Federal Reserve Bank
of Minneapolis

Winter 1993

# Quarterly Review

Federal Reserve Bank of Minneapolis

# Quarterly Review Vol. 17, No. 1

ISSN 0271-5287

This publication primarily presents economic research aimed at improving policymaking by the Federal Reserve System and other governmental authorities.

*Any views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.*



The *Quarterly Review* is published by the Research Department of the Federal Reserve Bank of Minneapolis. Subscriptions are available free of charge.



Direct all comments and questions to

*Quarterly Review*
Research Department
Federal Reserve Bank of Minneapolis
P.O. Box 291
Minneapolis, Minnesota 55480-0291
(612-340-2341 / FAX 612-340-2366).

# Explaining Financial Market Facts: The Importance of Incomplete Markets and Transaction Costs*

S. Rao Aiyagari
Research Officer
Research Department
Federal Reserve Bank of Minneapolis

The average annual real return on short-term (90-day) U.S. Treasury bills over the period 1949–78 is less than 1 percent. On stocks over the same period, this return is about 7 percent.[1] These two facts have stimulated a lengthy discussion in the economics and finance literature, beginning with Mehra and Prescott 1985. The facts lead to the following natural questions: Did individuals anticipate these returns in making their consumption, saving, and portfolio decisions over the historical period? And will these average returns persist?

If individuals did not anticipate these returns, then we would expect their future consumption, saving, and portfolio behavior to change in light of these facts, thereby changing the average returns in the future. One possibility, for example, is that stocks have proved to be a lot more attractive compared to T-bills than people expected, thereby driving people to try to switch their portfolios in favor of stocks and away from T-bills. We would expect the market result of such behavior to be a decrease in the return on stocks and an increase in the return on T-bills.

Economists have attempted to explain the facts regarding average returns on the assumption that these returns were anticipated, rather than unanticipated. One reason for this assumption is that it is hard to believe that people systematically misperceive the average returns over long periods of time. (This is basically a weak form of rational expectations.) Another reason for the assumption is that if we were to allow ourselves the freedom to attribute misperceptions to people, then any observed pattern of average returns would be consistent with behavior. The view that the average returns were anticipated suggests that these returns will persist.

Given that the returns were anticipated, a natural way to explain the facts regarding average returns is to appeal to the different risk characteristics of stocks and T-bills. Presumably, stocks have higher average returns than T-bills in order to compensate the holders for bearing the additional risk. The simplest model that captures this explanation is the standard intertemporal model of asset pricing due to Lucas (1978). A key characteristic of this model is that it assumes *complete frictionless markets*. That is, individuals can buy and sell in all markets at given prices without being subject to borrowing or short-sales constraints and without incurring any transaction costs.

This is, indeed, the approach Mehra and Prescott (1985) tried. However, they found it impossible to generate the observed average returns using this model. Reasonably parameterized versions of the model predicted too small an equity premium (the excess average return on stocks over

*Isolated portions of this article are borrowed from an article in the *Journal of Monetary Economics* (June 1991, vol. 27, no. 3, pp. 311–31): "Asset Returns With Transactions Costs and Uninsured Individual Risk" by S. Rao Aiyagari and Mark Gertler, with the permission of Elsevier Science Publishers B. V. (North-Holland). 

[1]These figures are from Labadie 1989, p. 289. The stock return refers to the return on Standard & Poor's 500 common stock price index (the S&P 500 index).

the return on short-term T-bills) and too high a risk-free rate (the average real return on short-term T-bills).[2] The largest equity premium that Mehra and Prescott could generate from the model was 0.35 percent per year; the corresponding risk-free rate was about 4 percent per year. These results led Mehra and Prescott (1985, p. 145) to conclude that these return puzzles cannot be "accounted for by models that abstract from transactions costs, liquidity constraints and other frictions absent in the Arrow-Debreu set-up."

A number of researchers have attempted to save the complete frictionless markets framework. (See, for example, Epstein and Zin 1987; Nason 1988; Rietz 1988; Cecchetti, Lam, and Mark 1989; Labadie 1989; Weil 1989; and Constantinides 1990.) Even though these approaches have met with some limited success, in my view they are unsatisfactory because they continue to rely on the complete frictionless markets approach. The thrust of this paper is that any model that relies on the complete frictionless markets approach is bound to be highly unsatisfactory for explaining a number of empirical observations concerning the behavior of individual consumptions, wealths, portfolios, and asset market transactions. In my view, it is not sensible to divorce an explanation of asset returns from these other facts. I will argue that deviating from the complete frictionless markets framework is, therefore, both necessary and fruitful for solving the return puzzles as well as explaining these other phenomena. One possible reason that researchers have stuck to the complete frictionless markets approach is the considerable ease of obtaining qualitative and quantitative predictions from such models. However, while the task of analyzing models with incomplete markets and transaction costs is much more difficult, it is not impossible. (See, for example, Aiyagari and Gertler 1991.)

The plan of this article is as follows. I will first demonstrate the failure of the standard complete frictionless markets model to account for the low risk-free rate and the large equity premium. Then I will detail the variety of other empirical failures of this model concerning individual consumptions, wealths, portfolios, and asset market transactions. Then I will amend the standard model, by prohibiting some markets and introducing transaction costs, and explain how these amendments can move the model's predictions in the right directions: they lower the risk-free rate and enlarge the equity premium.[3] I will also show that the amended model is qualitatively consistent with several other facts concerning individual consumptions, wealths, portfolios, and asset market transactions that are anomalies in the context of the standard complete frictionless markets model. I will conclude with some suggestions for further work which I think will improve the match between theory and facts.

## A Failed Approach

In this section, I give a simple exposition of the standard complete frictionless markets model's failure to explain the level of the risk-free rate and the size of the equity premium. The exposition will serve to point up various other failures of this model and also to suggest ways of amending the model to simultaneously address the return puzzles as well as the model's other failures.

### *The Standard Model*

To begin the exposition, consider an economy with a large number (say, $I$) of individuals (indexed by $i$) who choose consumption over many periods and face uncertainty.

#### □ *Uncertainty*

A convenient way of representing uncertainty is via the concept of *states of the world.* A state of the world at a particular date corresponds to one of the many possible events that may occur at that date. These might be events like, there will (or will not) be an oil embargo, a war will (or will not) break out in the Middle East, Russia will (or will not) suffer an economic collapse, nuclear fusion will (or will not) become practical, a particular individual will (or will not) become sick, and so on. I will focus attention on two successive dates, labeled *1* and *2*, which should be thought of as representing a two-period segment of an ongoing economy. Assume that the state of the world is known at date 1 when decisions regarding consumption and asset trading are being made, but, of course, the state of the world at date 2 is uncertain. Let there be $J$ possible states of the world at date 2 (indexed by $j$), and let $\pi_j$ be the probability that state $j$ will occur (given the history of the world until and including date 1).

#### □ *Preferences*

Let $c_{i1}$ and $c_{i2}(j)$ be the consumptions of individual $i$ at date 1 and at date 2 in state $j$. An individual's preferences

---

[2]T-bills are only nominally risk free if they are held to maturity. To the extent that there is uncertainty regarding inflation, they are not risk free in real terms. However, over short periods of time, the inflation uncertainty is fairly small; therefore, we may regard short-term T-bills as essentially risk free in real terms as well.

[3]To see if this approach can fully solve the puzzles, of course, we must analyze the approach quantitatively as well as qualitatively. For a start on that sort of analysis, see Aiyagari and Gertler 1991.

over consumption at the two dates and in different states are represented by the following expression:

$$(1) \qquad U_i(c_{i1}) + \sum_j \pi_j U_i[c_{i2}(j)]/(1+\rho)$$

where $U_i(\cdot)$ represents the utility that the consumer gets from consumption and $\rho$ is the utility discount rate. To get total utility from consumption at the two dates and in all possible states of the world, then, add the utility from consumption at date 1 to the expected value of utility at date 2 (given by the sum of the utility in each state times the probability of that state) discounted by the utility discount rate $\rho$. The utility discount rate embodies the assumption that each unit of utility at date 2 is worth only $1/(1+\rho)$ units at date 1. If $\rho$ is positive (which is the natural assumption), this implies that the consumer is impatient since a consumer would rather have a unit of utility today than tomorrow. For this reason, $\rho$ is also referred to as the *rate of impatience* or the *rate of time preference*. The total utility expression in (1) should, again, be thought of as representing the consumer's preferences over the two-period segment that we are considering.

