Abstract

Asset prices contain information about the probability distribution of future states and the stochastic discounting of these states. Without additional assumptions, probabilities and stochastic discounting cannot be separately identified. To understand this identification challenge, we extract a positive martingale component from the stochastic discount factor process using Perron–Frobenius theory. When this martingale is degenerate, probabilities that govern investor beliefs are recovered from the prices of Arrow securities. When the martingale component is not trivial, using this same approach recovers a probability measure, but not the one that is used by investors. We refer to this outcome as “misspecified recovery.” We show that the resulting misspecified probability measure absorbs long-term risk adjustments. Many structural models of asset prices have stochastic discount factors with martingale components. Also empirical evidence on asset prices suggests that the recovered measure differs from the actual probability distribution. Even though this probability measure may fail to capture investor beliefs, we conclude that it is valuable as a tool for characterizing long-term risk pricing.

*We thank Fernando Alvarez, David Backus, Ravi Bansal, Anmol Bhandari, Peter Carr, Ing-Haw Cheng, Mikhail Chernov, Kyle Jurado, Stavros Panagakis, Martin Schneider, Kenneth Singleton, Johan Walden, Wei Xiong and the anonymous referees for useful comments.
1 Introduction

Asset prices are forward looking and encode information about investor beliefs. This leads researchers and policy makers to look at financial market data to gauge the views of the private sector about the future of the macroeconomy. It has been known, at least since the path-breaking work of Arrow, that asset prices reflect a combination of investor risk aversion and probability distributions used to assess risk. In this paper we ask what can be learned from the Arrow prices about investor beliefs in the presence of market determined stochastic discount factors that encode compensations for risk exposures. Data on asset prices alone are not sufficient to identify both the stochastic discount factor and transition probabilities without imposing additional restrictions. This additional information could be time series evidence on the evolution of the Markov state, or it could be information on the market determined stochastic discount factors.

The mathematical apparatus implied by Perron–Frobenius theory gives a way how to construct transition probabilities for Markov processes. Hansen and Scheinkman (2009) showed that application of this theory identifies a probability measure that reflects the long-term implications for risk pricing. They point out that while being substantively interesting, this measure is typically distinct from the physical probability measure. In contrast, Ross (2013) uses Perron–Frobenius theory to claim a recovery result — a full identification of the transition probabilities from asset prices. In this paper we connect the two results to make clear the special assumptions that are needed to guarantee that the transition probabilities recovered using Perron–Frobenius theory equal the actual transition probabilities.

In the general case, the ratio of the recovered to the true probability measure is manifested as a non-trivial martingale component in the stochastic discount factor process. Several of the structural models of asset pricing used in macroeconomics imply a non-trivial martingale component and existing empirical evidence suggests that this martingale component is quantitatively important.

Section 2 illustrates the challenge of identifying the correct probability measure from asset market data in a finite-state space environment. While the finite-state Markov environment is too constraining for many applications, the discussion in this section provides an overview of some of the main results in this paper. In particular, we show that:

- the Perron–Frobenius approach recovers a probability measure that reflects long-term pricing;
- the relative density of the Perron–Frobenius probability to the physical probability gives rise to a martingale component to the stochastic discount factor process;
- the stochastic discount factor process used by Ross (2013) implies that this martingale component is constant.

To place these results in a substantive context, we provide prototypical examples of asset pricing models in which this martingale is nontrivial.

In subsequent sections we establish these insights in greater generality, a generality rich enough to include many existing structural Markovian models of asset pricing. The framework for this
analysis, which encompasses discrete- and continuous-time models and allows for continuous state spaces, is introduced in Section 3. In Section 4, we pose the fundamental identification problem: data on asset prices can only identify the stochastic discount factor up to an arbitrary strictly positive martingale, and thus the physical probability measure remains unidentified without imposing additional restrictions or using additional data.

In Section 5, we extend the analysis of Ross (2013) to this more general setting. Provided we impose an additional ergodicity condition, the Perron–Frobenius approach identifies a unique probability measure. By connecting to the results in Hansen and Scheinkman (2009), we demonstrate in Section 6 that the Perron–Frobenius approach leads to the construction of a martingale component to the stochastic discount factor process, a component that must be identically equal to one for Ross (2013)’s assumption to hold.

In Section 7, we show the consequences of using the probability measure recovered by the use of the Perron–Frobenius theory when making inferences on the risk-return tradeoff. We show how the recovered probability measure absorbs long-term risk adjustments. Section 8 illustrates the impact of a martingale component to the stochastic discount factor in a workhorse asset pricing model that features long-run risk. These findings are aligned with a growing empirical literature that identifies the magnitude of the martingale component nonparametrically. In Section 9, we provide a unifying discussion of the empirical approaches that quantify the impact of the martingale components when an econometrician does not use the full array of Arrow prices. Section 10 concludes.

## 2 Illustrating the identification challenge

Given a complete set of one-period prices of Arrow securities, what can we learn about stochastic discount factors and the probabilities that reflect investor beliefs? Typically data on such asset prices alone cannot separately identify the one-period stochastic discount factor and the one-period transition probability. Following Hansen and Scheinkman (2014), in this section we find it revealing to use a simple finite-state Markov chain environment to illustrate the identification challenge.

There are multiple approaches for extracting probabilities from asset prices. For instance, risk neutral probabilities (e.g., see Ross (1978) and Harrison and Kreps (1979)) and closely related forward measures are frequently used in financial engineering. More recently, Perron–Frobenius Theory has been applied by Backus et al. (1989), Hansen and Scheinkman (2009) and Ross (2013) to the study of asset pricing with the latter two references featuring the construction of an associated probability measure. Hansen and Scheinkman (2009) and Ross (2013) have rather different interpretations of this measure, however. Ross (2013) identifies this measure with investors’ beliefs while Hansen and Scheinkman (2009) use it to characterize long-term contributions to risk pricing. Under this second interpretation, Perron–Frobenius Theory features an eigenvalue that dominates over long investment horizons and the resulting probability measure targets the valuation of assets that pay off in the far future as a point of reference. In this section we illustrate the constructions of the alternative probability using matrices associated with finite-state Markov
chains and we explore some simple example economies to understand better the construction of a probability measure based on Perron–Frobenius Theory.

Let $X$ be a discrete time, $n$-state Markov chain with transition matrix $P = [p_{ij}]$. We identify state $i$ with a coordinate vector $u^i$ with a single one in the $i$-th entry. Suppose the analyst infers the prices of one-period Arrow claims from data. We represent this input as a matrix $Q = [q_{ij}]$ that maps a payoff tomorrow specified as a function of tomorrow’s state into a price today. Since there are only a finite number of states, the payoff and price can both be represented as vectors. In particular, $q_{ij}$ is the price in state $X_t = u^i$ of a security that pays one unit of consumption in state $X_{t+1} = u^j$.

Notice that $Q$ has $n \times n$ entries. $P$ has $n \times (n - 1)$ free entries because row sums have to add up to one. The realized one-period stochastic discount factors are entries of a matrix $S = [s_{ij}]$ where $s_{ij}$ discounts one unit of consumption in state $u^j$ tomorrow given that the current state is $u^i$. The discounting is state-dependent to adjust valuation to uncertainty in the next-period payout. Risk adjustments are encoded by this state dependence because each future state may be discounted differently. In general the stochastic discount factor introduces $n \times n$ free parameters $s_{ij}$, $i, j = 1, \ldots, n$. The Arrow prices are the products:

$$q_{ij} = s_{ij}p_{ij}. \quad (1)$$

Suppose that we specify probabilities $p_{ij}$ subject to three restrictions. First they are nonnegative; second $p_{ij} = 0$ if, and only if, $q_{ij} = 0$; and third the row sums of the resulting matrix are all equal to one. Then $P = [p_{ij}]$ is a valid transition probability matrix; however, there are typically many specifications of such transition probabilities that are consistent with the Arrow prices. For each specification, we can construct stochastic discount factors depicted by $s_{ij}$’s that satisfy (1).

Additional data is needed to identify $P = [p_{ij}]$ and $S = [s_{ij}]$ from asset prices $Q = [q_{ij}]$ without imposing additional restrictions. For instance under the assumption of rational expectations, time series evidence reveals the evolution of the Markov state as captured by the transition probabilities. This evidence gives a separate way to identify $P$, and then Arrow prices reveal the state-specific discount factors $S$.

Alternatively, equation (1) reveals transition probabilities when we impose restrictions on the stochastic discount factors. Such constructions provide valuable tools for asset pricing even when these probabilities are not necessarily the same as those used by investors. In what follows we consider two alternative restrictions:

i) **Risk-neutral pricing:**

$$s^*_{ij} = \tilde{q}_i$$

for positive numbers $\tilde{q}_i$, $i = 1, 2, \ldots, n$. This restriction exploits the pricing of one-period discount bonds.

---

\[1\]The simple counting requires some qualification when $Q$ has zeros. For instance, when $q_{ij} = 0$, then $p_{ij} = 0$ in order to prevent arbitrage opportunities. In this case the numerical value assigned to $s_{ij}$ is inconsequential.
ii) *Long-term pricing:*

\[ \tilde{s}_{ij} = \exp(\eta) \frac{m_j}{m_i} \]  

(2)

for positive numbers \( m_i, i = 1, 2, ..., n \) and a real number \( \eta \) that is typically negative. The \( m_i \)'s need only be specified up to a scale factor and the resulting vector can be normalized conveniently. As we will show later, this restriction exploits the pricing of long-term discount bonds.

In both cases we reduce the number of free parameters in the matrix \( S \) from \( n^2 \) to \( n \), making identification of the probabilities possible. As we will see, each approach has an explicit economic interpretation, but the matrices of transition probabilities that are recovered do not necessarily coincide with those used by investors or with the actual Markov state dynamics. In the first case the difference between the inferred and true probabilities reflects a martingale that determines the one-period risk adjustments in financial returns. As we show below, in the second case the difference between the inferred and true probabilities reflects a martingale that determines long-term risk adjustments in pricing stochastically growing cash flows.

### 2.1 Risk-neutral probabilities

Risk-neutral probabilities are used extensively in the financial engineering literature. These probabilities are a theoretical construct used to absorb the local or one-period risk adjustments and are determined by positing a fictitious “risk-neutral” investor. The resulting stochastic discount factor of this investor is therefore given by

\[ s^*_i,j = \bar{q}_i, \]

reflecting the fact that all states \( j \) tomorrow are discounted equally. In order to satisfy pricing restrictions (1), the risk-neutral transition probabilities must be given by

\[ p^*_ij = \frac{q_{ij}}{\bar{q}_i}. \]

Since rows of a probability matrix have to sum up to one, it necessarily follows that \( \bar{q}_i = \sum_j q_{ij} \), which can be interpreted as the price of a one-period discount bond in state \( i \).

The risk-neutral probabilities \( [p^*_ij] \) can always be constructed and used in conjunction with discount factors \( [s^*_ij] \). By design the discount factors are independent of state \( j \), reflecting the absence of risk adjustments conditioned on the current state. In contrast, one-period discount bond prices can still be state-dependent and dependence is absorbed into the subjective discount factor of fictitious risk neutral investor. While it is widely recognized that the risk-neutral distribution is distinct from the actual probability distribution, some have argued that the risk-neutral dynamics remain interesting for macroeconomic forecasting precisely because they do embed risk adjustments.\(^2\)

\(^2\)Narayana Kocherlakota, President of the Federal Reserve Bank of Minneapolis, during a speech to the Mitsui Financial Symposium in 2012 asks and answers: “How can policymakers formulate the needed outlook for marginal net benefits? ... I argue that policymakers can achieve better outcomes by basing their outlooks on risk-neutral
When short-term interest rates are state dependent, forward measures are sometimes used in valuation. Write
\[(Q)^t = [q^t_{ij}]\]
where the \(q^t_{ij}\) are the Arrow prices for a \(t\)-period investment horizon (the \(t\) on \(q^t_{ij}\) is a superscript). The \(t\)-period forward probability measure given the current state \(u^i\) is
\[P^*_t = \left[ \frac{q^t_{ij}}{\sum_{j=1}^{n} q^t_{ij}} \right].\]
The denominator used for scaling the Arrow prices is now the price of a \(t\)-period discount bond. While the forward measure is of direct interest,
\[P^*_t \neq (P^*)^t, \quad (3)\]
when one-period bond prices are state dependent. Variation in one-period interest rates contribute to risk adjustment over longer investment horizons, and as a consequence the construction of risk-neutral probabilities is horizon-dependent.

### 2.2 Long-term pricing

We follow Backus et al. (1989) in studying long-term pricing of cash flows associated with fixed income securities using Perron–Frobenius theory. When there exists a \(\lambda > 0\) such that the matrix \(\sum_{t=0}^{\infty} \lambda^t Q^t\) has all entries that are strictly positive, the largest eigenvalue \(\exp(\tilde{\eta})\) of \(Q\) is unique and positive and the associated right eigenvector \(\tilde{e}\) has strictly positive entries. We denote the \(i^{th}\) entry of \(\tilde{e}\) as \(\tilde{e}_i\). Typically, \(\tilde{\eta} < 0\) to reflect time discounting of future payoffs over long investment horizons.

Recall that we may evaluate \(t\)-period claims by applying the matrix \(Q\) for \(t\) times in succession. From the Perron–Frobenius theory for positive matrices:
\[
\lim_{t \to \infty} \exp(-\tilde{\eta} t) Q^t f = (f \cdot \tilde{e}^*) \tilde{e}
\]
where \(\tilde{e}^*\) is the corresponding positive left eigenvector of \(Q\). Applying this formula, the large \(t\) approximation to the annualized state-dependent return for holding a long-term security payable in state \(j\) in \(t\) periods given current state \(i\) is
\[
-\tilde{\eta} - \frac{1}{t} \log(\tilde{e}^*_j) = \frac{1}{t} \log \tilde{e}_i.
\]
Thus the average rate of discount on this long-horizon cash flow is \(-\tilde{\eta}\) and the eigenvector entries provide a \(1/t\) refinement for large \(t\).

The eigenvector $\tilde{e}$ and the associated eigenvalue also provide a way to construct a probability transition matrix given $Q$. Form

$$\tilde{p}_{ij} = \exp(-\tilde{\eta})q_{ij}\tilde{e}_j/\tilde{e}_i. \tag{4}$$

Notice that

$$\sum_{j=1}^{n} \tilde{p}_{ij} = \exp(-\tilde{\eta}) \frac{1}{\tilde{e}_i} \sum_{j=1}^{n} q_{ij}\tilde{e}_j = 1.$$ 

Thus $\tilde{P} = [\tilde{p}_{ij}]$ is a transition matrix. Moreover,

$$q_{ij} = \exp(\tilde{\eta})\tilde{e}_i/\tilde{e}_j \tilde{p}_{ij}.$$ 

Thus we have used the eigenvector $\tilde{e}$ and the eigenvalue $\tilde{\eta}$ to construct a stochastic discount factor that satisfies (2) together with a probability measure that satisfies (1). The probability measure constructed in this fashion absorbs the compensations for exposure to long-term components of risk. Conversely, if one starts with an $\tilde{S}$ that satisfies both (1) and (2) then it is straightforward that the vector with entries $\tilde{e}_i = 1/m_i$ is an eigenvector of $Q$.\footnote{To see this, notice that $\tilde{s}_{ij}(1/m_j) = \exp(\eta)(1/m_i)$. The implied probabilities are given by $\tilde{p}_{ij} = q_{ij}/\tilde{s}_{ij}$. Pre-multiplying by the probabilities $\tilde{p}_{ij}$, summing over $j$, and stacking into the vector form, we obtain $Q\tilde{e} = \exp(\tilde{\eta})\tilde{e}$ for a vector $\tilde{e}$ with entries $\tilde{e}_i = 1/m_i.$}

If we start with the $t$-period Arrow prices in the matrix $Q^t = (Q)^t$ instead of the one-period Arrow prices in the matrix $Q$, then

$$Q^t\tilde{e} = \exp(\tilde{\eta}t)\tilde{e}$$

for the same vector $\tilde{e}$ and $\tilde{\eta}$. The implied matrix $\tilde{P}^t$ constructed from $Q^t$ satisfies:

$$\tilde{P}^t = (\tilde{P})^t$$

in contrast to the corresponding result (3) for the risk-neutral probabilities.

