

January 2008

Household Heterogeneity and Asset Trade: Resolving the Equity Premium Puzzle in Three Countries*

Narayana Kocherlakota

University of Minnesota, FRB-Minneapolis, and NBER

Luigi Pistaferri

Stanford University and CEPR

ABSTRACT

We consider two asset pricing models with agents who have privately observable skill shocks. The first is a standard incomplete markets model. In the second, agents sign lifetime work/insurance contracts with intermediaries who then trade assets. For each model, we find a stochastic discount factor in terms of the cross-sectional distribution of consumption. We use this representation to evaluate the models' predictions for the equity premium in three countries (US, UK and Italy). The first model fares poorly in all countries. The second model is consistent with the equity premium in all three countries when the CRRA is roughly 5.

*This paper originally circulated under the title, "Asset Pricing Implications of Pareto Optimality with Private Information." Kocherlakota acknowledges the support of NSF 0305833 and NSF 0606695. The paper has benefited from comments from participants in seminars at a number of schools. We thank Sonam Sherpa for research assistance, and Fernando Alvarez, Tim Cogley, Andrea Eisfeldt, Peter Hansen, Dirk Krueger, Felix Kubler, Jonathan Parker, and Adriano Rampini for their comments. We thank Orazio Attanasio and Guglielmo Weber for their help with the FES and SFB data respectively. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

In the United States, over the period 1980-2004, the average return to stocks was about 8% 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 condition of optimality

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

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 regularity conditions, the unconditional expectation in (1) can be consistently estimated using a time-series average. Kocherlakota (1996) documents that the sample analog of this first order is not satisfied in United States data (from 1889-1979) except for implausibly large values of γ .

The equity premium puzzle implicitly assumes that per-capita consumption 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 the second model, households first commit to a lifetime risk-sharing and work arrangement with some intermediary (like 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.

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

$$(2) \quad E \left[\frac{C_{-\gamma,t+1}}{C_{-\gamma,t}} (R_{t+1}^s - R_{t+1}^b) \right] = 0$$

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

$$(3) \quad E \left[\frac{C_{\gamma,t}}{C_{\gamma,t+1}} (R_{t+1}^s - R_{t+1}^b) \right] = 0$$

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 formulae 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. Because of 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 wealth effects in labor supply, it is cheap to provide incentives to these low-consumption skilled people. This cheapness in providing incentives means that intermediaries do not put a high value on stocks if they pay off in these 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 based on the availability of household consumption data for a sufficiently long period of time) by constructing sample analogs 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 (like the FES and SFB), without any need for a panel component.

We then apply GMM to the sample analogs of (1)-(3) to estimate γ in each country. We find that in all three countries, there is no value of γ in the interval $[0, 10]$ that satisfies sample analogs of either (1) or (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. In contrast, the incomplete markets model and representative agent models are not able to rationalize the equity premium for any value of γ in the interval $[0, 10]$.

2. Prior Literature

Our paper is related to Ligon (1998), who tests the risk-sharing implications of Pareto optimality with moral hazard. His approach is as follows. He uses consumption data from South Indian villages (the ICRISAT data set). 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:

$$(4) \quad E_t\{(c_{it+1}/c_{it})^a\} = 1$$

Under the former hypothesis of constrained Pareto optimality, a equals γ . Under the latter hypothesis of risk-free borrowing and lending, a equals $-\gamma$. Using the Generalized Method of Moments, 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 non-trivial movements in expected asset returns. As well, we do not need to assume that individual productivity shocks are 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 consumptions; we do not need to have panel data on consumption. Finally, our empirical analysis is more robust to measurement error than is his.¹

Our work is also related to recent papers using data from the CEX to evaluate incomplete markets models of asset pricing. In recent work, Cogley (2002), Brav, Constantinides, and Geczy (BCG) (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 who 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 BCG, 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 these IMRS's 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

¹Ligon (1996) discusses some ideas for dealing with measurement error, persistence, and relaxing risk-neutrality in Section 6. This working paper is the basis for Ligon (2005).

necessarily the same. One needs at least some panel data to implement the average IMRS SDF empirically, while one only needs 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, BCG (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 as the average is constructed using samples of agents with *larger* asset positions.²

Balduzzi and Yao (2004) 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 \$2000 or more in total financial assets.³

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 real exchange rates.⁴ We look at the implications of the models in the UK and Italy, not just the United States. Finally, with the exception of BCG, these other papers rely on Taylor series approximations of the relevant stochastic discount factors. 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.

²Vissing-Jorgensen (2002) does not look at restrictions involving multiple assets (like the equity premium restriction (1)). Instead, she examines the asset pricing equations for bonds and stocks separately.

³Balduzzi and Yao (2004) provide no specific definition of what they mean by "financial assets". When we examine the impact of financial market participation in what follows, we use the definitions of Attanasio, Banks and Tanner (2002) and Vissing-Jorgensen (2002).

⁴Kocherlakota 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 this paper predate the first versions of Kocherlakota and Pistaferri (2007).

3. 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 $0, 1, 2, \dots, T$; 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

$$\sum_{t=1}^T \beta^{t-1} \{c_t^{1-\gamma}/(1-\gamma) - v(l_t)\}, 0 < \beta < 1, 0 < \gamma$$

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 shocks 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 $\theta^t = (\theta_1, \dots, \theta_t)$ and the history of public shocks $z^t = (z_1, \dots, z_t)$. This implies that his choices in period t can only be a function 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

$$(5) \quad y_t = \phi_t(\theta^t, z^t)l_t$$

$$(6) \quad \phi_t : \Theta^t \times Z^t \rightarrow (0, \infty)$$

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 $Var(\phi_t(\theta^t, z^t)|z^t)$ 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 .⁵

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

$$(7) \quad c_t : \Theta^t \times Z^t \rightarrow R_+$$

$$(8) \quad y_t : \Theta^t \times Z^t \rightarrow [0, \bar{y}]$$

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 :

$$(9) \quad \sum_{\theta^t \in \Theta^t} c_t(\theta^t, z^t)\pi(\theta^t) \leq \sum_{\theta^t \in \Theta^t} y_t(\theta^t, z^t)\pi(\theta^t)$$

Because θ_t is only privately observable, allocations must respect incentive-compatibility conditions. (The following definitions correspond closely to those in Golosov, Kocherlakota and Tsyvinski (2003).) A *reporting strategy* $\sigma : \Theta^T \times Z^T \rightarrow \Theta^T \times Z^T$, where σ_t is (θ^t, z^t) -measurable and $\sigma(\theta^T, z^T) = (\theta^T, z^T)$. An agent can alter his consumption and labor by changing his reporting strategy. Suppose the allocation is (c, y) .

⁵Attanasio and Davis (1996) document that the cross-sectional dispersion of consumption increased in the 1980's 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.

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 $\sigma^t(\theta^t, z^t) = (\sigma_s(\theta^s, z^s))_{s=1}^t$, 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

$$V(\sigma_{TT}; c, y) \geq V(\sigma; c, y) \text{ for all } j \text{ and all } \sigma.$$

An allocation which is incentive-compatible and feasible is said to be incentive-feasible.

4. 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 η -th non-central moment of the cross-sectional distribution of consumption.

A. Incomplete Markets Equilibrium

Suppose that at each date $t = 0, 1, 2, \dots, (T - 1)$, all agents can buy and sell a complete set of z^{t+1} -contingent claims to consumption.⁶ Fix agents' labor choices at their individually optimal levels. Then, the choice problem of a typical agent is

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

⁶The SDF derived in this subsection would still be valid even if agents traded a proper subset of these claims.

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

$$\sum_{z_1} q_1(z_1)W_1(\theta_1, z_1) \leq W_0$$

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

W_0 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 :

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

If we take expectations over θ^t , we get:

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

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 γ -th 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 (Goloso and Tsyvinski (2006)). This trading protocol is similar to that described in Goloso and Tsyvinski (2006) (and 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 equilibrium in this trading protocol. 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.

$$\begin{aligned}
& \min_{c, W} \sum_{z_1} q_1(z_1) W_1(z_1) \\
& \text{s.t. } \sum_{\theta^t} \pi(\theta^t) c_t(\theta^t, z^t) + \sum_{z_{t+1}} q_{t+1}(z_{t+1}|z^t) W_{t+1}(z^{t+1}) \leq \sum_{\theta^t} \pi(\theta^t) y_t^*(\theta^t, z^t) + W_t(z^t), t \geq 2 \\
& \sum_{t=1}^T \beta^{t-1} \sum_{z^t} \Psi(z^t) \sum_{\theta^t} \pi(\theta^t) c_t(\theta^t, z^t)^{1-\gamma} / (1-\gamma) = u_* \\
& V(c, y^*; \sigma_{TT}) \geq V(c, y^*; \sigma) \text{ for all } \sigma.
\end{aligned}$$

We can derive a useful necessary condition of insurer optimality as follows. Suppose (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} :

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

This perturbation is akin to that in Kocherlakota (2005).⁷ It 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

⁷See also Rogerson (1985).

