SDF Bounds*

Piotr Orłowski[†] Alireza Tahbaz-Salehi[‡]
Fabio Trojani[§] Andrea Vedolin[¶]

November 2024

Abstract

This paper develops a framework for testing asset-pricing models by deriving restrictions on the (marginal and joint) distributions of stochastic discount factors (SDFs). The framework takes an arbitrary set of empirical or theoretical restrictions—such as observable returns and Euler equations—as primitives and yields necessary conditions on the joint distribution of SDFs that must hold for the model to align with the specified restrictions. Applying our results to international asset-pricing models, we show that observed asset prices impose non-trivial constraints on the comovement of SDFs across countries.

*We are grateful to Jingye Wang, Gurdip Bakshi, Fousseni Chabi-Yo, Mike Chernov, Jens Jackwerth, Oleg Itskhoki, and Stig Lundeby for helpful comments.

[†]HEC Montréal, piotr.orlowski@hec.ca.
[‡]Northwestern University and CEPR, alirezat@kellogg.northwestern.edu.
[§]University of Geneva and Swiss Finance Institute, fabio.trojani@unige.ch.
[¶]Boston University, NBER, and CEPR, avedolin@bu.edu.

1 Introduction

According to what is sometimes referred to as "the fundamental theorem of asset pricing," absence of arbitrage opportunities implies the existence of a stochastic discount factor (SDF) that prices all assets by simultaneously discounting the future and adjusting for risk. This result, coupled with the fact that an asset-pricing model can be represented by an SDF, has turned understanding the properties of SDFs into one of the cornerstones of modern asset-pricing literature.

A key pillar of this approach has been to characterize the set of admissible SDFs that are consistent with observed asset return data. A classical example is the seminal work of Hansen and Jagannathan (1991), who provide a bound on the volatility of an SDF in terms of the Sharpe ratios of observed returns. This result not only can be used to evaluate the performance of asset-pricing models, but also provides information on how to modify the SDF to be consistent with data. Subsequent work has refined and extended the volatility bound of Hansen and Jagannathan (1991) to higher-order moments (Snow, 1991) and the dynamic properties of SDFs (Alvarez and Jermann, 2005). Not surprisingly then, these bounds have become part of the standard toolbox for assessing various asset-pricing models.

While extremely powerful, the HJ bound and its various descendants only impose restrictions on the marginal distribution of a single SDF. This is despite the fact that many applications in macroeconomics and finance are fundamentally about the joint distribution of multiple SDFs. For example, in international finance, movements in exchange rates are tightly linked to the comovement of SDFs of different countries (Brandt, Cochrane, and Santa-Clara, 2006). Similarly, the term structure of risk is determined by the relationship between SDFs that price assets over different horizons (Hansen and Scheinkman, 2009).

In this paper, we fill this gap by developing a systematic approach to testing asset-pricing models, with a particular focus on obtaining testable restrictions on the joint distribution of multiple SDFs. Our framework takes an arbitrary set of empirical and/or theoretical restrictions—such as observable returns and Euler equations—as primitives and returns a family of necessary conditions on the joint distribution of SDFs that have to be satisfied if the model is to be consistent with the a priori specified restrictions.

We develop our results in the context of a general asset-pricing framework in which a vector of returns, $\mathbf{R} = (R_1, \ldots, R_n)$, is priced by vector of pricing kernels, $\mathbf{M} = (M_1, \ldots, M_k)$. Depending on the application, $(M_1, \ldots, M_k)$ may represent the stochastic discount factors of different countries, single-period pricing kernels that price returns over different horizons, or the distinct frequency components of a pricing kernel as in, e.g., Alvarez and Jermann (2005).

Our main objective is to obtain testable restrictions on the joint distribution of $(M_1, \ldots, M_k)$ in terms of (i) observed statistical properties of asset returns and (ii) the necessary theoretical conditions that the joint distribution of SDFs and returns must satisfy (such as no-arbitrage conditions). We therefore start our analysis by explicitly specifying all empirical and theoretical restrictions that the underlying asset-pricing model needs to satisfy. We encode such restrictions by the means of what we refer to as the "observable" set of moments: this set contains all joint moments of $\mathbf{M}$ and $\mathbf{R}$ whose values are pinned down either because they can be directly measured using return data or because of an a priori specified theoretical restriction on the model. An example of the latter type of restriction would be $\mathbb{E}[M_i R_j] = 1$ if the model requires the i-th SDF to price the j-th return. Not surprisingly then, the exact specification of the observable set depends on the particular application.

Having specified the observable set, we then derive testable restrictions on the unobservable moments of $\mathbf{M}$ and $\mathbf{R}$ in terms of their observable moments (whose values are known by definition). Specifically, we use the convexity of the joint cumulant-generating function of log SDFs and log returns and Jensen's inequality to obtain upper and lower bounds on the unobservable moments of $\mathbf{M}$ and $\mathbf{R}$. These bounds, which we refer to as *Jensen bounds*, have two key properties. First, the bounds are, by construction, only in terms of observable moments of SDFs and returns. As such, they can be constructed using moments that are either directly measurable from empirically observable data or whose values are pinned down by theoretical restrictions, such as an Euler equation. Second, the bounds are necessary conditions: they have to be satisfied if an asset-pricing model is to be consistent with the restrictions encoded in the observable set. Together, these two properties imply that, just like the HJ bound, the family of Jensen bounds can be used to test whether an asset-pricing model is consistent with the data.

A key characteristic of our Jensen bound-based approach to testing asset-pricing models is its generality: it provides a systematic method for testing an asset-pricing model given an arbitrary set of empirical and/or theoretical restrictions on the model. Furthermore, our methodology can be used to obtain testable restrictions on the joint distribution of multiple SDFs—such as tests for comovement of SDFs of different countries—a result with no counterpart in the prior literature.[1]

After establishing a few basic properties of Jensen bounds, we turn to applying our results to three canonical asset-pricing environments, obtaining closed-form expressions for our bounds in each case. As a first exercise (and to show how to operationalize our more general results), we consider a simple textbook environment, where a single SDF prices a set of assets. We show that, in such a simple setting, the family of Jensen bounds provides upper and lower bounds on various moments of the SDF in terms of the moments of the return distribution, thus nesting the well-known variance bound of Hansen and Jagannathan (1991) and entropy bounds of Alvarez and Jermann (2005) as special cases. As a second characterization result, we apply our results to an international asset-pricing environment and obtain upper and lower bounds on the joint moments of stochastic discount factors of two countries. We show how these results provide non-trivial restrictions on the comovement of the two SDFs (as measured by, say, their covariance and coskewness). Finally, we apply our general results to a dynamic asset-pricing environment and obtain explicit bounds on the joint moments of the permanent and transitory components of the SDF process.

We conclude the paper by showcasing our theoretical results in the context two empirical applications. In our first application, we revisit the exchange rate volatility puzzle. When financial markets are complete, changes in the depreciation rate of a currency are equal to the difference in SDFs. This result immediately gives rise to the exchange rate volatility puzzle, which posits that in order to match the relatively low volatility of exchange rates and high Sharpe ratios in asset markets, the correlation between international SDFs must be almost perfect, see, e.g., Brandt, Cochrane, and Santa-Clara (2006). However, the low correlation observed among macroeconomic quantities internationally implies that the correlation of international SDFs is too low. Using our theoretical framework, we can

[1]One notable exception is Bakshi and Chabi-Yo (2012), who obtain a lower bound on the variance of the ratio of permanent and transitory components of SDFs.

calculate bounds on the correlation of SDFs for different market structures. More specifically, since our framework only relies on no arbitrage in financial pricing but makes no assumptions about the market structure such as the degree of completeness, our bounds are informative beyond the standard case of complete markets. Moreover, our setting also allows for varying degrees of integration, from fully integrated markets where investors can trade all assets internationally, to settings where investors can only trade the short-term bond, i.e., can trade carry, see, e.g., Lustig and Verdelhan (2019). Our findings can be summarized as follows. When agents are allowed to trade various assets internationally, our lower bound implies that correlations are almost of the order of 90%. Our exercise shows that the counterfactual properties of exchanges rates implied by complete market models extends to theoretical models with incomplete financial markets or varying degrees of financial market integration.

In our second application, we use our bounds on the joint moments of SDFs to test the seminal international asset-pricing model of Colacito and Croce (2013). These authors show that a frictionless international economy with correlated long-run growth prospects can account for the forward premium puzzle as well as other asset-pricing anomalies. To test this model, we simulate the SDFs for two countries (the U.S. and the U.K.) using the calibration of Colacito and Croce (2013) and investigate whether the resulting joint distribution of SDFs violates our upper and lower Jensen bounds, both of which we calculate from return data directly. Our results indicate that while the model satisfies the univariate HJ bounds, the Jensen bounds are violated over a large region. Additionally, our family of tests indicates that the model is rejected because it does not produce enough negative co-skewness between the two SDFs. Our findings highlight a potential trade-off between allowing for high SDF volatilities to match univariate moments and joint moments of SDFs.

Related Literature. Our paper contributes to a large literature which studies properties of stochastic discount factors in arbitrage-free markets. The pioneering work of Hansen and Jagannathan (1991) derives a lower bound on the volatility of SDFs and characterizes the admissible set of SDFs in a model-free way to help diagnose asset pricing models. Subsequent work such as Snow (1991), Stutzer (1995), Bansal and Lehmann (1997), Alvarez and Jermann (2005), Backus et al. (2014), and Almeida and Garcia (2017), among many others, extend the volatility bounds to higher-order moments. Other extensions identify minimum distance modifications of a candidate SDF needed for it to be consistent with the data, see, e.g., Hansen and Jagannathan (1997), Almeida and Garcia (2012), and Gosh et al. (2017), among others.

We contribute to this literature by developing a systematic approach to testing asset-pricing models. Our framework takes an arbitrary set of empirical and/or theoretical restrictions—such as observable returns and Euler equations—as primitives and returns a family of necessary conditions on the joint distribution of SDFs that have to be satisfied if the model is to be consistent with the a priori specified restrictions. As such, our main results nest the bounds in the aforementioned papers as special cases. Furthermore, our methodology can be used to obtain testable restrictions on the joint distribution of multiple SDFs—such as tests for comovement of SDFs of different countries, the term structure of risks, or the joint properties of long- and short-run SDF components—a result with no counterpart in the prior literature.

Our paper also contributes to the literature that studies the effect of different market structures on exchange rates. Lustig and Verdelhan (2019) study exchange rates puzzles in preference-free incomplete market models and conclude that incomplete spanning does not help address key exchange rate puzzles. Beyond incomplete spanning, Sandulescu, Trojani, and Vedolin (2021) explore the degree of market integration for exchange rate puzzles and find that segmented markets populated by financial intermediaries are most appropriate to explain exchange rate behavior. Chernov, Haddad, and Itskhoki (2023) derive a set of restrictions on exchange rates and ask what type of deviations from complete markets are most promising to explain exchange rate behavior. They find that settings that impose little structure on financial markets are most helpful in addressing puzzles in international financial markets.

Our paper is different from these as we propose a simple test on the joint moments of SDFs in settings with minimal structural assumptions about the economy other than no arbitrage pricing. Our test allows us to diagnose various market structures and their implications for asset return distributions beyond the context of international markets but more generally also in settings with multi-horizon pricing or across different frequency domains.

Outline. The rest of the paper is organized as follows. Section 2 presents our framework and formally defines the set of observable joint moments of SDFs and returns. Section 3 contains our main results, where we use the convexity of the cumulant generating function and Jensen's inequality to obtain tight theoretical bounds on the joint distribution of SDFs and returns. In Section 4, we apply our general results to three canonical asset-pricing applications and obtain closed-form expressions for our bounds in each case. Section 5 presents our empirical applications. All proofs and some additional mathematical details are provided in the Appendix.

2 Framework

We consider a general asset-pricing framework in which a vector of returns is priced by another vector of pricing kernels. Specifically, consider an economy consisting of n different assets with gross returns $\mathbf{R} = (R_1, \ldots, R_n)$, and let $\mathbf{M} = (M_1, \ldots, M_k)$ denote a vector of positive random variables, which, with some abuse of terminology, we refer to as stochastic discount factors (SDFs). Depending on the application, $(M_1, \ldots, M_k)$ may represent the stochastic discount factors of different countries, single-period pricing kernels that price returns over different horizons, or the distinct frequency components of a pricing kernel as in, e.g., Alvarez and Jermann (2005).

