1\right)^3\right\} \tag{4.5}$$

The term $g_{t+1}\equiv I^{-1}\sum_{i=1}^{I} g_{i,t+1}$ is the sample mean of the consumption growth rate across consumers; $I^{-1}\sum_{i=1}^{I}\left(\frac{g_{i,t+1}}{g_{t+1}}-1\right)^2$ is the squared coefficient of variation of the consumption growth rate across consumers; and $I^{-1}\sum_{i=1}^{I}\left(\frac{g_{i,t+1}}{g_{t+1}}-1\right)^3$ is the skewness of the cross sectional variation of consumption growth.

The theory requires that the idiosyncratic income shocks have three properties in order to explain the returns on financial assets. First, they must be *uninsurable*. If there exist a complete set of markets, the equilibrium of a heterogeneous-household, full-information economy is isomorphic in its pricing implications to the equilibrium of a representative-household, full-information economy and household consumption growth cannot do better than aggregate consumption growth in explaining the returns. To see this, note that, if a complete set of markets exists, then the heterogeneous households are able to equalize, state by state, their marginal rates of substitution. In particular for any state s at time $t+1$, there exists a state-specific but consumer-independent parameter λ_s such that $g_{i,t} = g_t = \lambda_s$, $i = 1,...,I$. The SDF in equations (4.5) simplifies into

$$m_{t+1} = \beta\, g_{t+1}^{-A} \quad . \tag{4.6}$$

or, equivalently, into[18]

$$m_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-A} . \tag{4.7}$$

Therefore, if there exist a complete set of markets, the equilibrium of a heterogeneous-household, full-information economy is isomorphic in its pricing implications to the equilibrium of a representative-household, full-information economy and household consumption growth cannot do better than aggregate consumption growth in explaining the returns.

Second, the theory requires that the idiosyncratic income shocks be *persistent*. If the shocks are transient, then households can smooth their consumption by borrowing or by drawing down their savings.[19]

---

[18] When there is a complete set of markets, equation (4.7) follows from the fact that, for any state s, $\frac{C_{t+1}}{C_t} = \frac{I^{-1}\sum_{i=1}^{I} c_{t+1}^i}{I^{-1}\sum_{i=1}^{I} c_t^i} = \lambda_s = g_{t+1}$. I expect that the SDF given by equation (4.7) is less susceptible to observation error than the SDF given by equation (4.6).

[19] Aiyagari and Gertler (1991) and Heaton and Lucas (1996) find that consumers facing *transient* shocks come close to the complete-markets rule of complete risk sharing even with transaction costs and/or borrowing costs, *provided that the supply of bonds is not restricted to an unrealistically low level*.

Third, the conditional variance, or some higher moment of the income shocks, must be counter-cyclical. In equation (4.5), the SDF is monotone increasing in the conditional variance. If the conditional variance is counter-cyclical, then the SDF is counter-cyclical and co-varies negatively with the market return, even though aggregate consumption has low covariance with the market return. In principle, this negative co-variation gives rise to an equity premium that is absent in a complete market. Even if the conditional variance is not counter-cyclical, the cyclical or counter-cyclical behavior of some higher moment of the income shocks may generate a high equity premium. For example, in equation (4.5), the SDF is monotone decreasing in the conditional skewness. If the conditional skewness is cyclical, then the SDF is counter-cyclical and co-varies negatively with the market return, giving rise to an equity premium.

A good example of a major uninsurable income shock is job loss. Job loss is *uninsurable* because unemployment compensation is inadequate. Layoffs have *persistent* implications on household income, even though the laid-off workers typically find another job quickly.[20] Layoffs are *counter-cyclical* as they are more likely to occur in recessions.

The first implication of the theory is an explanation of the counter-cyclical behavior of the equity risk premium: the risk premium is highest in a recession because the stock is a poor hedge against the uninsurable income shocks, such as job loss, that are more likely to arrive during a recession.

The second implication is an explanation of the unconditional equity premium puzzle: even though per capita consumption growth is poorly correlated with stocks returns, investors require a hefty premium to hold stocks over short-term bonds because stocks perform poorly in recessions, when the investor is most likely to be laid off.