Hansen and Scheinkman (2009) and Ross (2013) both use this approach to construct a probability distribution, but they interpret it differently. Hansen and Scheinkman (2009) study multi-period pricing by compounding stochastic discount factors. They use the probability ratios for $\tilde{p}_{ij}$ given by (4) and consider the following decomposition:

$$q_{ij} = \left[ \exp(\tilde{\eta})\frac{\tilde{e}_i}{\tilde{e}_j} \tilde{p}_{ij} \right] p_{ij} = \exp(\tilde{\eta}) \left( \frac{\tilde{e}_i}{\tilde{e}_j} \right) h_{ij} p_{ij}.$$ 

Hence,

$$s_{ij} = \exp(\tilde{\eta}) \left( \frac{\tilde{e}_i}{\tilde{e}_j} \right) h_{ij} \tag{5}$$

where

$$h_{ij} = \tilde{p}_{ij}/p_{ij}.$$
provided that \( p_{ij} > 0 \). When \( p_{ij} = 0 \) the construction of \( h_{ij} \) is inconsequential.

Notice that \( h_{ij} \) cannot be written as \( h_{ij} = \exp(\tilde{\eta})\hat{e}_i/\hat{e}_j \) for some number \( \tilde{\eta} \) and a vector \( \tilde{e} \) with positive entries, and thus the decomposition in (5) is unique. For if \( h_{ij} = \exp(\tilde{\eta})\hat{e}_i/\hat{e}_j \) for some number \( \tilde{\eta} \) and a vector \( \tilde{e} \) with positive entries, there would exist another Perron–Frobenius eigenvector for \( Q \) with entries given by \( \hat{e}_i\hat{e}_j \) and an eigenvalue \( \exp(\tilde{\eta} + \hat{\eta}) \). The Perron–Frobenius Theorem guarantees that there is only one eigenvector with strictly positive entries (up to scale) implying that \( \tilde{e} \) must be a vector of constants and \( \hat{\eta} = 0 \). In particular we have that:

\[
s_{ij} = \exp(\tilde{\eta}) \left( \frac{\tilde{e}_i}{\tilde{e}_j} \right) \text{ for some vector } \tilde{e} \iff h_{ij} \equiv 1.
\]

Hansen and Scheinkman (2009) and Hansen (2012) discuss how the decomposition of the the one-period stochastic discount factor displayed on the right-hand side of (5) can be used to study long-term valuation. The third term, which is a ratio of probabilities, is used as a change of probability measure in their analysis.

In contrast, Ross (2013) uses (4) to identify the transition probability used by investors by presuming that \( s_{ij} = \exp(\eta)m_j/m_i \) for some vector \( m \) with positive entries. The uniqueness of the decomposition (5) implies that \( m_i = 1/\tilde{e}_i \) and \( h_{ij} = 1 \) for all \((i,j)\), which implies that the transition probabilities \( P \) and \( \tilde{P} \) coincide. In the next sections we address these issues under much more generality by allowing for continuous-state Markov processes. As we will see, some additional complications emerge.

### 2.3 Compounding one-period stochastic discounting

To anticipate the counterpart to (5) for a more general probability model, consider the stochastic discount factor process \( S = \{S_t : t = 0, 1, 2, \ldots \} \) that is obtained by compounding the one-period stochastic discount factor. In general, \( S_t \) depends on the history of the state from 0 to \( t \) since the increment between \( t \) and \( t+1 \) is given by:

\[
\frac{S_{t+1}}{S_t} = X_t' S X_{t+1}.
\]

Similarly, we define \( H = [h_{ij}] \) and

\[
\frac{H_{t+1}}{H_t} = X_t' H X_{t+1}.
\]

Then

\[
\frac{S_{t+1}}{S_t} = \exp(\tilde{\eta}) \left( \frac{\tilde{e} \cdot X_t}{\tilde{e} \cdot X_{t+1}} \right) \left( \frac{H_{t+1}}{H_t} \right),
\]

and, compounding over time and initializing \( S_0 = 1 \), we obtain

\[
S_t = \exp(\tilde{\eta}t) \left( \frac{\tilde{e} \cdot X_0}{\tilde{e} \cdot X_t} \right) \left( \frac{H_t}{H_0} \right)
\] (6)
with

\[ \frac{H_t}{H_0} = \prod_{\tau=1}^{t} X'_{\tau-1} \mathbf{H} X_{\tau}. \]

The initial distribution of \( X_0 \) together with the transition matrix \( \mathbf{P} \) define a probability \( P \) over realizations of the process \( X \). Because \( h_{ij} \) is obtained as a ratio of probabilities, \( H \) is a positive martingale under \( P \) for any positive specification of \( H_0 \) as a function of \( X_0 \). Thus from (6), the eigenvalue \( \eta \) contributes an exponential function of \( t \) and the eigenvector contributes a function of the Markov state to the stochastic discount factor process. In addition there is a martingale component, whose logarithm has stationary increments.

If we use the positive martingale \( H \) to induce a change of measure, we obtain the probability \( \tilde{P} \):

\[ \tilde{P}(X^t = x^t) = P(X^t = x^t) H_0 \prod_{\tau=1}^{t} x_{\tau-1} \mathbf{H} x_{\tau}. \]

for alternative possible realizations \( x^t = (x_0, x_1, ..., x_t) \) of \( X^t = (X_0, X_1, ..., X_t) \). In this formula, we presume that \( EH_0 = 1 \), and we use \( H_0 \) in order for \( \tilde{P} \) to include a change in the initial distribution of \( X_0 \). Thus the random variable \( H_0 \) modifies the distribution of \( X_0 \) under \( \tilde{P} \) vis-à-vis \( P \), and \( \tilde{P} \) specifies the altered transition probabilities. Most of our analysis conditions on \( X_0 \), in which case the choice of \( H_0 \) is inconsequential and \( H_0 \) can be set to one for simplicity.\(^4\)

The assumption of Ross (2013) amounts to requiring that \( H_t / H_0 \equiv 1 \) or equivalently that \( S \) has no martingale component. In this case,

\[ S_t = \exp(\eta t) \left( \frac{m \cdot X_t}{m \cdot X_0} \right) \]

where the entries of the vector \( m \) are given by \( m_i = 1/\bar{e}_i \) and \( \eta = \bar{\eta} \).

### 2.4 Examples

The behavior of underlying shocks is of considerable interest when constructing stochastic equilibrium models. There is substantial time series literature on the role of permanent shocks in multivariate analysis and there is a related macroeconomic literature on models with balanced growth behavior, allowing for stochastic growth. The martingale components to stochastic discount factors characterize durable components to risk adjustments in valuation over alternative investment horizons. As we will see, one source for these durable components are permanent shocks to the macroeconomic environment. But valuation models have other sources for this durability including investor preferences. The following examples illustrate that even in this \( n \)-state Markov-chain context it is possible to obtain a non-trivial martingale component for the stochastic discount factor. In particular, \( \tilde{P} \neq P \) and as a consequence the recovered transition matrix \( \tilde{P} \) is a distorted version of the “true” transition matrix \( P \).

\(^4\)The choice of \( H_0 \) will come into play in Section 5 and where \( H_0 \) will be chosen to assure stationarity of the process \( X \) under the probability measure \( \bar{P} \).
As these examples highlight, the martingale component to valuation could emerge for two reasons. First there may be stochastic growth in the equilibrium consumption. Alternatively, even if the logarithm of consumption is a time-invariant function of a stationary and ergodic Markov process added to a deterministic time trend, some specifications of investor preferences may suffice to induce a martingale component.

**Example 2.1.** Consumption-based asset pricing models assume that the stochastic discount factor process is a representation of investors’ preferences over consumption. Suppose that the growth rate of equilibrium consumption is stationary and that investor preferences can be depicted using a power utility function. Thus the marginal rate of substitution is

\[
\exp(-\delta) \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} = \phi(X_{t+1}, X_t).
\]

With this formulation, we may write

\[
s_{ij} = \phi(X_{t+1} = u_j, X_t = u_i).
\]

Stochastic growth in consumption as reflected in the entries \(s_{ij}\) will induce a martingale component to the stochastic discount factor. An exception occurs when \(C_t = \exp(g_c t)(c \cdot X_t)\) for some vector \(c\) with strictly positive entries and a constant \(g_c\), and hence

\[
\frac{C_{t+1}}{C_t} = \exp(g_c) \left( \frac{c \cdot X_{t+1}}{c \cdot X_t} \right).
\]

Here \(g_c\) governs the deterministic growth in consumption. In this case,

\[
s_{ij} = \exp(-\delta - \gamma g_c) \left( \frac{c_j}{c_i} \right)^{-\gamma}
\]

implying that \(\bar{\eta} = -(\delta + \gamma g_c)\) and \(\bar{e}_j = (c_j)^\gamma\). This special case is featured by Ross (2013), but once growth is stochastic the stochastic discount factor has a martingale component that reflects this stochastic contribution.

This special case is featured in Ross (2013), but here, except for a deterministic trend, consumption is stationary. Once the consumption process is exposed to permanent shocks, the stochastic discount factor inherits a martingale component that reflects this stochastic contribution.

**Example 2.2.** Again let \(C_t = \exp(g_c t)(c \cdot X_t)\) be a trend-stationary consumption process where \(c\) is an \(n \times 1\) vector that represents consumption in individual states of the world. The (representative) investor is now endowed with recursive preferences of Kreps and Porteus (1978) and Epstein and Zin (1989). We consider a special case of unitary elasticity of substitution. The continuation value for these preferences satisfies the recursion

\[
V_t = \left[1 - \exp(-\delta)\right] \log C_t + \frac{\exp(-\delta)}{1 - \gamma} \log E_t[\exp((1 - \gamma)V_{t+1})],
\]

(7)
where $\gamma$ is a risk aversion coefficient and $\delta$ is a subjective rate of discount. For this example, $V_t = V(t, X_t = u^i) = v_i + g_c t$ where $v_i$ is the continuation value for state $X_t = u^i$ net of a time trend. Let $v$ be the vector with entry $i$ given by $v_i$ and $\exp[(1 - \gamma)v]$ be the vector with entry $i$ given by $\exp[(1 - \gamma)v_i]$. The (translated) continuation values satisfy the fixed-point equation:

$$v_i = [1 - \exp(-\delta)] \log c_i + \frac{\exp(-\delta)}{1 - \gamma} \log [P_i \exp((1 - \gamma)v)] + \exp(-\delta) g_c$$  \hspace{1cm} (8)

where $P_i$ is the $i$-th row of the transition matrix $P$. This equation gives the current-period continuation values as a function of the current-period consumption and the discounted risk-adjusted future continuation values. We are led to a fixed-point equation because of our interest in an infinite-horizon solution. Given the solution $v$ of this equation, denote $v^* = \exp((1 - \gamma)v)$.

The implied stochastic discounting is captured by the following equivalent depictions:

$$s_{ij} = \exp(- (\delta + g_c)) \left( \frac{c_i}{c_j} \right) \left( \frac{v^*_i}{P_i v^*} \right),$$

or, compounding over time,

$$S_t = \exp(- (\delta + g_c) t) \left( \frac{c \cdot X_0}{c \cdot X_t} \right) \left( \frac{H_t}{H_0} \right)$$  \hspace{1cm} (9)

where

$$\frac{H_{t+1}}{H_t} = \frac{X_{t+1} \cdot v^*}{X_t \cdot (P v^*)}.$$ 

The process $H$ is a martingale. Perron–Frobenius theory applied to $P$ implies that $P v^* = v^*$ if, and only if $v^*$ has constant entries. As long as the solution $v$ of equation (8) satisfies $v_i \neq v_j$ for some pair $(i,j)$, we conclude that $P v^* \neq v^*$.

For this example

$$q_{ij} = p_{ij} \exp(- (\delta + g_c)) \left( \frac{c_i}{c_j} \right) \left( \frac{v^*_j}{P_i v^*} \right).$$

Solving

$$Q \bar{e} = \exp(\bar{\eta}) \bar{e}$$

for a vector $e$ with positive entries yields

$$\bar{e}_j = c_j, \quad j = 1, \ldots, n$$

$$\bar{\eta} = -(\delta + g_c).$$

This Perron–Frobenius solution (4) does not recover the transition matrix $P$ but rather $\bar{P}$ given by

$$\bar{p}_{ij} = \frac{v^*_j}{P_i v^*}.$$ 

The recovered transition matrix $\bar{P}$ is therefore distorted by the risk-adjustment that arises from
fluctuations in the continuation value $v$. In particular, when $\gamma > 1$, transition probabilities $\tilde{p}_{ij}$ are overweighted for low continuation value states $v_j$ next period. When $\gamma = 1$, the two transition matrices coincide because $v^*$ is necessarily constant across states.\footnote{The correct transition matrix is also recovered in the limiting case as $\delta \to 0$. In this case, the continuation value $v_j$ converges to a constant independent of the state $j$ due to the trend stationarity specification of the consumption process, and the risk adjustment embedded in the potential fluctuations of the continuation values becomes immaterial.}

Consider the Kreps and Porteus (1978) model from Example 2.2 and suppose that $\gamma \neq 1$. Without knowing the probability transition matrix, this model is indistinguishable from an entirely similar model except with $\gamma = 1$. When $\gamma = 1$ the economy is a special case of that given by Example 2.1. This equivalence between models is evident by endowing investors with beliefs captured by the transition matrix $\tilde{P}$ for the economy with $\gamma = 1$. This same logic extends directly to comparisons between two other values of $\gamma$, neither of which is equal to unity.

3 General framework

We now introduce a framework which encompasses a large class of relevant asset pricing models. Consistent intertemporal pricing together with a Markovian property lead us to use a class of stochastic processes called multiplicative processes. These processes are built from the underlying Markov process in a special way and will be used to model stochastic discount factors. Alternative structural economic models will imply restrictions on the stochastic discount factors that we postulate. In Section 4 we model stochastic discount factors as multiplicative functionals when framing the fundamental identification problem. In Section 5 we will discuss a sufficient condition for identification, while in Section 6 we will apply a decomposition result for multiplicative functionals from Hansen and Scheinkman (2009) to connect the probability measures $P$ and $\tilde{P}$ introduced in the preceding section.

We start a probability space $\{\Omega, F, P\}$ and a set of indices $T$ (either the non-negative integers or the non-negative reals). On this probability space, there is an $n$-dimensional, stationary Markov process $X = \{X_t : t \in T\}$ and a $k$-dimensional process $W$ with increments that are jointly stationary with $X$ and initialized at $W_0 = 0$. Write $\mathcal{F} = \{\mathcal{F}_t : t \in T\}$ for the (completed) filtration generated by histories of $W$ and $X_0$. We restrict the process $W$ to be a martingale relative to $\mathcal{F}$. The increments represent shocks to the economic dynamics and could be independently distributed over time. As such, in discrete time we suppose that

$$X_{t+1} = f(X_t, \Delta W_{t+1})$$

where $\Delta W_{t+1} \equiv W_{t+1} - W_t$.

We exploit the Markov structure of $X$ by assuming that the distribution of $(X_{t+1}, \Delta W_{t+1})$ conditioned on $\mathcal{F}_t$ depends only on $X_t$. In particular, $Y = (X, W)$ is a first-order Markov process as is $X$. The composite system $Y$ has a triangular structure because only $X_t$ is the relevant state
vector for \((X_{t+1}, \Delta W_{t+1})\). We consider processes \(M\) of the form:

\[
\log M_{t+1} - \log M_t = g(X_t, \Delta W_{t+1}).
\] (10)

We impose the analogous structure when \(X\) is a continuous-time diffusion. The process \(W\) is now an underlying \(n\) dimensional Brownian motion and suppose \(X_0\) is independent of \(W\) and \(\mathcal{F}\) the filtration associated with Brownian motion augmented to include date zero information revealed by \(X_0\). Then \(X\) and \(\log M\) processes evolve according to:

\[
d X_t = \mu(X_t) dt + \sigma(X_t) dW_t \\
\log M_t = \beta(X_t) dt + \alpha(X_t) \cdot dW_t.
\]

Notice that the conditional distribution of \((X_{t+\tau}, W_{t+\tau} - W_t)\) conditioned on \(\mathcal{F}_t\) depends only on \(X_t\) analogous to the assumption that we imposed in the discrete-time specification.\(^7\)

In both discrete and continuous time, we assume

**Assumption 3.1.** The process \(X\) is stationary and ergodic under \(P\).

For specifications given in discrete time by (10) and in continuous time by (12), the process \(\log M\) has stationary increments by construction, and thus \(M\) will display geometric growth or decay along stochastic trajectories. It will often suffice to initialize \(\log M_0\) at zero but sometimes we will allow it to depend on \(X_0\). We refer to the process \(M\) as a multiplicative functional of \(Y\).\(^8\)

Our discrete and continuous-time specifications of \(M\) restrict the class of processes to those for which the probability of \(\log M_{t+\tau} - \log M_t\) conditioned on \(\mathcal{F}_t\) depends only on \(X_t\) for any \(\tau > 0\). This paper focuses exclusively on multiplicative functionals further restricted as follows.

