Asset Pricing Implications of Pareto Optimality with Private Information

Narayana Kocherlakota

University of Minnesota, Federal Reserve Bank of Minneapolis, and National Bureau of Economic Research

Luigi Pistaferri

Stanford University, National Bureau of Economic Research, and Centre for Economic Policy Research

We compare the empirical performance of a standard incomplete markets asset pricing model with that of a novel model with constrained Pareto-optimal allocations. We represent the models' stochastic discount factors in terms of the cross-sectional distribution of consumption and use these representations to evaluate the models' empirical implications. The first model is inconsistent with the equity premium in the United States, United Kingdom, and Italy. The second model is consistent with the equity premium and the risk-free rate in all three countries if the coefficient of relative risk aversion is roughly 5 and the quarterly discount factor is less than 0.5.

I.   Introduction

In the United States, over the period 1980–2004, the average return to
stocks was about 8 percent per year higher than the average return to
Treasury bills. The *equity premium puzzle* (Mehra and Prescott 1985) is
that this gap in average returns is too large to be justified by the size
of the covariance between stock returns and per capita consumption
growth. One way to see the difficulty is to consider the necessary con-
dition of optimality,

$$E\left[\left(\frac{C_{1,t+1}}{C_{1,t}}\right)^{-\gamma}(R_{t+1}^{s} - R_{t+1}^{b})\right] = 0, \tag{1}$$

for a representative household. Here, γ is the household's coefficient
of relative risk aversion, $C_{1,t}$ is per capita consumption in period t,
R_{t+1}^{s} is the realized gross real return to stocks in period $t+1$, and
R_{t+1}^{b} is the gross real return to Treasury bills. The restriction (1) is
implied by what we call the *representative agent* model. Under weak reg-
ularity conditions, the unconditional expectation in (1) can be consis-
tently estimated using a time-series average. Kocherlakota (1996) doc-
uments that the sample analogue of this first-order condition is not
satisfied in U.S. data (from 1889 to 1979) except for very large values
of γ.

   The equity premium puzzle implicitly assumes that per capita con-
sumption growth is a good proxy for individual household consumption
growth. In reality, households' consumptions grow at different rates. In
this paper, we explore the equity premium implications of two different
models of asset trade among (ex post) heterogeneous households. In
both models, the source of this ex post heterogeneity is that households
are hit by skill shocks with realizations that are private information to
them. In the first model, households can trade only securities that pay
off contingent on publicly observable aggregate shocks. We call this
model the *incomplete markets* model. In this model, each household's
consumption satisfies the Euler equation (1).

   In the second model, households first commit to a lifetime risk-shar-
ing and work arrangement with some intermediary (such as a family,
employer, or insurance company). The intermediaries structure this
contract so as to provide incentives and insurance at minimal cost. The
intermediaries then trade the same set of securities as in the first model;
households are not allowed to trade assets on their own. In this latter
model, unlike the first one, equilibrium allocations are Pareto optimal
(given the private information friction). Hence, we term this model the
private information Pareto-optimal (PIPO) model. The control of
households' asset market behavior by intermediaries implies that house-

hold consumption does *not* satisfy (1). Instead, household consumption satisfies an alternative restriction known as the inverse Euler equation (Golosov, Kocherlakota, and Tsyvinski 2003).

We derive the two models' implications for the equity premium by aggregating these household-level restrictions. We show that, regardless of the underlying stochastic process generating skills, in the incomplete markets model, equilibrium quantities and returns must satisfy the restriction

$$E\left[\frac{C_{-\gamma,t+1}}{C_{-\gamma,t}}\left(R^s_{t+1} - R^b_{t+1}\right)\right] = 0, \tag{2}$$

where $C_{\eta,t}$ is the ηth noncentral moment of the cross-sectional distribution of consumption in period t. In the PIPO model, equilibrium quantities and returns must satisfy the restriction

$$E\left[\frac{C_{\gamma,t}}{C_{\gamma,t+1}}\left(R^s_{t+1} - R^b_{t+1}\right)\right] = 0. \tag{3}$$

Again, this restriction is a robust one, in the sense that it does not depend on the underlying stochastic process generating skills. If all households' consumptions always grow at the same rate, then these model implications are both equivalent to the first-order condition of a standard representative household. However, if households' consumption growth rates differ, (2) and (3) differ from each other and from the representative household's first-order condition (1). Importantly, we show that both formulas are still valid with mismeasured consumption, for a wide class of measurement error processes.

We demonstrate that the incomplete markets and PIPO models differ in the way that consumption inequality affects the pricing of assets. The incomplete markets model implies that the equity premium is high if the *left* tail of the cross-sectional consumption distribution is heavier when stock returns are low. Intuitively, the left tail grows heavier if there is a lot of downside idiosyncratic risk. For precautionary savings reasons, agents value bonds more highly if they pay off well in these states compared to stocks.

In contrast, if $\gamma > 1$ (the empirically relevant case), the PIPO model implies that the equity premium is high if the *right* tail of the cross-sectional consumption distribution is heavier when stock returns are high. This implication is driven by incentive considerations. If the right tail of the cross-sectional consumption distribution is very heavy, then many skilled people have very low consumption. (Think about the extreme case in which virtually all consumption is in the hands of one person.) Because of the wealth effects in labor supply, it is cheap to

provide incentives to these low-consumption skilled people.[1] This cheapness in providing incentives means that intermediaries do not put a high value on stocks if they pay off in those states in which the consumption distribution has a heavy right tail.

We evaluate these differing empirical implications for three different countries, United States, United Kingdom, and Italy (chosen on the basis of the availability of household consumption data for a sufficiently long period of time), by constructing sample analogues of (1)–(3). We use observations from the Consumer Expenditure Survey (CEX) in the United States, the Family Expenditure Survey (FES) in the United Kingdom, and the Survey of Family Budgets (SFB) in Italy to form estimates of cross-sectional moments of consumption at each date in each country. Note that an important advantage of our stochastic discount factors is that they can be estimated using repeated cross-sectional data (such as the FES and SFB), without any need for a panel component.

We then apply generalized method of moments (GMM) to the sample analogues of (1)–(3) to estimate γ in each country. We find that in all three countries, there is no value of γ that satisfies sample analogues of either (1) or (2).[2] In contrast, the estimates of γ that we obtain using (3) in each country are clustered around the value 5. In fact, when using (3), we are unable to reject the null hypothesis that the value of γ is the same across countries. We conclude that if households have a coefficient of relative risk aversion of about 5, the PIPO model is able to rationalize the magnitude of the equity premium in all three countries.

We extend this analysis by jointly estimating γ and a value for households' common discount factor β. To do so, we use the above implications together with the models' implications for the pricing of short-term government debt. In all three countries, the resulting estimate of β from the PIPO model is less than 0.52. At the quarterly frequency, such an estimate strikes us as implausibly low.

II. Prior Literature

Our paper is related to Ligon (1998), which tests the risk-sharing implications of Pareto optimality with moral hazard. Ligon's approach is as follows. He uses consumption data from South Indian villages (the data set from the International Crops Research Institute for the Semi-

---

[1] Our representations for the stochastic discount factors presume that preferences are separable between consumption and leisure. Such preferences immediately imply that there are nonzero wealth effects in labor supply.

[2] The only exception is the representative agent model in the U.S. case, where, consistent with previous research, we find that a very high value of γ (53 in our case) satisfies the sample analogue of (1).

Arid Tropics). He assumes that there is a risk-neutral banker outside the villages, agents in the village have the same discount rate as the interest rate offered by the outside banker, and all agents have coefficient of relative risk aversion $\gamma > 0$. He asks if the allocation of risk within the village is better described as being Pareto optimal, given moral hazard, or as the result of risk-free borrowing and lending. He answers this question by estimating the parameter a from the following moment restriction:

$$E_t\left\{\left(\frac{c_{it+1}}{c_{it}}\right)^a\right\} = 1. \tag{4}$$

Under the former hypothesis of constrained Pareto optimality, $a = \gamma$. Under the latter hypothesis of risk-free borrowing and lending, $a = -\gamma$. Using GMM, he estimates a to be positive and interprets this as demonstrating the relative empirical relevance of constrained Pareto optimality.

Our approach bears some similarity to Ligon's. But there are important differences. First, our theoretical analysis is more general than his. We allow for aggregate shocks and do not assume that there is a risk-neutral outsider. Hence, we are able to allow for nontrivial movements in expected asset returns. Also, we do not need to assume that individual productivity shocks are independent and identically distributed (i.i.d.) over time (as he does). This assumption of i.i.d. productivity shocks is at odds with the data (Meghir and Pistaferri 2004). Second, our testable implications are in terms of the cross-sectional consumption distribution, not individual consumption; we do not need to have panel data on consumption. Finally, our empirical analysis is more robust to measurement error than his.[3]

Our work is also related to recent papers using data from the CEX to evaluate incomplete markets models of asset pricing. In recent work, Brav, Constantinides, and Geczy (2002), Cogley (2002), and Vissing-Jorgensen (2002) use data from the CEX to test the hypothesis that asset prices and household consumption are consistent with an incomplete markets equilibrium. These papers basically proceed as follows. They select all households from the CEX that have two or more periods of observations. They next construct an intertemporal marginal rate of substitution (IMRS) in a given period for each household with observations for that period and the prior one (for Brav et al., a period is a quarter, and for Vissing-Jorgensen, a period is a half year). Finally, they construct a theoretically valid stochastic discount factor by averaging

---

[3] Ligon (1996) discusses some ideas for dealing with measurement error, persistence, and relaxing risk neutrality in his Sec. 6. This working paper is the basis for Ligon (2005).

these IMRSs across households. (Henceforth, we term this the *average IMRS* stochastic discount factor, and we use the acronym SDF to refer to stochastic discount factors more generally.)

Note that the average IMRS SDF is not the same as the incomplete markets SDF ($\beta C_{-\gamma,t+1}/C_{-\gamma,t}$) described in the introduction. The average IMRS SDF used in the prior literature is the average of the ratios of marginal utilities. Our incomplete markets SDF is instead the ratio of averages of marginal utilities. In an incomplete markets economy, with no binding borrowing constraints, both SDFs are valid, but they are not necessarily the same. One needs at least some panel data to implement the average IMRS SDF empirically, whereas one needs only a time series of cross sections to implement the incomplete markets SDF.

The findings of this recent work are somewhat mixed. Cogley (2002) argues that the average IMRS SDF does not provide much additional explanatory power over the representative agent SDF in terms of the equity premium. In contrast, Brav et al. (2002) find that the average IMRS SDF does a good job of rationalizing the equity premium. These differences could be explained by differences in the sample period used, sample selection, and the nature of the approximation adopted. Vissing-Jorgensen (2002) considers different samples of households depending on the size of their position in the asset market. She finds that the (log-linearized) average IMRS SDF is a valid SDF for *smaller* values of γ when the average is constructed using samples of agents with *larger* asset positions.[4]

Balduzzi and Yao (2007) also use data from the CEX. Like us, they examine the performance of the incomplete markets SDF, not the average IMRS SDF. They find that if the coefficient of relative risk aversion is around 8 or higher, a log-linearized version of the SDF can account for the equity premium if they restrict attention to households that have $2,000 or more in total financial assets.