Our main objective is to obtain testable restrictions on the joint distribution of $(M_1, \ldots, M_k)$ in terms of (i) observed statistical properties of asset returns and (ii) the necessary theoretical conditions that the joint distribution of $\mathbf{M}$ and $\mathbf{R}$ must satisfy (such as no-arbitrage conditions). To state this problem formally, let

$$\mathbb{E}[\mathbf{M}^m \mathbf{R}^r] = \mathbb{E}\left[\prod_{i=1}^{k} M_i^{m_i} \prod_{j=1}^{n} R_j^{r_j}\right] \tag{1}$$

denote the joint moments of SDFs and asset returns, where $m = (m_1, \ldots, m_k)$ and $r = (r_1, \ldots, r_n)$

are collections of real scalars. Note that set $\{\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] : (m,r) \in \mathbb{R}^{k+n}\}$ uniquely identifies the joint distribution of $\mathbf{M}$ and $\mathbf{R}$. As a result, any restriction on the joint distribution of returns and SDFs can be naturally expressed in terms of their joint moments in equation (1). We thus have the following definition:

Definition 1. Given collections of SDFs, $\mathbf{M} = (M_1, \dots, M_k)$, and returns, $\mathbf{R} = (R_1, \dots, R_n)$, the *set of observable (joint) moments of* $\mathbf{M}$ *and* $\mathbf{R}$ is

$$\mathcal{O} = \{(m,r) \in \mathbb{R}^{k+n} : \mathbb{E}[\mathbf{M}^m\mathbf{R}^r] \text{ is observable}\}.$$

As already mentioned, a given moment is "observable" if either (i) it can be directly measured using observable statistical properties of asset returns or (ii) its value is pinned down by theoretical conditions that the asset-pricing model needs to satisfy (such as no-arbitrage conditions). As such, set $\{\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] : (m,r) \in \mathcal{O}\}$ encodes all empirical and theoretical restrictions on the joint moments of $\mathbf{M}$ and $\mathbf{R}$.

With Definition 1 in hand, we can now formally express the main objective of this paper: our goal is to obtain tight and testable theoretical bounds on unobservable moments $\{\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] : (m,r) \notin \mathcal{O}\}$ in terms of the set of observable moments, $\{\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] : (m,r) \in \mathcal{O}\}$.[2]

Before proceeding any further, however, it is useful to illustrate how set $\mathcal{O}$ encodes the empirical and theoretical restrictions on the joint distribution of $\mathbf{M}$ and $\mathbf{R}$ by the means of a few examples. We will use these examples in subsequent sections to showcase our results.

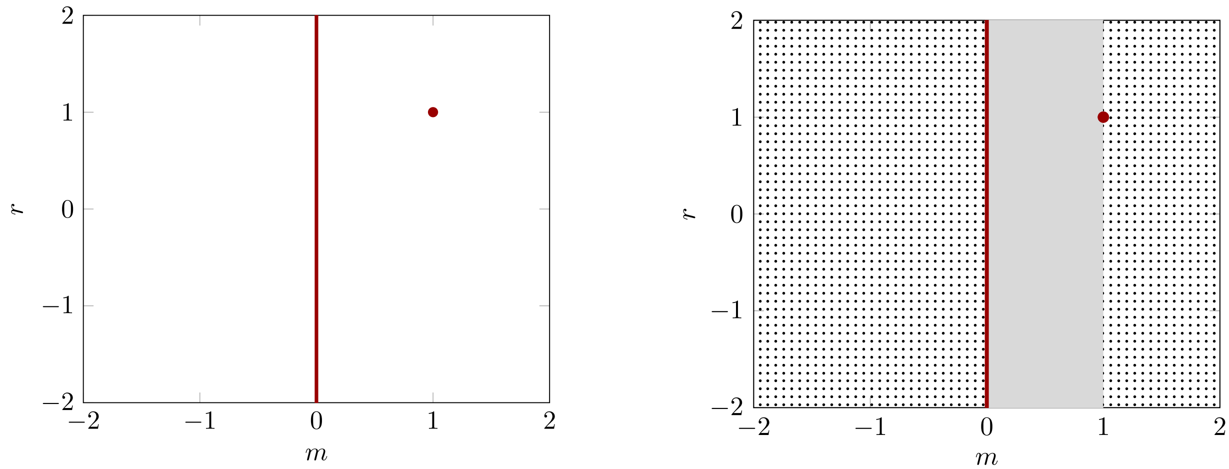
Example 1 (univariate case)**.** As the simplest possible example, consider a univariate special case of our framework, in which an arbitrary return, R, is priced by a single SDF, M. This means that the joint distribution of $\mathbf{M}$ and $\mathbf{R}$ can be summarized by $\{\mathbb{E}[M^mR^r] : (m,r) \in \mathbb{R}^2\}$.

When the return distribution is empirically observable, $\mathbb{E}[R^r]$ can be directly measured using return data. As a result, $(0,r)$ belongs to the observable set for any given $r \in \mathbb{R}$. In addition, the assumption that the SDF prices the return means that $\mathbb{E}[MR] = 1$, which means that $(1,1)$ also belongs to the observable set. Putting the above together implies that the set of observable moments is given by $\mathcal{O} = \{(0,r) : r \in \mathbb{R}\} \cup \{(1,1)\}$. Figure 1(a) depicts this set. The vertical line represents the moments that can be measured empirically, while the single point captures the no-arbitrage condition that M and R need to jointly satisfy.

Example 2 (international financial markets)**.** Next, consider an economy consisting of two countries, labeled 1 and 2, each with its own currency. Investors in each country can trade assets denominated in their respective currencies. Let $\mathbf{R} = (R_1, R_2)$, where R_i denotes an arbitrary return that can be traded by investors in country i. Similarly, let $\mathbf{M} = (M_1, M_2)$, where M_i is the SDF that prices R_i from the perspective of investors in country i (i.e., in country i's currency).

In this economy, the joint distribution of SDFs and returns can be summarized by $\{\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] : (m,r) \in \mathbb{R}^4\}$, where $m = (m_1, m_2)$ and $r = (r_1, r_2)$. When the joint distribution of returns (R_1, R_2) is

[2] With a slight abuse of terminology, we use the term "observable set" to refer to both sets $\mathcal{O}$ and $\{\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] : (m,r) \in \mathcal{O}\}$. We also remark that, depending on the application, the expectation operator in Definition 1 can be a conditional or unconditional expectation operator. However, to avoid notational clutter, we do not explicitly index $\mathcal{O}$ by the information set corresponding to the expectation operator.



(a) observable set, $\mathcal{O}$

(b) sets $\overline{\mathcal{O}}$ (shaded region) and $\underline{\mathcal{O}}$ (dotted region)

Figure 1. Panel (a) depicts the set of observable moments, $\mathcal{O}$, corresponding to Example 1. Panel (b) additionally depicts sets $\overline{\mathcal{O}}$ (shaded region) and $\underline{\mathcal{O}}$ (dotted region) as defined in equations (4) and (9), respectively.

empirically observable, then $\mathbb{E}[R_1^{r_1} R_2^{r_2}]$ can be measured using return data for any $(r_1, r_2) \in \mathbb{R}^2$. As a result, the set of observable moments is given by

$$\mathcal{O} = \{(0, 0, r_1, r_2) : (r_1, r_2) \in \mathbb{R}^2\} \cup \{(1, 0, 1, 0), (0, 1, 0, 1)\}, \tag{2}$$

where the last two elements of the set reflect the pricing restrictions $\mathbb{E}[M_1 R_1] = 1$ and $\mathbb{E}[M_2 R_2] = 1$, respectively.

Example 3 (permanent-transitory decomposition of SDFs)**.** As our final example, consider a discrete-time dynamic economy with stochastic discount factor process $\{M_{t,t+\tau}\}_{t\geq 1}$ that satisfies $\mathbb{E}_t[M_{t,t+\tau} R_{t,t+\tau}] = 1$ for all t and all $\tau \geq 1$, where $R_{t,t+\tau}$ is the gross return on an arbitrary traded asset over the period from t to $t + \tau$ and $\mathbb{E}_t[\cdot]$ denotes the information set at time t. Alvarez and Jermann (2005) show that the stochastic discount factor in such an economy can be decomposed into permanent and transitory components:

$$M_{t,t+\tau} = M^P_{t,t+\tau} M^T_{t,t+\tau}.$$

The permanent component in the above decomposition is a martingale ($\mathbb{E}_t[M^P_{t,t+\tau}] = 1$ almost surely), whereas the transitory component is equal to the reciprocal of the return to holding a discount bond of (asymptotically) long maturity from date t to date $t + \tau$: $M^T_{t,t+\tau} = 1/R^{(\infty)}_{t,t+\tau}$.

Setting $\mathbf{M} = (M^P_{t,t+\tau}, M^T_{t,t+\tau})$ and $\mathbf{R} = R_{t,t+\tau}$ and assuming that the joint distribution of $R_{t,t+\tau}$ and the infinite-maturity bond can be measured using return data, it follows that the observable set in this application is given by

$$\mathcal{O} = \{(0, m_2, r,) : (m_2, r) \in \mathbb{R}^2\} \cup \{(1, 1, 1)\} \cup \{(1, 0, 0)\}. \tag{3}$$

The first subset above captures empirical restrictions on the joint moments of SDFs and returns imposed by the observability of the joint distribution of $(R_{t,t+\tau}, R^{(\infty)}_{t,t+\tau})$, wheres the last two elements reflect theoretical restrictions: $\mathbb{E}_t[M^P_{t,t+\tau} M^T_{t,t+\tau} R_{t,t+\tau}] = 1$ and $\mathbb{E}_t[M^P_{t,t+\tau}] = 1$, respectively.

3 Jensen Bounds

In this section, we provide tight theoretical bounds on the joint distribution of SDFs and returns in the general asset-pricing framework outlined in Section 2. Specifically, given a collection of SDFs $\mathbf{M} = (M_1, \ldots, M_k)$ and returns $\mathbf{R} = (R_1, \ldots, R_n)$ and a set of observables, $\mathcal{O}$, we obtain upper and lower bounds on unobservable moments $\{\mathbb{E}[\mathbf{M}^m \mathbf{R}^r] : (m, r) \notin \mathcal{O}\}$ in terms of observable moments $\{\mathbb{E}[\mathbf{M}^m \mathbf{R}^r] : (m, r) \in \mathcal{O}\}$. Since our bounds are expressed solely in terms of observable moments, they can be used to test whether an asset-pricing model is consistent with (i) empirically measurable quantities, such as asset returns, and (ii) the necessary theoretical conditions that the model has to satisfy.

Our main characterization results rely on the convexity of the joint cumulant-generating function of log SDFs and log returns (defined below) and Jensen's inequality. We therefore refer to our family of bounds as *Jensen bounds*. As we will discuss in detail in Section 4, these bounds extend well-known bounds such as the HJ bound and the entropy bound of Alvarez and Jermann (2005), among others. Importantly, our results go beyond these prior bounds by allowing us to obtain restrictions on the joint distribution of $(M_1, \ldots, M_k)$ in addition to their marginals.

3.1 Jensen Upper Bounds

We start by characterizing our family of upper bounds. Let $\phi : \mathbb{R}^k \times \mathbb{R}^n \to \mathbb{R}$ denote the joint *cumulant-generating function* (or the CGF) of log SDFs and log returns, defined as $\phi(m, r) = \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$.[3] Additionally, let $\overline{\mathcal{O}}$ denote the convex hull of the set of observable moments:

$$\overline{\mathcal{O}} = \left\{ \sum_{i=1}^{n} \lambda_i x_i \,\middle|\, x_1, \ldots, x_n \in \mathcal{O}, \ \sum_{i=1}^{n} \lambda_i = 1, \text{ and } \lambda_1, \ldots, \lambda_n \geq 0 \right\}, \tag{4}$$

where recall that set $\mathcal{O}$ encodes the (empirical and theoretical) restrictions on the joint distribution of $\mathbf{M}$ and $\mathbf{R}$. We have the following result:

Theorem 1. *Let $\mathbf{M}$ and $\mathbf{R}$ denote collections of SDFs and asset returns, respectively, and let $\mathcal{O}$ denote the corresponding set of observable moments. Then,*

$$\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] \leq \overline{\mathrm{J}}_{\mathcal{O}}(m, r) \qquad \textit{for all } (m, r) \in \overline{\mathcal{O}}, \tag{5}$$

[3]More generally, the CGF of a random vector $(X_1, \ldots, X_n)$ is defined as the logarithm of its moment-generating function, whenever it exists. Specifically, $\phi(t_1, \ldots, t_n) = \log \mathbb{E}[e^{t_1 X_1 + \cdots + t_n X_n}]$.

where

$$\overline{\mathrm{J}}_{\mathcal{O}}(m,r) = \inf_{F} \quad \int_{\mathcal{O}} \log \mathbb{E}[\mathbf{M}^{\tilde{m}}\mathbf{R}^{\tilde{r}}]\, \mathrm{d}F(\tilde{m},\tilde{r}) \tag{6}$$

$$\text{s.t.} \quad \int_{\mathcal{O}} \tilde{m}\, \mathrm{d}F(\tilde{m},\tilde{r}) = m \tag{7}$$

$$\int_{\mathcal{O}} \tilde{r}\, \mathrm{d}F(\tilde{m},\tilde{r}) = r, \tag{8}$$

and F is a probability distribution with support over $\mathcal{O}$.

This result provides an upper bound on $\log \mathbb{E}[\mathbf{M}^m\mathbf{R}^r]$ for any (m,r) that is in the convex hull of the set of observable moments, $\mathcal{O}$. Importantly, as is evident from the optimization problem in (6)–(8), this upper bound, $\overline{\mathrm{J}}_{\mathcal{O}}(m,r)$, can be calculated using moments that are either directly measurable from empirically observable data or whose values are pinned down by theoretical restrictions, such as an Euler equation.

