

# Arbitrage Free Dispersion\*

Piotr Orłowski

András Sali

Fabio Trojani

Swiss Finance Institute

Swiss Finance Institute

Swiss Finance Institute

University of Lugano<sup>†</sup>

University of Lugano<sup>‡</sup>

University of Geneva<sup>§</sup>

February 2, 2016

---

\*We thank for insightful comments Caio Almeida (discussant), Torben Andersen, Tomas Björk, Magnus Dahlquist, Semyon Malamud, Olivier Scaillet, Paul Schneider, Paolo Sodini, Viktor Todorov and the participants of the 2015 SFI research days, the Second International Workshop in Financial Econometrics in Salvador, Brazil, and the seminars series at Bocconi University, Northwestern University, Stockholm School of Economics and University of Geneva. We are grateful to the Swiss Finance Institute, Project "Term Structures and Cross Sections of Asset Risk Premia", for financial support.

<sup>†</sup>Piotr Orłowski, University of Lugano, Via Buffi 13, 6900 Lugano, Switzerland. Email: [Piotr.Orlowski@usi.ch](mailto:Piotr.Orlowski@usi.ch)

<sup>‡</sup>András Sali, University of Lugano, Via Buffi 13, 6900 Lugano, Switzerland. Email: [Andras.Sali@usi.ch](mailto:Andras.Sali@usi.ch)

<sup>§</sup>Fabio Trojani, SFI Senior Chair, University of Geneva, Bd du Pont d'Arve 40, 1211 Geneva 4, Switzerland. Email: [Fabio.Trojani@alphacruncher.com](mailto:Fabio.Trojani@alphacruncher.com)

# Arbitrage Free Dispersion

First version: December 2014

This version: February 2, 2016

## Abstract

We develop a theory of arbitrage-free dispersion (AFD) that characterizes the testable restrictions of asset pricing models. AFD measures Jensen's gap in the cumulant generating function of pricing kernels and returns. It implies a wide family of model-free dispersion constraints, which extend dispersion and co-dispersion bounds in the literature and are applicable with a unifying approach in multivariate and multiperiod settings. Empirically, the dispersion of stationary and martingale pricing kernel components in the benchmark long-run risk model yields a counterfactual dependence of short- vs. long-maturity bond returns and is insufficient for pricing optimal portfolios of market equity and short-term bonds.

**Keywords:**— Arbitrage-Free Dispersion, Cumulant Generating Function, Convexity, Convex Inequalities, Jensen's Gap, Pricing Kernel Bounds, Entropy, Long-Run Risk Models, Tests of Asset Pricing Models

*Knowledge is never used up. It increases by diffusion and grows by dispersion.*

Daniel J. Boorstin

# 1 Introduction

Arbitrage-free markets are characterized by tight relations between unobservable pricing kernels, observable asset payoffs and their arbitrage-free price. These relations constrain the joint stochastic properties of pricing kernels and asset returns along several dimensions, which are informative about the market price of the relevant uncertain economic states and the set of arbitrage-free prices of untraded assets in incomplete markets. In this paper, we introduce a new systematic approach for testing asset pricing models, which parsimoniously aggregates the observable asset pricing information from multivariate arbitrage-free markets and comprehensively characterize tight arbitrage-free constraints on the joint distribution of potentially multivariate pricing kernels and asset returns. We complement and extend the existing literature, by introducing a new systematic approach to test multivariate pricing kernel specifications on multiple investment horizons.

We first summarize the observable arbitrage-free information by a well-defined subset of observed values on the joint cumulant generating function (CGF) of pricing kernels and asset returns. To illustrate, in the simplified setting of a single pricing kernel  $M$  and a traded return  $R$ , the joint CGF is defined by

$$\mathcal{K}_{MR}(m, r) := \log E[M^m R^r] ; (m, r) \in \mathbb{R}^2 , \quad (1)$$

and the asset pricing constraint for the traded return is  $\mathcal{K}_{MR}(1, 1) = \log E[MR] = 0$ . When the marginal distribution of returns is also observable, then  $\mathcal{K}_{MR}(0, \cdot)$  is observed and the observable information on  $\mathcal{K}_{MR}$  is summarized by the values of the CGF on the observable set  $\mathcal{O}_{\mathcal{K}_{MR}} := \{(m, r) \in \text{dom}(\mathcal{K}_{MR}) : m = 0 \text{ or } (m, r) = (1, 1)\}$ .

Second, we derive a broad class of multivariate arbitrage-free inequalities between observable and unobservable regions of the joint CDF, which are a consequence of the convexity of cumulant generating functions. In this way, the observable information on traded asset returns restricts the class of joint

distributions for pricing kernels and returns that are consistent with arbitrage-free markets.<sup>1</sup> We also show that while convex arbitrage-free inequalities hold and are computable for general multivariate settings, their application to lower dimensional settings provides a direct derivation of a large class of pricing kernel bounds in the literature. For instance, convexity of  $\mathcal{K}_{MR}$  in equation (1) implies for any  $\alpha \in (0, 1)$  the bound:

$$\begin{aligned} \log E[M^\alpha] &= \mathcal{K}_{MR}(\alpha, 0) \leq \alpha \mathcal{K}_{MR}(1, 1) + (1 - \alpha) \mathcal{K}_{MR}(0, -\alpha/(1 - \alpha)) \\ &= \log E[R^{-\alpha/(1-\alpha)}]^{1-\alpha} , \end{aligned} \tag{2}$$

which naturally implies  $\mathcal{E}(M) := -\log E[M/E[M]] \geq \log(E[R]E[M])$ , i.e., the entropy pricing kernel bound in [Bansal and Lehmann \(1997\)](#), [Alvarez and Jermann \(2005\)](#), [Liu \(2013\)](#) and [Backus, Chernov, and Zin \(2014\)](#), among others.<sup>2</sup>

Third, we show that convex arbitrage-free inequalities are interpretable as arbitrage-free constraints on the multivariate dispersion of pricing kernels and returns, where dispersion is defined using a family of so-called Jensen's gaps generated by the multivariate CGF. In this sense, violations of certain convex inequalities or pricing kernel bounds are equivalent to an insufficient arbitrage-free dispersion in some regions of the multivariate state space. To illustrate, given CGF (1) and a non degenerate prior distribution  $\pi$ , Jensen's inequality induces the Jensen's gap:

$$\mathcal{J}_\pi(M, R) := E_\pi[\mathcal{K}_{MR}(m, r)] - \mathcal{K}_{MR}(E_\pi[(m, r)]) \geq 0 , \tag{3}$$

with equality if and only if  $(M, R) = E[(M, R)]$ . Jensen's gap  $\mathcal{J}_\pi(M, R)$  is a measure of multivariate dispersion and directly constrains the arbitrage-free CGF whenever  $E_\pi[\mathcal{K}_{MR}(m, r)]$  or  $\mathcal{K}_{MR}(E_\pi[(m, r)])$

<sup>1</sup>For instance, it constrains the marginal distribution of pricing kernels, by bounding the range of admissible nonlinear moments. Similarly, it constrains the range of admissible prices for untraded nonlinear payoffs, such as the prices of payoffs with nonlinear exposure in the underlying return.

<sup>2</sup>The bound follows by continuity, taking limits as  $\alpha \rightarrow 1$  in the convex arbitrage-free inequality (2).

is observable.<sup>3</sup> In this way, we identify unique model-free constraints for the joint distribution of pricing kernels and returns that are directly interpretable as specification constraints for their joint dispersion properties.

Fourth, we parsimoniously incorporate arbitrage-free dispersion constraints into lower and upper bounds on the CGF of pricing kernels and returns. Precisely, we introduce the upper (lower) arbitrage-free CGF  $\mathcal{K}_{MR}^U$  ( $\mathcal{K}_{MR}^L$ ), which is defined as the smallest (largest) observable upper (lower) bound implied by convex arbitrage-free inequalities on any arbitrage-free CGF, and we naturally derive upper (lower) arbitrage-free CGFs from basic arbitrage-free dispersion properties. To illustrate, given an arbitrage-free CGF evaluated at point  $(m_\star, r_\star)$  and a prior with support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  such that  $(m_\star, r_\star) = E_{\pi_\star}[(m, r)]$ , inequality (3) gives:

$$E_{\pi_\star}[\mathcal{K}_{MR}(m, r)] \geq \mathcal{K}_{MR}(m_\star, r_\star) . \quad (4)$$

As the left side of this inequality is observable, we obtain the following upper arbitrage-free CGF  $\mathcal{K}_{MR}^U$  evaluated in  $(m_\star, r_\star)$ :<sup>4</sup>

$$\mathcal{K}_{MR}^U(m_\star, r_\star) := \inf_{\pi_\star} E_{\pi_\star}[\mathcal{K}_{MR}(m, r)] \geq \mathcal{K}_{MR}(m_\star, r_\star) . \quad (5)$$

Upper and lower arbitrage-free CGFs are observable in a model-free way and naturally constrain the joint distribution of pricing kernels and returns. We show in a number of important cases that these constraints are the tightest attainable, given an observed arbitrage-free price system. Thus, upper and lower CGFs meaningfully synthesize the available model-free information on the distribution of pricing kernels and returns that can be inferred from statistical return observation and existing arbitrage-free relations.

We make use of our theory of AFD to specify a coherent diagnostics approach for systematically

<sup>3</sup>For instance, the inequality  $E_\pi[\mathcal{K}_{MR}(m, r)] \geq \mathcal{K}_{MR}(E_\pi[(m, r)])$  implied by a Bernoulli prior with  $\pi(1, 1) = \alpha \in (0, 1)$  and  $\pi(0, -\alpha/(1-\alpha)) = 1-\alpha$  directly generates pricing kernel bound (2). We discuss in detail below the properties of  $\mathcal{J}_\pi(M, R)$  as a measure of multivariate dispersion.

<sup>4</sup>The infimum on the left hand side of this inequality is over priors with support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  such that  $(m_\star, r_\star) = E_{\pi_\star}[(m, r)]$ .

testing the multivariate dispersion properties of asset pricing models. We apply this diagnostics approach to test the multivariate dispersion properties of transient and martingale pricing kernel components in the benchmark [Bansal and Yaron \(2004\)](#)-model of Long Run Risks (LRR), where the transient component is approximated empirically from the observed return of a long-maturity bond. While the marginal dispersion properties of pricing kernel components are broadly consistent with the data, multivariate dispersion tests provide sharper evidence of a model failure, in terms of (i) a counterfactual dependence of long vs. short maturity bond returns and (ii) an insufficient dispersion for pricing the optimal returns of a class of power utility investors invested in short-term bonds and aggregate market equity.

The remainder of the paper proceeds as follows. In [Section 2](#), we introduce the joint CGF of pricing kernels and asset returns, showing that a number of interesting model settings – including multiple SDF components, countries and time horizons – can naturally be incorporated in this framework. [Section 3](#) exploits the convexity of the joint CGF to specify general multivariate dispersion measures in a variety of relevant asset pricing contexts. [Section 4](#) shows how they induce natural model-free bounds on the arbitrage-free CGF and the dispersion of pricing kernels and returns, motivating a systematic diagnostics approach for testing the multivariate dispersion implications of asset pricing models. [Section 5](#) addresses in more detail concrete asset pricing contexts and explicitly computable dispersion constraints in these settings. It demonstrates the sharpness of pricing kernel dispersion bounds resulting from our approach in a number of important cases and delivers testable multivariate dispersion constraints for the benchmark Long Run Risk (LRR) model in [Bansal and Yaron \(2004\)](#). [Section 6](#) characterizes and tests empirically the multivariate arbitrage-free dispersion properties of recent empirical parameterizations of LRR models. [Section 7](#) concludes.

## 2 Arbitrage Free Cumulant Generating Function

Under weak assumptions, arbitrage-free markets imply tight relations between unobservable pricing kernels and the arbitrage-free price of marketed asset payoffs. Such constraints are conveniently

summarized by particular restrictions on the joint CGF of pricing kernels and payoffs.

## 2.1 Definition

We introduce the joint CGF of pricing kernel and asset returns using a general multivariate structure, in which uncertainty is generated by a vector of returns priced by another vector of pricing kernel components. Examples of pricing kernel components are vectors of domestic and foreign pricing kernels in international arbitrage-free markets, the vector of single-period pricing kernels that price returns over different horizons or the distinct frequency components of a pricing kernel, as, e.g., in [Alvarez and Jermann \(2005\)](#).

**Definition 1** (CGF of Pricing Kernel and Returns). *Given a set of strictly positive pricing kernel components  $M := (M_1, \dots, M_{d_1})$  and a set of positive marketed gross returns  $R := (R_1, \dots, R_{d_2})$ , the arbitrage-free cumulant generating function of  $(M, R)$  is the function  $\mathcal{K}_{MR} : \mathbb{R}^{d_1+d_2} \rightarrow \overline{\mathbb{R}}$ , defined for any  $m := (m_1, \dots, m_{d_1})$  and  $r := (r_1, \dots, r_{d_2})$  by*

$$\mathcal{K}_{MR}(m, r) := \log E[M^r R^r] := \log E \left[ \prod_{i=1}^{d_1} M_i^{m_i} \prod_{j=1}^{d_2} R_j^{r_j} \right]. \quad (6)$$

The marginal pricing kernel (returns) CGF is defined by  $\mathcal{K}_M(\cdot) := \mathcal{K}_{MR}(\cdot, 0)$  ( $\mathcal{K}_R(\cdot) := \mathcal{K}_{MR}(0, \cdot)$ ).

The joint CGF uniquely identifies the joint distribution of pricing kernel components and returns.<sup>5</sup>

Therefore, it also identifies the arbitrage-free pricing system implied by a (parametric or nonparametric) specification of an asset pricing model. Conversely, the empirically observable arbitrage-free pricing restrictions are naturally summarized by the values of an arbitrage-free CGF on a corresponding subset of points  $(m, r) \in \mathbb{R}^{d_1+d_2}$ . Such restrictions generate natural specification constraints for asset pricing models.

**Definition 2** (Observable Arbitrage-Free CGF). *An arbitrage-free CGF is observable in  $(m, r) \in \mathbb{R}^{d_1+d_2}$ , whenever  $\mathcal{K}_{MR}(m, r)$  is known through the statistical observation of asset returns or the prices*

---

<sup>5</sup>Throughout the paper we assume that the joint cumulant generating function is finite in a non-degenerate open domain containing the origin.

of extant payoffs. We denote by  $\mathcal{O}_{\mathcal{K}_{MR}} := \{(m, r) \in \mathbb{R}^{d_1+d_2} : \mathcal{K}_{MR}(m, r) \text{ is observed}\}$  the set of observable points of an arbitrage-free CGF and call  $\mathcal{K}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}}$  the observable arbitrage-free CGF.

## 2.2 Observability and Marginal CGF

Whenever we can assume statistical observability of the return distribution, the marginal return CGF is observable, i.e.,  $(0, r) \in \mathcal{O}_{\mathcal{K}_{MR}}$  for any  $r \in \mathbb{R}^{d_2}$ . In general, the marginal CGF of a pricing kernel is never empirically completely observable. Whenever the price  $B$  of a risk-free zero bond is observable, then  $\mathcal{K}_M(1) = \log E[M] = \log B$  and  $(1, 0) \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Additional points on the marginal CGF may be directly observable due to normalizing conditions. For instance, when a pricing kernel is decomposed into the product of a transient and a permanent martingale component  $M^T$  and  $M^P$ , the martingale condition  $\mathcal{K}_{M^T M^P}(0, 1) = \log E[M^P] = 0$  yields  $(0, 1, 0) \in \mathcal{O}_{\mathcal{K}_{M^T M^P R}}$ ; see again [Alvarez and Jermann \(2005\)](#), among others.

## 2.3 Univariate Return and Pricing Kernel ( $d_1 = d_2 = 1$ )

Given the statistical observability of a univariate return distribution, the additional set of observable points depends on the structure of observable arbitrage-free constraints. Whenever a risk-free bond is priced, then obviously  $(1, 0) \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Whenever a risky return is also priced, then  $\mathcal{K}_{MR}(1, 1) = \log E[MR] = 0$  and  $(1, 1) \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Figure 1(a) plots the observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$  generated by statistical return information and such arbitrage-free constraints. This set is not convex. It contains the vertical line with abscissa in  $m = 0$ , i.e., the domain of the observable moment generating function of returns, and only two additional points of the vertical line with abscissa in  $m = 1$ , which is the region collecting information about the risk-neutral distribution of returns. The sparsity of points in this region reflects the large degree of market incompleteness of this setting, in which no nonlinear payoff on underlying return  $R$  is traded. The other extreme is a market extended to complete option trading, allowing to trade any smooth nonlinear function of  $R$  using portfolios of out-of-the-money options. In such a case,



we have for any  $p \in \mathbb{R}$ , following Carr and Madan (1998) and Schneider and Trojani (2013):

$$\mathcal{K}_{MR}(1, p) = \log E[MR^p] = \log[p + (1 - p)B + p(p - 1) \int_0^\infty K^{p-2} O(R, K) dK] ,$$

where  $O(R, K)$  is the arbitrage-free price of an out-of-the-money option on  $R$  with strike price  $K > 0$ .

Figure 1(b) illustrates the enrichment of the (non convex) set of empirically observable CGF points generated by complete option markets, which now includes the vertical line with abscissa in  $m = 1$ .

## 2.4 Transient and Persistent Pricing Kernel Components ( $d_1 = 2$ and $d_2 = 1$ )

Whenever the long end of the yield curve is observable, the pricing kernel decomposition  $M^P M^T$  into transient and persistent components produces additional observable constraints. Following Alvarez and Jermann (2005),  $M^P$  does not affect the price of infinitely long maturity zero coupon bonds and  $R_\infty = 1/M^T$ , where  $R_\infty$  is the return of the infinitely long maturity zero coupon bond. Therefore,

$$\mathcal{K}_{MR}(m_T, 0, r) = \log E[(M^T)^{m_T} R^r] = \log E[R_\infty^{-m_T} R^r] , \quad (7)$$

where  $M := (M^T, M^P)$  and  $(m_T, 0, r) \in \mathcal{O}_{\mathcal{K}_{MR}}$  for each  $(m_T, r) \in \mathbb{R}^2$ . The pricing constraints for the short term zero bond and the risky return imply  $(1, 1, 0), (1, 1, 1) \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Non convex set  $\mathcal{O}_{\mathcal{K}_{MR}}$  is illustrated in Figure 2. The observable points resulting from arbitrage-free pricing relations are marked as violet circles ( $\circ$ ), while those that are observable statistically span the red  $(m_T, r)$  plane. Two such points, corresponding to the CGF values  $\mathcal{K}_{MR}(0, 0, 1) = \log E[R]$  and  $\mathcal{K}_{MR}(1, 0, 0) = \log E[R_\infty^{-1}]$ , are marked as violet squares ( $\square$ ).

## 2.5 Domestic and Foreign Pricing Kernels ( $d_1 = 2$ and $d_2 = 2$ )

In an international context, the pricing kernel components can be domestic and foreign pricing kernels  $M_d$  and  $M_f$ , pricing domestic and foreign risk-free bonds and risky returns. The empirically observable points follow as in the previous examples, so that points  $(1, 0, 0, 0), (0, 1, 0, 0), (1, 0, 1, 0)$  and  $(0, 1, 0, 1)$

are all elements of  $\mathcal{O}_{\mathcal{K}_{MR}}$ , where  $M := (M_d, M_f)$  and  $R := (R_d, R_f)$  is the vector of domestic and foreign risky returns. Whenever domestic and foreign option markets are complete, it also follows  $(1, 0, r_d, 0), (0, 1, 0, r_f) \in \mathcal{O}_{\mathcal{K}_{MR}}$  for any  $(r_d, r_f) \in \mathbb{R}^2$ . Additional observable constraints can emerge from the spot exchange rate market, as the exchange rate return  $R_e := (M_f/M_d) \cdot \mathcal{E}$  needs to satisfy the additional arbitrage-free conditions:

$$1 = E[M_d R_e] = E[M_f \mathcal{E}] ; 1 = E[M_f (1/R_e)] = E[M_d (1/\mathcal{E})] .$$

When domestic and foreign markets are complete,  $R_e = M_d/M_f$  and the joint CGF of pricing kernels and returns characterizes the observable arbitrage-free restrictions from domestic, foreign and spot exchange rate markets. Further CGF constraints arise when complete exchange rate option markets allow the trading of smooth functions of  $R_e$ . Indeed, in this case  $(1 - p, p, 0, 0) \in \mathcal{O}_{\mathcal{K}_{MR}}$  for every  $p \in \mathbb{R}$ , since

$$\mathcal{K}_{MR}(1 - p, p, 0, 0) = \log E[M_d R_e^p] = \log[p + (1 - p)B_d + p(p - 1) \int_0^\infty K^{p-2} O(R_e, K) dK] .$$

A non degenerate exchange rate component  $\mathcal{E}$  due to market incompleteness can also be naturally incorporated into our framework, by means of an arbitrage-free joint CGF of variables  $(M_d, M_f, R_d, R_f, \mathcal{E})$  and the corresponding non convex set of observable points.

