Using Common Features to Construct a Preference-Free Estimator of the Stochastic Discount Factor[1]

Fabio Araujo
Department of Economcis
Princeton University

Marcelo Fernandes
Graduate School of Economics – EPGE
Getulio Vargas Foundation

João Victor Issler[2]
Graduate School of Economics – EPGE
Getulio Vargas Foundation

February 13, 2004.

[1]This paper is a modified version of Fabio Araujo's M.A. dissertation – Araujo(2003), supervised by João Victor Issler. We are grateful for the comments and suggestions given by committee members by the time of defense, in particular to those given by Marco Bonomo and Walter Novaes. This paper also benefited from comments of the participants of the conference "Common Features in Maastricht," where it was presented. In particular we acknowledge the comments and suggestions of Robert F. Engle, Alain Hecq, Marco Lippi, and Simon Potter. The usual disclaimer applies.

[2]Corresponding author: Graduate School of Economics, Getulio Vargas Foundation, Praia de Botafogo 190 s. 1100, Rio de Janeiro, RJ 22253-900, Brazil.

Abstract

We propose a novel estimator for the stochastic discount factor (SDF) in a panel-data context. Under general conditions it depends exclusively on appropriate averages of asset returns, and its computation is a direct exercise, as long as one has enough observations to fit our asymptotic results. We identify the SDF using the fact that it is the "common feature" in every asset return of the economy. Moreover, it does not depend on any assumptions about preferences, or on consumption data, which allows testing directly different preference specifications, as well as the existence of the equity-premium puzzle. Preliminary results are encouraging.

1 Introduction

Despite their shortcomings, finance theories, such as the Capital Asset Pricing model (CAPM) – and its variants – and the Arbitrage-Pricing Theory (APT), have been work horses in finance and macroeconomics for a long time. Using mostly time-series (aggregate) data, early research has shown that these models failed to explain data regularities generating some important puzzles; see Hall (1978), Flavin (1981), Hansen and Singleton (1982, 1983, 1984), Mehra and Prescott (1985), Campbell (1987), Campbell and Deaton (1989), Epstein and Zin (1991), and Fama and French (1992, 1993). However, subsequent research in macro, using panel data, has convincingly shown that at least some of the early rejections were due to aggregation problems; see Runkle (1991), Attanasio and Browning (1995), and Attanasio and Weber (1995). Recent research confirmed that the gap between some of these models and the data narrows considerably once their cross-sectional dimension is taken into account. This has happened both in the finance and in the macro literatures: Lettau and Ludvigson (2001) propose "Resurrecting the Consumption-Based CAPM" (CCAPM) and Mulligan (2002) sustains that the problem with the low estimates of the intertemporal elasticity of substitution was the lack of proper cross-sectional aggregation for asset returns used in regressions, which has led him to use of the return on aggregate capital. Regarding the evolution of the literature over time, it seems that only considering the time-series dimension of the data may be the cause of the early rejections of these models.

In this paper, we propose a novel estimator for the stochastic discount factor (SDF), or pricing kernel, that exploits both the time-series and the cross-sectional dimensions of the data. Under very general conditions it depends exclusively on appropriate averages of asset returns. This makes its computation a simple and direct exercise, as long as one has enough time-series and cross-section observations to fit our asymptotic results. The identification strategy employed here to recover the SDF relies on one of its basic properties, following from the linearized version of the set of euler equations in the representative consumer's optimization problem – it is the "common feature," in the sense of Engle and Kozicki (1993), in every asset return of the economy. In the CCAPM context, this happens because the representative consumer must equate current asset prices to expected future payoffs appropriately discounted by the intertemporal marginal rates of substitutions. However, in the context of a representative consumer with a single good, the latter is the same for every asset in the economy.

Our SDF estimator does not depend on any assumptions about preferences, or on consumption data, being in this sense preference-free, which enables its use to test directly different preference specifications which are commonly used in the finance and in the macro literatures. Moreover, since our approach does not assume *a priori* that any type of finance theory is appropriate, it could be used more generally to test directly the implications of some of them. Because of the close relationship between the SDF and the *risk-free* rate, a consistent estimator for the latter can be based on a consistent estimator for the former, allowing the discussion of important issues in finance, such as the equity-premium and the risk-free rate puzzles.

Our approach is related to research done in three different fields. From econometrics, it is related to the common-features literature after Engle and Kozicki (1993) and to the latest addition to it in Engle and Marcucci (2003). It is also related to the spirit of the work on common factors of Geweke (1977), Stock and Watson (1989, 1993) and Forni et al. (2000). From finance, it is related to the stochastic-discount-factor approach initiated by Harrison and Kreps (1979) and by Hansen and Jagannathan (1991), who use observables to examine whether different preference specifications were admissible. It is also related to work that employs factor models within the CCAPM framework, perhaps best exemplified by Fama and French (1992, 1993) and the latest addition of Lettau and Ludvigson (2001). From macroeconomics, it is related to the work on aggregation bias in estimating the intertemporal elasticity of substitution in consumption by Attanasio and Weber (1995) and to the lack of proper cross-sectional aggregation of asset returns that motivates the work of Mulligan (2002).

The next Section presents basic theoretical results using the CCAPM and our estimation techniques, discussing first consistency and then efficiency in estimation. Section 3 shows how to use our estimator to evaluate the CCAPM, using formal and informal statistical methods. Section 4 presents a simple empirical illustration (not a full empirical application) of our theoretical results, and Section 5 concludes.

2 Economic Theory and SDF Estimation

2.1 A Simple Consistent Estimator

Harrison and Kreps (1979) and Hansen and Jagannathan (1991) describe a general framework to asset pricing that relies on the Pricing Equation[1]:

$$\mathsf{E}_t\left\{m_{t+1}x_{i,t+1}\right\} = p_{i,t}, \qquad i \in \{1,\ldots,N\}, \text{ or} \tag{1}$$

$$\mathsf{E}_t\left\{m_{t+1}R_{i,t+1}\right\} = 1, \qquad i \in \{1,\ldots,N\} \tag{2}$$

where $\mathsf{E}_t(\cdot)$ denotes the conditional expectation given the information available at time t, m_t is the stochastic discount factor, $p_{i,t}$ denotes the price of the i-th asset at time t, $x_{i,t+1}$ denotes the payoff of the i-th asset in $t+1$, $R_{i,t+1} = \frac{x_{i,t+1}}{p_{i,t}}$ denotes the gross return of the i-th asset in $t+1$, and N is the number of assets in the economy.

Equation (2) is the central pillar of our estimator and a basic assumption present in virtually all studies in finance and macroeconomics dealing with asset pricing and intertemporal substitution. It is important to stress that (2) entails very little theoretical structure: it is essentially equivalent to the "law of one price" – where securities with identical payoffs in all states of the world must have the same price. There are no assumptions about aggregation or the existence of complete markets. There is also no need to specify a preference representation for (2) to hold, or the presence of a representative consumer.

The existence of a SDF m_{t+1} that prices assets in (1) is obtained under very mild conditions. In particular, there is no need to assume a complete set of security markets. However, the discussion about uniqueness of m_{t+1} is more subtle. Under the assumption of complete markets, and a representative consumer, there will be a unique SDF m_{t+1} pricing all assets, which is an element of the payoff space. However, if markets are incomplete, i.e., if they do not span the entire set of contingencies, there will be an infinite number of stochastic discount factors m_{t+1} pricing all traded securities. Despite that, there will still exist a unique discount factor m^*_{t+1}, which is an element of the payoff space, pricing all traded securities. Moreover, any discount factor m_{t+1} can be decomposed as the sum of m^*_{t+1} and an error term orthogonal to payoffs, i.e., $m_{t+1} = m^*_{t+1} + \nu_{t+1}$, where $\mathsf{E}_t\left(\nu_{t+1}x_{i,t+1}\right) = 0$. The important fact here is that the pricing implications of any m_{t+1} are the same as those of m^*_{t+1}.

[1] See also Ross(1978), Rubinstein(1976) and Hansen and Richard(1987).

The discussion about existence an uniqueness of m_{t+1} is directly related to whether or not any estimation strategy can identify the SDF in an econometric sense. Of course, under the assumptions of complete markets, and a representative consumer, it will be possible for econometric techniques to identify the SDF in (2), which exists and is unique. This happens because econometric techniques usually deliver unique estimates. Under incomplete markets, however, where there is an infinite number of SDFs pricing assets in (1), this poses a problem for econometric techniques, which will only be able identify the SDF up to an error term, i.e., will be able to identify m^*_{t+1}.

