Long-term Yields Implied by Stochastic Discount Factor Decompositions

FERNANDO CORDEIRO*
CAIO ALMEIDA*

Abstract
We use the framework developed by Christensen (2017) and Hansen and Scheinkman (2009) to study the long-term interest rates in the US and Brazil. We apply a nonparametric estimator to US and Brazilian data to identify how the yield of a long-term zero-coupon bond responds to the initial state of the economy. Using a flexible specification for the state process leads to an interesting non-linear response of the yield to changes in the initial state. As a by-product of our work, we assess the performance of Christensen’s estimator using Monte Carlo simulations based on two widely adopted asset pricing models (rare disasters and habit formation). Adopting US and Brazilian data on aggregate consumption and dividends, we also provide an empirical analysis based on the different asset pricing models analyzed and further including the Epstein-Zin model with unitary elasticity of substitution adopted by Christensen (2017). While long-term yields for the rare disasters and habit models appear to be too high, we identify plausible values for those yields when adopting the Epstein–Zin model, respectively of 5.7% (US) and 5.4% (Brazil) per year.

Keywords:  Nonparametric stochastic discount factors, martingales, Markov processes, asset pricing models
JEL Codes:  C1, E1, G1

1 Introduction
Understanding the trade-offs for investment decisions with payoffs in the long-run is a recurrent challenge of policy implementation. Most notably, such inquiry is salient to climate policies, public finance, and sustainable energy policies, as is pointed out by Gollier and Hammitt (2014). Even when it is clear that a long-run policy should be implemented, the question remains as of how much we should give up now to get an uncertain payoff in the distant future. Therefore, a critical aspect to be considered in this scenario is which discount rate one should

* We would like to thank the anonymous referee, Timothy Christensen and Marcelo Moreira for important comments and suggestions. The second author acknowledges financial support from CNPq. This study was partially financed by the Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001.
* Fundação Getulio Vargas, Escola Brasileira de Economia e Financas (FGV/EPGE), Rio de Janeiro, Brasil.
fernandolpcordeiro@gmail.com
use to evaluate distant and uncertain payoffs. In the past, most of the finance literature was concerned with developing and testing a theory of risk-return trade-offs in the short-run. More recently, some approaches were developed to understand and analyze the risk-return trade-offs of long-dated assets. For instance, Hansen, Heaton, and Li (2008), Lettau and Wachter (2007), and Gollier (2008) developed works along these lines. In this work, we address the long-term interest rate estimation using the theoretical framework advanced by Hansen and Scheinkman (2009). The estimation procedure we follow was devised by Christensen (2017) and allows the estimation of a long-term interest rate as well as the approximation of long-term yields.

In particular, Hansen and Scheinkman (2009) (henceforth HS) introduced an elegant operator approach to analyze the price of long-dated assets, exploring the solution of a Perron–Frobenius problem. Specifically, the authors show that the stochastic discount factor (henceforth SDF) can be decomposed into permanent and transitory components, where the permanent component is a martingale and the transitory component is related to the return of holding a discount bond of asymptotically long maturity. Moreover, they find a representation of the components using the eigenfunction and principal eigenvalue of a pricing operator. One can approximate the price of long-run state-dependent payoffs using the eigenfunction and eigenvalue. More broadly, this result is part of a research program that seeks to identify a term structure of risk premia, i.e., compensation for risk exposure at different investment horizons.

Under certain model specifications, one can find analytical solutions to study long-run risk-returns trade-offs. For instance, Hansen et al. (2008) propose a log-linear model where the authors derive a decomposition of long-run expected returns into the sum of a risk-free component, a measure of long-run exposure to risk and the price of long-run risk. Therefore, under this specific model, it is possible to study the intertemporal composition of risk prices transparently. However, for more complex models with nonlinearities in the state process, such approach becomes intractable. The strength of HS results is that they allow for a very flexible specification for the state process, assuming it is a stationary and ergodic Markov process.

More recently, Christensen (2017) introduced a nonparametric estimation of the solution to the Perron–Frobenius eigenfunction problem of HS. It does not impose any tight parametric restriction on the stochastic process governing the state variable. Therefore, there is no explicit assumption regarding the functional form of the eigenfunction. The solution to the Perron–Frobenius problem
is estimated by a sieve approach, where an infinite-dimensional functional is approximated by a low-dimensional eigenvector problem. Using this framework, a researcher can recover the time series of the permanent and transitory components and estimate the associated long-term yield.

In this work, we evaluate Christensen’s estimator performance under different asset pricing models, namely the rare disasters model from Barro (2006) and the habit model from Campbell and Cochrane (1999). The simulation exercise done by Christensen (2017) consists of a standard consumption-based model with representative consumer endowed with power-utility or recursive preferences, where the only state variable is the log-consumption growth, which follows a stationary AR(1) process. While this is a valid and interesting example, it does not correspond to the current finance literature most recent practices. In particular, it is known that such a consumption growth dynamics with power-utility is incapable of resolving the equity premium puzzle of Mehra and Prescott (1985), and that recursive preferences with IES equal to one provides a somewhat restrictive asset pricing model. Consequently, most of the empirical finance literature prefers to work with variations of the disaster or habit models, as well as the long-run risk model from Bansal and Yaron (2004). However, the long-run risk model is out of the scope of this work.\(^1\)

In order to assess the estimation quality, we derive analytical solutions for the Perron–Frobenius problem in these two models. Following, Christensen (2017) we represent the bias and RMSE using the \(L^2\) norm with respect to the stationary measure of the state process. To implement this bias measure for complex non-Gaussian cases, we employ a kernel estimation to approximate the density of the stationary measure. Generally, since the approximation theory is derived under a \(L^2\) norm, it will appear better in regions where there is more probability mass than in regions where there is less mass. Consequently, we find that for models like Barro’s rare disaster, where with a small probability a jump occurs and the state vector goes far away from its mean, the methodology requires large sample sizes to get to a reasonable approximation quality, and the approximation is not uniform. This “non-uniformity” of the quality of the estimates over different regions of the state space is to be expected from the \(L^2\) theory.\(^2\) Moreover, since the disaster shocks are asymmetric with negative

\(^1\) Bansal and Yaron (2004) clearly state that an IES greater than one is critical for achieving their results. On the other hand, Christensen’s framework for recursive preferences is developed, for now, only for the particular case of IES equal to unity.

\(^2\) If one is concerned about the estimation quality in the tails, one should construct a weighting that assigns much more mass to the tails and modify the estimator and theory so as to establish
skewness, they naturally make the approximation to have an asymmetric behavior around the mean.

Adopting US and Brazilian data on aggregate consumption and dividends, we also provide an empirical analysis based on the different asset pricing models analyzed and further including the Epstein–Zin model with unitary elasticity of substitution adopted by Christensen (2017). While long-term yields for the rare disasters and habit models appear to be too high, we identify plausible values for those yields when adopting the Epstein–Zin model, respectively of 5.7% (US) and 5.4% (Brazil) per year.

2 Related Literature

Alvarez and Jermann (2005) devise a multiplicative decomposition of the SDF into transitory and permanent component. One of the components is a martingale and, consequently, is denoted as the permanent component. The other component is referred to as the transitory component. The authors establish the existence and uniqueness of the decomposition without using the pricing operators. The main result of the paper is a lower-bound on the volatility of the permanent component, which is given by

$$\frac{E\left(\log \frac{R_{t+1}}{P_{t+1,t}}\right) - E\left(\frac{R_{t+1,\infty}}{R_{t+1,1}}\right)}{E\left(\log \frac{R_{t+1}}{R_{t+1,1}}\right) + L\left(\frac{1}{R_{t+1,1}}\right)},$$

(1)

where $R_{t+1}$ is the gross return on a generic portfolio held from $t$ to $t + 1$, and $R_{t+1,k}$ is the gross return from holding from time $t$ to time $t + 1$ a claim to one unit of the numeraire to be delivered at time $t + k$:

$$R_{t+1,k} := \frac{P_{t+1}[Y_{t+k}]}{P_{t}[Y_{t+k}]},$$

(2)

where $P_{t}[Y_{t+k}]$ is the price at time $t$ of the state-contingent payoff $Y_{t+k}$ to be paid at time $t + k$, and $L(x_{t}) := \log E[x_{t}] - E[\log x_{t}]$. Using data from the US market, they estimate the bound (1) through various approximations for the means on the bound formula. For instance, using different portfolios to estimate $R_{t+1}$ and asymptotically equivalent measures for the term premium $E(\log R_{k}/R_{1})$. For every different approach, they reject the hypothesis that the SDF has no permanent component and find the bound to be between 0.8 and 1. Then,
Alvarez and Jermann estimate an upper bound for the transitory component. For data availability reasons, these estimations are much less informative than the estimation for the permanent component. Nonetheless, the results suggest that the volatility of the transitory component is a small fraction of the volatility of the SDF itself. Moreover, the authors argue that movements in the aggregate price level have small impacts in the SDF. Consequently, permanent components are primarily driven by real variables. However, this observation is only valid when the consumer price index accurately reflects the properties of the price level faced by asset market participants.