## 2.6 Multi-Period Pricing Kernels and Returns ( $d_1 = d_2$ )

Given an investment horizon consisting of  $d = d_1 = d_2$  periods, we can easily incorporate multi-period arbitrage-free information into our framework. Let  $\{M_i : i = 1, \dots, d\}$  ( $\{R_i : i = 1, \dots, d\}$ ) be a sequence of single-period pricing kernels (risky returns) for pricing time  $i$  payoffs at time  $i - 1$  (priced at time  $i - 1$  and paying off at time  $i$ ). Whenever risk-free bond prices  $B_i$  for maturity  $i = 1, \dots, d$

are observed, then arbitrage-free CGF yields:

$$\mathcal{K}_{MR}(\iota_i, 0_{2d-i}) = \log E \left( \prod_{k=1}^i M_k \right) = \log B_i , \quad (8)$$

where  $\iota_i$  is a vector of ones in  $\mathbb{R}^i$  and  $0_{2d-i}$  a vector of zeros in  $\mathbb{R}^{2d-i}$ , i.e.,  $(\iota_i, 0_{2d-i}) \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Similarly,

$$\mathcal{K}_{MR}(\iota_i, 0_{d-i}, \iota_i, 0_{d-i}) = \log E \left[ \prod_{k=1}^i M_k R_k \right] = 0 , \quad (9)$$

i.e.,  $(\iota_i, 0_{d-i}, \iota_i, 0_{d-i}) \in \mathcal{O}_{\mathcal{K}_{MR}}$ .

### 3 Dispersion Measured by Jensen's Gap

The convexity properties of arbitrage free CGF's are directly linked to the dispersion of pricing kernel and returns. Therefore, asset pricing restrictions are naturally formulated using appropriate measures of dispersion. In this way, we obtain a general unifying approach for testing asset pricing models.<sup>6</sup>

#### 3.1 Jensen's Gap and Multivariate Dispersion

We propose to measure multivariate dispersion by a family of Jensen gaps implied by the joint CGF.

##### 3.1.1 Definition and Main Properties

**Definition 3.** (i) For a prior distribution  $\pi$  on  $\mathbb{R}^{d_1+d_2}$  and a joint CGF  $\mathcal{K}_{MR}$ , the Jensen's gap under prior  $\pi$  is:

$$\mathcal{J}_\pi(M, R) := E_\pi[\mathcal{K}_{MR}(m, r)] - \mathcal{K}_{MR}(E_\pi[(m, r)]) \geq 0 . \quad (10)$$

(ii) The marginal Jensen's gaps implied by prior  $\pi$  are defined similarly:  $\mathcal{J}_\pi(M) := E_\pi[\mathcal{K}_M(m)] - \mathcal{K}_M(E_\pi[m])$  and  $\mathcal{J}_\pi(R) := E_\pi[\mathcal{K}_R(r)] - \mathcal{K}_R(E_\pi[r])$ .

---

<sup>6</sup>Given that we measure dispersion by the convexity of a CGF, dispersion constraints implied by our approach are particularly appropriate to test log-linear, or nearly log-linear, models. Similarly, dispersion measured by the convexity of a CGF extends from Gaussian to non Gaussian settings in a transparent way.

$\mathcal{J}_\pi(M, R)$  ( $\mathcal{J}_\pi(M)$  and  $\mathcal{J}_\pi(R)$ ) measures (measure) the multivariate (marginal) dispersion of pricing kernels and returns, consistently with a number of useful properties, collected in Proposition 1. Among these, additivity under independent experiments is a useful property, e.g., when studying a term structure of dispersion in pricing kernels and returns.

**Proposition 1.** *Jensen's gap  $\mathcal{J}_\pi(M, R) \geq 0$  is a dispersion measure with the following properties:*

1.  $\mathcal{J}_\pi(M, R) = 0$ , whenever  $(M, R)$  is a degenerate random vector.
2. For any prior  $\pi$  with support not included in a strict subspace of  $\mathbb{R}^{d_1+d_2}$ ,  $\mathcal{J}_\pi(M, R) = 0$  if and only if  $(M, R)$  is a degenerate random vector.
3. Given stochastically independent pricing kernel components and returns, it follows:

$$\mathcal{J}_\pi(M, R) = \mathcal{J}_\pi(M) + \mathcal{J}_\pi(R) . \quad (11)$$

4. Given two independent random vectors  $(M, R)$  and  $(N, Q)$  in  $\mathbb{R}^{d_1+d_2}$ , it follows:

$$\mathcal{J}_\pi(M \times N, R \times Q) = \mathcal{J}_\pi(M, R) + \mathcal{J}_\pi(N, Q) , \quad (12)$$

where  $M \times N := (M_1N_1, \dots, M_{d_1}N_{d_1})$  and  $R \times Q := (R_1Q_1, \dots, R_{d_2}Q_{d_2})$ .

5.  $\mathcal{J}_\pi(M, R)$  is positively homogenous of degree 0, in the sense that for any  $\lambda_M := (\lambda_{M1}, \dots, \lambda_{Md_1})$  and  $\lambda_R := (\lambda_{M1}, \dots, \lambda_{Md_2})$ :  $\mathcal{J}_\pi(\lambda_M \circ M, \lambda_R \circ R) = \mathcal{J}_\pi(M, R)$ .<sup>7</sup>
6. For any prior  $\pi$  on  $\mathbb{R}^{d_1+d_2}$ , having prior covariance matrix  $\text{Var}_\pi(m, r)$ :

$$\mathcal{J}_\pi(M, R) = \frac{\text{tr}(\Sigma \text{Var}_\pi(m, r))}{2} , \quad (13)$$

whenever  $(\log M, \log R)$  is a vector of Gaussian variables with covariance matrix  $\Sigma$ .

<sup>7</sup> $\circ$  denotes the Hadamard product. For two  $n \times m$  matrices  $A, B$  with elements  $A_{ij}, B_{ij}$ , matrix  $A \circ B$  is the  $n \times m$  matrix with elements  $(A \circ B)_{ij} = A_{ij}B_{ij}$ .

Property 1. implies a necessary requirement for a measure of multivariate dispersion, i.e., that it is zero whenever  $(M, R)$  is a degenerate random vector. This property follows from the linearity of the joint CGF in such a case. From Property 2, a strictly positive Jensen's gap is generated by the strict convexity of the joint CGF of multivariate non degenerate random variables, whenever prior  $\pi$  is not concentrated on a strict subspace of  $\mathbb{R}^{d_1+d_2}$ . In this case,  $\mathcal{J}_\pi(M, R) = 0$  if and only if random vector  $(M, R)$  is degenerate. Property 3. is an additive decomposition of the joint Jensen's gap into the sum of marginal Jensen's gaps, when pricing kernel and returns are stochastically independent. Property 4. implies additivity of  $\mathcal{J}_\pi(M, R)$  under independent experiments, a desirable aggregation property that generalizes the additivity of univariate dispersion measures such as entropy; see, e.g., [Backus, Chernov, and Zin \(2014\)](#). Property 5. implies scale invariance, while Property 6. gives the expression for  $\mathcal{J}_\pi(M, R)$  in the benchmark case of jointly Gaussian log pricing kernel and returns.

### 3.1.2 Jensen's Gap Dimension

The support of prior  $\pi$  in equation (10) introduces a degree of flexibility in  $\mathcal{J}_\pi(M, R)$ , which can be used to localize dispersion on particular regions of a multivariate subspace. An obvious localization is on the marginals of  $(M, R)$ . More generally, a prior with support in a subspace of dimension  $d < d_1+d_2$  can measure the dispersion of particular linear combinations of pricing kernel and returns.

**Definition 4.** (i) Given a prior distribution with prior covariance matrix  $Var_\pi(m, r)$  such that  $0 < tr(Var_\pi(m, r)) < \infty$ , we call  $d_\pi := rank(Var_\pi(m, r))$  the dimension of  $\mathcal{J}_\pi(M, R)$ . (ii) Given a Jensen's gap of dimension  $d_\pi$ , a standardized Jensen's gap of dimension  $d_\pi$  is defined by

$$\mathcal{D}_\pi(M, R) := \frac{\mathcal{J}_\pi(M, R)}{tr(Var_\pi(m, r))}. \quad (14)$$

When  $d_\pi < d_1 + d_2$  the prior  $\pi$  is concentrated on a  $d_\pi$ -dimensional subspace of  $\mathbb{R}^{d_1+d_2}$ . As a consequence, if  $\mathcal{J}_\pi(M, R) = 0$  the components of random vector  $(\log M, \log R)$  are related by an affine deterministic relationship. According to Proposition 1, a standardized Jensen's gap satisfies

the convenient normalization  $\mathcal{D}_\pi(M, R) = \sigma^2/2$ , whenever  $(\log M, \log R)$  is an iid vector of Gaussian variables with variance  $\sigma^2$ . In presence of deviations from Gaussianity, the leading contribution to  $\mathcal{D}_\pi(M, R)$  is approximatively given by

$$\mathcal{D}_\pi(M, R) \approx \frac{\text{tr}(\mathcal{K}_{MR}''(E_\pi[m], E_\pi[r])\text{Var}_\pi(m, r))}{2\text{tr}(\text{Var}_\pi(m, r))}, \quad (15)$$

where  $\mathcal{K}_{MR}''(\cdot, \cdot)$  is the Hessian of  $\mathcal{K}$ .<sup>8</sup> This approximation is exact for Gaussian random vectors and likely sufficiently accurate for priors with moderate degree of multivariate skewness and kurtosis.

### 3.2 Jensen's Gap and Entropy Measures

Jensen's gaps extend well-known concepts of dispersion in the literature, such as several useful measures of entropy and co-entropy.<sup>9</sup> They also induce a broad family of new concrete measures, such as, e.g., generalized entropy and generalized co-entropy, dispersion measures linked to Chernoff (1952) information, or asymmetric measures of co-dispersion. We illustrate their properties in the context of simple asset pricing settings.

---

<sup>8</sup>Higher order contributions can be computed following the multivariate cumulant expansion approach in Jammalamadaka, Rao, and Terdik (2006). Note that

$$\mathcal{D}_\pi(M, R) \approx \sum_{i=1}^{d_\pi} \frac{\lambda_{i\pi}}{\text{tr}(\text{Var}_\pi(m, r))} \cdot \langle q_{i\pi}, \mathcal{K}_{MR}''(E_\pi[m], E_\pi[r])q_{i\pi} \rangle, \quad (16)$$

with the nonzero eigenvalues  $\lambda_{i\pi}$  and the corresponding orthonormal eigenvectors  $q_{i\pi}$  of  $\text{Var}_\pi(m, r)$ , denoting by  $\langle \cdot, \cdot \rangle$  the Euclidean scalar product. It follows that whenever dimension  $d_\pi$  is strictly less than  $d_1 + d_2$ , some linear combinations of the columns of  $\mathcal{K}_{MR}''(E_\pi[m], E_\pi[r])$  do not contribute to the right hand of approximation (16). For instance, in the one dimensional case,  $\mathcal{D}_\pi(M, R)$  can be approximated by a single quadratic form of matrix  $\mathcal{K}_{MR}''(E_\pi[m], E_\pi[r])$ .

<sup>9</sup>See, e.g., Alvarez and Jermann (2005), Backus, Chernov, and Martin (2011), Backus, Chernov, and Zin (2014), Backus, Boyarchenko, and Chernov (2014), Bakshi and Chabi-Yo (2014) and Chabi-Yo and Colacito (2013), among others.

### 3.2.1 Univariate Return and Pricing Kernel ( $d_1 = d_2 = 1$ )

*Generalized Entropy.* Given a Bernoulli prior  $\pi_\alpha$  with mass  $\alpha \in (0, 1)$  in  $(m, r) = (1, 0)$  and mass  $1 - \alpha$  in  $(m, r) = (0, 0)$ , we obtain:

$$\begin{aligned} \mathcal{D}_{\pi_\alpha}(M, R) &= \frac{\alpha \mathcal{K}_{MR}(1, 0) + (1 - \alpha) \mathcal{K}_{MR}(0, 0) - \mathcal{K}_{MR}(\alpha, 0)}{\alpha(1 - \alpha)} \\ &= \frac{\log E[(M/E(M))^\alpha]}{\alpha(\alpha - 1)} =: \mathcal{E}_\alpha(M), \end{aligned} \quad (17)$$

i.e., the  $\alpha$ -Rényi (1960) entropy of the stochastic discount factor.  $\mathcal{E}_\alpha(M)$  is a standardized Jensen's gap of dimension  $d_\pi = 1$ . Its geometric interpretation based on dispersion measure  $\mathcal{D}_{\pi_\alpha}(M, R)$  is illustrated in Figure 5. As  $\alpha \rightarrow 0$ , we obtain  $\mathcal{E}_0(M) = E[-\log(M/E(M))]$ , i.e., the stochastic discount factor entropy in Alvarez and Jermann (2005), Backus, Chernov, and Martin (2011), Backus, Chernov, and Zin (2014), among others.

*Co-Generalized Entropy.* When returns and pricing kernel are stochastically independent, Proposition 1 implies  $\mathcal{E}_\alpha(MR) = \mathcal{E}_\alpha(M) + \mathcal{E}_\alpha(R) =: \mathcal{E}_\alpha^\perp(MR)$ . We can define the  $\alpha$ -Rényi co-entropy by

$$\mathcal{C}_\alpha(M, R) := \frac{\alpha - 1}{\alpha} \left( \mathcal{E}_\alpha(MR) - \mathcal{E}_\alpha^\perp(MR) \right) = \frac{1}{\alpha^2} \log \left( \frac{E[(MR)^\alpha]}{E[M^\alpha]E[R^\alpha]} \right), \quad (18)$$

which for  $\alpha \rightarrow 0$  yields the co-entropy introduced in Backus, Boyarchenko, and Chernov (2014) and Bakshi and Chabi-Yo (2014), among others. Note that whenever  $(\log M, \log R)$  is jointly normally distributed, Proposition 1 implies  $\mathcal{C}_\alpha(M, R) = \text{Cov}(\log M, \log R)$ .

*Power Generalized Entropy.* Given a Bernoulli prior  $\pi_\alpha$  with mass  $\alpha \in (0, 1)$  in  $(m, r) = (p, 0)$  ( $p \in \mathbb{R}$ ) and mass  $1 - \alpha$  in  $(m, r) = (0, 0)$ , we obtain

$$\mathcal{D}_{\pi_\alpha}(M, R) = \frac{\log E[(M^p/E(M^p))^\alpha]}{p^2\alpha(\alpha - 1)} = \frac{1}{p^2} \mathcal{E}_\alpha(M^p) =: \mathcal{E}_\alpha^p(M), \quad (19)$$

i.e.,  $\mathcal{D}_{\pi_\alpha}(M, R)$  is proportional to the  $\alpha$ -Rényi (1960) entropy of the  $p$ -th power of the pricing kernel.  $\mathcal{E}_\alpha^p(M)$  is a standardized Jensen's gap of dimension  $d_\pi = 1$ . For  $\alpha \rightarrow 0$  we obtain  $4\mathcal{E}_0^2(M) = \mathcal{E}_0(M^2)$ ,

i.e., the entropy of the squared pricing kernel adopted in [Bakshi and Chabi-Yo \(2014\)](#) to specify tractable multivariate pricing kernel bounds.<sup>10</sup>

*Dimension*  $d_\pi > 1$ . Most dispersion measures rely on a dimension  $d_\pi = 1$ . Jensen's gaps of dimension  $d_\pi > 1$  are easily constructed. To illustrate, consider a prior  $\pi_{\alpha,\beta}$  with mass  $\alpha > 0$  in  $(m, r) = (1, 0)$ , mass  $\beta > 0$  in  $(m, r) = (0, 1)$  and mass  $1 - (\alpha + \beta)$  in  $(m, r) = (0, 0)$ . It then follows:

$$\begin{aligned} \mathcal{D}_{\pi_{\alpha,\beta}}(M, R) &= \frac{\alpha \mathcal{K}_{MR}(1, 0) + \beta \mathcal{K}_{MR}(0, 1) - \mathcal{K}_{MR}(\alpha, \beta)}{\alpha(1 - \alpha) + \beta(1 - \beta)} \\ &= \frac{\log E[(M/E(M))^\alpha (R/E[R])^\beta]}{\alpha(\alpha - 1) + \beta(\beta - 1)} =: \mathcal{E}_{\alpha,\beta}(M, R). \end{aligned} \quad (20)$$

$\mathcal{E}_{\alpha,\beta}(M, R)$  is a standardized Jensen's gap of dimension  $d_\pi = 2$  and a proper measure of bivariate dispersion. Independence of pricing kernel and returns additionally implies

$$\mathcal{E}_{\alpha,\beta}(M, R) = \frac{\alpha(\alpha - 1)\mathcal{E}_\alpha(M) + \beta(\beta - 1)\mathcal{E}_\beta(R)}{\alpha(\alpha - 1) + \beta(\beta - 1)} =: \mathcal{E}_{\alpha,\beta}^\perp(M, R), \quad (21)$$

i.e.,  $\mathcal{E}_{\alpha,\beta}^\perp(M, R)$  equals a convex combination of [Rényi \(1960\)](#) entropies. A measure of co-dispersion that (i) is zero if and only if  $M$  and  $R$  are stochastically independent and (ii) equals  $Cov(\log M, \log R)$  when returns and pricing kernel are jointly log normal, then naturally follows:

$$\mathcal{C}_{\alpha,\beta}(M, R) := \frac{\alpha(\alpha - 1) + \beta(\beta - 1)}{\alpha\beta} (\mathcal{E}_{\alpha,\beta}(M, R) - \mathcal{E}_{\alpha,\beta}^\perp(M, R)) = \frac{1}{\alpha\beta} \log \left( \frac{E[M^\alpha R^\beta]}{E[M^\alpha] E[R^\beta]} \right).$$

$\mathcal{C}_{\alpha,\beta}(M, R)$  is in general not a symmetric measure of co-dispersion. This property can be useful, e.g., to characterize pricing kernel and return dependence while explicitly accounting for the asymmetric role of pricing kernels and individual asset returns in arbitrage-free markets.<sup>11</sup>

<sup>10</sup>Using power generalized entropy and Proposition 1, it is also possible to specify convenient measures of co-dispersion, which are consistent with Pearson's measure of correlation in the log Gaussian case.

<sup>11</sup>By construction,  $\mathcal{C}_{\alpha,\alpha}(M, R) = \mathcal{C}_\alpha(M, R)$ , illustrating that lower-dimensional dispersion or co-dispersion measures are special cases of higher-dimensional dispersion or co-dispersion measures.



### 3.2.2 Domestic and Foreign Pricing Kernels ( $d_1 = 2$ and $d_2 = 2$ )

*Chernoff (1952) Information.* Given a Bernoulli prior  $\pi_\alpha$  with mass  $\alpha \in (0, 1)$  in  $(m, r) = (1, 0, 0, 0)$  and mass  $1 - \alpha$  in  $(0, 1, 0, 0)$ , we obtain:

$$\begin{aligned} \mathcal{J}_{\pi_\alpha}(M, R) &= \alpha \mathcal{K}_{MR}(1, 0, 0, 0) + (1 - \alpha) \mathcal{K}_{MR}(0, 1, 0, 0) - \mathcal{K}_{MR}(\alpha, 1 - \alpha, 0, 0) \\ &= -\log E[(M_d/E(M_d))^\alpha (M_f/E(M_f))^{1-\alpha}] =: -\log CC_\alpha(M_d, M_d) , \end{aligned}$$

with the  $\alpha$ -Chernoff (1952) coefficient  $CC_\alpha(M_d, M_d)$ . The optimal Chernoff (1952) coefficient

$$CC_{\alpha^*}(M_d, M_f) := \min_{\alpha \in (0, 1)} CC_\alpha(M_d, M_f) , \quad (22)$$

is a symmetric measure of similarity between pricing kernels, while Chernoff (1952) information (or Chernoff divergence):

$$CI_*(M_d, M_f) := -\ln CC_{\alpha^*}(M_d, M_f) = \max_{\alpha \in (0, 1)} \mathcal{J}_{\pi_\alpha}(M, R) , \quad (23)$$

is a symmetric measure of discrepancy between pricing kernels.

## 4 Informative and Observable Arbitrage Free Dispersion

The convexity of the joint CGF imposes constraints on the dispersion properties of pricing kernels and returns. Therefore, we make use of Jensen's gaps to characterize the testable dispersion properties that any asset pricing model needs to satisfy.

### 4.1 Definition

If quantities  $E_\pi[\mathcal{K}_{MR}(m, r)]$  and  $\mathcal{K}_{MR}(E_\pi[(m, r)])$  in Definition 3 are directly computable from the known values of  $\mathcal{K}_{MR}$  on observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$ , then  $\mathcal{J}_\pi(M, R)$  is observable and directly produces verifiable constraints on the convexity features of any arbitrage-free CGF. The situation is different

when either  $E_\pi[\mathcal{K}_{MR}(m, r)]$  or  $\mathcal{K}_{MR}(E_\pi[(m, r)])$  is unobservable.