Our strategy to derive the main results in this paper will be constructive. We start by using the very restrictive assumption that $m_{t+1}R_{i,t+1}$ is *log-Normal* and Homoskedastic, which is later relaxed in different directions. Surprisingly, the initial results obtained under these assumptions are later confirmed under much more general conditions. In order to make our final results applicable to a wide range of asset returns, we carefully match our final assumptions on their behavior to the stylized facts of these extensively investigated data.

It is well known that for any *log-Normal* random variable z_{t+1}:

$$\ln \mathsf{E}_t\left(z_{t+1}\right) = \mathsf{E}_t\left(\ln z_{t+1}\right) + \frac{1}{2}\mathsf{E}_t\left\{\left[\ln z_{t+1} - \mathsf{E}_t\left(\ln z_{t+1}\right)\right]^2\right\}. \tag{3}$$

Also, we can always decompose any random variable $\ln z_{t+1}$ as the sum of the space spanned by its conditional expectation and an unpredictable error term:

$$\ln z_{t+1} = \mathsf{E}_t\left(\ln z_{t+1}\right) + \varepsilon_{i,t+1}. \tag{4}$$

Using the two properties in (3) and (4), allows rewriting (2) as:

$$\ln R_{i,t+1} = -\ln m_{t+1} - \ln \chi^i_{m,t} + \varepsilon_{i,t+1,} \qquad i \in \{1, \ldots, N\} \tag{5}$$

where:

$$\begin{aligned}
\ln \chi^i_{m,t} &= \frac{1}{2}\left(\delta^2_{m,t} + 2\delta_{im,t} + \delta^2_{i,t}\right) \\
\delta^2_{m,t} &\equiv \mathsf{E}_t\left[\ln m_{t+1} - \mathsf{E}_t\left(\ln m_{t+1}\right)\right]^2 \\
\delta^2_{i,t} &\equiv \mathsf{E}_t\left[\ln R_{i,t+1} - \mathsf{E}_t\left(\ln R_{i,t+1}\right)\right]^2 \\
\delta_{im,t} &\equiv \mathsf{E}_t\left\{\left[\ln m_{t+1} - \mathsf{E}_t\left(\ln m_{t+1}\right)\right]\left[\ln R_{i,t+1} - \mathsf{E}_t\left(\ln R_{i,t+1}\right)\right]\right\} \\
\mathsf{E}_t\left(\ln m_{t+1} + \ln R_{i,t+1}\right) &= -\frac{1}{2}\left(\delta^2_{m,t} + 2\delta^i_{m,t} + \delta^2_{i,t}\right).
\end{aligned}$$

Notice that, by construction, $\mathsf{E}_t \varepsilon_{i,t+1} = 0$. As a consequence, $\mathsf{E}\varepsilon_{i,t+1} = 0$ as well, where $\mathsf{E}(\cdot)$ denotes the unconditional expectation operator. This implies that the cross-sectional distribution of $\varepsilon_{i,t+1}$ will also have a zero mean, which is a key ingredient to prove consistency of our estimator.

We now state our first basic result:

Proposition 1 *If the sequence $\{m_t R_{i,t}\}$ with $(i, t) \in \{1, ..., N\} \times \{1, ..., T\}$ is conditionally homoskedastic $\forall i$ and log-Normal $\forall t$, the SDF m_t can be consistently estimated for all t, as $N, T \to \infty$, at the same rate, using:*

$$\widehat{m}_t = \frac{\overline{R}_t^G}{\frac{1}{T}\sum_{t=1}^{T}\left(\overline{R}_t^G \overline{R}_t^A\right)},$$

where $\overline{R}_t^G = \prod_{i=1}^{N}\left[(R_{i,t})^{-\frac{1}{N}}\right]$ and $\overline{R}_t^A = \frac{1}{N}\sum_{i=1}^{N} R_{i,t}$ are respectively the geometric and arithmetic averages of all asset returns.

Proof. See Appendix. ■

Remark 1 *Under the assumptions of a complete set of security markets, and a representative consumer (innocuous), the SDF m_t is identified and a consistent estimator of it will be given by $\widehat{m}_t$ in Proposition 1. Under incomplete markets, we can only identify m_t^*, $m_t = m_t^* + \nu_t$, $\mathsf{E}_{t-1}(\nu_t x_{i,t}) = 0$, which has the same pricing implications of m_t. A consistent estimator of m_t^* will be given by $\widehat{m}_t$ in Proposition 1.*

There are two interesting features of $\widehat{m}_t$. First, it is a simple function of the geometric and the arithmetic average of asset returns, which makes its computation straightforward. Second, no more than (2), log-Normality, and Homoskedasticity of $\{m_t R_{i,t}\}$ were assumed in constructing it. In particular, no assumptions whatsoever about preferences were needed. In this sense, the estimator $\widehat{m}_t$ is model-free, and can be later used to test or validate different preference specifications in the same spirit of Hansen and Jagannathan (1991).

From (5) and (20) it becomes clear that $\ln m_t$ is the "common feature," in the sense of Engle and Kozicki (1993), of all asset returns: every term in it is indexed by i, with the exception of $\ln m_{t+1}$. It will generate all the "serial correlation common feature" in asset returns and, for any two assets i and j, $\ln R_{i,t+1} - \ln R_{j,t+1}$ will not have serial correlation, which makes $(1, -1)$ a "cofeature vector" for all asset pairs, which can be further exploited to construct alternative estimators of $\ln m_t$.

The assumption of conditional homoskedasticity plays a key role in the proof, since it implies that $\widehat{m}_t$ and m_t differ by a multiplicative constant χ_m, which otherwise would be time varying. At this stage, because of the overwhelming empirical evidence of heteroskedastic returns, e.g., Bollerslev, Engle and Nelson (1994), it seems natural to first relax this assumption. As discussed in our second basic result below, we may allow for as many heteroskedastic returns as one wants. Indeed, if the number of heteroskedastic returns is bounded by $N^{1-\delta}$, with $\delta > 0$, however small, our result in Proposition 1 is still unaltered.

Proposition 2 *If the sequence $\{m_t R_{i,t}\}$ with $(i, t) \in \{1, ..., N\} \times \{1, ..., T\}$ log-Normal $\forall t$, and Homoskedastic apart from a subset, whose number of elements is bounded by $N^{1-\delta}$, with $\delta > 0$, all with unconditional variance uniformly bounded in N, then the SDF m_t can be consistently estimated for all t, as $N, T \to \infty$, at the same rate, using the same expression in Proposition 1:*

$$\widehat{m}_t = \frac{\overline{R}_t^G}{\frac{1}{T}\sum_{t=1}^{T}\left(\overline{R}_t^G \overline{R}_t^A\right)}.$$

Proof. See Appendix. ■

A key element of the proof of Proposition 2 is that the proportion of assets with homoskedastic (heteroskedastic) returns is fixed as the number of assets grows to infinity. In principle, for large N, this requires having a large number of assets whose returns are homoskedastic, which may be a restrictive condition, especially if we consider high frequency data such as daily, weekly or even monthly observations.

An alternative to Proposition 2 is to work with time-aggregated data, where conditional heteroskedasticity fades away; see Drost and Nijman (1993) and Drost and Werker (1996). In the panel-data context above, time aggregation can be implemented considering an increasing large number of cross-section observations, while keeping fixed the number of time-series observations, with the time span S and the level of time aggregation h growing at the same rate, so as to keep the number of time-series observations constant. Hence, $N \to \infty$, $S \to \infty$ and $h = S/T$ with T fixed, however large.

Time-varying second moments are considered here by assuming that the error term $\varepsilon_{i,t}$ in (5) follows a discrete-time square-root stochastic autoregressive volatility (SR-SARV) process of order p with respect to an increasing filtration $J_{i,t} = \sigma(\varepsilon_{i,\tau}, F_\tau;\ \tau \le t)$, $t \in \mathsf{Z}$, as in Meddahi and Renault (2002). Hence, $\{\varepsilon_{i,t};\ t \in \mathsf{Z}\}$ is a stationary square-integrable process such that $\mathsf{E}\left(\varepsilon_{i,t+1} \mid J_{i,t}\right) = 0$ and the conditional variance process $f_{i,t+1|t} = \mathsf{V}\left(\varepsilon_{i,t+1} \mid J_{i,t}\right)$ is a marginalization of a stationary VAR(1) of dimension p:

$$
\begin{aligned}
f_{i,t+1|t} &= a_i' F_t > 0 \\
F_{t+1} &= \Lambda + \Gamma\, F_t + U_{t+1},
\end{aligned}
$$

where $\mathsf{E}\left(U_{t+1} \mid J_t\right) = 0$, $a_i \in \mathsf{R}^p$, $\Lambda \in \mathsf{R}^p$, and all the eigenvalues of Γ have modulus smaller than one.