In HS, the SDF decomposition is revisited from an operator point of view, yielding insights on the behavior of long-term prices. More precisely, risk-return trade-offs in the long-run are examined in the limiting behavior of the family of operators. The authors develop a general theory for the analysis of long-run risk-return relationship in continuous-time with a Markovian state process. Broadly, they study the valuation of cash flows that grow stochastically over time. The decomposition is characterized by the eigenfunction and eigenvalue that solve the Perron–Frobenius of a pricing operator.

In a different context, Ross (2015) also uses the Perron–Frobenius problem, but to identify (or “recover”) investor’s beliefs. In Borovička, Hansen, and Scheinkman (2016), the authors shed light on the connection between the results from Ross (2015) and HS. Specifically, it is shown that when the martingale term from the Perron–Frobenius problem is not degenerate (identically equal to one), the recovered probabilities differ from subjective probabilities. Hence, such assuming the two types of probabilities are identical can lead to invalid inferences.

3 The Theory

In this section, we present the most relevant results on the SDF decomposition and the nonparametric estimation. Most of the content is very similar to the presentation given by Christensen (2017), whenever possible the notation is the same.

3.1 Theoretical Setup

In the model, the economy evolves in discrete time $t \in \{1,2,\ldots\} := T$. The state of the economy is described by a stochastic process $X$. Formally, consider a probability space $(\Omega,\mathcal{F},\mathbb{P})$ on which there is a time homogeneous, strictly stationary and ergodic Markov process $X = \{X_t:t \in T\}$, which takes values in
In absence of arbitrage opportunities, there exists a positive stochastic discount factor process $M = \{M_t; t \in T\}$ which assigns prices to state contingent claims. Hence, the price at time $t$ of a random payoff $\psi(X_{t+k})$ at date $t+k$ is given by

$$P_t[\psi(X_{t+k})] = \mathbb{E}\left[ \frac{M_{t+k}}{M_t} \psi(X_{t+k}) \mid \mathcal{I}_t \right],$$

where $\mathcal{I}_t$ represents all the information available to the investor up to date $t$. Assuming that only payoffs depending on future values of the state are considered, and trading at intermediate dates is allowed, then the SDF is a positive multiplicative functional of $X$. This implies that one can use the representation $M_{t+1}/M_t = m(X_t, X_{t+1})$, for a positive function $m$. Given the Markov structure of the state process, one can define the one-period pricing operator $\mathbb{M}$ by

$$(\mathbb{M} \psi)(x) := \mathbb{E}[m(X_t, X_{t+1}) \psi(X_{t+1}) \mid X_t = x]. \quad (3)$$

Following Christensen (2017), we assume that $\mathbb{M}$ is a bounded linear operator on the Hilbert space $L^2 := \{ \psi: \mathcal{X} \to \mathbb{R} \text{ such that } \int \psi^2 dQ < \infty \}$. The Perron–Frobenius problem introduced by HS is to find a function $\phi$ and scalar $\rho$ such that

$$\mathbb{M} \phi = \rho \phi, \quad (4)$$

where the eigenvalue $\rho$ is a positive real number that is equal to the spectral radius of $\mathbb{M}$, and the eigenfunction $\phi$ is positive $Q$-almost everywhere. The existence of such solution is an extension of the Krien–Rutman theorem, which also states that the adjoint of $\mathbb{M}$, $\mathbb{M}^*$, has $\rho$ as an eigenvalue and positive eigenvector $\phi^*$:

$$\mathbb{M}^* \phi^* = \rho \phi^*. \quad (5)$$

A fascinating fact is that $\mathbb{M}^*$ can be characterized using the time-reversed Markov process:

$$(\mathbb{M}^* \psi)(x) := \mathbb{E}[m(X_t, X_{t+1}) \psi(X_t) \mid X_{t+1} = x],$$

Next, HS define the components

$$\frac{M_{t+k}^P}{M_t^P} := \rho^{-k} \frac{M_{t+k}}{M_t} \frac{\phi(X_{t+k})}{\phi(X_t)} \quad \text{and} \quad \frac{M_{t+k}^T}{M_t^T} := \rho^k \frac{\phi(X_t)}{\phi(X_{t+k})}. \quad (6)$$

---

3 See page 182 of HS

4 See, e.g., Rosenblatt (1971). On the next section we provide a small motivation to this fact by exploring the case when $\mathcal{X}$ is finite.
Moreover, the authors show that there might exist multiple solutions to the Perron–Frobenius problem. However, only one solution is suitable for the long-term analysis. Specifically, this solution has a martingale term that induces a change of measure under which \( X \) is stochastically stable. Under certain regularity conditions, one can find the following approximation:

\[
\lim_{k \to \infty} \rho^{-k} M_k \psi(x) = \tilde{E} \left[ \frac{\psi(X_t)}{\phi(X_t)} \right] \phi(x),
\]

where the expectation \( \tilde{E}[\cdot] \) is taken under a certain measure \( \tilde{Q} \), which is absolutely continuous with respect to \( Q \). Christensen (2017) establishes further conditions under which \( \tilde{Q} \) is characterized by the change of measure \( d\tilde{Q}/dQ = \phi \phi^* \). In addition, \( y = -\log \rho \) can be interpreted as the long-run yield. Another quantity of interest is the entropy of the permanent component, which is given by

\[
L := \log \mathbb{E} [ M_{t+1}^P / M_t^P ] - \mathbb{E} [ \log ( M_{t+1}^P / M_t^P ) ].
\]

Alvarez and Jermann (2005) show that \( L \) is bounded below by the expected excess return of any traded asset relative to a discount bond of asymptotical maturity. Following Christensen (2017), \( L \) is estimated as

\[
\hat{L} = \log \hat{\rho} - \frac{1}{n} \sum_{t=0}^{n-1} \log m(X_t, X_{t+1}).
\]

3.2 Example: Gaussian AR(1)

Consider an economy consisting of a representative agent who maximizes a time-additive utility function \( \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(C_t) \), where \( u(C) := C^{1-\gamma} / (1 - \gamma) \), i.e. power-utility. The state variable is log consumption growth \( X_t = g_t := \log(C_t / C_{t-1}) \), which evolves as a Gaussian AR(1) process:

\[
g_{t+1} - \mu = \kappa (g_t - \mu) + \sigma e_{t+1}, \quad e_t \overset{\text{iid}}{\sim} \mathcal{N}(0, 1),
\]

where \( |\kappa| < 1 \). The SDF is given by

\[
m(X_t, X_{t+1}) = \beta (C_{t+1} / C_t)^{-\gamma} = \beta \exp(-\gamma g_{t+1}).
\]

Thus, a solution to the Perron–Frobenius problem should be of the form

\[
(M \phi)(x) = \mathbb{E} \left[ \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \phi(g_{t+1}) \mid g_t = x \right] = \rho \phi(x).
\]
Let’s use an undetermined coefficients approach to find the solution. Consider a solution of the form \( \phi(x) = \exp(ax) \). For the sake of exposition, we use the notation \( E_t \) to denote the expectation conditional on \( g_t \). Then, equation (9) becomes

\[
E_t[\beta \exp(-\gamma g_{t+1}) \exp(a g_{t+1})] = \rho \exp(a g_t)
\]
\[
E_t[\exp((a - \gamma) g_{t+1})] = \beta^{-1} \rho \exp(a g_t).
\]

Now, using the fact that \( g_{t+1} \mid g_t \sim \mathcal{N}(\mu(1 - \kappa) + \kappa g_t, \sigma^2) \) we get

\[
E_t[\exp((a - \gamma) g_{t+1})] = \exp\left((a - \gamma)\mu(1 - \kappa) + (a - \gamma)\kappa g_t + \frac{1}{2} \sigma^2 (a - \gamma)^2\right).
\]

Thus, the solution must satisfy \( (a - \gamma) \kappa = a \), which implies \( a = -\gamma \kappa / (1 - \kappa) \).