Theorem 1 is a consequence of Jensen's inequality and the observation that the CGF, $\phi(m,r) = \log \mathbb{E}[\mathbf{M}^m\mathbf{R}^r]$, is a convex function of (m,r). To see the logic behind Theorem 1, note that when (m,r) is in the convex hull of the observable set, it can be expressed as a convex combination of points $(\tilde{m},\tilde{r}) \in \mathcal{O}$, as captured by the constraints in (7) and (8). Therefore, Jensen's inequality implies that $\log \mathbb{E}[\mathbf{M}^m\mathbf{R}^r]$ is upper bounded by the convex combination of points in $\{\log \mathbb{E}[\mathbf{M}^{\tilde{m}}\mathbf{R}^{\tilde{r}}] : (\tilde{m},\tilde{r}) \in \mathcal{O}\}$ with the same weights. Naturally, taking the infimum over all such convex combinations results in the tightest possible upper bound. This is exactly what the optimization problem in (6) does. Put differently, $\overline{\mathrm{J}}_{\mathcal{O}}(m,r)$ is the largest convex extension of $\log \mathbb{E}[\mathbf{M}^m\mathbf{R}^r]$ to the convex hull of set $\mathcal{O}$ (Peters and Wakker, 1986).

To see the bite of Theorem 1 in the context of a concrete example, consider the simple special case outlined in Example 1, where a given SDF, M, prices an arbitrary return, R, so that $\mathbb{E}[MR] = 1$. The observable set in this application is depicted in the panel (a) of Figure 1, with its convex hull depicted as the shaded region in panel (b). By definition, any point $(m,r) \notin \mathcal{O}$ is not observable. But that does not mean that $\mathbb{E}[M^m R^r]$ can take any arbitrary value. Rather, the empirical and theoretical restrictions captured by the observable set—together with the convexity of the CGF—impose an upper bound on the value of $\log \mathbb{E}[M^m R^r]$ for any (m,r) in the shaded region. This upper bound is what is characterized by Theorem 1.

3.2 Jensen Lower Bounds

Our next result provides a lower bound on the joint moments of $\mathbf{M}$ and $\mathbf{R}$. Given the set of observable moments, $\mathcal{O}$, define

$$\underline{\mathcal{O}} = \left\{ \sum_{i=1}^{p} \lambda_i x_i \,\middle|\, x_1, \ldots, x_p \in \mathcal{O},\ \sum_{i=1}^{p} \lambda_i = 1,\ \text{and } \lambda_i > 0 \geq \lambda_j\ \forall j \neq i \right\}. \tag{9}$$

Set $\underline{\mathcal{O}}$ is similar to the convex hull of the observable set defined in equation (4) in that its elements are linear combinations of elements of $\mathcal{O}$ with weights that add up to one, except that only one of the weights is positive. This means that a point belongs to $\underline{\mathcal{O}}$ if a convex combination of that point and elements of $\mathcal{O}$ falls inside set $\mathcal{O}$.[4] We have the following result:

Theorem 2. *Let M and R denote collections of SDFs and returns and let $\mathcal{O}$ denote the corresponding set of observable moments. Then,*

$$\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] \geq \underline{\mathrm{J}}_{\mathcal{O}}(m,r) \qquad \textit{for all } (m,r) \in \underline{\mathcal{O}}, \tag{10}$$

where

$$\underline{\mathrm{J}}_{\mathcal{O}}(m,r) = \sup_{\substack{\lambda \geq 1, F \\ (\hat{m},\hat{r}) \in \mathcal{O}}} \quad (1-\lambda)\int_{\mathcal{O}} \log \mathbb{E}[\mathbf{M}^{\tilde{m}}\mathbf{R}^{\tilde{r}}]\, \mathrm{d}F(\tilde{m},\tilde{r}) + \lambda \log \mathbb{E}[\mathbf{M}^{\hat{m}}\mathbf{R}^{\hat{r}}] \tag{11}$$

$$\text{s.t.} \quad (1-\lambda)\int_{\mathcal{O}} \tilde{m}\, \mathrm{d}F(\tilde{m},\tilde{r}) + \lambda \hat{m} = m \tag{12}$$

$$(1-\lambda)\int_{\mathcal{O}} \tilde{r}\, \mathrm{d}F(\tilde{m},\tilde{r}) + \lambda \hat{r} \ = r \tag{13}$$

and F is a probability distribution with support over $\mathcal{O}$.

Theorem 2 is the lower bound counterpart to Theorem 1: it provides a bound on certain unobservable moments of SDFs and returns in terms of their observable moments. Also similar to Theorem 1, it is a consequence of Jensen's inequality and the fact that the CGF, $\phi(m,r) = \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$, is convex in (m,r). The difference between the two results is also noteworthy: whereas $\overline{\mathrm{J}}_{\mathcal{O}}(m,r)$ in Theorem 1 is the largest convex extension of $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$ to the convex hull of the observable set, $\underline{\mathrm{J}}_{\mathcal{O}}(m,r)$ in Theorem 2 is a partial extension of $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$ to outside of the convex hull of the observable set. The difference between the regions over which the two bounds are operational can be seen most easily in the context of Example 1 and by comparing the shaded and the dotted regions in Figure 1(b).

Note that even though we expressed Theorems 1 and 2 in terms of unconditional expectations, similar results can be obtained for conditional expectations: as long as one replaces all expectation operators with the same conditional expectation, both results would go through with no other change.

As a final remark, we note that the fact that the bounds in Theorems 1 and 2 are expressed as solutions to optimization problems is due to the fact that, in general, and depending on the application, the observable set $\mathcal{O}$ may take a fairly complex form. While one can always calculate these bounds numerically, in many relevant instances, it is possible to obtain closed-form expressions. We defer such characterization results to Section 4.

3.3 Basic Properties of Jensen Bounds

We conclude this section by establishing a few basic properties of the bounds in Theorems 1 and 2. While fairly intuitive, these properties illustrate how the empirical or theoretical restriction on the joint distribution of SDFs and returns, as summarized by the observable set, are incorporated into $\overline{\mathrm{J}}$ and $\underline{\mathrm{J}}$.

[4] See the proof of Theorem 2 for more details.

Proposition 1. *If* $(m, r) \in \mathcal{O}$*, then* $\underline{\mathrm{J}}_{\mathcal{O}}(m, r) = \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] = \overline{\mathrm{J}}_{\mathcal{O}}(m, r)$*.*

This result shows that both bounds are tight over the set of observable moments. This is, of course, by construction: $\overline{\mathrm{J}}_{\mathcal{O}}(m, r)$ and $\underline{\mathrm{J}}_{\mathcal{O}}(m, r)$ are extensions of the CGF, $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$, beyond the observable set. As such, they must coincide with $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$ (and one another) over the observable set.

Proposition 2. *Suppose* $\mathcal{O} \subseteq \mathcal{O}_\star$*. Then,*

$$\overline{\mathrm{J}}_{\mathcal{O}_\star}(m, r) \leq \overline{\mathrm{J}}_{\mathcal{O}}(m, r) \qquad \textit{for all } (m, r) \in \overline{\mathcal{O}} \tag{14}$$

$$\underline{\mathrm{J}}_{\mathcal{O}_\star}(m, r) \geq \underline{\mathrm{J}}_{\mathcal{O}}(m, r) \qquad \textit{for all } (m, r) \in \underline{\mathcal{O}}. \tag{15}$$

This result, which is yet another immediate consequence of Theorems 1 and 2, shows that enlarging the set of observable moments can only result in tighter bounds. Thus, any additional empirical or theoretical restriction on the joint distribution of SDFs and returns is incorporated into $\overline{\mathrm{J}}$ and $\underline{\mathrm{J}}$ by making both bounds tighter.

4 Characterization Results

The upper and lower Jensen bounds presented in Theorems 1 and 2 are expressed as solutions to optimization problems. This reflects the fact that, in general, and depending on the application, the observable set $\mathcal{O}$ may take a fairly complex form. While one can always calculate these bounds numerically, in many instances, it is possible to obtain closed-form expressions for $\overline{\mathrm{J}}$ and $\underline{\mathrm{J}}$. In this section, we provide such characterization results for three canonical applications: (i) a univariate case along the lines of Example 1, (ii) international financial markets, and (iii) the permanent-transitory decomposition of SDFs in dynamic economies. In addition to being useful on their own right, these characterization results illustrate how to operationalize the family of Jensen bounds in various contexts.

4.1 Univariate Case

As our starting point, we consider the simple univariate case outlined in Example 1, where an SDF, M, prices an arbitrary return, R, i.e., $\mathbb{E}[MR] = 1$. Recall from the discussion in Example 1 that if moments of R can be measured using return data, the observable set is given by $\mathcal{O} = \{(0, r) : r \in \mathbb{R}\} \cup \{(1, 1)\}$ (depicted in panel (a) of Figure 1). Using Theorems 1 and 2 to bound $\log \mathbb{E}[M^m R^r]$, while setting $m = \alpha$ and $r = 0$, leads to the following result:

Proposition 3. *Let* R *denote an arbitrary return priced by SDF,* M*, so that* $\mathbb{E}[MR] = 1$*.*

(a) If $\alpha \in (0, 1)$*, then*

$$\log \mathbb{E}[M^\alpha] \leq (1 - \alpha) \log \mathbb{E}[R^{\alpha/(\alpha-1)}]. \tag{16}$$

(b) If $\alpha < 0$ or $\alpha > 1$, then

$$\log \mathbb{E}[M^{\alpha}] \geq (1-\alpha) \log \mathbb{E}[R^{\alpha/(\alpha-1)}]. \tag{17}$$

This result provides upper and lower bounds on various moments of the SDF in terms of the moments of the return distribution. The bound in (16) is the Jensen upper bound of Theorem 1, while the bound in (17) is the Jensen lower bound of Theorem 2. That these bounds have to hold for any return means that they can be used to test the "viability" of an asset-pricing model—as summarized by its implied SDF—using return data.

The bounds in Proposition 3 nest multiple well-known bounds as special cases. First, note that setting $\alpha = 2$ in (17) implies that $\mathbb{E}[M^2] \geq 1/\mathbb{E}[R^2]$, which is equivalent to the HJ bound.[5] Next, observe that dividing both sides of (16) by α and taking the limit as $\alpha \to 0$ implies that $\mathbb{E}[\log M] \leq -\mathbb{E}[\log R]$. This is the entropy bound of Alvarez and Jermann (2005). More generally, the bounds in (16) and (17) together recover the entire family of bounds characterized by the prior work of Snow (1991).

While the bounds in (16) and (17) are not novel, Proposition 3 plays three roles for our purposes. First, it establishes the relationship between the Jensen bounds and earlier well-known bounds, such as the HJ and entropy bounds. Second, it illustrates how to operationalize the results in Theorem 1 and 2 in the simplest possible setting. Finally, this proposition, which bounds the marginal moments of M, provides us with a natural benchmark for our subsequent results, where we obtain restrictions on the joint moments of multiple SDFs.

4.2 International Financial Markets

In this subsection, we use our general results in Theorems 1 and 2 to obtain novel bounds on the joint moments of stochastic discount factors of multiple countries in an international asset-pricing setting.

The setting is similar to the one outlined in Example 2. Consider an economy consisting of two countries, labeled 1 and 2, each with its own currency, where investors in each country can trade assets denominated in their respective currencies. Let $\mathbf{R} = (R_1, R_2)$ denote an arbitrary pair of returns that can be traded by investors in the two respective countries. Similarly, let $\mathbf{M} = (M_1, M_2)$, where M_i is the SDF that prices R_i from the perspective of investors in country i (i.e., in country i's currency). This means that $\mathbb{E}[M_1 R_1] = \mathbb{E}[M_2 R_2] = 1$. As described in Example 2, if the joint distribution of R_1 and R_2 is empirically observable, the theoretical and empirical restrictions on the joint distribution of $\mathbf{M}$ and $\mathbf{R}$ are encoded by the observable set in (2). Applying Theorems 1 and 2 to this setting to bound $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$ for $m = (\alpha, \beta)$ and $r = (0, 0)$ leads to the following result:

Proposition 4. *Let R_1 and R_2 denote arbitrary returns that can be traded by investors in the countries 1 and 2, respectively.*

(a) If $\alpha, \beta \geq 0$ and $\alpha + \beta \in (0, 1)$, then

$$\log \mathbb{E}[M_1^{\alpha} M_2^{\beta}] \leq (1-\alpha-\beta) \log \mathbb{E}[(R_1^{\alpha} R_2^{\beta})^{1/(\alpha+\beta-1)}]. \tag{18}$$

[5] See Appendix A for more details.

(b) If either (i) $\alpha, \beta < 0$ or (ii) $\alpha + \beta > 1$ and $\min\{\alpha, \beta\} < 0$, then

$$\log \mathbb{E}[M_1^\alpha M_2^\beta] \geq (1 - \alpha - \beta) \log \mathbb{E}[(R_1^\alpha R_2^\beta)^{\frac{1}{\alpha+\beta-1}}]. \tag{19}$$

This result provides upper and lower bounds on the joint moments of the two countries' SDFs in terms of return distributions. Specifically, $\log \mathbb{E}[M_1^\alpha M_2^\beta]$ has to satisfy the bounds in (18) and (19) for any pair of returns, R_1 and R_2, that are tradable by investors in their respective countries. Thus, as in the univariate bounds in Proposition 3, these bounds, which can be estimated given the time series of financial assets, can serve as tests of asset-pricing models.