Our work is novel relative to these other papers in several respects. Like Kocherlakota and Pistaferri (2007), we consider the implications of the PIPO model as well as the implications of the more traditional incomplete markets formulation. However, our focus is on asset pricing rather than on real exchange rates.[5] We look at the implications of the models in the United Kingdom and Italy, not just the United States. Finally, with the exception of Brav et al., these other papers rely on

[4] Vissing-Jorgensen (2002) does not look at restrictions involving multiple assets (such as the equity premium restriction [1]). Instead, she examines the asset pricing equations for bonds and stocks separately.

[5] Kocherlakota and Pistaferri (2007) show that the PIPO model does a better job of accounting for real exchange rate movements than the incomplete markets or representative agent models. That paper builds on this one, in the sense that the first versions of the present paper predate the first versions of that paper. Kocherlakota (1998) analyzes a calibrated two-period version of the PIPO model.

Taylor series approximations of the relevant SDFs. The errors in these approximations may lead to biases in the results. As opposed to dealing with potential outliers in an ad hoc fashion (by discarding data or by using approximations to the theory), we instead deal with them by placing no restriction on the marginal distribution of the measurement errors.

III. Environment

In this section, we describe the environment. The description is basically the same as that in Kocherlakota (2005) and Kocherlakota and Pistaferri (2007).

Consider an economy with $T + 1$ periods. Time is indexed 0, 1, 2, ..., T; period 0 is a contracting/trading period in which no production or consumption takes place. There is a unit measure of agents. A typical agent has expected utility preferences, with cardinal utility function

$$\sum_{t=1}^{T} \beta^{t-1} \left[\frac{c_t^{1-\gamma}}{1-\gamma} - v(l_t) \right], \quad 0 < \beta < 1, \ 0 < \gamma.$$

Here, c_t is the agent's period t consumption. The variable l_t represents the agent's labor.

There are two kinds of shocks in the economy: public aggregate shocks and private idiosyncratic shocks. The first kind of shock works as follows. Let Z be a finite set, and let Ψ be a probability density over Z^T that assigns positive probability to all elements of Z^T. At the beginning of period 1, an element z^T of Z^T is drawn according to Ψ. The random vector z^T is the sequence of public aggregate shocks; z_t is the realization of the shock in period t.

The idiosyncratic shocks work as follows. Let Θ be a finite set and let π be a probability density defined over Θ^T. At the beginning of period 1, an element of Θ^T is drawn for each agent according to the density π. Conditional on z^T, the draws are independent across agents, and π is the same for all realizations of z^T; we require $\pi(\theta^T) > 0$ for all θ^T in Θ^T. We assume that a law of large numbers applies across agents: conditional on any z^T, the measure of agents in the population with type θ^T is given by $\pi(\theta^T)$.

Any given agent learns the realization of z_t and his own θ_t at the beginning of period t and not before. Thus, at the beginning of period t, the agent knows his own private history $\theta^t = (\theta_1, \ldots, \theta_t)$ and the history of public shocks $z^t = (z_1, \ldots, z_t)$. This implies that his choices in period t can be a function only of this history.

The individual-specific and aggregate shocks jointly determine skills. In period t, an agent produces output y_t according to the function

$$y_t = \phi_t(\theta^t, z^t)l_t, \tag{5}$$

$$\phi_t : \Theta^t \times Z^t \to (0, \infty). \tag{6}$$

We assume that an agent's output in either sector is observable at time t, but his labor input is known only to him. We refer to $\phi_t(\theta^t, z^t)$ as an agent's skill in history (θ^t, z^t).

An important element of our analysis is the flexible specification of the stochastic process generating skills. This flexibility takes two forms. First, we are agnostic about the time-series properties of the skill shocks. This generality is crucial, given the current empirical debate about the degree of persistence of individual wages. In particular, we are able to allow for the possibility that individual skills may be at once persistent and stochastic. Both aspects seem to be important empirically. Second, it has been argued by Storesletten, Telmer, and Yaron (2001) that the cross-sectional variance of wages is higher in recessions than in booms. Thus, the cross-sectional variance of skills varies with aggregate conditions. We can capture this possibility in our setting, because $\mathrm{Var}\,[\phi_t(\theta^t, z^t)|z^t]$ may depend on z_t. The idea here is that the range of ϕ, as a function of θ^t, can be allowed to depend on z_t.[6]

We define an allocation in this society to be $(c, y) = (c_t, y_t)_{t=1}^T$, where

$$c_t : \Theta^t \times Z^t \to R_+ \tag{7}$$

and

$$y_t : \Theta^t \times Z^t \to [0, \bar{y}]. \tag{8}$$

Here, $y_t(\theta^t, z^t)$ is the amount of output that an agent with shock history θ^t produces, given that the public shock history is z^t. We define an allocation (c, y) to be *feasible* if for all t, z^t,

$$\sum_{\theta^t \in \Theta^t} c_t(\theta^t, z^t)\pi(\theta^t) \le \sum_{\theta^t \in \Theta^t} y_t(\theta^t, z^t)\pi(\theta^t). \tag{9}$$

Because θ_t is only privately observable, allocations must respect incentive-compatibility conditions. (The following definitions corre-

---

[6] Attanasio and Davis (1996) document that the cross-sectional dispersion of consumption increased in the 1980s in the United States along with the publicly observable change in the cross-sectional dispersion of wages. Sometimes this finding is interpreted as being evidence that individuals cannot insure themselves against publicly observable shocks. But, as Attanasio and Davis themselves point out, these movements are also consistent with the hypothesis that the increase in the cross-sectional variance of *measured* wages was associated with an increase in the variance of private information about skills. Again, we can specify our function ϕ_t so as to capture this possibility.

spond closely to those in Golosov et al. [2003].) A *reporting strategy* $\sigma : \Theta^T \times Z^T \to \Theta^T \times Z^T$, where σ_t is (θ^t, z^t)–measurable and $\sigma(\theta^T, z^T) = (\bar{\theta}^T, z^T)$ for some $\bar{\theta}^T$. An agent can alter his consumption and labor by changing his reporting strategy. Suppose that the allocation is (c, y). If an agent uses reporting strategy σ, his consumption in history (θ^t, z^t) is

$$c_t(\sigma^t(\theta^T, z^T)),$$

where $\sigma^t(\theta^T, z^T) = (\sigma_s(\theta^T, z^T))^t_{s=1}$, and his labor is

$$\frac{y_t(\sigma^t(\theta^T, z^T))}{\phi_t(\theta^T, z^T)} .$$

Let $V(\sigma; c, y)$ be the ex ante utility that an agent receives from using strategy σ. Let σ_{TT} be the truth-telling strategy $\sigma_{TT}(\theta^T, z^T) = (\theta^T, z^T)$ for all θ^T, z^T. Then, an allocation (c, y) is *incentive compatible* if

$$V(\sigma_{TT}; c, y) \geq V(\sigma; c, y) \quad \text{for all } \sigma \text{ in } \Sigma.$$

An allocation that is incentive compatible and feasible is said to be incentive feasible.

IV. Asset Pricing

Given this definition of the environment, we now examine the properties of asset prices in two different trading setups. In the first setup, agents trade a complete set of z^t-contingent securities directly with one another. In the second setup, intermediaries trade a complete set of z^t-contingent securities on behalf of the agents. In this latter formulation, the intermediaries directly insure the agents against the realization of the θ shocks, given that they are private information. (We think of these intermediaries as being employers, insurers, or families.) We derive useful necessary conditions of equilibrium in both settings. In these necessary conditions, the cross-sectional moments of consumption play a key role. As in the introduction, we use the notation $C_{\eta,t}$ to refer to the ηth noncentral moment of the cross-sectional distribution of consumption.

A. Incomplete Markets Equilibrium

Suppose that at each date $t = 0, 1, 2, \ldots, T - 1$, all agents can buy and sell a complete set of z^{t+1}-contingent claims to consumption. Fix agents' labor choices at their individually optimal levels, and let $W_{t+1}(\theta^t, z^{t+1})$ denote the portfolio of z^{t+1}-contingent claims to consumption purchased

by an agent with skill history θ^t. Then, the choice problem of a typical agent is

$$\max_{c,W} E_0 \sum_{t=1}^T \frac{\beta^{t-1}c_t^{1-\gamma}}{1-\gamma}$$

subject to

$$c_t(\theta^t,\ z^t) + \sum_{z_{t+1}} q_{t+1}(z_{t+1}|z^t)W_{t+1}(\theta^t,\ z^{t+1}) \le y_t(\theta^t,\ z^t) + W_t(\theta^{t-1},\ z^t),$$

$$\sum_{z_1} q_1(z_1)W_1(z_1) \le W_0,$$

$$W_{T+1}(\theta^T,\ z^T) \ge 0 \quad \text{for all } (\theta^T,\ z^T),$$

$$W_0 \text{ given,}$$

where $q_{t+1}(z_{t+1}|z^t)$ is the history z^t price of z^{t+1}-contingent consumption. Given this choice problem, we can readily show that agents must satisfy the following first-order condition with respect to W and c:

$$q_{t+1}(z_{t+1}|z^t)c_t(\theta^t,\ z^t)^{-\gamma} = \beta\Psi(z_{t+1}|z^t)\sum_{\theta_{t+1}} \pi(\theta_{t+1}|\theta^t)c_{t+1}(\theta^{t+1},\ z^{t+1})^{-\gamma}. \quad (10)$$

If we take expectations over θ^t, we get

$$q_{t+1}(z_{t+1}|z^t)\sum_{\theta^t} \pi(\theta^t)c_t(\theta^t,\ z^t)^{-\gamma} =$$

$$\beta\Psi(z_{t+1}|z^t)\sum_{\theta^{t+1}} \pi(\theta^{t+1})c_{t+1}(\theta^{t+1},\ z^{t+1})^{-\gamma}, \quad t \ge 1.$$

This expression involves unconditional expectations of marginal utilities. However, the law of large numbers implies that the fraction of agents in history z^t with skill history θ^t is given by the unconditional probability $\pi(\theta^t)$. Hence, we can apply a law of large numbers to conclude that the negative γth moment of the cross-sectional distribution of consumption in public history z^t is

$$C_{-\gamma,t}(z^t) = \sum_{\theta^t} \pi(\theta^t)c_t(\theta^t,\ z^t)^{-\gamma}.$$

We can conclude that, in equilibrium, the price q_{t+1} satisfies

$$q_{t+1}(z_{t+1}|z^t) = \beta\Psi(z_{t+1}|z^t)\frac{C_{-\gamma,t+1}(z^{t+1})}{C_{-\gamma,t}(z^t)}.$$

B.   *Private Information—Pareto-Optimal Equilibrium*

It is well known from the work of Green (1987) and others that there are other incentive-compatible allocations that Pareto dominate the above equilibrium allocation. In this subsection, we describe an alternative trading mechanism, for which it is known that the equilibrium allocation is (constrained) Pareto optimal. This trading protocol is similar to that described originally in Atkeson and Lucas (1992). In period 0, insurers compete with one another by offering contracts. These contracts specify consumption and output as a function of agents' reports. The insurers then trade a complete set of z^{t+1}-contingent claims to consumption at each date t. However, the agents do not engage in this trade.