### □ Complete Markets

I now introduce the concept of *complete markets*. Imagine that at date 1 the consumer can purchase claims to state $j$ consumption in period 2 at the price $q_j$ in terms of date 1 consumption. As an example, the consumer might wish to purchase (at some price) claims to a certain number of gallons of gas at date 2 if there is a Middle East war and purchase (at a possibly different price) claims to a possibly different number of gallons of gas at date 2 if there is no Middle East war. It is important to understand that the price is paid at date 1 and that what the consumer gets in return is a *contingent* claim; that is, the claim to a certain number of gallons of gas in the event of a Middle East war is fulfilled if and only if that event occurs. Markets are said to be *complete* if such markets exist for all possible events at date 2 and beyond.

### □ Consumer Choice

When such markets exist, it is very easy to partially characterize the consumer's choice of consumptions at the two dates and in different states. This follows from the familiar principle of equalizing the ratio of marginal utilities between consumption at date 1 and consumption at date 2 in some state $j$ to the price of date 2 state $j$ consumption. From (1) this leads to

$$(2) \qquad \pi_j MU_i[c_{i2}(j)]/[(1+\rho)MU_i(c_{i1})] = q_j$$

where $MU$ denotes marginal utility. Implicit in equation (2) is the assumption that everyone is free to buy as well as sell as many claims to date 2 consumption in state $j$ as the individual can afford. That is, individuals do not face borrowing or short-sales constraints. Further, all claims are always fulfilled; that is, there is no default or bankruptcy.

### □ Asset Returns

I now show how the above framework allows for a very easy way to describe asset returns.

Essentially, an asset at date 1 can be represented by its returns at date 2 (per unit of consumption at date 1) in each of the possible states.[4] Therefore, an asset may be described by $\{r_j, j = 1,2,...,J\}$, where $r_j$ is the return on the asset at date 2 in state $j$. If the asset's return varies from state to state, it is a *risky* asset (like a stock). If its return is the same in every state, it is *risk free* (like a default-free bond). I will denote by $r^*$ the return on the risk-free asset at date 2 in every state $j$; the risk-free rate is defined to be $r^*$. The *risk premium* on an asset is, then, given by the excess of the expected return on the asset compared to the risk-free rate, that is, by $(\Sigma_j \pi_j r_j) - r^*$.

Asset returns and contingent claim prices are related as follows:

$$(3) \qquad \sum_j q_j(1+r_j) = 1.$$

The above relation arises from the following argument. A consumer can pay one unit of consumption at date 1 and acquire an asset that pays $1 + r_j$ in state $j$ at date 2. Alternatively, the consumer can purchase claims to $1 + r_j$ units of date 2 consumption in state $j$ for each $j$. Either way, the consumer has the same pattern of consumption across states at date 2. The cost of the former option is the right side of (3). The cost of the latter option is the left side of (3). Therefore, in the absence of any transaction costs, borrowing, or short-sales constraints, the equality (3) must obtain.[5]

---

[4]Note that this description includes long-lived assets which may pay returns beyond date 2. In this case, I include in the asset's return at date 2 (in some state) the price at which it will sell.

[5]The reason for insisting on the absence of borrowing or short-sales constraints is the following. Suppose, to take an example, that $\Sigma_j q_j(1+r^*) < 1$. Then the consumer could borrow one unit of consumption and purchase claims to $1 + r^*$ units of date 2 consumption in every state. The consumer will have $1 - \Sigma_j q_j(1+r^*)$ units of current consumption left over and still be able to pay off the loan at date 2 (from the proceeds

### Risk Aversion

I now introduce the concept of *risk aversion.* A consumer is risk averse if the individual prefers to receive a constant consumption $c$ in each state $j$ rather than to receive a pattern of consumption $\{c_j, j = 1,2,...,J\}$ across states which has an expected value, given by $\Sigma_j \pi_j c_j$, equal to $c$. A convenient utility function which exhibits risk aversion is given by

(4) $U(c) = (c^{1-\mu}-1)/(1-\mu)$

where $\mu \geq 0$ is known as the *relative risk-aversion coefficient.*[6] If $\mu$ is zero, then utility is linear in consumption and the consumer is indifferent between receiving $c$ for sure and having a pattern of consumption across states with the expected value $c$. Both of these yield the same level of utility, and the consumer is said to be *risk neutral.* If $\mu$ is positive, then the consumer always prefers the sure thing (since it yields higher utility) and is said to be *risk averse.*

For this type of utility function, marginal utility is given by $c^{-\mu}$. Using this, we can rewrite (2) as follows:

(5) $\pi_j[c_{i2}(j)/c_{i1}]^{-\mu_i}/(1+\rho) = q_j$

where $\mu_i$ is consumer $i$'s relative risk-aversion coefficient.

As an example of applying these equations, suppose that an individual is risk neutral; that is, $\mu_i$ is zero for some $i$. It follows from (5) that $q_j = \pi_j/(1+\rho)$. Plugging this into (3), we find that the expected return on every asset (and, hence, also the risk-free rate) must equal $\rho$. Risk premiums are zero when someone is risk neutral.

### A Representative Consumer

A simple trick now allows us to do away with the large number of possibly heterogeneous consumers and replace them with a single representative consumer. Define

(6) $c_1 = (c_{11} \times c_{21} \times c_{31} \times ... \times c_{I1})^{1/I}$

(7) $c_2(j) = [c_{12}(j) \times c_{22}(j) \times c_{32}(j) \times ... \times c_{I2}(j)]^{1/I}$

(8) $1/\mu = [(1/\mu_1) + (1/\mu_2) + (1/\mu_3) + ... + (1/\mu_I)]/I.$

Note that $c_1$ and $c_2(j)$ are the geometric means of individual consumptions at date 1 and at date 2 in state $j$, respectively. Therefore, we may think of these as per-capita consumptions. Further, $\mu$ is the harmonic mean of the individual risk-aversion coefficients and represents a sort of average risk-aversion coefficient. It is easy to show that using (6)–(8) we can rewrite (5) in terms of these means, or averages, as follows:

(9) $[c_{i2}(j)/c_{i1}]^{-\mu_i} = [c_2(j)/c_1]^{-\mu} = (1+\rho)q_j/\pi_j$

(10) $\pi_j[c_2(j)/c_1]^{-\mu}/(1+\rho) = q_j.$

Equations (10) and (3) show that as far as asset returns are concerned, we may as well imagine that the economy consists of a single representative agent who consumes the per-capita consumption in the economy and has a risk-aversion coefficient $\mu$.

From (10) we can obtain a fundamental equation relating the returns on assets to the growth rate of per-capita consumption. To do this, substitute for $q_j$ from (10) into (3) to get

(11) $\sum_j \{\pi_j[c_2(j)/c_1]^{-\mu}/(1+\rho)\}(1+r_j) = 1.$

Now note that $\Sigma_j \pi_j X_j$, where $X$ is any random variable taking the value $X_j$ in state $j$, is simply the expected value of $X$, which I will denote by $E(X)$. Letting $X_j$ be $(1+r_j) \times [c_2(j)/c_1]^{-\mu}$, we can rewrite the above equation to obtain the following fundamental equation relating the returns on assets to the growth rate of per-capita consumption:

(12) $E[(1+r)(c_2/c_1)^{-\mu}] = 1 + \rho.$

In (12), $r$ is the possibly random return on an asset (say, stocks), and $c_2/c_1$ is the possibly random gross growth rate in per-capita consumption.

We can now see the factors that determine the level of the risk-free rate and the size of the risk premium.

---

of contingent claims). That is, the consumer can make a pure arbitrage profit provided there is no restriction on borrowing. Similarly, if $\Sigma_j q_j(1+r^*) > 1$, the consumer can make a pure arbitrage profit provided there are no restrictions on short-selling contingent claims.

[6] This particular choice of the utility function is dictated by the facts concerning long-run growth. The facts are that over long periods of time, consumption of a typical individual and the real wage have grown at about the same rate, whereas individual hours worked has not changed much. If the utility function has the form $c^{\eta}v(T-n)$, where $T$ is individual time endowment and $n$ is individual hours worked (so that $T-n$ is individual leisure time), then the behavior of consumption, hours worked, and the real wage will be consistent with the growth facts. To see this, note that the marginal condition for hours worked versus leisure implies that $cv'(T-n)/[\eta v(T-n)] = w$, where $w$ is the real wage. Clearly, this condition is consistent with $c$ and $w$ growing at a common rate while $n$ does not change much. It turns out that the above utility function is the only one that is consistent with the growth facts. If we abstract from the choice of hours worked (which is not the focus of this article) by fixing $n$, then the resulting utility function defined over individual consumption has the form assumed above.