Using the expression in the last equation yields

\[
E_t[\exp((a - \gamma) g_{t+1})] = \exp\left(-\gamma \mu - \frac{\gamma \kappa}{1 - \kappa} g_t + \frac{1}{2} \frac{\sigma^2 \gamma^2}{(1 - \kappa)^2}\right)
\]
\[
= \exp\left(-\gamma \mu + \frac{1}{2} \frac{\sigma^2 \gamma^2}{(1 - \kappa)^2}\right) \phi(g_t).
\]

Therefore, we have \( \rho = \beta \exp\left(-\gamma \mu + \frac{1}{2} \frac{\sigma^2 \gamma^2}{(1 - \kappa)^2}\right) \). Now let’s find \( \phi^* \) using the time-reversed pricing operator. Let \( X^* = \{g_t^*: 0 \leq t \leq m\} \) denote the state process with time running backwards, starting from a distant date \( m \). In this case we can write

\[
g_t^* - \mu = \kappa (g_{t+1}^* - \mu) + \sigma e_t^* , \quad e_t^* \overset{iid}{\sim} \mathcal{N}(0, 1).
\]

So the problem is

\[
\mathbf{M}^* \phi^* (g_{t+1}^*) = \mathbb{E} [\beta \exp(-\gamma g_{t+1}^*) \phi(g_t^*) \mid g_{t+1}^*] = \rho \phi(g_{t+1}^*).
\]

Again we conjecture a solution of the form \( \phi^*(x) = \exp(a^* x) \)

\[
E_{t+1}[\beta \exp(-\gamma g_{t+1}^*) \exp(a^* g_t)] = \rho \exp(a^* g_{t+1}^*)
\]
\[
E_{t+1}[\exp(a^* g_t)] = \beta^{-1} \rho \exp((a^* + \gamma) g_{t+1}^*).
\]
Now, using the fact that $g_t^* | g_{t+1}^* \sim N(\mu(1 - \kappa) + \kappa g_{t+1}^*, \sigma^2)$ we get

$$E_{t+1}[\exp(a^* g_t)] = \exp\left(a^* \mu(1 - \kappa) + a^* \kappa g_{t+1}^* + \frac{\sigma^2}{2} a^* \right).$$

Thus, a solution should satisfy $a^* \kappa = a^* + \gamma$, which implies $a^* = -\gamma / (1 - \kappa)$. Replacing the expression for $a^*$ in the last equation, we get

$$E_{t+1}[\exp(a^* g_t)] = \exp\left(-\frac{\gamma \mu}{1 - \kappa} + \frac{1}{2} \frac{\sigma^2 \gamma^2}{(1 - \kappa)^2}\right) = \beta^{-1} \rho \exp\left(-\frac{\gamma \kappa}{1 - \kappa} g_{t+1}^*\right).$$

Finally, because $\phi$ and $\phi^*$ are defined up to scale, we follow Christensen (2017) and normalize them so that $E[\phi(g_t)\phi^*(g_t)] = 1$ and $\|\phi\|_2 = 1$. In this way, the eigenfunctions can be used to define a change of measure $d\bar{Q}/dQ = \phi \phi^*$. To do so, we seek a constant $b$ such that

$$\|\phi\|^2 = \int \phi(x)^2 dQ = E\left[\exp\left(2b - 2 \frac{\gamma \kappa}{1 - \kappa} g_t\right)\right] = \exp\left(2b - 2 \frac{\gamma \kappa \mu}{1 - \kappa} + \frac{2\gamma^2 \kappa^2 \sigma^2}{(1 - \kappa)^2}\right) = 1.$$

Therefore, we have $b = \frac{\gamma \kappa \mu}{1 - \kappa} - \frac{\gamma^2 \kappa^2 \sigma^2}{(1 - \kappa)^2}$. Next, we seek $b^*$ such that

$$E[\phi(g_t)\phi^*(g_t)] = E\left[\exp\left(b^* + b - \gamma \frac{1 + \kappa}{1 - \kappa} g_t\right)\right] = \exp\left(b^* + b - \gamma \frac{1 + \kappa}{1 - \kappa} \mu + \frac{1}{2} \frac{\sigma^2 \gamma^2 (1 + \kappa)^2}{(1 - \kappa)^2}\right) = \exp\left(b^* - \frac{\gamma \mu}{1 - \kappa} + \frac{\sigma^2 \gamma^2}{2(1 - \kappa)^2} \frac{1 + 2k - k^2}{1 - \kappa}\right) = 1.$$

Hence, the final solution is given by

$$\phi(x) = \exp\left(\frac{\gamma \kappa \mu}{1 - \kappa} - \frac{\gamma^2 \kappa^2 \sigma^2}{(1 - \kappa)^2(1 - \kappa^2)} - \frac{\gamma \kappa}{1 - \kappa} x\right), \tag{10}$$

$$\phi^*(x) = \exp\left(\frac{\gamma \mu}{1 - \kappa} - \frac{\sigma^2 \gamma^2 (1 + 2k - k^2)}{2(1 - \kappa^2)(1 - \kappa)^2} - \frac{\gamma}{1 - \kappa} x\right). \tag{11}$$
3.3 Example: IID

Now consider the same setting from the last example, but each $g_t$ is iid with a given distribution $W$ with enough finite moments. Then, for any function $f \in L^2$ and $x \in X$, the pricing operator satisfies

$$Mf(x) = \mathbb{E}[\beta \exp(-\gamma g_{t+1}) f(g_{t+1}) | g_t = x] = \mathbb{E}[\beta \exp(-\gamma g_{t+1}) f(g_{t+1})].$$

Notice that $Mf$ does not vary with $x$. Thus, a solution for the Perron–Frobenius problem should be of the form $\phi = \alpha \mathds{1}$, where $\alpha \in \mathbb{R}$:

$$M\phi(x) = \mathbb{E}[\beta \exp(-\gamma g_{t+1}) \alpha \mathds{1}(g_{t+1}) | g_t = x] = \rho \alpha \mathds{1}(x).$$

Hence, the eigenvalue is given by $\rho = \mathbb{E}[\beta \exp(-\gamma g_{t+1})]$. Now, for the time-reversed problem we conjecture a solution of the form $\phi^*(x) = \exp(a^* x)$:

$$\mathbb{E}_{t+1}[\beta \exp(-\gamma g_{t+1}^*) \exp(a^* g_t)] = \rho \exp(a^* g_{t+1}^*)$$
$$\mathbb{E}_{t+1}[\exp(a^* g_t)] = \beta^{-1} \rho \exp((a^* + \gamma) g_{t+1}^*)$$
$$\mathbb{E}[\exp(a^* g_t)] = \mathbb{E}[\exp(-\gamma g_{t+1})] \exp((a^* + \gamma) g_{t+1}^*).$$

Clearly, $a^* = -\gamma$ is a solution. Since $\|\phi\|_2 = 1$, we only need to rescale $\phi^*$ to obtain $\mathbb{E}[\phi \phi^*] = 1$. Therefore, we set $\phi^*(x) = \exp(b - \gamma x)$, where

$$b = -\log(\mathbb{E}[\exp(-\gamma g_t)]).$$

3.4 Nonparametric estimation

This section summarizes the estimators of the Perron–Frobenius eigenvalue and eigenfunctions proposed by Christensen (2017), where a thorough exposition is given, as well as large-sample results for these estimators. The following exposition follows the original text very closely.

The functions $\phi$ and $\phi^*$ are estimated using a sieve approach, where the solution is obtained by using the “projected operator” into a low-dimensional subspace spanned by a finite number of basis functions. Specifically, let $b_{k_1}, \ldots, b_{k_k} \in L^2$ be a sequence of linearly independent basis functions (e.g. polynomials or splines). Let $B_k \subset L^2$ be the subspace spanned by $b_{k_1}, \ldots, b_{k_k}$, and let $\Pi_k : L^2 \to B_k$ denote the orthogonal projection onto $B_k$. The projected eigenfunction problem is