(c^*, W^*) is truly optimal for the insurer, the perturbation must not be cheaper.

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

$$\begin{aligned}
& \min_{\varepsilon, \delta, \eta} -\varepsilon\pi(\theta^t) + \delta q_{t+1}(\bar{z}_{t+1}|\bar{z}^t) \\
& \text{s.t. } (c_t^*(\theta^t, \bar{z}^t) - \varepsilon)^{1-\gamma}/(1-\gamma) + \beta\Psi(\bar{z}_{t+1}|\bar{z}^t)(c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1}) + \eta(\theta^{t+1}))^{1-\gamma}/(1-\gamma) \\
& = c_t^*(\theta^t, \bar{z}^t)^{1-\gamma}/(1-\gamma) + \beta\Psi(\bar{z}_{t+1}|\bar{z}^t)c_{t+1}^*(\theta^{t+1}, \bar{z}^{t+1})^{1-\gamma}/(1-\gamma) \text{ for all } \theta^{t+1} \\
& \sum_{\theta^{t+1}} \pi(\theta^{t+1})\eta(\theta^{t+1}) = \delta
\end{aligned}$$

The necessary conditions associated with this optimum are:

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

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

We can use these necessary conditions to obtain a useful representation for the stochastic discount factor in this trading structure. Substituting the second condition into the first, we get:

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

If we add over θ^t , and using 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 γ -th cross-sectional moment of consumption. Obviously, this argument can be generalized to any z^{t+1} .

C. Three Models and Their Implications for the Equity Premium

In this subsection, we use the above 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 which 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 which has price 1 in history z^t . The absence of arbitrage opportunities guarantees that its return must satisfy:

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

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

$$1 = E(SDF_{t+1} R_{t+1}^i | z^t)$$

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

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

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:

$$(10) \quad 0 = E[SDF_{t+1} (R_{t+1}^s - R_{t+1}^b)]$$

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 (10) using time-series averages. The *equity premium puzzle* says that it is difficult to find a model of the stochastic discount factor that satisfies the sample analog of (10). One way to see the problem is to write (10) as:

$$(11) \quad 0 = E(R_{t+1}^s - R_{t+1}^b) + \frac{1}{E(SDF_{t+1})} Cov[SDF_{t+1}, (R_{t+1}^s - R_{t+1}^b)]$$

The expected difference between stock returns and Treasury bill returns, at least as estimated via time-series averages in the United States data, is on the order of 7-8% per year. To resolve the equity premium puzzle, we need to find a stochastic discount factor that covaries sufficiently negatively with the realized equity premium so as to offset the large first term on the right hand side of (11).

The previous subsection suggests two candidate stochastic discount factors (SDFs). The first is what we will term the incomplete markets (INC) SDF:

$$(12) \quad SDF_{t+1}^{INC}(z^{t+1}; \beta, \gamma) = \beta \frac{C_{-\gamma, t+1}(z^{t+1})}{C_{-\gamma, t}(z^t)}$$

The second is what we will term the private information-Pareto optimal (PIPO) stochastic discount factor:

$$(13) \quad SDF_{t+1}^{PIPO}(z^{t+1}; \beta, \gamma) = \beta \frac{C_{\gamma,t}(z^t)}{C_{\gamma,t+1}(z^{t+1})}$$

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

$$(14) \quad SDF_{t+1}^{RA}(z^{t+1}; \beta, \gamma) = \beta \frac{C_{1,t+1}(z^{t+1})^{-\gamma}}{C_{1,t}(z^t)^{-\gamma}}$$

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 (10), unless γ is implausibly 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 example, consider the INC SDF. When γ is large, the incomplete markets SDF is high when the left tail of the consumption distribution is growing heavier. Hence, the incomplete markets SDF can resolve the equity premium puzzle if the left tail of the consumption distribution grows heavier when realized excess stock returns tend to be low. Intuitively, an increase in the heaviness of the left tail represents an increase in idiosyncratic risk. If stock returns are low when idiosyncratic risk is high, stocks are a poorer instrument for precautionary savings than bonds, and households will demand a higher return for holding stocks.

Now instead consider the PIPO SDF. If $\gamma > 1$ (the empirically relevant case), the SDF is large when the right tail of the consumption distribution is heavy. This model can resolve the equity premium puzzle if excess stock returns are positively correlated with growth in the heaviness of the right tail of the consumption distribution. The intuition behind this prediction is somewhat subtle. In the PIPO model, the insurers are trading assets on behalf of their insurees, and so pricing is determined by the insurers' demand for the various assets. If the right tail of the consumption distribution is very heavy, then there are many consumption-poor skilled households. Because of wealth effects in labor supply, it is cheap to provide incentives to these people, and so intermediaries have a low shadow cost of generating resources when the right tail is heavy. For this reason, they especially value (that is, demand low returns for) assets that have large payoffs when consumption inequality is low. We can conclude that, with the PIPO SDF, the equity premium is high if the

realized equity premium covaries positively with growth in the heaviness of the right tail of the consumption distribution.

It is important to distinguish the INC SDF from the one employed by Brav, Constantinides, and Geczy (BCG) (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:

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

which is the average of the agents' intertemporal marginal rates of substitution. Like the incomplete markets SDF, this average IMRS discount factor is also a valid SDF in an incomplete markets equilibrium with non-binding 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.

D. 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 in consumption is that the data are typically measured with error. This measurement error usually creates difficulties when one applies the Generalized Method of Moments to estimate Euler equations of the form:

$$\beta E_t [(c_{t+1}/c_t)^{-\gamma} R_{t+1}] = 1$$

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 $c^* : \Theta^T \times Z^T \rightarrow R_+^T$ be the true allocation of consumption. We allow c^* to be measured with error as follows. Let $(\nu_1, \nu_2, \dots, \nu_T)$ 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 $\widehat{c}_t(\theta^t, z^t, \nu_t) = \exp(\nu_t)c_t^*(\theta^t, z^t)$ to be measured consumption. Define also:

$$(15) \quad \widehat{C}_{\eta,t} = E\{\widehat{c}_t^\eta | z^t\}$$

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

$$(16) \quad \widehat{C}_{\eta,t} = E[c_t^{*\eta} \exp(\gamma \nu_t) | z^t]$$

$$(17) \quad = E[\exp(\gamma \nu_t) | z^t] E(c_t^{*\eta} | z^t)$$

$$(18) \quad = E[\exp(\gamma \nu_t)] C_{\eta,t}$$

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

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

$$(19) \quad \frac{\widehat{C}_{\eta,t}(z^t)}{\widehat{C}_{\eta,t+1}(z^{t+1})} = \frac{C_{\eta,t}(z^t)}{C_{\eta,t+1}(z^{t+1})}$$

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 is stationary over time. These assumptions about the nature of the measurement error are not wholly innocuous. On the other hand, 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.⁸

5. Empirical Analysis: Preliminaries

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

⁸There is no evidence from validation consumption studies that can tell us whether the assumption we make about the nature of the measurement error are truly restrictive. Evidence from validation wage and income studies (Bound and Krueger, 1991) have found that: (a) measurement error appears serially correlated, (b) independent of schooling, and (c) negatively correlated with the true measure. The latter finding will, of course, invalidate our empirical strategy.