### The Risk-Free Rate

Applying (12) to the risk-free rate, we obtain

$$(13) \quad 1 + r^* = (1+\rho)/E[(c_2/c_1)^{-\mu}].$$

We can now derive the following approximate formula for the risk-free rate using (13) and some simplifying assumptions:[7]

$$(14) \quad r^* = \rho + \mu g - \mu^2\sigma^2/2.$$

In this formula, $\sigma$ is the standard deviation (S.D.) of per-capita consumption growth and $g$ is the average per-capita consumption growth. Clearly, $r^*$ depends on the value of $\rho$ (positively), on the average growth rate of per-capita consumption (positively), on the variance in the growth rate of per-capita consumption (negatively), and on the risk-aversion coefficient (ambiguously).

The reason for the positive dependence of $r^*$ on $\rho$ is the following. For a given growth rate of per-capita consumption, the higher is $\rho$ (that is, the more heavily individuals discount the future), the greater is individuals' preference for current consumption over future consumption. Therefore, the interest rate must be higher to make people accept the given growth rate and not borrow to have more current consumption.

The reason for the positive dependence of $r^*$ on the average growth rate of per-capita consumption is the following. For a given $\rho$, the higher is the growth rate of consumption, the higher is future consumption relative to current consumption. Therefore, the higher must the interest rate be to prevent people from trying to borrow in order to convert future consumption into current consumption.

The reason for the negative dependence of $r^*$ on the variance of per-capita consumption growth is the following. When there is uncertainty regarding the future, individuals typically attempt to save more. This increases lending and lowers the interest rate. The effect of the risk-aversion coefficient $\mu$ is ambiguous for the following reason. For small values of $\mu$, the positive effect of the average per-capita consumption growth dominates the negative effect of the variance in per-capita consumption growth. Hence, higher values of $\mu$ raise the risk-free rate. However, for large values of $\mu$, the latter effect dominates. Consequently, higher values of $\mu$ lower the risk-free rate.

### The Risk Premium

Now we ask what factors determine the risk premium on a risky asset. Intuitively, we know that since consumers are risk averse, they prefer to avoid random variations in consumption. If an asset has a pattern of returns such that it yields high returns when consumption is high and low returns when consumption is low, then holding such an asset tends to exacerbate consumption variability relative to holding the risk-free asset, which yields the same returns regardless of whether consumption is high or low. To compensate consumers for holding such an asset, its expected return has to be higher than the risk-free rate; that is, it must yield a positive risk premium. This same argument suggests that the risk premium on an asset need not be positive simply because its return is uncertain. Indeed, if an asset yields a pattern of returns such that its return is high when consumption is low and its return is low when consumption is high, then holding such an asset mitigates consumption variability relative to holding the risk-free asset. Consumers are willing to hold such an asset even if its expected return is lower than the risk-free rate. That is, such an asset will have a negative risk premium. Therefore, what matters for determining the risk premium on an asset is how the pattern of returns covaries with consumption. In particular, even if an asset is risky, as long as the return on the asset is uncorrelated with consumption, its expected return must equal the risk-free rate.

To see this point clearly, let $X$ denote $(c_2/c_1)^{-\mu}$, and note that $X$ is random since $c_2$ is random. Further, let $E(r)$ denote the expected return on the risky asset. Combining (12) and (13), we have

$$(15) \quad E(r) - r^* = -\text{cov}(r,X)/E(X).$$

Equation (15)[8] shows clearly that if an asset's return is positively correlated with consumption growth, so that $\text{cov}(r,X)$ is negative, then the asset will command a positive risk premium. We can obtain a more useful expression for the risk premium which permits us to see the role of the risk-aversion coefficient by making some simplifying assumptions and approximations. The expression is

---

[7]Assume that $\ln(c_2/c_1) = g + \varepsilon$, where $\varepsilon$ is a normally distributed random variable with mean zero and variance $\sigma^2$ and $g$ is average per-capita consumption growth. With this assumption, $E[(c_2/c_1)^{-\mu}] = \exp[-\mu g+\mu^2\sigma^2/2]$. Note that $\ln(c_2/c_1) \cong (c_2/c_1) - 1$, which is the growth rate in per-capita consumption. Also, approximate $\ln(1+r^*)$ and $\ln(1+\rho)$ by $r^*$ and $\rho$, respectively. The simplification is obtained by taking logs of both sides in (13) and using the above approximations and results.

[8]To see that equation (15) results from the combination of (12) and (13), note that (12) and (13) imply that $E(rX) = r^*E(X)$. Then note that $E(rX) = \text{cov}(r,X) + E(r)E(X)$.

given below:[9]

(16) $E(r) - r^* = \mu \text{cov}[\ln(1+r), \ln(c_2/c_1)].$

This expression shows clearly that the magnitude of the risk premium on an asset depends positively on the coefficient of relative risk aversion and on how strongly the return on the asset covaries with consumption growth.

### *The Model Faces the Facts*

We can now use equations (14) and (16) to understand why the empirically observed level of the risk-free rate and size of the equity premium are puzzling in light of the complete frictionless markets theory of the determination of asset returns described above.

Let's start with the risk-free rate. The growth rate of U.S. per-capita consumption (services plus nondurables) has averaged about 1.7 percent annually during 1950–88 with standard deviation equal to 0.0124. Empirically plausible values of $\mu$ are typically around 2 and seldom greater than 10. Estimated values of $\rho$ are in a range from 1 to 5 percent annually. Plugging these numbers into (14) shows that the annual risk-free rate ought to be at least 4.4 percent and possibly well over 10 percent. In fact, the annual risk-free rate is close to zero.

With regard to the equity premium, the magnitude of $\text{cov}[\ln(1+r), \ln(c_2/c_1)]$ is always less than $\text{S.D.}[\ln(1+r)] \times \text{S.D.}[\ln(c_2/c_1)]$. The standard deviation of stock returns has been about 7 percent annually over the postwar period.[10] As noted in the previous paragraph, the standard deviation of per-capita consumption growth has been about 0.0124. Plugging these numbers into (16) and using a value of 2 for the risk-aversion coefficient, we find that the predicted equity premium is about 0.17 percent. Even if we use a risk-aversion coefficient of 10, our predicted equity premium is still only about 0.85 percent annually. In fact, the observed equity premium is about 6 percent annually.

An alternative way to pose the return puzzles is to ask what value of $\mu$ would make the theoretical predictions match the facts. It turns out that typically the required values of $\mu$ are very large. To get the risk-free rate to be zero, we would need a value for $\mu$ of about 240.[11] To get the equity premium to be 6 percent, we would need a value for $\mu$ greater than 70. It seems very unlikely that consumers are so extremely risk averse. To get a feel for the degree of risk aversion involved, one can calculate that a consumer with a relative risk-aversion coefficient of 100 would prefer to take one unit of consumption for sure rather than face a lottery involving a 40 percent chance of getting 0.99 units of consumption and a 60 percent chance of getting 1 million units of consumption.

Yet another way to pose the return puzzles is to ask what the mean and the variance of the per-capita consumption growth rate would have to be to match the observed level of the risk-free rate and size of the equity premium. (See Hansen and Jagannathan 1991.) The resulting numbers turn out to be grossly counterfactual.

## Related Failures

Now I will detail many other empirical failures of the complete frictionless markets model. I will show that all of these empirical failures are related to this model's implication that individual consumptions move too closely with each other and with per-capita consumption. In the following section, I will show that introducing market incompleteness will avoid this implication and lead to a lower risk-free rate, introducing transaction costs will enlarge the equity premium, and adding both of these features will lead to better predictions for a variety of other facts.

### *The Model's Other Predictions . . .*

I will describe the following predictions of the complete frictionless markets theory:

- Individual consumptions ought to be perfectly correlated with each other and with per-capita consumption.
- Each individual's consumption ought to fluctuate as much as anyone else's and as much as per-capita consumption.
- An individual's position in society's wealth distribution should not vary much over time or across states of the world.