$$(\Pi_k M)\phi_k = \rho_k \phi_k,$$ (12)
where $\phi_k$ is a positive function and $\rho_k$ is the largest real eigenvalue of $\Pi_k \mathbf{M}$. Because $\phi_k \in B_k$, there exists a unique $c_k \in \mathbb{R}^k$ such that $\phi_k(x) = b^k(x)'c_k$, where $b^k(x) = (b_{k_1}(x), \ldots, b_{k_k}(x))'$. Moreover, if $(\Pi_k \mathbf{M})f(x) = b^k(x)'c$, then by the projection theorem $c$ minimizes the distance
\[
\| (\mathbf{M}f) - (\Pi_k \mathbf{M})f \|_2 = \mathbb{E} \left[ (b^k(X_t) - b^k(X_t)'c)^2 \right].
\] (13)
The first order condition of this problem yields
\[
c = \mathbb{E} \left[ b^k(X_t)b^k(X_t)' \right]^{-1} \mathbb{E} \left[ b^k(X_t)(\mathbf{M}f)(X_t) \right].
\] (14)
Therefore, a solution to the Perron–Frobenius should satisfy
\[
\rho_k c_k = \mathbb{E} \left[ b^k(X_t)b^k(X_t)' \right]^{-1} \mathbb{E} \left[ b^k(X_t)(\mathbf{M})b^k(X_t)'c_k \right] = \mathbb{E} \left[ b^k(X_t)b^k(X_t)' \right]^{-1} \mathbb{E} \left[ b^k(X_t)m(X_t, X_{t+1})b^k(X_{t+1})'c_k \right].
\] (15)
Notice that it is possible to write equation (15) in matrix form
\[
\rho_k c_k = \mathbf{G}_k^{-1} \mathbf{M}_k c_k,
\] (16)
where $\mathbf{G}_k = \mathbb{E} \left[ b^k(X_t)b^k(X_t)' \right]$ and $\mathbf{M}_k = \mathbb{E} \left[ b^k(X_t)m(X_t, X_{t+1})b^k(X_{t+1})' \right]$. Since $\phi^*$ is the eigenfunction associated with the adjoint of $\mathbf{M}$, the approximate solution solution for $\phi^*$ is $b^k(x)'c^*_k$ where $c^*_k$ is the eigenvector of $\mathbf{G}_k^{-1} \mathbf{M}_k'$ associated to the eigenvalue $\rho_k$. Therefore, $(\rho_k, c_k, c^*_k)$ solve the generalized eigenvector problems
\[
\mathbf{M}_k c_k = \rho_k \mathbf{G}_k c_k \quad \mathbf{c}^*_k \mathbf{M}_k = \rho_k \mathbf{c}^*_k \mathbf{G}_k.
\] (17)
To estimate $\phi$, $\phi^*$ and $\rho$, we solve the sample counterpart of (17):
\[
\hat{\mathbf{M}} \hat{c}_k = \hat{\rho} \hat{\mathbf{G}} \hat{c}_k \quad \hat{\mathbf{c}}^*_k \hat{\mathbf{M}} = \hat{\rho} \hat{\mathbf{c}}^*_k \hat{\mathbf{G}}.
\] (18)
Given the assumptions about the stochastic process $X$ and a sample of observations $\{X_0, \ldots, X_n\}$, a natural estimator for $\mathbf{G}_k$ is
\[
\hat{\mathbf{G}}_k := \frac{1}{n} \sum_{t=0}^{n-1} b^k(X_t)b^k(X_t)'.
\] (19)

5 See, e.g., sections 3.3 and 3.10 of Luenberger (1969).
Assuming the SDF factor \( m(X_t, X_{t+1}) \) is observable, one can estimate \( M_k \) as

\[
\hat{M}_k := \frac{1}{n} \sum_{t=0}^{n-1} b^k(X_t)m(X_t, X_{t+1})b^k(X_{t+1})'.
\] (20)

To implement the estimation, we use Hermite Polynomials\(^7\) as basis functions, which can be described by \( He_0(x) = 1, He_1(x) = x \) and the recursion

\[
He_{n+1}(x) = xHe_n(x) - nHe_{n-1}(x).
\]

4 Simulations

In this section, we assess the accuracy of the nonparametric estimation proposed by Christensen (2017) in different asset pricing models. To do so, we conduct Monte Carlo simulations for an adaptation of the rare disasters model from Barro (2006) and the habit model of Campbell and Cochrane (1999). Following, Christensen (2017) we represent the bias as \( \| \hat{\phi} - \phi \|_2 \), where \( \hat{\phi} \) is the pointwise average of the estimators across simulations. To implement this bias measure to complex non-gaussian cases, we employ a kernel estimation to approximate the density of \( Q \), the stationary measure. Similarly, the RMSE is computed in the following way. For each replication \( i = 1, \ldots, N \), we estimate the \( L^2 \) distance \( \| \hat{\phi}_i - \phi \|_2^2 \), then we take \( RMSE = \left( N^{-1} \sum_{i=1}^{N} \| \hat{\phi}_i - \phi \|_2^2 \right)^{1/2} \).

4.1 Rare Disasters Model

The economic environment consists of a representative agent who maximizes a time-additive utility function \( E_0 \sum_{t=0}^{\infty} \beta^t u(C_t) \), where \( u(C) := C^{1-\gamma} / (1 - \gamma) \). Under standard assumptions, the SDF is given by \( M_{t+1} = \beta( C_{t+1} / C_t)^{1-\gamma} \). We follow Backus, Chernov, and Martin (2011) and model the logarithm of consumption growth \( g_t := \log(C_t / C_{t-1}) \) as

\[
g_{t+1} = w_{t+1} + z_{t+1},
\]

where the components \((w_t, z_t)\) are independent of each other and over time. The first component is such that \( w \sim N(\mu, \sigma^2) \), whereas the second component is conceived to model disaster jumps. Specifically, \( z_t \) follows a Poisson mixture of normals with \( z_t | j_t \sim N(j\theta, j\delta^2) \) and \( j_t \) following a Poisson distribution with parameter \( \omega \).

\(^7\) See, e.g., pages 773–782 of Abramowitz and Stegun (1964) for more details and properties
First, let’s find an analytical solution to the Perron–Frobenius problem in this setting. Because $g_t$ is iid, we know that $\phi(x) = 1(x)$. So the problem simplifies to

$$E_t[\beta \exp(-\gamma g_{t+1})] = \rho = \beta E_t[\exp(-\gamma w_{t+1})] E_t[\exp(-\gamma z_{t+1})]$$

$$= \beta \exp\left(-\gamma \mu + \frac{1}{2} \sigma^2 \gamma^2\right) E_t[\exp(-\gamma z_{t+1})].$$

The second term can be written as

$$E_t[\exp(-\gamma z_{t+1})] = \sum_{n=0}^{\infty} E_t[\exp(-\gamma z_{t+1}) | j = n] P(j = n)$$

$$= \sum_{n=0}^{\infty} \exp\left(-\gamma n \theta + \frac{1}{2} \gamma^2 n \delta^2\right) \frac{e^{-\omega \omega^2 n}}{n!}.$$

Now, for the time-reversed problem, we know from section 3.3 that $\phi^*(x) = \exp(-\gamma x)$ is a valid solution. Finally, to get the normalization $E[\phi(g_t) \phi^*(g_t)] = 1 = E[\phi^*(g_t)]$, we look for $b^*$ such that

$$E[b^* \exp(-\gamma g_t)] = b^* \exp\left(-\gamma \mu + \frac{1}{2} \sigma^2 \gamma^2\right) E[\exp(-\gamma z_{t+1})] = 1.$$

Hence, the final solution is given by

$$\phi^*(x) = \exp\left(\gamma \mu - \frac{1}{2} \sigma^2 \gamma^2\right) E[\exp(-\gamma z_{t+1})]^{-1} \exp(-\gamma x).$$

In order to implement the simulation, we follow Backus et al. (2011) and set the parameters as $\beta = 0.994$, $\gamma = 5.19$, $\mu = 0.023$, $\sigma = 0.01$, $\omega = 0.01$, $\theta = -0.3$, and $\delta = 0.15$. We generate 50,000 samples of length 400, 800, 1,600, and 3,200, and use a Hermite Polynomial basis of dimension $k = 6$. Choosing the smoothing parameter $k$ is not straightforward. For high values of $k$ it might be the case that the matrix $G_k$ is numerically singular. On the other hand, for small values of $k$, the approximation is poor.

Varying the jump intensity $\omega$ in the distribution of $g_t$ notably changes the estimator’s precision as shown in Figure 1. In particular, by observing Figures 1(c) and 1(d) one can notice that, as is expected with the approximation under the $L^2$ norm, the approximation’s quality is not uniform: for values distant from the stationary mean of the process, the approximation gets poorer. It is also interesting to notice that the approximation is not symmetric: the quality is
Figure 1. Simulation results for disaster model with a Hermite polynomial basis with $k = 6$. Graphs display pointwise 90% confidence intervals for $\hat{\phi}$ and $\hat{\phi}^*$ across simulations (light, medium and dark correspond to $n = 800, 1,600$ and $3,200$ respectively; and the true $\phi$ and $\phi^*$ are plotted as solid black lines).
considerably worse for points above the mean than for points below the mean. In addition, from Table 1 one can notice that the RMSE for \( \phi^* \) is significantly larger than for \( \phi \), especially for a small sample size. The estimator of the long-term interest rate \( \hat{y} \) has a small negative bias for sample size equal to or greater than 800. Likewise, the RMSE of \( \hat{y} \) is relatively small but still larger than the ones found in the simulations of Christensen (2017).