Modelling $\varepsilon_{i,t}$ as a SR-SARV process is appealing for two reasons. First, discrete- and continuous-time versions of SR-SARV processes are consistent with each other, because the exact discretization of the continuous-time SR-SARV belongs to the class of discrete-time SR-SARV models. This is interesting since continuous-time models of asset returns play a major role in asset pricing. Second, the SR-SARV process encompasses many popular volatility models used in the financial econometric literature, e.g., the GARCH(1,1) and GARCH diffusion processes.

To consider how the level of temporal aggregation affects the conditional variance, we first establish some notation. High frequency observations are on $\varepsilon_{i,t}$ and the filtration reads $J_{i,t} = \sigma(\varepsilon_{i,\tau}, F_\tau;\ \tau \le t)$, with $t = 1, \ldots, T$. Because (log) returns and the SDF are flow variables, low frequency observations are on $\varepsilon^{(h)}_{i,th} \equiv \sum_{j=0}^{h-1} \varepsilon_{i,th-j}$. The filtration $J^{(h)}_{i,th}$ for the time-aggregated process is then $\sigma\left(\varepsilon^{(h)}_{i,\tau}, F_{\tau h};\ \tau \le t\right)$. The result below shows that the SR-SARV model is closed under time aggregation.

Proposition 3 *Let $\{\varepsilon_{i,t};\ t \in Z\}$ follow a SR-SARV(p) process with respect to the increasing filtration $J_{i,t} = \sigma(\varepsilon_{i,\tau}, F_\tau;\ \tau \le t)$, with conditional variance $f_{i,t+1|t} = e_i' F_t > 0$. For a given integer h, the process $\varepsilon^{(h)}_{i,th} \equiv \sum_{j=0}^{h-1} \varepsilon_{i,th-j}$ also follows a SR-SARV(p) process with respect to the increasing filtration $J^{(h)}_{i,th} = \sigma\left(\varepsilon^{(h)}_{i,\tau}, F_{\tau h};\ \tau \le t\right)$. In particular:*

$$
f^{(h)}_{i,th+h|th} = \mathsf{V}\left(\varepsilon^{(h)}_{i,th+h} \,\middle|\, J_{i,th}\right) = e_i' \left[A^{(h)} F_{th} + B^{(h)}\right],
$$

where $A^{(h)} = \sum_{j=0}^{h-1} \Gamma^{h-j-1}$ and $B^{(h)} = \sum_{j=0}^{h-1} \left(\sum_{k=0}^{h-j-2} \Gamma^k\right) \Lambda$. In the event that $e_i' A^{(h)} \neq 0$, then $f^{(h)}_{i,th} = e_i^{(h)\prime} F^{(h)}_{th}$ with $e_i^{(h)} = A^{(h)\prime} e_i$ and:

$$
F^{(h)}_{th} = F_{th} + e_i^{(h)} \left(e_i^{(h)\prime} e_i^{(h)}\right)^{-1} e_i' B^{(h)}.
$$

Moreover, $F_{th}^{(h)}$ *is a VAR(1) with an autoregressive matrix* $\Gamma^{(h)} = \Gamma^h$*.*

Proof. See Meddahi and Renault (2002). ■

The fact that the autoregressive matrix $\Gamma^{(h)}$ is exponential on the level h of time aggregation implies that the persistence increases exponentially with the frequency. Hence, conditional heteroskedasticity vanishes as the frequency decreases. Therefore, for a large enough number of assets N, time span S, and a high enough aggregation level h (low enough frequency of observations), the conditional homoskedasticity assumption may be justified, if not for all assets at least to subset of them, which shows that the assumptions in either Propositions 1 or 2 are feasible in this context.

Regardless of which of the two results are used – either Proposition 1 or 2 – the next proposition suggests that an estimate of the risk-free rate is straightforward once we have a consistent estimator of the SDF:

Proposition 4 *Using* $\widehat{m}_{t+1}$ *as in Propositions 1 or 2 above offers a consistent estimate of the risk-free rate,* R_{t+1}^f*:*

$$\widehat{R_{t+1}^f} = \frac{1}{\mathsf{E}_t\left(\widehat{m}_{t+1}\right)}.$$

Proof. See Appendix. ■

Although $\widehat{R_{t+1}^f}$ will be a consistent estimate of R_{t+1}^f, its computation is not as straightforward as that of $\widehat{m}_t$, since to implement it we need to compute the conditional expectation of $\widehat{m}_t$. For that we need an econometric model. It may be a very general non-parametric econometric model, for example, but we will need an econometric model nevertheless. Despite that, we still do not need any assumptions on preferences to compute $\widehat{R_{t+1}^f}$, or any finance theory, which shows that this estimator is "model free" in the sense given above.

A key assumption made above to obtain $\widehat{m}_t$ in Propositions 1 and 2 was that $\{m_t R_{i,t}\}$ with $(i,\, t) \in \{1, ..., N\} \times \{1, ..., T\}$ was log-Normal $\forall t$. This assumption was imposed because its is algebraic convenience, since log-Normal returns have a conditional mean that is a function of the first two moments of the associated Normal distribution. In general, $\mathsf{E}_t\{m_{t+1} R_{i,t+1}\}$ will also be a function of higher moments as well.

Relaxing log-Normality does not present a problem in the context above. The only difference is how we interpret $\ln \chi_{m,t}^i$ in (5). Under log-Normality, $\ln \chi_{m,t}^i$ will include only the second moments and cross-moments of $\ln m_{t+1}$ and of $\ln R_{i,t+1}$. Without log-Normality, it will include higher-order moments as well. Hence, we can regard (5) as a result of a full functional expansion on $m_{t+1}R_{i,t+1}$. As long as all of these higher-order moments are finite and time invariant, we are back to Proposition 1, and the only difference is that $\ln \chi_m^i$ now captures not just the effect of the variances and covariances but of the higher-order moments as well. Even if we allow for finite but time-varying higher-order moments, we can still use Proposition 2. Again, the only difference is how to interpret the first term of $\ln \chi_{m,t}$ in (21).

2.2 A Simple Efficient Estimator (Incomplete)

Taking logs of the both sides of the Pricing Equation (2), and further applying a Taylor expansion yields, for every $i \in \{1, \ldots, N\}$,

$$\mathsf{E}_t(\ln m_{t+1} + \ln R_{i,t+1}) + \frac{1}{2}\mathsf{V}_t(\ln m_{t+1} + \ln R_{i,t+1}) \simeq 0, \tag{6}$$

where $\mathsf{V}_t(\cdot)$ denotes the conditional variance given the available information at time t. Notice that (6) holds exactly only under log-Normality. However, the approximation error is negligible as long as the higher-order moments are time-invariant.

As in Proposition 3, we further assume that the sampling frequency is low enough, which allows writing,

$$\ln R_{i,t+1} + \ln m_{t+1} = \mathsf{E}_t(\ln R_{i,t+1} + \ln m_{t+1}) + \epsilon_{i,t+1}, \tag{7}$$

where $\epsilon_{i,t+1}$ has mean zero and a constant variance σ_i^2, which is the conditional variance $\mathsf{V}_t(\ln m_{t+1} + \ln R_{i,t+1})$.

Notice that, under log-Normality, $\epsilon_{i,t+1} \sim N(0, \sigma_i^2)$, which allows

From (6) and (7), it follows that:

$$\ln R_{i,t+1} + \ln m_{t+1} = -\frac{1}{2}\sigma_i^2 + \epsilon_{i,t+1}. \tag{8}$$

In the context of panel-data regression, (8) corresponds to a standard unobserved fixed-effects model with no explanatory variables other than time dummies. The coefficients of the time dummies then provide the estimates for the log of the SDF, whereas the fixed-effects capture the individual heterogeneity that stem from the variances of the log-returns.

Remark 2 *The approximation given by (6) is exact only under log-normality. If the log-returns display skewness, for instance, then the fixed effects are not necessarily negative. We therefore interpret nonnegative fixed effects as departures from log-normality.*

Averaging (8) over $t = 1, \ldots, T$ yields the cross-section equation

$$\frac{1}{T}\sum_{t=1} \ln R_{i,t} = -\frac{1}{T}\sum_{t=1} \ln m_t - \frac{1}{2}\sigma_i^2 + \frac{1}{T}\sum_{t=1} \epsilon_{i,t}. \tag{9}$$

Subtracting (9) from (8) gives way to the fixed-effects transformed equation

$$\begin{aligned}\ln R_{i,t+1} - \frac{1}{T}\sum_{t=1} \ln R_{i,t} &= -\left(\ln m_{t+1} - \frac{1}{T}\sum_{t=1} \ln m_t\right) + \epsilon_{i,t+1} - \frac{1}{T}\sum_{t=1} \epsilon_{i,t} \\ &= -\sum_{\tau=1} \ln m_\tau \left(\delta_{\tau,t} - \frac{1}{T}\right) + \epsilon_{i,t+1} - \frac{1}{T}\sum_{t=1} \epsilon_{i,t},\end{aligned}$$

where $\delta_{\tau,t}$ ($\tau = 1, \ldots, T$) denotes the indicator function that takes value one at time τ and zero otherwise.