**Definition 5.** (i) Jensen's gap  $\mathcal{J}_\pi(M, R)$  in Definition 3 is an informative arbitrage-free dispersion whenever (1)  $\pi$  has support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  or (2)  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ . (ii) Inequalities

$$\{\mathcal{J}_\pi(M, R) \geq 0 : \text{prior } \pi \text{ is such that (1) holds}\} , \quad (24)$$

define the set of observable arbitrage-free dispersion constraints of Type (1). (iii) Inequalities

$$\{\mathcal{J}_\pi(M, R) \geq 0 : \text{prior } \pi \text{ is such that (2) holds}\} , \quad (25)$$

define the set of observable dispersion constraints of Type (2). (iv) An informative arbitrage-free dispersion  $\mathcal{J}_\pi(M, R)$  is observable, whenever both  $\pi$  has support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  and  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ .

Informative dispersion restricts the range of possible values of an arbitrage-free CGF in unobservable parts of its domain.<sup>12</sup> Observable dispersion uniquely constrains the convexity properties of the observable arbitrage-free CGF. Further observable constraints on the convexity of the arbitrage-free CGF can be obtained using observable differences of informative dispersions.

**Definition 6.** Given priors  $\pi_1$  and  $\pi_2$  with support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  and such that  $E_{\pi_1}[(m, r)] = E_{\pi_2}[(m, r)]$ ,  $\Delta\mathcal{J}_{\pi_1, \pi_2}(M, R) := \mathcal{J}_{\pi_1}(M, R) - \mathcal{J}_{\pi_2}(M, R)$  is called an observable arbitrage-free excess dispersion.

## 4.2 Implications of Observable Dispersion and Excess Dispersion

Observable dispersions and excess dispersions naturally reflect the observability of the arbitrage-free CGF in particular regions of its domain. In order to avoid obvious model misspecifications, the model-implied and the arbitrage-free CGF have to coincide on the observable set. Therefore, the model-implied CGF convexity has to be consistent with the observable dispersion and excess dispersion. In all other cases, a dispersion violation is directly observed and a new model specification is necessary.

<sup>12</sup>This is in contrast to the observable parts of the domain, where the arbitrage-free CGF is fully identified.

**Definition 7.** Given model  $\mathbb{M}$ , an observable dispersion (excess dispersion) violation arises whenever  $\mathcal{J}_\pi(M, R) \neq \mathcal{J}_\pi^{\mathbb{M}}(M, R)$  ( $\Delta\mathcal{J}_{\pi_1, \pi_2}(M, R) \neq \Delta\mathcal{J}_{\pi_1, \pi_2}^{\mathbb{M}}(M, R)$ ) for some observable arbitrage-free dispersion  $\mathcal{J}_\pi(M, R)$  (excess dispersion  $\Delta\mathcal{J}_{\pi_1, \pi_2}(M, R)$ ).

#### 4.2.1 Transient vs. Persistent Pricing Kernel Components and Horizon Dependence

Despite the fact that observable dispersion and excess dispersion depend on directly observable convexity properties of the arbitrage-free CGF, under plausible assumptions they already characterize important properties of asset prices, including the joint dependence of permanent and transient pricing kernel components or the horizon dependence of zero-coupon bond prices.

**Example 1** (Dependence of transient and persistent pricing kernel components). Following [Bakshi and Chabi-Yo \(2014\)](#), the covariance between transient and persistent pricing kernel components is observable when the return  $R_\infty$  of the infinite maturity bond is observable:  $\text{cov}(M^T, M^P) = B - E[1/R_\infty]$ . A suitable monotonic transformation and an appropriate scaling of this equality, gives:

$$\frac{1}{2} \log (\text{cov}(M^T/E(M^T), M^P) + 1) = E_{\pi_1}[\mathcal{K}_{M^T M^P}(m^T, m^P)] - E_{\pi_2}[\mathcal{K}_{M^T M^P}(m^T, m^P)], \quad (26)$$

where  $\pi_1$  ( $\pi_2$ ) has mass 1/2 in  $(m^T, m^P) = (1, 1)$  ( $(m^T, m^P) = (1, 0)$ ) and mass 1/2 in  $(m^T, m^P) = (0, 0)$  ( $(m^T, m^P) = (0, 1)$ ). Since  $E_{\pi_1}[(m^T, m^P)] = E_{\pi_2}[(m^T, m^P)] = (1/2, 1/2)$ , equation (26) induces an observable excess dispersion  $\Delta\mathcal{J}_{\pi_1, \pi_2}(M^T, M^P)$ . In contrast to  $\text{cov}(M^T, M^P)$ , excess dispersion (26) is independent of the scale of  $M^T$ , i.e., its first moment. In this sense, it implies a definition of an excess dispersion violation that is robust with respect to an incorrect measurement of the first moment of  $1/R_\infty$ .

**Example 2** (Horizon dependence). Given a vector  $M = (M_1, \dots, M_n)$  of strictly stationary single-period stochastic discount factors  $M_i$ , pricing at time  $i - 1$  payoffs paid at time  $i$ , [Backus, Chernov,](#)

and Zin (2014) measure horizon dependence as

$$H(n) := \frac{1}{n} \mathcal{E}_0 \left( \prod_{i=1}^n M_i \right) - \mathcal{E}_0(M_1) = \frac{\ln(B_n)}{n} - \ln(B_1), \quad (27)$$

where  $B_i$  ( $i = 1, \dots, n$ ) is the price of a zero bond with maturity  $i$ . Therefore,  $H(n)$  is a measure of the (negative) slope of the yield curve at horizon  $n$ . Denoting by  $\iota(e_i)$  the vector of ones (the  $i$ -th unit vector) in  $\mathbb{R}^n$ , we can write  $H(n)$  as an arbitrage-free excess dispersion:<sup>13</sup>

$$H(n) = \frac{\mathcal{K}_M(\iota)}{n} - \frac{\sum_{i=1}^n \mathcal{K}_M(e_i)}{n} = \Delta \mathcal{J}_{\pi_1, \pi_2}(M), \quad (28)$$

using a prior  $\pi_1$  ( $\pi_2$ ) with mass  $1/n$  in  $m = \iota$  and mass  $1 - 1/n$  in  $m = 0_n$  (with uniform mass  $1/n$  in each unit vector  $e_i$ ). Thus, horizon dependence can be understood as a particular convexity requirement, along the main diagonal in the domain of the arbitrage-free CGF of  $M$ .

#### 4.2.2 Implications for Observable Model-Implied CGFs

Intuitively, the absence of observable dispersion or excess dispersion violations must constrain the convexity of the arbitrage-free CGF on observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$  quite strongly. Indeed, in that case the observable model-implied CGF  $\mathcal{K}_{MR}^{\mathbb{M}}|_{\mathcal{O}_{\mathcal{K}_{MR}}}$  is already uniquely identified, up to a linear transformation, as is stated precisely in the next proposition, proven in the Supplemental Appendix.

**Proposition 2.** *If there are no observable arbitrage free dispersion or excess dispersion violations, then  $\mathcal{K}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}} - \mathcal{K}_{MR}^{\mathbb{M}}|_{\mathcal{O}_{\mathcal{K}_{MR}}}$  is linear, i.e., there exists a vector  $e \in \mathbb{R}^{d_1+d_2}$  such that  $\mathcal{K}_{MR}(o) = \mathcal{K}_{MR}^{\mathbb{M}}(o) + e' \cdot o$  for every  $o = (m, r) \in \mathcal{O}_{\mathcal{K}_{MR}}$ .*

From Proposition 2, the absence of observable dispersion or excess dispersion violations implies an observable model-implied CGF that is uniquely identified, up to a possibly inappropriate scaling of pricing kernel components or returns. Inappropriate scaling can be corrected by rescaling, e.g., when the scaling discrepancy is concentrated in the marginal distribution of pricing kernel components. In

---

<sup>13</sup>Note that  $E_{\pi_1}[m] = E_{\pi_2}[m] = \iota/n$ .

contrast, rescaling does not correct observable dispersion or excess dispersion violation, as Jensen's gap is homogenous of degree zero (see point 5. of Proposition 1). This motivates the next definition.

**Definition 8** (Scaling Discrepancy). *Whenever  $\mathcal{K}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}} - \mathcal{K}_{MR}^M|_{\mathcal{O}_{\mathcal{K}_{MR}}}$  is linear, we say that there is a pure scaling discrepancy between observable model-implied and arbitrage-free CGFs.*

In summary, a discrepancy between observable arbitrage-free and model-implied CGFs can emerge either from a scaling discrepancy or from a dispersion or excess dispersion violation. Dispersion violations in the marginal CGF of returns can in principle be corrected quite precisely, when the return distribution is observable with sufficient accuracy. In contrast, correcting dispersion violations in the marginal CGF of the pricing kernel is more challenging, because dispersion is only sparsely observable along that dimension.

### 4.3 Implications of Constraints of Type (1) and Upper Arbitrage Free CGF

When informative arbitrage-free dispersion is unobservable, dispersion constraints of Type (1) imply testable constraints for the arbitrage-free CGF on the convex hull of all observable points, defined by:

$$\bar{\mathcal{O}}_{\mathcal{K}_{MR}} := \{E_{\pi}[(m, r)] : \text{prior } \pi \text{ has support in } \mathcal{O}_{\mathcal{K}_{MR}}\} . \quad (29)$$

In this way, the arbitrage-free CGF is restricted also in not directly observable regions of its domain.

Positivity of Jensen's gap yields for any  $(m_{\star}, r_{\star}) \in \bar{\mathcal{O}}_{\mathcal{K}_{MR}}$  the upper bound:

$$\mathcal{K}_{MR}(m_{\star}, r_{\star}) \leq E_{\pi}[\mathcal{K}_{MR}(m, r)] , \quad (30)$$

where the right hand side of this inequality is observable because  $\pi$  has support on  $\mathcal{O}_{\mathcal{K}_{MR}}$ . The tightest such upper bound follows from the infimum over all priors with support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  and such that  $(m_{\star}, r_{\star}) = E_{\pi}[(m, r)]$ . This motivates the concept of an upper arbitrage-free CGF.

**Definition 9** (Upper Arbitrage Free CGF). *The upper arbitrage free CGF is the function  $\mathcal{K}_{MR}^U$  :*

$\mathbb{R}^{d_1+d_2} \rightarrow \mathbb{R} \cup \{+\infty\}$ , defined for any  $(m_\star, r_\star) \in \mathbb{R}^{d_1+d_2}$  by:<sup>14</sup>

$$\mathcal{K}_{MR}^U(m_\star, r_\star) := \inf_{\pi} \{E_{\pi}[\mathcal{K}_{MR}(m, r)]\} , \quad (31)$$

where the infimum is over priors with support on  $\mathcal{O}_{\mathcal{K}_{MR}}$  and such that  $(m_\star, r_\star) = E_{\pi}[(m, r)]$ .

The upper arbitrage-free CGF is a convex upper extension of  $\mathcal{K}_{MR}$  to  $\mathbb{R}^{d_1+d_2}$ , which coincides with  $\mathcal{K}_{MR}$  on the observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$  and defines a finite upper bound for  $\mathcal{K}_{MR}$  on the convex hull  $\overline{\mathcal{O}_{\mathcal{K}_{MR}}}$ .<sup>15</sup> By definition,  $\mathcal{K}_{MR}^U$  is computable from the set  $\mathcal{O}_{\mathcal{K}_{MR}}$  of observable points and implies for any  $(m_\star, r_\star) \in \overline{\mathcal{O}_{\mathcal{K}_{MR}}}$  a nontrivial inequality for the specification of an arbitrage-free CGFs:

$$\mathcal{K}_{MR}(m_\star, r_\star) \leq \mathcal{K}_{MR}^U(m_\star, r_\star) . \quad (32)$$

Figure 3(a) illustrates the convex domain  $\overline{\mathcal{O}_{\mathcal{K}_{MR}}}$  of finite upper arbitrage-free CGF values, generated by the following empirically observable CGF points:  $\mathcal{K}_{MR}(1, 0) = \log B$ ,  $\mathcal{K}_{MR}(1, 1) = 0$  and  $\mathcal{K}_{MR}(0, r) = \log E[R^r]$  for  $r \in (0, 1)$ .  $\overline{\mathcal{O}_{\mathcal{K}_{MR}}}$  is convex and closed. Outside this region, the upper bound on  $\mathcal{K}_{MR}$  generated by Type (1) dispersion constraints is trivial. Naturally, a violation of bound (32) in regions where it is non-trivial provides useful information for the specification of asset pricing models.

**Definition 10.** *Given a model  $\mathbb{M}$  and unobservable point  $(m_\star, r_\star) \in \overline{\mathcal{O}_{\mathcal{K}_{MR}}} \setminus \mathcal{O}_{\mathcal{K}_{MR}}$ , an arbitrage-free dispersion violation of Type (1) arises whenever  $\mathcal{K}_{MR}^{\mathbb{M}}(m_\star, r_\star) > \mathcal{K}_{MR}^U(m_\star, r_\star)$ .*

#### 4.4 Implications of Constraints of Type (2) and Lower Arbitrage Free CGF

Given an informative unobservable arbitrage-free dispersion, dispersion constraints of Type (2) imply a second set of observable constraints for an arbitrage-free CGF:

$$\mathcal{K}_{MR}(E_{\pi}[(m, r)]) \leq E_{\pi}[\mathcal{K}_{MR}(m, r)] , \quad (33)$$

<sup>14</sup>By definition,  $\inf_{\emptyset} E_{\pi}[\mathcal{K}_{MR}(m, r)] := +\infty$ .

<sup>15</sup>See Peters and Wakker (1987), among other, for the properties of finite convex extensions of a convex function.

where the left hand side of this inequality is observable when  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Inequality (33) is a lower bound for the expected arbitrage-free CGF under any prior  $\pi$  with observable expectation  $E_\pi[(m, r)]$ . From inequality (33), we obtain directly computable lower bounds for the arbitrage-free CGF in unobservable regions of its domain. Given unobservable point  $(m_\star, r_\star)$ , consider a prior  $\pi$  with support in  $\mathcal{O}_{\mathcal{K}_{MR}} \cup \{(m_\star, r_\star)\}$  and such that  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ . Inequality (33) then gives:

$$\mathcal{K}_{MR}(E_\pi[(m, r)]) \leq E_\pi[\mathcal{K}_{MR}(m, r)1_{\mathcal{O}_{\mathcal{K}_{MR}}}(m, r)] + \pi(m_\star, r_\star)\mathcal{K}_{MR}(m_\star, r_\star), \quad (34)$$

where  $1_{\mathcal{O}_{\mathcal{K}_{MR}}}$  is the indicator function of set  $\mathcal{O}_{\mathcal{K}_{MR}}$ . In (34) all quantities but  $\mathcal{K}_{MR}(m_\star, r_\star)$  are observable and the following lower bound holds:

$$\mathcal{K}_{MR}(m_\star, r_\star) \geq \frac{\mathcal{K}_{MR}(E_\pi[(m, r)]) - E_\pi[\mathcal{K}_{MR}(m, r)1_{\mathcal{O}_{\mathcal{K}_{MR}}}(m, r)]}{\pi(m_\star, r_\star)}. \quad (35)$$

The tightest such lower bound is given by the supremum of the right hand side of inequality (35) over priors with support in  $\mathcal{O}_{\mathcal{K}_{MR}} \cup \{(m_\star, r_\star)\}$  and such that  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ . This motivates the concept of a lower arbitrage-free CGF.

**Definition 11** (Lower Arbitrage Free CGF). *The lower arbitrage-free CGF is the function  $\mathcal{K}_{MR}^L : \mathbb{R}^{d_1+d_2} \rightarrow \mathbb{R} \cup \{-\infty\}$ , defined for any  $(m_\star, r_\star) \in \mathbb{R}^{d_1+d_2}$  by:<sup>16</sup>*

$$\mathcal{K}_{MR}^L(m_\star, r_\star) := \sup_{\pi} \left\{ \frac{\mathcal{K}_{MR}(E_\pi[(m, r)]) - E_\pi[\mathcal{K}_{MR}(m, r)1_{\mathcal{O}_{\mathcal{K}_{MR}}}(m, r)]}{\pi(m_\star, r_\star)} \right\}, \quad (36)$$

where the supremum is over priors with support in  $\mathcal{O}_{\mathcal{K}_{MR}} \cup \{(m_\star, r_\star)\}$  and such that  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ .

The lower arbitrage-free CGF is a convex lower extension of  $\mathcal{K}_{MR}$  to  $\mathbb{R}^{d_1+d_2}$ , which coincides with  $\mathcal{K}_{MR}$  on the observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$ . By construction,  $\mathcal{K}_{MR}^L$  is observable from the set  $\mathcal{O}_{\mathcal{K}_{MR}}$  of

<sup>16</sup>By definition,  $\sup_{\emptyset} E_\pi[\mathcal{K}_{MR}(m, r)] := -\infty$ .

empirically observable points and implies a nontrivial inequality on the arbitrage-free CGF:

$$\mathcal{K}_{MR}(m_\star, r_\star) \geq \mathcal{K}_{MR}^L(m_\star, r_\star) . \quad (37)$$

This lower bound is finite whenever a prior exists with support in  $\mathcal{O}_{\mathcal{K}_{MR}} \cup \{(m_\star, r_\star)\}$  and such that  $E_\pi[(m, r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$ <sup>17</sup>. The set of points where  $\mathcal{K}_{MR}^L$  is finite follows from the following identity:

$$(m_\star, r_\star) = (E_\pi[(m, r)] - E_\pi[(m, r)1_{\mathcal{O}_{MR}}(m, r)]) / \pi(m_\star, r_\star) . \quad (38)$$

We denote by  $\underline{\mathcal{O}}_{\mathcal{K}_{MR}} \subset \mathbb{R}^{d_1+d_2}$  the domain on which  $\mathcal{K}_{MR}^L$  is finite. Figure 3(b) illustrates the non-convex domain  $\underline{\mathcal{O}}_{\mathcal{K}_{MR}}$  of finite lower arbitrage-free CGF values, generated by the following observable CGF points:  $\mathcal{K}_{MR}(1, 0) = \log B$ ,  $\mathcal{K}_{MR}(1, 1) = 0$  and  $\mathcal{K}_{MR}(0, r) = \log E[R^r]$  for  $r \in (0, 1)$ .  $\underline{\mathcal{O}}_{\mathcal{K}_{MR}}$  is not convex, but is closed. Outside this region, the lower bound on  $\mathcal{K}_{MR}$  generated by Type (2) dispersion constraints is trivial. A violation of a nontrivial bound (37) thus provides additional information for the specification of asset pricing models.