In what follows, we derive necessary conditions of a PIPO equilibrium. Let y^* be the output process specified by the equilibrium contract. As above, let $q_1(z_1)$ be the period 0 price of z_1-contingent consumption and $q_{t+1}(z_{t+1}|z^t)$ be the price of z^{t+1}-contingent consumption in history z^t. We know that in the contract market, the agents receive some equilibrium level of utility. The profit-maximizing insurance companies structure contracts so as to minimize the costs of providing that utility. Hence, the equilibrium consumption contract c^* solves the problem

$$\min_{c,W} \sum_{z_1} q_1(z_1) W_1(z_1)$$

subject to

$$\sum_{\theta^t} \pi(\theta^t) c_t(\theta^t,\ z^t)\ +\sum_{z_{t+1}} q_{t+1}(z_{t+1}|z^t) W_{t+1}(z^{t+1}) \leq$$

$$\sum_{\theta^t} \pi(\theta^t) y_t^*(\theta^t,\ z^t)\ +\ W_t(z^t),\quad t \geq 2,$$

$$\sum_{t=1}^{T} \beta^{t-1} \sum_{z^t} \Psi(z^t) \sum_{\theta^t} \pi(\theta^t) \frac{c_t(\theta^t,\ z^t)^{1-\gamma}}{1-\gamma}\ =\ u_*,$$

$$V(c,\ y^*;\ \sigma_{TT}) \geq V(c,\ y^*;\ \sigma)\quad \text{for all } \sigma.$$

We can derive a useful necessary condition of insurer optimality as follows. Suppose that $(c^*,\ W^*)$ solves the insurer's problem. Fix a public history $\bar{z}^{t+1}$. Consider a perturbation in which we lower $c_t^*(\theta^t,\ \bar{z}^t)$ by ε and raise $W_{t+1}^*(\bar{z}^{t+1})$ by δ. We can then use this extra payoff in period $t + 1$ to raise $c_{t+1}^*(\theta^{t+1},\ \bar{z}^{t+1})$ by $\eta(\theta^{t+1})$, where

$$\sum_{\theta^{t+1}} \eta(\theta^{t+1}) \pi(\theta^{t+1})\ =\ \delta$$

and for all θ^{t+1}

$$u(c_t^*(\theta^t, \bar{z}^t) - \varepsilon) + \beta\Psi(\bar{z}_{t+1}|\bar{z}^t)u(c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1}) + \eta(\theta^{t+1})) =$$

$$u(c_t^*(\theta^t, \bar{z}^t)) + \beta\Psi(\bar{z}_{t+1}|\bar{z}^t)u(c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1})).$$

This perturbation is akin to that in Kocherlakota (2005), which extends the approach originally taken by Rogerson (1985). The perturbation is budget feasible for the insurer, does not violate the incentive constraints of the agents, and delivers the same ex ante utility to the agents. Hence, if (c^*, W^*) is truly optimal for the insurer, the perturbation must not be cheaper.

We can reinterpret this verbal argument mathematically. It tells us that setting ε, δ, and η equal to zero must solve the problem

$$\min_{\varepsilon,\delta,\eta} -\varepsilon\pi(\theta^t) + \delta q_{t+1}(\bar{z}_{t+1}|\bar{z}^t)$$

subject to

$$\frac{[c_t^*(\theta^t, \bar{z}^t) - \varepsilon]^{1-\gamma}}{1-\gamma} + \beta\Psi(\bar{z}_{t+1}|\bar{z}^t)\frac{[c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1}) + \eta(\theta^{t+1})]^{1-\gamma}}{1-\gamma} =$$

$$\frac{c_t^*(\theta^t, \bar{z}^t)^{1-\gamma}}{1-\gamma} + \beta\Psi(\bar{z}_{t+1}|\bar{z}^t)\frac{c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1})^{1-\gamma}}{1-\gamma} \quad \text{for all } \theta^{t+1},$$

$$\sum_{\theta^{t+1}} \pi(\theta^{t+1})\eta(\theta^{t+1}) = \delta.$$

The necessary conditions associated with this optimum are

$$c_t^*(\theta^t, \bar{z}^t)^{-\gamma}\sum_{\theta_{t+1}} \lambda(\theta^{t+1}) = \pi(\theta^t)$$

and

$$q_{t+1}(\bar{z}_{t+1}|\bar{z}^t)\pi(\theta^{t+1}) = \beta\lambda(\theta^{t+1})\Psi(\bar{z}^{t+1}|\bar{z}^t)c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1})^{-\gamma} \quad \text{for all } \theta^{t+1},$$

where $\lambda(\theta^{t+1})$ is the multiplier on the first constraint.

We can combine these necessary conditions to obtain a useful representation for the SDF in this trading structure. Substituting the second condition into the first, we get

$$\pi(\theta^t)c_t^*(\theta^t, \bar{z}^t)^\gamma = \frac{q_{t+1}(\bar{z}_{t+1}|\bar{z}^t)}{\Psi(\bar{z}_{t+1}|\bar{z}^t)}\beta^{-1}\sum_{\theta_{t+1}} \pi(\theta^{t+1})c_{t+1}^*(\theta^{t+1}, \bar{z}_{t+1})^\gamma \qquad (11)$$

for all θ^t. If we add over θ^t and use the law of large numbers as before, we obtain

$$q_{t+1}(\bar{z}_{t+1}|\bar{z}^t) = \frac{\beta C_{\gamma,t}(\bar{z}^t)\Psi(\bar{z}_{t+1}|\bar{z}^t)}{C_{\gamma,t+1}(\bar{z}^{t+1})},$$

where $C_{\gamma,t}(z^t) = \sum \pi(\theta^t) c_t(\theta^t, z^t)^\gamma$ is the γth cross-sectional moment of consumption. Obviously, this argument can be generalized to any z^{t+1}.

C. *Failure of Consumption Smoothing*

In a PIPO equilibrium, individual consumption obeys the restriction (11). Suppose that Z is a singleton. Then, this restriction implies that

$$c_t^*(\theta^t)^\gamma = \beta^{-1} q_{t+1} E\{c_{t+1}^*(\theta^{t+1})^\gamma | \theta^t\}, \tag{12}$$

where q_{t+1} is the inverse of the gross rate of return. This restriction (12) is sometimes referred to as the inverse Euler equation. For incentive reasons, we would expect that consumption in period $t + 1$ should respond to new information contained in the realization of θ_{t+1}. Hence, $\text{Var}(c_{t+1}^* | \theta^t) > 0$, and so, because of Jensen's inequality, (12) is inconsistent with the usual Euler equation:

$$q_{t+1} c_t^*(\theta^t)^{-\gamma} = \beta E\{c_{t+1}^*(\theta^{t+1})^{-\gamma} | \theta^t\}. \tag{13}$$

Note that (13) is equivalent to (10) when Z is a singleton. In other words, incomplete markets equilibrium and PIPO equilibrium are inconsistent with one another, as long as there is consumption risk at some date.

In the remainder of this subsection, we provide some basic intuition for why consumption smoothing (i.e., [13]) is not satisfied by equilibrium contracts in the PIPO market structure. For simplicity, assume that Θ has two elements and Z has one. Suppose that an intermediary offers a consumption contract c so that

$$u'(c_t(\theta^t)) = \beta R \pi_H u'(c_{t+1}(\theta^t, \theta_H)) + \beta R \pi_L u'(c_{t+1}(\theta^t, \theta_L))$$

and

$$u'(c_{t+1}(\theta^t, \theta_H)) < u'(c_{t+1}(\theta^t, \theta_L)).$$

(Here, π_i represents the probability of $\theta_{t+1} = \theta_i$, conditional on θ^t.) This contract provides perfect smoothing but limited insurance in period $t + 1$. Intuitively, the intermediary should be able to lower its costs by offering a contract that pays the second-order cost of reducing smoothing to get the first-order benefit of improving insurance.

To be more specific, suppose that the intermediary changes the contract by lowering $c_{t+1}(\theta^t, \theta_L)$ by ε_L, lowering $c_{t+1}(\theta^t, \theta_H)$ by ε_H, and raising

$c_t(\theta^t)$ by ε. It chooses $(\varepsilon_H, \varepsilon_L, \varepsilon)$ so that there exists some δ with the property that

$$
\begin{aligned}
\delta &= -\beta u(c_{t+1}(\theta^t, \theta_H) - \varepsilon_H) + \beta u(c_t(\theta^t, \theta_H) \\
&= -\beta u(c_{t+1}(\theta^t, \theta_L) - \varepsilon_L) + \beta u(c_t(\theta^t, \theta_L)) \\
&= -u(c_t(\theta^t)) + u(c_t(\theta^t) + \varepsilon).
\end{aligned}
$$

(In words, the utility loss in period $t + 1$ is the same across the two states and equals the utility gain in period t.) Because u is concave, such a contract must necessarily have the property that $\varepsilon_H > \varepsilon_L$. Hence, while it necessarily worsens smoothing, the alternative contract provides superior risk sharing in period t.

We can approximate the changes as

$$
\varepsilon \approx \frac{\delta}{u'(c_t(\theta^t))},
$$

$$
\varepsilon_L \approx \frac{\delta\beta^{-1}}{u'(c_{t+1}(\theta^t, \theta_L))},
$$

$$
\varepsilon_H \approx \frac{\delta\beta^{-1}}{u'(c_{t+1}(\theta^t, \theta_H))}.
$$

If we apply Jensen's inequality to the perfect smoothing condition, we know that

$$
\frac{\delta}{u'(c_t(\theta^t))} > \frac{\beta^{-1}R^{-1}\pi_H\delta}{u'(c_{t+1}(\theta^t, \theta_H))} + \frac{\beta^{-1}R^{-1}\pi_L\delta}{u'(c_{t+1}(\theta^t, \theta_L))},
$$

which implies that

$$
-\varepsilon + R^{-1}\pi_H\varepsilon_H + \beta^{-1}R^{-1}\pi_L\varepsilon_L > 0.
$$

The new contract lowers the intermediary's costs.

This argument demonstrates that the intermediary is able to provide better risk sharing to the agent in period $t + 1$ by front-loading consumption into period t. The superior contract does presume that the intermediary can monitor the agent's savings. The intermediary uses this monitoring ability to reduce the agent's consumption in period $t + 1$ in response to any extra asset holdings on the part of the agent. This assumption captures the idea that providers of insurance against skill shocks (such as governments and families) do have access to information about wealth and do use that information in insurance provision.