A few remarks are in order. First, note that, in deriving Proposition 4, we are not imposing any restrictions on the market structure: international financial markets can be fully integrated (in the sense that any asset that can be traded by investors in one country can be traded by investors in the other country), they can be fully segmented, or anything in between. Changes in the market structure simply changes the set of tradable returns R_1 and R_2, without altering the functional forms of inequalities (18) and (19).[6]

Next, note that inequalities (18) and (19) reduce to their univariate counterparts (16) and (17) if either α or β is equal to zero. But as long as both exponents are non-zero, the bounds in Proposition 4 represent non-trivial restrictions on the joint distribution of M_1 and M_2; a result with no counterpart in the prior literature. Figure 2 illustrates the regions over which these bounds are operational and compares them to the corresponding regions for the univariate bounds of Proposition 3. As is evident from the figure, an international-asset pricing model may well satisfy the HJ bound (or even the entire family of univariate bounds of Proposition 3 on the marginal distributions), while violating the multivariate bounds of Proposition 4 on the joint distribution of SDFs.

The implications of Proposition 4 for the joint distribution of the two SDFs can be seen most transparently when $\log M_1$ and $\log M_2$ are jointly normally distributed (as in, e.g., Lustig and Verdelhan (2019)). Under such an assumption, inequalities (18) and (19) provide non-trivial restrictions on the covariance of log SDFs in terms of log excess returns. To see this, let

$$\Delta(\alpha, \beta) = (1 - \alpha - \beta) \log \mathbb{E}[(e^{\alpha \mathrm{rx}_1} e^{\beta \mathrm{rx}_2})^{1/(\alpha+\beta-1)}], \tag{20}$$

where $\mathrm{rx}_i = \log R_i - \log R_{fi}$ denotes the log excess return of an arbitrary asset available to investors in country i and R_{fi} is the corresponding risk-free return. We have the following result:

Corollary 1. *Suppose $\log M_1$ and $\log M_2$ are jointly normally distributed and let $\Delta(\alpha, \beta)$ be given by (20).*

(a) If $\alpha, \beta \geq 0$ and $\alpha + \beta \in (0, 1)$, then

$$\frac{1}{2}\alpha(\alpha - 1)\operatorname{var}(\log M_1) + \frac{1}{2}\beta(\beta - 1)\operatorname{var}(\log M_2) + \alpha\beta \operatorname{cov}(\log M_1, \log M_2) \leq \Delta(\alpha, \beta). \tag{21}$$

[6] This is analogous to the HJ bound, which has to hold for any tradable return R: while expanding the set of tradable returns tightens the HJ bound, the statement of the inequality itself remains unchanged.

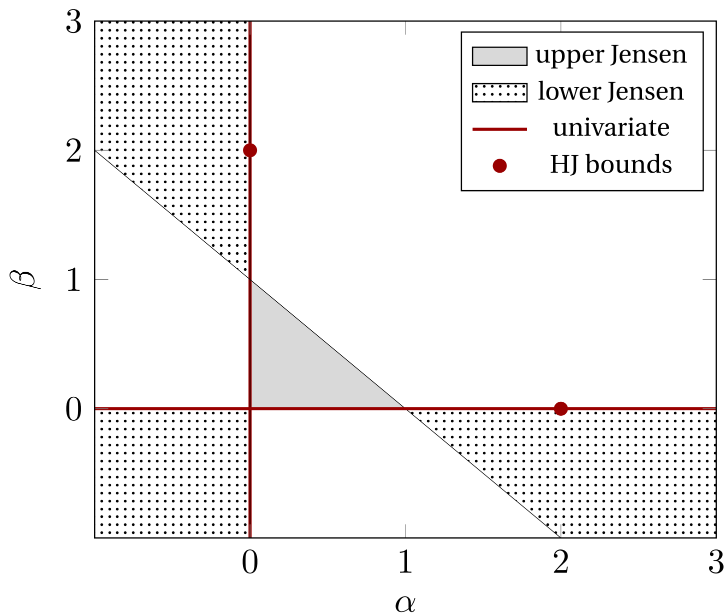


Figure 2. This figure depicts the regions for upper and lower Jensen bounds for $\log \mathbb{E}[M_1^\alpha M_2^\beta]$ in the (α, β) space, where M_1 and M_2 are the SDFs of countries 1 and 2 in the international economy in Subsection 4.2.

(b) If either (i) $\alpha, \beta < 0$ *or (ii)* $\alpha + \beta > 1$ *and* $\min\{\alpha, \beta\} < 0$*, then*

$$\frac{1}{2}\alpha(\alpha-1)\,\mathrm{var}(\log M_1) + \frac{1}{2}\beta(\beta-1)\,\mathrm{var}(\log M_2) + \alpha\beta\mathrm{cov}(\log M_1, \log M_2) \geq \Delta(\alpha, \beta). \tag{22}$$

Corollary 1 establishes that, when the two SDFs are log-normally distributed, different values of α and β in (21) and (22) (or equivalently in (18) and (19)) translate into different bounds on weighted sums of variances and covariance of $\log M_1$ and $\log M_2$. When either α or β is equal to zero, the term corresponding to the covariance drops out, turning (21) and (22) into bounds on the marginals. But when both α and β are different from zero, the return distribution imposes non-trivial bounds on the covariance of $\log M_1$ and $\log M_2$, an object that measures the degree of risk-sharing between the two countries (Brandt, Cochrane, and Santa-Clara, 2006).

The bounds in (18) and (19) are informative even beyond the log-normal benchmark discussed above. However, in that case, the bounds in Proposition 4 impose restrictions not only on variances and covariance of $\log M_1$ and $\log M_2$, but also on their higher-order (joint) cumulants, such as their skewness and coskewness. We return to this discussion when we study an empirical application of our results in Section 5.

4.3 Permanent-Transitory Decomposition of SDFs

As our final characterization result, we apply Theorems 1 and 2 to a dynamic economy to construct bounds on the joint moments of permanent and transitory components of SDF processes.

The setting is similar to the one outlined in Example 3. Consider a discrete-time dynamic economy with stochastic discount factor process $\{M_{t,t+\tau}\}_{t\geq 1}$ that satisfies $\mathbb{E}_t[M_{t,t+\tau}R_{t,t+\tau}] = 1$ for all t and all

$\tau \geq 1$, where $R_{t,t+\tau}$ is the gross return on an arbitrary traded asset over the period from t to $t+\tau$ and $\mathbb{E}_t[\cdot]$ denotes the information set at time t. Alvarez and Jermann (2005) and Hansen and Scheinkman (2009) show that the SDF can be decomposed into permanent and transitory components:

$$M_{t,t+\tau} = M^P_{t,t+\tau} M^T_{t,t+\tau} \tag{23}$$

The permanent component in the above decomposition is a martingale, whereas the transitory component is equal to the reciprocal of the return to holding a discount bond of (asymptotically) long maturity from date t to date $t+\tau$: $M^T_{t,t+\tau} = 1/R^{(\infty)}_{t,t+\tau}$.

To obtain bounds on the joint moments of the permanent and transitory components of the SDF, we first cast the above setting as a special case of our general framework in Section 2. Specifically, we set $\mathbf{M} = (M^P_{t,t+\tau}, M^T_{t,t+\tau})$ and $\mathbf{R} = R_{t,t+\tau}$. Assuming that the joint distribution of $R_{t,t+\tau}$ and the infinite-maturity bond can be measured using return data, we can apply Theorems 1 and 2 to bound $\log \mathbb{E}_t[\mathbf{M}^m \mathbf{R}^r]$ for $m = (\alpha, \beta)$ and $r = 0$. The next proposition summarizes the result:

Proposition 5. *Let $M^P_{t,t+\tau}$ and $M^T_{t,t+\tau}$ denote the permanent and transitory components of the SDF process (23), respectively. Also, $R_{t,t+\tau}$ be an arbitrary return such that $\mathbb{E}_t[M_{t,t+\tau}R_{t,t+\tau}] = 1$.*

(a) If $\alpha \in (0,1)$, then

$$\log \mathbb{E}_t[(M^P_{t,t+\tau})^\alpha (M^T_{t,t+\tau})^\beta] \leq (1-\alpha) \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{\frac{\beta-\alpha}{\alpha-1}} (R_{t,t+\tau})^{\frac{\alpha}{\alpha-1}}]. \tag{24}$$

(b) If $\alpha > 1$ or $\alpha < 0$, then

$$\log \mathbb{E}_t[(M^P_{t,t+\tau})^\alpha (M^T_{t,t+\tau})^\beta] \geq (1-\alpha) \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{\frac{\beta-\alpha}{\alpha-1}} (R_{t,t+\tau})^{\frac{\alpha}{\alpha-1}}]. \tag{25}$$

5 Applications

In this section, we showcase our general framework and the characterization results of Sections 3 and 4 by considering two empirical applications. As our first application, we apply our framework to calculate bounds on the correlation of international SDFs which is intimately linked to the variance of exchange rates. According to the exchange rate volatility puzzle, the relatively low volatility of exchange rates together with the high volatility of SDFs needed to match high Sharpe ratios in asset returns, implies SDF correlations that are almost perfect. In complete markets, SDF correlations correspond to correlations of consumption growth across countries which are significantly lower. In the following, we study bounds on SDF correlations without imposing any assumptions on market structure.

5.1 Application 1: Bounds on SDF Correlations

To describe the comovement of SDFs, let us first introduce the following measures of SDF dispersion:

$$C_{12}(\alpha,\beta) := \frac{\log \mathbb{E}[M_1^\alpha M_2^\beta] - \log \mathbb{E}[M_1^\alpha] - \log \mathbb{E}[M_2^\beta]}{\alpha\beta}\,, \tag{26}$$

and

$$D_1(\alpha) := \frac{2}{\alpha(\alpha-1)}(\log \mathbb{E}[M_1^\alpha] - \alpha \log \mathbb{E}[M_1])\;;\; D_2(\beta) := \frac{2}{\beta(\beta-1)}(\log \mathbb{E}[M_2^\beta] - \beta \log \mathbb{E}[M_2])\,. \tag{27}$$

Note that in the case where the two SDFs are jointly log-Normal, the expression in equation (26) directly corresponds to the covariance of the two SDFs. Now, recall from Proposition 4 that we can upper and lower bound joint moments of two SDFs, corresponding to SDFs of different countries. It therefore follows from part (b) for any $\alpha, \beta < 0$, that

$$C_{12}(\alpha,\beta) \geq \frac{1}{\alpha\beta}\left[(1-\alpha-\beta)\log \mathbb{E}[(R_1^\alpha R_2^\beta)^{\frac{1}{\alpha+\beta-1}}] - \log \mathbb{E}[M_1^\alpha] - \log \mathbb{E}[M_2^\beta]\right],$$

where R_1 and R_2 are the returns traded by investors in country 1 and 2, respectively. Moreover,

$$\begin{aligned}\log \mathbb{E}[M_1^\alpha] &= \frac{\alpha(\alpha-1)}{2} D_1(\alpha) + \alpha \log R_1^f\,,\\ \log \mathbb{E}[M_2^\beta] &= \frac{\beta(\beta-1)}{2} D_2(\beta) + \beta \log R_2^f\,.\end{aligned}$$

Here, note that for any admissible marginal dispersion $D_1(\alpha), D_2(\beta)$ of an SDF, quantities $\log \mathbb{E}[M_1^\alpha], \log \mathbb{E}[M_1^\beta]$ are uniquely determined by the bond pricing equation. Therefore, to bound SDF correlation, we get the following lower bound:

$$R_{12}(\alpha,\beta) := \frac{C_{12}(\alpha,\beta)}{\sqrt{D_1(\alpha)D_2(\beta)}} \tag{28}$$

$$\geq \frac{\frac{1}{\alpha\beta}\left[(1-\alpha-\beta)\log \mathbb{E}[(R_1^\alpha R_2^\beta)^{\frac{1}{\alpha+\beta-1}}] - \log \mathbb{E}[M_1^\alpha] - \log \mathbb{E}[M_2^\beta]\right]}{\sqrt{D_1(\alpha)D_2(\beta)}} \tag{29}$$

We can now directly implement expression (29) using returns data for the US and the UK. In particular, we assume that investors can trade both the short-term bond and the equity in each country, i.e., markets are fully integrated. Figure 3 depicts the attained values for different powers α and β. The result reveals that the tightest possible lower bound is 86%. Recall that in order to use our bounds, we did not make any assumptions about market structure such as the degree of completeness. The high value implied by our bounds implies that surely, lowering the degree of spanning in financial markets is not helpful in addressing the exchange rate volatility puzzle, echoing the findings in Lustig and Verdelhan (2019).