**Definition 12.** *Given a model  $\mathbb{M}$  and unobservable point  $(m_\star, r_\star) \in \underline{\mathcal{O}}_{\mathcal{K}_{MR}} \setminus \mathcal{O}_{\mathcal{K}_{MR}}$ , an arbitrage-free dispersion violation of Type (2) arises whenever  $\mathcal{K}_{MR}^{\mathbb{M}}(m_\star, r_\star) < \mathcal{K}_{MR}^L(m_\star, r_\star)$ .*

## 4.5 Implications of Dispersion Constrains for Dispersion Bounds

Violations of Type (1) and (2) can be generated both by an inappropriate model dispersion or an inappropriate model scaling of pricing kernel or returns. Often, scale-invariant dispersion bounds easily follow from the upper and lower CGF's in Definition 4. Indeed, for any two priors  $\pi, \pi'$ , it

<sup>17</sup> $\mathcal{K}_{MR}^L$  is also related to the minimal convex extension of convex function  $\mathcal{K}_{MR}$ ; see Dragomirescu and Ivan (1992), among others.



directly follows from the definitions:

$$\mathcal{D}_\pi(M, R) \geq \left( \frac{E_\pi[\mathcal{K}_{MR}^L(m, r)] - \mathcal{K}_{MR}^U(E_\pi[(m, r)])}{\text{tr}(\text{Var}_\pi(m, r))} \right)^+ =: \mathcal{D}_\pi^L(M, R), \quad (39)$$

$$\mathcal{D}_{\pi'}(M, R) \leq \frac{E_{\pi'}[\mathcal{K}_{MR}^U(m, r)] - \mathcal{K}_{MR}^L(E_{\pi'}[(m, r)])}{\text{tr}(\text{Var}_{\pi'}(m, r))} =: \mathcal{D}_{\pi'}^U(M, R), \quad (40)$$

where  $(x)^+ := \max(x, 0)$  is the positive part of  $x$ . Dispersion bounds (39) and (40) are not binding when the right hand side equals 0 and  $+\infty$ , respectively. A necessary condition for a binding bound (39) is that  $\pi$  has support in  $\underline{\mathcal{O}}_{\mathcal{K}_{MR}}$  and  $E_\pi[(m, r)] \in \overline{\mathcal{O}}_{\mathcal{K}_{MR}}$ . If  $\pi$  has support in  $\mathcal{O}_{\mathcal{K}_{MR}}$ , this bound directly reflects Type (1) constraints. Whenever two such priors with identical mean exist, then dispersion bound (39) is always binding. Bound (40) is binding if and only if prior  $\pi'$  has support in  $\overline{\mathcal{O}}_{\mathcal{K}_{MR}}$  and  $E_{\pi'}[(m, r)] \in \underline{\mathcal{O}}_{\mathcal{K}_{MR}}$ .<sup>18</sup>

## 4.6 Diagnostic Tests of Asset Pricing Models

A general approach for testing asset pricing models can rely on a test of the null hypothesis:

$$\mathcal{H}_0(m^*, r^*) : \mathcal{K}_{MR}^L(m^*, r^*) \leq \mathcal{K}_{MR}^M(m^*, r^*) \leq \mathcal{K}_{MR}^U(m^*, r^*), \quad (41)$$

over a range of relevant arguments  $(m^*, r^*) \in \mathbb{R}^{d_1+d_2}$ . On observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$ , these inequalities are equalities and the relevant (composite) null hypothesis is:

$$\mathcal{H}_0(\mathcal{O}_{\mathcal{K}_{MR}}) : \mathcal{K}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}} = \mathcal{K}_{MR}^M|_{\mathcal{O}_{\mathcal{K}_{MR}}}. \quad (42)$$

On unobservable set  $\mathcal{O}_{\mathcal{K}_{MR}}^c$  null hypothesis (41) depends on true inequalities. Consistently with the discussion in Section 4.3, the inequality on the right hand side of this null hypothesis is binding on

<sup>18</sup>In general, bounds (39) and (40) are not both always binding for the same prior  $\pi$ :  $\mathcal{D}_\pi^L(M, R) \leq \mathcal{D}_\pi(M, R) \leq \mathcal{D}_\pi^U(M, R)$ . A necessary condition is that  $\pi$  has support in  $\underline{\mathcal{O}}_{\mathcal{K}_{MR}} \cap \overline{\mathcal{O}}_{\mathcal{K}_{MR}}$  and is such that  $E_\pi[(m, r)] \in \underline{\mathcal{O}}_{\mathcal{K}_{MR}} \cap \overline{\mathcal{O}}_{\mathcal{K}_{MR}}$ . Such a situation can arise when the observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$  is not extremal.

set  $\overline{\mathcal{O}}_{\mathcal{K}_{MR}}$  and the resulting (composite) null hypothesis is:

$$\mathcal{H}_0(\overline{\mathcal{O}}_{\mathcal{K}_{MR}}) : \mathcal{K}_{MR}^{\mathbb{M}}|_{\overline{\mathcal{O}}_{\mathcal{K}_{MR}}} \leq \mathcal{K}_{MR}|_{\overline{\mathcal{O}}_{\mathcal{K}_{MR}}} . \quad (43)$$

Similarly, the inequality on the left hand side of null hypothesis (41) is binding on set  $\underline{\mathcal{O}}_{\mathcal{K}_{MR}}$  and the resulting (composite) null hypothesis is:

$$\mathcal{H}_0(\underline{\mathcal{O}}_{\mathcal{K}_{MR}}) : \mathcal{K}_{MR}|_{\underline{\mathcal{O}}_{\mathcal{K}_{MR}}} \leq \mathcal{K}_{MR}^{\mathbb{M}}|_{\underline{\mathcal{O}}_{\mathcal{K}_{MR}}} . \quad (44)$$

#### 4.6.1 Model Diagnostics Based on Null Hypothesis $\mathcal{H}_0(\mathcal{O}_{\mathcal{K}_{MR}})$

A diagnostics test of null hypothesis

$$\mathcal{H}_0(\mathcal{O}_{\mathcal{K}_{MR}}) = \bigcap_{(m^*, r^*) \in \mathcal{O}_{\mathcal{K}_{MR}}} \mathcal{H}_0(m^*, r^*) , \quad (45)$$

tests the specification of a CGF in observable parts of the domain of the arbitrage-free CGF. Such a test is naturally complemented by a test of the observable model-implied dispersion and excess dispersion properties, which is a test of the (composite) null hypothesis:

$$\mathcal{J}_{\pi}^{\mathbb{M}}(M, R) = \mathcal{J}_{\pi}(M, R) , \quad (46)$$

$$\Delta \mathcal{J}_{\pi_1, \pi_2}^{\mathbb{M}}(M, R) = \Delta \mathcal{J}_{\pi_1, \pi_2}(M, R) , \quad (47)$$

for any observable arbitrage-free dispersion and excess dispersion. Consistent with Proposition 2, this two-step approach isolates potential rejections of  $\mathcal{H}_0(\mathcal{O}_{\mathcal{K}_{MR}})$  due to an observable scaling mismatch from those due to an inappropriate observable model dispersion or excess dispersion.

#### 4.6.2 Model Diagnostics Based on Null Hypotheses $\mathcal{H}_0(\overline{\mathcal{O}}_{\mathcal{K}_{MR}})$ and $\mathcal{H}_0(\underline{\mathcal{O}}_{\mathcal{K}_{MR}})$

A diagnostics test of null hypotheses

$$\mathcal{H}_0(\overline{\mathcal{O}}_{\mathcal{K}_{MR}}) = \bigcap_{(m^*, r^*) \in \overline{\mathcal{O}}_{\mathcal{K}_{MR}}} \mathcal{H}_0(m^*, r^*) ; \quad \mathcal{H}_0(\underline{\mathcal{O}}_{\mathcal{K}_{MR}}) = \bigcap_{(m^*, r^*) \in \underline{\mathcal{O}}_{\mathcal{K}_{MR}}} \mathcal{H}_0(m^*, r^*) , \quad (48)$$

tests the specification of an asset pricing model in unobservable parts of the domain of the arbitrage-free CGF. In order to isolate a potential violation of  $\mathcal{H}_0(\overline{\mathcal{O}}_{\mathcal{K}_{MR}})$  or  $\mathcal{H}_0(\underline{\mathcal{O}}_{\mathcal{K}_{MR}})$  due to inappropriate scaling from those due to inappropriate unobservable dispersion or excess dispersion, it is convenient to complement these tests by a set of scale invariant dispersion tests, based on the dispersion bounds introduced in Section 4.5. Using dispersion bounds, a scale independent diagnostics test for asset pricing models can rely on a test of the inequalities:

$$\mathcal{D}_{\pi}^L(M, R) \leq \mathcal{D}_{\pi}^M(M, R) \leq \mathcal{D}_{\pi}^U(M, R) , \quad (49)$$

over a range of relevant priors  $\pi$  implying a binding dispersion bound. When two priors  $\pi_1, \pi_2$  with support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  exist, such that  $E_{\pi_1}[(m, r)] = E_{\pi_2}[(m, r)]$  and  $E_{\pi_1}[\mathcal{K}_{MR}(m, r)] > E_{\pi_2}[\mathcal{K}_{MR}(m, r)]$ , a binding lower bound in the LHS of inequality (49) is easily available, using the convexity of  $\mathcal{K}_{MR}^U$ :

$$\mathcal{D}_{\pi_1}^L(M, R) \geq \frac{E_{\pi_1}[\mathcal{K}_{MR}(m, r)] - \mathcal{K}_{MR}^U(E_{\pi_1}[(m, r)])}{tr(Var_{\pi_1}(m))} \geq \frac{E_{\pi_1}[\mathcal{K}_{MR}(m, r)] - E_{\pi_2}[\mathcal{K}_{MR}(m, r)]}{tr(Var_{\pi_1}(m))} > 0 .$$

In contrast, a binding upper bound in the RHS of inequalities (49) emerges only when observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$  is not extremal. The observable set implied by the [Bansal and Yaron \(2004\)](#) long-run risk model introduced in Section 6 is extremal. Therefore, our empirical analysis of the arbitrage free dispersion properties of long-run risk models in Section 6 is based on tests of lower dispersion bounds.

## 5 Explicit Pricing Kernel Bounds Induced by Dispersion Constraints

We derive with a unifying dispersion approach explicit model-free pricing kernel bounds induced by constraints of Type (1) or (2). This approach allows an easy derivation of existing sharp univariate bounds in the literature. Moreover, it is suitable for a natural extension to economies with multiple pricing kernel components, for which we obtain new sharp pricing kernel bounds. We illustrate our approach in the benchmark economy with a single pricing kernel. We then extend it to pricing kernels with transient and persistent components and to international economies with domestic and foreign pricing kernels.

### 5.1 Univariate Pricing Kernel Bounds

Given a univariate pricing kernel ( $d_1 = 1$ ) and  $d_2$  risky returns, we consider an arbitrage-free joint CGF restricted by complete return observability ( $(0, r) \in \mathcal{O}_{\mathcal{K}_{MR}}$  for any  $r \in \mathbb{R}^{d_2}$ ), a risk-free bond with price  $B$  ( $(1, 0_{d_2}) \in \mathcal{O}_{\mathcal{K}_{MR}}$ ) and risky returns  $R = (R_1, \dots, R_{d_2})$  ( $(1, e_i) \in \mathcal{O}_{\mathcal{K}_{MR}}$  for each unit vector  $e_i$  in  $\mathbb{R}^{d_2}$ ).

#### 5.1.1 Dispersion Constraints of Type (1) and Entropy Bounds

For any  $\alpha \in (0, 1)$ , a first set of dispersion constraints of Type (1) follows using a Bernoulli prior  $\pi$  with mass  $\alpha \in (0, 1)$  on  $(1, e_i)$  and mass  $1 - \alpha$  on  $(0, -\frac{\alpha}{1-\alpha}e_i)$ . Indeed, from the dispersion constraint

$$\mathcal{J}_\pi(M, R) = E_\pi[\mathcal{K}_{MR}(m, r)] - \mathcal{K}_{MR}(E_\pi[m, r]) \geq 0, \quad (50)$$

we have  $\mathcal{K}_{MR}(1, e_i) = 0$  and  $(0, -\frac{\alpha}{1-\alpha}e_i) \in \mathcal{O}_{\mathcal{K}_{MR}}$ , i.e., prior  $\pi$  has support in  $\mathcal{O}_{\mathcal{K}_{MR}}$ . Consequently, inequality (50) defines an arbitrage-free dispersion constraint of Type (1). Explicit calculations give:

$$\frac{1}{1-\alpha} \log E[M^\alpha] = \frac{1}{1-\alpha} \mathcal{K}_{MR}(\alpha, 0_{d_2}) \leq \mathcal{K}_{MR}(0, -\frac{\alpha}{1-\alpha}e_i) = \log E[R_i^{-\alpha/(1-\alpha)}]. \quad (51)$$

Equivalently, this is a lower bound on the Rényi (1960) entropy of the pricing kernel:

$$\mathcal{E}_\alpha(M) = \frac{1}{\alpha(\alpha-1)} \log E[(M/E(M))^\alpha] \geq -\frac{1}{\alpha} \log E[R_i^{-\alpha/(1-\alpha)}] - \frac{1}{\alpha-1} \log E[M]. \quad (52)$$

As  $\alpha \rightarrow 0$ , this is the entropy bound in, e.g., Alvarez and Jermann (2005):  $\mathcal{E}_0(M) \geq \log E[R_i E(M)]$ .

### 5.1.2 Dispersion Constraints of Type (2) and Entropy Bounds

For any  $\alpha > 1$ , a dispersion constraint of Type (2) follows using a prior with mass  $1/\alpha \in (0, 1)$  in  $(\alpha, 0_{d_2})$  and mass  $(\alpha-1)/\alpha$  in  $(0, \frac{\alpha}{\alpha-1}e_i)$ . Indeed, since  $E_\pi[m, r] = (1, e_i) \in \mathcal{O}_{\mathcal{K}_{MR}}$ , we obtain:

$$0 = \mathcal{K}_{MR}(1, e_i) \leq \frac{1}{\alpha} \mathcal{K}_{MR}(\alpha, 0_{d_2}) + \frac{\alpha-1}{\alpha} \mathcal{K}_{MR}\left(0, -\frac{\alpha}{1-\alpha}e_i\right), \quad (53)$$

which is pricing kernel bound (51) for  $\alpha > 1$ . Such pricing kernel bounds for  $\alpha > 0$  are equivalent to the pricing kernel bounds derived in Liu (2013) using Hölder-type inequalities. For  $\alpha < 0$ , we obtain a second set of constraints of Type (2) using a prior with mass  $1/(1-\alpha) \in (0, 1)$  in  $(\alpha, 0_{d_2})$  and mass  $-\alpha/(1-\alpha)$  in  $(1, e_i)$ . Indeed, as  $E_\pi[m, r] = (0, -\frac{\alpha}{1-\alpha}e_i) \in \mathcal{O}_{\mathcal{K}_{MR}}$  and  $\mathcal{K}_{MR}(1, e_i) = 0$ , we have:

$$\mathcal{K}_{MR}(0, -\frac{\alpha}{1-\alpha}e_i) \leq \frac{1}{1-\alpha} \mathcal{K}_{MR}(\alpha, 0_{d_2}), \quad (54)$$

which is the reversed pricing kernel bound (51) for  $\alpha < 0$ . Equivalently,

$$\mathcal{E}_\alpha(M) \leq \frac{(1-\alpha) \log E[R_i^{-\alpha/(1-\alpha)}] - \alpha \log E[M]}{\alpha(\alpha-1)}. \quad (55)$$

These bounds for  $\alpha < 0$  are equivalent to the bounds derived in Snow (1991) using Hölder-type inequalities. In summary, we have obtained the following proposition.

**Proposition 3.** *For any  $\alpha \in \mathbb{R}$ , the following dispersion constraints hold:*

$$\frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha-1)} \geq -\frac{\mathcal{K}_{R_i}(\alpha/(\alpha-1))}{\alpha}; \quad i = 1, \dots, d_2. \quad (56)$$

Figure 4 illustrates for  $d_1 = d_2 = 1$  the construction of the above pricing kernel bounds. For  $\alpha = 1/2$ , we apply a constraint of Type (1) with observable points  $(m, r) = (1, 1)$  and  $(m, r) = (0, -1)$ . In this case, the unobserved point  $(m, r) = (0, 1/2)$  lies in the convex hull of the observable points. For  $\alpha = 2$  ( $\alpha = -1$ ), we apply a constraint of Type (2) with observable points  $(m, r) = (1, 1)$  and  $(m, r) = (0, 2)$  (points  $(m, r) = (1, 1)$  and  $(m, r) = (0, 1/2)$ ). In these last two cases, the unobserved point  $(m, r) = (0, 2)$  ( $(m, r) = (0, -1)$ ) lies outside of the convex hull of the observable points.

### 5.1.3 Bound Tightness

An important question is whether the pricing kernel bounds resulting from the dispersion constraints in Proposition 3 are tight, in the sense that they are the sharpest bounds implied by the arbitrage-free constraints on returns  $(R_0, R_1, \dots, R_{d_2})$ , where  $R_0$  is the risk-free return.<sup>19</sup> Given the joint distribution of returns, the tightest lower bound on  $\frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha-1)}$  is the one implied by the minimum divergence pricing kernel in Almeida and Garcia (2011).<sup>20</sup> We find that the pricing kernel bounds implied by arbitrage-free dispersion constraints are sharp, after a maximization of the right hand side on inequality (56) over all returns of portfolios with weights  $1 - \sum_{i=1}^{d_2} \lambda_i, \lambda_1, \dots, \lambda_{d_2}$  in returns  $R_0, \dots, R_{d_2}$ .

**Proposition 4** (Bound Tightness). *Let  $M^*$  be the solution of the minimal divergence problem in Almeida and Garcia (2011):*

$$\inf_M \left\{ \frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha-1)} \right\}, \quad (59)$$

<sup>19</sup>Bansal and Lehmann (1997) and Alvarez and Jermann (2005), among others, show that the tightest pricing kernel entropy bound, which is obtained for  $\alpha \rightarrow 0$  in our setting, is the one generated by the return of the growth optimal portfolio. By construction, that bound is equivalent to the bound generated for  $\alpha = 0$  by arbitrage-free dispersion constraints incorporating the observability of the return on the growth optimal portfolio.

<sup>20</sup>This follows from the equivalence of the optimization problem:

$$\inf_M \left\{ \frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha-1)} \right\} \text{ s.t. } \mathcal{K}_{MR}(1, e_i) = 0 \text{ (} i = 0, \dots, d_2 \text{)}, \quad (57)$$

with the minimal divergence problem:

$$\inf_M \left\{ \frac{E[M^\alpha] - E[M]^\alpha}{\alpha(\alpha-1)} \right\} \text{ s.t. } E[MR_i] = 1 \text{ (} i = 0, \dots, d_2 \text{)}, \quad (58)$$

in Almeida and Garcia (2011).

subject to  $M > 0$  and  $\mathcal{K}_{MR}(1, e_i) = 0$  for  $i = 0, \dots, d_2$ . Consider the following maximization problem:

$$\sup_{\lambda} \left\{ -\frac{\mathcal{K}_{R_{\lambda}}(\alpha/(\alpha-1))}{\alpha} \right\}, \quad (60)$$

where  $R_{\lambda} = \sum_{i=1}^{d_2} \lambda_i R_i + (1 - \sum_{i=1}^{d_2} \lambda_i) R_0$ . (with the constraint  $R_{\lambda} > 0$ ). Given the solution  $\lambda^*$  to this problem, it follows:

$$\frac{\mathcal{K}_{M^*}(\alpha)}{\alpha(\alpha-1)} = -\frac{\mathcal{K}_{R_{\lambda^*}}(\alpha/(\alpha-1))}{\alpha}, \quad (61)$$

and the minimal divergence stochastic discount factor is given by  $M^* = R_{\lambda^*}^{-1/(1-\alpha)} / E[R_{\lambda^*}^{-\alpha/(1-\alpha)}]$ .

Proposition 4 shows that the tightest lower bound on  $\frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha-1)}$ , which is compatible with a stochastic discount factor pricing returns  $R_0, \dots, R_{d_2}$ , follows from a single arbitrage-free dispersion constraint for portfolio return  $R_{\lambda^*}$ .<sup>21</sup> The bound tightness in Proposition 4 also implies closed-form upper and lower arbitrage free CGF of the pricing kernel, with explicit domains  $D_U = [0, 1]$  and  $D_L = [0, 1]^c$  where these functions take finite values:

$$\mathcal{K}_M(\alpha) \geq \mathcal{K}_M^L(\alpha) = (1-\alpha)\mathcal{K}_R(\alpha/(\alpha-1)) > -\infty; \alpha \in D_L, \quad (62)$$

$$\mathcal{K}_M(\alpha) \leq \mathcal{K}_M^U(\alpha) = (1-\alpha)\mathcal{K}_R(\alpha/(\alpha-1)) < \infty; \alpha \in D_U. \quad (63)$$

## 5.2 Multivariate Pricing Kernel Bounds

Using the same general dispersion approach as in the previous section, we now address multivariate settings with multiple pricing kernel components.

<sup>21</sup>Proposition 4 also implies that while the pricing kernel bounds in Snow (1991) and Liu (2013) derived from univariate pricing constraints are not sharp in general, they are after an optimization with respect to the family of portfolio returns generated by the priced underlying assets in an arbitrage free market. This follows directly from the equivalence of the minimum divergence stochastic discount factor bound in Almeida and Garcia (2011) and the optimized dispersion bound in Proposition 4.

### 5.2.1 Dispersion Constraints on Transient and Persistent Pricing Kernel Components

In the context of Section 2.4, we directly obtain a family of dispersion constraints of Type (1), using for given  $0 < \beta < 1$  a prior with mass  $\beta$  in  $(m, r) = (1, 1, 1) \in \mathcal{O}_{\mathcal{K}_{MR}}$  and mass  $1 - \beta$  in  $(m, r) = (-(\beta - \alpha), 0, -\beta)/(1 - \beta) \in \mathcal{O}_{\mathcal{K}_{MR}}$ . In this way, we have for any  $\alpha \in \mathbb{R}$ :

$$\frac{\mathcal{K}_{M^T M^P}(\alpha, \beta)}{\beta(\beta - 1)} \geq -\frac{\mathcal{K}_{M^T R}\left(-\frac{\beta - \alpha}{1 - \beta}, -\frac{\beta}{1 - \beta}\right)}{\beta} = -\frac{\mathcal{K}_{R_\infty R}\left(\frac{\beta - \alpha}{1 - \beta}, -\frac{\beta}{1 - \beta}\right)}{\beta}. \quad (64)$$

This bound constraints the joint distribution of transient and persistent pricing kernel components and coincides with the univariate pricing kernel bound (56) when  $\beta = \alpha$ . Note that a violation of bound (64) can arise from an inappropriate model-implied scaling of  $M^T = 1/R_\infty$ . A scale-independent bound equivalent to bound (64) is given in the next proposition, proven in the Supplemental Appendix.