**Restriction 3.2.** The multiplicative functional \(M\) satisfies (10) for some function \(g\) for discrete-time models, or (12) for a pair of functions \((\beta, \alpha)\) for continuous-time models.

The multiplicative functionals we consider are strictly positive. For two such functionals \(M^1\) and \(M^2\), the product \(M^1 M^2\) and the reciprocal \(1/M^1\) are also strictly positive multiplicative functionals. Moreover, these operations also preserve Restriction 3.2 since they correspond to adding or taking the negative of the respective coefficients in expressions (10) and (12).\(^6\)

\(^6\)While the Brownian information specification (11)–(12) abstracts from jumps, these can be included without changing the implications of the analysis, see Hansen and Scheinkman (2009).

\(^7\)If in fact \(\sigma(x)' \sigma(x)\) is nonsingular for all \(x\), we may solve for the Brownian increment

\[
d W_t = [\sigma(x)' \sigma(x)]^{-1} \sigma(x)' [d X_t - \mu(X_t) dt],
\]

implying that the Brownian motion history is revealed by the \(X\) history and \(\mathcal{F}\) is also the filtration associated with the diffusion \(X\).

\(^8\)To be consistent with the existing literature, we should call this construct an extended multiplicative functional. As defined formally in Appendix A, a multiplicative functional is initialized at \(M_0 = 1\) whereas we allow for an initialization that depends on \(X_0\). That is, \(M\) is an extended multiplicative functional if \(\frac{dM}{Mt}\) is a multiplicative functional, where \(M_0\) depends only on \(X_0\). For simplicity, in this paper we drop the reference to extended.
We will use such processes to depict stochastic discount factors and positive martingales. In the following two examples, we show how multiplicative functionals relate to the framework analyzed in Section 2.

**Example 3.3.** Let $X$ be a discrete-time, $n$-state Markov chain and set $\Delta W_{t+1} = X_{t+1} - E(X_{t+1}|X_t)$. Construct

$$
\log M_{t+1} - \log M_t = (X_t)'L_1 \Delta W_{t+1} + (X_t)'L_0 X_t,
$$

where $L_1$ and $L_0$ are $n \times n$ matrices. By exponentiating the right-hand side and evaluating it for all possible coordinate vector combinations for $X_t$ and $X_{t+1}$, it may be shown that

$$
\frac{M_{t+1}}{M_t} = (X_t)'M X_{t+1}
$$

where $M$ is a matrix with positive entries. This multiplicative functional $M$ was used in Section 2 in depicting stochastic discount factors and positive martingales.

**Example 3.4.** Let $X$ be a discrete-time, $n$-state Markov chain, and $\Delta W_{t+1}$ a $k$-dimensional standard normally distributed random vector independent of $F_t$ and $X_{t+1}$. Set

$$
\log M_{t+1} - \log M_t = X_t \cdot [\bar{\beta} + \bar{\alpha}(\Delta W_{t+1})]
$$

where $\bar{\beta}$ is a vector of length $n$, and $\bar{\alpha}$ is an $n \times k$ matrix. This is a Markov switching model with state dependence in both the conditional mean and in the exposure to the normally distributed shocks.

Since the product of multiplicative functionals is a multiplicative functional, we may combine the two examples to build a more general multiplicative functional. Moreover, the composite functional can be represented in accordance to formula (10).

### 3.1 Stochastic discount factors

Formally, a stochastic discount factor process $S$ is a positive multiplicative functional with finite first moments (conditioned on $X_0$) such that the date $\tau$ price of any bounded $F_\tau$ measurable claim $\phi_t$ for $t > \tau$ is:

$$
\Pi_{\tau,t}(\phi_t) = E\left[\frac{S_t}{S_\tau} \phi_t \mid S_\tau\right].
$$

As a consequence, for a bounded claim $\psi(X_t)$ that depends only on the current Markov state, the price is

$$
[Q_t \psi](x) \doteq E[S_t \psi(X_t) \mid X_0 = x].
$$

We view $Q_t$ as the pricing operator for payoff horizon $t$. By construction,

$$
\Pi_{\tau,t}[\psi(X_t)] = [Q_{t-\tau} \psi](X_t).
$$
This operator $Q_t$ is at least well defined for bounded functions of the Markov state, but often for an even larger class of such functions depending on the tail behavior of the stochastic discount factor $S_t$.

The multiplicative property of $S$ allows us to price consistently at intermediate dates. In discrete time we may build the $t$-period operator $Q_t$ by applying the one-period operator $Q_1$ $t$ times in succession. In continuous time, $\{Q_t : t \in T\}$ forms what is called a semigroup of operators. The counterpart to a one-period operator is a generator of this semigroup that governs instantaneous valuation and which acts as a time derivative of $Q_t$ at $t = 0$. Thus in discrete time it suffices to study the one-period operator and in continuous time the generator of the family of operators $\{Q_t : t \in T\}$.

3.2 Multiplicative martingales and probability measures

Alternative probability measures equivalent to $P$ are built using strictly positive martingales. Given an $\mathcal{F}$-martingale $H$ that is strictly positive with $E(H_0) = 1$, define a probability $P^H$ such that if $A \in \mathcal{F}_\tau$ for some $\tau \geq 0$,

$$P^H(A) = E(1_A H_\tau).$$

The Law of Iterated Expectations guarantees that these definitions are consistent, that is, if $A \in \mathcal{F}_\tau$ and $t > \tau$ then

$$P^H(A) = E(1_A H_t) = E(1_A H_\tau).$$

Now suppose that $H$ is a multiplicative martingale, a multiplicative functional that is also a martingale with respect to the filtration $\mathcal{F}$. In addition, we restrict the functional to satisfy Restriction 3.2. For the martingale restriction to be satisfied with discrete-time specification (10), impose

$$E \left( \exp \left[ g(X_t, \Delta W_{t+1}) \right] | X_t = x \right) = 1.$$

Under the implied change of measure, the probability distribution for $(X_{t+1}, \Delta W_{t+1})$ conditioned on $\mathcal{F}_t$ continues to depend only on $X_t$. For the continuous-time specification (12), the drift term of the martingale $H$ satisfies\(^\text{10}\)

$$\beta(x) = -\frac{1}{2} \alpha(x) \cdot \alpha(x).$$

Under the change of measure, $W_t$ has a drift given by $\alpha(X_t)$. Thus in both discrete and continuous time, $Y$ will remain a Markov process and the triangular nature of $Y = (X, W)$ will be preserved. In general, under the probability implied by $H$, the process $W$ will no longer be a martingale and stationarity restrictions for the process $X$ may no longer be preserved. We will have more to say

\(^9\)There is a different stochastic discount factor process that we could use for much of our analysis. Let $\mathcal{F}$ denote the (closed) filtration generated by $X$. Compute $S_t = E \left[ S_t | \mathcal{F}_t \right]$. Then $S_t$ is a stochastic discount factor process pertinent for pricing claims that depend on the history of $X$. It is also a multiplicative functional constructed from $X$.

\(^\text{10}\)This restriction implies that $H$ is a local martingale and additional restrictions may be required to ensure that $H$ is a martingale.
about the potential failure to preserve stationarity later in the paper.

While we normalize \( S_0 = 1 \), we will not do the same for \( H_0 \). For some of our subsequent discussion, we use \( H_0 \) to alter the initial distribution of \( X_0 \) in a convenient way. Thus we allow \( H_0 \) to depend on \( X_0 \), but we restrict it to have expectation equal to unity. In the next section we study the connections between stochastic discount factor processes and alternative probability distributions modeled using multiplicative martingales.

4 Fundamental identification problem

Asset prices as depicted by equation (13) depend simultaneously on stochastic discount factor processes and on investor beliefs. A stochastic factor process is thus only well defined for a given probability. If we happen to misspecify investor beliefs, this misspecification can be offset by altering the stochastic discount factor. In this section we formalize this identification problem and we consider potential restrictions on the stochastic discount factors that can solve this challenge.

Definition 4.1. The pair \((S, P)\) explains asset prices if equation (13) gives the date zero price of any bounded, \( \mathcal{F}_t \) measurable claim \( \phi_t \) payable at any time \( t \in T \).

Consider now a multiplicative martingale \( H \) satisfying Restriction 3.2 and the associated probability measure \( P^H \) defined through (3.2). Similarly let \( S \) be a multiplicative functional satisfying Restriction 3.2 and initialized at \( S_0 = 1 \). We define:

\[
S^H = S \frac{H_0}{H}.
\]

The following proposition is immediate:

Proposition 4.2. Suppose that \( H \) is a multiplicative martingale satisfying Restriction 3.2 with \( E(H_0) = 1 \) and \( S \) is a multiplicative functional satisfying Restriction 3.2 with \( S_0 = 1 \). If the pair \((S, P)\) explains asset prices then the pair \((S^H, P^H)\) also explains asset prices. Moreover, \( S^H \) is a multiplicative functional satisfying Restriction 3.2 and \( S^H_0 = 1 \).

This proposition captures the notion that stochastic discount factors are only well-defined for a given probability distribution. When we change the probability distribution, we typically must change the stochastic discount factor to represent the same asset prices. In other words we have multiple ways to represent prices \( \Pi_t \):

\[
\Pi_{\tau,t}(\phi_t) = E\left[ \left( \frac{S_t}{S_\tau} \right) \phi_t \mid \mathcal{F}_\tau \right] = E^H\left[ \left( \frac{S^H_t}{S^H_\tau} \right) \phi_t \mid \mathcal{F}_\tau \right]
\]

where \( E^H \) is the expectation operator associated with \( P^H \). Suppose that this probability distribution used by investors is not known to an external analyst. Given that \( H \) can be any positive multiplicative martingale, we are left with a fundamental identification problem. From the Arrow prices alone we cannot distinguish \((S, P)\) from \((S^H, P^H)\). Restriction 3.2 does not resolve this
identification challenge. Thus we cannot recover investor beliefs from the Arrow prices alone. To achieve identification of investor beliefs, either we have to restrict the stochastic discount factor process $S$ or we have to restrict the probability distribution used to represent the valuation operators $\Pi_{\tau,t}$ for $\tau \leq t \in T$.

There are multiple ways we might address this lack of identification. First, we might impose rational expectations, observe time series data, and let the Law of Large Numbers for stationary distributions determine the probabilities. Then observations for a complete set of Arrow securities allow us to identify $S$. See Hansen and Richard (1987) for an initial discussion of the stochastic discount factors and the Law of Large Numbers, and see Hansen and Singleton (1982) for an econometric approach that imposes a parametric structure on the stochastic discount factor and avoids assuming that the analyst has access to data on the complete set of Arrow securities.

Alternatively, we may restrict the stochastic discount factor process further. For instance, risk-neutral pricing, introduced in Section 2.1, restricts the stochastic discount factor to be predetermined or locally predictable. Thus for a discrete-time specification:

$$\log S_{t+1} - \log S_t = \log(\bar{q}(X_t))$$

where $\bar{q}(X_t)$ is the price of one-period discount bond. When this restriction is used, typically there is no claim that the resulting probability distribution is the same as that used by investors.

A different restriction imposes a special structure on $S$:

**Restriction 4.3.** Let

$$S_t = \exp(-\delta t) \frac{m(X_t)}{m(X_0)}$$

for some positive function $m$ and some real number $\delta$.

Later in this section we will give an economic example that satisfies this restriction, but as the examples in Section 2.4 show this restriction is not needed for the existence of a Markov structure in pricing. For further discussion see Section 6. Ross (2013) proved an identification result under Restriction 4.3 when the dynamics of $X$ are driven by a finite-state Markov chain as in Section 2. In the next section we show that a strengthening of Restriction 4.3 is sufficient to guarantee that the Arrow prices identify the stochastic discount factor and a probability distribution associated with that stochastic discount factor in the more general framework introduced in Section 3. As we will see, the specification in Restriction 4.3 curtails long-term risk pricing in ways that may be unappealing.

## 5 A recovery result

Our results for finite-state Markov economies turn out too special in a particular way that is important to our analysis. For general Markov processes, even ones that are stationary, Restriction 4.3 is not sufficient to identify the pair $(S, P)$ from the Arrow prices embedded in the pricing operators
In this section we establish a more general recovery result requiring an additional ergodicity condition imposed on the recovered measure. Prior to developing this more general result, we illustrate with an example why such a restriction is required.

5.1 An illustration of recovery failure

Example 5.1. Suppose that $X$ is a Feller square root process:

$$dX_t = -\kappa (X_t - \bar{\mu}) dt + \bar{\sigma} \sqrt{X_t} dW_t$$

where $\kappa > 0$, $\bar{\mu} > 0$ and $\kappa \bar{\mu} \geq \frac{1}{2} (\bar{\sigma})^2$. With these restrictions, the process $X$ is strictly positive.

Let $m(x) = \exp (\zeta x)$ so that the stochastic discount factor satisfies

$$S_t = \exp [-\delta t + \zeta (X_t - X_0)],$$

or, in differential form,

$$d \log S_t = [-\delta - \zeta \kappa (X_t - \bar{\mu})] dt + \zeta \bar{\sigma} \sqrt{X_t} dW_t.$$ 

We choose a multiplicative martingale of the form:

$$H_t = \exp \left[ - \int_0^t \xi(X_s) dW_s - \frac{1}{2} \int_0^t \xi(X_s)^2 ds \right].$$

and thus $S^H_t = S_t H_0 / H_t$ must satisfy:

$$d \log S^H_t = \left( -\delta - \zeta \kappa (X_t - \bar{\mu}) + \frac{1}{2} [\xi(X_t)]^2 \right) dt + \left[ \zeta \bar{\sigma} \sqrt{X_t} + \xi(X_t) \right] dW_t.$$ 

We will show that there is a choice of $\xi$ such that

$$S_t \neq S^H_t = \exp (-\tilde{\delta} t) \frac{\exp (\tilde{\zeta} X_t)}{\exp (\zeta X_0)},$$

or

$$d \log S^H_t = \left[ -\tilde{\delta} - \tilde{\zeta} \kappa (X_t - \bar{\mu}) \right] dt + \tilde{\zeta} \bar{\sigma} \sqrt{X_t} dW_t.$$ 

Equating drift and volatility coefficients in (5.1) and (5.1) yields

$$\xi(x) = \left( \tilde{\zeta} - \zeta \right) \bar{\sigma} \sqrt{x},$$

and

$$-\zeta \kappa x + \frac{1}{2} \left( \tilde{\zeta} - \zeta \right)^2 \sigma^2 x = -\tilde{\zeta} \kappa x.$$
Consequently, the solution for $S^H$ satisfies

$$
\tilde{\zeta} = \zeta - \frac{2\kappa}{\sigma^2},
$$

and the martingale $H$ is characterized by

$$
\xi(x) = -\frac{2\kappa}{\sigma} \sqrt{x}.
$$

Under the probability $P^H$ the Brownian motion $W$ has a local drift $-\xi(X_t)$ and hence the local drift of $X$ is

$$
-\kappa (x - \bar{\mu}) + 2\kappa x = \kappa x + \kappa \bar{\mu}.
$$

Thus the dynamics for $X$ are

$$
dX_t = \kappa (X_t + \bar{\mu}) dt + \sigma \sqrt{X_t} d\tilde{W}_t
$$

where $\tilde{W}_t = W_t + \int_0^t \xi(X_s) ds$ is a Brownian motion under $P^H$.

Observe that in this example, under $P$, the distribution of the process $X$ converges to a stationary distribution, independent of the initial conditions. However, under $P^H$ the realizations of $X_t$ diverge as $t \to \infty$ with probability one. It turns out that, as we show in the next section, adding an assumption that $X$ is stationary and ergodic allows us to uniquely identify the stochastic discount factor and associated probability.

### 5.2 Recovery under ergodicity

As we change probability measures, stationarity and ergodicity will not necessarily be satisfied, but checking for this stability under the probability $P^H$ will be featured in our analysis. Thus in establishing a recovery result, we impose the following restriction on the stochastic evolution of $X$ under the probability distribution $P^H$.

**Restriction 5.2.** The process $X$ is stationary and ergodic under $P^H$.

Stationarity requires the choice of an appropriate $H_0 = h(X_0)$ that induces a stationary distribution under $P^H$ for the Markov process $X$.\footnote{Given $H_t/H_0$, the random variable $H_0 = h(X_0)$ must satisfy the equation:

$$
E \left[ \psi(X_t) \left( \frac{H_t}{H_0} \right) h(X_0) \right] = E [\psi(X_0) h(X)]
$$

for any bounded (Borel measurable) $\psi$ and any $t \in T$.} As we change probability measures, stationarity and ergodicity will not necessarily be satisfied, but checking for this stability under the probability $P^H$ will be featured in our analysis.