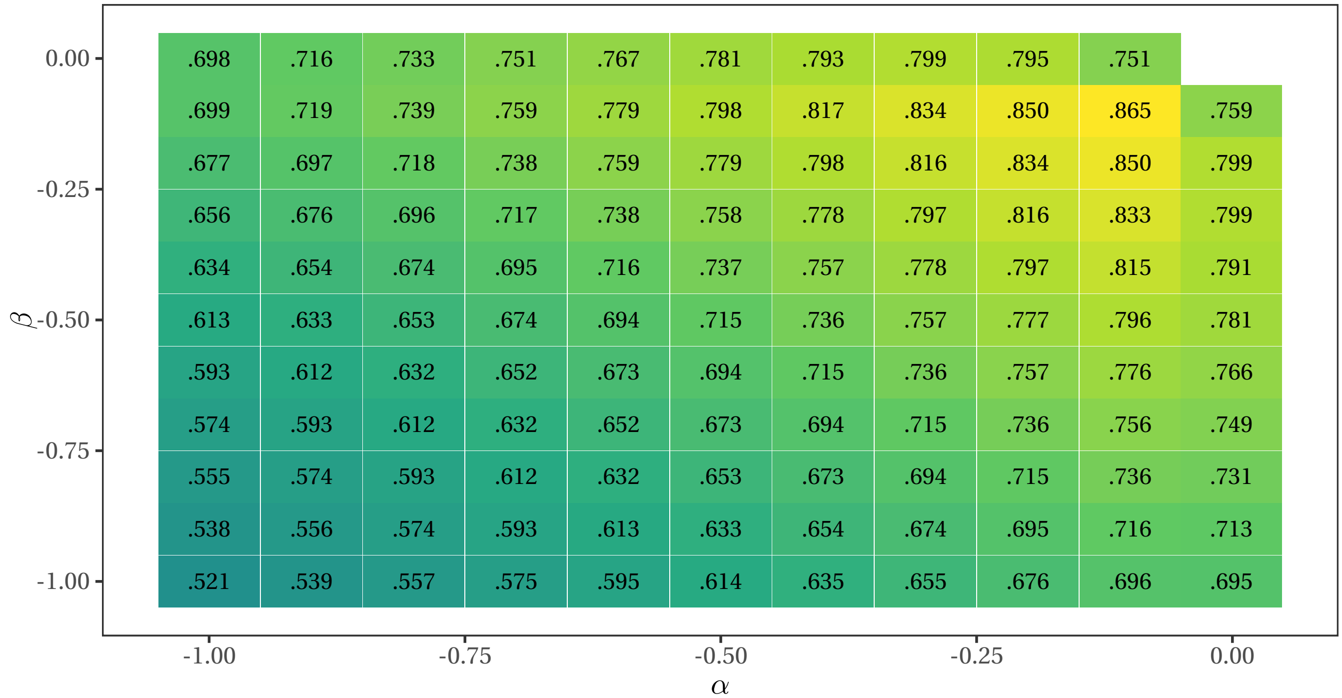


Figure 3. This figure depicts lower bounds using equation (29). α and β indicate, respectively, the power of the SDFs of country 1 (the U.S.) and country 2 (the U.K.) .

5.2 Application 2: Diagnosing International Asset Pricing Models

As our main application, we illustrate how our results in Proposition 4 can be used to test international asset-pricing models. We do so by applying our results to the model of Colacito and Croce (2013), who study a general equilibrium two-country economy with long-run risk in consumption and Epstein and Zin (1989) preferences. Colacito and Croce (2013) show that such a model can account for the forward premium puzzle as well as other asset-pricing anomalies.[7]

The economy consists of two countries, labeled 1 and 2, each endowed with a single good. Endowments of the goods follow cointegrated dynamics:

$$\log x_{1,t} = \mu_1 + \log x_{1,t-1} + z_{1,t-1} + \tau(\log x_{2,t-1} - \log x_{1,t-1}) + \epsilon_{1,t}$$
$$\log x_{2,t} = \mu_2 + \log x_{2,t-1} + z_{2,t-1} - \tau(\log x_{2,t-1} - \log x_{1,t-1}) + \epsilon_{2,t},$$

where $x_{i,t}$ is the time-t endowment of the good in country i, τ measures the degree of cointegration, and $z_{1,t}$ and $z_{2,t}$ are highly persistent AR(1) processes:

$$z_{i,t} = \rho_i z_{i,t-1} + \varepsilon_{i,t}.$$

Due to their long-lasting impact on the growth rate of endowments, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are known as long-run shocks, whereas $\epsilon_{1,t}$ and $\epsilon_{2,t}$ are referred to as short-run shocks. The vector of shocks $\xi_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \epsilon_{1,t}, \epsilon_{2,t})$ is i.i.d. over time and normally distributed with mean 0 and covariance matrix Σ.

[7] Also see Colacito and Croce (2011).

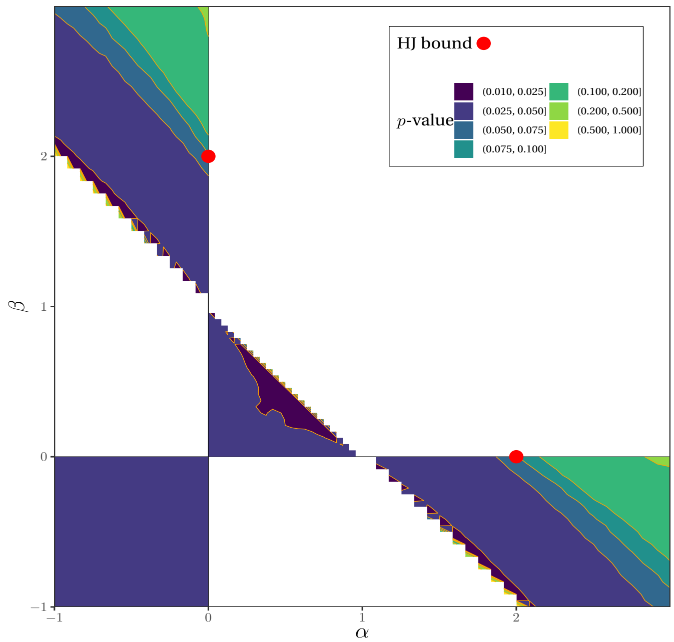


Figure 4. This figure depicts bootstrapped p-values for testing Proposition 4 for the Colacito and Croce (2013) model. α and β indicate, respectively, the power of the SDFs of country 1 (the U.S.) and country 2 (the U.K.) in equations (18) and (19). The red dot indicates the univariate HJ bound.

The representative household in country $i \in \{1, 2\}$ has Epstein-Zin preferences given by

$$U_{i,t} = \left\{(1-\delta)(C_{i,t})^{1-1/\psi} + \delta \mathbb{E}_t \left[(U_{i,t+1})^{1-\gamma}\right]^{\frac{1-1/\psi}{1-\gamma}}\right\},$$

where γ parametrizes the degree of relative risk aversion and ψ is the intertemporal elasticity of substitution (IES), and the consumption bundle. This implies that country i's SDF is given by

$$M_{i,t} = \delta \left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-1/\psi} \left(\frac{(U_{i,t})^{1-\gamma}}{\mathbb{E}_{t-1}[(U_{i,t})^{1-\gamma}]}\right)^{\frac{1/\psi-\gamma}{1-\gamma}}. \tag{30}$$

Finally, we assume that the consumption bundle in each country is a Cobb-Douglas aggregate of the two goods: $C_{i,t} = c_{ii,t}^{\kappa} c_{ij,t}^{1-\kappa}$, where $c_{ii,t}$ and $c_{ij,t}$ denote the consumption of goods i and j by the representative household in country i and $\kappa > 1/2$ captures the degree of home bias in the economy.

Testing Proposition 4 proceeds in three steps. First, we simulate time-series of SDFs according to (30) to calculate the left-hand side expressions of parts (a) and (b) of Proposition 4 for different powers α and β. More specifically, to simulate U.S. and U.K. SDFs, we use the following parameters (see Colacito and Croce (2013) and Colacito et al. (2018)): risk aversion (γ) is 6.5, IES (ψ) is 1.6, subjective discount factor (δ) is 0.98, home bias (κ) is 0.97, mean of endowment growth is 2%, short-run risk volatility is 1.87%, long-run risk volatility is 6%, long-run risk correlation is 0.4, and the cross-correlation of long-run shocks is 0.985. For powers α and β, we follow Figure 2.

Second, to calculate the right-hand side of parts (a) and (b) of Proposition 4, we optimize the bounds using pricing restrictions in each country for the risk-free rate and the equity. To this end, we use data from Datastream for short-term interest rates and MSCI equity indices for the U.S. and U.K. Third, we obtain bootstrapped p-values for each estimate and each power α and β by resampling from the distribution of asset returns. Bootstrapped p-values from 20,000 samples are presented in Figure 4.

We first notice that the model passes the univariate HJ bounds as indicated by the red dot along each axis. More specifically, p-values are 7.28% for $\alpha = 2$, $\beta = 0$ and 7.22% for $\alpha = 0$, $\beta = 2$. However, we find that the model is everywhere rejected when considering the lower bound as p-values are everywhere below 5% or even below 2.5%. Interestingly, we also find rejections in the negative quadrant where p-values are consistently below 5%. To understand why it is more difficult to pass the bounds in the region $\alpha, \beta < 0$, recall from equation (19) that potentially the model does not generate enough negative coskewness between M_1 and M_2. It is, however, easy to introduce more joint skewness into domestic and foreign SDFs by adding rare global disasters in consumption as for example in Farhi and Gabaix (2016) or Barro and Jin (2021).

6 Conclusion

This paper develops a parsimonious model-free framework for testing asset pricing models. Our main result provides testable bounds on the joint distribution of pricing kernels in the presence of empirical (such as observability of certain returns) and theoretical (such as no arbitrage) restrictions. We call this set of restrictions, the observable set. Making use of the convexity of the joint cumulant-generating function of SDFs and returns and Jensen's inequality allows us to obtain both upper and lower bounds on the unobservable moments of the joint distribution of pricing kernels and returns. Since the bounds only depend on observables, they can be easily calculated from the data, allowing for diagnosing asset pricing models.

Different from earlier literature which has focused exclusively on univariate settings, our methodology can be used to obtain testable bounds for the joint distribution of multiple SDFs. One prime candidate for testing our framework is in the context of international finance, as exchange rates are tightly linked to how SDFs in different countries comove with one another. To illustrate our methodology, we take a canonical international finance model and calculate lower and upper bounds for different moments of domestic and foreign SDFs. We find that while the standard HJ bounds would not have uncovered any violation of the bounds, our multivariate bounds indicate that the model can be rejected in several regions.

Our results in this paper rely on a dichotomous framework, where a given (joint) moment of SDFs and returns is either perfectly observable or not observable at all. However, in many instances, it is reasonable to expect that certain moments are observable but only imperfectly, for example, due to measurement error, a misspecified SDF, or frictions that can lead to non-zero pricing errors. Incorporating these non-zero pricing errors into our framework is an important direction for future research.

A Appendix: Relation to Other Bounds in the Literature

In this appendix, we discuss how the various bounds developed throughout the paper relate to bounds in the prior literature. Specifically, we show that the family of Jensen bounds in Propositions 3–5 nest the HJ bound, the entropy bounds of Alvarez and Jermann (2005), and the variance bounds of Bakshi and Chabi-Yo (2012) as special cases.

A.1 The HJ Bound

According to the well-known bound of Hansen and Jagannathan (1991), the ratio of the standard deviation of a stochastic discount factor to its mean exceeds the Sharpe ratio attained by any portfolio that is priced by that SDF. Specifically, if R_f denotes the risk-free rate, then

$$\frac{\text{stdev}(M)}{\mathbb{E}[M]} \geq \frac{\mathbb{E}[R - R_f]}{\text{stdev}(R - R_f)}, \tag{A.1}$$

for any R such that $\mathbb{E}[MR] = 1$. In what follows, we show that the HJ bound can be derived as a special case of the family of bounds in Proposition 3 by setting $\alpha = 2$ in inequality (17).

Setting $\alpha = 2$ in (17) implies that $\mathbb{E}[M^2] \geq 1/\mathbb{E}[\tilde{R}^2]$ for any portfolio of tradable returns $\tilde{R}$ that is priced by the SDF M, i.e., for all $\tilde{R}$ such that $\mathbb{E}[M\tilde{R}] = 1$. In particular, $\mathbb{E}[M^2] \geq 1/\mathbb{E}[(\lambda R + (1-\lambda)R_f)^2]$ for any portfolio weight $\lambda \in \mathbb{R}$ and any return R such that $\mathbb{E}[MR] = 1$. Maximizing the right-hand side of this inequality over all λ implies that

$$\mathbb{E}[M^2] \geq \max_{\lambda \in \mathbb{R}} \frac{1}{\mathbb{E}[(\lambda R + (1-\lambda)R_f)^2]} = \frac{1}{R_f^2} \frac{\mathbb{E}[(R - R_f)^2]}{\mathbb{E}[(R - R_f)^2] - (R_f - \mathbb{E}[R])^2}.$$

Since $R_f = 1/\mathbb{E}[M]$, we can rewrite the above inequality as

$$\frac{\text{var}(M)}{\mathbb{E}^2[M]} \geq \frac{(R_f - \mathbb{E}[R])^2}{\mathbb{E}[(R - R_f)^2] - (R_f - \mathbb{E}[R])^2} = \frac{\mathbb{E}^2[R - R_f]}{\text{var}(R - R_f)}.$$

Taking square roots of both sides of the above inequality then leads to (A.1). Therefore, the HJ bound can be obtained as a special case of Proposition 3 by setting $\alpha = 2$.