---

[9]To obtain (16), let $1 + r = [1 + E(r)]\exp[u - \sigma^2(u)/2]$, where $u$ is normally distributed with mean zero and variance $\sigma^2(u)$. Using the previous assumption regarding $\ln(c_2/c_1)$, we can write

$$\begin{aligned}(1+r^*)E(X) &= E[(1+r)X] \\ &= [1 + E(r)]E\{\exp[u - \sigma^2(u)/2 - \mu g - \mu\varepsilon]\} \\ &= [1 + E(r)]\exp[-\mu g + \mu^2\sigma^2/2 - \mu\text{cov}(u,\varepsilon)] \\ &= [1 + E(r)]E(X)\exp[-\mu\text{cov}(u,\varepsilon)].\end{aligned}$$

Canceling $E(X)$ from both sides, taking logs, and using the approximation $\ln(1+z) \cong z$ for small $z$, we obtain equation (16).

[10]This figure is calculated from the tables in Labadie 1989, p. 289, for the annual real return on the S&P 500 index during 1949–78.

[11]Note that very large values of $\mu$ will reduce the risk-free rate because of the negative effect of the term $-\mu^2\sigma^2/2$. See equation (14).

- Every individual must hold at least some amount of risky assets which have expected returns higher than the risk-free rate.
- Individuals have no need to engage in trading in asset markets, and any pattern of transaction volumes and transaction velocities among various securities is consistent with the theory.

To see the above implications of the theory, I repeat equation (9) here:

$$(17)\quad [c_{i2}(j)/c_{i1}]^{-\mu_i} = [c_2(j)/c_1]^{-\mu} = (1+\rho)q_j/\pi_j.$$

It can be seen from the above equation that all individuals' date 2 consumptions move up or down together in step with date 2 per-capita consumption. If $c_2(j)$ is higher than $c_2(k)$, where $j$ and $k$ are any two possible states of the world, then $c_{i2}(j)$ is higher than $c_{i2}(k)$ for all individuals $i$. Therefore, individual consumption growth rates must be perfectly correlated with each other as well as with the per-capita consumption growth rate.

Equation (17) also implies that if the risk-aversion coefficients (the $\mu_i$'s) are not too different across individuals, then each individual's consumption growth fluctuates as much as anyone else's and as much as per-capita consumption growth. As an example, an oil embargo must affect each individual's consumption in the same way and to the same extent as it does per-capita consumption. Further, events which affect an individual's personal circumstances (like health and employment) but have a negligible impact on per-capita consumption must also have a negligible impact on that individual's consumption. The reason is that the individual would have purchased claims to consumption in the event of a loss of health or employment in order to lessen the risk of consumption loss, and other individuals (who are not affected by the change in this individual's health or employment) would have sold such claims and shared a little bit of the risk.

Yet another implication of (17) is (again assuming that the risk-aversion coefficients are not too dissimilar) that there cannot be any rags-to-riches or riches-to-rags kinds of stories of individual fortunes or misfortunes. If an individual has higher consumption today than another individual, then the first individual will have higher consumption tomorrow (whatever the state of the world) than the second individual. Consequently, an individual who has higher wealth today than another individual will have higher wealth tomorrow (regardless of the state of the world) than the other individual. An individual's position in society's wealth distribution remains frozen over time and across states of the world.

Another implication of the complete frictionless markets theory is that every individual will hold at least some amounts of risky assets which have expected returns that are favorable, or exceed the risk-free rate. (This, by the way, is known as *Arrow's theorem*.) To understand this implication, consider an individual whose entire wealth (including human wealth) is currently held in the form of the risk-free asset. As a consequence, this person's second-period consumption is completely risk free; that is, $c_{i2}(j)$ is independent of $j$. Now consider the impact on the person's total utility if the individual were to purchase a small amount of the risky asset. This impact is given by

$$(18)\quad -MU_i(c_{i1}) + MU_i(c_{i2})\sum_j \pi_j(1+r_j)/(1+\rho).$$

The first term here is the loss in utility from reduced current consumption (due to the amount spent in purchasing the risky asset), and the second-term is the gain in utility from increased second-period consumption. The gain consists of the sum of the increases in second-period consumption in each state (which is $1+r_j$ in state $j$) times the probability of occurrence of each state, weighted by the marginal utility of date 2 consumption (which is independent of the state) and discounted by the rate of time preference $\rho$. The gain in total utility can be rewritten as

$$(19)\quad MU_i(c_{i1})\left[-1 + \sum_j \pi_j(1+r_j)/(1+r^*)\right]$$

since

$$(20)\quad 1 + r^* = (1+\rho)MU_i(c_{i1})/MU_i(c_{i2}).$$

Now note that the gain in total utility must be positive since, by assumption, the expected return on the risky asset exceeds the risk-free rate. Hence, the individual ought to be willing to purchase at least a small amount of the risky asset. In fact, if individual risk-aversion coefficients are not too different, then all individuals should hold approximately the same portfolios except for scale. That is, the proportions of wealth held in each asset must be roughly the same for all individuals since they all have roughly similar attitudes toward risk.

Yet another implication of the complete frictionless markets theory is that there is no need for individuals to engage in asset trading. As an example, there is nothing to gain from having a stock market in which individuals can buy and sell shares of stocks. Instead, individuals can simply hold on to their stock portfolio, receive the dividends, and buy or sell the appropriate amounts of various contingent claims to achieve the desired pattern of consumption over time and across states of the world. Consequently, the theory offers no explanation either for the specific volumes of trading in various asset markets or for the specific pattern of transaction velocities (turnover rates) of different types of assets.

### *. . . vs. The Facts*

Even casual observation suggests that every one of the above implications is grossly counterfactual. Casual empiricism as well as formal evidence indicates that individual consumptions are much more variable than aggregate consumption (Barsky, Mankiw, and Zeldes 1986 and Deaton 1991); further, individual consumptions are not very highly correlated, either with each other or with aggregate consumption. Individual specific circumstances appear to affect individual consumptions far more than is suggested by the theory.

The following evidence, described by Carroll (1991), indicates that individual wealth holdings are highly volatile. According to Avery and Kennickell (1989), 60 percent of U.S. households were in a different wealth decile in 1985 than in 1982. Approximately 30 percent moved up, and 30 percent moved down. Only people in the topmost and the bottommost deciles were more likely to stay put than to move to another decile. It would be hard to explain the movement of large fractions of households across the wealth distribution over such a short period of time (suggesting that the movement is not due to age and life cycle–related factors) if markets were complete and operated without frictions. Avery et al. (1984) present evidence to the effect that in the United States the ratio of median wealth to median income is higher for individuals in occupations with greater income uncertainty, for example, farmers and the self-employed. This suggests that the risks of such occupations are not being shared as the theory suggests.

Mankiw and Zeldes (1991) present evidence to the effect that only about 25 percent of U.S. households own any stocks. This seems to be in striking contrast to the theory's prediction that all individuals would hold at least some amount of stocks since the expected return on stocks is so much higher than the risk-free rate.

Further evidence reveals similar contradictions between the theory and the evidence concerning portfolios. According to evidence presented by Avery, Elliehausen, and Kennickell (1988), the ownership of stocks is highly concentrated at the top end of the wealth distribution, whereas the ownership of liquid assets is concentrated at the bottom end. They say, for example, that in 1963 the top 1 percent of U.S. wealth holders owned about 60 percent of all equity but only about 10 percent of all liquid assets. In contrast, the bottom 90 percent of households owned about 53 percent of all liquid assets and only about 9 percent of all equity. Greenwood (1983, pp. 34–35) presents similar evidence: in 1973 the top 5 percent of U.S. wealth holders owned about 85 percent of all corporate stocks and about 60 percent of all debt instruments. Finally, Kessler and Wolff (1991, p. 263) report that in the United States in 1983, the lowest wealth quintile's portfolio was over 80 percent liquid assets (currency, demand deposits, and time deposits), only about 9 percent financial securities and corporate stocks, and only about 3 percent other real estate (that is, not including housing) and unincorporated business. In contrast, the highest wealth quintile's portfolio was only about 15 percent liquid assets, about 22 percent financial securities and corporate stocks, and over 42 percent other real estate and unincorporated business.