Under the assumption that $\epsilon_{i,t+1} \sim N(0, \sigma_i^2)$, a fully-efficient estimate of the (log of the) SDF can be obtained by using pooled OLS, so as to retrieve the estimated series of the log-SDF by stacking the fixed-effects estimates of the coefficients of the time dummies. This will be equivalent to maximum likelihood. If normality is not assumed, this estimate will still be consistent but not fully-efficient.

2.3 Comparisons with the Literature

As far as we are aware of, studies in finance and macroeconomics dealing with the SDF do not try to obtain a direct estimate of it as we propose above. Usually, the SDF is estimated indirectly as a function of consumption data, through the use of a parametric function to represent preferences; see Hansen and Singleton (1982, 1984), Brown and Gibbons (1985) and Epstein and Zin (1991).

Hansen and Jagannathan (1991) avoid dealing with a direct estimate of the SDF, but note that the SDF has its behavior (in particular its variance) bounded by two restrictions. The first is the moment restriction:

$$\mathsf{E}_t\left\{m_{t+1}R_{i,t+1}\right\} = 1, \qquad i \in \{1, \ldots, N\}. \tag{10}$$

The second is the restriction that m is always positive, since, for a non-satiated representative consumer with utility function $U(\cdot)$, $U'(\cdot) > 0$, and discount factor β, $0 < \beta < 1$, $m_{t+1} = \beta\frac{U'(c_{t+1})}{U'(c_t)}$, where c_t is aggregate consumption.

Hansen and Jagannathan exploit the fact that it is always possible to project m onto the space of payoffs. Denoting by m^* the least-squares projection of m onto the space of payoffs, they show that it is directly related to the portfolio with the smallest second moment R^* by:

$$m^*_{t+1} = \frac{R^*_{t+1}}{\mathsf{E}_t\left\{\left(R^*_{t+1}\right)^2\right\}}. \tag{11}$$

It is then straightforward to express m^* only as a function of observables:

$$m^*_{t+1} = \iota' \left[\mathsf{E}_t\left(R_{t+1}R'_{t+1}\right)\right]^{-1} R_{t+1}, \tag{12}$$

where ι is a $N \times 1$ vector of ones.

Although they do not discuss it at any length, equation (12) shows that it is possible to identify m^*_{t+1} in the Hansen and Jagannathan framework. A similar expression for m_{t+1} could be obtained under the assumption of a complete set of security markets, when all the information on m will be contained on payoffs. In this case, (12) will identify m_{t+1}.

It is important to compare our results above with those of Hansen and Jagannathan. If one regards, (12) as a means to identify either m or m^*, which Hansen and Jagannathan did not discuss at all, it is apparent that both approaches face similar identification problems: they identify m_{t+1}only up to an error term, unless there is a full set of security markets. However, using (12) has important limitations that are not present in our approach. First, it is obvious from (12) that a conditional econometric model is needed to implement an estimate for m^*_{t+1}, since one has to compute the conditional moment $\mathsf{E}_t\left(R_{t+1}R'_{t+1}\right)$, something not present in our direct estimate. Second, while our method benefits from an increasing number of assets ($N \to \infty$), the use of (12) will suffer numerical problems in computing an estimate of $\left[\mathsf{E}_t\left(R_{t+1}R'_{t+1}\right)\right]^{-1}$. These problems will have two different sources: possible singularities in $\mathsf{E}_t\left(R_{t+1}R'_{t+1}\right)$, as the number of assets becomes large, and instability in inverting a high-dimensional matrix. Third, in the use of (12) one has to worry about how to impose the constraint that $m > 0$. Of course, our estimator in Proposition 1 and 2 faces an identical problem. However, this constraint can be imposed by employing instead the log-linearized version of (10), where we can use the fact that the exponential function has a positive range to get a positive estimate for m.

Although our approach exploits the panel data structure of asset returns in constructing an estimator for m_t, being in this sense disaggregate, it is related to the approach of Mulligan (2002), where return data is super-aggregated to compute the return to aggregate capital. For algebraic convenience, in illustrating the similarities between these two approaches, we use the log-utility assumption for preferences – where $m_{t+j} = \beta \frac{c_t}{c_{t+j}}$ – as well as the assumption of no production in the economy.

Since asset prices are the expected present-value of the dividend flows, and since with no production dividends are equal to consumption in every period, the price of the portfolio representing aggregate capital $\bar{p}_t$ is:

$$\bar{p}_t = \mathsf{E}_t \left(\sum_{i=1}^{\infty} \beta^i \frac{c_t}{c_{t+i}} c_{t+i} \right) = \frac{\beta}{1-\beta} c_t.$$

Hence, the return on aggregate capital $\overline{R}_{t+1}$ is given by:

$$\overline{R}_{t+1} = \frac{\bar{p}_{t+1} + c_{t+1}}{\bar{p}_t} = \frac{\beta c_{t+1} + (1-\beta) c_{t+1}}{\beta c_t} = \frac{c_{t+1}}{\beta c_t} = \frac{1}{m_{t+1}}, \tag{13}$$

which is the reciprocal of the SDF. Therefore there is a duality between the approach in Mulligan and ours' in the context above.

Taking logs of both sides of (13):

$$\ln \overline{R}_{t+1} = -\ln m_{t+1},$$

shows that the common feature in (5) is indeed the return of aggregate capital. Hence, we can decompose the return on every asset in the economy into an aggregate return and idiosyncratic terms:

$$\ln R_{i,t+1} = \ln \overline{R}_{t+1} - \ln \chi^i_{m,t} + \varepsilon_{i,t+1}, \qquad i \in \{1, \ldots, N\}. \tag{14}$$

Of course, it may not be so simple to derive this duality result under more general conditions but it can still be thought of as an approximation. Although similar in spirit, the work of Mulligan and ours' follow very different paths in empirical implementation: here our goal is to extract $-\ln m_{t+1}$ from a large data set of asset returns, whereas Mulligan uses national-account data to construct the return to aggregate capital. Because national-account data is prone to be measured with error, which will be increased as the level of aggregation increases, the approach taken by Mulligan may generate measurement error in the estimate of $\overline{R}_t$. However, our approach may avoid these problems for two reasons. First, we work with asset return data, which is more reliable than national-account data. Second, averaging returns in the way we propose factors out idiosyncratic measurement error in our estimate of m_t.

Factor models within the CCAPM framework have a long tradition in finance and in financial econometrics; see, for example, Fama and French (1993), Lettau and Ludvigson (2001), and Engle and Marcucci (2003). Fama and French and Lettau and Ludvigson propose respectively a three- and a two-factor model where, in the former, factors are related to firm size, book-to-market equity and the aggregate stock market, and, in the latter, with a time-varying risk premium and deviations to the long-run consumption-wealth ratio. In Engle and Marcucci, the focus is on common volatility for asset returns, where "common conditional variances" imply the existence of linear combinations of two (or perhaps more) heteroskedastic returns, $\ln R_{i,t+1} - \alpha \ln R_{j,t+1}$, such that the time-varying variance term drops out. Hence,

$$\ln \chi^i_{m,t} - \alpha \ln \chi^j_{m,t},$$

is not time-varying.

Compared to these papers our approach is to focus on only one factor (or feature) in (5), $-\ln m_{t+1}$ – the SDF. Although there may be higher-order factors coming from superior moments as in Fama and French, Lettau and Ludvigson, and Engle and Marcucci, we disregard their dynamic structure. Nevertheless, we do not disregard their effect on the mean of m_t, which are included in $\ln \chi^i_{m,t}$.

In our view, the important question is whether or not different finance theories are at odds with the data once we consider only the first order factor $-\ln m_{t+1}$, i.e., a parsimonious representation. The discussion in Cochrane (2001, ch. 7) shows that increasing the number of factors for the CCAPM does not necessarily generate a better model, since the risk of overfitting and of instability across different samples is always there. Our effort was to find a parsimonious factor model, where the factor has a straightforward macroeconomic explanation – it is the stochastic discount factor for all assets, or the return on aggregate capital.