A. Micro Data *The US CEX*

The microeconomic data for the US are drawn from the 1980-2004 Consumer Expenditure Survey (CEX). The CEX provides a continuous and comprehensive flow of data on the buying habits of American consumers. The data are collected by the Bureau of Labor Statistics and used primarily for revising the CPI. Consumer units are defined as members of a household related by blood, marriage, adoption, or other legal arrangement, single person living alone or sharing a household with others, or two or more persons living together who are financially dependent. 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, such as food, personal care, and household supplies. The Interview sample is in the form of a rotating panel, and it follows survey households for a maximum of 5 quarters, although only inventory and basic sample data are collected in the first quarter (these data are not publicly available.) The data base covers about 95% of all expenditure, with the exclusion of expenditures for housekeeping supplies, personal care products, and non-prescription drugs. Following most previous research, our analysis below uses only the Interview sample.

The CEX collects information on a variety of socio-demographic variables, including characteristics of members, characteristics of housing unit, geographic information, inventory of household appliances, work experience and earnings of members, unearned income, taxes, and other receipts of consumer unit, credit balances, assets and liabilities, occupational expenses and cash contributions of consumer unit. Expenditure is reported in each interview (after the first) and refers to the months of the previous quarter. Thus, a household interviewed in April 1980 reports expenditure for January, February, and March 1980. Income is reported in the second and fifth interview, and it refers to the previous twelve months. Holdings of financial assets are reported only in the fifth interview.

We refer the reader to the Appendix for step-to-step details on sample selection and consumption definition. Our sample selections are aimed at eliminating the most severe reporting errors in consumption.⁹

⁹An alternative (or a further sample selection) is to remove observations in the tails of the cross-sectional distribution of

We end up discarding about 25% of observations through our selection procedure.¹⁰ The definition of total nondurable consumption is similar to Attanasio and Weber (1995). It includes food (at home and away from home), alcoholic beverages and tobacco, heating fuels and utilities, transports (including gasoline), personal care, clothing and footwear, entertainments, other services (including domestic services). It excludes expenditure on various durables, housing (furniture, appliances, etc.), education and health.

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 de-seasonalized by simple multiplicative regression adjustments. Finally, we convert it into adult-equivalent consumption data.¹¹ Given the overlapping panel nature of the CEX, each month a certain number of households enter the panel and an approximately equal number leave it. Monthly consumption data are aggregated to form quarterly consumption data for each household in the sample. Then, we aggregate across households to form moments of the quarterly consumption distribution. Note that households start their second interview (when consumption data are first collected) in different months. Thus, some households’ second interview covers the months of January through March, some other households’ second interview will have data for the months of February through April, and so forth. By the very design of the CEX, no households contribute multiple observations to adjacent overlapping quarters. In other words, a household that contributes data to January-March 1980 will not contribute data for February-April (or March-May). Its next contribution, if that exists, will be for April-June 1980.

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 (see Panel A of Table 1 for descriptive statistics on average household expenditure). Some of the discrepancy is undoubtedly due to differences in covered population and definitional issues. But the amount of underestimation

consumption. Our sample selection is likely already removing some of these observations. A sample selection of this form drops extreme errors, but also genuine observations (the very rich or the very poor). This is undesirable in the context of the theory we are studying. For similar reasons, we do not use a Taylor series approximation to the various SDFs.

¹⁰The starting sample has an average of 1911 households in any given (overlapping) quarter. Our final sample has an average of 1412 household per (overlapping) quarter.

¹¹The number of adult equivalents is defined as $(A + \alpha K)^\beta$ where A is the number of adults (aged 18 or more), K the number of kids, and α and β parameters. We set $\alpha = 0.7$ and $\beta = 0.65$ (following recommendations contained in Citro and Michaels, 1995, which in turn draws from Betson, 1990). Similar results are obtained if we use a more sophisticated Engel approach.

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.¹² 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 et al (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 has maintained representativeness of the US population over time, and attrition has not changed much since the redesign of the survey of the early 1980s.¹³

The UK FES

The micro data for the UK come from the 1975-1999 Family Expenditure Survey (FES). The FES, conducted by the Office for National Statistics, surveys a random sample of households in the UK. The FES is primarily a survey of household expenditure 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 Consumer Price Index. However, with time the survey has become multi-purpose, providing a wealth of information on household economic and social variables, such as household composition, size, social class, occupation, etc.

The FES has been in operation since 1957 and up to and including 1993, data is available by calendar year. From 1994-1995 data is available by financial year (April-March). The basic unit of the survey is the household. Starting with 2000-01, a household is defined as a group of people living at the same address with common housekeeping, i.e., sharing household expenses such as food and bills, *or* sharing a living room. Before 2000, the definition of household required both common housekeeping *and* a shared living room. On average, about 7,000 households are interviewed each year. Each individual in the household aged 16 or more

¹²See Attanasio, Battistin, and Ichimura (2004).

¹³Given these differences between the CEX data and the NIPA data, it is useful to check whether similar results are obtained using the latter. To this purpose, we also estimated the parameters in the representative agent SDF using aggregate NIPA PCE data. We obtain NIPA PCE data from the NIPA Table 2.8.5, which reports Personal Consumption Expenditures by major type of product (durable goods, non durable goods, and services) on a monthly basis (all the NIPA tables can be found at <http://www.bea.gov/bea/dn/nipaweb/index.asp>). The data are collected by the Bureau of Economic Analysis. Our measure of consumption is Personal Consumption Expenditures on nondurable goods (this is comparable to the measure of consumption we construct in the CEX, where services from durables are missing). The data are seasonally adjusted at annual rates, deflated using the same monthly CPI we use to deflate CEX data, and divided by the US population (midperiod estimates). These adjustments mimic those implemented for the micro CEX data as to ensure comparability. The monthly data so obtained are summed to form overlapping quarterly consumption data, the same data construction criterion used in the CEX (thus, consumption in 1980:3 refers to January-March 1980, consumption in 1980:4 to February-April 1980, and so on). However, changing the measure of consumption in this way had little impact on our results for the representative agent case (the results are available on request).

is asked to keep diary records of daily expenditure for two weeks. Information about regular expenditure, 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 is collected throughout the year to cover seasonal variations in expenditures.

The FES consists of three main modules, a Household Schedule, an Income Schedule, and Diary Records. The Household Schedule is taken at the main interview. Information for most of the questions is obtained from the head of household (or housewife). Information is collected about the household, the sex and age of each member, and also details about the type and size of the household accommodation. The main part of the questionnaire relates to expenditure both of a household and individual nature, but the questions are mainly confined to expenses of a recurring nature.

The Income Schedule data are collected for each household spender. The schedule is concerned with income, national insurance contributions and income tax. Information collected includes: employment status and recent absences from work, earnings of an employee, self-employed earnings, National Insurance contributions, pensions and other regular allowances, occasional benefits, social security benefits and other types, investment income, miscellaneous earnings of a “once-only” character, tax paid directly to Inland Revenue or refunded, income of a child.

The Diary Records cover fourteen days and each household member aged 16 or more is asked to record all expenditure made during this period. As for the CEX, we refer the reader to the Appendix for step-to-step details on sample selection and consumption definition. The definition of total nondurable consumption is as in Attanasio and Weber (1993), and it is meant to be comparable to the one for the CEX.

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 (in the same basis as the US CPI). Then, data are de-seasonalized by simple multiplicative regression adjustments. Finally, we convert it into adult-equivalent consumption data. Given that the data we have are average weekly expenditure, we multiply FES expenditure by 13 to form quarterly consumption

data for each household in the sample.¹⁴

As stated above, the FES randomly selects households throughout the year. We identify households interviewed in a certain month as those completing their two-week reporting in that month. We treat their (scaled-up) non-durable consumptions as being a random sample of household consumption over the prior quarter.

The Italian SFB

Micro data for Italy are drawn from the 1985-2002 Survey on Family Budgets (SFB), which is conducted every year by ISTAT (the Central Statistics Office). The main purpose of the SFB is to obtain information about the level and the composition of consumption of Italian households. 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 seven-day period. Information is also collected about own-consumption, goods and services provided in kind by employers in lieu of pay, and imputed rents. Data on purchases made for reasons other than consumption are typically not collected (i.e., tax payments, work-related purchases, etc.). Consumption data are collected for large sub-aggregates, such as food, housing, furniture, apparel, health, transports and communications, leisure, entertainment, education, and other goods and services. The survey also includes demographic information about all the members of the household, information about the occupied dwelling, income and savings. Income and savings are poorly measured and, since 1997, only available 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, beside 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 $SDF_{1997:1}^{PIPO}$, $SDF_{1997:2}^{PIPO}$, and $SDF_{1997:3}^{PIPO}$. We follow similar logic for obtaining estimates under

¹⁴As far as reliability is concerned, great care is taken in collecting information from households and comprehensive checks are applied during processing, so that errors in recording and processing are minimized. The main factors that affect the reliability of the survey results are sampling variability, non-response bias and some incorrect reporting of certain items of expenditure and income. Procedures are in place to ensure that users are provided with high quality data. For example, quality control is carried out to ensure that any outliers are genuine, and checks are made on any unusual changes in average spending compared with the previous year.

the other two models.