**Table 1.** Bias and RMSE found in simulations with a Hermite Polynomial sieve with \( k = 6 \).

<table>
<thead>
<tr>
<th></th>
<th>( n )</th>
<th>( \hat{\phi} )</th>
<th>( \hat{\phi}^* )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{y} )</th>
<th>( \hat{L} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td>400</td>
<td>0.0034</td>
<td>0.0163</td>
<td>0.0001</td>
<td>0.0012</td>
<td>-0.0013</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0.0040</td>
<td>0.0094</td>
<td>0.0016</td>
<td>-0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>0.0040</td>
<td>0.0112</td>
<td>0.0033</td>
<td>-0.0023</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>3200</td>
<td>0.0026</td>
<td>0.0111</td>
<td>0.0027</td>
<td>-0.0021</td>
<td>0.0021</td>
</tr>
<tr>
<td>RMSE</td>
<td>400</td>
<td>0.0720</td>
<td>1.0983</td>
<td>0.0536</td>
<td>0.0493</td>
<td>0.0437</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>0.0576</td>
<td>0.1168</td>
<td>0.0522</td>
<td>0.0461</td>
<td>0.0430</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>0.0489</td>
<td>0.0909</td>
<td>0.0531</td>
<td>0.0450</td>
<td>0.0434</td>
</tr>
<tr>
<td></td>
<td>3200</td>
<td>0.0369</td>
<td>0.0922</td>
<td>0.0439</td>
<td>0.0374</td>
<td>0.0364</td>
</tr>
</tbody>
</table>

### 4.2 Habit Model

Now consider a similar economic environment, where a representative agent maximizes a time-additive utility function 

\[
E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, H_t),
\]

with

\[
u(C, H) := \frac{(C - H)^{-\gamma} - 1}{1 - \gamma},
\]

where \( C_t \) is the level of consumption and \( H_t \) is the level of habit. In addition, the surplus consumption ratio is defined as \( S_t := (C_t - H_t)/C_t \). The level of habit \( H_t \) is determined implicitly by the dynamics of the log surplus consumption ratio, \( s_t := \log S_t \), which follows a heteroskedastic AR(1) process:

\[
s_{t+1} = (1 - \kappa)\bar{s} + \kappa s_t + \lambda(s_t)(g_{t+1} - \mu),
\]

where \( \bar{s} \) is the mean surplus consumption ratio, \( \kappa \) is the persistence of the process, and \( \lambda(s) \) is the sensitivity function. Furthermore, log consumption growth is modeled as an iid normal process:

\[
g_t := \mu + \nu_t, \quad \text{where } \nu_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2).
\]
The functional form of $\lambda$ is conceived by Campbell and Cochrane (1999) to guarantee three conditions: (1) the risk-free interest rate is constant; (2) habit is predetermined at the steady state $s_t = \bar{s}$; and (3) habit moves nonnegatively with consumption everywhere. To achieve these conditions, the authors establish $\bar{S} = \sigma \sqrt{\gamma/(1 - \kappa)}$ and propose

$$
\lambda(s_t) = \left( \frac{\sqrt{1 - 2(s_t - \bar{s})}}{S} - 1 \right) I\{s_t \leq s_{\text{max}}\},
$$

where $I\{\cdot\}$ is an indicator function and $s_{\text{max}} := \bar{s} + \frac{1 - \bar{s}^2}{2}$ is the largest value that guarantees a nonnegative $\lambda$. In this setting, the SDF is given by

$$
\frac{M_{t+1}}{M_t} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{S_{t+1}}{S_t} \right)^{-\gamma}.
$$

Let the state process be $X_t = (g_t, s_t)$. Even though the variable $s_t$ is not observable, we use it as one of the state variables for an initial exploration of the model. Constructing the state process in this way allows for tractable analytical solutions and, therefore, the computation of RMSE. The Perron–Frobenius problem in this setting is given by

$$
E_t \left[ \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{S_{t+1}}{S_t} \right)^{-\gamma} \phi(g_{t+1}, s_{t+1}) \right] = \rho \phi(g_t, s_t). \tag{21}
$$

In the Appendix A we show that one can find the analytic solutions:

$$
\phi(x, y) = \exp(\eta + \gamma y),
$$

$$
\phi^*(x, y) = \beta \rho^{-1} \exp(-\eta - \gamma(x + y)),
$$

$$
\rho = \beta \exp\left(-\gamma \mu + \frac{\sigma^2 \gamma^2}{2}\right),
$$

where $\eta = -\log(E[\exp(2\gamma s_t)])/2$. 

In order to implement the simulation, we follow Campbell and Cochrane (1999) and set the parameters as $\beta = 0.89$, $\gamma = 2$, $\mu = 1.89\%$, $\sigma = 1.5\%$ and $\kappa = 0.87$. Figure 2 shows the simulated SDF for this model. It is interesting to notice how the log surplus is negatively correlated with the SDF, which means that states with a low surplus are priced relatively higher by the consumer. Also note that the time-varying feature of the volatility of the SDF appears to have a
Figure 2. Time series generated from simulated data using Campbell and Cochrane (1999) calibration. First panel shows the simulated $g_t$, the second panel displays simulated $s_t$, and third shows the simulated SDF.

similar pattern to the SDFs obtained in the nonparametric estimation literature.\(^8\)

As in the last section, we generate 50,000 samples of length 400, 800, 1,600, and 3,200. Following Christensen (2017) we employ a sparse tensor product to construct the bivariate basis functions. First, we create univariate Hermite polynomial bases for $g_t$ and $s_t$ with $k = 7$. Then, we make a sparse tensor product basis, discarding any tensor product polynomial whose total degree is order eight or higher. In order to set the univariate sieve dimension $k$, we ran simulations for multiple values and chose the value displaying the overall smallest RMSE.\(^9\)

Inspecting Table 2, it is clear that the nonparametric estimation of this model is significantly more challenging than the rare disasters model. First, the RMSE for the eigenfunctions is an order of magnitude higher than in the univariate rare disasters model. Second, one can notice that increasing the sample size eight-fold does not lead to a substantial improvement in the

---

\(^8\) See, e.g., Almeida, Ardison, and Garcia (2018).

\(^9\) In the online appendix we show the results for simulations varying the univariate sieve dimension.
Table 2. Bias and RMSE found in simulations for the Habit model with normalization $||\phi||_2 = 1 = ||\phi^*||_2$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\hat{\phi}$</th>
<th>$\hat{\phi}^*$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{y}$</th>
<th>$\hat{L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0.3460</td>
<td>0.2724</td>
<td>0.0057</td>
<td>-0.0057</td>
<td>0.0057</td>
</tr>
<tr>
<td>800</td>
<td>0.3344</td>
<td>0.1905</td>
<td>0.0056</td>
<td>-0.0057</td>
<td>0.0057</td>
</tr>
<tr>
<td>1600</td>
<td>0.3303</td>
<td>0.1360</td>
<td>0.0062</td>
<td>-0.0063</td>
<td>0.0063</td>
</tr>
<tr>
<td>3200</td>
<td>0.3331</td>
<td>0.1022</td>
<td>0.0087</td>
<td>-0.0088</td>
<td>0.0088</td>
</tr>
</tbody>
</table>

The $\phi$ estimator performance, as opposed to the rare disasters model. On the other hand, the RMSE for the long-term interest rate $y$ is roughly in the same level as in the rare disasters model, decreasing as the sample size increases from 400 to 800 and then the RMSE increases as the sample size doubles from 800 and 1,600. Such behavior illustrates the fact that the sieve dimension should slowly increase with the sample size to attain better results. Figures 3(a) and 3(b) show that, as is expected from the $L^2$ approximation theory, the eigenfunctions estimators perform well, on average, on regions near the stationary mean of the process and the approximation is worse for extreme points.

For this model, we use the normalization $||\phi||_2 = 1 = ||\phi^*||_2$. In the Appendix A, we show that there are three potential normalizations but only two scalars to be chosen. This leaves us the possibility of choosing $||\phi^*||_2 = 1 = E[\phi\phi^*]$

Figure 3. Contour plots of pointwise difference between estimated and true eigenfunction.
or \( \| \phi \|_2 = 1 = \mathbb{E}[\phi \phi^*] \) instead. In the Appendix B, we show that the choice of normalization substantially impacts the approximation quality for the eigenfunctions in this case.