If $X$ satisfies Restriction 5.2 then it satisfies a Strong Law of Large Numbers.\footnote{E.g., Breiman (1982), Corollary 6.23 on page 115.} If $\psi$ has finite
expected value then

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} \psi(X_t) = E^H \psi(X_0)
\]

almost surely when \( T \) is the set of nonnegative integers (discrete time). The process \( X \) also obeys another version of Law of Large Numbers that considers convergence in means: \(^{13}\)

\[
\lim_{N \to \infty} E^H \left[ \left| \frac{1}{N} \sum_{t=1}^{N} \psi(X_t) - E^H \psi(X_0) \right| \right] = 0,
\]

As a consequence of both versions of the Law of Large Numbers, times series averages of conditional expectations also converge:

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E^H [\psi(X_t) | X_0] = E^H \psi(X_0)
\]

almost surely in discrete time. Corresponding results hold in continuous time.

We now show that the pair \((S^H, P^H)\) that explains asset prices and satisfies Restriction 4.3 is unique, provided that stationarity and ergodicity are preserved under the measure \( P^H \). Formally,

**Proposition 5.3.** Suppose \((S, P)\) explains asset prices and \( S \) satisfies Restriction 4.3. Let \( H \) be a positive multiplicative martingale such that \((S^H, P^H)\) also explains asset prices and \( X \) is stationary and ergodic under \( P^H \). If \( S^H \) also satisfies Restriction 4.3, then \( H \equiv 1 \).

The proof of this theorem is similar to the proof of a related uniqueness result in Hansen and Scheinkman (2009) and is detailed in Appendix B.\(^{14}\) Proposition 5.3 gives an identification result, which is a counterpart of the recovery result in Ross (2013). Under the conditions of this proposition, remarkably we may infer both the stochastic discount factor and a probability distribution associated with that stochastic discount factor from the Arrow prices. Notice that the identification result given in Proposition 5.3 does not resolve the fundamental identification problem from Proposition 4.2. In general, there still exists a whole class of pairs \((S^H, P^H)\) that explain asset prices in the sense of Definition 4.1. Proposition 5.3 shows uniqueness only in the class of stochastic discount factors satisfying Restriction 4.3 and probability measures under which \( X \) is ergodic.\(^{15}\)

The recovery result presented in this section relies on two restrictions. In addition to restricting \( X \) to be ergodic under \( P^H \), we also followed Ross (2013) by imposing Restriction 4.3. In the next section, we relax Restriction 4.3. Instead we postulate a stochastic discount factor \( S \) in the

\(^{13}\)Ibid., Corollary 6.25 on page 117.

\(^{14}\)Hansen and Scheinkman (2009) and Hansen (2012) used an implication of the SLLN in their analysis.

\(^{15}\)In a continuous-time Brownian information setup, alternative conditions on the boundary behavior of the underlying Markov process also uniquely identify a probability measure. These conditions utilize linkages of the Perron–Frobenius theorem to the Sturm–Liouville problem in the theory of second-order differential equations. Carr and Yu (2012) and Dubynskiy and Goldstein (2013) impose conditions on reflecting boundaries, while Walden (2014) analyzes natural boundaries. While technical conditions like these may deliver a unique solution, as we will see they do not resolve the fundamental identification problem once we relax Restriction 4.3.
class of multiplicative functionals introduced in Section 3 and ask what does the Perron–Frobenius approach actually recover.

6 What is recovered?

In this section we develop the implications of dispensing with Restriction 4.3. As we pointed out in Example 5.1, an ergodicity condition must be imposed if one hopes to identify a single probability measure even when Restriction 4.3 holds. For this reason we maintain the ergodicity Restriction 5.2. We start by showing that, in this case, the Perron–Frobenius approach recuperates a stochastic discount factor that in general differs from the actual stochastic discount factor by a martingale component. This martingale is identically equal to one if and only if Restriction 4.3 holds. The inferred probability \( \tilde{P} \) would thus, except when Restriction 4.3 holds, be distinct from the original probability \( P \). In the remaining part of this section, we provide additional insights on the role of the martingale component. We compare this martingale to the ‘permanent shock’ extracted from log \( S \) using familiar time series methods, and we discuss classes of models that share the same martingale component. In Section 7 we then investigate the implications of the alternative probability \( \tilde{P} \) when used to represent long-term values.

6.1 Perron–Frobenius approach to valuation

Suppose there exists a function \( \tilde{e}(x) \) that solves the following Perron–Frobenius problem:

**Problem 6.1** (Perron–Frobenius). Find a scalar \( \tilde{\eta} \) and a function \( \tilde{e} > 0 \) such that for every \( t \in T \),

\[
[Q_t \tilde{e}](x) = \exp(\tilde{\eta} t) \tilde{e}(x).
\]

A solution to this problem will necessarily satisfy the conditional moment restriction:

\[
E \left[ S_t \tilde{e}(X_t) \mid \mathcal{F}_\tau \right] = \exp \left[ (t - \tau) \tilde{\eta} \right] S_\tau \tilde{e}(X_\tau)
\] (14)

for \( t \geq \tau \).

When the state space is finite as in Section 2, functions of \( x \) can be represented as vectors in \( \mathbb{R}^n \), and the operator \( Q_1 \) can be represented as a matrix \( Q \). In this case, the existence and uniqueness (up to scale) of a solution to Problem 6.1 is well understood. Existence and uniqueness are more complicated in the case of general state spaces. Hansen and Scheinkman (2009) present sufficient conditions for the existence of a solution, but even in examples commonly used in applied work, multiple (scaled) positive solutions are a possibility. See Hansen and Scheinkman (2009), Hansen (2012) and our subsequent discussion for such examples. When we have a solution of the Perron–Frobenius problem, we follow Hansen and Scheinkman (2009), and define a process \( \tilde{H} \) that
satisfies:

\[ \frac{\tilde{H}_t}{\tilde{H}_0} = \exp(-\tilde{\eta}t)S_t \frac{\tilde{e}(X_t)}{\tilde{e}(X_0)}. \]

The process \( \tilde{H} \) is a \( \mathcal{F} \)-martingale under the probability measure \( P \):

\[
E \left[ \tilde{H}_t \mid \mathcal{F}_\tau \right] = \frac{\exp(-\tilde{\eta}t)}{\tilde{e}(X_0)} E \left[ S_t \tilde{e}(X_t) \mid \mathcal{F}_\tau \right] \tilde{H}_0 = \frac{\exp(-\tilde{\eta}\tau)}{\tilde{e}(X_0)} S_\tau \tilde{e}(X_\tau) \tilde{H}_0 = \tilde{H}_\tau,
\]

where in the second equality we used equation (14).

The process \( \tilde{H} \) inherits much of the mathematical structure of the original stochastic discount factor process \( S \) and is itself a multiplicative martingale. For instance, if \( S \) has the form given by equation (10), then in discrete time:

\[
\log \tilde{H}_{t+1} - \log \tilde{H}_t = g(X_t, \Delta W_{t+1}) + \log \tilde{e}(X_{t+1}) - \log \tilde{e}(X_t) - \tilde{\eta} = \tilde{g}(X_t, \Delta W_{t+1})
\]

where we have used the fact that \( X_{t+1} = f(X_t, \Delta W_{t+1}) \).

As we noted, such a construction may not be unique. (For instance, see Example 5.1.) If, however, we restrict \( \tilde{H} \) so that \( X \) under the implied probability distribution \( \tilde{P} \) is stationary and ergodic (Restriction 5.2), then there is at most one such \( \tilde{H} \). The proof of this uniqueness result is essentially the same as that of Proposition 5.3, and is a minor extension of a result reported in Hansen and Scheinkman (2009). While this approach recovers a single probability \( \tilde{P} \) associated with the martingale \( \tilde{H} \), there is no claim that this constructed probability is the one that generates the data.

If \( \tilde{H} \equiv 1 \), then \( S \) necessarily satisfies Restriction 4.3. On the other hand, if \( S \) satisfies Restriction 4.3, then \( \tilde{e}(x) = \frac{1}{m(x)} \) solves the Perron–Frobenius Problem 6.1 and the associated \( \tilde{H} \equiv 1 \). Thus we obtain:

**Proposition 6.2.** \( S \) satisfies Restriction 4.3 if, and only if, there exists a solution \( \tilde{e} > 0 \) to the Perron–Frobenius problem and the associated \( \tilde{\eta} \) such that:

\[
\frac{\tilde{H}_t}{\tilde{H}_0} = \exp(-\tilde{\eta}t)S_t \frac{\tilde{e}(X_t)}{\tilde{e}(X_0)} \equiv 1.
\]

This proposition reveals the recovery result as a special case of the decomposition (6.1) of Hansen and Scheinkman (2009). Recall that the primitive of our analysis is the pair \((S, P)\) that explains asset prices where the probability measure \( P \) satisfies the stochastic stability condition. If the Perron–Frobenius Problem 6.1 yields a solution \( \tilde{H} \equiv 1 \), the uniqueness result implies that \( P \) is the single stationary and ergodic measure that will be recovered. In the general case, the stochastic discount factor \( S \) contains a martingale component \( \tilde{H} \neq 1 \) for which the associated measure \( \tilde{P} \) satisfies the ergodicity Restriction 5.2. In this case the recovery procedure alone will *not* identify the underlying probability measure. Instead, it will recover the measure \( \tilde{P} \) implied by \( \tilde{H} \). The
associated stochastic discount factor that explains asset prices is
\[ \tilde{S}_t = S_t \frac{H_0}{H_t} = \exp(\tilde{\eta}t) \frac{e(X_0)}{e(X_t)}, \]
which satisfies Restriction 4.3 by construction but differs from the stochastic discount factor \( S \) associated with \( P \) by the martingale \( \tilde{H} \).

While we have studied recovery using \( P \) as a starting point and then constructed \( \tilde{P} \), our choice of \( P \) for this purpose could in fact be altered and \( \tilde{P} \) would remain the same. Suppose we start with a positive multiplicative martingale \( H' \) satisfying Restriction 3.2 and an associated probability \( P' \). Construct
\[ S'_t = S_t \frac{H'_0}{H'_t}, \]
which is also a multiplicative functional satisfying Restriction 3.2. By construction \( (S', P') \) explains asset prices. The Perron–Frobenius approach applied to \( (S', P') \) will recover the same \( (\tilde{S}, \tilde{P}) \). Given the flexibility in the starting point, we cannot infer \( (S, P) \) from \( (\tilde{S}, \tilde{P}) \) and must appeal to other evidence and restrictions for identification.

6.2 An illustration of what is recovered

In the previous discussion, we described two issues arising in the recovery procedure. First, the positive candidate solution for \( \tilde{e}(x) \) may not be unique. Our Restriction 5.2 allows us to pick the single solution that preserves stationarity and ergodicity. Second, and more importantly, even this unique choice may not uncover the true probability distribution if there is a martingale component in the stochastic discount factor. The following example shows that in a simplified version of a stochastic volatility model one always recovers an incorrect probability distribution.

Example 6.3. Consider a stochastic discount factor model with state-dependent risk prices. Suppose that
\[ d \log S_t = \tilde{\beta} dt - \frac{1}{2} X_t (\tilde{\alpha})^2 dt + \sqrt{X_t} \tilde{\alpha} dW_t \]
where \( X \) has the square root dynamics given in Example 5.1 and \( \tilde{\beta} < 0 \). By design, the instantaneous short-term interest rate is constant. Guess a solution for a positive eigenfunction:
\[ \tilde{e}(x) = \exp(\nu x). \]
Since \( \{\exp(-\tilde{\eta}t)S_t \tilde{e}(X_t) : t \geq 0\} \) is a martingale, its local mean should be zero:
\[ \tilde{\beta} - \frac{1}{2} (\tilde{\alpha})^2 x - \nu \kappa x + \nu \kappa \mu + \frac{1}{2} x (\nu \sigma + \tilde{\alpha})^2 - \tilde{\eta} = 0. \]

16Example 3.1 in Hansen and Scheinkman (2014) points out this connection in a finite-state Markov chain setting.
In particular, the coefficient on \( x \) should satisfy
\[
v \left[ -\kappa + \frac{1}{2} v (\bar{\sigma})^2 + \bar{\sigma} \bar{\alpha} \right] = 0.
\]

There are two solutions: \( v = 0 \) and
\[
v = \frac{2\kappa - 2\bar{\alpha} \bar{\sigma}}{(\bar{\sigma})^2}
\]
(15)

For this example, the risk neutral dynamics for \( X \) corresponds to the solution \( v = 0 \) and the instantaneous risk-free rate is constant and equal to \( -\bar{\beta} \). The resulting \( X \) process remains a square root process, but with \( \kappa \) replaced by
\[
\kappa_n = \kappa - \bar{\alpha} \bar{\sigma}.
\]

Although \( \kappa \) is positive, \( \kappa_n \) could be positive or negative. If \( \kappa_n > 0 \), then the Perron–Frobenius problem that we feature extracts the risk-neutral dynamics, but this is distinct from the actual probability evolution for \( X \). Suppose instead that \( \kappa_n < 0 \). This occurs when \( \kappa < \bar{\sigma} \bar{\alpha} \). In this case we choose \( v \) according to (15), implying that \( \kappa \) is replaced by
\[
\kappa_{pf} = -\kappa + \bar{\sigma} \bar{\alpha} = -\kappa_n > 0.
\]

The resulting dynamics are distinct from both the risk-neutral dynamics and the original dynamics for the process \( X \).

This example was designed to keep the algebra simple, but there are straightforward extensions that are described in Hansen (2012). In this example there is a single shock that shifts both the stochastic discount factor and the state variable \( X \) that govern volatility. The so-called “local risk price” for this shock is given by \(-\sqrt{X_t} \bar{\alpha}\). There is a straightforward extension that extends the dimension of the Brownian increment and reproduces analogous findings. In addition, a predictable component can be included in the local mean for the log \( S \). Hansen (2012) includes these generalizations using a model with affine dynamics of the type featured by Duffie and Kan (1994). Borovička et al. (2011) analyze the continuous-time Campbell and Cochrane (1999) model and discuss the multiplicity of solutions to the associated Perron–Frobenius problem and the appropriate selection of a solution. Borovička et al. (2011) and Hansen (2012) perform analogous calculations for a continuous-time version of the Bansal and Yaron (2004) model.

Multiplicity of solutions to the Perron–Frobenius problem is prevalent in models with continuous states. In confronting this multiplicity, Hansen and Scheinkman (2009) show that the eigenvalue \( \bar{\eta} \) that leads to a stochastically stable probability measure \( \bar{P} \) gives a lower bound on the set of eigenvalues associated with strictly positive eigenfunctions. For a univariate continuous-time Brownian motion setup, Walden (2014) and Park (2014) construct positive solutions \( e \) for every candidate eigenvalue \( \eta > \bar{\eta} \). However, none of these solution pairs \((e, \eta)\) leads to a probability measure that satisfies Restriction 5.2.

In what follows we describe two channels by which martingale components are implied by
6.3 Permanent shocks

Alvarez and Jermann (2005) and Bakshi and Chabi-Yo (2012) use empirical estimates of bounds on the magnitude of the martingale component as evidence for the impact of permanent shocks to the stochastic discount factor process. To understand this link more fully consider an alternative martingale extraction that is familiar from time series analysis. Recall that for the multiplicative functional $S$ of the type we constructed in Section 3, $\log S_t$ has stationary increments. When identifying shocks and assessing their impact, empirical macroeconomists often construct models in terms of logarithms to impose positivity and to exploit the resulting stationarity in increments. Permanent shocks are identified by extracting a martingale component from the time series expressed in logarithms. This leads us to use directly the model $\log S_t$.

Hansen (2012) analyzes the connection between martingale extractions in levels versus logarithms for a class of continuous-time models. Here we briefly review and expand on this analysis.

The martingale component of $\log S$ can be computed from the “additive” decomposition

$$\log S_t = \hat{\eta}t + \log \hat{H}_t - \log \hat{H}_0 - \tilde{c}(X_t) + \tilde{e}(X_0)$$

(16)

where $\log \hat{H}$ is a martingale. The increment of $\log \hat{H}$ reveals a permanent shock up to a scale normalization. For the Markov diffusion model this martingale increment can be expressed as $\tilde{v}(X_t) \cdot dW_t$. Similarly, for positive martingale component $\tilde{H}$,

$$\frac{d\tilde{H}_t}{\tilde{H}_t} = \tilde{v}(X_t) \cdot dW_t,$$

but in general $\tilde{v} \neq \tilde{\upsilon}$. Nevertheless, when $\log S$ has a non-degenerate martingale component, then $S$ has a multiplicative martingale component and vice versa. When $M$ is globally log-normal, the increments will coincide; but outside the log-normal environment this link may cease to be direct.