A.2 Entropy Bounds

Alvarez and Jermann (2005) and Backus, Chernov, and Zin (2014) show that the entropy of an SDF is lower bounded by (log) excess return of any tradable asset:

$$L_t(M_{t,t+1}) \geq \mathbb{E}_t[\log R_{t,t+1}] - \log R_{t,t+1}^{(1)}, \tag{A.2}$$

$M_{t,t+1}$ is the one-period SDF, $L_t(x) = \log \mathbb{E}_t[x] - \mathbb{E}_t[\log x]$ is the conditional entropy of x, and $R_{t,t+1}^{(1)}$ is the return on a one-period risk-free bond. Alvarez and Jermann (2005, Propositions 2 and 3) additionally establish that the permanent and transitory components of the SDF process satisfy the following inequalities:

$$L_t(M^P_{t,t+1}) \geq \mathbb{E}_t[\log R_{t,t+1}] - \mathbb{E}_t[\log R^{(\infty)}_{t,t+1}] \tag{A.3}$$

$$L(M^T_{t,t+1})/L(M_{t,t+1}) \leq \frac{L(1/R^{(\infty)}_{t,t+1})}{\mathbb{E}[\log(R_{t,t+1}/R^{(1)}_{t,t+1})] + L(1/R^{(1)}_{t,t+1})}, \tag{A.4}$$

where $M^P_{t,t+1}$ and $M^T_{t,t+1}$ are, respectively, the permanent and transitory components of $M_{t,t+1}$, $L(x) = \log \mathbb{E}[x] - \mathbb{E}[\log x]$ is the (unconditional) entropy of x, and $R^{(\infty)}_{t,t+1}$ is the return to holding a discount bond of (asymptotically) long maturity from date t to date $t+1$. In what follows, we show that these inequalities can be derived as special cases of the Jensen bounds in Proposition 5.

We start with (A.2). Setting $\alpha = \beta$ in (24) implies that $\log \mathbb{E}_t[M^{\alpha}_{t,t+1}] \leq (1-\alpha)\log \mathbb{E}_t[R^{\alpha/(\alpha-1)}_{t,t+1}]$. Dividing both sides of this inequality by $\alpha > 0$ and taking the limit as $\alpha \downarrow 0$ implies that $\mathbb{E}_t[\log M_{t,t+1}] \leq -\mathbb{E}_t[\log R_{t,t+1}]$. Subtracting $\log \mathbb{E}_t[M_{t,t+1}]$ from both sides of this inequality and noting that $\mathbb{E}_t[M_{t,t+1}] = 1/R^{(1)}_{t,t+1}$ then establishes (A.2).

We next turn to (A.3). Set $\beta = 0$ in (24) to obtain $\log \mathbb{E}_t[(M^P_{t,t+1})^{\alpha}] \leq (1-\alpha)\mathbb{E}_t[(R_{t,t+1}/R^{(\infty)}_{t,t+1})^{\alpha/(\alpha-1)}]$. Dividing both sides of this inequality by $\alpha > 0$ and taking the limit as $\alpha \downarrow 0$ implies that $\mathbb{E}_t[\log M^P_{t,t+1}] \leq \mathbb{E}_t[\log R^{(\infty)}_{t,t+1}] - \mathbb{E}_t[\log R_{t,t+1}]$. But note that, since $\mathbb{E}_t[M^P_{t,t+1}] = 1$, the left-hand side of this inequality is equal to $-L_t(M^P_{t,t+1})$, thus establishing (A.3).

Finally, we show that (A.4) can also be obtained as a special case of Proposition 5. To this end, first recall that the inequality in (24) also holds with unconditional expectations. Setting $\alpha = \beta$, dividing both sides by α, and taking the limit as $\alpha \downarrow 0$ then implies that $L(M_{t,t+1}) \geq \mathbb{E}[\log R_{t,t+1}] - \log \mathbb{E}[R^{(1)}_{t,t+1}]$, where we are using the fact that $\mathbb{E}_t[M_{t,t+1}] = 1/R^{(1)}_{t,t+1}$. Adding and subtracting $\mathbb{E}[\log R^{(1)}_{t,t+1}]$ to the right-hand side of this inequality leads to

$$L(M_{t,t+1}) \geq \mathbb{E}[\log(R_{t,t+1}/R^{(1)}_{t,t+1})] + L(1/R^{(1)}_{t,t+1}).$$

On the other hand, note that, $M^T_{t,t+1} = 1/R^{(\infty)}_{t,t+1}$, which means that $L(M^T_{t,t+1}) = L(1/R^{(\infty)}_{t,t+1})$. Combining this observation with the above inequality then establishes (A.4).

A.3 Variance Bound on the Permanent Component of SDF

Bakshi and Chabi-Yo (2012, Propositions 1) show that the permanent component of an SDF process must satisfy the following inequality:

$$\text{var}_t(M^P_{t,t+1}) \geq \frac{(1 - \mathbb{E}_t[R_{t,t+1}/R^{(\infty)}_{t,t+1}])^2}{\text{var}_t(R_{t,t+1}/R^{(\infty)}_{t,t+1})}, \tag{A.5}$$

where $R^{(\infty)}_{t,t+1}$ denotes the return to holding a discount bond of (asymptotically) long maturity from date t to date $t+1$. In what follows, we show that inequality (A.5) can be derived as a special case of the family of bounds in Proposition 5.

As a first observation, note that setting $\alpha = 2$ and $\beta = 0$ in inequality (25) implies that $\mathbb{E}_t[(M^P_{t,t+1})^2] \geq 1/\mathbb{E}_t[(\tilde{R}_{t,t+1}/R^{(\infty)}_{t,t+1})^2]$ for any portfolio of one-period returns, $\tilde{R}_{t,t+1}$, that is priced by the SDF. Let $R_{t,t+1} = \lambda\tilde{R}_{t,t+1} + (1-\lambda)R^{(\infty)}_{t,t+1}$. Hence, $\mathbb{E}_t[(M^P_{t,t+1})^2] \geq 1/\mathbb{E}_t[(\lambda R_{t,t+1}/R^{(\infty)}_{t,t+1} + 1 - \lambda)^2]$ for any $\lambda \in \mathbb{R}$. Maximizing both sides of this inequality over $\lambda \in \mathbb{R}$ therefore implies that

$$\mathbb{E}_t[(M^P_{t,t+1})^2] \geq \max_{\lambda}\ 1/\mathbb{E}_t[(\lambda R_{t,t+1}/R^{(\infty)}_{t,t+1} + 1 - \lambda)^2] = \frac{\mathbb{E}_t[(R_{t,t+1}/R^{(\infty)}_{t,t+1} - 1)^2]}{\text{var}_t(R_{t,t+1}/R^{(\infty)}_{t,t+1} - 1)}.$$

Subtracting 1 from both sides of the above inequality then establishes (A.5).

B Appendix: Proofs

This appendix contains the proofs of the results presented in the main body of the paper. We start with a simple lemma that serves as the basis of proofs of Theorems 1 and 2.

Lemma B.1. *The cumulant-generating function of log SDFs and log returns, $\phi(m,r) = \log\mathbb{E}[\mathbf{M}^m\mathbf{R}^r]$, is jointly convex in (m,r).*

Proof. Let $m, m' \in \mathbb{R}^k$ and $r, r' \in \mathbb{R}^n$ and suppose $\lambda \in [0,1]$. Since $\mathbf{M}$ and $\mathbf{R}$ are non-negative, Hölder's inequality implies that

$$\mathbb{E}[\mathbf{M}^{\lambda m + (1-\lambda)m'}\mathbf{R}^{\lambda r + (1-\lambda)r'}] = \mathbb{E}[(\mathbf{M}^m\mathbf{R}^r)^\lambda(\mathbf{M}^{m'}\mathbf{R}^{r'})^{1-\lambda}] \leq (\mathbb{E}[\mathbf{M}^m\mathbf{R}^r])^\lambda\,(\mathbb{E}[\mathbf{M}^{m'}\mathbf{R}^{r'}])^{1-\lambda}.$$

Taking logarithms from both sides of the above inequality establishes the result. □

Proof of Theorem 1

Fix a point (m,r) in the convex hull of the observable set, i.e., $(m,r) \in \mathcal{O}$. By definition, there exists a probability distribution F with support over $\mathcal{O}$ such that

$$m = \int_{\mathcal{O}} \tilde{m}\,\mathrm{d}F(\tilde{m},\tilde{r}) \qquad \text{and} \qquad r = \int_{\mathcal{O}} \tilde{r}\,\mathrm{d}F(\tilde{m},\tilde{r}). \tag{B.1}$$

Next, recall from Lemma B.1 that the cumulant-generating function $\phi(m,r) = \log\mathbb{E}[\mathbf{M}^m\mathbf{R}^r]$ is jointly convex in (m,r). Therefore, by Jensen's inequality

$$\log\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] \leq \int_{\mathcal{O}} \log\mathbb{E}[\mathbf{M}^{\tilde{m}}\mathbf{R}^{\tilde{r}}]\,\mathrm{d}F(\tilde{m},\tilde{r}).$$

Taking the infimum from both sides of the above inequality over probability distributions F that satisfy the constraints in (B.1) implies that $\log\mathbb{E}[\mathbf{M}^m\mathbf{R}^r] \leq \bar{\mathrm{J}}_{\mathcal{O}}(m,r)$. □

Proof of Theorem 2

We first show that if $x \in \underline{\mathcal{O}}$, then there exists a convex combination of x and points in $\mathcal{O}$, with a strictly positive weight on x, that falls inside set $\mathcal{O}$. To establish this claim, suppose $x \in \underline{\mathcal{O}}$. Therefore, there exist $x_1, \ldots, x_p \in \mathcal{O}$ and weights $\lambda_1, \ldots, \lambda_p$ such that $x = \sum_{j=1}^{p} \lambda_j x_j$ and λ_1 is the only positive weight. Consequently, $x_1 = (1/\lambda_1)x - \sum_{j=2}^{p}(\lambda_j/\lambda_1)x_j$, which means that there exists a convex combination of x and points in $\mathcal{O}$—with a strictly positive weight on x—that falls inside set $\mathcal{O}$. This establishes the above claim.

With the above representation of $\underline{\mathcal{O}}$ in hand, we now proceed to proving Theorem 2. Fix $(m,r) \in \underline{\mathcal{O}}$. This means there is a constant $\theta \in (0,1]$, a probability distribution F with support over $\mathcal{O}$, and $(\hat{m},\hat{r}) \in \mathcal{O}$ such that

$$\hat{m} = \theta m + (1-\theta)\int_{\mathcal{O}} \tilde{m}\, \mathrm{d}F(\tilde{m},\tilde{r}) \qquad \text{and} \qquad \hat{r} = \theta r + (1-\theta)\int_{\mathcal{O}} \tilde{r}\, \mathrm{d}F(\tilde{m},\tilde{r}). \tag{B.2}$$

Next, recall from Lemma B.1 that the cumulant-generating function $\phi(m,r) = \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$ is jointly convex in (m,r). Therefore, by Jensen's inequality

$$\log \mathbb{E}[\mathbf{M}^{\hat{m}} \mathbf{R}^{\hat{r}}] \leq \theta \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] + (1-\theta)\int_{\mathcal{O}} \log \mathbb{E}[\mathbf{M}^{\tilde{m}} \mathbf{R}^{\tilde{r}}]\, \mathrm{d}F(\tilde{m},\tilde{r}).$$

Consequently,

$$\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] \geq \lambda \log \mathbb{E}[\mathbf{M}^{\hat{m}} \mathbf{R}^{\hat{r}}] + (1-\lambda)\int_{\mathcal{O}} \log \mathbb{E}[\mathbf{M}^{\tilde{m}} \mathbf{R}^{\tilde{r}}]\, \mathrm{d}F(\tilde{m},\tilde{r})$$

where $\lambda = 1/\theta \geq 1$. Taking the supremum from both sides of the above inequality over $\lambda \geq 1$, points $(\hat{m},\hat{r}) \in \mathcal{O}$, and probability distributions F that jointly satisfy the constraints in (B.2) then implies that $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] \geq \underline{\mathrm{J}}_{\mathcal{O}}(m,r)$. □

Proof of Proposition 1

We start by establishing that if $(m,r) \in \mathcal{O}$, then $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] = \overline{\mathrm{J}}_{\mathcal{O}}(m,r)$. Since (m,r) trivially belongs to the convex hull of $\mathcal{O}$, Theorem 1 guarantees that $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] \leq \overline{\mathrm{J}}_{\mathcal{O}}(m,r)$. It is therefore sufficient to establish that $\overline{\mathrm{J}}_{\mathcal{O}}(m,r) \leq \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$. To this end, let F be a distribution with point mass of one on (m,r). This distribution trivially satisfies constraints (7) and (8). Therefore, the definition of $\overline{\mathrm{J}}_{\mathcal{O}}(m,r)$ in equation (6) implies that $\overline{\mathrm{J}}_{\mathcal{O}}(m,r) \leq \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$. Putting the two inequalities together establishes that $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] = \overline{\mathrm{J}}_{\mathcal{O}}(m,r)$ for all $(m,r) \in \mathcal{O}$.