5 Implications to Long-term interest rates

In this section, we explore the implications of HS and Christensen (2017) framework to long-run interest rates from two different perspectives. First, we examine the long-run interest rate \( y = -\log \rho \) obtained from data simulated from the disaster and habit model with calibrations designed to match specific US data characteristics, as well as structural estimations using US and Brazilian data. Second, we use the long-run approximation (6) to estimate the yield of a long-term zero-coupon bond as a function of the initial state of the economy. Such approximation is computed in the following way. Let \( P_{0,\tau}(x) \) be the price of a zero-coupon bond at time \( t = 0 \) that pays one unit at time \( t = \tau \) when the current state variable is \( X_0 = x \). If \( \tau \) is very large, then (6) implies that

\[
\rho^{-\tau} P_{0,\tau}(x) \approx \hat{\mathbb{E}}\left[ \frac{1}{\phi(X_t)} \right] \phi(x).
\]

Assuming the identification conditions for the change of measure \( d\tilde{Q}/dQ = \phi \phi^* \) are valid, we get

\[
P_{0,\tau}(x) \approx \rho^{-\tau} \mathbb{E}[\phi^*(X_t)]\phi(x).
\]  

(25)

Because the bond’s yield is given by \( (1/P_{0,\tau})^{1/\tau} - 1 \), it is possible to estimate it as

\[
\left[ \rho^{-\tau} \left( \frac{1}{n} \sum_{t=0}^{\tau} \hat{\phi}^*(X_t) \right) \right]^{-1/\tau} - 1.
\]  

(26)

In this way, we estimate how the yield from a long-run zero-coupon bond varies with the state process \( X_0 = x \). Observing figures 5(a) and 5(b), one can notice that by allowing for a flexible approach to the state process we find that the yield varies non-linearly with \( X_0 = x \).

5.1 U.S. Data

In the models used in the last section, we derived analytical solutions for \( \rho \) and \( y \). We report these values for a disaster model following the calibration proposed by Backus et al. (2011) and a habit model calibrated by Campbell and Cochrane (1999). In addition, using simulated data from these models, we report in Table 3 intervals containing the 2.5% and 97.5% quantiles of the estimators.
Table 3. Values for calibrated habit and disaster models. Intervals display 2.5% and 97.5% percentiles of the estimators values across simulations.

<table>
<thead>
<tr>
<th></th>
<th>( n )</th>
<th>( \hat{\rho} )</th>
<th>( \hat{\gamma} )</th>
<th>( \hat{L} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disaster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True value</td>
<td>0.9326</td>
<td>0.0698</td>
<td>0.0400</td>
<td></td>
</tr>
<tr>
<td>95% interval</td>
<td></td>
<td>(0.8849,0.1.0267)</td>
<td>(−0.0263,0.1.1223)</td>
<td>(0.0017,0.1.1219)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>(0.8935,1.0027)</td>
<td>(−0.0027,0.1126)</td>
<td>(0.0076,0.1059)</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>(0.9021,0.9982)</td>
<td>(0.0018,0.1030)</td>
<td>(0.0142,0.1018)</td>
</tr>
<tr>
<td></td>
<td>3200</td>
<td>(0.9092,0.9781)</td>
<td>(0.0221,0.0952)</td>
<td>(0.0199,0.0840)</td>
</tr>
<tr>
<td><strong>Habit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True value</td>
<td>0.8574</td>
<td>0.1539</td>
<td>0.0004</td>
<td></td>
</tr>
<tr>
<td>95% interval</td>
<td></td>
<td>(0.8548,0.9715)</td>
<td>(0.0289,0.1569)</td>
<td>(−0.0060,0.1211)</td>
</tr>
<tr>
<td></td>
<td>800</td>
<td>(0.8555,0.9452)</td>
<td>(0.0564,0.1561)</td>
<td>(−0.0030,0.0973)</td>
</tr>
<tr>
<td></td>
<td>1600</td>
<td>(0.8561,0.9630)</td>
<td>(0.0377,0.1554)</td>
<td>(−0.0013,0.1168)</td>
</tr>
<tr>
<td></td>
<td>3200</td>
<td>(0.8565,1.0000)</td>
<td>(0.0000,0.1549)</td>
<td>(−0.0003,0.1534)</td>
</tr>
</tbody>
</table>

across simulations. Finally, we estimate the yield of a zero-coupon bond as a function of the initial state using the procedure described above and present it in figures 4(a) and 4(b). Inspecting Table 3, one can find that the calibrated models imply sizeable long-run interest rates—approximately 7% and 15.4% for the disaster and habit models, respectively. We consider these estimates of little value since such models were calibrated to match short-run features of data. However, it is interesting to notice that the approximation of the yield of a long-run zero-coupon bond is close to the value of \( y \), as can be seen in figures 4(a) and 4(b).

Moving to structural estimation, we report the results from Christensen (2017) for the sake of comparison. The model and estimation procedure can be briefly summarized in the following way. The economy is modeled with a representative agent with Epstein and Zin (1989) preferences and unit elasticity of intertemporal substitution (EIS). In the first case, the state process is given by consumption growth alone, and in the second case, the state process is two-dimensional with consumption growth and earnings growth as variables. The data are quarterly and composed of 277 observations from 1947:Q1 to 2016:Q1. First, the preference parameters \( \beta \), \( \gamma \) and a transformation of the continuation value function (\( \chi \) in Christensen’s notation) are estimated from the state process data and the time series of seven asset returns. The estimation employs a procedure proposed by Ai and Chen (2003) for models with conditional moment restrictions and unknown functions. Then, using \( \hat{\beta} \), \( \hat{\gamma} \) and \( \hat{\chi} \), the eigenfunctions
and eigenvalues are estimated as described in subsection 3.4. The results are reported in Table 4, where we can see two types of estimation. In the first type, on the left panel, $\beta$ and $\gamma$ are estimated from the data and used to estimate $\rho$ and $y$. However, the author notes that the estimate for the long-run yield of 1.9% per quarter (around 7.8% per year) is too large and can be explained by the low value of $\hat{\beta}$. To understand how the estimates of $\rho$ and $y$ respond to other reasonable values of $\beta$ and $\gamma$, the author estimates the model fixing $\beta = 0.99$ and $\gamma = 20, 25,$ and $30$. In this way, one gets values of $\hat{y}$ around 1.4% per quarter (about 5.7% per year), which is a more plausible estimate. Finally, in figures 5(a) and 5(b) we show the approximation of the annual yield of a long-term zero-coupon bond using the full structural estimation for specifications with a unidimensional and bidimensional state process. Again, the values are close to $\hat{y}$ and show an interesting non-linear response to the initial state of the economy.

5.2 Brazilian Data

Now we employ the same methodology used by Christensen (2017) for Brazilian data. We use a dataset assembled by Brandão and Almeida (2018) with a sample from 1996:Q2 to 2015:Q4. Consumption is sourced from IBGE’s data on family consumption from the national accounts. Financial data was gathered through Economatica software. Aggregate dividends were computed by taking real dividends from all firms listed in Ibovespa at least once during the sample period. In each year, firms without trade were dropped from the sample. Then,
Table 4. Results reported by Christensen (2017) for US data. Left panel: Estimates of $\rho$, $y$ and $L$ corresponding to $(\hat{\beta}, \hat{\gamma}, \hat{\lambda})$. Right panel: estimates of $\rho$, $y$ and $L$ corresponding to pre-specified $(\beta, \gamma)$ and estimated $(\hat{\lambda}, \hat{\chi})$. 90% bootstrap confidence intervals are in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>$X_t = (g_t, d_t)$</th>
<th>$X_t = g_t$</th>
<th>$X_t = (g_t, d_t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}$</td>
<td>0.9812 (0.9733, 0.9902)</td>
<td>0.9817 (0.9726, 0.9893)</td>
<td>0.9859 (0.9851, 0.9872)</td>
</tr>
<tr>
<td>$\hat{\gamma}$</td>
<td>0.0190 (0.0098, 0.0270)</td>
<td>0.0184 (0.0107, 0.0277)</td>
<td>0.0142 (0.0129, 0.0150)</td>
</tr>
<tr>
<td>$\hat{L}$</td>
<td>0.0193 (0.0000, 0.0381)</td>
<td>0.0215 (0.0000, 0.0426)</td>
<td>0.0128 (0.0090, 0.0185)</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>0.9851 (0.9784, 0.9926)</td>
<td>0.9853 (0.9771, 0.9921)</td>
<td>0.99 (0.9908, 0.9924)</td>
</tr>
<tr>
<td>$\hat{\gamma}$</td>
<td>24.4712 (19.9850, 29.9700)</td>
<td>27.4838 (20.9850, 33.9700)</td>
<td>20</td>
</tr>
<tr>
<td>$\hat{\lambda}$</td>
<td>0.8999 (0.8146, 0.9922)</td>
<td>0.8872 (0.7927, 0.9888)</td>
<td>0.9154 (0.9008, 0.9324)</td>
</tr>
</tbody>
</table>

(a) Yield for zero-coupon bond for different long-term maturities $\tau$ for $X_t = g_t$. Estimation using US data. (b) Yield for zero-coupon bond for long-term maturity $\tau$ for $X_t = (g_t, d_t)$. Estimation using US data.