**Proposition 5** (Bound Tightness with SDF Decomposition). *Given the physical distribution  $\mathbb{P}$  of pricing kernel components and returns, let equivalent measure  $\mathbb{T}$  be defined by the Radon-Nikodym derivative  $\frac{d\mathbb{T}}{d\mathbb{P}} := \frac{(M^T)^\gamma}{E[(M^T)^\gamma]}$  for some  $\gamma \in \mathbb{R}$ .<sup>22</sup> Then  $M^\mathbb{T} := M(M^T)^{-\gamma}E[(M^T)^\gamma]$  is a stochastic discount vector with respect to measure  $\mathbb{T}$  and the following bound is sharp for any  $\beta \in \mathbb{R}$ :*

$$\frac{\mathcal{K}_{M^\mathbb{T}}(\beta)}{\beta(\beta - 1)} \geq -\frac{\mathcal{K}_R^\mathbb{T}(-\beta/(1 - \beta))}{\beta}. \quad (65)$$

*This bound is equivalent to the bound*

$$\frac{\mathcal{K}_{M^T M^P}(\gamma + (1 - \gamma)\beta, \beta)}{\beta(\beta - 1)} \geq -\frac{\mathcal{K}_{M^T R}(\gamma, -\beta/(1 - \beta))}{\beta}, \quad (66)$$

*in the sense that the difference of the LHS and the RHS of inequalities (65) and (66) is identical.*

The parameter choice  $\gamma = \frac{\alpha - \beta}{1 - \beta}$  in Proposition 5 implies bound (64) for any  $\alpha, \beta \in \mathbb{R}$  such that  $\mathcal{K}_M((\alpha - \beta)/(1 - \beta))$  is well-defined. Note while bound (65) is equivalent to bound (66), it is also robust to the scale of  $M^T$ . In this sense, a violation of bound (65) by an asset pricing model has the

<sup>22</sup>This change of measure is well-defined if and only if the marginal CGF of  $M^T$  in  $\gamma$  is well-defined.



desirable property of being robust to an inappropriate model-implied scaling of  $M^T$ , deriving, e.g., from an inappropriate empirical measurement of long term real bond returns.

### 5.2.2 Chernoff (1952) Entropy Bounds on Domestic and Foreign Pricing Kernels

In the context of Section 3.2.2, inequality  $\min(x, y) \leq x^\alpha y^{1-\alpha}$  yields for any  $\alpha \in (0, 1)$  the following Chernoff (1952) information bound on the average minimal pricing kernel:<sup>23</sup>

$$E[\min(M_d/E[M_d], M_f/E[M_f])] \leq \exp(-\mathcal{CI}_*(M_d, M_f)) . \quad (67)$$

When markets are complete, the forward exchange rate return  $F_e = (M_f/E[M_f])/(M_d/E[M_d])$  implies:

$$E \left[ \frac{M_d}{E[M_d]} \max(0, 1 - F_e) \right] \geq 1 - \exp(-\mathcal{CI}_*(M_d, M_f)) \approx \mathcal{CI}_*(M_d, M_f) , \quad (68)$$

i.e., the (forward) price of an at-the-money put option on the (forward) exchange rate is a tight upper bound on the Chernoff information of domestic and foreign pricing kernels.<sup>24</sup>

### 5.2.3 Dispersion Constraints of Type (1) on Domestic and Foreign Pricing Kernels

In a  $d$ -country economy with pricing kernel components  $M = (M_1, \dots, M_d)$  pricing the gross returns  $R = (R_1, \dots, R_d)$ , the following pricing constraints hold for  $i = 1, \dots, d$ :

$$\mathcal{K}_{M_i R_i}(1, 1) = \mathcal{K}_{MR}([e'_i, e'_i]) = 0 . \quad (69)$$

Given strictly positive vector  $\alpha = (\alpha_1, \dots, \alpha_d)$  such that  $\|\alpha\|_1 := \sum_{i=1}^d \alpha_i < 1$ , we also have:

$$[\alpha, 0_d] = \left( 1 - \sum_{i=1}^d \alpha_i \right) \left[ 0_d, -\frac{\alpha}{1 - \sum_{i=1}^d \alpha_i} \right] + \alpha_i \sum_{i=1}^d [e_i, e_i] . \quad (70)$$

<sup>23</sup>Recall Definition (22) of Chernoff information. We make use of inequality  $\min(x, y) \leq x^\alpha y^{1-\alpha}$  for  $x, y \geq 0$  and  $\alpha \in (0, 1)$ .

<sup>24</sup>From the symmetry of Chernoff information, the same bound applies for the (forward) price of an at-the-money put option on the (forward) exchange return  $1/F_e$ .

Dispersion constraints of Type (1) then directly imply following multivariate pricing kernel bound:

$$\frac{\mathcal{K}_M(\alpha)}{(\sum_{i=1}^d \alpha_i - 1) \prod_{i=1}^d \alpha_i} \geq - \frac{\mathcal{K}_R\left(\alpha / (\sum_{i=1}^d \alpha_i - 1)\right)}{\prod_{i=1}^d \alpha_i}. \quad (71)$$

This bound is a natural multivariate version of bound (51). Moreover, it can be optimally sharpened in economies with multiple domestic and foreign returns. Precisely, let pricing kernel  $M_i$  price returns  $R_{i0}, \dots, R_{iN_i}$  in market  $i = 1, \dots, d$ , where  $R_{i0}$  is the  $i$ -th risk-free return. Then, the optimization of lower bound (71) over portfolios of these returns provides in Proposition 6 the sharpest bound.<sup>25</sup>

**Proposition 6** (Multivariate Bound Tightness). *Consider for  $\alpha \in \mathbb{R}_{++}^d$  such that  $\|\alpha\|_1 < 1$  the pricing kernel vector  $M^* = (M_1^*, \dots, M_d^*)$  that solves the following minimum divergence problem.<sup>26</sup>*

$$\inf_M \left\{ \frac{\mathcal{K}_M(\alpha)}{(\sum_{i=1}^d \alpha_i - 1) \prod_{i=1}^d \alpha_i} \right\}, \quad (72)$$

subject to the following moment conditions, indexed by  $i = 1, \dots, d$  and  $k_i = 0, \dots, N_i$ :

$$\mathcal{K}_{M_i R_{ik_i}}(1, 1) = 0. \quad (73)$$

Further, consider the solution  $\lambda^*$  of the maximization problem:

$$\sup_\lambda \left\{ - \frac{\mathcal{K}_{R_\lambda}(\alpha / (\sum_{i=1}^d \alpha_i - 1))}{\prod_{i=1}^d \alpha_i} \right\}, \quad (74)$$

where  $R_\lambda = (R_{1\lambda_1}, \dots, R_{d\lambda_d})$  and  $R_{i\lambda_i} = \sum_{k_i=1}^{N_i} \lambda_{ik} R_{ik} + (1 - \sum_{k_i=1}^{N_i} \lambda_{ik}) R_{i0}$  is the return of a portfolio of returns (with constraint  $R_{i\lambda_i} > 0$ ) denominated in the  $i$ -th domestic currency. It then follows:

$$\frac{\mathcal{K}_{M^*}(\alpha)}{(\sum_{i=1}^d \alpha_i - 1) \prod_{i=1}^d \alpha_i} = - \frac{\mathcal{K}_{R_{\lambda^*}}(\alpha / (1 - \sum_{i=1}^d \alpha_i))}{\prod_{i=1}^d \alpha_i}. \quad (75)$$

<sup>25</sup>The proof is collected in the Supplemental Appendix.

<sup>26</sup> $\mathbb{R}_{++}^d$  denotes the  $d$ -dimensional strictly positive cone.

The optimal pricing kernel  $M^* := (M_1^*, \dots, M_d^*)$  has components given explicitly by:

$$M_i^* = \frac{\left[ R_{i\lambda_i^*}^{1-\sum_{j \neq i} \alpha_j} \prod_{j \neq i}^d R_{j\lambda_j^*}^{\alpha_j} \right]^{1/(\sum_{j=1}^d \alpha_j - 1)}}{E \left[ \prod_{j=1}^d R_{j\lambda_j^*}^{\alpha_j} \right]} ; i = 1, \dots, d . \quad (76)$$

From Proposition 6, the tightest lower bound on  $\mathcal{K}_M(\alpha)/((\sum_{i=1}^d \alpha_i - 1) \prod_{i=1}^d \alpha_i)$ , which is compatible with a  $d$ -dimensional vector  $M$  of pricing kernels for the given sets of returns, is obtained from a single dispersion constraint of Type (1) applied to the vector of portfolio returns  $R_{\lambda^*} = (R_{1\lambda_1^*}, \dots, R_{d\lambda_d^*})$ . By construction, the tightness result in Proposition 6 also identifies in closed-form the upper arbitrage-free CGF on domain  $D := \{\alpha \in \mathbb{R}_{++}^d : \|\alpha\|_1 < 1\}$ :

$$\mathcal{K}_M(\alpha) \leq \mathcal{K}_M^U(\alpha) = \left( 1 - \sum_{i=1}^d \alpha_i \right) \mathcal{K}_{R_{\lambda^*}} \left( \alpha / (1 - \sum_{i=1}^d \alpha_i) \right) ; \alpha \in D . \quad (77)$$

This result also induce an obvious generalization of univariate dispersion bounds. For instance, a prior  $\pi_\alpha$  with mass  $\alpha_i$  on observable point  $[e'_i, 0'_d]$  ( $i = 1, \dots, d$ ) and mass  $1 - \sum_{i=1}^d \alpha_i$  on point  $0_{2d}$  implies the lower dispersion bound

$$\begin{aligned} \mathcal{D}_{\pi_\alpha}(M) &= \frac{E_{\pi_\alpha}[\mathcal{K}_M(m)] - \mathcal{K}_M(E_{\pi_\alpha}[m])}{\sum_{i=1}^d \alpha_i (1 - \alpha_i)} \\ &\geq \frac{-\sum_{i=1}^d \alpha_i \log R_{i0} + (\sum_{i=1}^d \alpha_i - 1) \mathcal{K}_{R_{\lambda^*}}(\alpha / (\sum_{i=1}^d \alpha_i - 1))}{\sum_{i=1}^d \alpha_i (1 - \alpha_i)} , \end{aligned} \quad (78)$$

which is a natural multivariate extension of the univariate entropy bound (52). This bound is computable from the marginal distribution of returns across different markets and it yields restrictions on both the marginal distribution of pricing kernel components and their joint dependence.

It is useful to recall that no assumption on market completeness has been made in Proposition 6, such as assumptions about the structure of the exchange rates between domestic and foreign markets. Obviously, the optimal bound in Proposition 6 is sharper whenever the set of returns  $R_{i1}, \dots, R_{iN_i}$  is wider in each market. Using exchange rate markets, the set of domestic asset returns is naturally

extended, by adding to each set of domestic returns the set of foreign returns converted in domestic currency with the corresponding exchange rate return.

### 5.3 Relation to Other Bounds in the Literature

[Bakshi and Chabi-Yo \(2012\)](#) investigate the variance of the permanent / transitory component stochastic discount factor. Clearly the variance of the permanent component of the SDF can be written as:

$$\text{Var}(M^P) = \mathbb{E}[(M^P)^2] - \mathbb{E}^2[M^P] = \exp(\kappa(2)) - 1 \quad (79)$$

which is just a monotone transformation of the divergence measure  $\mathcal{D}_H$  with  $H$  binomially distributed over  $\{0,2\}$ . Finding the optimal value for  $\mathcal{D}_H$  by optimizing over  $\lambda$  as in [Proposition 4](#) provides the entropy bound given in [Bakshi and Chabi-Yo \(2012\)](#). Similar conclusion holds for the transitory component as well.

[Bakshi and Chabi-Yo \(2014\)](#) introduce the “entropy” of  $M^2$  as:

$$L(M^2) = \log(\mathbb{E}(M^2)) - \mathbb{E}(\log(M^2)) = \frac{1}{4} \lim_{\alpha \rightarrow 0} \mathcal{D}_{H_\alpha} \quad (80)$$

where  $H_\alpha$  is a binomial distribution taking values in  $\{0, 2\}$  with probabilities  $(1-\alpha)$  and  $\alpha$  respectively. For each  $\alpha$  we can optimize over  $\lambda$  to achieve the maximal value of  $\mathcal{D}_{H_\alpha}$  given the lower bound [\(71\)](#). Taking the limit as  $\alpha \rightarrow 0$  results in the lower bound given in [Bakshi and Chabi-Yo \(2014\)](#).

## 6 Arbitrage Free Dispersion in Long Run Risk Models

We characterize the AFD of [Bansal and Yaron \(2004\)](#) long-run risk (LRR) model, based on the recent model estimation in [Bansal, Kiku, and Yaron \(2012\)](#), which incorporates temporal aggregation.

## 6.1 LRR Model

The LRR model is based on a representative agent with recursive preferences maximizing the life-time utility,

$$V_t = \left[ (1 - \delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left( E_t[V_{t+1}^{1-\gamma}] \right)^{1/\theta} \right]^{\frac{\theta}{1-\gamma}}, \quad (81)$$

where  $C_t$  is consumption at time  $t$ ,  $0 < \delta < 1$  the time preference rate,  $\gamma$  the parameter of relative risk aversion and  $\theta := \frac{1-\gamma}{1-\psi}$ , with  $\psi$  is the elasticity of inter-temporal substitution (IES). Utility maximization is subject to the budget constraint  $W_{t+1} = (W_t - C_t)R_{c,t+1}$ , where  $R_{c,t+1}$  is the return on invested wealth, and consumption growth  $\Delta c_{t+1}$  satisfies the LRR dynamics:

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1}, \\ x_{t+1} &= \rho x_t + \psi_e \sigma_t e_{t+1}, \\ \sigma_{t+1}^2 &= \sigma_0^2 + \nu(\sigma_t^2 - \sigma_0^2) + \sigma_w w_{t+1}. \end{aligned} \quad (82)$$

with  $(\eta_{t+1}, e_{t+1}, w_{t+1}) \sim IIN(0, I_3)$ . The resulting (single-period) pricing kernel is

$$M_{t,t+1} = \delta^\theta (C_{t+1}/C_t)^{-\theta/\psi} R_{c,t+1}^{\theta-1}. \quad (83)$$

We borrow from the headline estimated specification with monthly aggregation intervals in [Bansal, Kiku, and Yaron \(2012\)](#), making use of the parameter estimates in the right-hand-side panel of Table II in their paper, and study the joint AFD properties of transient and persistent pricing kernel components in the LRR model.<sup>27</sup> [Bansal, Kiku, and Yaron](#) estimate the parameters in the LRR model using annual time series of real consumption, the stock market portfolio and real risk-free rates, in the sample period from 1930 to 2009.<sup>28</sup> In order to factorize the pricing kernel in the LRR model into

<sup>27</sup>This specification is not rejected by the overidentification tests in [Bansal, Kiku, and Yaron \(2012\)](#).

<sup>28</sup>Consumption represents per-capita real consumption expenditures on nondurables and services from NIPA tables. Aggregate stock market data consist of annual observations of returns, dividends, and prices of the CRSP value-weighted portfolio of all stocks traded on the NYSE, AMEX, and NASDAQ. The ex-ante real risk-free rate is constructed from a projection of the ex-post real rate on the current nominal yield and inflation over the previous year. Market data are converted to real using the consumer price index (CPI) from the BLS.

transient and persistent components, we follow [Alvarez and Jermann \(2005\)](#) and make use of proxies based on the return of long-maturity bonds. This data is available from CRSP's Fixed Term Indices dataset at a monthly frequency. Consistently with the aggregation procedure in [Bansal, Kiku, and Yaron \(2012\)](#), we aggregate these returns to an annual frequency and convert them to real returns using the CPI from the BLS.

## 6.2 Joint CGF of Transient and Persistent Pricing Kernel Components

We factorize the annual pricing kernel in the LRR model as  $M_{t+1} = M_{t+1}^T M_{t+1}^P$ , where permanent component  $M^P$  is such that  $E_t[M_{t+1}^P] = 1$ . Following [Alvarez and Jermann \(2005\)](#), we identify the transient component using the annual return on the infinite maturity bond:  $M_{t+1}^T = 1/R_{\infty,t+1}$ . We calculate the model-implied CGF of  $M = (M^T, M^P)$  by Monte Carlo simulation. Precisely, using monthly aggregation steps  $s = 1, \dots, 12$  we simulate  $N = 10^6$  annual paths of state dynamics (82) in the LRR model, based on parameter estimates from Table II in [Bansal, Kiku, and Yaron \(2012\)](#). Along each of the  $N$  annual simulated paths, we calculate on a monthly frequency the time series of single-period stochastic discount factors (83) and model-implied long-maturity bond prices, from which we obtain the annual pricing kernel  $M_{t+1}(u) = \prod_{s=1}^{12} M_{s,s+1}(u)$  and the annual long maturity bond returns  $R_{\infty,t+1}(u)$ ,  $u = 1, \dots, N$ .<sup>29</sup> Given the simulated distribution of  $M_{t+1}^T(u) = 1/R_{\infty,t+1}(u)$  and  $M_{t+1}(u)$  realizations, we calculate the simulated realization of  $M_{t+1}^P(u) = M_{t+1}(u)/M_{t+1}^T(u)$ . Finally, for powers  $(t, p)$  in domain  $D = (0, 1)^2$ , we compute the Monte Carlo CGF estimate as:

$$\mathcal{K}_M^{\mathbb{M}}(p, t) = \log E^{\mathbb{M}} \left[ (M_{t+1}^T)^t (M_{t+1}^P)^p \right] \approx \frac{1}{N} \sum_{u=1}^N (M_{t+1}^T(u))^t (M_{t+1}^P(u))^p, \quad (84)$$

where  $\mathbb{M}$  emphasizes the model-implied character of this CGF. The CGF  $\mathcal{K}_M^{\mathbb{M}}|_D$  in the LRR model is plotted in the left panel of Figure 9. It features a pronounced convexity along the  $p$ -axis and a much

<sup>29</sup>The yields of discount real bonds are affine functions of the state variables in the LRR model after a log-linearization. These affine functions can be calculated recursively as bond maturity increases. In this way, we obtain the price of the infinite maturity bond numerically for a sufficiently long maturity, avoiding to solve the eigenfunction problem implied by Perron-Frobenius theorem; see, e.g., [Bakshi and Chabi-Yo \(2014\)](#). We follow [Bansal, Kiku, and Yaron \(2012\)](#) and log-linearize around the mean value of the price-consumption ratio. This provides a fixed-point problem that is solved numerically.

flatter profile along the  $t$ -axis. Consistent with the intuition in, e.g, [Alvarez and Jermann \(2005\)](#), these convexity properties induce a significant dispersion in the permanent pricing kernel component and a much lower dispersion in the transient pricing kernel component of the LRR model.

### 6.3 Observable Set $\mathcal{O}_{\mathcal{K}_{MR}}$ for the LRR Model

Given a suitable risky portfolio return  $R_\lambda$ , we obtain the observable arbitrage-free CGF using following observable set  $\mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ .

**Assumption 1** (Observable Set). Set  $\mathcal{O}_{\mathcal{K}_{MR_\lambda}}$  is defined by the following observable points:

- (1) Restriction at the origin:  $(0, 0, 0) \in \mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ .
- (2) Martingale normalization:  $(0, 1, 0) \in \mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ .
- (3) Pricing of short-term bond:  $(1, 1, 0) \in \mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ .
- (4) Pricing of risky return:  $(1, 1, 1) \in \mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ .
- (5) Statistically observable risky and long-horizon bond returns:  $(t, 0, r) \in \mathcal{O}_{\mathcal{K}_{MR_\lambda}}$  for  $(t, r) \in \mathbb{R}^2$ .

Figure 2 illustrates the observable set  $\mathcal{O}_{\mathcal{K}_{MR_\lambda}}$  implied by assumptions (1)-(5). The vertical red plane in  $(m^T, r)$  coordinates is generated by observability assumption (5). This assumption follows from the CGF condition:

$$\mathcal{K}_{MR_\lambda}(t, 0, r) = \log E[(M^T)^t R_\lambda^r] = \log E[R_\infty^{-t} R_\lambda^r], \quad (85)$$

which reflects [Alvarez and Jermann \(2005\)](#) identification  $M^T = 1/R_\infty$ . The remaining points in the graph, highlighted with purple circles, correspond to assumptions (1)-(4). To generate risky portfolio return  $R_\lambda$ , we consider distinct sets of benchmark assets. Return  $R_{\lambda_A} := R_0 + \lambda_A(R_M - R_0)$  (Set A = Mkt + Bond) is the return of a portfolio invested in the short-term zero bond and the aggregate equity market. Return  $R_{\lambda_B} := R_0 + \lambda_{B1}(R_1 - R_0) + \lambda_{B2}(R_2 - R_0)$  is the return of a portfolio invested in the short-term zero bond and two size-sorted portfolios with book-to-market

ratio in the top 50% quantile of the CRSP universe (Set B = S-G + L-G + Bond).<sup>30</sup> Finally, return  $R_{\lambda_C} := R_0 + \sum_{i,j=1}^2 \lambda_{Cij}(R_{ij} - R_0)$  is the return of a portfolio invested in the short-term bond and four double-sorted portfolios with respect to size and book-to-market (Set C = S-G + L-G + S-V + L-V + Bond).<sup>31</sup> We focus on the restrictions implied by Assumption 1 for the marginal arbitrage-free CGF of pricing kernel vector  $M = (M^T, M^P)$  on domain  $D = (0, 1)^2$ . As  $D \subset \overline{\mathcal{O}}_{\mathcal{K}_{MR_\lambda}}$ , the upper arbitrage-free CGF  $\mathcal{K}_{MR_\lambda}^U$  is finite on  $D$  and provides a useful observable upper bound for  $\mathcal{K}_M$  on this domain. In order to obtain the sharpest upper bound  $\mathcal{K}_{MR_{\lambda^*}}^U = \inf_{\lambda} \mathcal{K}_{MR_\lambda}^U$  in our tests of the LRR model, we optimize the portfolio composition  $\lambda$ , with respect to weights  $\lambda_u$  corresponding to benchmark investment sets  $u = A$ ,  $u = B$  and  $u = C$ , respectively.