[20] The empirical evidence is sensitive to the model specification. Heaton and Lucas (1996) modeled the income process as *univariate* and provided empirical evidence from the Panel Study on Income Dynamics (PSID) that the idiosyncratic income shocks are transitory. Storesletten, Telmer, and Yaron (2001) modeled the income process as *bivariate* and provided empirical evidence from the PSID that the idiosyncratic income shocks have a highly persistent component that becomes more volatile during economic contractions. Storesletten, Telmer, and Yaron (2000) corroborated the latter evidence by studying household consumption over the life cycle.

Since the proposition demonstrates the existence of equilibrium in frictionless markets, it implies that the Euler equations of household (but not necessarily of per capita) consumption must hold. Furthermore, since the given price processes have embedded in them whatever predictability of returns by the price-dividend ratios, dividend growth rates, and other instruments that the researcher cares to ascribe to returns, the equilibrium price processes have this predictability built into them by construction.

Brav, Constantinides, and Geczy (2002) provided empirical evidence of the importance of uninsurable idiosyncratic income risk on pricing. They estimated the RRA coefficient and tested the set of Euler equations of household consumption on the premium of the value-weighted and the equally weighted market portfolio return over the risk free rate and on the premium of value stocks over growth stocks.[21] They did not reject the Euler equations of household consumption with RRA coefficient between two and four, although they rejected the Euler equations of per capita consumption with any value of the RRA coefficient. A RRA coefficient between two and four is economically plausible.

Open questions remain that warrant further investigation. According to the theory in Constantinides and Duffie (1996), periods with frequent and large uninsurable idiosyncratic income shocks are associated with both dispersed cross-sectional distribution of the household consumption growth and low stock returns. An interesting empirical question is which moments of the cross-sectional distribution of the household consumption growth capture the dispersion. Brav, Constantinides, and Geczy (2002) found that, in addition to the mean and variance, the *skewness* of the cross-sectional distribution is important in explaining the equity premium.

Krebs (2002) provided a theoretical justification as to why it is possible that neither the variance nor the skewness, but higher moments of the cross-sectional distribution are important in explaining the equity premium. He extended the

[21] In related studies, Jacobs (1999) studied the PSID database on food consumption; Cogley (2002) and Vissing-Jorgensen (2002) studied the CEX database on broad measures of consumption; Jacobs and Wang (2004) studied the CEX database by constructing synthetic cohorts; and Ait-Sahalia, Parker, and Yogo (2004) instrumented the household consumption with the purchases of certain luxury goods.

Constantinides and Duffie (1996) model by introducing rare idiosyncratic income shocks that drive consumption close to zero. In his model, the conditional variance and skewness of the idiosyncratic income shocks are nearly constant over time. Despite this, Krebs demonstrated that the original proposition of Constantinides and Duffie remains valid, that is, there exist household income processes, consistent with given aggregate income and dividend processes, such that equilibrium equity and bond price processes match the given equity and bond price processes. Essentially, he provided a theoretical justification as to why it may be hard to empirically detect the rare but catastrophic shocks in the low-order cross-sectional moments of household consumption growth.

A promising direction for future research is to address the relation between the equity return and the higher-order cross-sectional moments of household consumption with Monte Carlo methods. Another promising direction is to instrument the hard-to-observe time-series changes in the cross-sectional distribution with Labor Bureau statistics.

5. Concluding remarks

I examined the observed asset returns and conclude that the evidence does not support the case for abandoning the rational economic model. I argued that the standard model is greatly enhanced by relaxing some of its assumptions. In particular, I argued that we go a long way toward addressing market behavior by recognizing that consumers face uninsurable and idiosyncratic income shocks, for example, the loss of employment. The prospect of such events is higher in economic downturns and this observation takes us a long way toward understanding both the unconditional moments of asset returns and their variation along the business cycle.

I also argued that life cycle considerations are important and often overlooked in finance. Borrowing constraints become important when placed in the context of the life cycle. The fictitious representative consumer that holds all the stock market and bond market wealth does not face credible borrowing constraints. Young consumers, however, do face credible borrowing constraints. I traced their impact on the equity premium, the demand for bonds—who holds bonds if the equity premium is so high? —and on the limited participation of consumers in the capital markets.

Finally, I argued that by relaxing the assumption of convenience that preferences are time separable drives a wedge between the preference properties of risk aversion and intertemporal elasticity of substitution, within the class of von Neumann Morgenstern preferences. Further work along these lines may enhance our understanding of the price behavior along the business cycle with credibly low risk aversion coefficient.