Lastly, the vast amount of trading that takes place daily in stock markets and other financial markets would appear to be difficult to reconcile with the theory. In addition, there is a specific pattern to turnover rates among different types of assets; liquid assets (like savings and money market accounts) generally have much higher turnover rates than less liquid assets like stocks.[12] This fact suggests that transaction costs involved in borrowing and trading stocks are likely quite important.[13]

---

[12]According to Aiyagari and Gertler (1991), the ratio of shares sold over a year to the average number of shares listed for the year is about 0.5. Further, a substantial fraction of the volume is accounted for by institutional traders, which own about half of the outstanding shares. Turnover by households, who own the other half, is negligible. As a comparison, the equivalent turnover statistic is about 3 for savings accounts and about 7 for bank money market funds, indicating a substantially higher transaction velocity.

[13]Aiyagari and Gertler (1991) mention three basic kinds of (pretax) costs involved in trading stocks: brokerage commission costs, buy/sell spreads, and time involved in acquiring knowledge and in record keeping. At a deeper level, the existence of these costs reflects the informational frictions involved in trading heterogeneous assets like stocks. In addition, tax considerations are likely to be a factor since capital gains levies are based on realization rather than accrual. Restrictions on borrowing and short-selling are ubiquitous. The borrowing rate typically exceeds the lending rate substantially, and individuals are typically unable to borrow much against future earnings. For example, in the United States, the historical difference between the credit card rate and the risk-free rate is larger than 8 percent. Borrowing to buy stocks is subject to limits and margin requirements. See Aiyagari and Gertler 1991 for a detailed discussion of these costs.

## A Better Approach

Having thus far argued that the complete frictionless markets model fails not only in explaining asset returns but also in explaining individual consumptions, wealths, portfolios, and asset market transactions, I will now make two amendments to the simple standard model: prohibit some markets and introduce transaction costs. First I will explain the intuition behind how introducing these features might solve the equity premium and risk-free rate puzzles. Then I will describe how these same features can also successfully overturn several of the other empirical failures of the complete frictionless markets model. In doing so, I will be showing how departing from the complete frictionless markets framework provides a superior explanation not only of asset returns but also of a variety of other phenomena.

### *Overview*

In my amended model, individuals face idiosyncratic shocks to labor income. I assume that markets for claims on labor income do not exist. This assumption implies that individuals must *self-insure,* that is, buy and sell assets to smooth consumption. The model has two kinds of securities: stocks and short-term government bonds (T-bills). A key distinction between the two is that, by assumption, stocks are costly to trade while T-bills are not. T-bills can be thought of either as being directly held by households or as being costlessly repackaged by an intermediary which in turn issues costlessly tradable securities to its depositors. In the model, however, stocks cannot be repackaged. Regardless of whether T-bills are directly or indirectly held by households, in the model these bonds have an edge over stocks as a vehicle for self-insurance.

Introducing nontraded individual income risks lets the model generate a low risk-free rate—in equilibrium, a risk-free rate potentially well below the rate that would prevail with complete frictionless markets. Similarly, introducing costs of trading stocks along with uninsured individual risks lets the model enlarge the equity premium. Individuals trade in securities because of the need for self-insurance. In equilibrium, therefore, costs of trading become relevant to pricing securities. And since exchanging T-bills is costless, stocks must pay an added premium—a transaction/liquidity premium—to be competitive.

### *An Amended Model With Incomplete Markets . . .*

Consider a simple world in which a large number of infinitely lived individuals receive (perishable) labor income in each period.[14] Let $i$ index an individual and $t$ denote time (taking values 0,1,2,...), and use $y_t^i$ and $c_t^i$ to denote an individual's time $t$ labor income and consumption, respectively. Each person $i$ has preferences over consumption given by

$$(21) \quad E_0\left\{\sum\nolimits_{t=1}^{\infty}\beta^t U(c_t^i)\right\}$$

for $\beta = 1/(1+\rho)$ and $\rho > 0$, where we continue to assume that the utility function $U(c)$ is of the form $(c^{1-\mu}-1)/(1-\mu)$.

Suppose that labor income is random and is uncorrelated across individuals. Therefore, per-capita labor income (which equals per-capita consumption) is certain, while individual labor income is random. Let $y_t$ denote per-capita labor income, and suppose that it grows at the constant rate $g$ over time and that an individual's share in per-capita income, denoted $\theta_t^i$ ($\equiv y_t^i/y_t$), follows a stationary Markov process over time. Note that, as a consequence, an individual's income grows at the average rate $g$.

Since individual labor income is risky, individuals would like to purchase insurance against the possibility of receiving low labor income. One way to organize such an insurance market would be for individuals to pay a premium of $\bar{y}_t - y_t$ each period, where $\bar{y}_t$ is the maximum possible labor income in period $t$, and receive an insurance payment of $\bar{y}_t - y_t^i$ if their labor income is $y_t^i$. Note that such an insurance scheme is actuarially fair and provides complete insurance.[15] An individual's labor income, after the premium is paid and the insurance payment received, is always equal to per-capita labor income $y_t$.

Now suppose that such insurance markets do not exist. Instead, restrict individuals to trade only via credit markets in which they can borrow (up to some preestablished credit limit) or lend at a fixed interest rate. Also assume that the government has outstanding an amount of T-bills which is a constant proportion of per-capita income, that the T-bills are costlessly intermediated by money market mutual funds, and that individuals can hold money market accounts. Assume that government consumption is zero and that interest payments on T-bills are financed by a lump-sum tax which is a constant proportion of per-capita income and is identical across individuals.

---

[14]This is a slightly modified version of the model that was used in Aiyagari and Gertler 1991.

[15]This simplified insurance arrangement is equivalent to one in which individuals buy and sell claims contingent on the labor income realizations of all individuals in the economy.

The budget constraint that an individual faces is given by the following:

$$(22) \quad c_t^i + a_{t+1}^i = y_t^i + (1+r)a_t^i - \tau y_t$$

$$(23) \quad a_{t+1}^i \geq -ky_t$$

$$(24) \quad c_t^i \geq 0$$

where $a_t^i$ denotes liquid asset (money market account) holdings, $\tau$ is the ratio of taxes to per-capita income, and $ky_t$ is the credit limit. (Since an individual's income is growing at the average rate $g$, it seems reasonable to let the credit limit increase with income.) It is convenient to rewrite the individual budget constraint in terms of $\hat{c}_t^i$ ($\equiv c_t^i/y_t$) and $\hat{a}_{t+1}^i$ ($\equiv a_{t+1}^i/y_t$) by dividing equations (22)–(24) by $y_t$ as follows:

$$(25) \quad \hat{c}_t^i + \hat{a}_{t+1}^i = \theta_t^i + (1+r)\hat{a}_t^i/(1+g) - \tau$$

$$(26) \quad \hat{a}_{t+1}^i \geq -k$$

$$(27) \quad \hat{c}_t^i \geq 0.$$

The government budget constraint can be written as follows:

$$(28) \quad (r-g)\hat{b}/(1+g) = \tau$$

where $\hat{b}$ is the constant ratio of outstanding T-bills to per-capita income.[16] In equilibrium, the ratio of liquid assets held (net of borrowing) to per-capita income must equal $\hat{b}$, the ratio of T-bills to per-capita income.

### □ *Lowering the Risk-Free Rate*

If complete frictionless insurance markets existed, each individual's labor income would be the same as per-capita labor income and, consequently, each individual's consumption would have to be constant across states of the world and equal per-capita consumption. [See equation (17) and the ensuing discussion.] Consequently, the risk-free rate would have to equal $\rho + \mu g$ since per-capita consumption is nonrandom. [See equation (14).][17] However, with only credit markets, the return would have to be less than this. The reason is as follows.

With only credit markets, an individual can try to smooth consumption over time, that is, achieve the same average consumption next period as today by borrowing if today's income is low or lending if today's income is high. Note that since the shocks are idiosyncratic, if someone receives low income today (and needs to borrow), then someone else receives high income today (and wants to lend). But it would not be possible to smooth consumption across different states of the world tomorrow. This is because lending yields a fixed return tomorrow regardless of whether tomorrow's income is high or low. However, in an economy like the one above that goes on for many periods, one can borrow again (or lend again) when tomorrow comes if one's income turns out to be low (or high). This way, by repeatedly borrowing and lending, one can try to achieve smooth consumption across time as well as across states of the world.