#### 6.4 Omibus Diagnostics Tests of Null Hypothesis $\mathcal{H}_0(\overline{\mathcal{O}}_{\mathcal{K}_{MR_{\lambda^*}}})$

Motivated by the diagnostics testing approach of Section 4.6.1, we first focus on dispersion constraints of Type (1) for the LRR model, over the domain  $D \subset \overline{\mathcal{O}}_{\mathcal{K}_{MR_{\lambda^*}}}$ :

$$\mathcal{H}_0(\overline{D}) : \mathcal{K}_M^M |_D \leq \mathcal{K}_{MR_{\lambda^*}}^U |_D . \quad (86)$$

The estimated upper arbitrage free CGF  $\hat{\mathcal{K}}_{MR_{\lambda^*}}^U |_D$  for data sets A and C is presented in the middle and the right panel of Figure 9, respectively. Similar to the model-implied CGF, the upper CGFs imply a pronounced convexity along the  $p$ -axis and a flatter profile along the  $t$ -axis, supporting a dominating dispersion of permanent relative to transient pricing kernel components in the data. However, shapes and levels of the upper and model-implied CGFs are also different in a number of cases.

In order to rely on an accurate finite-sample inference on the LRR model, we develop a suitable bootstrap procedure for estimating bootstrap confidence intervals about point estimate  $\hat{\mathcal{K}}_{MR_{\lambda^*}}^U(p, t, 0)$  for each  $(p, t) \in D$ .<sup>32</sup> Conservative bootstrap  $p$ -values for the test of null hypothesis (86) are presented

<sup>30</sup>For brevity, S (L) stays for **S**mall (**L**arge) stocks and G (V) for **G**rowth (**V**alue) stocks.

<sup>31</sup>We obtain these stock return data from Kenneth French's website. [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)

<sup>32</sup>Details on this bootstrap procedure are provided in the Supplemental Appendix, available from the authors upon request.



in Figure 10. Using dataset A, we obtain no significant violation of this null hypothesis over the vast majority of domain  $D$ . In contrast, using datasets  $B$  and  $C$ , we progressively obtain wider regions of rejection of null hypothesis (86), also in the interior of domain  $D$ , at standard significance levels. While the violations highlighted by investment sets B and C indicate a possible misspecification of the joint CGF of  $(M^T, M^P)$  in the LRR model, it is useful to recall that null hypothesis (86) is not robust with respect to the model-implied scale of  $M^T = 1/R_\infty$ . Therefore, a violation of this hypothesis is not directly interpretable as a (scale-invariant) dispersion violation or as an observable scale discrepancy between the model-implied and the arbitrage-free CGF. Similarly, the non rejection of null hypothesis (86) using investment set A could be the consequence of a low power of the omnibus test in the joint presence of a dispersion violation and an observable scaling discrepancy. Consistently with the concepts developed in Section 4, we address the separation of model violations due to inappropriate dispersion from those due to a scaling discrepancy in the next sections.

### 6.5 Diagnostics Test of a Scaling Discrepancy in $M^T = 1/R_\infty$

Due to the normalization of the permanent pricing kernel component in the LRR model, a potential model-implied scaling discrepancy can only arise from an inappropriate scale of  $M^T = 1/R_\infty$ . Empirically, a scaling discrepancy in  $M^T$  can follow from the fact that in the data there is no exact analogue of the long maturity bond return  $R_\infty$ . For instance, a natural empirical proxy for  $R_\infty$  is the return  $R_{LT}$  of a bond with the highest observed maturity. Thus, a discrepancy between scales  $E[1/R_\infty]$  and  $E[1/R_{LT}]$  can be motivated by a limitation of proxy  $1/R_{LT}$  for measuring the scale of  $1/R_\infty$ , rather than by a broader empirical violation of the LRR model.

We can address the properties of  $1/R_{LT}$  as an empirical proxy for  $1/R_\infty$ , by comparing in Figure 7 their marginal CGFs. We find that the marginal CGF of  $1/R_\infty$  in the LRR model (plotted in blue) is outside the 95% pointwise bootstrap confidence interval around the empirical marginal CGF of  $1/R_{LT}$  (plotted in black). However, a rescaled transient component  $\widetilde{M}_T := K/R_\infty$  ( $K > 0$ ) such that  $E^M[\widetilde{M}^T] = E[1/R_{LT}]$  produces a marginal model-implied CGF well-inside the bootstrap confidence

intervals.<sup>33</sup> Therefore, while a violation of Type (1) due to scaling can be explained by the limitations of proxy  $1/R_{LT}$  for measuring the scale of  $1/R_\infty$ , a violation due to inappropriate dispersion cannot be explained by the limitations of proxy  $1/R_{LT}$  for measuring the dispersion of  $1/R_\infty$ . Indeed, we can directly measure the dispersion of  $M_T = 1/R_\infty$ , e.g., using the generalized entropy  $\mathcal{E}_\alpha(M^T)$  in Section 3.2.1, based on a Bernoulli prior  $\pi_\alpha$  such that  $\pi_\alpha(1, 0, 0) = \alpha \in (0, 1)$  and  $\pi_\alpha(0, 0, 0) = 1 - \alpha$ :

$$\mathcal{D}_{\pi_\alpha}(M^T) = \frac{E_{\pi_\alpha}[\mathcal{K}_{M^T}(t)] - \mathcal{K}_{M^T}(E_{\pi_\alpha}[t])}{\alpha(1-\alpha)} = \frac{\alpha\mathcal{K}_{M^T}(1) - \mathcal{K}_{M^T}(\alpha)}{\alpha(1-\alpha)} = \mathcal{E}_\alpha(M^T), \quad (87)$$

which is observable under Assumption 1 (5). Figure 6 collects the results of a diagnostics test of null hypothesis  $\mathcal{H}_0 : \mathcal{E}_\alpha(1/R_\infty) = \mathcal{E}_\alpha(1/R_{LT})$ , for parameters  $\alpha \in (0, 1)$ . We find that the model-implied dispersion of  $1/R_\infty$  is well inside the 95% bootstrap confidence interval of the estimated dispersion  $\widehat{\mathcal{E}}_\alpha(1/R_{LT})$ , confirming that the dispersion of  $1/R_\infty$  well reproduces the empirical dispersion of  $1/R_{LT}$ .

## 6.6 Diagnostics Tests of Observable Excess Dispersion

From Definition 6, an excess dispersion  $\Delta\mathcal{J}_{\pi_1, \pi_2}$  is observable when two priors with identical mean and support in  $\mathcal{O}_{\mathcal{K}_{MR}}$  exist. In the LRR model, we obtain a family of observable excess dispersions for bivariate vector  $M = (M^T, M^P)$ , using priors  $\pi_1, \pi_2$  such that  $\pi_1$  has support in  $\{(1, 1), ((t-p)/(1-p), 0)\}$ ,  $\pi_2$  has support in  $\{(0, 1), (t/(1-p), 0)\}$ , and  $\pi_1(1, 1, 0) = \pi_2(0, 1, 0) = p \in (0, 1)$ . The resulting observable excess dispersion is:

$$\Delta\mathcal{J}_{\pi_1, \pi_2} = p \log B - (1-p) \log \left( E \left[ R_\infty^{-\frac{t}{1-p}} \right] / E \left[ R_\infty^{-\frac{t-p}{1-p}} \right] \right), \quad (88)$$

where we used Assumption 1 (5) to write  $M^T = 1/R_\infty$ . Similar to the observable covariance measure  $\text{cov}(M^T, M^P) = B - E[1/R_\infty]$  in Bakshi and Chabi-Yo (2014), excess dispersion (88) is not in general robust with respect to the scaling of  $1/R_\infty$ . However, for  $t = p = 1/2$  we obtain the scale invariant

<sup>33</sup>A rescaling from  $M^T$  to  $\widetilde{M}^T$  is equivalent to modifying the model-implied CGF by a linear function such that  $\mathcal{K}_{M^T}^M(1) = \mathcal{K}_{\widetilde{M}^T}^M(1)$ .

observable excess dispersion in Example 1:

$$\Delta\mathcal{J}_{\pi_1, \pi_2}(R_\infty) := \frac{1}{2} \log(\text{cov}(R_\infty/E[R_\infty], M^P) + 1) . \quad (89)$$

Table 1 presents the results of a direct test of null hypothesis  $\mathcal{H}_0(\Delta\mathcal{J}) : \Delta\mathcal{J}_{\pi_1, \pi_2}(R_{LT}) = \Delta\mathcal{J}_{\pi_1, \pi_2}^M(R_\infty)$  based on excess dispersion (89). We find that the model-implied excess dispersion is negative, while its point estimate  $\widehat{\Delta\mathcal{J}}_{\pi_1, \pi_2}(R_{LT})$  in the data is positive. The difference is statistically significant, as the model-implied excess dispersions does not fit within the 95% bootstrap confidence interval about  $\widehat{\Delta\mathcal{J}}_{\pi_1, \pi_2}(R_{LT})$ . As excess dispersion (89) is robust with respect to the scale of  $M_T$ , such a violation is independent of the model's ability to fit average long-maturity bond returns and reflects a deeper failure of the LRR model in explaining the unconditional slope of the yield curve.

## 6.7 Diagnostics Tests of Marginal Lower Dispersion Bounds

As illustrated by Figure 8, for any point  $(t_0, p_0) \in D := (0, 1)^2$  we can obtain multiple dispersion constraints of Type (1), using distinct priors  $\pi$  with support in  $\mathcal{O}_{\mathcal{K}_M}$  and mean  $E_\pi[(t, p)] = (t_0, p_0)$ , which exclusively use information from the permanent and transient pricing kernel components. Given set  $\Pi(t_0, p_0)$  of such priors and  $\pi_0 \in \Pi(t_0, p_0)$ , following observable lower dispersion bound emerges:

$$\begin{aligned} \mathcal{D}_{\pi_0}(M) &= \frac{E_{\pi_0}[\mathcal{K}_M(t, p)] - \mathcal{K}(E_{\pi_0}[(t, p)])}{\text{tr}(\text{Var}_{\pi_0}(t, p))} \\ &\geq \frac{E_{\pi_0}[\mathcal{K}_M(t, p)] - \inf_{\pi \in \Pi(t_0, p_0)} E_\pi[\mathcal{K}_M(t, p)]}{\text{tr}(\text{Var}_{\pi_0}(t, p))} =: \mathcal{D}_{\pi_0, \Pi(t_0, p_0)}^L(M) > 0 , \end{aligned} \quad (90)$$

where the last strict inequality holds if the infimum is not attained in  $\pi_0$ .

Consistently with the graphical description in Figure 8, we construct set  $\Pi(t_0, p_0)$  by including the following priors. First, we include prior  $\pi_1$  with support in  $\{(1, 1), ((t_0 - p_0)/(1 - p_0), 0)\}$  and prior  $\pi_2$  with support in  $\{(0, 1), (t_0/(1 - p_0), 0)\}$  such that  $\pi_1(1, 1) = \pi_1(0, 1) = p_0$  (see Figure 8(a)). Second, for any  $0 < t_0 \leq p_0 < 1$  (any  $1 > t_0 > p_0 > 1$ ) we include priors  $\pi_3$  with support in  $\{(0, 0), (0, 1), (1, 1)\}$  (in  $\{(0, 0), (1, 0), (1, 1)\}$ ) such that  $\pi_3(0, 0) = (1 - t_0)$  and  $\pi_3(0, 1) = t_0 - p_0$  ( $\pi_1(0, 0) = (1 - t_0)$  and

$\pi_1(1, 0) = t_0 - p_0$ ); see Figure 8(b) (Figure 8(c)).

The left and the middle panels of Figure 11 plot for any  $(t_0, p_0) \in D$  the estimated lower bound  $\widehat{\mathcal{D}}_{\pi_0, \Pi(t_0, p_0)}^L(M)$  and the model-implied dispersion  $\mathcal{D}_{\pi_0}^M(M)$ , respectively, where prior  $\pi_0 = \arg \sup_{\pi \in \Pi(t_0, p_0)} E_\pi[\mathcal{K}_M(t, p)]$  is selected to ensure a non trivial lower dispersion bound (90). We find that typically the estimated lower bound is much lower than the model-implied dispersion, with bootstrap  $p$ -values for the test of null hypothesis  $\mathcal{D}_{\pi_0}^M(M) \geq \mathcal{D}_{\pi_0, \Pi(t_0, p_0)}^L(M)$ , in the right panel of Figure 11, which never produce a rejection even at large significance levels. This evidence shows that the information in the marginal distribution of pricing kernel components, which is generated by the return of short-term and long-maturity bonds, is insufficient to reject the CGF specification of the LRR model in unobservable parts of its domain.

## 6.8 Diagnostics Tests of Joint Lower Dispersion Bounds

Marginal lower dispersion bound (90) can be substantially improved, based on the information generated by traded portfolio return  $R_\lambda$ , giving rise to joint lower dispersion bounds. Indeed, for  $(t_0, p_0) \in D$  we can obtain additional multiple dispersion constraints of Type (1), using the joint arbitrage-free CGF of pricing kernels and returns and a prior  $\pi$  with support in  $\mathcal{O}_{\mathcal{K}_{MR_\lambda}}$  such that  $(t_0, p_0, 0) = E_\pi[(m, r)]$ .

For instance, using prior  $\pi_1$  in the previous section, we obtain:

$$\begin{aligned} \mathcal{D}_{\pi_1}(M) &= \frac{E_{\pi_1}[\mathcal{K}_M(t, p)] - \mathcal{K}_M(t_0, p_0)}{\text{tr}(\text{Var}_{\pi_1}(t, p))} \\ &\geq \frac{E_{\pi_1}[\mathcal{K}_M(t, p)] - \mathcal{K}_{MR_\lambda}^U(t_0, p_0, 0)}{\text{tr}(\text{Var}_{\pi_1}(t, p))} =: \mathcal{D}_{\pi_1}^L(M, R_\lambda) > 0, \end{aligned} \quad (91)$$

where  $\mathcal{D}_{\pi_1}^L(M, R_\lambda)$  is observable empirically. By construction,  $\mathcal{K}_{MR_\lambda}^U(t_0, p_0, 0)$  is finite and has the following explicit expression from Proposition 5: <sup>34</sup>

$$\begin{aligned}\mathcal{K}_{MR_\lambda}^U(t_0, p_0, 0) &= (1 - p_0)\mathcal{K}_{M^T R_\lambda}\left(\frac{t_0 - p_0}{1 - p_0}, -\frac{p_0}{1 - p_0}\right) \\ &= (1 - p_0)\mathcal{K}_{R_\infty R_\lambda}\left(\frac{p_0 - t_0}{1 - p_0}, -\frac{p_0}{1 - p_0}\right).\end{aligned}\quad (92)$$

By optimizing portfolio weight vector  $\lambda$  as in Proposition 4, we obtain the sharpest dispersion bound:

$$\mathcal{D}_{\pi_1}(M) \geq \mathcal{D}_{\pi_1}^L(M, R_{\lambda^*}), \quad (93)$$

where  $\lambda^* = \arg \inf_\lambda \mathcal{K}_{MR_\lambda}^U(t_0, p_0, 0)$ . This bound is also robust with respect to the scale of  $R_\infty$ , whenever portfolio return  $R_{\lambda^*}$  does not depend on a position in the long-maturity bond, giving rise to the scale-independent null hypothesis: <sup>35</sup>

$$\mathcal{H}_0 : \mathcal{D}_{\pi_1}^M(M) \geq \mathcal{D}_{\pi_1}^L(M, R_{\lambda^*}). \quad (94)$$

Figure 12 summarizes the test results of null hypothesis (94), using investment sets A, B and C and for  $(t_0, p_0) \in D$ . For each dataset, we obtain a substantial variability of the test  $p$ -values over domain  $D$ . The test results are especially interesting in the left panel of Figure 12(a), which corresponds to the very simple investment set A, consisting of the short-term risk-free bond and the aggregate equity market. In this panel, we obtain a large number of rejections of the lower dispersion bound (94) at standard significance levels below 5%, basically for all points above the main diagonal in  $(t, p)$  coordinates. In this way, we reject the LRR model by using portfolios of asset returns that were used to estimate the model and by additionally incorporating only the information generated by the

<sup>34</sup> Finiteness of  $\mathcal{K}_{MR_\lambda}^U(t_0, p_0, 0)$  follows from the fact that  $(t_0, p_0, 0)$  is in the convex hull of set  $\{(t, 0, r) : t, r \in \mathbb{R}\} \cup \{(1, 1, 1), (1, 1, 0)\} \subset \mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ . Figure 2 illustrates the observable points in set  $\mathcal{O}_{\mathcal{K}_{MR_\lambda}}$ .

<sup>35</sup> Scale invariance of the bound follows from the fact that  $\mathcal{K}_{R_\infty}\left(-\frac{t-p}{1-p}\right) - \mathcal{K}_{R_\infty R}\left(-\frac{t-p}{1-p}, -\frac{p}{1-p}\right)$  does not depend on the scale of  $R_\infty$ . Including the long-maturity bond in the investment set makes the lower dispersion bound sharper, but it obscures whether a potential dispersion violation is due to an inappropriate convexity or an inappropriate scaling of the model-implied CGF.

long-maturity bond return, in order to construct an observable proxy of the transient pricing kernel component. Note that a multivariate dispersion test, able to exploit the decomposition of the pricing kernel into transient and persistent components, is important in order to uncover these violations. Indeed, the  $p$ -value of the dispersion test for  $p, t \rightarrow 0$ , which corresponds to a univariate test of the standard entropy bound (52), does not imply a model rejection at the 5% significance level.

For investment set B, the rejection area expands to almost all of domain  $D$ , except regions with either high  $p$  and small  $t$  or high  $t$  and small  $p$ . A further expansion of the rejection region to virtually all of  $D$  arises for investment set C, when book-to-market-sorted returns are taken into account. This finding highlights a difficulty of the LRR model in explaining the dispersion properties of growth and value returns, which complements the model's ability to produce a "cross-section" of value premia close to the extant CAPM, as documented by [Bansal, Kiku, and Yaron \(2012\)](#). Finally, when return  $R_\lambda^*$  in null hypothesis (93) depends on the return of the long-maturity bond, we get in Figure 12(b) even stronger violations for all investment sets, which additionally reflect the weak properties of proxy  $1/R_{LT}$  for measuring the scale of  $1/R_\infty$ , shown in Section 6.5.

## 7 Conclusion

We introduce a general theory of arbitrage-free dispersion (AFD) that characterizes the testable properties of multivariate asset pricing models. We measure multivariate dispersion by a family of Jensen's gaps that directly reflect the convexity properties of the arbitrage-free cumulant generating function (CGF) of pricing kernels and returns. We show that the observable asset pricing restrictions and the statistical information on asset returns produce tight constraints on the arbitrage-free CGF and its AFD, which are helpful to test multivariate pricing kernel specifications with a unifying approach. While our approach naturally incorporates existing AFD constraints in the literature based on univariate pricing kernel bounds, we show that it naturally extends to general multivariate pricing kernel specifications, incorporating, e.g., transient and persistent pricing kernel components, domestic and

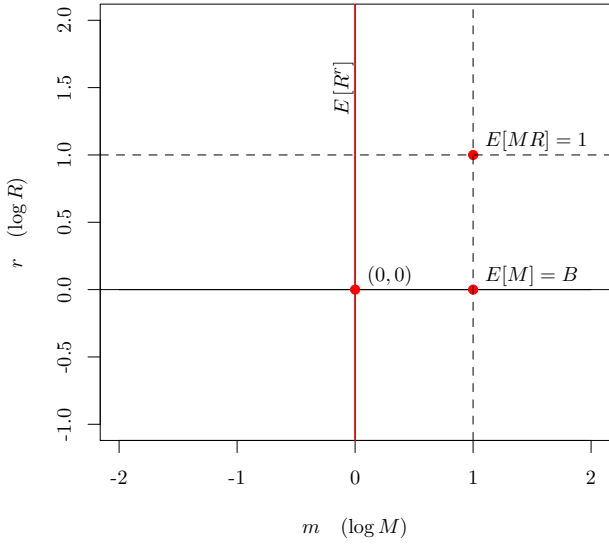
foreign state prices or horizon dependence.<sup>36</sup> For general multivariate specifications, we systematically develop a wide family of testable AFD constraints, for which we derive closed-form expressions and optimality properties in a number of concrete model settings. Using a recent estimation of the [Bansal and Yaron \(2004\)](#) model in the literature, we empirically test the multivariate AFD properties of Long Run Risk models, focusing on the joint model-implied distribution of permanent and transient pricing kernel components. We find that while the arbitrage-free and the model-implied CGF both imply a dominating degree of dispersion associated with the permanent component, the joint model-implied dispersion implies a counterfactual dependence of short vs. long-maturity bond returns and is insufficient for pricing the return of a simple portfolio of short-term riskless bonds and market equity. Such dispersion violations are robust with respect to the quality of empirical proxies for long maturity bond returns and are sharper when pricing the return of optimal portfolios invested in double-sorted size and value stocks.