This idea has empirical counterparts. Two important sources of insurance against large idiosyncratic productivity shocks are the govern-

ment and informal social networks such as families and/or friends. In the United States, the government explicitly conditions social insurance on asset holdings. Thus, an unemployed household with large asset holdings pays more income taxes and so receives less net payments from the government than an unemployed household with low asset holdings. Even more directly, the government does not make welfare payments to households with sufficiently high asset holdings. For example, to be eligible for Supplemental Security Income, food stamp benefits, welfare benefits under the Temporary Assistance for Needy Families program, or Medicaid, applicants generally must meet both an income test and an explicit asset test. Specifically, the food stamp asset limit is $2,000 ($3,000 for households with an elderly or disabled member); for SSI, the limits are $2,000 for a single individual and $3,000 for a couple.

In terms of intrafamilial insurance, there is some evidence that parents' transfers to their heirs depend on the heirs' wealth. McGarry and Schoeni (1995) use microeconomic data from the Health and Retirement Survey (HRS) to show that "equal transfers to all children are the exception rather than the rule" (S204). More important for our purpose, they find in multivariate controlled regressions that "parents transfer less to wealthy children, and thus less to children with housing wealth than those without" (S207).[7] At a more introspective level, it seems intuitive that a family would be more willing to help out a freshly unemployed member if that member were homeless.

D.   Three Models and Their Implications for the Equity Premium

In this subsection, we use our earlier results to derive the key testable implications for the equity premium. We begin with some basics from asset pricing theory. Consider an arbitrary financial asset i that has a realized gross return $R_{t+1}^i(z^{t+1})$ in history z^{t+1}, given an investment of one unit of consumption in history z^t. This asset can be regarded as a bundle of state-contingent claims to consumption that has price 1 in history z^t. The absence of arbitrage opportunities guarantees that its return must satisfy

$$1 = \sum_{z_{t+1}} q_{t+1}(z_{t+1}|z^t) R_{t+1}^i(z^t, z_{t+1}).$$

---

[7] Light and McGarry (2004) use instead a set of direct questions regarding expectations of inter vivos transfers asked to all mothers in the HRS. Mothers who report that they intend to divide their estates unequally are asked to explain why. Of the mothers providing a response, slightly less than 20 percent "refer to their childrens' financial needs in explaining why some will receive a larger bequest than others (e.g., 'the oldest son has more assets than the youngest son')," whereas another 9 percent cite asset tests, e.g., "my daughter can't have over $2000 or she will lose her state benefits" (1675).

For the purposes of empirical implementation, it is typically convenient to rewrite this expression as

$$1 = E(\text{SDF}_{t+1} R_{t+1}^i | z^t),$$

where $E(\cdot | z^t)$ is the expectation conditional on z^t. Here, as above, we define

$$\text{SDF}_{t+1}(z^{t+1}) \equiv \frac{q_{t+1}(z_{t+1} | z^t)}{\Psi(z_{t+1} | z^t)}$$

to be a stochastic discount factor. Now consider two particular assets, stocks and Treasury bills. Let R_{t+1}^s be the realized return for stocks and R_{t+1}^b be the realized return for Treasury bills. The *realized excess stock return* is the gap between these realized returns. In equilibrium, it must be true that

$$0 = E[\text{SDF}_{t+1}(R_{t+1}^s - R_{t+1}^b)], \tag{14}$$

where we have used the law of iterated expectations to eliminate the conditioning on z^t.

Under appropriate regularity conditions, we can estimate the unconditional expectation in (14) using time-series averages. The *equity premium puzzle* says that it is difficult to find a model of the SDF that satisfies the sample analogue of (14). One way to see the problem is to write (14) as

$$0 = E(R_{t+1}^s - R_{t+1}^b) + \frac{1}{E(\text{SDF}_{t+1})} \text{Cov}[\text{SDF}_{t+1}, (R_{t+1}^s - R_{t+1}^b)]. \tag{15}$$

The expected difference between stock returns and Treasury bill returns, at least as estimated via time-series averages in the U.S. data, is on the order of 7–8 percent per year. To resolve the equity premium puzzle, we need to find an SDF that covaries sufficiently negatively with the realized equity premium so as to offset the large first term on the right-hand side of (15).

The previous subsection suggests two candidate SDFs.[8] The first is what we will term the incomplete markets (INC) SDF:

$$\text{SDF}_{t+1}^{\text{INC}}(z^{t+1}; \beta, \gamma) = \beta \frac{C_{-\gamma, t+1}(z^{t+1})}{C_{-\gamma, t}(z^t)}. \tag{16}$$

---

[8] To derive these SDFs, we assume in both types of market structures that asset traders trade a complete set of z'-contingent securities. This assumption is not essential. The SDFs in (16) and (17) satisfy (14) as long as the traders can exchange stocks and Treasury bills at each date.

We do not explore the hypothesis that the SDF in country j does a good job of explaining the equity premium in country i (appropriately adjusted for exchange rate movements). By not doing so, we are implicitly assuming that asset traders in country j face nontrivial costs of trading bonds and stocks in country i.

The second is what we will term the PIPO SDF:

$$\text{SDF}_{t+1}^{\text{PIPO}}(z^{t+1}; \ \beta, \ \gamma) \ = \ \beta \frac{C_{\gamma,t}(z^t)}{C_{\gamma,t+1}(z^{t+1})}. \tag{17}$$

Both of these SDFs are interesting only insofar as there is nontrivial heterogeneity in household consumption growth. In particular, suppose that Θ is a singleton. Then, both SDFs collapse to the usual representative agent (RA) SDF:

$$\text{SDF}_{t+1}^{\text{RA}}(z^{t+1}; \ \beta, \ \gamma) \ = \ \beta \frac{C_{1,t+1}(z^{t+1})^{-\gamma}}{C_{1,t}(z^t)^{-\gamma}}, \tag{18}$$

where $C_{1,t}$ represents the first moment of the cross-sectional consumption distribution.

It is well known that the RA SDF does not covary sufficiently with the realized equity premium to satisfy (14), unless γ is very large (see Kocherlakota 1996). The PIPO and INC SDFs have the potential to perform better than the RA SDF, because they allow consumption *inequality* to affect the SDF. For either model to predict a high equity premium, it must be true that the heterogeneity in consumption leads stocks to be less valuable relative to bonds than predicted by the representative agent model. In the incomplete markets model without borrowing constraints, an asset's value (compared to the representative agent model) is driven by agents' precautionary demands for that asset. There is more precautionary demand for assets that pay off exactly when there is a lot of downside individual-level consumption risk (the left tail of the consumption distribution is heavy). There is little precautionary demand when there is a lot of upside risk (the right tail is heavy). For this reason, the equity premium is higher if stocks pay off relatively little when $C_{-\gamma,t+1}$ is high.

The pricing of assets in the PIPO model is instead driven by the dependence of the insurer's marginal value of extra consumption on aggregate states. The key to this marginal valuation is that, because of the incentive problem, the insurer must allocate any extra consumption in a given state across households in a way that raises each household's *utility* by the same amount. For example, suppose that the aggregate state is such that the right tail of the consumption distribution is very heavy, so that there is a small number of extremely rich agents. Then, concavity plus the equal-utility-increase requirement means that the insurer must spend most of the extra resources on those rich agents. The average utility across all households will not go up by very much. It follows that the insurer does not value assets that pay off when the right tail of the consumption distribution is very heavy, and so the equity premium is high if stocks pay off more when $C_{\gamma,t+1}$ is high.

It is important to distinguish the INC SDF from the one employed by Brav et al. (2002) and Cogley (2002). Those papers make the same assumptions about market structure (incomplete markets with non-binding borrowing constraints) and derive the following SDF:

$$\beta E[c_{t+1}^{\mathrm{INC}}(\theta^{t+1},\ z^{t+1})^{-\gamma} c_t^{\mathrm{INC}}(\theta^t,\ z^t)^{\gamma}|z^{t+1}],$$

which is the average of the agents' IMRSs. Like the incomplete markets SDF, this average IMRS discount factor is also a valid SDF in an incomplete markets equilibrium with nonbinding borrowing constraints. We use our incomplete markets SDF instead of the average IMRS SDF because to estimate the latter, we would need observations of consumption over time for a given household. This panel structure is unavailable in the United Kingdom or Italy.

E.   Measurement Error in Consumption

Our empirical strategy will be to estimate the relevant cross-sectional moments in the PIPO and INC SDFs using cross-sectional household data on consumption. One of the difficulties with using cross-sectional data on consumption is that the data are typically measured with error. This measurement error usually creates difficulties when one applies GMM to estimate Euler equations of the form

$$\beta E_t\!\left[\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} R_{t+1}\right] = 1.$$

Measurement error in the level of consumption can bias the level of measured household consumption growth upward or downward and so can contaminate the estimates of β and γ in unknown ways. In this subsection, we show that the INC and PIPO SDFs are robust to a wide class of measurement error processes.

In particular, let $c*$ be the true allocation of consumption. We allow $c*$ to be measured with error as follows. Let $(\nu_1, \nu_2, \ldots, \nu_T)$ be a collection of random variables with joint probability measure μ_ν over the Borel sets in R^T. At the beginning of period 1, after the public shock sequence z^T is drawn, a realization ν^T is drawn according to μ_ν for each agent; conditional on z^T, the draws of ν^T and θ^T are independent from each other and are independent across agents. Note too that ν^T is independent of z^T (because it is drawn from μ_ν for all z^T); however, the measurement error is allowed to have arbitrary serial correlation.

Define $\hat{c}_t(\theta^t,\ z^t,\ \nu_t) = \exp{(\nu_t)}c_t*(\theta^t,\ z^t)$ to be measured consumption. Define also

$$\hat{C}_{\eta,t} = E\{\hat{c}_t^\eta|z^t\} \tag{19}$$

to be the ηth moment of cross-sectional measured consumption, in public history z^t. From the definition of measured consumption, we know that

$$
\begin{aligned}
\hat{C}_{\eta,t} &= E[c_t^{*\eta}\exp(\eta\nu_t)|z^t] \\
&= E[\exp(\eta\nu_t)|z^t]E(c_t^{*\eta}|z^t) \\
&= E[\exp(\eta\nu_t)]C_{\eta,t},
\end{aligned}
\tag{20}
$$

where the penultimate equation comes from the independence of ν_t from θ^t, conditional on z^t.

Now suppose that $E[\exp(\eta\nu_t)] < \infty$ and ν_t is stationary. These assumptions imply that

$$
\frac{\hat{C}_{\eta,t}(z^t)}{\hat{C}_{\eta,t+1}(z^{t+1})} = \frac{C_{\eta,t}(z^t)}{C_{\eta,t+1}(z^{t+1})}
\tag{21}
$$

for all t, z^{t+1}. Thus, under these assumptions about ν, there is no error associated with computing the three SDFs with measured consumption as opposed to true consumption. Note that we have assumed that ν is independent across agents, independent from agents' true types, and stationary over time. These assumptions about the nature of the measurement error are not wholly innocuous. However, we do not have to make any assumptions at all about the magnitude of the measurement error, beyond assuming the finiteness of a particular moment, or impose any particular restrictions on its autocorrelation structure.[9]

V. Empirical Analysis: Preliminaries

In this section, we describe our data and empirical methodology.