Note that even if borrowing were prohibited (so that the credit limit $k$ were zero), one could achieve some degree of smoothness in consumption by building up or running down holdings of liquid assets. The larger the average amount of liquid assets held, the greater is the individual's ability to maintain smooth consumption in the face of fluctuating income. Such holdings of liquid assets are referred to as *precautionary holdings* since they are motivated by a desire to prevent the individual from suffering reduced consumption from a long series of low income realizations. Saving in order to build up such precautionary holdings of liquid assets is referred to as *precautionary saving*.[18]

To simplify matters, suppose that the credit limit is zero, so that there is no borrowing. In the absence of income uncertainty, an individual would not hold any liquid assets if the return were less than $\rho + \mu g$. This is because the individual would like to borrow, and since this is not allowed, he or she would simply run down holdings of liquid assets to zero.[19] However, if there is uncertainty in the individual's income, then this would not be sensible.

---

[16]This can be derived as follows: Let $B_t$ be the total amount of T-bills outstanding. Then the government budget constraint is $B_t(1+r) = \tau y_t + B_{t+1}$. Dividing this equation throughout by $y_t$, letting $B_t/y_{t-1} = B_{t+1}/y_t = \hat{b}$, and recalling that $y_t/y_{t-1} = 1 + g$, we obtain (28).

[17]Recall that equation (14) is an approximation derived from (13). The exact formula for the risk-free rate (in the absence of any randomness in consumption) is given by $1 + r = (1+\rho)(1+g)^{\mu}$.

[18]The common expression *saving for a rainy day* captures the essence of precautionary saving. The studies that contain analyses of models of precautionary saving with borrowing constraints include Schechtman and Escudero 1977; Bewley undated, 1984; and Clarida 1987, 1990. See Aiyagari 1992 for an exposition of such models and more extensive references.

[19]From (13) one can see that if borrowing is permitted and $r$ is less than $\rho + \mu g$, then $c_2/c_1$ must be less than $1 + g$. That is, individual consumption must grow slower than income. By borrowing, the individual increases current consumption at the expense of future consumption, thus lowering consumption growth.

The individual would maintain some positive holdings of liquid assets in order to serve as a buffer against low income realizations. The higher the return on the assets, the larger is the average holding of liquid assets (relative to per-capita income).

This relationship is depicted in the accompanying chart by the curve labeled $\hat{a}(r)$. The most important feature of this curve is that average liquid asset holdings tend to infinity as the risk-free rate approaches $\rho + \mu g$ from below. The reason for this is twofold. First, since there is always a probability (however small) of receiving a long string of low income realizations, it would not be possible to maintain a smooth consumption profile unless the individual had an infinite amount of liquid assets. Second, it is costless (at the margin) for the individual to acquire an additional unit of liquid assets. The loss in current utility, given by $MU$, is balanced by the expected discounted gain, $(1+r)MU'/(1+\rho)$, since $r = \rho + \mu g$ and $MU' = MU/(1+g)^{\mu}$.[20] These two factors lead the consumer to acquire a large amount of assets to maintain smooth consumption when $r = \rho + \mu g$.

The relation between the interest rate and the supply of liquid assets is also shown in the chart, by the curve labeled $\hat{b}(r)$. This relation is derived from the government budget constraint (28), which can be rewritten as

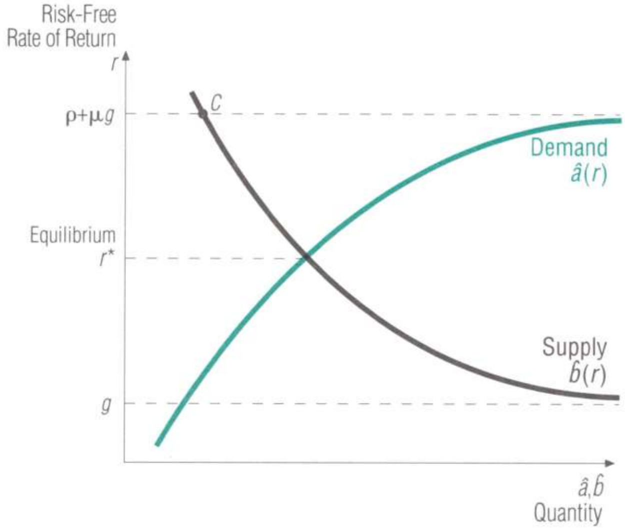
$$(29) \quad \hat{b}(r) = \tau(1+g)/(r-g).$$

In a steady-state equilibrium, the ratio of liquid assets to per-capita income that individuals desire to hold (net of borrowing) must equal the ratio of liquid assets to per-capita income supplied. As can be seen in the chart, the equilibrium risk-free rate, $r^*$, is below $\rho + \mu g$. The precautionary saving motive makes it desirable for individuals to maintain holdings of liquid assets even if the risk-free rate is below $\rho + \mu g$, whereas in the absence of income uncertainty (and, hence, a precautionary motive), individuals would not hold any liquid assets if the risk-free rate were below $\rho + \mu g$.[21] Thus, incomplete markets lower the risk-free rate via the precautionary motive.

What factors influence the amount of liquid assets held and the level of the risk-free rate? The greater the degree of income uncertainty, the larger the amount of liquid assets an individual needs to hold in order to provide a buffer. If income uncertainty increases, the $\hat{a}(r)$ curve will shift to the right, thereby lowering the risk-free rate and increasing liquid asset holdings. If individuals are permitted to borrow up to some limit, then they will reduce their liquid asset holdings. This is because they need not hold as large an amount of liquid assets since they can always borrow when they run out of liquid assets and receive low income realizations. In this case, the $\hat{a}(r)$ curve will shift to the left, raising the risk-free rate and reducing liquid asset holdings.

To see that the precautionary motive can potentially lower the risk-free rate significantly, recall equation (14) in which we now interpret $\sigma$ as the standard deviation of individual consumption growth rather than that of per-capita consumption growth. Because of incomplete markets, individual consumption growth can be considerably more volatile than per-capita consumption growth, and this greater volatility can lower the risk-free rate significantly. As an example, suppose that $\mu$ is 4 and $\sigma$ is 0.1, so

The Market for Liquid Assets†




†Recall that a security's return varies inversely with its price; that's why the demand and supply curves slope as they do.

[20]This follows from the fact that our choice of the utility function implies that $MU$ is given by $c^{-\mu}$ and that individual consumption will grow at the same rate $g$ as per-capita consumption if the individual is able to smooth consumption across states. See equation (21) and the discussion below.

[21]It follows that with complete markets, the steady-state equilibrium would be represented by the point $C$ in the chart. There the risk-free rate equals $\rho + \mu g$, and the amount of liquid assets held relative to per-capita income equals $\tau(1+g)/(\rho+\mu g-g)$ and, hence, is determined entirely by government policy without regard to the extent of uncertainty in individual incomes.

that individual consumption growth is about eight times more variable than per-capita consumption growth. It is easy to calculate that the last term in (14) can lead to a drop in the risk-free rate of about 8 percent.

### *. . . And With Transaction Costs*

Now consider what would happen if we introduced another type of asset (stocks) into this economy. Suppose that there exist $\bar{s}$ capital machines which costlessly produce output each period. The proceeds are distributed as dividends to shareholders who own the machines. There are $\bar{s}$ equity claims, which are tradable and perfectly divisible. One claim entitles the owner to $1/\bar{s}$ percent of the total output from all the machines each period. Assume that the output per machine is a constant proportion $d$ of per-capita income over time.[22]

Assume that the costs of trading stocks are proportional to the value of the trade.[23] Let $\alpha_b$ be the per-unit-of-value buying cost and $\alpha_s$ the per-unit-of-value selling cost. An individual $i$'s budget constraint is, then, modified to the following:

$$(30)\quad c_t^i + p_t(s_{t+1}^i - s_t^i) + a_{t+1}^i = y_t^i + s_t^i d y_t + (1+r)a_t^i - \tau y_t - \max[\alpha_b p_t(s_{t+1}^i - s_t^i), \alpha_s p_t(s_t^i - s_{t+1}^i)]$$

where $p_t$ is the price of equity. Short sales of stocks are not allowed, so that $s_t^i \geq 0$. The return to stocks $r_s$ is defined as follows:

$$(31)\quad 1 + r_s = (d_{t+1} + p_{t+1})/p_t.$$

### □ *Enlarging the Equity Premium*

In the absence of transaction costs, the return on stocks must equal the risk-free rate since stocks are also risk free. To see the role that transaction costs play in generating a spread between the returns to equity and liquid assets, consider an individual's decision whether to buy or sell stocks. There will be two levels of income, denoted $y_b(s,a)$ and $y_s(s,a)$, with $0 < y_s(s,a) < y_b(s,a)$, such that the individual sells stocks whenever income is below $y_s$, holds stocks when it is between $y_s$ and $y_b$, and buys stocks when it is above $y_b$. Notice that, in general, these regions will depend on the individual's initial holdings of stocks and liquid assets, $s$ and $a$, respectively.