I believe that the integration of the notions of *incomplete markets*, the *life cycle*, *borrowing constraints*, and other sources of *limited stock market participation* is a promising vantage point from which to study the prices of asset and their returns both theoretically and empirically within the class of rational asset pricing models.

REFERENCES

Abel, Andrew B., 1990, Asset prices under habit formation and catching up with the Joneses, *American Economic Review Papers and Proceedings* 80, 38-42.

Ait-Sahalia, Yacine, Jonathan A. Parker, and Motohiro Yogo, 2004, Luxury goods and the equity premium, *Journal of Finance*, forthcoming.

Aiyagari, Rao S., and Mark Gertler, 1991, Asset returns with transactions costs and uninsured individual risk, *Journal of Monetary Economics* 27, 311-331.

Altonji, Joseph G., Fumio Hayashi, and Laurence J. Kotlikoff, 1992, Is the extended family altruistically linked? *American Economic Review* 82, 1177-1198.

Attanasio, Orazio P., James Banks, and Sarah Tanner, 2002, Asset holding and consumption volatility, *Journal of Political Economy* 110, 771-792.

Attanasio, Orazio P., and Steven J. Davis, 1997, Relative wage movements and the distribution of consumption, *Journal of Political Economy* 104, 1227-1262.

Bakshi, Gurdip, and Zhiwu Chen, 1996, The spirit of capitalism and stock market prices, *American Economic Review* 86, 133-157.

Bansal, Ravi, and John W. Coleman, 1996, A monetary explanation of the equity premium, term premium, and risk-free rate puzzles, *Journal of Political Economy* 104, 1135-1171.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: a potential resolution of asset pricing puzzles, *Journal of Finance*, forthcoming.

Benartzi, Shlomo, and Richard H. Thaler, 1995, Myopic loss aversion and the equity premium puzzle, *Quarterly Journal of Economics* 110, 73-92.

Bertaut, Carol C., and Michael Haliassos, 1997, Precautionary portfolio behavior from a life-cycle perspective, *Journal of Economic Dynamics and Control* 21, 1511-1542.

Bewley, Truman F., 1982, Thoughts on tests of the intertemporal asset pricing model, Working paper, Northwestern University.

Blume, Marshall E., and Stephen P. Zeldes, 1993, The structure of stock ownership in the U.S., Working paper, University of Pennsylvania.

Bodie, Zvi, Robert C. Merton, and William F. Samuelson, 1992, Labor supply flexibility and portfolio choice in a life cycle model, *Journal of Economic Dynamics and Control* 16, 427-449.

Boldrin, Michel, Lawrence J. Christiano, and Jonas D. M. Fisher, 2001, Habit persistence, asset returns, and the business cycle, *American Economic Review* 91, 149-166.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy, 2002, Asset pricing with heterogeneous consumers and limited participation: Empirical evidence, *Journal of Political Economy* 110, 793-824.

Brav, Alon, and Christopher C. Geczy, 1995, An empirical resurrection of the simple consumption CAPM with power utility, Working paper, University of Chicago.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265-296.

Brown, Stephen J., William N. Goetzmann, and Stephen Ross, 1995, Survival, *Journal of Finance* 50, 853-873.

Cagetti, Marco, Lars Peter Hansen, Thomas Sargent, and Noah Williams, 2002, Robustness and pricing with uncertain growth, *Review of Financial Studies* 15, 363-404.

Campbell, John *Y.,* 2001, Asset pricing at the millennium, *Journal of Finance* 55, 1515-1567.

Campbell, John Y*.,* 2003, Consumption-based asset pricing, in George M. Constantinides, Milton Harris, and Rene Stulz, eds., *Financial Markets and Asset Pricing: Handbook of the Economics of Finance*, Volume 1B, Handbooks in Economics 21. Elsevier/North-Holland.

Campbell, John Y*.,* Joao F. Cocco, Francisco J. Gomes, and Pascal J. Maenhout, 2001, Investing retirement wealth: A life-cycle model, in, John Y. Campbell and Martin Feldstein, *Risk Aspects of Social Security Reform* (University of Chicago Press, Chicago).

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: a consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, NJ).

Cecchetti, S. G., P. Lam, and N. Mark (1994), The equity premium and the risk free rate: matching the moments, *Journal of Monetary Economics*, 31 21-46.