In what follows, we describe the post-1996 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 definition is extended to include all individuals who co-reside habitually with the household. Sampling is in two stages. The first stage is at the level of municipalities, and it is stratified. The second stage is at the level of households. The Italian territory is divided into 232 strata. Of these, 107 strata ("main strata") include a single municipality (typically, the largest cities). These municipalities are in the survey every month. The remaining 125 strata ("residual strata") include municipalities within the same region and are roughly of the same demographic size. From each residual stratum, three municipalities are drawn, to be surveyed on the first, second, and third month of a given quarter, respectively. Each quarter the process starts afresh.

The sampling design in the second stage is at the level of households. 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%, which compares well with other surveys, such as the FES in the UK or the CEX in the US. Two data collection techniques are used. The first is a standard face-to-face interview; the second is compilation of a diary where the household records all purchases made over the last 7 days. Each household is interviewed only once and then leaves the survey.

Our definition of consumption is as close as possible to the one for the US and the UK. It's the sum of food (at home and away from home), alcohol and beverages, tobacco, apparel, housing services, utilities, vehicle expenses, gasoline, public transports, personal hygiene, reading items, TV subscription, out-of-pocket health spending, education, communications, rented dwellings. Details about items included are in the Appendix.

As with 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 (in the same 2000 basis as the CPIs for the other two countries). Then, data are de-seasonalized 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.

As stated above, the SFB randomly selects households throughout the year. The survey keeps track of those households' consumption expenditures for a seven-days period. We identify households interviewed in a certain month as those completing their seven-days reporting in that month. We treat their (scaled-up) non-durable consumptions as being a random sample of household consumption over the prior quarter.

B. Asset Returns Data

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

The risk free rate R^b is obtained in the following way. First, we extract the one-month nominal returns on Treasury bills.¹⁵ 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 US 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 one-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 UK we use the FTSE All-Share Return Index.¹⁶ 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,

¹⁵For Italy the source is the Central Bank of Italy. For the UK, it is the Central Statistical Office, Annual Abstract of Statistics.

¹⁶For Italy the source is Banca Commerciale Italiana, for the UK is the Central Statistical Office Annual Abstract of Statistics (before 1989), and Eurostat (1989 onward).

historically, the gap between average stock returns and average Treasury bill returns is very large (on the order of 6% per year) and difficult to rationalize using standard representative agent asset pricing models. As in Kocherlakota (1996), we assess the candidate stochastic discount factors' ability to rationalize the equity premium by considering the restriction that:

$$(20) \quad E [SDF_t (R_t^s - R_t^b)] = 0$$

where as before SDF_t is the stochastic discount factor, R_t^s is the return 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 analogs of the various stochastic discount factors using these *cross-sectional* data. In particular, let:

$$(21) \quad \widehat{SDF}_t^{PIPO}(\gamma) = \frac{N_{t-3}^{-1} \sum_{i=1}^{N_{t-3}} c_{it-3}^\gamma}{N_t^{-1} \sum_{i=1}^{N_t} c_{it}^\gamma}$$

$$(22) \quad \widehat{SDF}_t^{INC}(\gamma) = \frac{N_t^{-1} \sum_{i=1}^{N_t} c_{it}^{-\gamma}}{N_{t-3}^{-1} \sum_{i=1}^{N_{t-3}} c_{it-3}^{-\gamma}}$$

$$(23) \quad \widehat{SDF}_t^{RA}(\gamma) = \frac{\left(N_t^{-1} \sum_{i=1}^{N_t} c_{it}\right)^{-\gamma}}{\left(N_{t-3}^{-1} \sum_{i=1}^{N_{t-3}} c_{it-3}\right)^{-\gamma}}$$

denote the sample analogs 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 US CEX the first available observation for $\widehat{SDF}_t^{PIPO}(\gamma)$ is for 1980:3, and it is constructed as the ratio of the γ -th moment of consumption for 1979:12 (calculated using all households reporting expenditure data for October-December 1979) and the γ -th 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 $\widehat{SDF}_t^j(\gamma)$ ($j = PIPO, INC, RA$) for the US, $T = 297$ for the UK, and $T = 210$ for Italy (after dropping the first three months of 1997 due to the change in the sampling design described above).

We then form sample analogs (in the *time series* dimension) of the restriction (20). For example, the sample analog of the left hand side of (20) in the PIPO model is:

$$(24) \quad T^{-1} \sum_{t=1}^T \frac{N_{t-3}^{-1} \sum_{i=1}^{N_{t-3}} c_{it-3}^{\gamma}}{N_t^{-1} \sum_{i=1}^{N_t} c_{it}^{\gamma}} (R_t^s - R_t^b)$$

We estimate the unknown parameter γ and evaluate the three discount factors by applying Generalized Method of Moments (GMM) to this moment condition. We estimate the value of γ that satisfies (20) separately for each country (results reported in Table 3), and then pooling all the country data together (Table 4).

Finally, it is worth discussing the non-standard inference problem we face. Note that time-series moment conditions like (24) are functions of cross-sectional non-linear 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 formulae to calculate standard errors (perhaps after accounting for serial correlation induced by the use of overlapping data). 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 BCG (2002).¹⁷ We compute our standard errors to account for serial correlation in returns and overlapping data using a correction of the form proposed by Hansen and Hodrick (1980).

6. Empirical Analysis: Results

We provide some simple summary statistics in Table 1 (for the micro data) and Table 2 (the time series data). Some demographic features are well known. US households are younger and more likely to be college educated. Italian (and to some extent US) households are aging rapidly. UK and especially Italian households have reduced fertility rates. The data show strong consumption growth in both the UK and Italy, and less so in the US.

Table 2 show that there is a large equity premium in all countries. The mean return to stocks is about

¹⁷An 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.

1.9% per quarter higher than the mean return to Treasury bills in the US. This sample estimate is higher than the 6.2% annual number averaged in the hundred years of data (1889-1978) studied by Mehra and Prescott. The standard deviation of stock returns is about 7.7% per quarter. In the UK, the premium is even higher (2.8% per quarter), while it is slightly lower in Italy (1.3% per quarter).

We also plot the PIPO stochastic discount factor in Figure 1 for $\gamma = 5$ and the Incomplete Markets SDF in Figure 2 for $\gamma = 2$ (as we will see, these are the values of γ that minimize the equity premium pricing error in the two models). For these 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.

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:

$$(25) \quad \bar{\varepsilon}_t^j(\gamma) = \frac{1}{T} \sum_{t=1}^T \widehat{SDF}_t^j(\gamma) (R_t^s - R_t^b)$$

for $j = PIPO$, INC , and RA . Equation (25) is the empirical analog of (20) and $\widehat{SDF}_t^j(\gamma)$ is defined by equations (21)-(23).

In Table 3 we estimate the coefficient of relative risk aversion γ that minimizes (25) by applying the Generalized Method of Moments 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 3. 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. Contrast these findings 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.¹⁸ Note also that the usual GMM estimate of the standard error is undefined because the

¹⁸BCG (2002) restrict attention to households with non-negative 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

(one) moment condition we use is not set equal to zero in sample. As for the representative agent model, we find that in the US 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 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.¹⁹

In Table 4 we impose that γ is the same across countries and consider the set of orthogonality conditions

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

where $e_{k,t}^j(\gamma) = SDF_{k,t}(\gamma, z^t) (R_{k,t}^s(z^t) - R_{k,t}^b(z^t))$ 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 $\boldsymbol{\varepsilon}_t(\gamma)$ 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 bad properties of GMM in small sample, we use the identity matrix as our weighting matrix.²⁰

Table 4 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%) for a value of γ of around 5. This confirms that the PIPO model provides the best fit for the equity premium data.

Our results for the incomplete markets SDF contrast with the results of BCG (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

Tables 3-5.

¹⁹For the US, we obtain qualitatively similar evidence if we use a market return that includes stocks listed on NASDAQ.