---

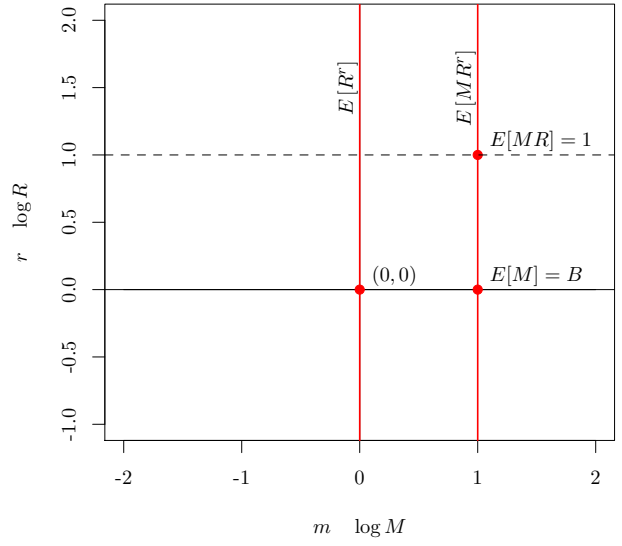
<sup>36</sup>Univariate pricing kernel bounds that are embedded in our AFD approach include [Bansal and Lehmann \(1997\)](#), [Alvarez and Jermann \(2005\)](#), [Bakshi and Chabi-Yo \(2012\)](#), [Liu \(2013\)](#), [Bakshi and Chabi-Yo \(2014\)](#) and [Backus, Chernov, and Zin \(2014\)](#), among others.

## Appendix I Figures





(a)  $\mathcal{O}_{\mathcal{K}_M R}$  with incomplete option markets.



(b)  $\mathcal{O}_{\mathcal{K}_M R}$  with complete option markets.

Figure 1: Observable sets in  $(M, R)$  space for  $d_1 = d_2 = 1$ . The red points and segments represent the tuples  $(m, r) \in \mathcal{O}_{\mathcal{K}_M R} \subset \mathbb{R}^2$  where the joint arbitrage-free CGF is observable, based on statistical return observations and asset pricing restrictions on the risk-free bond and the risky asset (panel (a)). In panel (b) the additional observable points on the  $m = 1$  line result from observing the prices of a continuum of options which can be used to replicate portfolios with final payoff  $R^r$ .

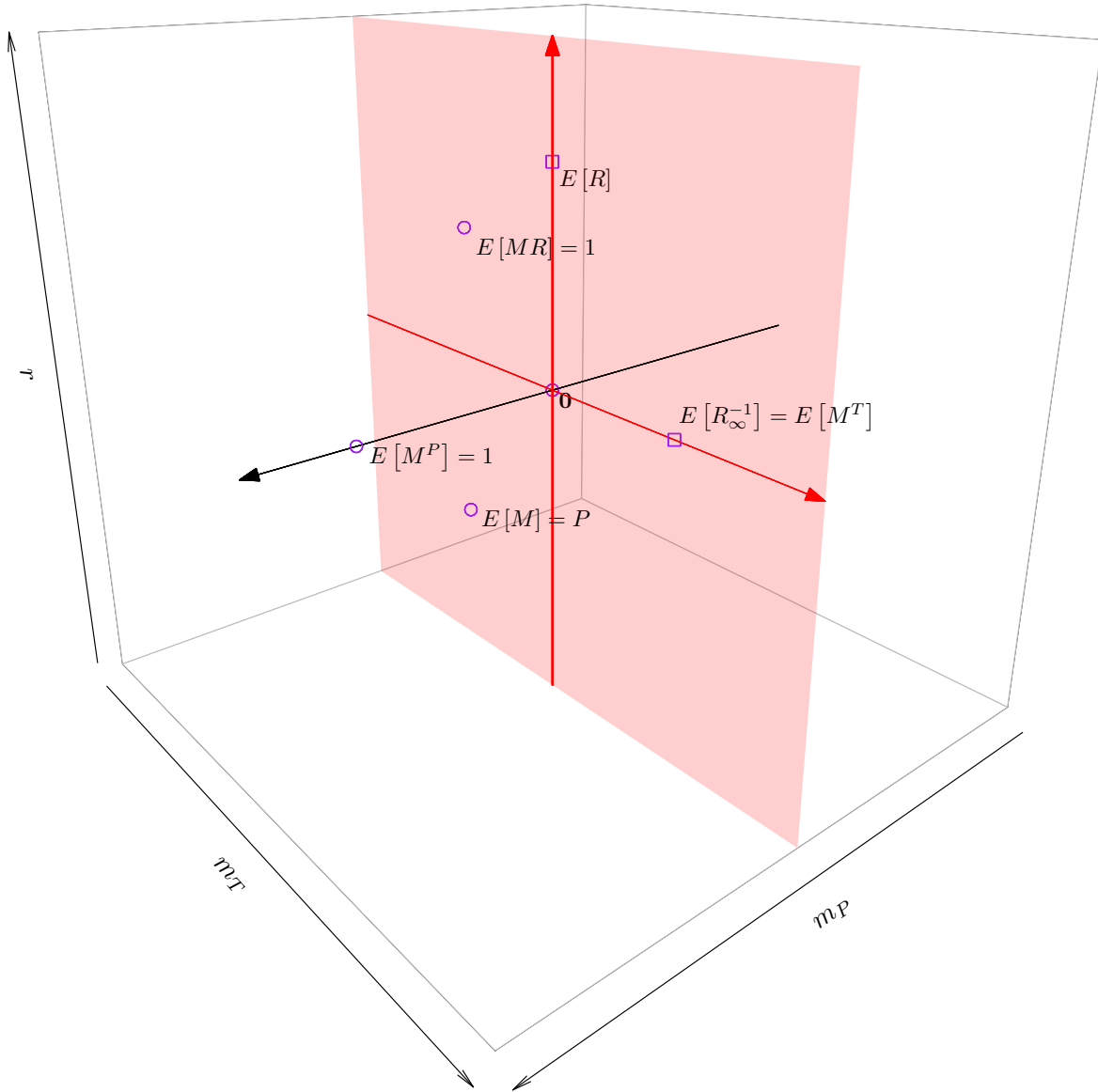
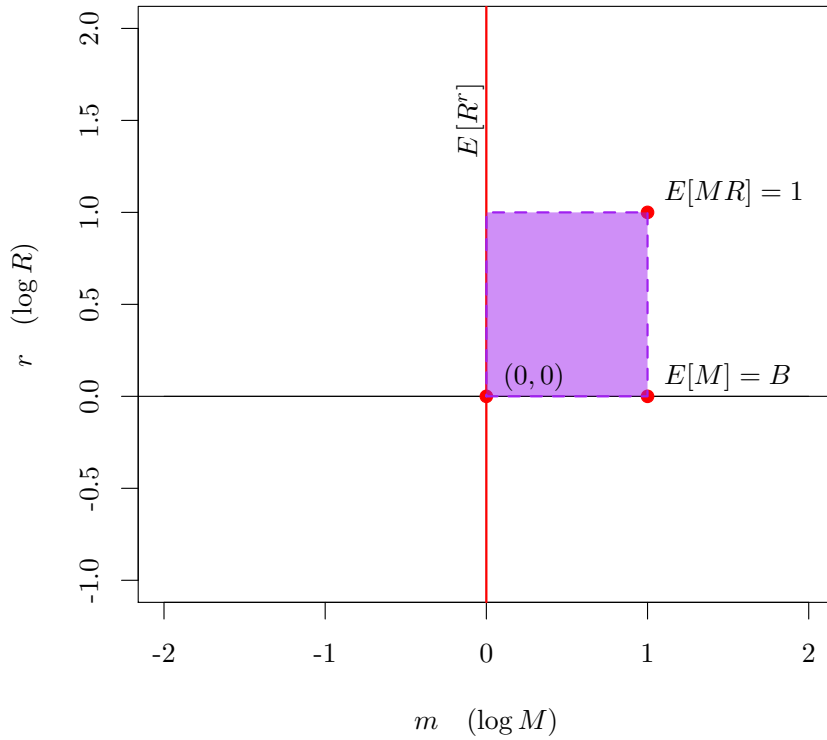
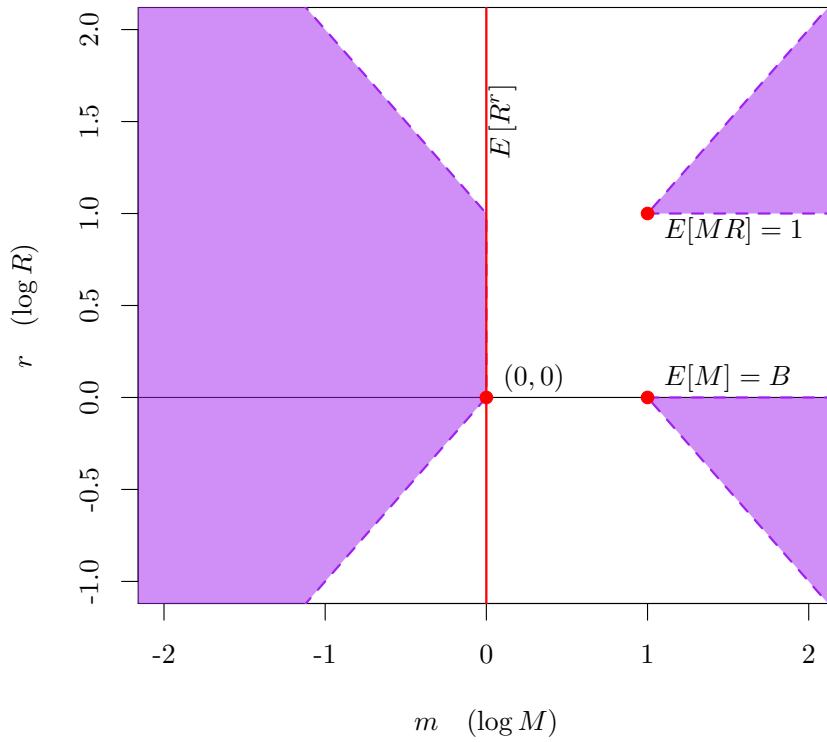


Figure 2: Observable points  $\mathcal{O}_{\mathcal{K}_{MR}}$  of the joint CGF in the single risky asset case with persistent and transient pricing kernel components. The red surface and the purple points represent the tuples  $(m, r) \in \mathcal{O}_{\mathcal{K}_{MR}} \subset \mathbb{R}^3$  where the joint CGF is observed based on statistical observations and asset pricing restrictions on the short-term risk-free bond, the long-term risk-free bond and the risky asset.



(a) Convex hull  $\bar{\mathcal{O}}$  in  $(M, R)$  space generated by the observable set  $\{(0, r) : 0 \leq r \leq 1\} \cup \{(1, 0), (1, 1)\}$ .  $\mathcal{K}^U$  is finite on this region.



(b) Set  $\mathcal{Q}$  in  $(M, R)$  space generated by the observable set  $\{(0, r) : 0 \leq r \leq 1\} \cup \{(1, 0), (1, 1)\}$ .  $\mathcal{K}^L$  is finite on this region.

Figure 3: Illustration of regions with finite  $\mathcal{K}^U$  and  $\mathcal{K}^L$  in a setting with a univariate pricing kernel  $M$  and a single priced return  $R$ , with an observed bond price  $B$ .

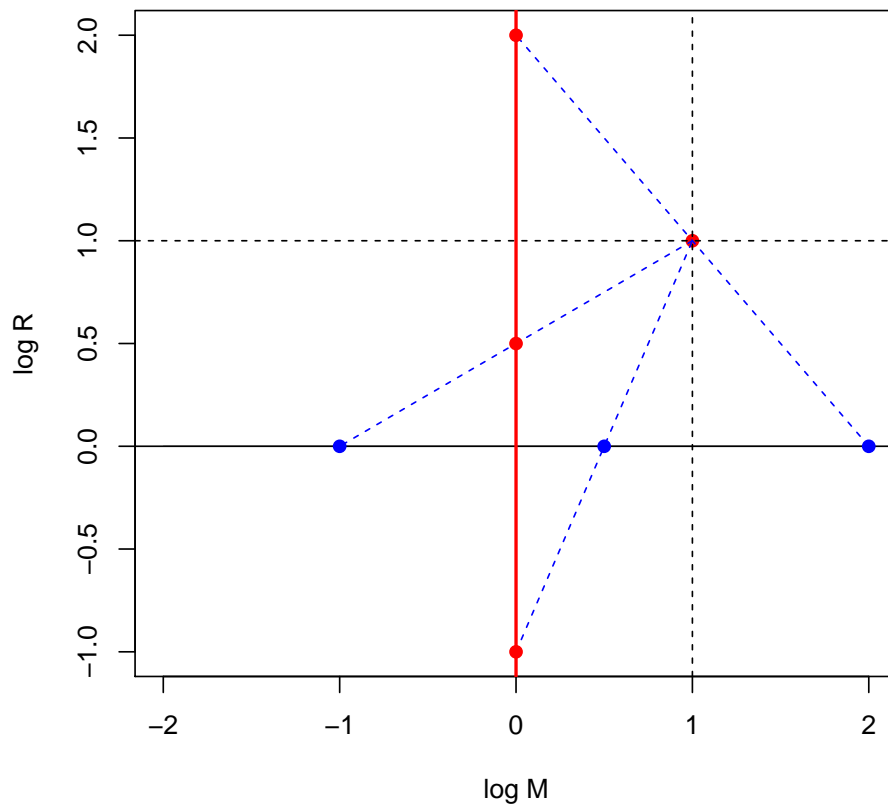
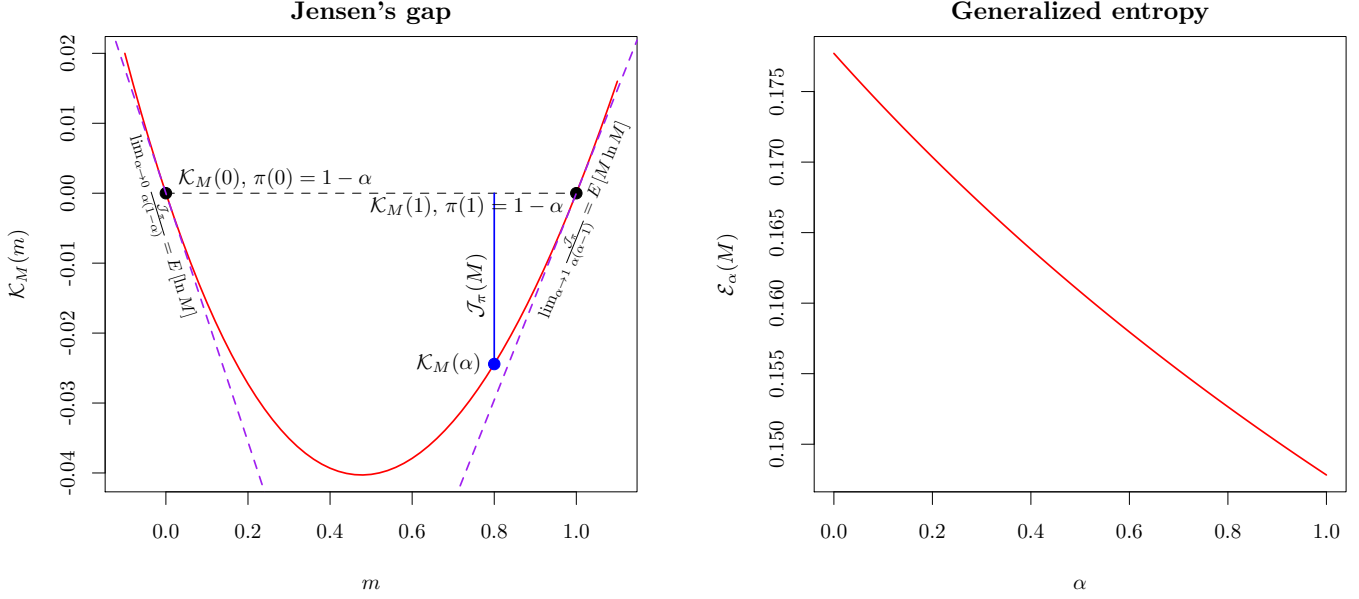


Figure 4: Formation of dispersion constraints of Type (1) and (2) in a setting with  $d_1 = d_2 = 1$ . A Type (1) constraint is used to bound the CGF value from above at unobservable point  $(1/2, 0)$  with the use of values at observable points  $(1, 1)$  and  $(0, -1)$ . Type (2) constraints are used to bound the value of the CGF from below at unobservable points  $(-2, 0)$  and  $(-1, 0)$ , with priors having support on points  $\{(1, 1), (0, 2)\}$  and  $\{(1, 1), (0, 1/2)\}$ , respectively.



(a) Illustration of Jensen's gap  $\mathcal{J}_\pi$  for stochastic discount factor (b) Dispersion  $\mathcal{D}_{\pi_\alpha}(M) = \mathcal{J}_{\pi_\alpha}/\alpha(1 - \alpha)$  calculated according to  $M$ .  $\mathcal{J}_\pi$  at  $\alpha$  is calculated with prior  $\pi$  such that  $\pi(0) = 1 - \alpha$  and equation (17).  $\pi(1) = \alpha$ . The vertical blue line at  $\alpha = 0.8$  is the gap. At  $\alpha = 0$  ( $\alpha = 1$ ) we take limits of the dispersion measure in panel 5(b) obtained by standardising  $\mathcal{J}_\pi$  by  $\alpha(1 - \alpha)$  and illustrate them as the negative slope (slope) of the CGF at 0 (1).

Figure 5: Illustration of Jensen's gap and generalized entropy of the SDF in equation (17). The SDF  $M$  is centered so that  $E[M] = 1$ . The standard deviation of  $M$  is set to 0.65.

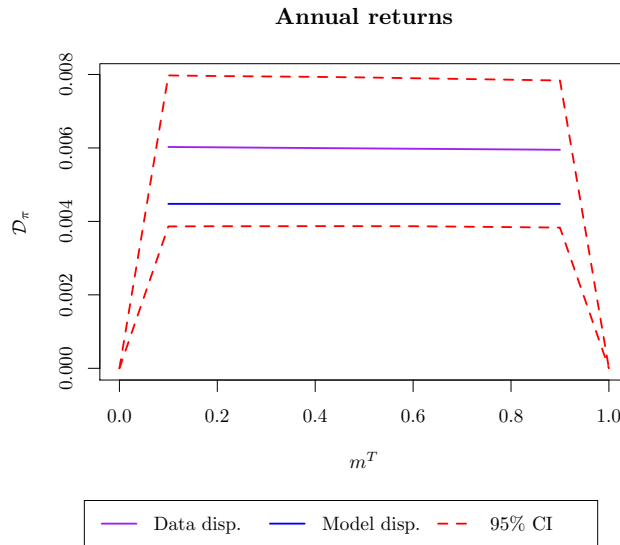


Figure 6: Observable dispersion of the transitory part of the SDF (87), calculated under the assumption that real returns on nominal long-maturity bonds are a proxy for real returns on infinite-maturity real bonds, which in turn are the inverse of the transitory component of the SDF, i.e.  $R_\infty = (M^T)^{-1}$ .