Figure 5. Plots of long-term zero-coupon bond yield. Pointwise average taken across simulations.

we sum the remaining dividends to generate the aggregate series. Annual population data was collected from IBGE as well, and linear methods were used to extrapolate the series to 2014 and 2015. To generate quarterly population data, we interpolate the annual series with cubic splines. The state variables used were real per capita consumption and dividend growth, which were deflated using the IGP-DI. In order to estimate the structural parameters from the conditional moment restrictions, we use returns on the Brazilian stock market Ibovespa and the short-term interest rate SELIC (settled by the central bank). Again, all series are deflated by the IGP-DI.

The estimation results are reported in Table 5. We can see that for the
one-dimensional specification, the long-term interest rate is 1.05% per quarter (about 4.3% per year), while in the two-dimensional specification the long-term interest rate is approximately 1.32% (around 5.4% per year). Fixing the value of beta and gamma, we found the long-term interest rate decreasing from 1.33% to 1.25% as gamma increased from 16.3 to 30.

Figure 6 show the estimated eigenfunctions and change of measure ($\phi, \phi^*$) for the univariate and bivariate specifications using Brazilian data. For the univariate specification, one can see that both functions have a similar shape, i.e., a concave decreasing function for low and mid values of $g$, and convex parabola for higher values of $g$. However, as is pointed by Christensen (2017) for the US data estimations, the scale of the plots indicate that $\phi$ is relatively flat and $\phi^*$ is significantly steeper. Therefore, the change of measure is dominated by $\phi^*$. This phenomenon is also present in the bivariate specification. On the other hand, the interpretation of the bivariate estimates is challenging as the contour plots show that the eigenfunction’s shape is not similar to the usual parametric specifications, and it displays an intricate nonlinear pattern. Roughly, $\phi^*$ highest values occur in the region of low $g$ and low $d$, while the lowest values occur in the region of mean values of $g$ and $d$.

In Figure 7, we can see that using the nonparametric approach, the long-run interest rate decreases nonlinearly with gamma, as opposed to the VAR specification. Figures 8(a) and 8(b) show how the approximated yield of a long-term zero-coupon bond responds to the initial state of the economy. In the one-dimensional specification, we find an interesting nonlinear and nonmonotonic curve. On the other hand, the two-dimensional specification displays a nonintuitive pattern, where areas of large consumption growth and
Figure 6. Brazilian data: plots of $\hat{\phi}(x)$ (upper panels), $\hat{\phi}^*(x)$ (middle panels) and the estimated change of measure $\hat{\phi}(x)\hat{\phi}^*(x)$ between the stationary distribution $Q$ and the distribution corresponding to $\tilde{E}$ under recursive preferences using the estimated preference parameters in the left panel of Table 5.
Long-term Yields Implied by Stochastic Discount Factor Decompositions

Figure 7. Brazilian data. Solid lines indicate nonparametric estimates of the quarterly long-run yield and correlation between $\hat{m}_{t+1}^P$ and $\hat{m}_{t+1}^T$ under recursive preferences with $X_t = (g_t, d_t)$, $\beta = \hat{\beta} = 0.9859$ for different values of $\gamma$. Dashed lines indicate parametric estimates obtained from fitting a Gaussian VAR(1) to $X_t = (g_t, d_t)$.

Figure 8. Brazilian data: Plots of long-term zero-coupon bond yield. Pointwise average taken across simulations.
average dividend growth display the largest long-term yield, while areas with extreme values of consumption and dividend growth represent lower yield.

### 6 Conclusion

In this article, we are interested in assessing long-term yields based on stochastic discount factor decompositions. To that end, we employ an empirical procedure proposed by Christensen (2017) to implement the HS methodology. We assess the performance of Christensen’s estimator through Monte Carlo simulations of the Disaster and Habit formation asset pricing models. In the Disaster model, the approximation is not symmetric: the precision is considerably worse for points above the mean than for points below the mean. In addition, the RMSE for $\phi^*$ is significantly larger than for $\phi$, especially for smaller samples. In the Habit model, estimation proved to be more challenging. Even when increasing the sample size five-fold does not lead to substantial improvements in terms of RMSE, as opposed to the Disaster model and to the Gaussian AR(1) example provided by Christensen. The performance of the estimator $\hat{\phi}^*$ is substantially worse than the performance of $\hat{\phi}$. Next steps in this research should address this anomalous behavior.

In a second set of results, given that the $L^2$ theory under which the estimator is derived naturally makes it sensible to disaster and habit formation models, we go back to the recursive utility framework with $EIS = 1$ to approximate the yield of a long-term zero-coupon bound using the SDF decomposition. The goal here is to analyze how the approximation of long-term yields varies according to the initial state of the economy $X_0 = x$. Since the estimation does not impose rigid restrictions on the law of motion of the state variable, we find that the yield varies non-linearly with $X_0 = x$. Using U.S. and Brazilian data, we follow Christensen’s estimation methodology and find that such yield approximation is very close to the long-run interest obtained from the eigenvalue of the Perron–Frobenius problem (the log of the eigenvalue solution). We estimate the Brazilian long-term real interest rate around 5.4% per year, with the corresponding long-term US real interest rate estimated by Christensen (2017) being around 5.7%. We conjecture that the slightly higher value of the US rate is due to a combination of two factors: A noisier dividend growth process available for the Brazilian economy and the potentially different assessment of stock returns in the two markets. While for the US Christensen (2017) adopts returns data on six value-weighted portfolios sorted on size and book-to-market values provided by Kenneth French, we adopt data only on the Ibovespa index returns. All in all, both numbers
appear to be plausible from an structural point of view. Further robustness tests adopting a broader dataset to represent the Brazilian stock market may be performed in future work.

This work can be extended in several ways. For example, one could examine in more details how the estimator performance varies with the sieve dimension $k$, optimally choosing this dimension based on a data-driven procedure. Also, one could construct alternative weighting procedures that assigns more mass to the tails and modify the estimator to work better in regions with smaller probability mass. Finally, as noted by Christensen, a natural extension is to first extract the SDF flexibly from panels of asset returns data and then apply the methodology to estimate the components as opposed to using a parametric model.

References
Appendix A  Habit model analytic solution

Let the state process be $X_t = (g_t, s_t)$. The Perron–Frobenius problem in this setting is given by

$$
\mathbb{E}_t \left[ \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{S_{t+1}}{S_t} \right)^{-\gamma} \phi(g_{t+1}, s_{t+1}) \right] = \rho \phi(g_t, s_t). \quad (27)
$$
Therefore, a solution should satisfy equation (27) becomes

\[
E_t[\beta \exp(-\gamma g_{t+1}) \exp((b - \gamma)s_{t+1} + \gamma s_t)]
\]

\[
= E_t[\beta \exp(-\gamma(\mu + v_{t+1})) \exp((b - \gamma)((1 - \kappa)\bar{s} + \kappa s_t + \lambda(\bar{s})v_{t+1}) + \gamma s_t)]
\]

\[
= E_t[\beta \exp\{((b - \gamma)\lambda(s_t) - \gamma)v_{t+1} + ((b - \gamma)\kappa + \gamma)s_t - \gamma \mu + (b - \gamma)(1 - \kappa)\bar{s})\}
\]

\[
= \beta \exp\left\{\frac{\sigma^2}{2}((b - \gamma)\lambda(s_t) - \gamma)^2\right\}\exp\{((b - \gamma)\kappa + \gamma)s_t - \gamma \mu + (b - \gamma)(1 - \kappa)\bar{s})\}.
\]

Therefore, a solution should satisfy \((b - \gamma)\kappa + \gamma = b\), which implies \(b = \gamma\). Then, equation (21) becomes

\[
M\phi(g_t, s_t) = \beta \exp\left\{-\gamma \mu + \frac{\sigma^2 \gamma^2}{2}\right\}\exp\{\gamma s_t\}.
\]