Next, we show that if $(m,r) \in \mathcal{O}$, then $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] = \underline{\mathrm{J}}_{\mathcal{O}}(m,r)$. It is straightforward to verify that $(m,r) \in \underline{\mathcal{O}}$. Therefore, Theorem 2 guarantees that $\log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r] \geq \underline{\mathrm{J}}_{\mathcal{O}}(m,r)$. To complete the proof it is therefore sufficient to show that $\underline{\mathrm{J}}_{\mathcal{O}}(m,r) \geq \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$. To this end, consider the optimization problem in (11)–(13) and set $\lambda = 1$, $(\hat{m},\hat{r}) = (m,r)$, and let F be an arbitrary probability distribution with support over $\mathcal{O}$. It is immediate that these choices satisfy all the constraints of the optimization problem. As a result, the definition of $\underline{\mathrm{J}}_{\mathcal{O}}(m,r)$ in equation (11) implies that $\underline{\mathrm{J}}_{\mathcal{O}}(m,r) \geq \log \mathbb{E}[\mathbf{M}^m \mathbf{R}^r]$. □

Proof of Proposition 2

We first establish the inequality in (14). Suppose $\mathcal{O} \subseteq \mathcal{O}_\star$ and fix a point $(m, r) \in \overline{\mathcal{O}}$. Clearly, $(m, r) \in \overline{\mathcal{O}}_\star$. Furthermore, the fact that $\mathcal{O} \subseteq \mathcal{O}_\star$ means that any probability distribution with support over $\mathcal{O}$ is also a probability distribution with support over $\mathcal{O}_\star$. Consequently, the feasible set in the the optimization problem in (6)–(8) under $\mathcal{O}_\star$ is a superset of the feasible set under $\mathcal{O}$. As a result, $\overline{\mathrm{J}}_{\mathcal{O}_\star}(m, r) \leq \overline{\mathrm{J}}_{\mathcal{O}}(m, r)$.

We next establish (15). Once again, suppose $\mathcal{O} \subseteq \mathcal{O}_\star$ and fix a point $(m, r) \in \underline{\mathcal{O}}$. Clearly, $(m, r) \in \underline{\mathcal{O}}_\star$. Furthermore, the fact that $\mathcal{O} \subseteq \mathcal{O}_\star$ means that any probability distribution with support over $\mathcal{O}$ is also a probability distribution with support over $\mathcal{O}_\star$. Consequently, the feasible set in the the optimization problem in (11)–(13) under $\mathcal{O}_\star$ is a superset of the feasible set under $\mathcal{O}$. As a result, it is immediate that $\underline{\mathrm{J}}_{\mathcal{O}_\star}(m, r) \geq \underline{\mathrm{J}}_{\mathcal{O}}(m, r)$. □

Proof of Proposition 3

Proof of part (a) Recall that the observable set in this setting is given by $\mathcal{O} = \{(0, r) : r \in \mathbb{R}\} \cup \{(1, 1)\}$. This implies that $(\alpha, 0) \in \overline{\mathcal{O}}$ if $0 < \alpha < 1$. Therefore, Theorem 1 implies that $\log \mathbb{E}[M^\alpha] \leq \overline{\mathrm{J}}_{\mathcal{O}}(\alpha, 0)$, where

$$\begin{aligned} \overline{\mathrm{J}}_{\mathcal{O}}(\alpha, 0) = \inf_{\pi \in [0,1], G} \quad & \pi \log \mathbb{E}[MR] + (1 - \pi) \int_{-\infty}^{\infty} \log \mathbb{E}[R^{\tilde{r}}] \, \mathrm{d}G(\tilde{r}) \\ \text{s.t.} \quad & \pi = \alpha \\ & \pi + (1 - \pi) \int_{-\infty}^{\infty} \tilde{r} \, \mathrm{d}G(\tilde{r}) = 0 \end{aligned}$$

and G is a probability distribution over the real line. Note that we are using the fact that a general distribution F with support over $\mathcal{O}$ can be expressed as point mass π on point $(1, 1)$ and a mass $1 - \pi$ over the set $\{(0, r) : r \in \mathbb{R}\}$. Since $\mathbb{E}[MR] = 1$, we have

$$\begin{aligned} \overline{\mathrm{J}}_{\mathcal{O}}(\alpha, 0) = \inf_{G} \quad & (1 - \alpha) \int_{-\infty}^{\infty} \log \mathbb{E}[R^{\tilde{r}}] \, \mathrm{d}G(\tilde{r}) \\ \text{s.t.} \quad & \int_{-\infty}^{\infty} \tilde{r} \, \mathrm{d}G(\tilde{r}) = \alpha/(\alpha - 1) \end{aligned}$$

Now the observation that $\log \mathbb{E}[R^{\tilde{r}}]$ is convex in $\tilde{r}$ implies that the distribution G that minimizes the objective function puts a mass of one on $\tilde{r} = \alpha/(\alpha - 1)$. Consequently, $\overline{\mathrm{J}}_{\mathcal{O}}(\alpha, 0) = (1 - \alpha) \log \mathbb{E}[R^{\alpha/(\alpha-1)}]$. Therefore, Theorem 1 implies that $\log \mathbb{E}[M^\alpha] \leq (1 - \alpha) \log \mathbb{E}[R^{\alpha/(\alpha-1)}]$. □

Proof of part (b) We state the proof assuming that $\alpha > 1$, as the proof for the case that $\alpha < 0$ is analogous. As a first observation, note that since the observable set is given $\mathcal{O} = \{(0, r) : r \in \mathbb{R}\} \cup \{(1, 1)\}$, it is immediate that $(\alpha, 0) \in \underline{\mathcal{O}}$. Therefore, Theorem 2 implies that $\log \mathbb{E}[M^\alpha] \geq \underline{\mathrm{J}}_{\mathcal{O}}(\alpha, 0)$, where

$$\begin{aligned} \underline{\mathrm{J}}_{\mathcal{O}}(\alpha,0) = \sup_{\substack{\lambda\geq 1,\pi,G \\ (\hat{m},\hat{r})\in\mathcal{O}}} \quad & (1-\lambda)\pi \log \mathbb{E}[MR] + (1-\lambda)(1-\pi)\int_{-\infty}^{\infty} \log \mathbb{E}[R^{\tilde{r}}]\,\mathrm{d}G(\tilde{r}) + \lambda \log \mathbb{E}[M^{\hat{m}} R^{\hat{r}}] \\ \text{s.t.} \quad & (1-\lambda)\pi + \lambda \hat{m} = \alpha \\ & (1-\lambda)\pi + (1-\lambda)(1-\pi)\int_{-\infty}^{\infty} \tilde{r}\,\mathrm{d}G(\tilde{r}) + \lambda\hat{r} \; = 0, \end{aligned}$$

π is a number between 0 and 1, and G is a probability distribution over the real line. To solve the above problem, note that the constraint that $(\hat{m}, \hat{r}) \in \mathcal{O}$ means that $\hat{m} \in \{0, 1\}$. But $\hat{m} = 0$ is inconsistent with the first constraint the fact that $\alpha, \lambda > 1$. Consequently, $\hat{m} = 1$, which means that $\hat{r} = 1$. Hence, we can rewrite the above problem as follows:

$$\begin{aligned} \underline{\mathrm{J}}_{\mathcal{O}}(\alpha,0) = \sup_{G} \quad & (1-\alpha)\int_{-\infty}^{\infty} \log \mathbb{E}[R^{\tilde{r}}]\,\mathrm{d}G(\tilde{r}) \\ \text{s.t.} \quad & \int_{-\infty}^{\infty} \tilde{r}\,\mathrm{d}G(\tilde{r}) = \alpha/(\alpha-1), \end{aligned}$$

where we are using the fact that $\mathbb{E}[MR] = 1$. Now the observation that $\log \mathbb{E}[R^{\tilde{r}}]$ is convex in $\tilde{r}$ implies that the distribution G that maximizes the objective function puts a mass of one on $\tilde{r} = \alpha/(\alpha - 1)$. Consequently, $\log \mathbb{E}[M^\alpha] \geq \underline{\mathrm{J}}_{\mathcal{O}}(\alpha, 0) = (1 - \alpha) \log \mathbb{E}[R^{\alpha/(\alpha-1)}]$. □

Proof of Proposition 4

Proof of part (a) Let $\mathbf{R} = (R_1, R_2)$ and $\mathbf{M} = (M_1, M_2)$, where R_i denotes an arbitrary return that can be traded by investors in country i and M_i is the SDF that prices R_i from the perspective of investors in country i. Recall from Example 2 that the observable set in this setting is given by (2). It is immediate that $(\alpha, \beta, 0, 0) \in \overline{\mathcal{O}}$ if (α, β) is in the unit simplex. We can therefore invoke Theorem 1, which implies that $\log \mathbb{E}[M_1^\alpha M_2^\beta] \leq \overline{\mathrm{J}}_{\mathcal{O}}(\alpha, \beta, 0, 0)$, where

$$\begin{aligned} \overline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0) = \inf_{G,\pi_1,\pi_2} \quad & (1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \log \mathbb{E}[R_1^{\tilde{r}_1} R_2^{\tilde{r}_2}]\,\mathrm{d}G(\tilde{r}_1,\tilde{r}_2) \\ \text{s.t.} \quad & \pi_1 \begin{bmatrix}1\\0\end{bmatrix} + \pi_2 \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix} \\ & \pi_1 \begin{bmatrix}1\\0\end{bmatrix} + \pi_2 \begin{bmatrix}0\\1\end{bmatrix} + (1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \begin{bmatrix}\tilde{r}_1\\ \tilde{r}_2\end{bmatrix}\,\mathrm{d}G(\tilde{r}_1,\tilde{r}_2) = \begin{bmatrix}0\\0\end{bmatrix}, \end{aligned}$$

π_1 and π_2 are non-negative numbers satisfying $\pi_1 + \pi_2 \leq 1$ and G is a probability distribution with support over $\mathbb{R}^2$. From the first constraint it is immediate that $\pi = \alpha$ and $\pi_2 = \beta$. As a result, the problem simplifies as follows:

$$\overline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0) = \inf_{G} \quad (1-\alpha-\beta)\int_{\mathbb{R}^2} \log \mathbb{E}[R_1^{\tilde{r}_1} R_2^{\tilde{r}_2}]\, \mathrm{d}G(\tilde{r}_1,\tilde{r}_2)$$

$$\text{s.t.} \quad (\alpha+\beta-1)\int_{\mathbb{R}^2} \begin{bmatrix}\tilde{r}_1\\ \tilde{r}_2\end{bmatrix} \mathrm{d}G(\tilde{r}_1,\tilde{r}_2) = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}.$$

Recall from Lemma B.1 that $\log \mathbb{E}[R_1^{\tilde{r}_1} R_2^{\tilde{r}_2}]$ is jointly convex in $(\tilde{r}_1,\tilde{r}_2)$. This, coupled with the fact that $\alpha+\beta<1$, implies that the distribution G that minimizes the objective function puts a mass of one on the point $(\tilde{r}_1,\tilde{r}_2) = (\frac{\alpha}{\alpha+\beta-1}, \frac{\beta}{\alpha+\beta-1})$. Consequently, $\overline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0) = (1-\alpha-\beta)\log \mathbb{E}[(R_1^{\alpha}R_2^{\beta})^{1/(\alpha+\beta-1)}]$, thus establishing the inequality in (18). □

Proof of part (b) We state the proof for case (ii), where $\alpha+\beta>1$ and $\min\{\alpha,\beta\}<0$. The proof for case (i) is analogous. Also, due to the symmetry in the problem, we can assume, without loss of generality, that $\alpha>1>0>\beta$.

Given that the observable set is given (2), it is immediate that $(\alpha,\beta,0,0)\in\underline{\mathcal{O}}$. Therefore, Theorem 2 implies that $\log \mathbb{E}[M_1^{\alpha}M_2^{\beta}] \geq \underline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0)$, where

$$\underline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0) = \sup_{\substack{\lambda\geq 1,\pi_1,\pi_2\\ G,(\hat{m},\hat{r})\in\mathcal{O}}} \quad (1-\lambda)(1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \log \mathbb{E}[R_1^{\tilde{r}_1} R_2^{\tilde{r}_2}]\, \mathrm{d}G(\tilde{r}_1,\tilde{r}_2) + \lambda \log \mathbb{E}[M_1^{\hat{m}_1} M_2^{\hat{m}_2} R_1^{\hat{r}_1} R_2^{\hat{r}_2}]$$

$$\text{s.t.} \quad (1-\lambda)\pi_1 \begin{bmatrix}1\\0\end{bmatrix} + (1-\lambda)\pi_2\begin{bmatrix}0\\1\end{bmatrix} + \lambda\begin{bmatrix}\hat{m}_1\\ \hat{m}_2\end{bmatrix} = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}$$

$$(1-\lambda)\pi_1 \begin{bmatrix}1\\0\end{bmatrix} + (1-\lambda)\pi_2\begin{bmatrix}0\\1\end{bmatrix} + (1-\lambda)(1-\pi_1-\pi_2)\int_{\mathbb{R}^2}\begin{bmatrix}\tilde{r}_1\\ \tilde{r}_2\end{bmatrix}\mathrm{d}G(\tilde{r}_1,\tilde{r}_2) + \lambda\begin{bmatrix}\hat{r}_1\\ \hat{r}_2\end{bmatrix} = 0,$$