3 Using our Estimator to Evaluate the CCAPM

3.1 Testing Preference Specifications within the CCAPM

An important question that can be addressed with our estimator of m_t is how to test and validate specific preference representations using it. Because $\widehat{m}_t$ is constructed from asset-return data alone, without any information about preferences whatsoever, it can be used to examine the appropriateness of different preference specifications.

First, even for consistent estimates, as is the case of $\widehat{m}_t$, we can always write:

$$\ln m_{t+1} = \ln \widehat{m}_{t+1} + \eta_{t+1}, \tag{15}$$

where η_{t+1} is the approximation error between $\ln m_{t+1}$ and its estimate $\ln \widehat{m}_{t+1}$. The properties of η_{t+1} will depend on the properties of m_{t+1} and $R_{i,t+1}$, and, in general, it will be serially dependent and heterogeneous.

Economic theory provides a number of preference representations that are frequently used in empirical studies. Table 1 below summarizes three of the most popular ones, where C_t denotes the external consumption, in the case of External Habit preferences, U_t represents the recursive utility function and B_t represents the optimal portfolio in the case of Kreps-Porteus preference representation.

Using (15) and the expressions for the (log) SDF in Table 1, we arrive at:

$$\ln \widehat{m}_{t+1} = \ln \beta - \gamma \Delta \ln c_{t+1} - \eta_{t+1}, \tag{16}$$

$$\ln \widehat{m}_{t+1} = \ln \beta - \gamma \Delta \ln c_{t+1} + \kappa \left(\gamma - 1\right) \Delta \ln c_t - \eta_{t+1}, \tag{17}$$

$$\ln \widehat{m}_{t+1} = \theta \ln \beta - \theta \gamma \Delta \ln c_{t+1} - \left(1 - \theta\right) \ln B_{t+1} - \eta_{t+1}, \tag{18}$$

which can be used to estimate the parameters of the CRRA, External Habit and Kreps-Porteus preference representations respectively, and to test these them using standard specification tests.

Perhaps the most appealing way of estimating (16), (17) and (18), simultaneously testing for over-identifying restrictions, is to use the generalized method of moments (GMM) proposed by Hansen (1982). Lagged values of regressors can be used as instruments in this case. Since (16) is nested into (17), we can also perform a redundancy test for $\Delta \ln c_t$ in (16).

3.2 Informal Testing

Writing in full the euler equation (5):

$$E_t \left[\beta \frac{U'(c_{t+1})}{U'(c_t)} R_{i,t+1} \right] = 1.$$

It is well known that it implies that assets with pro-cyclical (counter-cyclical) returns will pay a positive (negative) risk premium, since:

$$E_t\left\{R_{i,t+1}\right\} - R^f_{t+1} = -\frac{cov_t\left\{U'(c_{t+1}), R_{i,t+1}\right\}}{E_t\left\{U'(c_{t+1})\right\}}. \tag{19}$$

Therefore, on average, the sign of $R_{i,t+1} - R^f_{t+1}$ and of $\frac{c_{t+1}}{c_t} \cdot R_{i,t+1} - \overline{\frac{c_{t+1}}{c_t} \cdot R_{i,t+1}}$ must coincide, where the last term is demeaned. Of course, this is an informal test since signs can coincide but magnitudes can be very different. Nevertheless, this is an important instance where the CCAPM has failed in previous studies, and for that reason alone it is worth being investigated. Notice that we only need an estimate of the risk-free rate to implement such a test.

4 Empirical Illustration

4.1 Data

As a simple illustration of the approach described above – not a full empirical discussion – we apply our techniques to U.S. data, on quarterly frequency, from 1979:1 to 1998:4, all extracted from the DRI database. In computing $\widehat{m}_t$ using the results of Propositions 1 and 2, we employ 15 portfolio and individual returns: 6 of these portfolios returns are returns for international stock markets – Germany, Canada, France, UK, Italy and Japan; 6 are returns to domestic stock markets – the returns computed using the NYSE common price indices for finance, transportation and utility and the returns computed using the S&P common stock price indices for capital goods, composite and utilities. There is also the return on the 3-month T-Bill, the return on Gold and the national average contract mortgage rate. These returns cover a wide spectrum of alternative assets that are available to the typical U.S. household. Because some of these returns are portfolio returns, our final estimate of $\widehat{m}_t$ is ultimately an average of the returns of hundreds of different assets, which fits an important aspect of our asymptotic results.

In choosing data frequency, we preferred using quarterly data to reduce the proportion of heteroskedastic returns on overall returns, since it is well known that time aggregation reduces conditional heteroskedasticity. Consumption data used to test preferences is seasonally adjusted real total private consumption per-capita, following early research by Campbell (1987), among others.

4.2 Results

Figure 1 below shows our estimate of the SDF $\hat{m}_t$ for the period 1979:1 to 1998:4. It is close to unity most of the time and bounded by the interval [0.8, 1.2]. Moreover, it shows little signs of heteroskedasticity, which is perhaps a consequence of our choice of frequency. When we project our estimate on lagged returns, in order to get a conditional model for it, this group of regressors turn out only to be marginally significant. The adjusted R^2 of the conditional model is only 5%; see Table 2.

In constructing the estimate of the risk-free rate R^f_{t+1} we take the reciprocal of the predicted value of the conditional model in Table 2, which is plotted in Figure 2. Next, using $\hat{R}^f_{t+1}$, we perform the informal sign test of the CCAPM using (19): surprisingly, it correctly predicts the sign of the risk premium for all assets at least 80% of the time; see results in Table 3.

It is worth reporting that the average risk-free return for the period 1979:1 to 1998:4 is 7.2% per year. Of course, this is much higher than the return of the T-Bill, which was used as the risk-free rate by Mehra and Prescott (1985) to compute the equity premium. Once our estimate of the risk-free rate is considered, the average equity premium of 6.1% a year, computed by Mehra and Prescott, is reduced dramatically, which may hint that there is no equity-premium puzzle. It seems that the problem there lied in thinking about the T-Bill as a risk-free asset. Despite the fact that the T-Bill is a relatively safe asset, what constitutes a risk-free asset is its measurability with respect to the current information set – a property which most people would agree the T-Bill does not have. Despite this shortcoming, it is still interesting how one would classify the T-Bill. We suggest that it can be thought of as a hedge against the extremely unlikely event that "all goes wrong" in the U.S. or World economy.

Next we investigate the behavior of different preference representations in Table 1 by means of GMM estimation of equations (16), (17) and (18). For each equation, we use as a basic instrument list the two lags of $\ln \hat{m}_{t+1}$, two lags of $\ln \frac{c_{t+1}}{c_t}$, and two lags of $\ln B_{t+1}$, which is further reduced in two or three elements to check the robustness of initial results. Varying instruments did not change initial results at all, therefore we present only median estimates in Table 4 in order to save space.

The first thing to notice in Table 4 is that there is no evidence of rejection in over-identifying-restriction tests. Moreover, this is true not only for median estimates but to all estimates we produced. Second, results for the CRRA and the External Habit specifications yielded sensible estimates for the discount rate and the risk-aversion coefficient: $\hat{\beta} = 0.98$ and $\hat{\gamma} \in [2.1,\ 2.5]$. Compared to the estimates in Hansen and Singleton (1982, 1984), the estimates for $\hat{\gamma}$ are closer to what can be expected *a priori*. Our estimates *resurrect* the CCAPM, even with CRRA utility – notice that they are ten times those in Hansen and Singleton, who provided formal statistical evidence against the CCAPM. Moreover, they are in line with the panel-data estimates in Runkle (1991), which are about 2.2, and are consistent with the confidence interval of the intertemporal elasticity estimates of Attanasio and Weber (1995). They are slightly higher than the estimates in Mulligan (2002) with time-series data, which are in the range [0.5, 1.7]. Third, estimates for the Kreps-Porteus specification cannot be labelled "sensible," especially in light of an estimate of β higher than unity.

If we compare the results for the CRRA specification with that of the External Habit specification, because lagged consumption growth is not significant in the latter, we would prefer the former. Hence, one quantitative and qualitative result that emerges is that we cannot rule out the CRRA specification for the CCAPM for U.S. data: not only it is not rejected in direct specification testing, but it also yielded more parsimonious and sensible results compared to reasonable alternatives proposed in the literature. These results are really new, because they were obtained with aggregate data, with which the CCAPM has had an extensive record of rejections.

5 Conclusions

In this paper, we propose a novel estimator for the stochastic discount factor (SDF), or pricing kernel, that exploits both the time-series and the cross-sectional dimensions of the data. It depends exclusively on appropriate averages of asset returns, which makes its computation a simple and direct exercise. The identification strategy employed to recover the SDF relies on one of its basic properties, following from the linearized version of the set of euler equations in the representative consumer's optimization problem – it is the "common feature," in the sense of Engle and Kozicki (1993), in every asset return of the economy.