Recall that $\tilde{\eta}$ is the (negative) growth rate of $S$. The term $\hat{\eta}$ in (16) is the trend growth in $\log S$. The difference $\tilde{\eta} - \hat{\eta}$ between these terms is what is sometimes called the average relative entropy of $S$. (See for instance, Alvarez and Jermann (2005) and Backus et al. (2014) for use of this construct in closely related ways.) It reflects the magnitude of the martingale component $\tilde{H}$:

$$\tilde{\eta} - \hat{\eta} \equiv \lim_{t \to \infty} \frac{1}{t} [\log E(S_t \mid X_0 = x) - E(\log S_t \mid X_0 = x)]$$

$$= -\lim_{t \to \infty} \frac{1}{t} E \left[ \log \frac{\tilde{H}_t}{\hat{H}_t} - \log \tilde{H}_0 - \log \tilde{c}(X_t) + \log \tilde{e}(X_0) \mid X_0 = x \right]$$

$$= -E \left[ \log \frac{\tilde{H}_1}{\hat{H}_0} \right].$$

provided that $E[\log \tilde{e}(X_t)]$ is finite. Thus the difference between $\tilde{\eta}$ and $\hat{\eta}$ from two different martingale constructions reveals the average of the one-period log-likelihood ratio between the original
probability measure and the $\tilde{\cdot}$ probability measure. In Section 9, we will generalize the measurement of the magnitude of this martingale component beyond the entropy measure.

7 Long-term pricing

We now show that the probability measure identified by our application of Perron–Frobenius theory absorbs the risk adjustments over long investment horizons, and study the consequences of this risk adjustment.

The Perron–Frobenius theory features an eigenvalue $\bar{\eta}$ and the associated eigenfunction $\bar{e}$ which determine the limiting behavior of securities with payoffs far in the future. We exploit this domination to study long-term risk-return tradeoffs building on the work of Hansen and Scheinkman (2009) and Hansen et al. (2008) and long-term holding period returns building on the work of Alvarez and Jermann (2005). We will show that under the $\tilde{P}$ probability measure, risk-premia on long term cash flows that grow stochastically are zero; but not under the $P$ measure. We also show that the holding period return on long-term bonds is the increment in the stochastic discount factor under the $\tilde{P}$ measure.

For some of the results in this section, we impose the following refinement of ergodicity.

Restriction 7.1. The Markov process $X$ is aperiodic, irreducible and positive recurrent under the measure $\tilde{P}$.

We refer to this restriction as stochastic stability, and it implies that

$$\lim_{t \to \infty} \tilde{E} [f(X_t) \mid X_0 = x] = \tilde{E} [f(X_0)]$$

almost surely provided that $\tilde{E} [f(X_0)] < \infty$.\footnote{For discrete-time models, see Meyn and Tweedie (2009) Theorem 14.0.1 on page 334 for an even stronger conclusion. We prove the results in this section and appendices only for the discrete-time case. Analogous results for the continuous-time case would use propositions in Meyn and Tweedie (1993).} In this formula, we use the notation $\tilde{E}$ to denote expectations computed with the probability $\tilde{P}$ implied by $\tilde{H}$,\footnote{Note that we use $\tilde{P}$ and $\tilde{E}$ instead of the more cumbersome $P^{\tilde{H}}$ and $E^{\tilde{H}}$.}

7.1 Long-term yields

We first show that the characterization of the eigenvalue $\bar{\eta}$ in Section 2.2 extends to this more general framework. Consider,

$$[Q_t \psi] (x) = E [S_t \psi(X_t) \mid X_0 = x] = \exp(\bar{\eta} t) \bar{e}(x) \tilde{E} \left[ \frac{\psi(X_t)}{\bar{e}(X_t)} \right] \mid X_0 = x$$

for some positive payoff $\psi(X_t)$ expressed as a function of the Markov state. For instance, to price pure discount bonds we should set $\psi(x) \equiv 1.$
Consider the implied yield to this investment under the measure \( \tilde{P} \):

\[
\tilde{y}_t[\psi(X)](x) = \frac{1}{t} \log \tilde{E}[\psi(X_t) | X_0 = x] - \frac{1}{t} \log [Q_t \psi](x)
\]

Taking the limit as \( t \to \infty \):

\[
\lim_{t \to \infty} \tilde{y}_t[\psi(X)](x) = -\tilde{\eta} + \lim_{t \to \infty} \frac{1}{t} \log \tilde{E}[\psi(X_t) | X_0 = x]
\]

\[
\quad - \lim_{t \to \infty} \frac{1}{t} \log \tilde{E}\left[ \frac{\psi(X_t)}{e(X_t)} \bigg| X_0 = x \right]
\]

This expression suggests that \(-\tilde{\eta}\) is the long term yield maturing in the distant future, provided that the last two terms vanish. These last two terms vanish under the stochastic stability Restriction 7.1 provided that \( \tilde{E}[\psi(X_0)] < \infty, \quad \tilde{E}\left[ \frac{\psi(X_0)}{e(X_0)} \right] < \infty. \)

Given the logarithms of the conditional expectations are divided by \( t \), \(-\tilde{\eta}\) is the long-term yield under more general circumstances. See Appendix C for details.

When we use the original measure \( P \) instead of \( \tilde{P} \), the finite horizon yields will differ. The limiting yield would still equal \(-\tilde{\eta}\) under the original probability measure, provided that \( E[\psi(X_t)] < \infty. \)

In summary, stationary cash flow risk does not alter the long-term yield since \(-\tilde{\eta}\) is also the yield on a long-term discount bond. The limiting risk premium are zero under both probability measures because of the transient nature of the cash-flow risk.

So far we have studied yields on securities with payoffs \( \psi(X_t) \) at date \( t \). Next we introduce payoffs that are exposed to permanent shocks and consequently grow stochastically.

### 7.2 Long-term risk-return tradeoff

Consider a positive cash flow process \( G \) that grows stochastically over time. We modeled such a cash flow as a multiplicative functional satisfying Restriction 3.2 with payoff \( G_t \) at time \( t \). By design, the growth rate in logarithms fluctuates randomly depending on the Markov state and the shocks modeled as martingale increments. Given the multiplicative nature of \( G \), the impact of growth compounds over time. While we expect the long-term growth rate of \( G \) to be positive, the overall exposure to shocks will increase with the payoff date \( t \). Cash flow risk will no longer be transient as in Section 7.1. For convenience we initialize \( G_0 = 1. \)

The yield on the cash flow \( G \) under the recovered model \( (\tilde{S}, \tilde{P}) \) is

\[
\tilde{y}_t[G](x) = \frac{1}{t} \log \frac{\tilde{E}[G_t | X_0 = x]}{\tilde{E}[\tilde{S}_t G_t | X_0 = x]} = \frac{1}{t} \log \frac{E\left[ \frac{\tilde{G}_t}{H_0} G_t \bigg| X_0 = x \right]}{E\left[ \tilde{S}_t G_t \bigg| X_0 = x \right]} \quad (18)
\]
where
\[ \frac{\tilde{H}_t}{\tilde{H}_0} = \exp(-\eta t) \frac{\tilde{e}(X_t)}{e(X_0)}. \]

Using the formula for \( \tilde{H} \) in equation (18), we obtain:
\[ \tilde{y}_t[G](x) = \frac{1}{t} \log E \left[ \frac{\tilde{H}_t}{\tilde{H}_0} G_t \mid X_0 = x \right] - \frac{1}{t} \log E [S_t G_t \mid X_0 = x] \]
\[ = -\eta + \frac{1}{t} \log E \left[ S_t G_t \frac{\tilde{e}(X_t)}{e(X_0)} \mid X_0 = x \right] - \frac{1}{t} \log E [S_t G_t \mid X_0 = x], \]

To analyze the limiting yield, note that the multiplicative functional \( SG \) satisfies Restriction 3.2. Let \( \eta^* \) denote the Perron–Frobenius eigenvalue using \( SG \) in place of \( S \) when solving Problem 6.1. This entails solving:
\[ E [S_t G_t e^* (X_t) \mid X_0 = x] = \exp (\eta^* t) e^* (x) \]
and selecting the eigenvalue-eigenfunction pair such that the implied martingale induces stationarity and ergodicity, as we have argued previously. Imitating our calculation of yields on securities with stationary payoffs, \( \eta^* \) is the asymptotic growth rate of \( SG \):
\[ \lim_{t \to \infty} \frac{1}{t} \log E \left[ S_t G_t \frac{e(X_t)}{e(X_0)} \mid X_0 = x \right] = \frac{1}{t} \log E [S_t G_t \mid X_0 = x] = \eta^*, \]
provided that additional moment restrictions are imposed. See Appendix C for details. Thus
\[ \lim_{t \to \infty} \frac{1}{t} \log E [G_t \mid X_0 = x] = -\eta + \eta^* - \eta^* = -\eta \]

The limiting yield computed under \( \tilde{P} \) remains the same even after we have introduced stochastic growth in the payoff. In particular, the long-term risk premia on cash flows are zero under \( \tilde{P} \) even when the cash flows displays stochastic growth. This conclusion, however, is altered when we compute the yield under the original probability measure. Now the horizon \( t \) yield is:
\[ y_t[G](x) = \frac{1}{t} \log E [G_t \mid X_0 = x] - \frac{1}{t} \log E [S_t G_t \mid X_0 = x] \]

When \( S \) and \( G \) have non-trivial martingale components, the expected rate of growth of the multiplicative functional \( SG \) does not typically equal the sum of the expected rate of growth of \( S \) plus the expected rate of growth of \( G \). Hence the limiting yield of the payoff \( G \) under \( P \) differs from \( -\eta \).

While the long-term risk-premia on stochastically growing cash flows are zero under \( \tilde{P} \), these same long-term risk premia under \( P \) are often not degenerate. Essentially by construction, the probability measure associated with Perron–Frobenius Theory makes the long-term risk-return tradeoff vanish. Thus absorbing the martingale component of a stochastic discount factor into the \( \tilde{P} \) probability measure can have a particularly important impact on risk pricing over long investment
7.3 Forward measures and holding-period returns to long-term bonds

Holding-period returns on long-term bonds inform us about the solution to the Perron–Frobenius Problem (Problem 6.1) for the stochastic discount factor process. The long maturity limit of a holding period return \( R_{t,t+1}^\tau \) between period \( t \) and period \( t+1 \) on bond with maturity \( \tau \) is

\[
R_{t,t+1}^\infty = \lim_{\tau \to \infty} R_{t,t+1}^\tau = \lim_{\tau \to \infty} \left[ \frac{Q_{\tau-1}(X_{t+1})}{Q_\tau(X_t)} \right] = \exp(-\eta) \frac{\bar{e}(X_{t+1})}{\bar{e}(X_t)}
\]

almost surely provided that the stochastic stability Restriction 7.1 is imposed and

\[
\tilde{E}\left[ \frac{1}{\bar{e}(X_0)} \right] < \infty.
\]

Since prices of discount bonds at alternative investment horizons are used to construct forward measures, this same computation allows to characterize the limiting forward measure. As Hansen and Scheinkman (2014) argue, the limit of the forward probability measures defined in Section 2.1 above coincides with the measure recovered using Perron–Frobenius theory. To see why, consider the forward measure at date \( t \) for a maturity \( \tau \). We represent this measure using the positive random variable

\[
F_{t,t+\tau} = \frac{S_{t+\tau}/S_t}{E\left[ S_{t+\tau}/S_t \left| X_t \right. \right]},
\]

with conditional expectation one given the date-\( t \) Markov state \( X_t \). The associated conditional expectations computed using this forward measure are formed by first multiplying by \( F_{t,t+\tau} \) prior to computing the conditional expectations using the \( P \) measure. Thus \( F_{t,t+\tau} \) determines the conditional density of the forward measure with respect to the original measure. The implied one-period transition between date \( t \) and date \( t+1 \) is given by

\[
E\left( F_{t,t+\tau} \left| F_{t+1} \right. \right) = \left( \frac{S_{t+1}}{S_t} \right) \left( \frac{E\left[ S_{t+\tau}/S_{t+1} \left| F_{t+1} \right. \right]}{E\left[ S_{t+\tau}/S_t \left| F_t \right. \right]} \right) = \left( \frac{S_{t+1}}{S_t} \right) \left[ \frac{Q_{\tau-1}(X_{t+1})}{Q_\tau(X_t)} \right]
\]

which by the Law of Iterated Expectations has expectation equal to unity conditioned on date \( t \) information. Using our previous calculations, taking limits as the investment horizon \( \tau \) becomes arbitrarily large, the limiting transition distribution is determined by the random variable:

\[
\left( \frac{S_{t+1}}{S_t} \right) R_{t,t+1}^\infty = \frac{\bar{H}_{t+1}}{H_t}
\]

which reveals the martingale increment in the stochastic discount factor. This also shows that the limiting one-period transition constructed from the forward measure coincides with the Perron–Frobenius transition probability. Qin and Linetsky (2014a) characterize this limiting behavior under more general circumstances without relying on a Markov structure.
When the right-hand side of (19) is exactly one, the one-period stochastic discount factor is the inverse of the limiting holding-period return. This link was first noted by Kazemi (1992). More generally, it follows from this formula that

\[ E \left[ \log R_{t,t+1}^\infty \mid X_t = x \right] \leq E \left[ \log S_t - \log S_{t+1} \mid X_t = x \right], \tag{20} \]

since Jensen’s Inequality informs us that

\[ E \left[ \log \tilde{H}_{t+1} - \log \tilde{H}_t \mid X_t = x \right] \leq 0. \]

Bansal and Lehmann (1997) featured the maximal growth portfolio, that is the portfolio of returns that attains the right-hand side of (20). When \( \log \tilde{H}_{t+1} - \log \tilde{H}_t \) is identically zero, \( R_{t,t+1}^\infty \) coincides with the return on the maximal growth portfolio. As Bansal and Lehmann (1994) noted, the Kazemi (1992) result omits permanent components to the stochastic discount factor process. Nevertheless the one-period stochastic discount factor can still be inferred from that maximal growth portfolio provided that an econometrician has a sufficiently rich set of data on returns. Following Alvarez and Jermann (2005), we will use formula (19) in Section 9 when we discuss empirical methods and evidence for assessing the magnitude of the martingale component to the stochastic discount factor process.

Now suppose that we change measures and perform the calculations under \( \tilde{P} \) using stochastic discount factor:

\[ \tilde{S}_t = S_t \frac{\tilde{H}_0}{\tilde{H}_t}. \]

In this case

\[ \tilde{E} \left[ \log R_{t,t+1}^\infty \mid X_t = x \right] = \tilde{E} \left[ \log \tilde{S}_t - \log \tilde{S}_{t+1} \mid X_t = x \right]. \]

To interpret this finding, consider any one-period positive return \( R_{t,t+1} \). Since

\[ \tilde{E} \left[ \left( \frac{\tilde{S}_{t+1}}{\tilde{S}_t} \right) R_{t,t+1} \mid F_t \right] = 1, \]

applying Jensen’s Inequality,

\[ \tilde{E} \left[ \log R_{t+1} \mid F_t \right] \leq \tilde{E} \left[ \log \tilde{S}_t - \log \tilde{S}_{t+1} \mid F_t \right] \tag{21} \]

By construction the martingale component of the stochastic discount factor under the \( \tilde{P} \) probability measure is degenerate. As a consequence, the holding-period return on a long-term bond \( R_{t,t+1}^\infty \) coincides with the growth optimal return.

29
8 A quantitative example

We show that a well known structural model of asset pricing proposed by Bansal and Yaron (2004) implies a prominent martingale component. This model features growth-rate predictability and stochastic volatility in the aggregate consumption process. We utilize a continuous-time Brownian information specification described in Hansen et al. (2007) that is calibrated to the consumption dynamics postulated in Bansal and Yaron (2004).

We compare the implications of using the probability measure associated with the Perron–Frobenius extraction with the original probability measure and with the risk neutral measure. For this example the Perron–Frobenius extraction gives a very similar probability measure to the risk neutral measure but substantially different from the original probability measure. More generally, our aim in this section is to show that the differences in probability measure could be substantial rather than giving a definitive conclusion that they are. The latter conclusion would necessitate a confrontation with direct statistical evidence, an aspect that we discuss in Section 9.