²⁰We use the formula for the overidentifying test statistic reported in Cochrane (2001).

measurement error processes that we assume in this paper. However, there are two main differences between what we do and what BCG 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, BCG’s sample selection is different than ours: They only keep households who stay in the sample for three or more quarters (because they use the average IMRS SDF) and, to minimize measurement error, discard households who report extremely large increases or decreases in consumption from one quarter to another. Their sample selections end up discarding about 60% 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.²¹

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 5th and final interview, to report information on their current holdings of assets. Separate questions are asked about “savings accounts in banks, savings and loans, credit unions, and similar accounts” (saving 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, more, or less than those

²¹We constructed a subsample of the CEX using the selection criteria reported in their paper and were able to replicate most of the results of their paper. In this sample, the sample equity premium is eliminated using the PIPO discount factor when we set γ between 9 and 10.

held 12 months earlier, and in case they differ, to report the difference. This means 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 5th and 2nd interview. This means that we can construct the SDF for financial market participants only in one case, namely for data collected in the 2nd ($t - 3$) and 3rd (t) interview.²²

We adopt two alternative classifications of financial market participation at the beginning of period t . The first follows Attanasio, Banks and Tanner (2002) and assumes that it coincides with stockholding. The second follows Vissing-Jorgensen (2002) and equates it with holding of both stocks and bonds. (This is the most appropriate case for the equity premium equation; the equity premium first order condition is only necessarily satisfied for agents who participate both in stock markets and bond markets.) The definition of stockholder and bondholder is as in Vissing-Jorgensen (2002).²³ The results, reported in Table 5, show that accounting 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 5, 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. 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

²²Thus, in our experiment the numerator of the SDF uses only consumption data collected in the 3rd interview, while the denominator uses only data collected in the 2nd interview. Both refer to financial market participants at the beginning of the 2nd interview as reported during their 5th interview.

²³A household is classified as stockholder (bondholder) at the beginning of period t if it: (a) reports a positive amount held in stocks (bonds) in the 5th 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 (bond) held by less than the current amount, (d) reports an increase 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.

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 γ -th 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 γ -th 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 population, no value of γ that will zero out the equity premium.

We estimate γ by applying GMM to the moment condition (20) 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 $b = 1, 2, \dots, B$ 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% for the US sample and 19% for the UK 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.

7. 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 co-vary *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 co-vary *negatively* with consumption inequality. We estimate the two models in three different countries: United States, United Kingdom, and

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

Italy. In all three, the second (PIPO) model is consistent with the equity premium for an estimated coefficient of relative risk aversion of around 5. The first, incomplete markets, model is inconsistent with the equity premium in any of the three countries for any coefficient of relative risk aversion.

The intuition behind our finding 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. 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

- [1] ATKESON, A., and LUCAS, R. E. (1995), "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory* 66: 64-88.
- [2] ATTANASIO, O., BANKS, J., and TANNER, S. (2002), "Asset Holding and Consumption Volatility," *Journal of Political Economy* 110, 771-792.
- [3] ATTANASIO, O., BATTISTIN, E., and ICHIMURA, H. (2004), "What Really Happened to Consumption Inequality in the US?," NBER working paper 10338.
- [4] ATTANASIO, O., and DAVIS, S. (1996), "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy* 104, 1227-1262.
- [5] ATTANASIO, O., and WEBER, G. (1995), "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy* 103, 1121-1157.
- [6] BALDUZZI, P., and YAO, T. (2004), "Testing Heterogenous Agent Models: An Alternative Aggregation Approach," Boston College working paper.
- [7] BETSON, D. (1990), "Alternative Estimates of the Cost of Children from the 1980-1986 Consumer Expenditure Survey," US Department of Health and Human Services, Washington: D.C..
- [8] BOSWORTH, B., BURTLESS, B., and SABELHAUS, J. (1991), "The Decline in Saving: Evidence From Household Surveys," *Brooking Papers on Economic Activity*, 183-241.
- [9] BRAV, A., CONSTANTINIDES, G., and GECZY, C. (2002), "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy* 110, 793-824.
- [10] CITRO, C., and MICHAELS, R, eds. (1995), *Measuring Poverty: A New Approach (1995)*, National Academy Press, Washington: D.C.
- [11] COCHRANE, J. (2001), *Asset Pricing*, Princeton University Press, Princeton, NJ.

- [12] COGLEY, T. (2002), "Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey," *Journal of Monetary Economics* 49, 309-334.
- [13] GOLOSOV, M., KOCHERLAKOTA, N., and TSYVINSKI, A. (2003), "Optimal Indirect and Capital Taxation," *Review of Economic Studies* 70, 569-587.
- [14] HANSEN, L., and JAGANNATHAN, R. (1997), "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance* 52, 557-590.
- [15] HANSEN, L., and HODRICK, R.J. (1982), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88(5), 829-53.
- [16] KOCHERLAKOTA, N. (1996), "The Equity Premium: It's Still A Puzzle," *Journal of Economic Literature* 34, 42-71.
- [17] KOCHERLAKOTA, N. (2005), "Zero Expected Wealth Taxes: A Mirrlees Approach to Optimal Dynamic Taxation," *Econometrica* 73, 1587-1622.
- [18] KOCHERLAKOTA, N., and L. PISTAFERRI (2007), "Household heterogeneity and real exchange rates," *Economic Journal* 117, C1-25.
- [19] LIGON, E. (1996), "Risk-Sharing and Information: Theory and Measurement in Village Economies," unpublished working paper, UC-Berkeley.
- [20] LIGON, E. (1998), "Risk-Sharing and Information in Village Economies," *Review of Economic Studies* 65, 847-864.
- [21] LIGON, E. (2005), "Formal Markets and Informal Insurance," forthcoming, *International Review of Law and Economics*.
- [22] MANKIW, G., and S. ZELDES (1991), "The Consumption of Stockholders and Non-Stockholders," *Journal of Financial Economics* 29, 97-112.
- [23] MEGHIR, C., and PISTAFERRI, L., "Income Variance Dynamics and Heterogeneity," *Econometrica* 72, 1-32.

- [24] MEHRA, R., and PRESCOTT, E., “The Equity Premium: A Puzzle,” *Journal of Monetary Economics* 15, 145-161.
- [25] ROGERSON, W. (1985), “Repeated Moral Hazard,” *Econometrica*, 53, 69-76.
- [26] SEMENOV, A. (2004), “High-Order Consumption Moments and Asset Pricing,” York University working paper.
- [27] STORESLETTEN, K., TELMER, C., and YARON, A. (2001), “How Important are Idiosyncratic Shocks? Evidence from Labor Supply,” *American Economic Review Papers and Proceedings* 91, 413-417.
- [28] VISSING-JORGENSEN, A. (2002), “Limited Asset Market Participation and the Elasticity of Intertemporal Substitution,” *Journal of Political Economy* 110, 825-853.

Appendix

1. The US CEX

Each annual tape of the CEX contains four groups of core files, MTAB, ITAB, FMLY, and MEMB. There are also many auxiliary files, which are not used here. The FMLY, MEMB, MTAB, and ITAB files are organized by the calendar quarter of the year in which the data are collected. The FMLY files contain household characteristics, income, and summary level expenditures; the MEMB files contain member characteristics and income data; the MTAB files contain expenditures organized on a monthly basis at the Universal Classification Code (UCC) level; and the ITAB files contain income data converted to a monthly time frame and assigned to UCCs.

There are five quarterly data sets for each of these files. For example, in the 1980 tape, there are files running from the first quarter of 1980 through the first quarter of 1981. This is for the purpose of allowing computation of calendar year statistics. In fact, the MTAB 1980:Q1 file, say, contains expenditure information as reported by households interviewed in the first quarter of 1980. Since households report data for the previous three months, the 1980:Q1 file, say, has expenditure data from October 1979 to February 1980. The 1981:Q1 file is included in the 1980 tape because it contains information that spans the last three months of 1980. With just two exceptions (discussed below), the Y:Q1 file contained in the Y-1 tape is identical to the Y:Q1 file contained in the Y tape. The FMLY file for a given quarter has one record per household. Similarly, the MEMB file for a given quarter has one record per household member.