A. Micro Data

1. The U.S. CEX

The micro data for the United States come from the 1980–2004 Consumer Expenditure Survey. The data are collected by the Bureau of Labor Statistics (BLS) and used primarily for revising the consumer price index (CPI). Consumer units are defined as members of a house-

---

[9] There is no evidence from validation consumption studies that can tell us whether the assumption we make about the nature of the measurement error is truly restrictive. Evidence from validation wage and income studies (Bound and Krueger 1991) has found that measurement error (a) appears serially correlated, (b) is independent of schooling, and (c) is negatively correlated with the true measure of wage/income. The latter finding will, of course, invalidate our empirical strategy.

hold related by blood, marriage, adoption, or other legal arrangement. The definition of the head of the household in the CEX is the person or one of the persons who owns or rents the unit.

The CEX is based on two components, the diary, or record-keeping, survey and the interview survey. The diary sample interviews households for two consecutive weeks, and it is designed to obtain detailed expenditures data on small and frequently purchased items. The interview sample is in the form of a rotating panel, and it follows survey households for a maximum of 5 quarters. The database covers about 95 percent of all expenditures. Following most previous research, our analysis below uses only the interview sample.

The CEX collects information on a variety of sociodemographic variables. Expenditure is reported in each interview (after the first) and refers to the months of the previous quarter. Income is reported in the second and fifth interviews, and it refers to the previous 12 months. Holdings of financial assets are reported only in the fifth interview. We refer the reader to the online Appendix for further details on sample selection and consumption definitions.

We "deflate" consumption data to account for three phenomena: price differences over time, seasonal differences (i.e., month effects) within a year, and households' demographic differences at a certain point in time. Thus, nondurable consumption is first expressed in real terms using the CPI (all items; in 2000 dollars). Then, data are deseasonalized by simple multiplicative regression adjustments. Finally, we convert it into adult-equivalent consumption data.[10]

Recently, researchers have noted that for many commodities, the aggregation of CEX data matches poorly National Income and Product Accounts (NIPA) Personal Consumption Expenditure (PCE) data. Some of the discrepancy is undoubtedly due to differences in covered population and definitional issues. But the amount of underestimation of consumer expenditure is sometimes substantial, and it raises some important warning flags. Furthermore, there is evidence that the detachment between the CEX aggregate and the NIPA PCE has increased over time (see Attanasio, Battistin, and Ichimura 2004). At present, it is not clear why this is so and whether this is necessarily due to a worsening in the quality of the CEX. For example, Bosworth, Burtless, and Sabelhaus (1991) conclude that most of the discrepancy is explained by the failure of the CEX to sample the super-rich; others have suggested a greater incidence of attrition. According to the BLS, however, the CEX

---

[10] The number of adult equivalents is defined as $(A + \alpha K)^{\beta}$, where A is the number of adults (aged 18 or more), K is the number of kids, and α and β are parameters. We set $\alpha = 0.7$ and $\beta = 0.65$ (following recommendations contained in Citro and Michael [1995], which in turn draws from Betson [1990]). Similar results are obtained if we use a more sophisticated Engel approach.

has maintained representativeness of the U.S. population over time, and attrition has not changed much since the redesign of the survey of the early 1980s.[11]

2. The U.K. FES

The micro data for the United Kingdom come from the 1975–99 Family Expenditure Survey. The FES is conducted by the Office for National Statistics and is representative of the U.K. population. The FES is primarily a survey of household expenditures on goods and services and household income. Similarly to the CEX, the main goal of the survey was originally that of providing information on spending patterns for revising the CPI. However, with time the survey has become multipurpose, providing information on household economic and social variables.

The FES has been in operation since 1957, and up to and including 1993, data are available by calendar year. Since 1994, data are available by financial year (April–March). The basic unit of the survey is the household, defined as a group of people living at the same address with common housekeeping (sharing household expenses and/or a living room). Each individual in the household aged 16 or more is asked to keep diary records of daily expenditures for 2 weeks. Information about regular expenditures, such as rent and mortgage payments, is obtained from a household interview along with retrospective information on certain large, infrequent expenditures such as those on vehicles. Data are collected throughout the year to cover seasonal variations in expenditures. See the online Appendix for details on sample selection and consumption definitions.

Similarly to the CEX, the FES data we construct are "deflated" to account for price differences over time, seasonal differences (i.e., month effects) within a year, and households' demographic differences at a certain point in time. Thus, nondurable consumption is first expressed in real terms using the CPI (all items) described above (on the same basis as the U.S. CPI). Then, data are deseasonalized by simple multiplicative regression adjustments. Finally, we convert it into adult-equivalent consumption data. Given that the data we have are average weekly expenditures, we multiply FES expenditures by 13 to form quarterly consumption data for each household in the sample.

---

[11] Given the differences between the CEX and NIPA, we reestimated the model using the latter and found similar results for the representative agent case (results available on request).

3.   The Italian SFB

Micro data for Italy are drawn from the 1985–2002 Survey on Family Budgets, which is conducted every year by ISTAT (the Central Statistics Office). The SFB follows the standard international procedure of exploiting both information from recall questions for more durable items bought in the quarter prior to the interview and diary-based records of purchases carried out within a 7-day period. The survey also includes demographic information about all the members of the household. Income and savings are poorly measured and, since 1997, are available only in broad interval classes.

Micro data are available to researchers from 1985 onward. Since 1997 the SFB is conducted using sampling strategies recommended by Eurostat harmonization directives. In particular, besides the sampling design, the post-1996 SFB has different questionnaires and different procedures for revising and correcting the data. For these reasons, there is a structural break in the time series of consumption data from the SFB. We account for this break in the following way. For the PIPO case, our estimates are obtained dropping the observations on $SDF^{PIPO}_{1997:1}$, $SDF^{PIPO}_{1997:2}$, and $SDF^{PIPO}_{1997:3}$. We follow similar logic for obtaining estimates under the other two models.

In what follows, we describe the post-1996 (stratified) sampling design. The unit of observation is the household, which includes all individuals sharing a dwelling and related by blood, marriage, adoption, or foster care. The survey is a repeated cross section with about 32,000 households sampled throughout the year (or about 2,700 every month), drawn randomly from the civil registry of households. The response rate is around 84 percent, which compares well with the other two surveys we use. Each household is interviewed only once and then leaves the survey. See the online Appendix for details on sample selection and consumption definitions.

As with the CEX and the FES, we preliminarily "deflate" SFB data to account for price differences over time, seasonal differences (i.e., month effects) within a year, and households' demographic differences at a certain point in time. Thus, nondurable consumption is first expressed in real terms using the CPI (all items) described above (on the same 2000 basis as the CPIs for the other two countries). Then, data are deseasonalized by simple multiplicative regression adjustments. Finally, we convert it into adult-equivalent consumption data. Given that data are available to us as monthly averages, we multiply SFB expenditure by 3 to form quarterly consumption data for each household in the sample.

B.   Asset Returns Data

Our empirical analysis requires obtaining data on the stock market return (R^s) as well as the risk-free rate of return (R^b). We use returns data drawn from the Center for Research in Security Prices (CRSP) at the University of Chicago for the United States and the Global Financial Data Web site (http://www.globalfinancialdata.com/) for Italy and the United Kingdom.

The risk-free rate R^b is obtained in the following way. First, we extract the 1-month nominal returns on Treasury bills.[12] Then, we convert it in real terms dividing it by $1 + \pi$, where π is the monthly inflation rate obtained from the CPI (in 2000 currency), also used below. Finally, we obtain the quarterly return by compounding the monthly returns.

For the United States the market return R^s is the return on the CRSP value-weighted portfolio. It includes dividends and capital gains. It is based on the average 1-month nominal return of the pooled sample of stocks listed on the New York Stock Exchange and the American Stock Exchange. For Italy we use the BCI Global Return Index, whereas for the United Kingdom we use the FTSE All-Share Return Index.[13] These indexes include the changes in the price of the stock and the dividends that are paid to investors and then reinvested. We convert the nominal returns into real terms by dividing it by $1 + \pi$. Finally, we obtain the quarterly return by compounding the monthly returns. The difference $R_t^s - R_t^b$ is the equity premium.

C.   Methodology

Our estimation methodology is motivated by that originally described by Hansen and Singleton (1982). We want to investigate the equity premium puzzle of Mehra and Prescott (1985). They point out that, historically, the gap between average stock returns and average Treasury bill returns is very large (on the order of 6 percent per year) and difficult to rationalize using standard representative agent asset pricing models. As in Kocherlakota (1996), we assess the candidate SDFs' ability to rationalize the equity premium by considering the restriction that

$$E[\text{SDF}_t(R_t^s - R_t^b)] = 0, \qquad (22)$$

where as before SDF$_t$ is the stochastic discount factor, R_t^s is the return

[12] For Italy the source is the Central Bank of Italy. For the United Kingdom, it is the Central Statistical Office, Annual Abstract of Statistics.

[13] For Italy the source is Banca Commerciale Italiana; for the United Kingdom, it is the Central Statistical Office, Annual Abstract of Statistics (before 1989), and Eurostat (1989 onward).

to the stock market, and R_t^b is the return to the 90-day Treasury bill (all returns are real).

The micro data sets that we are using provide data of the form $\{\{c_{it}\}_{i=1}^{N_t}\}_{t=-2}^T$, where c_{it} is the consumption expenditure of household i for the quarter ending with month t (i.e., covering months $t-2$, $t-1$, and t). We define sample analogues of the various SDFs using these *cross-sectional* data. In particular, let

$$\widehat{\mathrm{SDF}}_t^{\mathrm{PIPO}}(\gamma) = \frac{N_{t-3}^{-1}\sum_{i=1}^{N_{t-3}} c_{it-3}^{\gamma}}{N_t^{-1}\sum_{i=1}^{N_t} c_{it}^{\gamma}}, \tag{23}$$

$$\widehat{\mathrm{SDF}}_t^{\mathrm{INC}}(\gamma) = \frac{N_t^{-1}\sum_{i=1}^{N_t} c_{it}^{-\gamma}}{N_{t-3}^{-1}\sum_{i=1}^{N_{t-3}} c_{it-3}^{-\gamma}}, \tag{24}$$

and

$$\widehat{\mathrm{SDF}}_t^{\mathrm{RA}}(\gamma) = \frac{(N_t^{-1}\sum_{i=1}^{N_t} c_{it})^{-\gamma}}{(N_{t-3}^{-1}\sum_{i=1}^{N_{t-3}} c_{it-3})^{-\gamma}} \tag{25}$$

denote the sample analogues of the PIPO, IM, and RA stochastic discount factors in period t (with $t = 1, \dots, T$). To reiterate, we use overlapping data, so t here indexes the last month of a given quarter. Thus, for example, in the case of the U.S. CEX, the first available observation for $\widehat{\mathrm{SDF}}_t^{\mathrm{PIPO}}(\gamma)$ is for 1980:3, and it is constructed as the ratio of the γth moment of consumption for 1979:12 (calculated using all households reporting expenditure data for October–December 1979) and the γth moment of 1980:3 (calculated using all households reporting expenditure data for January–March 1980). The last observation is for 2004:2. We have overall $T = 288$ observations on $\widehat{\mathrm{SDF}}_t^j(\gamma)$ ($j = $ PIPO, INC, RA) for the United States, $T = 297$ for the United Kingdom, and $T = 210$ for Italy (after dropping the first three months of 1997 because of the change in the sampling design described above).