Because our SDF estimator does not depend on any assumptions about preferences, or on consumption data, we are able to use it to test directly different preference specifications which are commonly used in the finance and in the macro literatures. We could also have tested directly alternative finance theories using it, but we did not. Our estimator offers an immediate estimate of the *risk-free* rate, allowing us to discuss important issues in finance, such as the equity-premium and the risk-free rate puzzles.

The techniques discussed above are applied to a small but representative data set of asset returns in the U.S. economy to illustrate the potential of applying them to a broader data set. Our estimate of the SDF $\widehat{m}_t$ is close to unity most of the time and bounded by the interval $[0.8, 1.2]$, showing little signs of heteroskedasticity. The estimate of the risk-free rate R^f_{t+1} performs well in informal and formal tests of the CCAPM: surprisingly, it correctly predicts the sign of the risk premium for all assets at least 80% of the time, and, using the CRRA specification, we cannot reject the model in standard over-identifying-restriction tests.

Estimates of the risk-aversion coefficient using $\widehat{m}_t$ are close to what can be expected *a priori* – about 2.2 – showing that we can *resurrect* the CCAPM, even with CRRA utility. Moreover, these estimates are in line with the estimates in Mulligan (2002) using time-series data, and to the panel-data estimates in Runkle (1991) and Attanasio and Weber (1995).

References

[1] Abel, A., 1990, "Asset Prices under Habit Formation and Catching Up with the Joneses,"*American Economic Review Papers and Proceedings,* 80, 38-42.

[2] Araujo, F., 2003, "Identificação do fator estocástico de descontos e algumas implicações sobre testes de modelos de consumo," M.A. dissertation, EPGE-FGV.

[3] Attanasio, O., and M. Browning, 1995, "Consumption over the Life Cycle and over the Business Cycle," *American Economic Review*, vol. 85(5), pp. 1118-1137.

[4] Attanasio, O., and G. Weber, 1995, "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, vol. 103(6), pp. 1121-1157.

[5] Bollerslev, Tim, Robert F. Engle and Daniel B. Nelson (1994), "ARCH Models," in Chapter 49 of *The Handbook of Econometrics*, Volume 4, Amsterdam: North-Holland.

[6] Brown, D.P. and Gibbons, M.R. (1985). "A Simple Econometric Approach for Utility-based Asset Pricing Models," *Journal of Finance*, 40(2): 359-81.

[7] Campbell, J., 1993, "Intertemporal Asset Pricing without Consumption Data," *American Economic Review,* 83, 487-512.

[8] Campbell, J., A. Lo and A. MacKinlay, 1997, *The Econometrics of Financial Markets.* Princeton: Princeton University Press.

[9] Campbell, J. and Deaton, A. (1989), "Why is Consumption so Smooth?". *Review of Economic Studies*, 56:357-374.

[10] Campbell, J. and Mankiw, G. (1989), "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence", *NBER Macroeconomics Annual*, 4:185-216.

[11] Campbell, J. (1987), "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis", *Econometrica*, 55(6):1249-73.

[12] Cochrane, J.H. (2001). *Asset Pricing,* Princeton and Oxford: Princeton University Press.

[13] Drost, F. C. and Nijman, T.E. (1993). "Temporal Aggregation of GARCH Processes". *Econometrica*, 61(4): 909-27.

[14] Drost, F.C. and Werker, B.J.M. (1996). "Closing the GARCH Gap: Continuous Time GARCH Modeling", *Journal of Econometrics*, 74(1): 31-57.

[15] Engle, R.F. and Kozicki, S. (1993). "Testing for Common Features", *Journal of Business and Economic Statistics*, 11(4): 369-80.

[16] Engle, R. and Marcucci, J. (2003), "Common Features: An Overview of Theory and Applications", forthcoming in the *Annals Issue of the Journal of Econometrics* on "Common Features."

[17] Epstein, L. and S. Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937-968.

[18] Epstein, L. and S. Zin, 1991, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation," *Journal of Political Economy*, 99, 263-286.

[19] Fama, E.F. and French, K.R. (1992)." The Cross-Section of Expected Stock Returns". *Journal of Finance*, 47(2): 427-65.

[20] Fama, E.F. and French, K.R. (1993). "Common Risk Factors in the Returns on Stock and Bonds". *Journal of Financial Economics*, 33(1): 3-56.

[21] Flavin, M. (1981), "The Adjustment of Consumption to Changing Expectations About Future Income". *Journal of Political Economy*, 89(5):974-1009.

[22] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), "The Generalized Dynamic Factor Model: Identification and Estimation", *Review of Economics and Statistics*, 2000, vol. 82, issue 4, pp. 540-554.

[23] Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time Series," Chapter 19 in D.J. Aigner and A.S. Goldberger (eds.) *Latent Variables in Socio-Economic Models*, Amsterdam: North Holland.

[24] Gourinchas, J.-O. and J. Parker(2002), "Consumption over the Lifecycle," forthcoming in *Econometrica.*

[25] Hall, R. (1978), "Stochastic Implications of the Life Cycle Permanent Income Hypothesis: Theory and Evidence". *Journal of Political Economy*, 86(6), 971-987.

[26] Hall, R., 1988, "Intertemporal Substitution in Consumption," *Journal of Political Economy,* 96, 221-273.

[27] Hall, R. and F. Mishkin, 1982, "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, vol. 50(2), pp. 461-481.

[28] Hansen, L., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

[29] Hansen, L. and K. Singleton, 1982,"Generalized Instrumental Variables Estimation of Nonlinear Expectations Models," *Econometrica*, 50(5), pp. 1269-1286.

[30] Hansen, L. and Singleton, K. (1983), "Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns". *Journal of Political Economy*, 91(2):249-265.

[31] Hansen, L. and K. Singleton, 1984, Erratum of the article "Generalized Instrumental Variables Estimation of Nonlinear Expectations Models," *Econometrica*, 52(1), pp. 267-268.

[32] Hansen, L. P. and Jagannathan, R. (1991), "Implications of Security Market Data for Models of Dynamic Economies", *Journal of Political Economy*, 99(2):225-262.

[33] Hansen, L.P. and Scott F. Richard, 1987, "The role of conditioning information in deducing testable restricions implied by dynamic asset pricing models," *Econometrica*, 55:587-613.

[34] Harrison, J.M. and Kreps, D.M. (1979), "Martingales and arbitrage in multiperiod securities markets". *Journal of Economic Theory*, 20, 381–405.

[35] Kreps, D. and E. Porteus, 1978, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica,* 46, 185-200.

[36] Lettau, M. and Ludvigson, S. (2001). "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," *Journal of Political Economy*, 109(6): 1238-87.

[37] Lucas, R., 1978, "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429-1445.

[38] Lusardi, A. (1996), "Permanent Income, Current Income and Consumption: Evidence from Two Panel Datasets", *Journal of Business and Economics Statistics*, 14(1):81-90.

[39] Meddahi, Nour and Eric Renault, 2002, "Temporal Aggregation of Volatility Models", *Journal of Econometrics*, forthcoming.

[40] Mehra, R., and E. Prescott, 1985, "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* 15, 145-161.

[41] Mulligan, C.B., 2002, "Capital, Interest, and Aggregate Intertemporal Substitution," NBER Working Paper # 9373, National Bureau of Economic Research.

[42] Ross, S.A. (1976), "The arbitrage pricing of capital asset pricing", *Journal of Economic Theory*, 13, pp. 341-360.

[43] Rubinstein, M. (1976), "The Valuation of Uncertain Income Streams and the Price of Options," *Bell Journal of Economics*, 7, 407-425.

[44] Runkle, D. (1991), "Liquidity Constraints and the Permanent Income Hypothesis: Evidence from Panel Data". *Journal of Monetary Economics*, 27(1):73-98.

[45] Stock, J. and Watson, M. (1989) "New Indexes of Leading and Coincident Economic Indicators", *NBER Macroeconomics Annual*, pp. 351-95.

[46] Stock, J. and Watson, M. (1993), Eds., "*New Research on Business Cycles, Indicators and Forecasting*" Chicago: University of Chicago Press, for NBER.