The CEX has three "household tracking" problems, detailed as follows. In 1980-81 households that re-entered the survey after missing an interview were assigned a new ID. While this is probably a minor proportion of the whole sample, it is for all purposes not a problem in our context, given that we do not focus on the longitudinal aspect of the survey. In 1986 the CEX changed its sample design. The consequence of this is that the core 1986:Q1 files contained in the 1985 tape survey different households than the core 1986:Q1 files contained in the 1986 tape. Indeed, issuing of household IDs starts from scratch beginning with the 1986 tape. Again, this is not a problem for us. We use both "samples", and so end up with a size that is larger than usual. Another sample design change occurs in 1996, but in this case some of the households that are surveyed in the core 1996:Q1 files contained in the 1995 tape appear also in the core 1996:Q1 files of

the 1996 tape, although with the same ID. Of course, we eliminate the duplicates.

We use the MTAB files from 1980:Q1 throughout 2004:Q1 to create monthly expenditure records for each household ever surveyed in the CEX. Since households report data for at most four quarters, there are between 3 and 12 observations per household. We merge this information with household characteristics from the FMLY file. The file so compiled contains 1,848,339 observations (where, to reiterate, each observation is a household/month data point).

Our measure of nondurable consumption is as in Attanasio and Weber (1995), and it is the sum of the following items (in parenthesis the UCC codes):²⁵ *Food at home* (790220, 790230, 190904), *Food away from home* (190901-190903, 790410, 790430, 800700), *Alcohol* (200900, 790310, 790320, 790420), *Apparel and footwear* (360110-420120), *Clothing services* (440110-440140, 440210, 440900), tobacco (630110, 630210), heating (250111-250904), Utilities: gas (260211-260214), Utilities: electricity (260111-260114), *Utilities: Water and sewerage* (270211-270214, 270411-270414, 270901-270904), *Public transportation* (530110-530902), *Vehicle expenses* (520110-520907), *Gasoline and oil* (470111-470212), *Vehicle maintenance and repairs* (470220-490900), *Parking fees* (220901-220902), *Newspapers and magazines* (590110-590212), *Books* (590220-590230), *Club membership fees* (620110-620115), *Ticket admissions* (620121-620310), *Miscellaneous entertainment expenses* (610900, 620330-620926), *Home rent* (210110-210902, 800710, 350110), *Home insurance* (220111-220122), *Home maintenance and repairs* (230111-230902, 330511, 340914, 790600), *Telephone and cable* (270000-270104, 270310), *Babysitting* (340210-340212), *Domestic services* (340310-340420), *Other home services* (340510-340530, 340906, 340911-340912, 340915), *Personal care* (650110-650900), *Rentals* (340610-340905, 340907-340908, 440150).

Here is a description of our sample selection. As said, we start with 1,848,339 monthly observations. We drop 164,125 observations for which our measure of total nondurable consumption is missing or zero. We further drop 4,258 observations for households that report zero food spending (at home and away from home) during an entire interview (3-month period). Households interviewed in a certain month are supposed to report consumption data only for the previous three months. We drop 4,116 observations reporting data for the same month in which they are interviewed. We next eliminate 398,047 observations corresponding

²⁵Erich Battistin at IFS kindly provided assistance in replicating Attanasio and Weber's aggregation procedures.

to households that are classified as "incomplete income respondents" or report less than three months of data for a given interview. At this point, we aggregate monthly data in (overlapping) quarters, indexed by the last month in the quarter. The resulting sample has 425,931 quarterly observations, corresponding to 147,412 households. We drop 9,635 observations corresponding to households that jump interviews (i.e., exit and re-enter the survey), and 5,498 observations corresponding to households living in college dorms. We end up with a final sample of 410,798 quarterly observations, or 140,364 households. The average number of household per (overlapping) quarter is 1,412. The median is 1,331.

2. The UK FES

The FES is composed of various modules: The Household schedule A (which asks questions about demographics, housing, durables), the Income schedule B (which asks questions about employment, wages, transfers, assets), the Diary of expenditure schedule D (which asks questions about expenditure on goods and services), the Checking schedule K and the Checking and outcome schedule K+L (which ask a number of consistency check questions), the England, Wales and Northern Ireland-only schedule M (which asks questions -especially about publicly provided housing- pertinent to all countries but Scotland), and the Northern Ireland schedule N.I. (which asks questions - especially about publicly provided housing- only pertinent to Northern Ireland). Household expenditure on nondurable goods and services, described below, uses information from schedules A, B, and D.

The expenditure codes in the FES have changed several times between 1968 and 2004. In what follows, we report the expenditure codes and their allocation to the various consumption categories for a single representative year, 1990. For other years, the allocation is similar although it is based on different expenditure codes.²⁶ The definition of nondurable consumption and services includes the following expenditure categories:²⁷ *Food at home* (101-137, 139-195, 198-199), *Food away from home* (840-857, 138, 196-197), *Alcohol* (260-289), *Tobacco* (211-213), *Fuel, light and power* (the sum of Coal and solid fuel (240,242,A321), Electricity (225,255,175,B222,B178⁻), Gas (226,254,B170,B221,B173⁻), and Other fu-

²⁶We thank Orazio Attanasio and Andrew Leicester for providing information on the allocation procedure.

²⁷The acronym "A" before an expenditure code indicates that the data come from the *Household* schedule A. The acronym "B" before an expenditure code indicates that the data come from the *Income* schedule B. A superscript ⁻ on an expenditure code means that the value is subtracted from the total. Expenditure codes without acronyms come from the *Diary of Expenditures* schedule D.

els (258,B017,B027)), *Telephone communications* (227,752,B166), *Housing* (the sum of Rent, Rates and community charges, Repair and maintenance (B102,B104,B107,B108), Own repair (232-238), Mortgage payments (B130,B198-B200), and Home insurance (B60,B110)), *Domestic services* (the sum of Domestic help (780), Repairs (782,788), and Laundry services (790,791)), *Clothing and footwear* (301-349), *Private transports* (the sum of Maintenance (510,513,514,545,546,548,549), Gasoline and oil (538,539,542), Vehicle tax and insurance (512,B187,B188,B179⁻), and others (555,556)), *Public transports* (the sum of Rail fares (550,551,B216,B218,B220), Bus and coach fares (552,B217), Air travel (553), School travel (B158)), *Entertainment* (the sum of TV license (229,760,768,B181), TV and cable rental (B195,B253,B254), Fees and admissions to events (753,755,761,763,764,765,B162), Hotels (B441-B452), and Holidays (754-759)), and *Miscellaneous items* (the sum of Household consumables (437,623,648,742,746), Petcare (731,732), Postage (751), Fees and subscriptions (219,220,228,770,772,796,797,799,805,806,807,B168,B180,B273,B280-B283), Gardening (733,734), and Books, newspapers and magazines (721-723)).

We use the FES files to create weekly expenditure records for each household ever surveyed in the FES. There is a single record for each household. In other words, there is no panel component in the FES. The file so compiled (after excluding households from Northern Ireland, less than 2% of the sample, and those for which our measure of total nondurable consumption is missing or zero) contains 172,163 observations (approximately 7,000 per year). We create quarterly expenditures by multiplying weekly expenditure by 13.

3. The Italy SFB

The micro data for Italy are released by ISTAT (the Central Statistical Office) under a confidentiality agreement.²⁸

The data available before 1997 come in semi-aggregate form. In particular, we define non-durable expenditure as the sum of: *Food at home* (the sum of the ISTAT original aggregates `panecer`, `zuccafte`, `carne`, `pesce`, `oliigr`, `latforuo`, `patfruor`), *Beverages* (`bevande`), *Food away from home* (`pastifc`), *Tobacco* (`tabacco`), *Apparel and footwear* (`vestcalz`), *Clothing services* (`lavander`), *Home rent* (`abitaz`), *Utilities* (`combenel`), *Car insurance* (`assauto`), *Telephone and cable* (`telefono`, `gettoni`, `abbontv`), *Personal care* (`sapone`, `barbriere`), *Gasoline and oil* (`benzina`), *Newspapers and magazines* (`giornali`), *Books* (`libri`),

²⁸We thank Guglielmo Weber with providing us with the data and notes regarding the aggregation procedures.

Home maintenance and repairs (piante, deters, and the difference between mobiliar and the sum of mobili, lenz, pento, cucine, frigo, lavatr, lavast, deters, servdom, lavander), *Domestic services* (servdom, servricr), *Public transportation* (autobus), and *Miscellaneous entertainment expenses* (alberg, cancel, and the difference between spetistr and the sum of giornali, libri, tassesc, radiotv, foto, artsport, piante, giochi, servricr, abbontv).