Arbitrage requires that any individual buying both stocks and bonds at time $t$ must be indifferent between acquiring either kind of asset at the margin. Therefore, for each person $i$ in this position at $t$, the following Euler conditions must hold (where the $i$ superscripts for agents are dropped for convenience):

$$(32)\quad MU = \beta(1+r)E(MU')$$

$$(33)\quad (1+\alpha_b)p_t MU = \beta\{\pi^b[d_{t+1} + (1+\alpha_b)p_{t+1}]E_b(MU') + \pi^s[d_{t+1} + (1-\alpha_s)p_{t+1}]E_s(MU') + \pi^h[d_{t+1} + (1+\lambda_h)p_{t+1}]E_h(MU')\}$$

where $\pi^b$, $\pi^s$, and $\pi^h$ are the probabilities the individual will be buying, selling, or holding stocks next period; $E_b$, $E_s$, and $E_h$ are the expectations conditional on buying, selling, or holding stocks next period; $MU$ and $MU'$ are the marginal utilities of consumption this period and the next, respectively;[24] and the number $\lambda_h$ satisfies

$$(34)\quad \alpha_b > \lambda_h > -\alpha_s.$$

The left side of equation (33) is the cost of buying stocks, and the right side is the expected marginal gain, after factoring in transaction costs. Note that the marginal gain depends on whether and how the individual will be adjusting his or her stock holdings in the subsequent period. The marginal value of stocks in the subsequent period is $(1+\alpha_b)p_{t+1}$ for someone who is buying stocks, and it is $(1-\alpha_s)p_{t+1}$ for someone who is selling. For someone holding, it lies between the buying and the selling price, at $(1+\lambda_h)p_{t+1}$. (In general, $\lambda_h$ will depend on whether the individual expects to be buying or selling down the road.) Everything else equal, the larger is $\pi^s$, the smaller is the expected marginal benefit from purchasing stocks. The unattractive aspect of turning around and selling the stocks in the subsequent period is having to incur the transaction cost.

In a steady-state equilibrium, $p_t$ will be proportional to total dividends and, hence, to per-capita income. Therefore, $p_{t+1}/p_t$ will equal $1 + g$. Dividing through equation (33) by $p_t$ and using (31), we can express (33) in the following form:

---

[22] I have abstracted from dividend uncertainty so that we can isolate the impact of the frictions we have introduced. Since there is no dividend risk, any spread between the returns on stocks and bonds is due only to the transaction costs operating in conjunction with the uninsured individual income risk.

[23] Proportional as well as fixed costs are considered in Aiyagari and Gertler 1991.

[24] Note that the Euler conditions for agents who are selling stocks or who are subject to borrowing or short-sale constraints will be different from (32) and (33).

$$(35)\quad (1+\alpha_b)MU = \beta\{\pi^b[1 + r_s + \alpha_b(1+g)]E_b(MU')$$
$$+ \pi^s[1 + r_s - \alpha_s(1+g)]E_s(MU')$$
$$+ \pi^h[1 + r_s + \lambda_h(1+g)]E_h(MU')\}$$
$$= \beta[(1+r_s)E(MU')$$
$$+ \pi^b\alpha_b(1+g)E_b(MU')$$
$$- \pi^s\alpha_s(1+g)E_s(MU')$$
$$+ \pi^h\lambda_h(1+g)E_h(MU')]$$
$$\leq \beta[(1+r_s)E(MU')$$
$$+ \pi^b\alpha_b(1+g)E_b(MU')$$
$$- \pi^s\alpha_s(1+g)E_s(MU')$$
$$+ \pi^h\alpha_b(1+g)E_h(MU')]$$
$$= \beta\{(1+r_s)E(MU')$$
$$+ \alpha_b(1+g)[E(MU') - \pi^sE_s(MU')]$$
$$- \pi^s\alpha_s(1+g)E_s(MU')\}$$
$$= \beta E(MU')[1 + r_s + \alpha_b(1+g)$$
$$- (\alpha_b+\alpha_s)(1+g)\pi^s$$
$$\times E_s(MU')/E(MU')].$$

Substituting for $MU$ from (32) into (35) and rearranging, we obtain the following inequality:

$$(36)\quad r_s - r \geq (r-g)\alpha_b + (1+g)\pi^s(\alpha_b+\alpha_s)E_s(MU')/E(MU').$$

Quite clearly, the transaction costs are responsible for the spread between the returns to stocks and bonds.[25] The spread is increasing in $\alpha_b$, $\alpha_s$, and $\pi^s$. Further, it is likely to be larger the more risk averse the individual; this is because sales of stocks are likely when consumption is low, which makes the utility measure of the transaction costs of selling (relatively) high. The lower bound for the spread equals $\pi^s(\alpha_b+\alpha_s)$, the probability of selling times the round-trip transaction cost.[26] This value arises (approximately) when individuals are risk neutral and when the shadow value of stocks for someone holding stocks is arbitrarily close to its upper bound, $(1+\alpha_b)p$.

In summary, the incompleteness of markets for insurance implies a lower risk-free rate of interest compared to that in the complete frictionless markets case. Further, the existence of trading costs for stocks in conjunction with the need to trade securities to smooth consumption can introduce a spread between the returns to stocks and bonds. In this way, the combination of incomplete markets and transaction costs can potentially provide an explanation for the low risk-free rate and the large equity premium.

### *Related Successes*

I will now show that the combination of these features can also account qualitatively for the behavior of individual consumptions, wealths, and portfolios and the pattern of transactions in asset markets.

The incompleteness of markets makes individual consumptions imperfectly correlated with each other as well as with aggregate consumption. This occurs because individuals who receive high labor income will be increasing their consumption and adding to their asset holdings whereas individuals who receive low labor income will be forced to reduce consumption somewhat and liquidate some assets.[27] Also, obviously, individual consumptions can fluctuate considerably more than aggregate consumption since individual consumptions respond to individual-specific circumstances that have no impact on aggregate consumption.

In addition, the incompleteness of markets leads to diversity among individuals in total wealth as well as in portfolios. An individual who has had a run of high labor income will have high wealth and enjoy high consumption compared to another individual who has suffered a run of bad luck. As a consequence, there is considerable mobility of individuals across the wealth and income distributions.

Individuals with high wealth are likely to hold relatively more stocks in their portfolios than individuals with low wealth since the former have a greater ability to cush-

---

[25]Note that the first term on the right side of (36) is an order of magnitude smaller than the other terms in the expression since it is a product of two small terms, the difference between the risk-free rate and the growth rate and the transaction cost incurred in buying stocks. Therefore, it is safe to neglect this term in the following discussion.

[26]The argument presumes that $E_s(MU') \geq E(MU')$; that is, marginal utility conditional on selling is at least as high as unconditional marginal utility, based on the idea that sales occur when consumption is low. Note that $\pi^s$ is the probability next period of selling the marginal unit of the stocks purchased this period, as opposed to selling all the stocks purchased this period. Note also that $\pi^s$ will vary across individuals depending on their portfolios.

[27]Note that in the present model in which there is no uncertainty in aggregate labor income, individual consumptions will be uncorrelated with each other since the shocks to labor income are idiosyncratic, that is, uncorrelated across individuals. If we also introduce some randomness in aggregate labor income, then individual consumptions will tend to move together with aggregate consumption whenever there is a shock to aggregate labor income. Thus, the combination of idiosyncratic and aggregate shocks will make individual consumptions positively, but less than perfectly, correlated with each other as well as with aggregate consumption.

ion their consumption in the event of low labor income by selling off liquid assets without incurring transaction costs.

Moreover, the turnover rate for liquid assets will be higher than that for stocks. Since the return on stocks is higher and trading stocks is subject to costs, the individual will engage in stock trading only as a last resort and will try for the most part to smooth consumption by buying and selling liquid assets.[28]

## Future Research

The basic model with incomplete markets and transaction costs outlined in this article could be extended in several obvious ways to try to further lower the risk-free rate and enlarge the equity premium.