Assume the date $t$ bivariate state vector takes the form $X_t = (X_{1,t}, X_{2,t})'$. In this model, $X_{1,t}$ represents predictable components in the growth rate of the multiplicative functional, and $X_{2,t}$ captures the contribution of stochastic volatility. The dynamics of $X$ specified in (11) have parameters $\mu(x)$ and $\sigma(x)$ given by

$$\mu(x) = \bar{\mu}(x - \iota) \quad \sigma(x) = \sqrt{x_2\bar{\sigma}}$$ (22)

where

$$\bar{\mu} = \begin{bmatrix} \bar{\mu}_{11} & \bar{\mu}_{12} \\ 0 & \bar{\mu}_{22} \end{bmatrix} \quad \bar{\sigma} = \begin{bmatrix} \bar{\sigma}_1 \\ \bar{\sigma}_2 \end{bmatrix}.$$ (23)

The parameters $\bar{\sigma}_1$ and $\bar{\sigma}_2$ are $1 \times 3$ row vectors. The vector $\iota$ is specified to be the vector of means of the state variables in a stationary distribution.

Multiplicative functionals $M$ satisfy Restriction 3.2 with parameters $\beta(x)$ and $\alpha(x)$ that are given by:

$$\beta(x) = \bar{\beta}_0 + \bar{\beta}_1 \cdot (x - \iota) \quad \alpha(x) = \sqrt{x_2\bar{\alpha}}.$$. (24)

In particular, the aggregate consumption process $C$ is a multiplicative functional parameterized by $(\beta_c, \alpha_c)$. Appendix D provides details on the calculations that follow.

We use three orthogonal shocks in our parameterization. The direct consumption shock is the component of the Brownian motion $W$ that is a direct innovation to the consumption process $\log C_t$. The growth rate shock is the Brownian component that serves as the innovation to the growth rate $X_{1,t}$, while the volatility shock is the innovation to the volatility process $X_{2,t}$.

We endow the representative investor with the recursive homothetic preferences featured in Example 2.2. We impose a unitary elasticity of substitution for convenience and thus use the continuous-time counterpart to (7). The continuous-time version of these preferences was developed in Duffie and Epstein (1992) and Schroder and Skiadas (1999). The discrete-time stochastic discount
factor has the form:

$$
\frac{S_{t+1}}{S_t} = \exp(-\delta) \left( \frac{C_t}{C_{t+1}} \right) \left( \frac{(V_{t+1})^{1-\gamma}}{E[(V_{t+1})^{1-\gamma} | F_t]} \right).
$$

The third term on the right-hand side has conditional expectation equal to unity and gives rise to a martingale component to \( S \). This insight carries over to continuous time and the continuous-time version of the stochastic discount factor evolves as:

$$
d \log S_t = -\delta dt - d \log C_t + d \log \tilde{S}_t
$$

where \( \tilde{S} \) is the continuous-time counterpart to the martingale \( H \) from equation (9).

The stochastic discount factor inherits the functional form (24) with parameters \((\beta_s, \alpha_s)\) derived in the Appendix. Since the consumption process \( C \) is modeled using a permanent shock, it also contains a martingale component (see Section 6.3) and equation (25) thus does not constitute a martingale decomposition obtained from the solution of the Perron–Frobenius Problem 6.1.

For the Perron–Frobenius probability extraction, we find a solution \((\tilde{\epsilon}, \tilde{\eta})\) to the Perron–Frobenius Problem 6.1 such that

$$
S_t = \exp(\tilde{\eta}t) \tilde{e}(X_0) \tilde{H}_t \tilde{H}_0 = \tilde{S}_t \tilde{H}_t \tilde{H}_0.
$$

and \( \tilde{H} \) implies a probability measure \( \tilde{P} \) that satisfies Restriction 5.2. We defer the details of the derivation of the model, including the solution of the martingale decomposition problem in continuous time, to Appendix D.\(^{19}\) A rather different motivation for the martingale \( \tilde{S} \) comes from literature on robustness concerns and asset pricing. For instance, see Anderson et al. (2003). In this case \( \tilde{S} \) is an endogenously determined probability adjustment for potential model misspecification. Other models of ambiguity aversion based on max-min utility also induce a martingale component to the stochastic discount factor.

In what follows, we will compare the implications of using the recovered pair \((\tilde{S}, \tilde{P})\) for the computation of risk premia and inference about the dynamics of the underlying state vector and consumption growth.

### 8.1 Alternative probability measures

Structural macro-finance models allow us to construct predictions about the future distribution of macroeconomic quantities and cash flows. The recovered pair \((\tilde{S}, \tilde{P})\) will lead to incorrect inference about these distributions as long as \( \tilde{P} \) is different from the true probability measure \( P \).

We show in the appendix that the martingale \( \tilde{H} \) associated with \( \tilde{P} \) takes the form

$$
\frac{d \tilde{H}_t}{\tilde{H}_t} = \sqrt{X_{2t+1}} \cdot dW_t
$$

\(^{19}\)The ergodicity and stationarity of \( X \) under the recovered measure \( \tilde{P} \), as well as the existence of a solution for the recursive utility stochastic discount factor, can always be checked for given parameters by a direct calculation. For instance, see the calculations in Borovička et al. (2014) for details.
where $\tilde{\alpha}_h$ is a vector that depends on the parameters of the model. This implies that we can write the joint dynamics of the state vector $X = (X_1, X_2)'$ as

$$
\begin{align*}
    dX_{1t} &= \left[\tilde{\mu}_{11} (X_{1t} - \tilde{\iota}_1) + \tilde{\mu}_{12} (X_{2t} - \tilde{\iota}_2)\right] dt + \sqrt{X_{2t} \tilde{\sigma}_1} \, d\tilde{W}_t \\
    dX_{2t} &= \tilde{\mu}_{22} (X_{2t} - \tilde{\iota}_2) dt + \sqrt{X_{2t} \tilde{\sigma}_2} \, d\tilde{W}_t
\end{align*}
$$

which has the same structure as (22)–(23) with a new set of coefficients $\tilde{\mu}_{i,j}$ derived in the Appendix. The process $\tilde{W}$ is a Brownian motion under $\tilde{P}$.

A second martingale associated with the risk-neutral dynamics behaves according to

$$
\frac{dH^*_t}{H^*_t} = \sqrt{X_{2t} \alpha^*_h} \cdot dW_t
$$

where the shock exposures to $-d \log C_t$ and $d \log \hat{S}_t$ both contribute to $\alpha^*_h$. The functional form for the risk-neutral dynamics also has the same structure as (22)–(23) with a different set of coefficients.

### 8.2 Forecasts with alternative probability measures

Structural macro-finance models allow us to construct predictions about the future distribution of macroeconomic quantities and financial cash flows. Probability measures extracted from asset market data can be used to forecast the future state of the macroeconomy and play a role in the discussion of public policy. While it is widely recognized that the risk-neutral distribution is distinct from the actual probability distribution, some have argued that the risk-neutral dynamics remain interesting as a forecasting model precisely because they do embed risk adjustments.\(^{20}\) As we argued in Section 2, the risk-neutral probability only adjusts for short-term risk when interest rates vary. Comparable risk adjustments are encoded in forward measures constructed for each horizon. As we argued in Section 7, the recovered probability measure $\tilde{P}$ absorbs long-term risk adjustments and is also often distinct from the actual probability distribution. Perhaps its use in forecasting can also be motivated. In this section we compare forecasts under the alternative distributions.

Figure 1 plots the joint stationary distribution of the state vector $X$ both under $P$ (left panel) and under $\tilde{P}$ (right panel). While the distribution in the left panel is the true distribution that is consistent with time series evidence, the distribution in the right panel is the one expected to be observed by a hypothetical investor with beliefs given by $\tilde{P}$. The distribution under $\tilde{P}$ exhibits a lower mean growth rate $X_1$ and a higher conditional volatility $X_2$ than the distribution under $P$. Moreover, the adverse states are correlated; low mean growth rate states are more likely to occur jointly with high volatility states. Bidder and Smith (2013) document similar distortions in

\(^{20}\)Narayana Kocherlakota, President of the Federal Reserve Bank of Minneapolis, during a speech to the Mitsui Financial Symposium in 2012 asks and answers: “How can policymakers formulate the needed outlook for marginal net benefits? . . . I argue that policymakers can achieve better outcomes by basing their outlooks on risk-neutral probabilities derived from the prices of financial derivatives.” See Hilscher et al. (2014) for a study of public debt using risk-neutral probabilities.
Figure 1: Stationary densities for the state vector $X = (X_1, X_2)'$ under the correctly specified probability measure $P$ (left panel) and the recovered probability measure $\tilde{P}$ (right panel). The dashed line in the right panel corresponds to the outermost contour for the distribution under the risk-neutral probability measure $P^*$. The parameterization of the model is $\bar{\beta}_{c,0} = 0.0015$, $\bar{\beta}_{c,1} = 1$, $\bar{\beta}_{c,2} = 0$, $\bar{\mu}_{11} = -0.021$, $\bar{\mu}_{12} = \bar{\mu}_{21} = 0$, $\bar{\mu}_{22} = -0.013$, $\bar{\alpha}_c = [0.0078 \ 0 \ 0]'$, $\bar{\sigma}_1 = [0 \ 0.00034 \ 0]$, $\bar{\sigma}_2 = [0 \ 0 \ -0.038]$, $t_1 = 0$, $t_2 = 1$, $\delta = 0.002$, $\gamma = 10$. Parameters are calibrated to monthly frequency.

a model with robustness concerns using the martingale $\hat{S}$ from equation (25).

The black dashed line in the right panel of Figure 1 gives the outermost contour line for the joint density under the risk-neutral dynamics. The distribution under the risk-neutral probability is remarkably similar to the $\tilde{P}$ state probabilities and very different from the physical probabilities.

The dynamics of the state vector together with the increments to the Brownian motion $W$ determine the probabilities of the aggregate consumption process. The recovered probability measure $\tilde{P}$ predicts the distribution of future consumption. In Figure 2, we plot the distribution of consumption growth under $P$ (blue band with solid lines) and $\tilde{P}$ (red band with dashed lines) conditional on the initial state $X_0 = (t_1, t_2)' = (0, 1)'$, which is the mean under the $P$ distribution.

As the right panel of Figure 2 indicates, this state corresponds to a rather atypical realization under $\tilde{P}$. The red band in the left panel of Figure 2 shows that over time, the consumption growth distribution under $\tilde{P}$ will decline toward the corresponding stationary distribution, which is negatively biased and wider relative to the true consumption growth distribution captured by the blue band. The right panel of Figure 2 plots the distribution of the logarithm of consumption and indicates that the differences in the predicted distribution under $P$ and $\tilde{P}$ are quantitatively very large, reflecting the large magnitude of the martingale $\tilde{H}$. We omit the graphs making the same comparison using $P^*$ instead of $\tilde{P}$ as the implications of $\tilde{P}$ are quite similar to those of $P^*$.

The similarity between the probability measures $\tilde{P}$ and $P^*$ emerges because the martingale component is known to dominate the behavior of the stochastic discount factor. See Hansen (2012)
and Backus et al. (2014) for evidence to this effect. Consider the extreme case in which the stochastic discount factor implies that the Perron–Frobenius eigenfunction is constant and the associated martingale implies that under the probability measure \( \tilde{P} \) the process \( X \) is ergodic. In this case \( \tilde{P} = P^* \). In this extreme case, the short-term interest rate is constant over time and the term structure is flat, neither of which is literally true for our parameterized recursive utility model. The martingale component is sufficiently dominant, however, to imply that \( \tilde{P} \) and \( P^* \) are close.

### 8.3 Asset pricing implications

Although both recovered pairs \((S, P)\) and \((\tilde{S}, \tilde{P})\) explain the same asset prices, the implications for the yields and holding period returns depend on the particular choice of this pair. In Section 7, we have shown that under \( \tilde{P} \), yields on risky cash flows in excess of the riskless benchmark converge to zero as the maturity of these cash flows increases. Since under \( \bar{P} \), the stochastic discount factor satisfies Restriction 4.3, equation (21) also shows that the expected logarithmic holding period return on the long-term bond is the highest such return in the economy when computed under \( \tilde{P} \).

The top left panel in Figure 3 plots the yields (18) on the payoff from the aggregate consumption process \( C_t \) as a function of the maturity of the payoff. The blue solid lines represent the quartiles of the yield distribution \( \gamma_t[C](x) \) corresponding to the stationary distribution of \( X_0 = x \) computed under \( P \). The red dashed lines represent the yields \( \bar{\gamma}_t[C](x) \) inferred by an investor who uses the recovered measure \( \tilde{P} \) to compute expected payoffs; but the distribution of these yields is plotted...
Figure 3: Yields and holding period returns under the true and recovered probability measure. The top row plots the annualized yields on cash flows corresponding to the aggregate consumption process (left panel) and on bonds (right panel) with different maturities. The blue bands with solid lines correspond to the distribution under $P$, while the red bands with dashed lines to the distribution under $\tilde{P}$. The lines represent quartiles of the distribution. The bottom row plots the median conditional expected logarithm of the one-year holding period returns computed under $P$ (left panel) and $\tilde{P}$ (right panel). The solid lines are returns for the consumption cash flows, dashed lines for the bonds, and the green dotted lines correspond to the maximum holding period return in the economy. The parameterization is as in Figure 1.

under the correct probability measure $P$ for the current state $X_0 = x$. The top right panel plots the corresponding bond yields.

Because the consumption process is negatively correlated with the martingale $\bar{H}$, the yields computed under $\bar{P}$ are downward biased. This bias becomes particularly large for long maturities but in fact, across all maturities, the yields on the consumption cash flows and on bonds are very close to each other. This is necessarily true by construction for long maturities as we showed in Section 7.2, but is also true throughout the term structure. The Perron–Frobenius martingale component thus accounts for virtually the whole risk premium (in excess of the maturity-matched bond) associated with the cash flows from aggregate consumption at all investment horizons.

The yields on bonds with different maturities are recovered correctly, and $y_t[1](x) = \tilde{y}_t[1](x)$. 

35
Equation (18) shows that the bias in yields arises from the covariance of the martingale $\tilde{H}$ with the cash flow. Since the payoff of the bond is constant, the covariance is zero in this case. This insight however does not apply to holding period returns on bonds, as we show next.

The bottom row of Figure 3 plots the conditional expected logarithmic holding period returns on bonds, together with the maximal conditional logarithmic holding period return. While there is substantial variation in these conditional expected returns, we only plot the median under the stationary distribution.

These graphs illustrate arguments in Section 7.3. The left panel plots the returns under the correct probability measure $P$. In particular, the blue band with solid lines in the left panel represents

$$E [\log R_{t,t+1}^\tau | \mathcal{F}_t] = E [\log \Pi_{t+1,t+\tau}(C_{t+\tau}) - \log \Pi_{t,t+\tau}(C_{t+\tau}) | \mathcal{F}_t].$$

across maturities $\tau$, while the light blue band with dashed lines shows the corresponding holding period returns on bonds. The term structure of holding period returns for the aggregate consumption process is flat because, as equation (25) indicates, the process $SC$ only contains a martingale component and a deterministic drift. The green band with dotted lines captures the growth optimal return.

In the absence of the martingale component $\tilde{H}$, the return on the long-term bond, $R_{t,t+1}^\infty$, coincides with the growth optimal return. The wedge between these two returns reflects the importance of the martingale component for pricing. In Section 9 we outline more systematic ways of measuring the magnitude of the martingale component.

The bottom right panel Figure 3 shows the same holding period returns, computed under the recovered measure $\tilde{P}$. Since the associated stochastic discount factor $\tilde{S}$ does not contain a martingale component, the return on the long term bond (the limit of the red dashed line) in this case corresponds to the growth optimal return (green dotted line). Under the correct measure $P$, the average annualized holding period return on the long-term bond is 0.38%, while it rises to 3.4% when computed under the recovered measure $\tilde{P}$.

Overall, the differences across these returns on different assets computed under $\tilde{P}$ are minimal, indicating again that absorbing the martingale $\tilde{H}$ into the probability measure leads to a substantial underestimation of the risk premia. Comparing the bottom two graphs in Figure 3, we observe very different consequences of the risk-return tradeoff inferred under the recovered model $\tilde{P}$.

Alvarez and Jermann (2005) and others have used finite-horizon bonds to approximate the return $R_{t,t+1}^\infty$ on the asymptotic bond. The bottom left graph shows that in this model, the holding period returns on bonds decline relatively quickly to their long-term counterpart, so that using a bond with maturity of 15 years as a proxy for the long-term bond would generate a quantitatively reasonable approximation.

All panels in Figure 3 then show the quantitatively large contribution of the martingale component to pricing in this model. Under the recovered model, the risk premia associated with the aggregate consumption process are negligible. The martingale component thus absorbs most of the risk adjustments, which the recovered model interprets as adverse biases in the dynamics of the
state vector, analyzed in Section 8.1.