Since 1997, with the revision of the survey, data are released to researchers with a level of details that matches that for CEX and FES. Thus, nondurable expenditure is defined as the sum of the following items (in parenthesis the ISTAT codes): *Food at home* (C_1101-C_1107, C_1701-C_1705, C_1801-C_1802, C_1201-C_1209, C_1297, C_1301-C_1304, C_1501-C_1505, C_1401-C_1406, C_1621-C_1627, C_1601-C_1609, C_1631, C_1806-C_1808), *Food away from home* (C_9801-C_9803), *Alcohol* (C_1803-C_1805), *Tobacco* (C_1901), *Apparel and footwear* (C_2101-C_2108, C_220-C_2203, C_2109, C_4503), *Clothing services* (C_2204, C_2110, C_4606), *Home rent* (C_3101-C_3102, C_3121-C_3122), *Other home services* (C_3407, C_3427, C_6305, C_4404, C_9497), *Home insurance* (C_3302, C_3322), *Heating* (C_3403-C_3406, C_3423-C_3426), *Utilities: gas* (C_3402, C_3422), *Utilities: electricity* (C_3401, C_3421), *Utilities: Water and sewerage* (C_3301, C_3321), *Home maintenance and repairs* (C_3201-C_3212, C_3221-C_3232, C_4108, C_4312-C_4313, C_7121, C_4202, C_7134), *Telephone and cable* (C_9301, C_9321, C_7205, C_9306-C_9307), *Babysitting* (C_4601), *Domestic services* (C_4602, C_4607), *Vehicle expenses* (C_6201, C_6207, C_6208), *Gasoline and oil* (C_6209-C_6210), *Vehicle maintenance and repairs* (C_6203-C_6205), *Parking fees* (C_6206, C_6211), *Public transportation* (C_6301-C_6304, C_6306-C_6307), *Ticket admissions* (C_7202, C_7204, C_7207-C_7208), *Miscellaneous entertainment expenses* (C_7201, C_7203, C_7113, C_9901-C_9903, C_9911-C_9913, C_9405, C_7116, C_7118-C_7119, C_7206), *Newspapers and magazines* (C_7301-C_7303), *Books* (C_7304), *Personal care* (C_4404, C_4603-C_4605, C_9101-C_9102, C_9104).

We drop households with total nondurable consumption that is missing or zero. The data are on a monthly basis. The final file contains 531,430 observations (approximately 30,000 per year). We create quarterly expenditures by multiplying monthly expenditure by 3. As explained in the text, we also drop observations for $SDF_{1997:1}^j$, $SDF_{1997:2}^j$, and $SDF_{1997:3}^j$ due to the 1997 sampling design change.

Table 1**Descriptive Statistics: Micro Data**

Panel A: Household data from the US CEX									
	1980	1983	1986	1989	1992	1995	1998	2001	2004
Age	46.17	46.10	46.44	47.25	47.70	48.11	48.50	48.62	49.07
Family size	2.72	2.61	2.62	2.60	2.55	2.58	2.47	2.57	2.54
# of kids	0.79	0.72	0.75	0.72	0.71	0.73	0.67	0.70	0.70
Proportion some college+	0.40	0.44	0.44	0.47	0.49	0.49	0.54	0.56	0.58
Quarterly income	8.11	8.95	9.34	10.00	9.62	9.72	10.30	11.20	11.53
Adult equiv. quarterly cons.	3.14	3.06	3.05	3.21	3.06	3.01	3.12	3.06	2.99
Household quarterly cons.	5.24	4.94	4.97	5.19	4.87	4.82	4.87	4.92	4.80
N	14,426	14,353	21,689	15,801	16,037	14,111	14,551	20,193	4,260

Panel B: Household data from the UK FES									
	1975	1978	1981	1984	1987	1990	1993	1996	1999
Age	50.15	50.17	50.33	50.63	50.01	50.50	50.20	50.04	50.37
Family size	2.80	2.71	2.72	2.62	2.52	2.47	2.47	2.45	2.36
# of kids	0.80	0.73	0.72	0.65	0.61	0.61	0.63	0.65	0.60
Proportion some college+									
Quarterly income	3.36	3.40	3.55	3.49	4.01	4.28	4.31	4.43	4.84
Adult equiv. quarterly cons.	1.81	1.82	1.88	1.94	2.19	2.42	2.42	2.38	2.57
Household quarterly cons.	2.93	2.88	2.99	3.04	3.32	3.61	3.59	3.52	3.70
N	7,058	6,872	7,395	6,909	7,265	6,907	6,844	6,375	6,508

Panel C: Household data from the Italy SFB

	1985	1987	1989	1991	1993	1995	1997	1999	2001
Age	51.56	51.95	52.34	52.58	53.18	53.77	54.07	54.75	55.42
Family size	2.99	2.95	2.91	2.89	2.83	2.80	2.86	2.81	2.73
# of kids	0.73	0.69	0.65	0.61	0.57	0.53	0.54	0.52	0.48
Proportion some college+	0.05	0.05	0.05	0.06	0.06	0.07	0.08	0.09	0.08
Quarterly income	10.56	11.25	11.89	12.74	11.87	11.86	8.45*	8.48*	8.68*
Adult equiv. quarterly cons.	4.81	5.21	5.51	5.87	5.45	5.60	6.09	6.11	6.28
Household quarterly cons.	8.26	8.86	9.35	9.92	9.05	9.22	10.11	10.00	10.10
N	32,701	34,758	33,624	32,148	34,273	34,443	22,361	20,929	23,918

Note: All monetary variables are deflated by the CPI (2000=100). For the US (UK) they are expressed in thousand dollars (pounds); for Italy they are expressed in million lira. The Adult equivalent quarterly consumption is also deseasonalized as described in the text. *denotes an imputation based on interval data.

Table 2
Descriptive Statistics: Time Series Data

Panel A: US									
	1980-82	1983-85	1986-88	1989-91	1992-94	1995-97	1998-00	2001-04	1980-04
r^b (mean)	0.87	1.23	0.62	0.62	0.14	0.65	0.60	0.05	0.59
r^b (st.dev.)	0.97	0.43	0.61	0.44	0.29	0.33	0.36	0.63	0.65
r^s (mean)	1.67	3.54	3.00	2.76	1.33	5.75	2.39	-0.13	2.52
r^s (st.dev.)	9.83	6.22	10.37	7.45	3.25	4.69	7.65	8.50	7.68
$corr(r^b, r^s)$	0.0044	0.0532	0.1712	0.7005	0.0676	0.2163	0.1249	0.1216	0.1894

Panel B: UK									
	1975-79	1980-82	1983-85	1986-88	1989-91	1992-94	1995-97	1998-99	1975-99
r^b (mean)	-1.04	0.70	1.28	1.31	1.46	1.05	0.76	0.96	0.67
r^b (st.dev.)	1.78	1.41	0.89	0.74	1.01	0.95	0.55	0.66	1.44
r^s (mean)	3.49	3.19	5.20	3.53	1.88	2.83	3.92	3.89	3.45
r^s (st.dev.)	14.83	7.57	6.28	12.01	8.77	7.80	3.46	8.56	9.73
$corr(r^b, r^s)$	-0.1416	0.1979	0.2758	0.1122	0.3167	0.2327	-0.0473	0.0736	0.0235

Panel C: Italy								
	1985-86	1987-88	1989-91	1992-94	1995-97	1998-00	2001-02	1985-02
r^b (mean)	1.65	1.44	1.58	1.78	1.31	0.46	0.28	1.23
r^b (st.dev.)	0.56	0.29	0.36	0.62	0.36	0.32	0.26	0.68
r^s (mean)	15.98	-2.89	-1.38	1.84	3.90	6.92	-6.46	2.49
r^s (st.dev.)	17.57	10.35	11.14	12.01	9.36	14.51	8.95	13.46
$corr(r^b, r^s)$	0.3144	-0.0624	0.3628	-0.2756	-0.2347	-0.0052	-0.5787	0.0523

Note: r^s and r^b are the return on stocks and Treasury bills, respectively. The table reports (overlapping) quarterly returns.