We then form sample analogues (in the time-series dimension) of the restriction (22). For example, the sample analogue of the left-hand side of (22) in the PIPO model is

$$T^{-1}\sum_{t=1}^{T} \frac{N_{t-3}^{-1}\sum_{i=1}^{N_{t-3}} c_{it-3}^{\gamma}}{N_t^{-1}\sum_{i=1}^{N_t} c_{it}^{\gamma}} (R_t^s - R_t^b). \tag{26}$$

We estimate the unknown parameter γ and evaluate the three discount factors by applying GMM to this moment condition. We estimate the value of γ that satisfies (22) separately for each country (results reported in table 2 below) and then pool all the country data together (table 3 below).

TABLE 1
DESCRIPTIVE STATISTICS

| | A. Macro Data | | | |
	$\mu(r^b)$	$\sigma(r^b)$	$\mu(r^s)$	$\sigma(r^s)$
United States	.55	.68	2.52	7.59
United Kingdom	.64	1.46	3.80	10.85
Italy	1.22	.68	2.68	13.54
	B. Micro Data			
	$\ln(C_{-2,t}/C_{-2,t-3})$	$\ln(C_{1,t}/C_{1,t-3})$	$\ln(C_{2,t}/C_{2,t-3})$	$\ln(C_{5,t}/C_{5,t-3})$
United States	.0039	−.0011	−.0009	.0212
United Kingdom	−.0043	.0039	.0089	.0340
Italy	−.0098	.0043	.0087	.0278

NOTE.—The variables r^b and r^s are the returns on stocks and Treasury bills, respectively (in percentage terms), and $C_{\eta,t}$ is the ηth central moment of the quarterly consumption distribution in period t. For the United States, the sample period for which we have micro data is 1980–2004; for the United Kingdom, it is 1975–99; and for Italy, it is 1985–2002.

Finally, it is worth discussing the nonstandard inference problem we face. Note that time-series moment conditions like (26) are functions of cross-sectional nonlinear moments of the data. If these cross-sectional moments were known, then we could simply plug in these known moments into the time-series moment conditions and then apply the usual time-series GMM formulas to calculate standard errors (perhaps after accounting for serial correlation induced by the use of overlapping data; see Hansen and Hodrick 1980). In reality, we do not know the true cross-sectional moments. In what follows, we will assume that the cross-sectional sample size increases at a rate that is sufficiently fast as to make this source of uncertainty inconsequential and treat sample estimates as if they were the true cross-sectional moments. This is the same approach followed by Brav et al. (2002).[14] Below, our inferences abstract from errors in the cross-sectional distribution.

VI. Empirical Analysis: Results

We provide some simple summary statistics in table 1. The table shows that there is a large equity premium in all countries over the sample periods for which we have micro data. The mean return to stocks is about 7.9 percent per year higher than the mean return to Treasury bills in the United States. This sample estimate is higher than the 6.2 percent annual number averaged in the 90 years of data (1889–1978)

[14] An alternative would be to assume that the cross-sectional sample size increases as fast as (or even less than) the time-series size. In this case, one could think of computing standard errors by the bootstrap. However, given the complications involved (for instance, the fact that when T goes to infinity, so does the number of parameters to estimate), we have decided to leave this as a topic for future research.
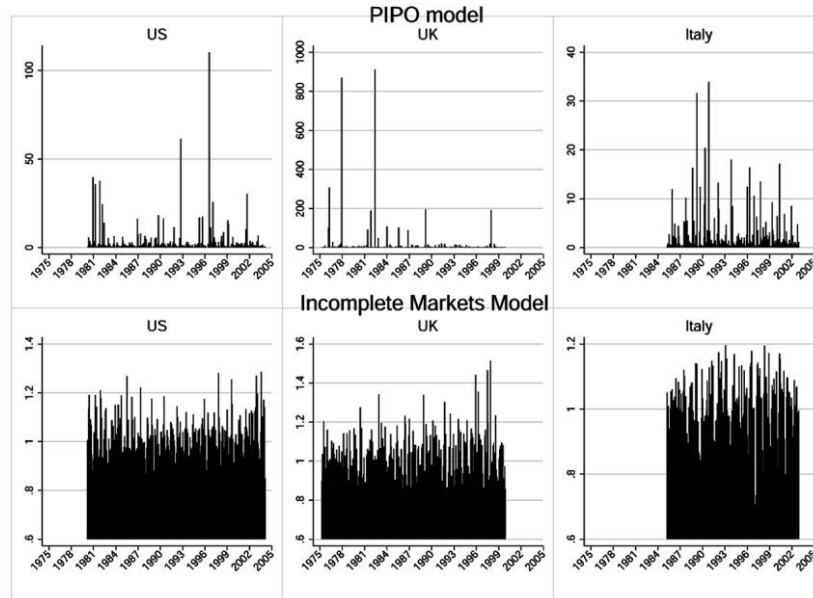
FIG. 1.—The PIPO SDF (for $\gamma = 5$) and the incomplete markets SDF (for $\gamma = 2$)

studied by Mehra and Prescott. The standard deviation of stock returns is about 30.4 percent per year. In the United Kingdom, the premium is even higher (12.6 percent per year), and it is slightly lower in Italy (5.8 percent per year).[15] We also report the growth of various moments of the cross-sectional distribution of quarterly consumption in each country. The data show higher consumption growth in the United Kingdom and Italy than in the United States. Volatility of consumption also grows more rapidly in those two countries. In all countries, high consumption moments (such as the fifth noncentral moment) grow rapidly.

In figure 1 we plot the PIPO stochastic discount factor for $\gamma = 5$ and the incomplete markets SDF for $\gamma = 2$ (as we will see, these are the values of γ that minimize the equity premium pricing error in the two models). The units in these graphs are quarterly (overlapping) as given by (23) and (24). For the chosen values of γ, the SDFs are highly variable. Of course, a valid SDF has to be more than variable: it must covary negatively with stock returns.

[15] Campbell (1999) reports a negative equity premium for Italy. Our results differ from his for two reasons. First, he uses the 3-month interbank rate, whereas we use the return on the 90-day Treasury bill. Second, and more important, Campbell focuses on a different period than we do (1971:2–1995:3 vs. 1985:1–2002:4).

TABLE 2
GMM Estimates of γ

	United States			United Kingdom			Italy		
	PIPO	INC	RA	PIPO	INC	RA	PIPO	INC	RA
γ	5.3326	1.5246	53.2607	4.9231	2.1969	4.8564	4.5357	2.3236	10.5233
	(1.3270)	(NA)	(21.1679)	(1.2451)	(NA)	(NA)	(3.1786)	(NA)	(NA)
$\bar{e}$	.0000	.0190	.0000	.0000	.0245	.0250	.0000	.0086	.0077

Note.—In this table, we report the estimates and standard errors (in parentheses) associated with estimating γ using the restriction that $e_t(\gamma)$ has expectation zero, where $e_t(\gamma) = \text{SDF}_t(\gamma)(R_t^s - R_t^b)$. The row $\bar{e}$ reports the sample mean of the pricing error at the estimated value of γ.

A.   The Equity Premium: Results

We now look at the ability of the various discount factors to rationalize the large equity premium in the data. Define the sample mean of the equity premium errors to be

$$\bar{e}^j(\gamma) = \frac{1}{T}\sum_{t=1}^{T} \widehat{\text{SDF}}_t^j(\gamma)(R_t^s - R_t^b) \tag{27}$$

for $j = $ PIPO, INC, and RA. Equation (27) is the empirical analogue of (22), and $\widehat{\text{SDF}}^j(\gamma)$ is defined by equations (23)–(25).

In table 2 we estimate the coefficient of relative risk aversion γ that minimizes (27) by applying GMM to the equity premium pricing error separately for the three countries. Note that since we have an exactly identified model, the choice of the weighting matrix is unimportant. The standard errors are corrected for serial correlation induced by the use of overlapping data as suggested by Hansen and Hodrick (1980).

There are two striking findings in table 2. First, the PIPO model is the only one that does consistently well across countries. Second, an estimate of the coefficient of relative risk aversion of about 5 would zero out the equity premium error in the PIPO case in all three countries. These findings can be compared with the ones we obtain in the incomplete markets case. The estimate of γ is lower than those found for the other two SDFs. However, this is misleading: the incomplete markets SDF can explain virtually none of the observed equity premium.[16] Another point to make is that the usual GMM estimate of the standard error is undefined because the (one) moment condition we use is not set equal to zero in the sample. As for the representative agent model, we find that in the U.S. case we can eliminate the equity premium pricing error with an estimate of γ equal to approximately 53. (This high estimate of γ, of course, is reminiscent of those obtained in prior work

[16] Brav et al. (2002) restrict attention to households with nonnegative financial wealth. When we use this smaller sample, in conjunction with the incomplete markets and complete markets SDFs, the point estimates are similar to what we obtain in tables 3–5.