9 Measuring the martingale component

In the previous section, we analyzed the pricing implications under the recovered probability measure in a well-known example of a structural asset pricing model. This model gives rise to a substantial martingale component and using this martingale to change the probability measure leads to large changes in expected returns and conditional distribution of the consumption growth. An alternative is to use a rich set of asset market data to deduce the Perron–Frobenius probability measure and then contrast that measure to the probability measure implied by historical returns and relevant conditioning information. Formal statistical methods can reveal if it is plausible that the two probability measures agree.

In this section we explore a different approach by studying what can be learned without using a complete set of asset returns with payoffs contingent on the relevant state variables. Thus we are less ambitious in terms of data requirements. Instead we consider methods for extracting evidence from asset market data about the magnitude of the martingale component in the stochastic discount factor process. We provide a unifying discussion of the literature and by so doing we add to the existing methods. This opens to door to new avenues for empirical research.

We build on the approach initiated by Hansen and Jagannathan (1991) aimed at nonparametric characterizations of stochastic discount factors without using a full set of Arrow prices. While full identification is not possible, data from financial markets remain informative. We draw on the pedagogically useful characterization of Almeida and Garcia (2013) and Hansen (2014), but adapt it to misspecified beliefs along the lines suggested in Ghosh et al. (2012) and Hansen (2014). In so doing we build on a key insight of Alvarez and Jermann (2005).

Consider strictly convex functions $\phi_{\theta}$ defined on the positive real numbers such as:

$$\phi_{\theta}(r) = \frac{1}{\theta(1 + \theta)} \left[ (r)^{1+\theta} - 1 \right]$$

for alternative choices of the parameter $\theta$. By design $\phi_{\theta}(1) = 0$ and $\phi_{\theta}'(1) = 1$. The function $\phi_{\theta}$ remains well defined for $\theta = 0$ and $\theta = -1$ by taking pointwise limits in $r$ as $\theta$ approaches these two values. Thus $\phi_{0}(r) = r \log r$ and $\phi_{-1}(r) = -\log r$. The functions $\phi_{\theta}$ are used to construct discrepancy measures between probability densities as in the work of Cressie and Read (1984).

We are interested in such measures as a way to quantify the martingale component to stochastic discount factors. Recall that

$$E \left[ \frac{\tilde{H}_{t+1}}{H_{t}} \mid X_{t} = x \right] = 1$$

and that $\tilde{H}_{t+1}/\tilde{H}_{t}$ defines a conditional density of the $\tilde{P}$ distribution relative to the $P$ distribution. This leads us to apply the discrepancy measures to $\tilde{H}_{t+1}/\tilde{H}_{t}$. 37
Since $\phi_\theta$ is strictly convex and $\phi_\theta(1) = 0$, from Jensen’s inequality:

\[
E \left[ \phi_\theta \left( \frac{H_{t+1}}{\bar{H}_t} \right) \bigg| X_t = x \right] \geq 0,
\]

with equality only when $\bar{H}_{t+1}/\bar{H}_t$ is identically one. There are three special cases that receive particular attention.

i) $\theta = 1$ in which case the implied measure of discrepancy is equal to one-half times the conditional variance of $\bar{H}_{t+1}/\bar{H}_t$;

ii) $\theta = 0$ in which case the implied measure of discrepancy is based on conditional relative entropy:

\[
E \left[ \left( \frac{\bar{H}_{t+1}}{\bar{H}_t} \right) \left( \log \bar{H}_{t+1} - \log \bar{H}_t \right) \bigg| X_t = x \right]
\]

which is the expected log-likelihood under the $\bar{P}$ probability measure.

iii) $\theta = -1$ in which case the discrepancy measure is:

\[
-E \left[ \log \bar{H}_{t+1} - \log \bar{H}_t \bigg| X_t = x \right]
\]

which the negative of the expected log-likelihood under the original probability measure.

We describe how to compute lower bounds for these discrepancy measures. We are led to the study lower bounds because we prefer not to compel an econometrician to use a full array of Arrow prices. Let $Y_{t+1}$ be a vector of asset payoffs and $Q_t$ the corresponding vector of prices. Recall the formula

\[
R_{t,t+1}^{\infty} = \exp(-\eta) \frac{e^{\mathcal{X}_{t+1}}}{e^{\mathcal{X}_t}},
\]

and thus

\[
\frac{S_{t+1}}{S_t} = \left( \frac{\bar{H}_{t+1}}{\bar{H}_t} \right) \left( \frac{1}{R_{t,t+1}^{\infty}} \right). \tag{26}
\]

As in Alvarez and Jermann (2005), suppose that the limiting holding-period return $R_{t,t+1}^{\infty}$ can be well approximated. In this case, one could test directly for the absence of the martingale component by assessing whether

\[
E \left[ \left( \frac{1}{R_{t,t+1}^{\infty}} \right) (Y_{t+1}) \bigg| X_t = x \right] = (Q_t)' .
\]

since in this case:

\[
\frac{S_{t+1}}{S_t} = \left( \frac{1}{R_{t,t+1}^{\infty}} \right)
\]

prices assets correctly.
More generally, we express the pricing restrictions as

$$E \left[ \left( \frac{\tilde{H}_{t+1}}{H_t} \right) \left( \frac{1}{R_{t,t+1}^{\infty}} \right) (Y_{t+1})' | X_t = x \right] = (Q_t)'$$

where $\tilde{H}$ is now treated as unobservable to an econometrician. To bound a discrepancy measure, let a random variable $J_{t+1}$ be a potential specification for the martingale increment:

$$J_{t+1} = \frac{\tilde{H}_{t+1}}{H_t}.$$

Solve

$$\lambda_\theta(x) = \inf_{J_{t+1} > 0} E [\phi_\theta (J_{t+1}) | X_t = x]$$

subject to the linear constraints:

$$E [J_{t+1} | X_t = x] - 1 = 0$$

$$E \left[ J_{t+1} \left( \frac{1}{R_{t,t+1}^{\infty}} \right) (Y_{t+1})' | X_t = x \right] - (Q_t)' = 0.$$

A strictly positive $\lambda_\theta(x)$ implies a nontrivial martingale component to the stochastic discount factor.

For the limiting continuous-time diffusion case, the choice of $\theta$ is irrelevant. Suppose that

$$d \log \tilde{H}_t = -\frac{1}{2} |\tilde{\alpha}(X_t)|^2 dt + \tilde{\alpha}(X_t) \cdot dW_t.$$

As a consequence,

$$\frac{d\tilde{H}_t}{\tilde{H}_t} = \tilde{\alpha}(X_t) \cdot dW_t.$$

Thus the local mean of $\left( \tilde{H} \right)^{\theta+1}$ is

$$\frac{\theta(\theta + 1)}{2} \left( \tilde{H}_t \right)^{\theta+1} |\tilde{\alpha}(X_t)|^2$$

and the discrepancy measure is $\frac{1}{2} |\tilde{\alpha}(X_t)|^2$ independent of $\theta$. The discrepancies for all values of $\theta$ are equal to one-half times the local variance of $\log \tilde{H}$. This equivalence of the discrepancy measures is special to the continuous-time diffusion model, however.

To compute $\lambda_\theta$ in practice requires that we estimate conditional distributions. There is an unconditional counterpart to these calculations obtained by solving:

$$\bar{\lambda}_\theta = \inf_{J_{t+1} > 0} E\phi_\theta (J_{t+1})$$

(27)
subject to:

\[ E[J_{t+1}] - 1 = 0 \]

\[ E\left[J_{t+1} \left( \frac{1}{R_{t,t+1}} \right) (Y_{t+1})' - (Q_t)' \right] = 0. \]

This bound, while more tractable, is weaker in the sense that \( \lambda_\theta \leq E\lambda_\theta(X_t) \). To guarantee a solution to optimization problem (27) it is sometimes convenient to include random variables \( J_{t+1} \) that are zero with positive probability. Since the aim is to produce bounds, this augmentation can be justified for mathematical and computational convenience. Although this problem optimizes over an infinite-dimensional family of random variables \( J_{t+1} \), the dual problem that optimizes over the Lagrange multipliers associated with the pricing constraint (28) is often quite tractable. See Hansen et al. (1995) and the literature on implementing generalized empirical likelihood methods for further discussion.

For the case in which \( \theta = 1 \), Hansen and Jagannathan (1991) study a mathematically equivalent problem by constructing volatility bounds for stochastic discount factors and deduce quasi-analytical formulas for the solution obtained when ignoring the restriction that stochastic discount factors should be nonnegative. Bakshi and Chabi-Yo (2012) apply the latter methods to obtain \( \theta = 1 \) bounds (volatility bounds) for the martingale component to the stochastic discount factor process. Similarly, Bansal and Lehmann (1997) study bounds on the stochastic discount factor process for the case in which \( \theta = -1 \) and show the connection with a maximum growth rate portfolio. Alvarez and Jermann (2005) apply these methods to produce the corresponding bounds for the martingale component to the stochastic discount factor process. Both Alvarez and Jermann (2005) and Bakshi and Chabi-Yo (2012) exploit equation (26) and approximate the return \( R_{t,t+1}^\infty \) in order to target their analysis to the martingale component. These papers provide empirical evidence in support of a substantial martingale component to the stochastic discount factor process. Bakshi and Chabi-Yo (2012) summarize results from both papers in their Table 1 and contrast differences in the \( \theta = 1 \) and \( \theta = -1 \) discrepancy measures. To our knowledge, the \( \theta = 0 \) discrepancy measure has not been used to quantify the magnitude of the martingale component to the stochastic discount factor processes.

**10 Conclusion**

Perron–Frobenius theory applied to Arrow prices identifies a martingale component to the market-determined stochastic discount factor process. This martingale component defines a distorted probability measure that absorbs long-term risk adjustments, in the same spirit as the risk-neutral probability measure absorbs one-period risk adjustments. Using asset prices observed at a single point in time, neither the stochastic discount factor nor its martingale component can be uniquely identified without further restrictions.

One identifying assumption, featured by Ross (2013), assumes that this martingale component
is identical to one. In this case, the probability measure recovered by the Perron–Frobenius theory coincides with the physical probability measure. If, however, the stochastic discount factor process includes a martingale component, then the use of the Perron–Frobenius eigenvalue and function recovers a long-term pricing measure that is distorted by this martingale component.\footnote{Recent working papers by Qin and Linetsky (2014a,b) provide additional results in a the continuous state space environment, with explicit connections to previous results in Hansen and Scheinkman (2009) and Hansen and Scheinkman (2014).}

Many structural models of asset pricing that are motivated by empirical evidence have non-trivial martingale components in the stochastic discount factors. These martingales characterize what probability is actually recovered by application of Perron–Frobenius theory. We illustrated this outcome in one such example economy in a model with long-run risk components to the macroeconomy when investors have non-separable recursive preferences. We also provided a unifying discussion of the empirical literature that derives non-parametric bounds for the magnitude of the martingale component and also finds a quantitatively large role of this martingale component for valuation. Finally we suggested ways how to further expand the set of testable implications in this literature.

In our previous work we showed how Perron–Frobenius theory helps us understand risk-return tradeoffs. This probability measure identified by Perron–Frobenius theory absorbs the long-term risk prices. Its naive use can distort the risk-return tradeoff in unintended ways, particularly at long investment horizons. The stochastic evolution of the macroeconomy is also altered. One might argue, however, that the dynamics under this probability measure are of interest precisely because this measure adjusts for the long-term riskiness of the macroeconomy. While we see value in using this probability measure prospectively, our analysis makes clear that the resulting forecasts are slanted in a particular but substantively interesting way.

The recovered probability measure provides insights into long-term pricing. The valuation operators we used in this paper applied to functions of a stationary Markov process. Hansen and Scheinkman (2009) and Hansen (2012) use Perron–Frobenius methods to study more general valuation problems in which cash flows can grow stochastically over time. We made reference to some of this analysis in Sections 7 and 8. These extensions are critical for many macro-finance applications because, empirically, many macro time series display stochastic growth.

The value of these methods extends beyond narrowly defined asset valuation. For instance, Hansen et al. (1999) and Alvarez and Jermann (2004) have shown how to use asset pricing methods for stochastically growing processes to measure the welfare cost of economic fluctuations. The probability measures recovered using the Perron–Frobenius theory could shed insights into these costs in economies in which shocks have long-term macroeconomic consequences.

Finally, long-term valuation is only a component to a more systematic study of pricing implications over alternative investment horizons. Recent work by Borovička et al. (2011) and Borovička et al. (2014) deduces methods that extend impulse response functions to characterize the pricing of exposures to shocks to stochastically growing cash flows over alternative investment horizons.
Appendix

A Multiplicative functional

The construct of a multiplicative functional is used elsewhere in the probability literature and in Hansen and Scheinkman (2009). Formally a multiplicative functional is a process $M$ that is adapted to $\mathcal{F}$, initialized at $M_0 = 1$ and, with a slight abuse of notation:

$$M_t(Y) = M_\tau(Y)M_{t-\tau}(\theta_\tau(Y)).$$

(29)

In formula (29), $\theta_\tau$ is the shift operator that moves the time subscript of $Y$ by $\tau$, that is, $(\theta_\tau(Y))_s = Y_{\tau+s}$.

We generalize this construct by building an extended multiplicative functional. The process $M$ is an extended multiplicative functional if $M_0$ is a strictly positive (Borel measurable) function of $X_0$ and $\{M_t/M_0 : t \in T\}$ is a multiplicative functional. This allows the process $M$ to be initialized at $M_0$ different from unity. In this paper we drop the use of the term extended for pedagogical convenience.

B Perron–Frobenius theory

Proof of Proposition 5.3. Write

$$S^H = \exp\left(-\delta t\right) \frac{\hat{m}(X_t)}{\hat{m}(X_0)},$$

for a positive function $\hat{m}$. Thus

$$\frac{H_t}{H_0} = \exp\left[-\left(\delta - \hat{\delta}\right) t\right] \frac{m(X_t)}{m(X_0)} = \exp(-\eta t) \frac{k(X_t)}{k(X_0)},$$

where $\eta = \delta - \hat{\delta}$ and $k = m/\hat{m}$. In what follows we consider the discrete-time case. The continuous-time case uses an identical approach, with the obvious changes.

First note that for a bounded function $f$ the Law of Large Numbers implies that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E^H[f(X_t) \mid X_0 = x] = E^H[f(X_0)]$$

(30)

and the same relationship holds also under the measure $P$. Consider three cases. First suppose that $\eta < 0$. Since $H$ is a martingale,

$$E[k(X_t) \mid X_0 = x] = \exp(\eta t)k(x)$$

(31)

Form

$$\hat{k}(x) = \min\{1, k(x)\} > 0, \text{ for all } x.$$

Since $\eta < 0$, the right-hand side of (31) converges to zero for each $x$ as $t \to \infty$. Thus,

$$0 = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E[k(X_t) \mid X_0 = x] \geq \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E\left[\hat{k}(X_t) \mid X_0 = x\right] = E\left[\hat{k}(X_0)\right] > 0.$$

Thus we have established a contradiction.

42
Next suppose that $\eta > 0$. Note that

$$E^H \left[ \frac{1}{k(X_t)} | X_0 = x \right] = \exp(-\eta t) \frac{1}{k(x)}.$$  \hspace{1cm} (32)

Form \[ \tilde{k}(x) = \min\left\{ 1, \frac{1}{k(x)} \right\} > 0, \text{ for all } x. \]

Since $\eta > 0$, the right-hand side of (32) converges to zero for each $x$ as $t \to \infty$. Thus,

$$0 = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E^H \left[ \frac{1}{k(X_t)} | X_0 = x \right] \geq \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E^H \left[ k(X_t) | X_0 = x \right] = E^H \left[ \tilde{k}(X_0) \right] > 0.$$

We have again established a contradiction.