A Proofs of Propositions in Section 2

Proof of Proposition 1. Averaging across assets:

$$\begin{aligned} \frac{1}{N}\sum_{i=1}^{N}\ln R_{i,t+1} &= -\ln m_{t+1} - \frac{1}{N}\sum_{i=1}^{N}\ln\chi_m^i + \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t+1} \\ \ln\left(\prod_{i=1}^{N}(R_{i,t+1})^{\frac{1}{N}}\right) &= -\ln(\chi_m m_{t+1}) + \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t+1} \end{aligned} \tag{20}$$

where $\ln\chi_m \equiv \frac{1}{N}\sum_{i=1}^{N}\ln\chi_m^i$ and therefore $\chi_m \equiv \prod_{i=1}^{N}\left[(\chi_m^i)^{\frac{1}{N}}\right]$.

Because the cross-sectional distribution of $\varepsilon_{i,t+1}$ has a zero mean, $\frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t+1} \xrightarrow{p} 0$, as $N\to\infty$. Hence, a consistent estimator for $\widetilde{m}_t \equiv \chi_m m_t$, as $N\to\infty$ is:

$$\widehat{\widetilde{m}}_t = \prod_{i=1}^{N}\left[(R_{i,t})^{-\frac{1}{N}}\right].$$

The only difference between $\widetilde{m}_t$ and m_t is a multiplicative constant that can be estimated consistently as follows. Multiply the Euler equation by χ_m to get:

$$\chi_m = \mathsf{E}_t\left\{\widetilde{m}_{t+1}R_{i,t+1}\right\} \quad \forall i = 1,...,N$$

Take the unconditional expectation and average out across N:

$$\chi_m = \frac{1}{N} \sum_{i=1}^{N} E\left\{\widetilde{m}_{t+1} R_{i,t+1}\right\}$$

Under the assumptions above, a consistent estimate for χ_m as $T \to \infty$ is:

$$\begin{aligned}
\widehat{\chi}_m &= \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{T} \sum_{t=1}^{T} \widehat{\widetilde{m}}_t R_{i,t} \right) \\
&= \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{\widetilde{m}}_t \frac{1}{N} \sum_{i=1}^{N} R_{i,t} \right) \\
&= \frac{1}{T} \sum_{t=1}^{T} \left[\left(\prod_{i=1}^{N} \left[(R_{i,t})^{-\frac{1}{N}} \right] \right) \left(\frac{1}{N} \sum_{i=1}^{N} R_{i,t} \right) \right]
\end{aligned}$$

Hence, a consistent estimator for m_t, as $N, T \to \infty$ is:

$$\begin{aligned}
\widehat{m}_t &= \frac{\widehat{\widetilde{m}}_t}{\widehat{\chi}_m} \\
&= \frac{\overline{R}_t^G}{\frac{1}{T} \sum_{t=1}^{T} \left(\overline{R}_t^G \overline{R}_t^A \right)}.
\end{aligned}$$

■

Proof of Proposition 2. Since $\widehat{\widetilde{m}}_t = \overline{R}_t^G$ is still a consistent estimator of $\widetilde{m}_t \equiv \chi_{m,t} m_t$, it remains to verify that the estimator of $\chi_{m,t}$ is consistent.

Without loss of generality, suppose that the first k_N processes are heteroskedastic. Then:

$$\ln \chi_{m,t} = \frac{1}{N} \sum_{i=1}^{k_N} \ln \chi_{m,t}^i + \left(\frac{N - k_N}{N} \right) \frac{1}{N - k_N} \sum_{i=k_N+1}^{N} \ln \chi_m^i. \tag{21}$$

Because these first k_N processes have uniformly bounded variance, there exist $0 < \overline{M} < \infty$ such that $\left| \ln \chi_{m,t}^i \right| \leq \overline{M}\ \forall (i,t)$. Hence, the first term above vanishes:

$$\frac{1}{N} \sum_{i=1}^{k_N} \ln \chi_{m,t}^i \leq \frac{1}{N} \sum_{i=1}^{k_N} \overline{M} \to 0, \text{ as } N \to \infty,$$

because $k_N \leq N^{1-\delta}$, with $\delta > 0$.

For the behavior of the second term, note that:

$$\frac{N - N^{1-\delta}}{N} \leq \frac{N - k_N}{N} \leq 1,\ N \geq 1.$$

Taking limits, when $N \to \infty$:

$$\lim_{N\to\infty} \frac{N - N^{1-\delta}}{N} \leq \lim_{N\to\infty} \frac{N - k_N}{N} \leq 1$$

$$1 = \lim_{N\to\infty} \frac{N}{N}(1 - N^{-\delta}) \leq \lim_{N\to\infty} \frac{N - k_N}{N} \leq 1.$$

Hence,

$$\lim_{N\to\infty} \frac{N - k_N}{N} = 1,$$

and $\ln \chi_{m,t}$ has the same asymptotic behavior of $\frac{1}{N-k_N} \sum_{i=k_{N+1}} \ln \chi_m^i$.

Therefore, we look for an estimator which have the same asymptotic behavior of:

$$\frac{1}{N - k_N} \sum_{i=k_{N+1}} \ln \chi_m^i.$$

Notice that we only use now the conditional homoskedastic returns, since $i \geq k_{N+1}$ in the formula above. Hence the problem is the same as before, and a consistent estimator of $\chi_{m,t}$ is given by:

$$\check{\chi}_m \equiv \frac{1}{T} \sum_{t=1} \left[\left(\prod_{i=k_N+1} \left[(R_{i,t})^{-\frac{1}{N-k_N}} \right] \right) \left(\frac{1}{N - k_N} \sum_{i=k_N+1} R_{i,t} \right) \right].$$

It is easy to see that, using the same argument as before, $\widehat{\chi}_m$ has the same asymptotic behavior of that of $\check{\chi}_m$. Hence, $\widehat{\chi}_m$ estimates $\chi_{m,t}$ consistently, and the result in Proposition 1 is still valid. ■

Proof of Proposition 3. By definition, R^f_{t+1} is measurable with respect to the sigma field generated by the information set of the representative consumer. Hence:

$$1 = \mathsf{E}_t \left\{ m_{t+1} R^f_{t+1} \right\} = R^f_{t+1} \mathsf{E}_t \{m_{t+1}\},\ \text{or}$$

$$R^f_{t+1} = \frac{1}{\mathsf{E}_t \{m_{t+1}\}},$$

which offers an immediate consistent estimator for the risk-free rate R^f_{t+1}:

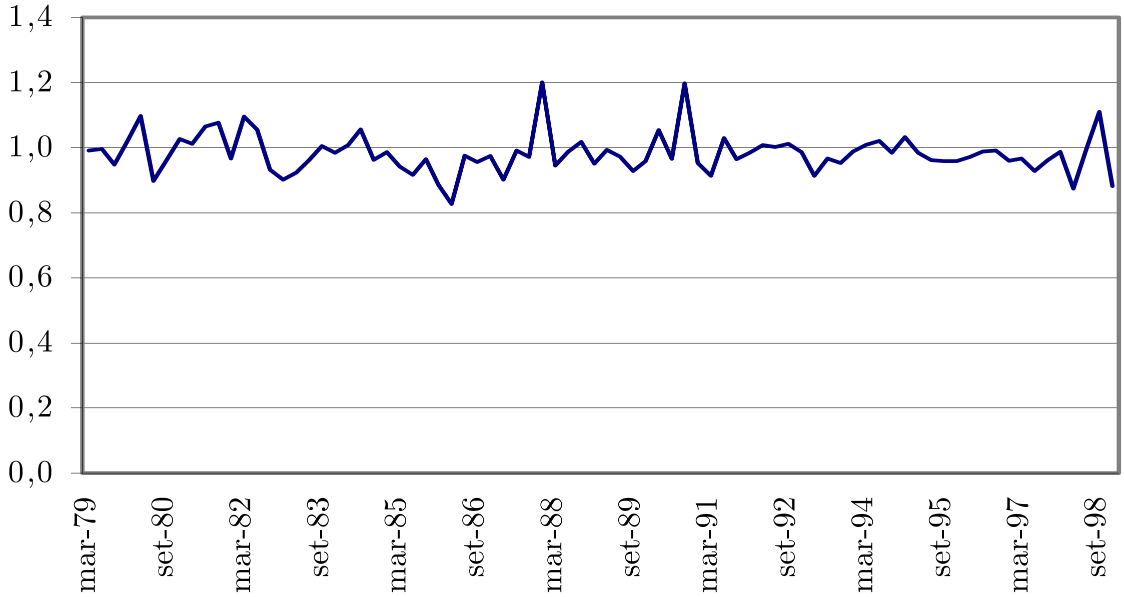
$$\widehat{R^f_{t+1}} = \frac{1}{\mathsf{E}_t \left\{ \widehat{m}_{t+1} \right\}}.$$