TABLE 3
JOINT GMM ESTIMATE OF γ

	PIPO	INC	RA
γ	4.9216	2.0491	7.5931
	(1.2211)	(12.2032)	(50.4559)
$\bar{e}_{\text{US}}$	.0088	.0191	.0182
$\bar{e}_{\text{UK}}$	.0002	.0246	.0252
$\bar{e}_{\text{IT}}$	−.0068	.0087	.0079
χ^2 statistic, overidentifying test	.1883	12.4015	8.1308
p-value	91%	0%	2%
χ^2 statistic equal γ	.0599	.0001	.0013
p-value	97%	99%	99%

NOTE.—In this table, we report the estimates and standard errors (in parentheses) associated with estimating γ using the restriction that $\varepsilon_t(\gamma)$ has expectation zero, where $\varepsilon_t(\gamma) = (e_{\text{US},t}(\gamma)\ e_{\text{UK},t}(\gamma)\ e_{\text{IT},t}(\gamma))'$ and $e_t(\gamma) = \text{SDF}_t(\gamma)(R_t^s - R_t^b)$. The rows $\bar{e}_j$ report the sample means of the pricing error at the estimated value of γ in country j. The χ^2 rows report, respectively, the value of the χ^2 statistic for the overidentifying restriction test (with two degrees of freedom), and the value of the χ^2-statistic for the test of equal γ across countries (with two degrees of freedom).

on the equity premium; see Kocherlakota 1996.) For the other two countries there is no estimate of γ that can zero out the equity premium pricing error.[17]

In table 3 we impose that γ is the same across countries and consider the set of orthogonality conditions

$$
E\begin{pmatrix} e_{\text{US},t}^j(\gamma) \\ e_{\text{UK},t}^j(\gamma) \\ e_{\text{IT},t}^j(\gamma) \end{pmatrix} = E(\varepsilon_t(\gamma)) = 0,
$$

where $e_{k,t}^j(\gamma) = \text{SDF}_{k,t}(\gamma, z^t)[R_{k,t}^s(z^t) - R_{k,t}^b(z^t)]$ is the pricing error for model j ($j = $ PIPO, INC, and RA) in country k ($k = $ US, UK, IT) and period t. We assume $E(e_{k,t}^j(\gamma)\ e_{r,s}^j(\gamma)) = 0$ for $k \neq r$ and all s, t. This means that the variance-covariance matrix of $\varepsilon_t(\gamma)$ is block diagonal, where each block is given by the variance-covariance matrix of the pricing error in the corresponding country. One of the advantages of imposing a common γ across countries is that we have overidentifying restrictions that we can use to test our model. We also test whether the assumption of no preference heterogeneity is valid in each of the three models. Given the well-known small-sample problems of optimal GMM, we use the identity matrix as our weighting matrix.[18]

Table 3 shows that for all models we cannot reject the hypothesis of preference homogeneity. However, while the incomplete markets model and the representative agent model are rejected according to the overidentifying restrictions test, the PIPO model is not (p-value of 91 percent) for a value of γ of around 5. This confirms that the PIPO model provides the best fit for the equity premium data.

[17] For the United States, we obtain qualitatively similar evidence if we use a market return that includes stocks listed on NASDAQ.

[18] We use the formula for the overidentifying test statistic reported in Cochrane (2001).

Our results for the incomplete markets SDF contrast with the results of Brav et al. (2002) and Semenov (2004) for the average IMRS SDF. They find that the sample equity premium is eliminated when γ is set to a relatively low value (less than 4). Measurement error cannot be the source of the discrepancy; recall that the average IMRS SDF is valid for return differentials like the equity premium under the same class of measurement error processes that we assume in this paper. However, there are two main differences between what we do and what Brav et al. do. First, as we stressed earlier, the incomplete markets SDF and the average IMRS SDF are distinct SDFs. The validity of the latter does not imply the validity of the former, although both should be valid in an incomplete markets equilibrium with no binding borrowing constraints. Second, Brav et al.'s sample selection is different from ours: They keep only households that stay in the sample for 3 or more quarters (because they use the average IMRS SDF) and, to minimize measurement error, discard households that report extremely large increases or decreases in consumption from one quarter to another. Their sample selections end up discarding about 60 percent of the households in the CEX. Finally, they use the sample period 1982:I–1996:I rather than the sample period 1980:I–2004:I.[19]

B.   Limited Participation

As pointed out by Mankiw and Zeldes (1991), the incomplete markets SDF may perform better if it is corrected for limitations on participation in financial markets. Attanasio, Banks, and Tanner (2002) and Vissing-Jorgensen (2002) use data from the United Kingdom and United States to investigate this hypothesis. They show that limited participation does improve the performance of incomplete markets models to account for the level of returns. However, they do not examine the interaction between limited participation and asset return *differentials* like the equity premium. Can accounting for limited participation make the incomplete markets SDF consistent with the equity premium?

To address this issue directly, we repeated our estimation procedure for the incomplete markets case considering only financial market participants. Since the FES and SFB do not include information about financial market participation, this experiment is run only on the CEX. Households in the CEX are asked, in their fifth and final interview, to report information on their current holdings of assets. Separate questions are asked about "savings accounts in banks, savings and loans,

---

[19] We constructed a subsample of the CEX using the selection criteria reported in their paper and were able to replicate most of their results. In this sample, the sample equity premium is eliminated using the PIPO discount factor when we set γ between 9 and 10.

TABLE 4
The Unexplained Equity Premium for Financial Market Participants, CEX

	Stockholders	Stockholders and Bondholders					
		$A > 0$	$A > 1,000$	$A > 2,000$	$A > 5,000$	$A > 10,000$	$A > 25,000$
γ	1.4630	1.0479	.7870	1.4571	.9218	1.0911	.9885
$\bar{e}$	.0163	.0164	.0165	.0158	.0167	.0165	.0169

Note.—In this table, we report the estimates associated with estimating γ using the restriction that $e_t(\gamma)$ has expectation zero, where $e_t(\gamma) = \text{SDF}_t(\gamma)(R_t^s - R_t^b)$. In all columns, standard errors are not defined. The row $\bar{e}$ reports the sample mean of the pricing error at the estimated value of γ. A denotes the amount held in stocks.

credit unions, and similar accounts" (savings accounts), "checking accounts, brokerage accounts, and other similar accounts" (checking accounts), "U.S. savings bonds" (bonds), and "stocks, mutual funds, private bonds, government bonds, or Treasury notes" (stocks). The survey also asks respondents to report if current holdings are the same as, more than, or less than those held 12 months earlier and, in case they differ, to report the difference. This means that we observe asset holdings at two dates.

The SDF between period $t-3$ and period t is valid only for financial market participants at the beginning of period $t-3$. To construct this SDF, one requires data on consumption at two dates ($t-3$ and t) and data on asset holdings at the beginning of $t-3$. In the CEX, data on consumption are available at each interview, but data on asset holdings are available only at the beginning of the fifth and second interviews. This means that we can construct the SDF for financial market participants only in one case, namely, for data collected in the second ($t-3$) and third (t) interviews.[20]

We adopt two alternative classifications of financial market participation at the beginning of period t. The first follows Attanasio et al. (2002) and assumes that it coincides with stockholding. The second follows Vissing-Jorgensen (2002) and equates it with holding both stocks and bonds. (This is the most appropriate case for the equity premium equation; the equity premium first-order condition is necessarily satisfied only for agents who participate in both stock markets and bond markets.) The definition of stockholder and bondholder is as in Vissing-Jorgensen (2002).[21] The results, reported in table 4, show that account-

[20] Thus, in our experiment the numerator of the SDF uses only consumption data collected in the third interview, whereas the denominator uses only data collected in the second interview. Both refer to financial market participants at the beginning of the second interview as reported during their fifth interview.

[21] A household is classified as a stockholder (bondholder) at the beginning of period t if it (a) reports a positive amount held in stocks (bonds) in the fifth interview and no change from 12 months earlier, (b) reports holding a lower amount than 12 months earlier, (c) reports an increase in the amount of stocks (bonds) held by less than the current amount, (d) reports an increase in in the amount of stocks (bonds) held but the current amount is missing, or (e) reports a positive difference between the current amount held in stocks (bonds) and the difference in holdings over the last 12 months.

TABLE 5
GMM ESTIMATES OF β AND γ USING THE EXCESS RETURN AND THE RISK-FREE RATE
LEVEL RETURN

	UNITED STATES			UNITED KINGDOM			ITALY		
	PIPO	INC	RA	PIPO	INC	RA	PIPO	INC	RA
γ	5.3326	1.5246	53.2607	4.9231	2.1969	4.8564	4.5357	2.3235	10.5232
	(1.3271)	(NA)	(21.1679)	(1.2451)	(NA)	(NA)	(3.1786)	(NA)	(NA)
β	.2319	.9905	.3950	.0850	.9858	.9844	.5282	.9855	.9399
	(.2778)	(NA)	(.2502)	(.1545)	(NA)	(NA)	(1.0199)	(NA)	(NA)
$\bar{e}_1$	.0000	.0190	.0000	.0000	.0245	.0250	.0000	.0086	.0077
$\bar{e}_2$	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

NOTE.—In this table, we report the estimates and standard errors (in parentheses) associated with estimating β and γ using the restriction that $e_t(\beta, \gamma)$ has expectation zero, where

$$e_t(\beta, \gamma) = \begin{pmatrix} \text{SDF}_t(1, \gamma)(R_t^s - R_t^b) \\ \text{SDF}_t(\beta, \gamma)R_t^b - 1 \end{pmatrix}.$$

The rows $\bar{e}_1$ and $\bar{e}_2$ report the sample means of the corresponding pricing errors at the estimated values of β and γ.

ing for limited financial market participation in this fashion does not substantively change our results for the incomplete markets SDF. Given that only some households are marginal in financial markets, we also changed the definition of participants and include in the group of participants only those with stockholding above a threshold. The results, also reported in table 4, are qualitatively similar (i.e., we find that the incomplete markets SDF cannot explain the observed equity premium) for increasing thresholds equal to $1,000, $2,000, $5,000, $10,000, and $25,000.

C. Risk-Free Rate

Above, we have evaluated the ability of the various SDFs to account for the difference in returns between stocks and bills in three different countries. The SDFs should also be able to account for the *levels* of returns. In particular, SDF_t should be consistent with the two population restrictions:

$$E[\text{SDF}_t(1, \gamma)(R_t^s - R_t^b)] = 0 \tag{28}$$

and

$$E[\text{SDF}_t(\beta, \gamma)R_t^b - 1] = 0. \tag{29}$$

The advantage of adding restriction (29) is that one can identify the intertemporal discount factor β, which is not identified from using (28) alone. We estimate the values of β and γ that set (28) and (29) as close as possible to zero, using each candidate SDF, separately for each country. The results of these estimations are reported in table 5. We find that in all models it is possible to find values of (β, γ) that zero out the

pricing error of the risk-free level return equation. However, the only model that consistently zeros out both pricing errors for plausible levels of γ is the PIPO model.

This additional explanatory power does not come without a cost, though. The estimate of β that we obtain in the PIPO model is lower than 0.5 (at a quarterly frequency) in all countries, and it is statistically insignificant. Such low estimates strike us as implausible.[22] It is intriguing to note that, when we take into account sampling error, the estimates of β and γ are highly similar across the three countries.

The models estimated above can be compared statistically only through the value of the pricing errors. To obtain a summary goodness of fit measure of the various specifications, we add the following overidentifying restriction to equations (28) and (29):

$$E[(\text{SDF}_t(\beta, \gamma)R_t^b - 1)R_{t-3}^b] = 0. \tag{30}$$

Here, it is important to note that the Treasury bill return is highly autocorrelated, so that (29) and (30) are both informative. We can then estimate the parameters β and γ using GMM and have an additional overidentifying restriction.