Finally, suppose $\eta = 0$. Then

$$E^H \left[ \frac{1}{k(X_t)} | X_0 = x \right] = \frac{1}{k(x)}$$

for all $x$. From (30):

$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E^H \left[ k^n(X_t) | X_0 = x \right] = E^H k^n(X_0).$$

for $k^n = \min\{1/k, n\}$. The same equality applies to the limit as $n \to \infty$ whereby $k^n$ is replaced by $1/k$. Consequently, $E^H \left[ \frac{1}{k(X_0)} \right] = \frac{1}{k(x)}$ for almost all $x$. Thus $k(x)$ is a constant, and $H_t/H_0 \equiv 1$ with probability one. \hfill \square

C  Long-term valuation limits

First we verify that the approximation results described in Section 7.1 hold under weaker conditions than Restriction 7.1. We assume instead:

$$\liminf_{t \to \infty} \tilde{E} \left[ \psi(X_t) \mid X_0 = x \right] > 0$$ \hspace{1cm} (33)

and that

$$\liminf_{t \to \infty} \tilde{E} \left[ \frac{\psi(X_t)}{e(X_t)} \mid X_0 = x \right] > 0.$$

Notice that these assumptions follows if for every bounded function $f$

$$\lim_{t \to \infty} \tilde{E} \left[ f(X_t) \mid X_0 = x \right] = \tilde{E} \left[ f(X_0) \right]$$ \hspace{1cm} (34)

almost surely. Meyn and Tweedie (2009) establish in Theorem 13.3.3 on page 327 that (34) holds for bounded functions $f$, provided $X$ is aperiodic and positive Harris recurrent under the measure $\tilde{P}$.

Furthermore, we assume that

$$\tilde{E} \left[ \frac{\psi(X_t)}{e(X_t)} \right] < \infty, \quad \tilde{E} \left[ \psi(X_t) \right] < +\infty,$$
and that $X$ satisfies Restriction 5.2 under $\tilde{P}$. Then,
\[
\frac{1}{N} \log \tilde{E}[\psi(X_N) \mid X_0 = x] \leq \frac{1}{N} \log \left( \tilde{E} \left[ \sum_{t=1}^{N} \psi(X_t) \mid X_0 = x \right] \right) = \frac{1}{N} \log N + \frac{1}{N} \log \left( \tilde{E} \left[ \frac{1}{N} \sum_{t=1}^{N} \psi(X_t) \mid X_0 = x \right] \right).
\]

Result (33) implies that the lim inf on the left-hand side converges to zero. The lim sup of the left-hand side also converges to zero. To verify this, note that the Law of Large Numbers and the resulting almost sure convergence extend to the time-series averages of conditional expectations,
\[
\lim_{N \to \infty} \tilde{E} \left[ \frac{1}{N} \sum_{t=1}^{N} \psi(X_t) \mid X_0 = x \right] = \tilde{E}[\psi(X_0)]
\]
almost surely. Hence both terms on the right-hand side of (35) converge to zero with probability one. Consequently, lim sup on the left-hand side does as well. Given both the lim inf and lim sup on the left-hand side of (35) to zero the left-hand side must converge to zero almost surely.

The same logic implies that
\[
\lim_{N \to \infty} \frac{1}{N} \log \tilde{E} \left[ \frac{\psi(X_N)}{e(X_N)} \mid X_0 = x \right] = 0,
\]
with probability one.22

Next introduce stochastic growth into the analysis as in Section 7.2.
\[
E \left( S_l G_l \left[ \frac{e(X_l)}{e(X_0)} \right] \mid X_0 = x \right) = \exp(t\eta^*) E^* \left[ \frac{e(X_l)e^*(X_0)}{e^*(X_l)} \mid X_0 = x \right]
\]
where $E^*$ is constructed using the martingale $H^*$ with increments:
\[
\frac{H^*_t}{H^*_0} = \exp(-\eta^*) S_l G_l \left[ \frac{e^*(X_l)}{e^*(X_0)} \right].
\]
Similarly, write
\[
E \left( S_l G_l \mid X_0 = x \right) = \exp(t\eta^*) E^* \left( \left[ \frac{e^*(X_0)}{e^*(X_l)} \right] \mid X_0 = x \right)
\]
Suppose that $H^*$ induces a probability measure under which Restriction 7.1 is satisfied. In addition suppose that
\[
E^* \left[ \frac{e(X_l)}{e^*(X_l)} \right] < \infty, \quad E^* \left[ \frac{1}{e^*(X_l)} \right] < \infty.
\]

22 For the continuous-time case, sample the Markov process at integer points in time. The process will remain stationary under $P$ but not necessary ergodic. The Law of Large Numbers still applies but with a limit that is the expectation conditioned on invariant events. The previous argument with these modifications establishes the limiting behavior.
From Restriction 7.1 we obtain,
\[
\lim_{t \to \infty} \frac{1}{t} E^* \left[ \frac{\tilde{e}(X_t)}{e^*(X_t)} \mid X_0 = x \right] = 0
\]
\[
\lim_{t \to \infty} \frac{1}{t} \log E^* \left[ \frac{1}{e^*(X_t)} \mid X_0 = x \right] = 0.
\]

Thus
\[
\lim_{t \to \infty} \frac{1}{t} \log E \left( S_t G_t \left[ \frac{\tilde{e}(X_t)}{\tilde{e}(X_0)} \right] \mid X_0 = x \right) =
\]
\[
= \eta^* + \lim_{t \to \infty} \frac{1}{t} \left[ \log e^*(x) - \log \tilde{e}(x) \right] + \lim_{t \to \infty} \frac{1}{t} E^* \left[ \frac{\tilde{e}(X_t)}{e^*(X_t)} \mid X_0 = x \right] = \eta^*,
\]
and similarly
\[
\lim_{t \to \infty} \frac{1}{t} \log E (S_t G_t \mid X_0 = x) =
\]
\[
= \eta^* + \lim_{t \to \infty} \frac{1}{t} \log e^*(x) + \lim_{t \to \infty} \frac{1}{t} \log E^* \left[ \frac{1}{e^*(X_t)} \mid X_0 = x \right] = \eta^*.
\]

D Derivations for the model with predictable consumption dynamics

In this section, we provide derivations for the model analyzed in Section 8. We will focus on the analysis of the Perron–Frobenius problem. A complete analysis of the model can be found in Hansen (2012) and in the appendix in Borovička et al. (2014).

D.1 Martingale decomposition

We solve the Perron–Frobenius problem
\[
E \left[ S_t \tilde{e} (X_t) \mid X_0 = x \right] = \exp (\eta t) \tilde{e} (x)
\]
where \( S \) is a multiplicative functional parameterized by coefficients \((\beta_s(x), \alpha_s(x))\). Since the problem holds for every \( t \), then it also holds in the limit (as long as it exists)
\[
\lim_{t \to 0} \frac{1}{t} \left[ E \left[ S_t \tilde{e} (X_t) \mid X_0 = x \right] - \exp (\eta t) \tilde{e} (x) \right] = 0.
\]
The limit yields the partial differential equation
\[
Se = \eta e
\]
where, in the general Brownian information setup given by equations (11)–(12), the infinitesimal generator \( S \) is given by
\[
S \tilde{e} = \left( \beta_s + \frac{1}{2} |\alpha_s|^2 \right) \tilde{e} + \tilde{e}_x \cdot (\mu + \sigma \alpha_s) + \frac{1}{2} \text{tr} [\tilde{e}_x \sigma \sigma'].
\]
Equation (D.1) is therefore a second-order partial differential equation, and we are looking for a solution in the form of a number \( \tilde{\eta} \) and a strictly positive function \( \tilde{e} \). Hansen and Scheinkman (2009) show that if there are multiple such solutions, then only the solution associated with the lowest value of \( \tilde{\eta} \) can generate ergodic
dynamics under the implied change of measure.

In the case of the long-run risk model introduced in Section 8 and parameterized by (22)–(24), we can guess

\[ \bar{e}(x) = \exp(\bar{e}_1 x_1 + \bar{e}_2 x_2) \]

which leads to the system of equations

\[
\begin{align*}
\eta &= \bar{\beta}_{s,0} - \bar{\beta}_{s,11} - \bar{\beta}_{s,12} \bar{\nu}_2 - \bar{\nu}_1 (\bar{\mu}_{11} + \bar{\mu}_{12} \bar{\nu}_2) - \bar{\nu}_2 \bar{\mu}_{22} \\
0 &= \bar{\beta}_{s,11} + \bar{\mu}_{11} \bar{\nu}_1 \\
0 &= \bar{\beta}_{s,12} + \frac{1}{2} |\bar{\alpha}_s|^2 + \bar{\nu}_1 (\bar{\mu}_{12} + \bar{\sigma}_1 \bar{\alpha}_s) + \frac{1}{2} \bar{\nu}_2^2 |\bar{\sigma}_1|^2 + \bar{\nu}_2 (\bar{\mu}_{22} + \bar{\sigma}_2 \bar{\alpha}_s + \bar{\nu}_1 \bar{\sigma}_1 \bar{\nu}_2) + \frac{1}{2} (\bar{\nu}_2)^2 |\bar{\sigma}_2|^2 
\end{align*}
\]

The coefficients \(\bar{\beta}_s\) and \(\bar{\alpha}_s\) will be determined below. The last equation is a quadratic equation for \(\bar{\nu}_2\) and we choose the solution for \(\bar{\nu}_2\) that leads to the smaller value of \(\bar{\eta}\). From the decomposition

\[ S_t = \exp(\bar{\eta} t) \frac{\bar{e}(X_0)}{\bar{e}(X_t)} \frac{\bar{H}_t}{\bar{H}_0} \]

we can extract the martingale \(\bar{H}\):

\[ d \log \bar{H}_t = d \log S_t - \bar{\eta} dt + d \log \bar{e}(X_t) \]

and thus

\[ \frac{d \bar{H}_t}{H_t} = \sqrt{X_{2t}} (\bar{\alpha}_s + \bar{\sigma}_1 \bar{\nu}_1 + \bar{\sigma}_2 \bar{\nu}_2) \cdot d W_t \]

This martingale implies a change of measure such that \(\bar{W}\) defined as

\[ d \bar{W}_t = -\sqrt{X_{2t}} \bar{\alpha}_h dt + d W_t \]

is a Brownian motion under the new measure \(\bar{P}\). Under the change of measure implied by \(\bar{H}\), we can write the joint dynamics of the model as

\[
\begin{align*}
\begin{bmatrix} d X_{1t} \\ d X_{2t} \end{bmatrix} &= \begin{bmatrix} \tilde{\mu}_{11} (X_{1t} - \bar{\nu}_1) + \tilde{\mu}_{12} (X_{2t} - \bar{\nu}_2) \\ \tilde{\mu}_{22} (X_{2t} - \bar{\nu}_2) \end{bmatrix} dt + \sqrt{X_{2t}} \bar{\sigma}_2 d \bar{W}_t \\
\end{align*}
\]

where

\[
\begin{align*}
\tilde{\mu}_{11} &= \bar{\mu}_{11} \\
\tilde{\mu}_{12} &= \bar{\mu}_{12} + \bar{\sigma}_1 \bar{\nu}_1 \\
\tilde{\mu}_{22} &= \bar{\mu}_{22} + \bar{\sigma}_2 \bar{\nu}_1 \\
\bar{\nu}_1 &= \bar{\nu}_1 + (\bar{\mu}_{11})^{-1} (\bar{\mu}_{12} \bar{\nu}_2 - \bar{\mu}_{12} \bar{\nu}_2). \\
\end{align*}
\]

Similarly, every multiplicative functional \(M\) with parameters given by (24) can be rewritten as

\[ d \log M_t = \left[ \tilde{\beta}_0 + \tilde{\beta}_{11} (X_{1t} - \bar{\nu}_1) + \tilde{\beta}_{12} (X_{2t} - \bar{\nu}_2) \right] dt + \sqrt{X_{2t}} \bar{\alpha} \cdot d \bar{W}_t \]
where

\[
\begin{align*}
\tilde{\beta}_0 &= \tilde{\beta}_0 + \tilde{\beta}_{11} (\tilde{\iota}_1 - \iota_1) + \tilde{\beta}_{12} (\tilde{\iota}_2 - \iota_2) + (\tilde{\alpha} \cdot \tilde{\alpha}) \tilde{\iota}_2 \\
\tilde{\beta}_{11} &= \tilde{\beta}_{11} \\
\tilde{\beta}_{12} &= \tilde{\beta}_{12} + \tilde{\alpha} \cdot \tilde{\alpha}.
\end{align*}
\]

D.2 Value function and stochastic discount factor for recursive utility

We choose a convenient choice for representing continuous values. Similar to the discussion in Schroder and Skiadas (1999), we use the counterpart to discounted expected logarithmic utility.

\[
dV_t = \mu_{v,t} dt + \sigma_{v,t} \cdot dW_t
\]

The local evolution satisfies:

\[
\mu_{v,t} = \delta V_t - \delta \log C_t - \frac{1 - \gamma}{2} |\sigma_{v,t}|^2
\]

(36)

When \( \gamma = 1 \) this collapses to the discounted expected utility recursion. Let

\[
V_t = \log C_t + v(X_t)
\]

and guess that

\[
v(x) = \tilde{v}_0 + \tilde{v}_1 \cdot x_1 + \tilde{v}_2 x_2
\]

We may compute \( \mu_{v,t} \) by applying the infinitesimal generator to \( \log C + v(X) \). In addition,

\[
\sigma_{v,t} = \alpha_c (X_t) + \sigma (X_t) \frac{\partial}{\partial x} v(X_t).
\]

Substituting into (36) leads to a set of algebraic equations

\[
\begin{align*}
\delta \tilde{v}_0 &= \tilde{\beta}_{c,0} - \iota_1 (\tilde{\beta}_{c,1} + \tilde{\mu}_{11} \tilde{v}_1) - \iota_2 (\tilde{\beta}_{c,2} + \tilde{\mu}_{12} \tilde{v}_1 + \tilde{\mu}_{22} \tilde{v}_2) \\
\delta \tilde{v}_1 &= \tilde{\beta}_{c,1} + \tilde{\mu}_{11} \tilde{v}_1 \\
\delta \tilde{v}_2 &= \tilde{\beta}_{c,2} + \tilde{\mu}_{12} \tilde{v}_1 + \tilde{\mu}_{22} \tilde{v}_2 + \frac{1}{2} (1 - \gamma) |\tilde{\alpha}_c + \tilde{\alpha}_1 \tilde{v}_1 + \tilde{\alpha}_2 \tilde{v}_2|^2
\end{align*}
\]

which can be solved for the coefficients \( \tilde{v}_i \). The third equation is a quadratic equation for \( \tilde{v}_2 \) that has a real solution if and only if

\[
D = \left[ \tilde{\mu}_{22} - \delta + (1 - \gamma) (\tilde{\alpha}_c + \tilde{\alpha}_1 \tilde{v}_1) \right] \left[ \tilde{\alpha}_2 \right]^2 - 2 (1 - \gamma) |\tilde{\alpha}_2|^2 \left( \tilde{\beta}_{c,2} + \tilde{\mu}_{12} \tilde{v}_1 + \frac{1}{2} (1 - \gamma) |\tilde{\alpha}_c + \tilde{\alpha}_1 \tilde{v}_1|^2 \right) \geq 0.
\]

In particular, the solution will typically not exist for large values of \( \gamma \). If the solution exists, it is given by

\[
\tilde{v}_2 = \frac{-[\tilde{\mu}_{22} - \delta + (1 - \gamma) (\tilde{\alpha}_c + \tilde{\alpha}_1 \tilde{v}_1)] \tilde{\alpha}_2 \pm \sqrt{D}}{(1 - \gamma) |\tilde{\alpha}_2|^2}
\]

The solution with the minus sign is the one that interests us.

The resulting stochastic discount factor has two components. One that is the intertemporal marginal rate of substitution for discounted log utility and the other is a martingale constructed from the continuation.
\[ d \log S_t = -\log \delta - d \log C_t + d \log \tilde{S}_t \]

where \( \tilde{S} \) is a martingale given by
\[ \frac{d \tilde{S}_t}{S_t} = \sqrt{X_{2,t}} (1 - \gamma) (\bar{\alpha}_c + \bar{\alpha}_1 \bar{v}_1 + \bar{\alpha}_2 \bar{v}_2)' dW_t. \]

This determines the coefficients \((\beta_s(x), \alpha_s(x))\) of the stochastic discount factor.

When we choose the ‘minus’ solution in equation (D.2), then \( \tilde{S} \) implies a change of measure that preserves ergodicity. Notice that while \( \tilde{S} \) is a martingale, it is distinct from \( \tilde{H}_t/\tilde{H}_0 \) as long as the consumption process itself contains a nontrivial martingale component.

### D.3 Conditional expectations of multiplicative functionals

In order to compute asset prices and their expected returns, we need to compute conditional expectations of multiplicative functionals \( M \) parameterized by (22)–(24). These conditional expectations are given by
\[ E \left[ M_t \mid X_0 = x \right] = \exp \left[ \theta_0(t) + \theta_1(t) \cdot x_1 + \theta_2(t) x_2 \right], \]

where parameters \( \theta_i(t) \) satisfy a system of ordinary differential equations derived in Hansen (2012) and in the appendix of Borovička et al. (2014).
References

Almeida, Caio and René Garcia. 2013. Robust Economic Implications of Nonlinear Pricing Kernels.