Now, for the time-reversed problem we conjecture \(\phi^*(x, y) = \exp(a^* x + b^* y)\) so

\[
E_{t+1}[\beta \exp(-\gamma g_{t+1}^*) \exp(-\gamma(s_{t+1}^* - s_t^*)) \phi^*(g_t^*, s_t^*)] = \rho \phi^*(g_{t+1}^*, s_{t+1}^*)
\]

becomes

\[
E_{t+1}[\exp(\gamma s_t^*) \exp(a^* g_t^* + b^* s_t^*)] = \frac{\rho}{\beta} \exp(\gamma(g_{t+1}^* + s_{t+1}^*)) \exp(a^* g_{t+1}^* + b^* s_{t+1}^*)
\]

\[
E_{t+1}[\exp(a^* g_t^* + (b^* + \gamma) s_t^*)] = \frac{\rho}{\beta} \exp((a^* + \gamma)g_{t+1}^* + (b^* + \gamma) s_{t+1}^*).
\]

The left-hand side of equation (29) can be written as

\[
E_{t+1}\left[\exp\left\{a^*(\mu + v_t^*) + (b^* + \gamma)((1 - \kappa)\bar{s} + \kappa s_{t+1}^* + \lambda(s_{t+1}^*)v_t^*)\right\}\right]
\]

\[
= E_{t+1}\left[\exp\left\{((\gamma + b^*)\lambda(s_{t+1}^* + a^*)v_t^*)\right\}\exp\left\{((\gamma + b^*)((1 - \kappa)\bar{s} + \kappa s_{t+1}^*) + a^* \mu\right\}\right]
\]

\[
= \exp\left\{((\gamma + b^*)\lambda(s_{t+1}^*) + a^*)^2 \frac{\sigma^2}{2}\right\}\exp\left\{((\gamma + b^*)((1 - \kappa)\bar{s} + \kappa s_{t+1}^*) + a^* \mu\right\}\exp\left\{((\gamma + a^*)g_{t+1}^* + (\gamma + b^*) s_{t+1}^*)\right\}
\]

On the other hand, the right-hand side of equation (29) can be written as

\[
\exp\left(-\gamma \mu + \frac{\sigma^2 \gamma^2}{2}\right) \exp((\gamma + a^*)g_{t+1}^* + (\gamma + b^*) s_{t+1}^*).
\]

Therefore, a solution should satisfy

\[
a = -\gamma,
\]

\[
(\gamma + b^*)\kappa = (\gamma + b^*),
\]

\[
(\gamma + b^*)\lambda(s_{t+1}^*) = 0.
\]
A simple solution is given by $b = -\gamma$.

So the unscaled solutions we have so far are:

\[
\phi(x, y) = \exp(\gamma y), \quad (30)
\]
\[
\phi^*(x, y) = \exp(-\gamma(x + y)), \quad (31)
\]
\[
\rho = \beta \exp\left(-\gamma \mu + \frac{\sigma^2 \gamma^2}{2}\right), \quad (32)
\]

To guarantee $\|\phi\|_2 = 1$, we seek for $\eta$ such that

\[
E\left[\exp(\eta + \gamma s_t)^2\right] = 1 = \exp(2\eta) E[\exp(2\gamma s_t)].
\]

Therefore, $\eta = -\log(E[\exp(2\gamma s_t)])/2$ is the solution we are looking for. The exact stationary (marginal) distribution of $s_t$ is not easily inferred because of the complexity added by the sensitivity function $\lambda$. For this reason, the numerical value of $\eta$ is approximated by replacing the population mean by the sample average over 200,000 replications. Finally, we need to rescale $\phi^*$ so $E[\phi^*\phi^*] = 1$ holds, i.e.,

\[
E[\exp(\eta + \gamma s_t) \exp(\theta - \gamma (g_t + s_t))] = 1
\]
\[
= \exp(\eta + \theta) E[\exp(-\gamma g_t)]
\]
\[
= \exp(\eta + \theta) \exp\left(-\gamma \mu + \frac{\gamma^2 \sigma^2}{2}\right)
\]
\[
= \exp(\eta + \theta) \rho \beta^{-1}.
\]

Thus, we set $\phi(x, y) = \beta \rho^{-1} \exp(-\eta - \gamma(x + y))$. Because there are three possible normalizations and only two scalars to be chosen, we also consider the normalization $\|\phi^*\|_2 = 1$ and $E[\phi^*\phi^*] = 1$, which yields $\theta = -\log(E[\exp(-2 \gamma (g_t + s_t))])$ and $\eta = -\theta + \gamma \mu - \gamma^2 \sigma^2/2$; and the normalization $\|\phi\|_2 = 1 = \|\phi^*\|_2$, which implies $\theta = -\log(E[\exp(-2 \gamma (g_t + s_t))])$ and $\eta = -\log(E[\exp(2\gamma s_t)])/2$. 


Appendix B  Habit model supplementary results

Table 6. Bias and RMSE found in simulations for the Habit model. For the normalization $\|\phi\|_2 = 1 = E[\phi\phi^*]$.  

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\hat{\phi}$</th>
<th>$\hat{\phi}^*$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{y}$</th>
<th>$\hat{L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0.3468</td>
<td>1.4133</td>
<td>0.0057</td>
<td>-0.0057</td>
<td>0.0057</td>
</tr>
<tr>
<td>800</td>
<td>0.3353</td>
<td>1.6917</td>
<td>0.0056</td>
<td>-0.0057</td>
<td>0.0057</td>
</tr>
<tr>
<td>1000</td>
<td>0.3340</td>
<td>1.8332</td>
<td>0.0059</td>
<td>-0.0060</td>
<td>0.0060</td>
</tr>
<tr>
<td>2000</td>
<td>0.3320</td>
<td>2.4565</td>
<td>0.0072</td>
<td>-0.0073</td>
<td>0.0073</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\hat{\phi}$</th>
<th>$\hat{\phi}^*$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{y}$</th>
<th>$\hat{L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0.3893</td>
<td>7.4863</td>
<td>0.0435</td>
<td>0.0397</td>
<td>0.0393</td>
</tr>
<tr>
<td>800</td>
<td>0.3651</td>
<td>6.5185</td>
<td>0.0400</td>
<td>0.0377</td>
<td>0.0375</td>
</tr>
<tr>
<td>1000</td>
<td>0.3623</td>
<td>7.0969</td>
<td>0.0407</td>
<td>0.0383</td>
<td>0.0382</td>
</tr>
<tr>
<td>2000</td>
<td>0.3563</td>
<td>6.4229</td>
<td>0.0485</td>
<td>0.0441</td>
<td>0.0440</td>
</tr>
</tbody>
</table>

(a) Contour plot of mean $\hat{\phi}(x)$ for Habit model  
(b) Contour plot of mean $\hat{\phi}^*(x)$ for Habit model

Figure 9. Simulation results for the normalization $\|\phi\|_2 = 1 = E[\phi\phi^*]$, using a Hermite polynomial basis with $k = 7$ for each variable. Graphs display the pointwise mean $\hat{\phi}$ and $\hat{\phi}^*$ estimated in a simulation with sample length equal to 2,000.
Figure 10. Contour plots of true solution for $\phi$ and $\phi^*$, for the normalization $\|\phi\|_2 = 1 = E[\phi\phi^*]$.

Table 7. Bias and RMSE found in simulations for the Habit model. For normalization $\|\phi^*\|_2 = 1 = E[\phi\phi^*]$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\hat{\phi}$</th>
<th>$\hat{\phi}^*$</th>
<th>$\hat{\rho}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>404.1470</td>
<td>0.2724</td>
<td>0.0057</td>
<td>-0.0057</td>
<td>0.0057</td>
</tr>
<tr>
<td>800</td>
<td>393.3920</td>
<td>0.1905</td>
<td>0.0056</td>
<td>-0.0057</td>
<td>0.0057</td>
</tr>
<tr>
<td>1600</td>
<td>381.7500</td>
<td>0.1360</td>
<td>0.0062</td>
<td>-0.0063</td>
<td>0.0063</td>
</tr>
<tr>
<td>3200</td>
<td>369.3257</td>
<td>0.1022</td>
<td>0.0087</td>
<td>-0.0088</td>
<td>0.0088</td>
</tr>
<tr>
<td>RMSE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>434.6703</td>
<td>0.3740</td>
<td>0.0435</td>
<td>0.0397</td>
<td>0.0393</td>
</tr>
<tr>
<td>800</td>
<td>437.7962</td>
<td>0.2727</td>
<td>0.0400</td>
<td>0.0377</td>
<td>0.0375</td>
</tr>
<tr>
<td>1600</td>
<td>405.8449</td>
<td>0.2166</td>
<td>0.0424</td>
<td>0.0393</td>
<td>0.0393</td>
</tr>
<tr>
<td>3200</td>
<td>393.3863</td>
<td>0.1862</td>
<td>0.0535</td>
<td>0.0482</td>
<td>0.0482</td>
</tr>
</tbody>
</table>