Because of the finite sample difficulties documented by Kocherlakota (1990) and others, we are unwilling to use the asymptotically optimal two-step procedure originally used by Hansen and Singleton (1982). Instead, we use the one-step GMM procedure described by Hansen and Jagannathan (1997). They suggest using the weighting matrix given by the inverse of the second moments of the payoffs to be priced. In our setting, that translates into $[E(xx')]^{-1}$, where x is a column vector given by $[(R_t^s - R_t^b)|R_t^b|R_t^b R_{t-3}^b]$. We choose this weighting matrix for two reasons. First, the weighting matrix is the same for all models, which allows for ready comparison across models, even nonnested ones as in our case. Second, Hansen and Jagannathan show that with this weighting matrix, the minimized GMM objective can be interpreted as a measure of distance between the candidate SDF and the set of true SDFs (the Hansen-Jagannathan distance).

The results for the United States and the United Kingdom, not reported for brevity but available in the online Appendix, largely confirm the results of table 5.[23] In a statistical sense, the PIPO stochastic discount factor does a better job than the other two SDFs in fitting the data. First, the sample means of the pricing errors associated with the esti-

---

[22] Within the context of a labor search model with firm and worker heterogeneity, Postel-Vinay and Robin (2002) estimate that workers have low discount factors (ranging between 0.12 and 0.55 annually). In their work on asset pricing using limited enforcement models, Alvarez and Jermann (2001) find that they need to set discount factors to around 0.8 annually to match the mean risk-free rate.

[23] The results for Italy are less clear-cut. See the online Appendix for a discussion.

mated PIPO discount factor are an order of magnitude smaller than those associated with the other two discount factors. Second, the over-identifying restriction is rejected for all models but PIPO. Finally, the estimated Hansen-Jagannathan distance is lower for the PIPO SDF than for the other two SDFs. However, from an economic point of view, the PIPO SDFs give less convincing evidence. The estimated value of γ is around 5, but the estimated value of β is again implausibly low (less than 0.3) in the United States and United Kingdom.

D. Power Exercise

One possible concern about our findings is that our tests of the null hypothesis of the PIPO model lack power against alternative hypotheses. One way to check the power of our tests is to consider artificially generated data in which the PIPO SDF should not, a priori, be able to explain the observed equity premium.

The way we implement this power exercise is as follows. At each date t, there is a cross-sectional sample of N_t households. For each date t, we draw a sample (with replacement) of size N_t of households. For given γ, this allows us to compute an estimate of the γth moment of household consumption for that date and that sample. We also have a time series of T excess stock returns. We draw a sample (with replacement) of size T from the time series of excess stock returns. We then randomly match the length T series of the estimated γth consumption moments with the length T series of excess returns. This random matching ensures that, in this artificial world, excess returns are stochastically independent from moments of household consumption. Given this independence, there is, in the population, no value of γ that will zero out the equity premium.

We estimate γ by applying GMM to the moment condition (22) on these artificially generated data. We save the estimate of γ ($\hat{\gamma}^b$) as well as the mean equity premium error ($\bar{e}(\hat{\gamma}^b)$). We repeat the exercise $b = 1, 2, \ldots, B$ times (we set $B = 500$). This gives us a distribution of $\hat{\gamma}^b$ and $\bar{e}(\hat{\gamma}^b)$. If our test lacks power, we should find that, under the alternative hypothesis of no connection between consumption and returns, the probability that the PIPO SDF zeros out the equity premium ($|\bar{e}(\hat{\gamma}^b)| < 10^{-5}$, say) for realistic values of $\hat{\gamma}^b$ ($0 \leq \hat{\gamma}^b \leq 7$, say) is high. In fact, we estimate this probability to be rather low (18 percent for the U.S. sample and 19 percent for the U.K. sample).[24] Thus, we conclude that our finding that the PIPO SDF provides the best fit for the equity premium data is unlikely to be generated by simply a "lucky" draw.

---

[24] We omit the Italian case because the break in the consumption series makes the implementation of the power exercise more complex.

VII.   Conclusions

This paper considers two distinct models of asset trade with heterogeneous agents and derives representations for the model SDFs in terms of the cross-sectional consumption distribution. The first model is a standard incomplete markets model. We show that in that model, the equity premium is higher if the stock returns covary *positively* with consumption inequality. The second model is a model in which agents sign lifetime insurance/work contracts with employers or some other intermediary, and those employers trade assets. In that model, the equity premium is higher if stock returns covary negatively with consumption inequality. We estimate the two models in three different countries: United States, United Kingdom, and Italy. The first, incomplete markets, model is inconsistent with the equity premium in the United States, United Kingdom, or Italy for any specification of the coefficient of relative risk aversion. In contrast, in all three countries, the second (PIPO) model is consistent with the equity premium for an estimated coefficient of relative risk aversion of around 5. However, the model is consistent with the short-term risk-free rate only for low discount factors (less than 0.5).

The intuition behind the PIPO model's success with the equity premium is simple. With high consumption inequality, there are many consumption-poor skilled agents. Such agents are easy to motivate, and so employers find it cheap to provide their employees with incentives in these high-inequality states. This basic argument means that in the PIPO model, the state price of consumption is higher when inequality is low.[25] It turns out empirically that consumption inequality—especially right-tail consumption inequality—tends to be rising exactly when stock returns are high. Stocks are not as valuable as bonds, and so the equity premium is high.

References

Alvarez, Fernando, and Urban Jermann. 2001. "Quantitative Asset Pricing Implications of Endogenous Solvency Constraints." *Rev. Financial Studies* 14:1117–51.
Atkeson, Andrew, and Robert E. Lucas. 1992. "On Efficient Distribution with Private Information." *Rev. Econ. Studies* 59:427–53.
Attanasio, Orazio, James Banks, and Sarah Tanner. 2002. "Asset Holding and Consumption Volatility." *J.P.E.* 110:771–92.

---

[25] In the PIPO model, right-tail skill inequality should drive movements in the pricing kernel. In the online Appendix, we provide suggestive evidence showing a strong correlation between the γth moment of the U.S. consumption distribution (we choose $\gamma = 5$ to be consistent with our empirical estimates) and the share going to the top 0.01 percent of the wage distribution, which is a measure of right-tail skill inequality (the wage fractile estimates come from Piketty and Saez [2003]).

Attanasio, Orazio, Erich Battistin, and Hide Ichimura. 2004. "What Really Happened to Consumption Inequality in the US?" Working paper no. 10338, NBER, Cambridge, MA.

Attanasio, Orazio, and Steven Davis. 1996. "Relative Wage Movements and the Distribution of Consumption." *J.P.E.* 104:1227–62.

Attanasio, Orazio, and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *J.P.E.* 103:1121–57.

Balduzzi, Pierluigi, and Tong Yao. 2007. "Testing Heterogeneous-Agent Models: An Alternative Aggregation Approach." *J. Monetary Econ.* 54:369–412.

Betson, David. 1990. "Alternative Estimates of the Cost of Children from the 1980–1986 Consumer Expenditure Survey." U.S. Dept. Health and Human Services, Washington, DC.

Bosworth, Barry, Gary Burtless, and John Sabelhaus. 1991. "The Decline in Saving: Evidence from Household Surveys." *Brooking Papers Econ. Activity,* pp. 183–241.

Bound, John, and Alan Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *J. Labor Econ.* 9 (1): 1–24.

Brav, Alon, George Constantinides, and Christopher C. Geczy. 2002. "Asset Pricing with Heterogeneous Consumers." *J.P.E.* 110:793–824.

Campbell, John Y. 1999. "Asset Prices, Consumption, and the Business Cycle." In *Handbook of Macroeconomics,* vol. 1, edited by John B. Taylor and Michael Woodford, 1231–1303. Amsterdam: North-Holland.

Citro, Constance F., and Robert T. Michael. 1995. *Measuring Poverty: A New Approach 1995.* Washington, DC: Nat. Acad. Press.

Cochrane, John. 2001. *Asset Pricing.* Princeton, NJ: Princeton Univ. Press.

Cogley, Timothy. 2002. "Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey." *J. Monetary Econ.* 49:309–34.

Golosov, Michael, Narayana Kocherlakota, and Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Rev. Econ. Studies* 70:569–87.

Green, Edward. 1987. "Lending and the Smoothing of Uninsurable Income." In *Contractual Agreements for Intertemporal Trade,* edited by E. Prescott and N. Wallace. Minneapolis: Univ. Minnesota Press.

Hansen, Lars Peter, and Robert J. Hodrick. 1980. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis." *J.P.E.* 88 (5): 829–53.

Hansen, Lars Peter, and Ravi Jagannathan. 1997. "Assessing Specification Errors in Stochastic Discount Factor Models." *J. Finance* 52:557–90.

Hansen, Lars Peter, and Kenneth J. Singleton. 1982. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica* 50:1269–86.

Kocherlakota, Narayana R. 1990. "On Tests of Representative Consumer Asset Pricing Models." *J. Monetary Econ.* 26:285–304.

———. 1996. "The Equity Premium: It's Still a Puzzle." *J. Econ. Literature* 34:42–71.

———. 1998. "The Effects of Moral Hazard on Asset Prices When Financial Markets Are Complete." *J. Monetary Econ.* 41:39–56.

———. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Optimal Dynamic Taxation." *Econometrica* 73:1587–1622.

Kocherlakota, Narayana, and Luigi Pistaferri. 2007. "Household Heterogeneity and Real Exchange Rates." *Econ. J.* 117:C1–C25.

Light, Audrey, and Kathleen McGarry. 2004. "Why Parents Play Favorites: Explanations for Unequal Bequests." *A.E.R.* 94 (5): 1669–81.

Ligon, Ethan. 1996. "Risk-Sharing and Information: Theory and Measurement in Village Economies." Working paper, Univ. California, Berkeley.

———. 1998. "Risk-Sharing and Information in Village Economies." *Rev. Econ. Studies* 65:847–64.

———. 2005. "Formal Markets and Informal Insurance." *Internat. Rev. Law and Econ.* 25 (1): 75–88.

Mankiw, Gregory, and Stephen Zeldes. 1991. "The Consumption of Stockholders and Non-stockholders." *J. Financial Econ.* 29:97–112.

McGarry, Kathleen, and Robert Schoeni. 1995. "Transfer Behavior in the Health and Retirement Study: Measurement and the Redistribution of Resources within the Family." *J. Human Resources* 30:S184–S226.

Meghir, Costas, and Luigi Pistaferri. 2004. "Income Variance Dynamics and Heterogeneity." *Econometrica* 72:1–32.

Mehra, Rajnish, and Edward Prescott. 1985. "The Equity Premium: A Puzzle." *J. Monetary Econ.* 15:145–61.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Q.J.E.* 118 (1): 1–39.

Postel-Vinay, Fabian, and Jean-Marc Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70:2295–2350.

Rogerson, William. 1985. "Repeated Moral Hazard." *Econometrica* 53:69–76.

Semenov, Andrei. 2004. "High-Order Consumption Moments and Asset Pricing." Working paper, York Univ.

Storesletten, Kjetil, Christopher Telmer, and Amir Yaron. 2001. "How Important Are Idiosyncratic Shocks? Evidence from Labor Supply." *A.E.R. Papers and Proc.* 91:413–17.

Vissing-Jorgensen, Annette. 2002. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *J.P.E.* 110:825–53.