## Model long term bond CGF vs. bootstrapped bound

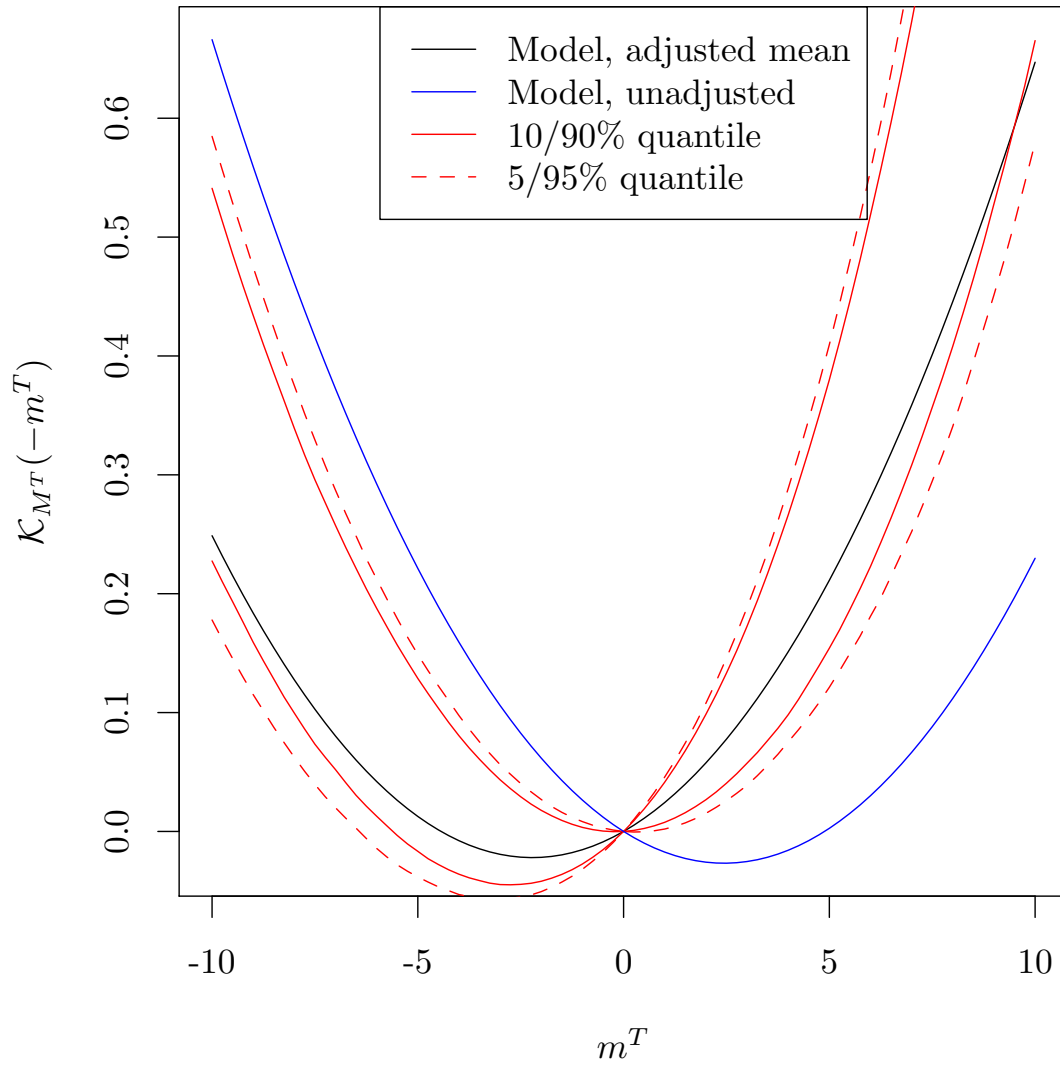
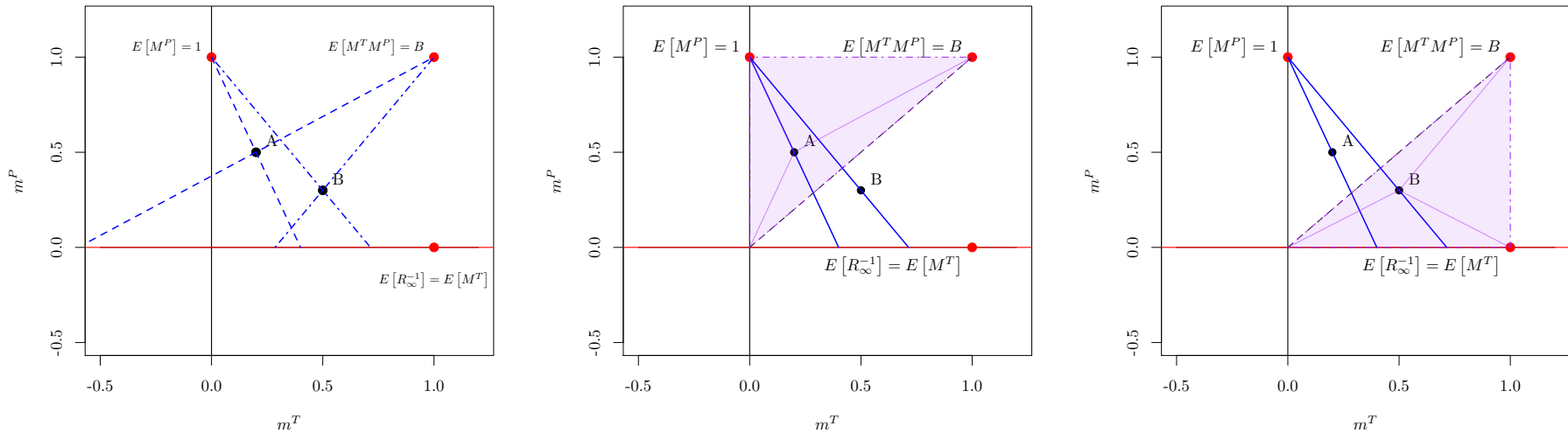


Figure 7: Model-based and estimated CGF confidence bounds of the transitory component of the stochastic discount factor, with  $M^T$  estimated with real returns on long-maturity bonds. Bootstrapped confidence bounds are denoted by red solid and dashed lines. The LRR model-implied CGF is given by the blue line. The black line depicts the LRR model implied CGF once the mean effective return is matched to the data mean (through a constant translation of logarithmic returns).



(a) Bounds in marginal CGF space – two-point based. (b) Bounds in marginal CGF space – above the diagonal. (c) Bounds in marginal CGF space – below the diagonal.

Figure 8: Dispersion bounds in the marginal CGF space. Red points and lines depict the observable set  $\mathcal{O}_{\mathcal{K}_{MR}}$ . Black points belong to the convex span of the observable set,  $\overline{\mathcal{O}}_{\mathcal{K}_{MR}}$ . In 8(a) the value of the CGF in each point is bounded in two ways: by taking the CGF value at (1, 1) (the log-bond price) or at (0, 1) (the martingale restriction), respectively, and the corresponding points on the  $m^T$  axis. In 8(b) and 8(c) the purple triangles provide two more ways of constructing Type (1) dispersion bounds. In order to construct  $\mathcal{K}_{MR}^U$  one has to pick the lowest available bound value.

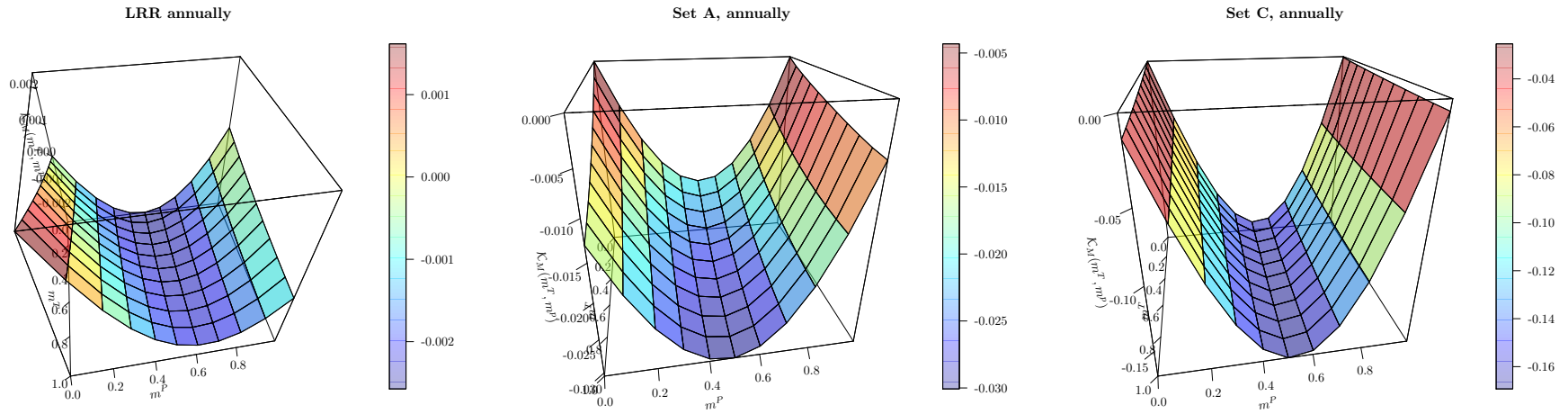


Figure 9: Model-implied marginal CGF of the SDF components (Bansal, Kiku, and Yaron (2012), leftmost panels), and estimates of  $\mathcal{K}_{M^T M^P}^U(m^T, m^P)$ ,  $(m^T, m^P) \in (0, 1) \times (0, 1)$ , obtained as in Proposition 4. The middle panel presents results where the portfolio of assets contains the market index, the single-period bond and the long-term bond (data set A). The rightmost panel presents results where the portfolio of assets contains, additionally, size- and value- sorted Fama-French portfolios (data set C). Data set description is available in Section 6.3.



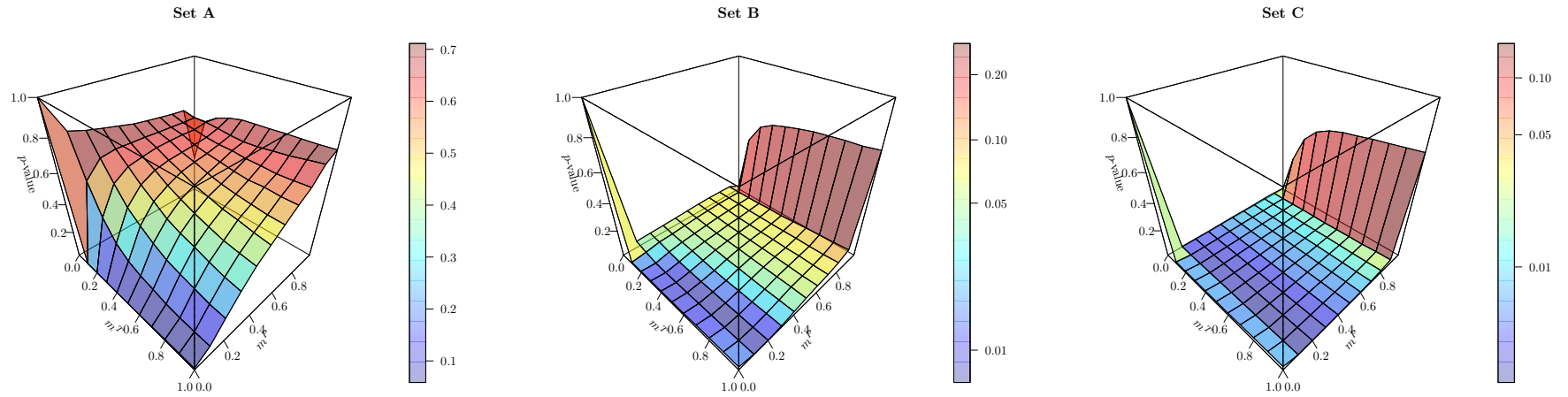


Figure 10: The omnibus test (see Section 6.4):  $p$ -values, for  $(m^T, m^P) \in (0, 1) \times (0, 1)$ , of the test whether the model-implied CGF evaluated at  $(m^T, m^P)$  attains lower values than the upper bound based on Proposition 4. The color scale on the four rightmost panels is logarithmic. Data set description is available in Section 6.3.

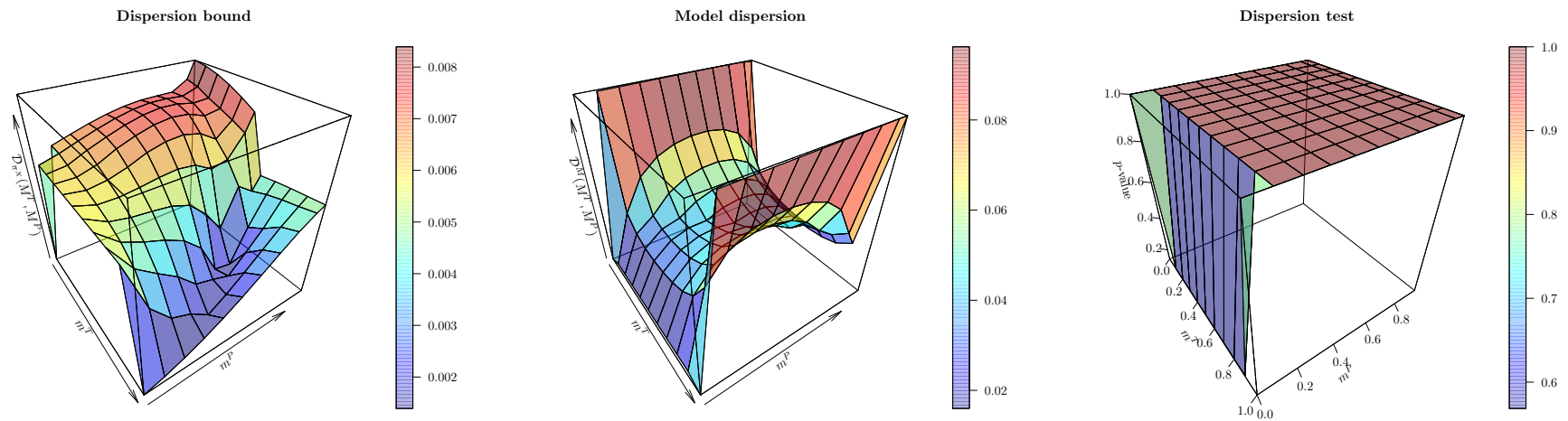
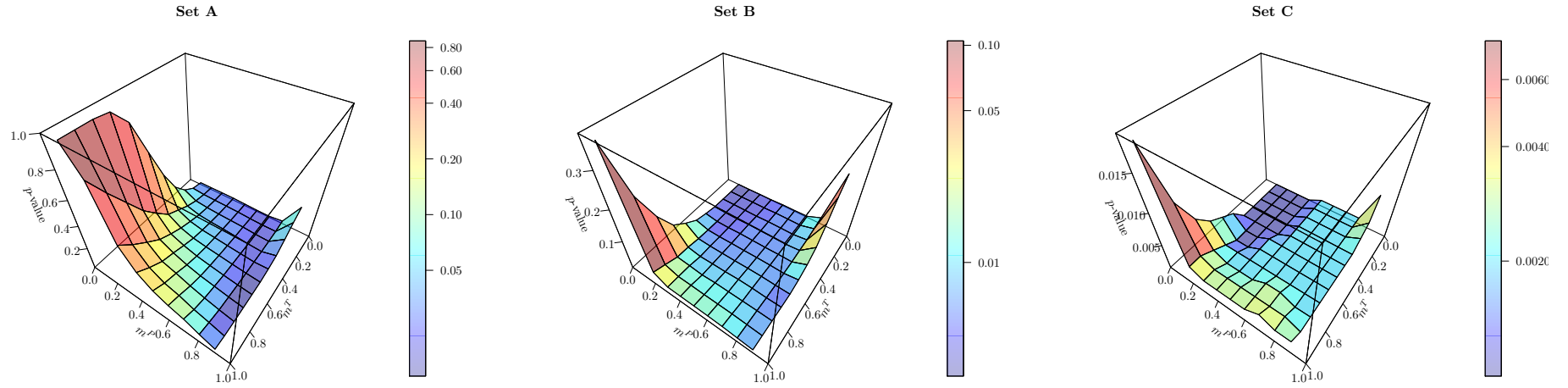
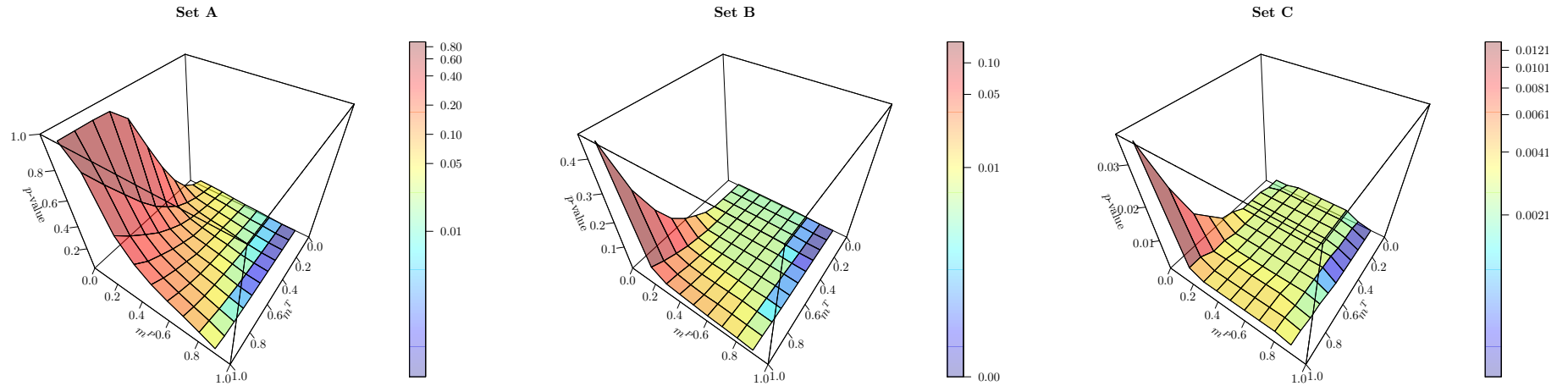


Figure 11: Dispersion test in the marginal SDF space (see Section 6.7):  $p$ -values, for  $(m^T, m^P) \in (0, 1) \times (0, 1)$ , of the test of the lower dispersion bound (90) using only observable SDF component information. Annual frequency. Data set description is available in Section 6.3.



(a) Dispersion test in  $(M, R)$  space for annual returns, without  $R_{LT}$  in the investor's portfolio;  $p$ -values.



(b) Dispersion test in  $(M, R)$  space for annual returns, with  $R_{LT}$  in the investor's portfolio;  $p$ -values.

Figure 12: Dispersion tests in  $(M, R)$  space (see Section 6.8):  $p$ -values. The null hypothesis (94) is tested for  $D_{\pi_{t,p}}$  as in (91) for  $(t, p) \in (0, 1) \times (0, 1)$ . Tests without  $R_{LT}$  in the investor's portfolio are immune to the mean level of  $\log R_{LT}$  in the data. Data set A considers a portfolio of the value-weighted stock index return and short-term bond. Data set B additionally takes size-sorted Fama-French portfolios. Data set C extends to size- and value-sorted Fama-French portfolios. Data set description is available in Section 6.3.

## Appendix II Tables

$1 - \alpha$	Annually
data	0.00302
model	-0.01017
95%	-0.007307 0.01371
99%	-0.01025 0.01703

Table 1: Excess dispersion in the [Bansal, Kiku, and Yaron \(2012\)](#) model and in the data. Model values calculated with the use of their best estimated model, whose parameters are reported in Table II of their paper. Data values calculated from sample ranging from 1946-03-30 to 2012-10-31. Confidence intervals calculated with the use of a time-series bootstrap (basic confidence interval type).

## References

- Almeida, C., and R. Garcia, 2011, “Robust economic implications of nonlinear pricing kernels,” .
- Alvarez, F., and U. J. Jermann, 2005, “Using asset prices to measure the persistence of the marginal utility of wealth,” *Econometrica*, 73(6), 1977–2016.
- Backus, D., N. Boyarchenko, and M. Chernov, 2014, “Term structures of asset prices and returns,” Work in progress.
- Backus, D., M. Chernov, and I. Martin, 2011, “Disasters Implied by Equity Index Options,” *The Journal of Finance*, 66(6), 1969–2012.
- Backus, D., M. Chernov, and S. Zin, 2014, “Sources of entropy in representative agent models,” *The Journal of Finance*, 69(1), 51–99.
- Bakshi, G., and F. Chabi-Yo, 2012, “Variance bounds on the permanent and transitory components of stochastic discount factors,” *Journal of Financial Economics*, 105(1), 191–208.
- , 2014, “New Entropy Restrictions and the Quest for Better Specified Asset Pricing Models,” *Fisher College of Business Working Paper*, (2014-03), 007.
- Bansal, R., D. Kiku, and A. Yaron, 2012, “Risks for the long run: Estimation with time aggregation,” working paper, National Bureau of Economic Research.
- Bansal, R., and B. N. Lehmann, 1997, “Growth-optimal portfolio restrictions on asset pricing models,” *Macroeconomic Dynamics*, 1(02), 333–354.
- Bansal, R., and Y. Yaron, 2004, “Risks for the long run: A potential resolution of asset pricing puzzles,” *Journal of Finance*, 59, 1481–1509.
- Carr, P., and D. Madan, 1998, “Towards a theory of volatility trading,” in R. Jarrow (ed) *Risk Book on Volatility*. New York: Risk, pp 417-27.
- Chabi-Yo, F., and R. Colacito, 2013, “The Term Structures of Co-Entropy in International Financial Markets,” *Charles A. Dice Center Working Paper*, (2013-17).
- Chernoff, H., 1952, “A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations,” *The Annals of Mathematical Statistics*, 23, 493–507.
- Dragomirescu, F., and C. Ivan, 1992, “The smallest convex extensions of a convex function,” *Optimization*, 24(3-4), 193–206.
- Jammalamadaka, S. R., T. Rao, and G. Terdik, 2006, “Higher Order Cumulants of Random VVector and Applications to Statistical Inference and Time Series,” *Sankhyā: The Indian Journal of Statistics (2003-2007)*, 68(2), 326–356.
- Liu, Y., 2013, “Index Option Returns and Generalized Entropy Bounds,” Available at SSRN 2149265.
- Peters, H., and P. Wakker, 1987, “Convex functions on non-convex domains,” *Economics letters*, 22(2-3), 251–255.

- Rényi, A., 1960, “On measures of information and entropy,” in *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*, pp. 547–561.
- Schneider, P., and F. Trojani, 2013, “Divergence and The Price of Uncertainty,” Working Paper.
- Snow, K. N., 1991, “Diagnosing asset pricing models using the distribution of asset returns,” *The Journal of Finance*, 46(3), 955–983.