Table 3
GMM estimates of γ

	US			UK			Italy		
	PIPO	IM	RA	PIPO	IM	RA	PIPO	IM	RA
γ	5.3326	1.5246	53.2607	4.9231	2.1969	4.8564	4.5357	2.3236	10.5233
(s.e.)	(1.3270)	n.a.	(21.1679)	(1.2451)	n.a.	n.a.	(3.1786)	n.a.	n.a.
\bar{e}_t	-3×10^{-9}	0.0190	-3×10^{-11}	-7×10^{-12}	0.0245	0.0250	-2×10^{-9}	0.0086	0.0077

Note: In this table, we report the estimates and standard errors associated with estimating γ using the restriction that $e_t(\gamma)$ has expectation zero, where

$$e_t(\gamma) = SDF_t(\gamma) (R_t^s - R_t^b)$$

The row \bar{e}_t reports the sample mean of the pricing error at the estimated value of γ .

Table 4

Joint GMM estimate of γ

	PIPO	IM	RA
γ	4.9216	2.0491	7.5931
(s.e.)	(1.2211)	(12.2032)	(50.4559)
$\bar{e}_{US,t}$	0.0088	0.0191	0.0182
$\bar{e}_{UK,t}$	0.0002	0.0246	0.0252
$\bar{e}_{IT,t}$	-0.0068	0.0087	0.0079
χ^2 -stat. OID test	0.1883	12.4015	8.1308
(p-value)	91%	0%	2%
χ^2 -stat. equal γ	0.0599	0.0001	0.0013
(p-value)	97%	99%	99%

Note: In this table, we report the estimates and standard errors associated with estimating γ using the restriction that $\varepsilon_t(\gamma)$ has expectation zero, where $\varepsilon_t(\gamma) = \left(e_{US,t}(\gamma) \ e_{UK,t}(\gamma) \ e_{IT,t}(\gamma) \right)'$ and $e_t(\gamma) = SDF_t(\gamma) (R_t^s - R_t^b)$. The rows $\bar{e}_{j,t}$ reports the sample mean of the pricing error at the estimated value of γ in country j . The row " χ^2 -stat. OID test" reports the value of the χ^2 -statistic for the overidentifying restriction test (with 2 degrees of freedom). The row " χ^2 -stat. equal γ " reports the value of the χ^2 -statistic for the test of equal γ across countries (with 2 degrees of freedom).

Table 5

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	0.7870	1.4571	0.9218	1.0911	0.9885
(s.e.)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
\bar{e}_t	0.0163	0.0164	0.0165	0.0158	0.0167	0.0165	0.0169

Note: In this table, we report the estimates and standard errors associated with estimating γ using the restriction that $e_t(\gamma)$ has expectation zero, where

$$e_t(\gamma) = SDF_t(\gamma) (R_t^s - R_t^b)$$

The row \bar{e}_t reports the sample mean of the pricing error at the estimated value of γ . A denotes the amount held in stocks.

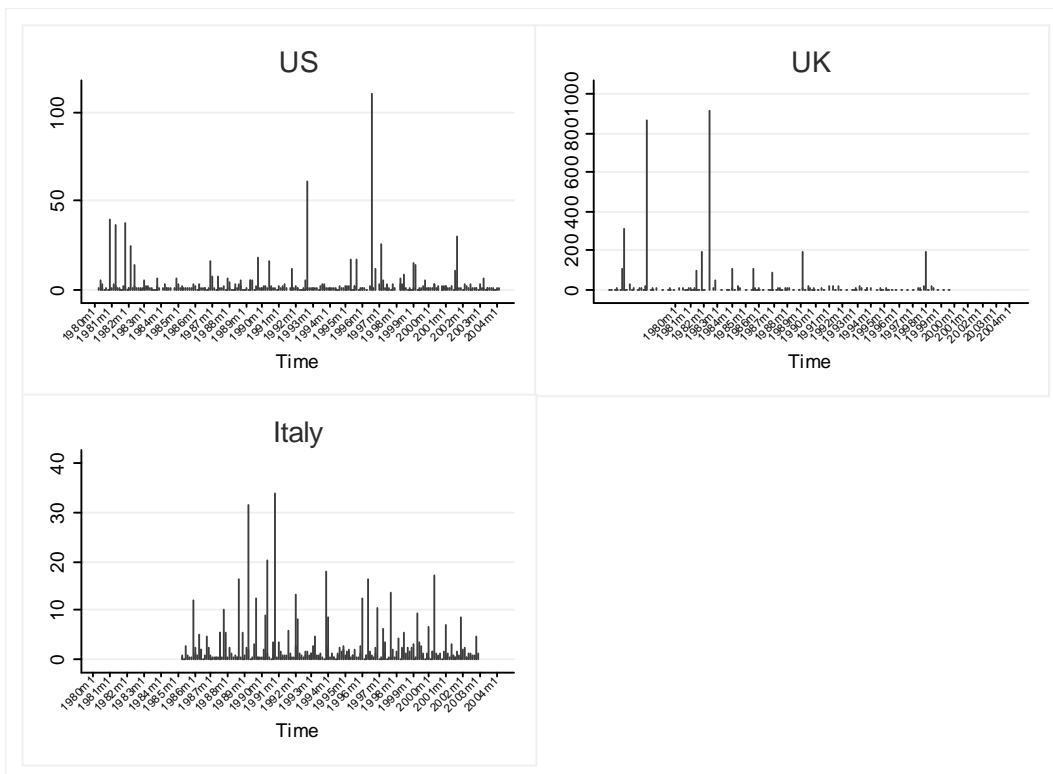


Figure 1: The PIPO stochastic discount factor (for $\gamma = 5$).

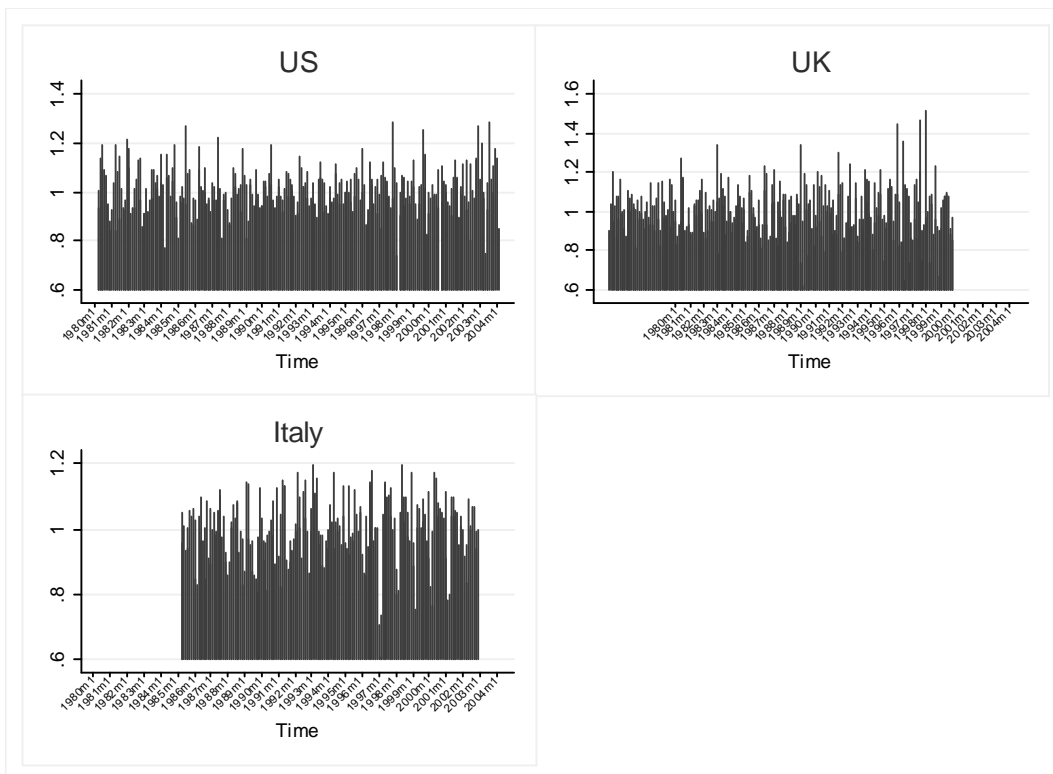


Figure 2: The Incomplete Markets stochastic discount factor (for $\gamma = 2$).