■

B Tables and Figures

Table 1: Preference Representations and Implied Stochastic Discount Factors

Preference Representation	Utility Function	SDF
CRRA	$U(c_t) = \frac{c_t^{1-\gamma}-1}{1-\gamma}$	$m_{t+1} = \beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}$
External Habit	$U(c_t) = \frac{\left(\frac{c_t}{C_{t-1}^{\kappa}}\right)^{1-\gamma}-1}{1-\gamma}$	$m_{t+1} = \beta\left(\frac{c_t}{c_{t-1}}\right)^{\kappa(\gamma-1)}\left(\frac{\gamma c_{t+1}}{c_t}\right)^{-\gamma}$
Kreps-Porteus	$U_t = \left[(1-\beta)c_t^{\rho} + \left(\mathrm{E}_t U_{t+1}^{\alpha}\right)^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}}$	$m_{t+1} = \left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right]^{\frac{\alpha}{\rho}}\left[\frac{1}{B_{t+1}}\right]^{1-\frac{\alpha}{\rho}}$



Stochastic Discount Factor -- USA

Figure 1

Table 2: Conditional Model for the SDF					
Adj.R^2	F-stat.	P-value	T	Serial Correlation at 5%?	ARCH Effect at 5%?
0.049	1.987	0.144	80	No	No

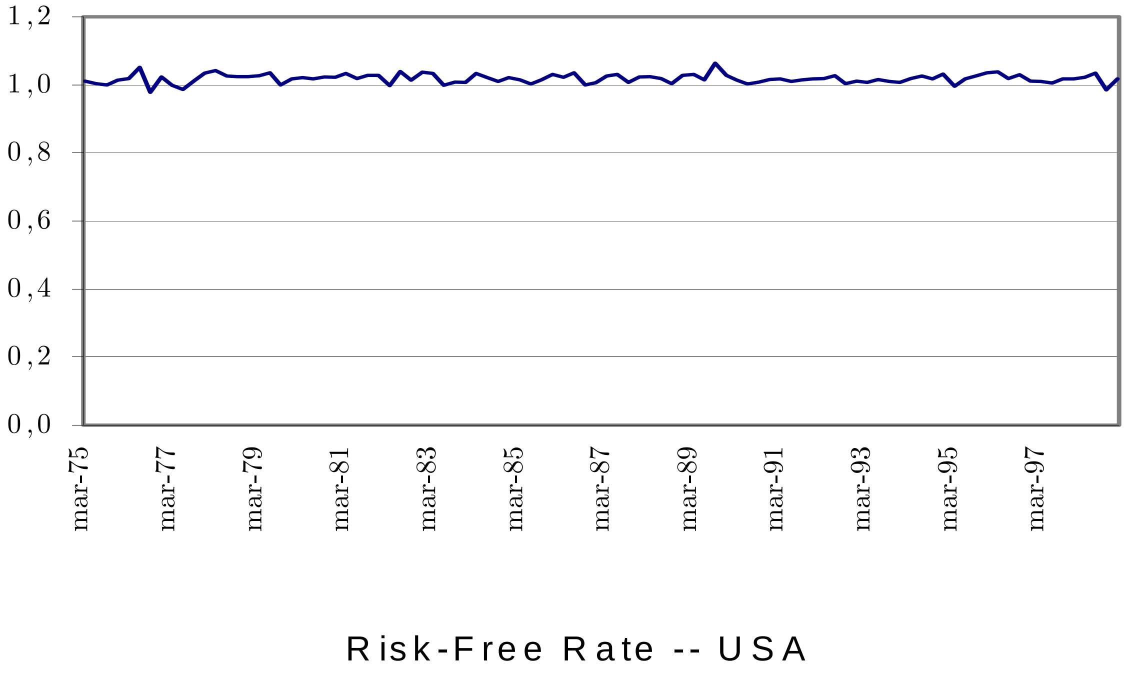


Figure 2

Agreement between the sign of the risk premium and the covariance between consumption growth and return

Positive	Negative	Undetermined
80.0%	6.7%	13.3%

Growth rates of consumption and risk-premia close to zero (+/- 0.1%) are considered undetermined

Table 3

Specification		CRRA	External Habit	Kreps-Porteus
		GMM	GMM	GMM
Constant	Coeff	-0.025	-0.026	0.002
	SE	0.007	0.005	0.023
	P-value	0.001	0.000	0.936
dlog(c)	Coeff	-2.087	-2.498	-0.596
	SE	1.009	0.605	0.619
	P-value	0.042	0.000	0.339
dlog(c(-1))	Coeff	-	0.125	-
	SE	-	0.791	-
	P-value	-	0.875	-
log(B)	Coeff	-	-	-0.902
	SE	-	-	0.800
	P-value	-	-	0.263
Serial-Correlation	LM test at 5%	No	No	No
ARCH LM test at 5%		No	No	No
Wald Test for	dlog(c(-1))	-	Yes	-
Redundancy 5%	log(B)	-	-	Yes
TJ test is Significant at 5%?		No	No	No
Number of Observations		77	77	77
	β	0.975	0.975	1.002
	SE	0.007	0.005	0.395
	γ	2.087	2.498	0.596
	SE	1.009	0.605	46.171
	κ	-	0.083	-
	SE	-	0.512	-
	θ	-	-	0.098
	SE	-	-	0.800

Median estimates for consumption regressor.

Table 4

B.1 Full Set of Estimates of Consumption Models

Table B.1.1: List of instruments in GMM estimation

Preferences		T	Instruments
CRRA	**1**	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) log(B(-1)) log(B(-2))
	2	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) log(B(-1))
	3	77	Cte log(m(-1)) dlog(c(-1)) dlog(c(-2)) log(B(-1))
Ext.Hab	**1**	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) dlog(c(-3)) log(B(-1)) log(B(-2))
	2	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) log(B(-1)) log(B(-2))
	3	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) dlog(c(-3)) log(B(-1))
	4	77	Cte log(m(-1)) dlog(c(-1)) dlog(c(-2)) dlog(c(-3)) log(B(-1))
K&P	**1**	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) log(B(-1)) log(B(-2))
	2	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) dlog(c(-2)) log(B(-1))
	3	77	Cte log(m(-1)) log(m(-2)) dlog(c(-1)) log(B(-1)) log(B(-2))
	4	77	Cte log(m(-1)) dlog(c(-1)) dlog(c(-2)) log(B(-1)) log(B(-2))

Table B.1.2: GMM estimates

Preferences		Cte	Prob	dlog(c)	Prob	dlog(c(-1))	Prob	log(B)	Prob	TJ-stat	Prob
CRRA	**1**	-0.024	0.000	-1.979	0.023	-	-	-	-	1.706	0.888
	2	-0.025	0.001	-2.087	0.042	-	-	-	-	1.667	0.797
	3	-0.025	0.001	-2.186	0.039	-	-	-	-	1.531	0.675
Ext Hab	**1**	-0.026	0.000	-2.498	0.000	0.125	0.875	-	-	1.959	0.855
	2	-0.024	0.000	-1.998	0.045	0.042	0.957	-	-	1.697	0.791
	3	-0.028	0.000	-2.962	0.005	0.531	0.596	-	-	1.690	0.793
	4	-0.027	0.000	-2.876	0.006	0.453	0.657	-	-	1.303	0.728
K&P	**1**	0.003	0.801	-0.602	0.312	-	-	-0.959	0.032	0.279	0.991
	2	0.005	0.703	-0.557	0.366	-	-	-1.048	0.039	0.014	1.000
	3	0.002	0.936	-0.596	0.339	-	-	-0.902	0.263	0.284	0.963
	4	0.006	0.701	-0.529	0.433	-	-	-1.074	0.064	0.110	0.991

C DRI Data Description

Table C.1: List of assets used to compute returns.

COMMODITIES PRICE:GOLD,LONDON NOON FIX,AVG OF DAILY RATE,$ PER OZ
INDEX RATE: NATIONAL AVERGE CONTRACT MORTGAGE RATE (%)
INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)
NYSE COMMON STOCK PRICE INDEX: FINANCE (12/31/65=50)
NYSE COMMON STOCK PRICE INDEX: TRANSPORTATION (12/31/65=50)
NYSE COMMON STOCK PRICE INDEX: UTILITY (12/31/65=50)
S&P S COMMON STOCK PRICE INDEX: CAPITAL GOODS (1941-43=10)
S&P S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
S&P S COMMON STOCK PRICE INDEX: UTILITIES (1941-43=10)
STOCK PRICE INDEX: CANADA
STOCK PRICE INDEX: FRENCE
STOCK PRICE INDEX: GERMANY
STOCK PRICE INDEX: ITALY
STOCK PRICE INDEX: JAPAN
STOCK PRICE INDEX: UNITED KINGDOM