π_1 and π_2 are non-negative numbers satisfying $\pi_1+\pi_2\leq 1$, and G is a probability distribution with support over $\mathbb{R}^2$. The constraint that $(\hat{m},\hat{r})\in\mathcal{O}$ means that $\hat{m}_1,\hat{m}_2\in\{0,1\}$. But $\hat{m}_1=0$ is inconsistent with the first constraint the fact that $\alpha,\lambda>1$. Therefore, it must be the case that $\hat{m}_1=1$, which in turn implies that $(\hat{m}_1,\hat{m}_2,\hat{r}_1,\hat{r}_2)=(1,0,1,0)$. We can therefore simplify the above problem as follows:

$$\underline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0) = \sup_{G} \quad (1-\alpha-\beta)\int_{\mathbb{R}^2} \log \mathbb{E}[R_1^{\tilde{r}_1} R_2^{\tilde{r}_2}]\, \mathrm{d}G(\tilde{r}_1,\tilde{r}_2)$$

$$\text{s.t.} \quad (\alpha+\beta-1)\int_{\mathbb{R}^2} \begin{bmatrix}\tilde{r}_1\\ \tilde{r}_2\end{bmatrix} \mathrm{d}G(\tilde{r}_1,\tilde{r}_2) = \begin{bmatrix}\alpha\\ \beta\end{bmatrix},$$

where we are using the fact that $\mathbb{E}[M_1R_1]=1$. Next, recall from Lemma B.1 that $\log \mathbb{E}[R_1^{\tilde{r}_1} R_2^{\tilde{r}_2}]$ is jointly convex in $(\tilde{r}_1,\tilde{r}_2)$. This, coupled with the assumption that $\alpha+\beta>1$, implies that the probability distribution G that maximizes the objective function puts the entire probability mass on the point $(\tilde{r}_1,\tilde{r}_2) = (\frac{\alpha}{\alpha+\beta-1}, \frac{\beta}{\alpha+\beta-1})$. Consequently, $\underline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0,0) = (1-\alpha-\beta)\log \mathbb{E}[(R_1^{\alpha}R_2^{\beta})^{1/(\alpha+\beta-1)}]$, thus establishing the inequality in (19). □

Proof of Proposition 5

Proof of part (a) Let $\mathbf{M} = (M^P_{t,t+\tau}, M^T_{t,t+\tau})$ and $\mathbf{R} = R_{t,t+\tau}$, where $M^P_{t,t+\tau}$ and $M^T_{t,t+\tau}$ are the permanent and transient components of the SDF, respectively, and $R_{t,t+\tau}$ is the return on an arbitrary traded asset over the period from t to $t+\tau$. Recall from Example 3 that the observable set in this setting is given by (3). It is immediate to verify that $(\alpha, \beta, 0) \in \overline{\mathcal{O}}$ for any $\alpha \in (0, 1)$. We can therefore invoke Theorem 1, which implies that $\log \mathbb{E}_t[(M^P_{t,t+\tau})^\alpha (M^T_{t,t+\tau})^\beta] \leq \overline{\mathrm{J}}_{\mathcal{O}}(\alpha, \beta, 0)$, where

$$\begin{aligned}
\overline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0) = \inf_{G,\pi_1,\pi_2} \quad & (1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{-\tilde{m}}(R_{t,t+\tau})^{\tilde{r}}]\, \mathrm{d}G(\tilde{m},\tilde{r}) \\
\text{s.t.} \quad & \pi_1 + \pi_2 = \alpha \\
& \pi_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \begin{bmatrix} \tilde{m} \\ \tilde{r} \end{bmatrix} \mathrm{d}G(\tilde{m},\tilde{r}) = \begin{bmatrix} \beta \\ 0 \end{bmatrix},
\end{aligned}$$

π_1 and π_2 are non-negative numbers satisfying $\pi_1+\pi_2 \leq 1$ and G is a probability distribution with support over $\mathbb{R}^2$. Using the first constraint, we can simplify the above problem as follows:

$$\begin{aligned}
\overline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0) = \inf_{G,\pi_1} \quad & (1-\alpha)\int_{\mathbb{R}^2} \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{-\tilde{m}}(R_{t,t+\tau})^{\tilde{r}}]\, \mathrm{d}G(\tilde{m},\tilde{r}) \\
\text{s.t.} \quad & \int_{\mathbb{R}^2} \begin{bmatrix} \tilde{m} \\ \tilde{r} \end{bmatrix} \mathrm{d}G(\tilde{m},\tilde{r}) = \frac{1}{\alpha-1} \begin{bmatrix} \pi_1 - \beta \\ \pi_1 \end{bmatrix}.
\end{aligned}$$

Recall from Lemma B.1 that the CGF is jointly convex in its arguments. This, coupled with the fact that $\alpha \in (0,1)$, implies that the distribution G that minimizes the objective function puts a mass of one on the point $(\tilde{m},\tilde{r}) = (\frac{\pi_1-\beta}{\alpha-1}, \frac{\pi_1}{\alpha-1})$. Consequently, after a change of variable, $\delta = \pi/\alpha$, we get

$$\log \mathbb{E}_t[(M^P_{t,t+\tau})^\alpha (M^T_{t,t+\tau})^\beta] \leq \inf_{\delta\in[0,1]} \quad (1-\alpha)\log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{\frac{\beta}{\alpha-1}}(R_{t,t+\tau}/R^{(\infty)}_{t,t+\tau})^{\frac{\alpha\delta}{\alpha-1}}] \tag{B.3}$$

for any arbitrary $R_{t,t+\tau}$ such that $\mathbb{E}_t[M_{t,t+\tau}R_{t,t+\tau}] = 1$. Setting $\delta = 1$ in (B.3) thus implies (24). □

Proof of part (b) We state the proof assuming that $\alpha > 1$, as the proof for the case that $\alpha < 0$ is analogous. Given that the observable set is given (3), it is immediate that $(\alpha,\beta,0) \in \underline{\mathcal{O}}$. Therefore, Theorem 2 implies that $\log \mathbb{E}_t[(M^P_{t,t+\tau})^\alpha (M^T_{t,t+\tau})^\beta] \geq \underline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0)$ for any $\alpha > 1$, where

$$\begin{aligned}
\underline{\mathrm{J}}_{\mathcal{O}}(\alpha,\beta,0) = \sup_{\substack{\lambda\geq 1,\pi_1,\pi_2 \\ G,(\hat{m},\hat{r})\in\mathcal{O}}} \quad & (1-\lambda)(1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{-\tilde{m}_2}(R_{t,t+\tau})^{\tilde{r}}]\mathrm{d}G(\tilde{m}_2,\tilde{r}) \\
& + \lambda \log \mathbb{E}_t[(M^P_{t,t+\tau})^{\hat{m}_1}(M^T_{t,t+\tau})^{\hat{m}_2} R^{\hat{r}}_{t,t+\tau}] \\
\text{s.t.} \quad & (1-\lambda)(\pi_1+\pi_2) + \lambda\hat{m}_1 = \alpha \\
& (1-\lambda)\pi_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (1-\lambda)(1-\pi_1-\pi_2)\int_{\mathbb{R}^2} \begin{bmatrix} \tilde{m}_2 \\ \tilde{r} \end{bmatrix} \mathrm{d}G(\tilde{m}_2,\tilde{r}) + \lambda \begin{bmatrix} \hat{m}_2 \\ \hat{r} \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix},
\end{aligned}$$

π_1 and π_2 are non-negative numbers satisfying $\pi_1 + \pi_2 \leq 1$, and G is a probability distribution with support over $\mathbb{R}^2$. The constraint that $(\hat{m}, \hat{r}) \in \mathcal{O}$ means that $\hat{m}_1 \in \{0, 1\}$. But $\hat{m}_1 = 0$ is inconsistent with the first constraint the fact that $\alpha, \lambda > 1$. Therefore, it must be the case that $\hat{m}_1 = 1$, which in turn implies that $\hat{m}_2 = \hat{r} \in \{0, 1\}$. As a result, we can simplify the above problem as follows:

$$\underline{\mathrm{J}}_{\mathcal{O}}(\alpha, \beta, 0) = \sup_{\substack{\lambda \geq 1, \pi_1 \in [0,1] \\ G, \hat{r} \in \{0,1\}}} (1-\alpha) \int_{\mathbb{R}^2} \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{-\tilde{m}_2} (R_{t,t+\tau})^{\tilde{r}}] \mathrm{d}G(\tilde{m}_2, \tilde{r})$$

$$\text{s.t.} \quad (\alpha - 1) \int_{\mathbb{R}^2} \begin{bmatrix} \tilde{m}_2 \\ \tilde{r} \end{bmatrix} \mathrm{d}G(\tilde{m}_2, \tilde{r}) = \begin{bmatrix} \lambda\hat{r} - \beta + (1-\lambda)\pi_1 \\ \lambda\hat{r} + (1-\lambda)\pi_1 \end{bmatrix}.$$

Once again, the convexity of the CGF, established in Lemma B.1, guarantees that the distribution G that maximizes the objective function assigns all the probability mass on a single point. This, coupled with Theorem 2, therefore implies that

$$\log \mathbb{E}_t[(M^P_{t,t+\tau})^{\alpha} (M^T_{t,t+\tau})^{\beta}] \geq \sup_{\substack{\hat{r} \in \{0,1\}, \lambda \geq 1 \\ \pi_1 \in [0,1]}} (1-\alpha) \log \mathbb{E}_t[(R^{(\infty)}_{t,t+\tau})^{\frac{\beta - \lambda\hat{r} - (1-\lambda)\pi_1}{\alpha - 1}} (R_{t,t+\tau})^{\frac{\lambda\hat{r} + (1-\lambda)\pi_1}{\alpha - 1}}]$$

Setting $\lambda = \alpha > 1$, $\pi_1 = 0$, and $\hat{r} = 1$ on the right-hand side of the above inequality then establishes (25). □

References

Almeida, Caio and René Garcia (2012), "Assessing misspecified asset pricing models with empirical likelihood estimators." *Journal of Econometrics*, 170, 519–537.

Almeida, Caio and René Garcia (2017), "Economic implications of nonlinear pricing kernels." *Management Science*, 63(10), 3361–3380.

Alvarez, Fernando and Urban J. Jermann (2005), "Using asset prices to measure the persistence of the marginal utility of wealth." *Econometrica*, 73(6), 1977–2016.

Backus, David, Mikhael Chernov, and Stanley E. Zin (2014), "Sources of entropy in representative agent models." *Journal of Finance*, 69(1), 51–99.

Bakshi, Gurdip and Fousseni Chabi-Yo (2012), "Variance bounds on the permanent and transitory components of stochastic discount factors." *Journal of Financial Economics*, 105(1), 191–208.

Bansal, Ravi and Bruce N. Lehmann (1997), "Growth-optimal portfolio restrictions on asset pricing models." *Macroeconomic Dynamics*, 108(1), 333–354.

Barro, Robert J. and Tao Jin (2021), "Rare events and long-run risks." *Review of Economic Dynamics*, 39, 1–25.

Brandt, Michael W., John H. Cochrane, and Pedro Santa-Clara (2006), "International risk sharing is better than you think, or exchange rates are too smooth." *Journal of Monetary Economics*, 53(4), 671–698.

Chernov, Mikhail, Valentin Haddad, and Oleg Itskhoki (2023), "What do financial markets say about the exchange rate?" *Working Paper, UCLA.*

Colacito, Riccardo and Mariano M. Croce (2011), "Risks for the long run and the real exchange rate." *Journal of Political Economy*, 119(1), 153–181.

Colacito, Riccardo and Mariano M. Croce (2013), "International asset pricing with recursive preferences." *Journal of Finance*, 68(6), 2651–2686.

Colacito, Riccardo, Mariano M. Croce, Federico Gavazzoni, and Robert Ready (2018), "Currency risk factors in a recursive multicountry economy." *Journal of Finance*, 73(6), 2719–2756.

Epstein, Larry and Stanley E. Zin (1989), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica*, 57(4), 937–969.

Farhi, Emmanuel and Xavier Gabaix (2016), "Rare disasters and exchange rates." *Quarterly Journal of Economics*, 131(1), 1–52.

Gosh, Anisha, Christian Julliard, and Alex P Taylor (2017), "What is the consumption-capm missing? an information-theoretic framework for the analysis of asset pricing models." *The Review of Financial Studies*, 30, 442–504.

Hansen, Lars Peter and Ravi Jagannathan (1991), “Implications of security market data for models of dynamic economies.” *Journal of Political Economy*, 99(2), 225–262.

Hansen, Lars Peter and Ravi Jagannathan (1997), “Asssessing specification errors in stochastic discount factor models.” *Journal of Finance*, 52(2), 552–590.

Hansen, Lars Peter and José A. Scheinkman (2009), “Long-term risk: An operator approach.” *Econometrica*, 77(1), 177–234.

Lustig, Hanno and Adrien Verdelhan (2019), “Does incomplete spanning in international financial markets help to explain exchange rates?” *American Economic Review*, 109(6), 2208–2244.

Peters, Hans J. M. and Peter P. Wakker (1986), “Convex functions on non-convex domains.” *Economics Letters*, 22(2–3), 251–255.

Sandulescu, Mirela, Fabio Trojani, and Andrea Vedolin (2021), “Model-free international stochastic discount factors.” *Journal of Finance*, 76(2), 935–976.

Snow, Karl N. (1991), “Diagnosing asset pricing models using the distribution of asset returns.” *Journal of Finance*, 46(3), 955–983.

Stutzer, Michael (1995), “A bayesian approach to diagnosis of asset pricing models.” *Journal of Econometrics*, 68(2), 367–397.