One of these extensions is to allow for additional heterogeneity among the agents in the model, in the form of stockholders versus nonstockholders.[29] This can be modeled as arising from either differential proportional costs or differential fixed costs of participating in or trading in the stock market.[30] The group of agents facing low costs will likely have portfolios consisting mostly of stocks, whereas the other group facing high costs will likely have portfolios consisting mostly of liquid assets. This feature will likely enlarge the equity premium arising from the model since the group that would like to hold more stocks due to the higher return faces high costs of participating in or trading in the stock market.

Another possible extension is to include a transaction motive for holding liquid assets in addition to precautionary considerations. Certainly a component of household holdings of savings and money market accounts stems from transaction needs. By enhancing the demand for liquid assets, this feature is likely to lower the risk-free rate arising from the model.

An implication of the incomplete markets and transaction cost model described in this article is that a reduction in transaction costs ought to shrink the equity premium. Many dramatic changes have occurred in financial markets in the United States and abroad over the last 15 years due to improvements in technology (especially in communication) and loosening of regulations. Many of these changes have resulted in lower costs of transacting in financial markets. An empirical investigation of whether these changes have resulted in a smaller equity premium seems worthwhile.

It would also be interesting to explicitly study the reasons for the absence of markets. In my incomplete markets model with transaction costs, no reason was given for the absence of insurance markets. One way to motivate the absence of such markets is to assume that individual labor income is *private information,* that is, cannot be known (at least costlessly) by anyone else. The insurance arrangement I outlined above is not feasible since the insurance company cannot verify the individual's labor income. The individual will always claim to have received the lowest labor income regardless of the actual labor income, and the market will collapse. Presumably, some other arrangement that respects the privacy of information regarding individual labor income will arise, but there is no reason why it should look exactly like the credit market arrangement. (For work along these lines, see Townsend 1979, Green 1987, Phelan and Townsend 1991, and Atkeson and Lucas 1992.) Therefore, the implications for asset returns are likely to be different from those arising from the credit market arrangement.

While incomplete markets models with transaction costs are considerably more difficult to analyze qualitatively and quantitatively than are complete frictionless markets models, the task is not impossible. In Aiyagari and Gertler 1991, a model of the type I have described here was used to conduct a number of quantitative experiments and study the effects on asset/income ratios and transaction velocities. However, the model abstracted from aggregate dividend risk. It would be highly desirable to allow for aggregate dividend risk, so that some portion of the equity premium reflects the riskiness of stock returns. There appear to be significant computational difficulties in taking this feature into account. Some research effort is currently being devoted to overcoming these difficulties, and I hope to be able to report on the progress of this work in the not-too-distant future.

---

[28]See Aiyagari and Gertler 1991 for a quantitative analysis of the type of model described here. That study uses specific functional forms and parameter values and reports the implications of the model for asset/income ratios and relative transaction velocities.

[29]Mankiw and Zeldes (1991) also emphasize (though for somewhat different reasons) the importance of distinguishing between individuals who regularly hold stocks and those not inclined to do so.

[30]As noted earlier, there appear to be dramatic differences in the portfolios held by individuals with different wealth levels. A substantial fraction of liquid assets is held by a group of households who own relatively little stocks, and the ownership of stocks is heavily concentrated. Recall, for example, that Avery, Elliehausen, and Kennickell (1988) estimate that in 1963 the bottom 90 percent of the U.S. wealth distribution held 53 percent of the total quantity of liquid assets but only 9 percent of the equity, while the top 1 percent held over 60 percent of the equity but only 10 percent of liquid assets.

# References

Aiyagari, S. Rao. 1992. Uninsured idiosyncratic risk and aggregate saving. Research Department Working Paper 502. Federal Reserve Bank of Minneapolis.

Aiyagari, S. Rao, and Gertler, Mark. 1991. Asset returns with transactions costs and uninsured individual risk. *Journal of Monetary Economics* 27 (June): 311–31.

Atkeson, Andrew, and Lucas, Robert E., Jr. 1992. On efficient distribution with private information. *Review of Economic Studies* 59 (July): 427–53.

Avery, Robert B.; Elliehausen, Gregory E.; Canner, Glenn B.; and Gustafson, Thomas A. 1984. Survey of consumer finances, 1983. *Federal Reserve Bulletin* 70 (September): 679–92.

Avery, Robert B.; Elliehausen, Gregory E.; and Kennickell, Arthur B. 1988. Measuring wealth with survey data: An evaluation of the 1983 survey of consumer finances. *Review of Income and Wealth* 34 (December): 339–69.

Avery, Robert B., and Kennickell, Arthur B. 1989. Measurement of household saving obtained from first differencing wealth estimates. Manuscript. Presented at the 21st General Conference of the International Association for Research in Income and Wealth, Lahnstein, Germany, August 20–26, 1989. Board of Governors of the Federal Reserve System.

Barsky, Robert B.; Mankiw, N. Gregory; and Zeldes, Stephen P. 1986. Ricardian consumers with Keynesian propensities. *American Economic Review* 76 (September): 676–91.

Bewley, Truman. Undated. Interest bearing money and the equilibrium stock of capital. Manuscript. Yale University.

__________. 1984. Notes on stationary equilibrium with a continuum of independently fluctuating consumers. Manuscript. Yale University.

Carroll, Christopher D. 1991. Buffer stock saving and the permanent income hypothesis. Working Paper 114. Board of Governors of the Federal Reserve System.

Cecchetti, Stephen; Lam, Pok-sang; and Mark, Nelson C. 1989. The equity premium and the risk free rate: Matching the moments. Manuscript. Ohio State University.

Clarida, Richard H. 1987. Consumption, liquidity constraints and asset accumulation in the presence of random income fluctuations. *International Economic Review* 28 (June): 339–51.

__________. 1990. International lending and borrowing in a stochastic, stationary equilibrium. *International Economic Review* 31 (August): 543–58.

Constantinides, George, M. 1990. Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy* 98 (June): 519–43.

Deaton, Angus. 1991. Saving and liquidity constraints. *Econometrica* 59 (September): 1221–48.

Epstein, Larry G., and Zin, Stanley E. 1987. Substitution, risk aversion, and the temporal behavior of asset returns: II: An empirical investigation. Manuscript. University of Toronto and Queen's University.

Green, Edward J. 1987. Lending and the smoothing of uninsurable income. In *Contractual arrangements for intertemporal trade,* ed. Edward C. Prescott and Neil Wallace, pp. 3–25. Minnesota Studies in Macroeconomics, Vol. 1. Minneapolis: University of Minnesota Press.

Greenwood, Daphne. 1983. An estimation of U.S. family wealth and its distribution from microdata, 1973. *Review of Income and Wealth* 29 (March): 23–44.

Hansen, Lars Peter, and Jagannathan, Ravi. 1991. Implications of security market data for models of dynamic economies. *Journal of Political Economy* 99 (April): 225–62.

Kessler, Denis, and Wolff, Edward N. 1991. A comparative analysis of household wealth patterns in France and the United States. *Review of Income and Wealth* 37 (September): 249–66.

Labadie, Pamela. 1989. Stochastic inflation and the equity premium. *Journal of Monetary Economics* 24 (September): 277–98.

Lucas, Robert E., Jr. 1978. Asset prices in an exchange economy. *Econometrica* 46 (November): 1429–45.

Mankiw, N. Gregory, and Zeldes, Stephen P. 1991. The consumption of stockholders and nonstockholders. *Journal of Financial Economics* 29 (March): 97–112.

Mehra, Rajnish, and Prescott, Edward C. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15 (March): 145–61.

Nason, James M. 1988. The equity premium and time-varying risk behavior. Finance and Economics Discussion Paper 11. Board of Governors of the Federal Reserve System.

Phelan, Christopher, and Townsend, Robert M. 1991. Computing multi-period information-constrained optima. *Review of Economic Studies* 58 (October): 853–81.

Rietz, Thomas A. 1988. The equity risk premium: A solution. *Journal of Monetary Economics* 22 (July): 117–31.

Schechtman, Jack, and Escudero, Vera L. S. 1977. Some results on "An income fluctuation problem." *Journal of Economic Theory* 16 (December): 151–66.

Townsend, Robert M. 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory* 21 (October): 265–93.

Weil, Philippe. 1989. The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24 (November): 401–21.