Chen, Xiaohong, and Sydney C. Ludvigson, 2004, Land of addicts? An empirical investigation of habit-based asset pricing models, working paper, New York University.

Cochrane, John H., 1991, A simple test of consumption insurance, *Journal of Political Economy* 99, 957-976.

Cochrane, John H., 1997, Where is the market going? Uncertain facts and novel theories, *Economic Perspectives* (Federal Reserve Bank of Chicago) 21, 3-37.

Cochrane, John H., 2001, *Asset Pricing* (Princeton University Press, Princeton, NJ).

Cochrane, John H., and Lars Peter Hansen, 1992, Asset pricing explorations for macroeconomics, in Olivier J. Blanchard and Stanley Fischer, eds., NBER, *Macroeconomics Annual* (MIT Press, Cambridge, MA).

Cogley, Timothy, 2002, Idiosyncratic risk and the equity premium: evidence from the consumer expenditure survey, *Journal of Monetary Economics*.

Constantinides, George M., 1982, Intertemporal asset pricing with heterogeneous consumers and without demand aggregation, *Journal of Business* 55, 253-267.

Constantinides, George M., 1990, Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519-543.

Constantinides, George M., 2002, Rational asset prices, *Journal of Finance* 57, 1567-1591.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra, 2002, Junior can't borrow: a new perspective on the equity premium puzzle*, Quarterly Journal of Economics* 117, 269-296.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra, 2004a, Junior must pay: pricing the implicit put in privatizing social security, *Annals of Finance*, forthcoming.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra, 2004b, Junior is rich: bequests as consumption, Working paper, University of Chicago.

Constantinides, George M., and Darrell Duffie, 1996, Asset pricing with heterogeneous consumers, *Journal of Political Economy* 104, 219-240.

Daniel, Kent, and David Marshall, 1997, The equity premium puzzle and the risk-free rate puzzle at long horizons, *Macroeconomic Dynamics* l, 452-484.

Danthine, Jean-Pierre, John B. Donaldson, and Rajnish Mehra, 1992, The equity premium and the allocation of income risk, *Journal of Economic Dynamics and Control* 16, 509-32.

Duesenberry, James S., 1949, *Income, Saving, and the Theory of Consumer Behavior* (Harvard University Press, Cambridge, MA).

Epstein, Larry G., and Stanley E. Zin, 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *Journal of Political Economy* 99, 263-286.

Ferson, Wayne E., and George M. Constantinides, 1991, Habit persistence and durability in aggregate consumption, *Journal of Financial Economics* 29, 199-240.

Ferson, Wayne E., and Campbell R. Harvey, 1992, Seasonality and consumption–based asset pricing, *Journal of Finance* 47, 511-552.

Friend, Irwin, and Marshall E. Blume, 1975, The demand for risky assets, *The American Economic Review* 65, 900-922.

Gabaix, Xavier, and David Laibson, 2001, The 6D bias and the equity premium puzzle, in Ben Bernanke and Ken Rogoff, eds., *NBER Macroeconomics Annual 2001* (MIT Press, Cambridge, MA).

Gallant, A. Ronald, Lars Peter Hansen, and George Tauchen (1990), Using conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution, *Journal of Econometrics* 45, 141-179.

Grossman, Sanford J., and Robert J. Shiller, 1981, The determinants of the variability of stock market prices, *American Economic Review Papers and Proceedings* 71, 222-227.

Guiso, Luigi, Michael Haliassos, and Tullio Jappelli, 2001, *Household Portfolios* (MIT Press, Cambridge, MA).

Haliassos, Michael, and Carol C. Bertaut, 1995, Why do so few hold stocks? *The Economic Journal* 105, 1110-1129.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225-262.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269-1288.

He, Hua, and David M. Modest, 1995, Market frictions and consumption-based asset pricing, *Journal of Political Economy* 103, 94-117.

Heaton, John C., 1995, An empirical investigation of asset pricing with temporally dependent preference specifications, *Econometrica* 63, 681-717.

Heaton, John C., and Deborah J. Lucas, 1996, Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy* 104, 443-487.

Heaton, John C., and Deborah J. Lucas, 1999, Stock prices and fundamentals, in Ben Bernanke and Julio Rotemberg, eds., *Macroeconomics Annual 1999,* National Bureau of Economic Research (MIT Press, Cambridge, MA).

Heaton, John, C., and Deborah J. Lucas, 2000, Portfolio choice and asset prices: The importance of entrepreneurial risk, *Journal of Finance* 55, 1163-.

Jacobs, Kris, 1999, Incomplete markets and security prices: Do asset-pricing puzzles result from aggregation problems? *Journal of Finance* 54, 123-163.

Jacobs, Kris, and Kevin Q. Wang, 2004, Idiosyncratic consumption risk and the cross-section of asset returns, *Journal of Finance*, forthcoming.

Jagannathan, Ravi, and Narayana R. Kocherlakota, 1996, Why should older people invest less in stocks than younger people? *Federal Bank of Minneapolis Quarterly Review* 20, 11-23.

Jorion, Philippe, and William N. Goetzmann, 1999, Global stock markets in the twentieth century, *Journal of Finance* 54, 953-980.

Kocherlakota, Narayana R., 1996, The equity premium: it's still a puzzle, *Journal of Economic Literature* 34, 42-71.

Krebs, Tom, 2002, Testable implications of consumption-based asset pricing with incomplete markets, *Journal of Mathematical Economics*, forthcoming.

Li, Yuming, 2001, Expected returns and habit persistence, *Review of Financial Studies* 14, 861-899.

Lucas, Robert, Jr., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429-1446.

Lynch, Anthony W., 1996, Decision frequency and synchronization across agents: Implications for aggregate consumption and equity returns, *Journal of Finance* 51, 1479-1497.

Mace, Barbara J., 1991, Full insurance in the presence of aggregate uncertainty, *Journal of Political Economy* 99, 928-956.

Mankiw, N. Gregory, 1986, The equity premium and the concentration of aggregate shocks, *Journal of Financial Economics* 17, 211-219.

Mankiw, N. Gregory, and Matthew D. Shapiro, 1986, Risk and return: consumption beta versus market beta, *Review of Economics and Statistics* 68, 452-459.

Mankiw, N. Gregory, and Stephen P. Zeldes, 1991, The consumption of stockholders and nonstockholders, *Journal of Financial Economics* 29, 97-112.

Marshall, Alfred, 1920, *Principles of Economics: An Introductory Volume* (Macmillan, London).

McGrattan, Ellen R., and Edward C. Prescott, 2001, Taxes, regulations, and asset prices, Working paper, Federal Reserve Bank of Minneapolis.

Mehra, Rajnish, 1998, On the volatility of stock prices: An exercise in quantitative theory, *International Journal of Systems Science* 29, 1203-1211.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145-161.

Mehra, Rajnish, and Edward C. Prescott, 1988, The equity premium: A solution? *Journal of Monetary Economics* 22, 133-136.

Mehra, Rajnish, and Edward C. Prescott, 2003, The equity premium in retrospect, in George M. Constantinides, Milton Harris, and Rene Stulz, eds., *Financial Markets and Asset Pricing: Handbook of the Economics of Finance*, Volume 1B, Handbooks in Economics 21. Elsevier/North-Holland.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112, 1-47.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887.

Negishi, Takashi, 1960, Welfare economics and existence of an equilibrium for a competitive economy, *Metroeconomica* 12, 92-97.

Pollak, Robert, A., 1970, Habit formation and dynamic demand functions, *Journal of Political Economy* 78, 745-763.

Poterba, James, M., 2001, The rise of the "equity culture": U.S. stockownership patterns, 1989-1998, Working paper, MIT.

Rietz, Thomas A., 1988, The equity risk premium: a solution, *Journal of Monetary Economics* 22, 117-131.

Rubinstein, Mark, 1976, The valuation of uncertain income streams and the pricing of options, *Bell Journal of Economics* 7, 407-425.

Ryder, Harl E., and Geoffrey M. Heal, 1973, Optimum growth with intertemporally dependent preferences, *Review of Economic Studies* 40, 1-43.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron, 2004, Consumption and risk sharing over the life cycle, *Journal of Monetary Economics*, forthcoming.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron, 2001, Asset pricing with idiosyncratic risk and overlapping generations, Working paper, Carnegie Mellon University.

Sundaresan, Suresh M., 1989, Intertemporally dependent preferences and the volatility of consumption and wealth, *Review of Financial Studies* 2, 73-89.

Vissing-Jorgensen, Annette, 2002, Limited asset market participation and the elasticity of intertemporal substitution, *Journal of Political Economy* 110, 825-853.

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401-421.