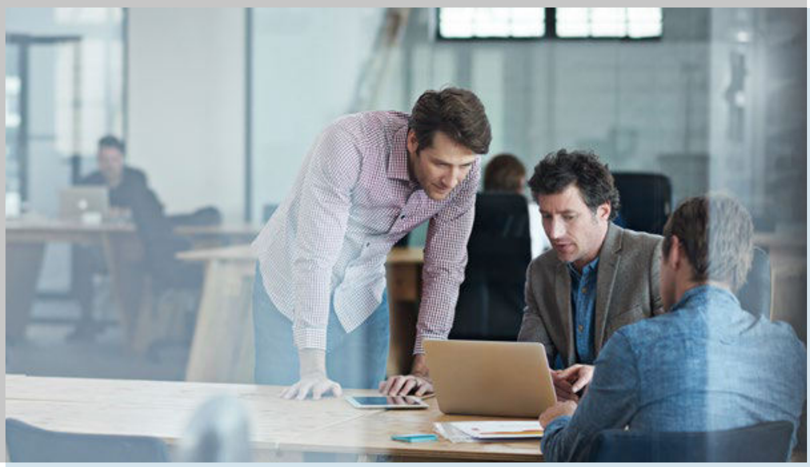
RESEARCH

# Swiss Finance Institute
# Research Paper Series
# N°19-20

## Arbitrage Free Dispersion



Piotr Orlowski
HEC Montreal

Andras Sali
Alphacruncher

Fabio Trojani
University of Geneva and Swiss Finance Institute

# Arbitrage Free Dispersion[*]

Piotr Orłowski
HEC Montréal[†]

András Sali
Alphacruncher AG[‡]

Fabio Trojani
Swiss Finance Institute
University of Geneva[§]

First version: December 2014
This version: August 10, 2018

**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

[*]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 partecipants 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.

[†]Piotr Orłowski, HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, H3T 2A7 Montréal, Canada. Email: piotr.orlowski@hec.ca

[‡]András Sali, Alphacruncher AG, Via Buffi 13, 6900 Lugano, Switzerland. Email: salia@alphacruncher.com

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

# 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

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

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 $\mathcal{O}_{\mathcal{K}_{MR}} := \{(m,r) \in dom(\mathcal{K}_{MR}) : m = 0 \text{ or } (m,r) = (1,1)\}$.

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.[1] 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,

[1]For instance, it constraints 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.

and Zin (2014), among others.[2]

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 contrains the arbitrage-free CGF whenever $E_\pi[\mathcal{K}_{MR}(m,r)]$ or $\mathcal{K}_{MR}(E_\pi[(m,r)])$ is observable.[3] 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 $\mathcal{K}^U_{MR}$ ($\mathcal{K}^L_{MR}$), 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:

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

As the left side of this inequality is observable, we obtain the following upper arbitrage-free CGF $\mathcal{K}^U_{MR}$ evaluated in $(m_\star, r_\star)$:[4]

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

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

[2]The bound follows by continuity, taking limits as $\alpha \to 1$ in the convex arbitrage-free inequality (2) .

[3]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 bound (2). We discuss in detail below the properties of $\mathcal{J}_\pi(M,R)$ as a measure of multivariate dispersion.

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

relations.

We make use of our theory of AFD to specify a coherent diagnostics approach for systematically 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 counterfactural 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 sharpeness 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 Appendix V characterize 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* $M := (M_1, \ldots, M_{d_1})$ *and a set of positive marketed gross returns* $R := (R_1, \ldots, R_{d_2})$*, the arbitrage-free cumulant generating function of* $(M, R)$ *is the function* $\mathcal{K}_{MR} : \mathbb{R}^{d_1+d_2} \to \overline{\mathbb{R}}$*, defined for any* $m := (m_1, \ldots, m_{d_1})$ *and* $r := (r_1, \ldots, r_{d_2})$ *by*

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

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

[5] Throughout the paper we assume that the joint cumulant generating function is finite in a non-degenerate open domain containing the origin.

### 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 constrains. 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 1a 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 (2015):

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

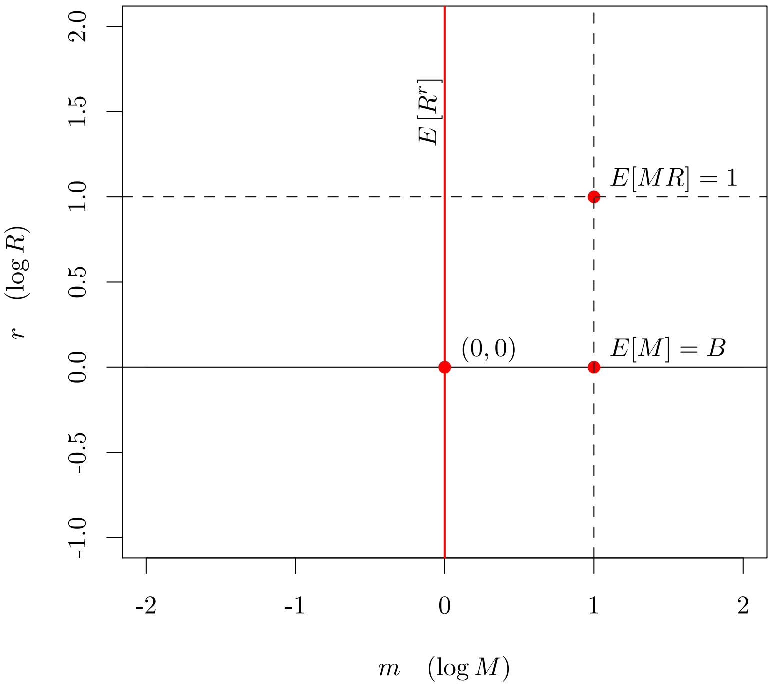
where $O(R,K)$ is the arbitrage-free price of an out-of-the-money option on $R$ with strike price $K > 0$. Figure 1b illustrates the enrichment of the (non convex) set of empirically observable CGF points generated by complete option markets, which now includes the vertical line with abscissa in $m = 1$.

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

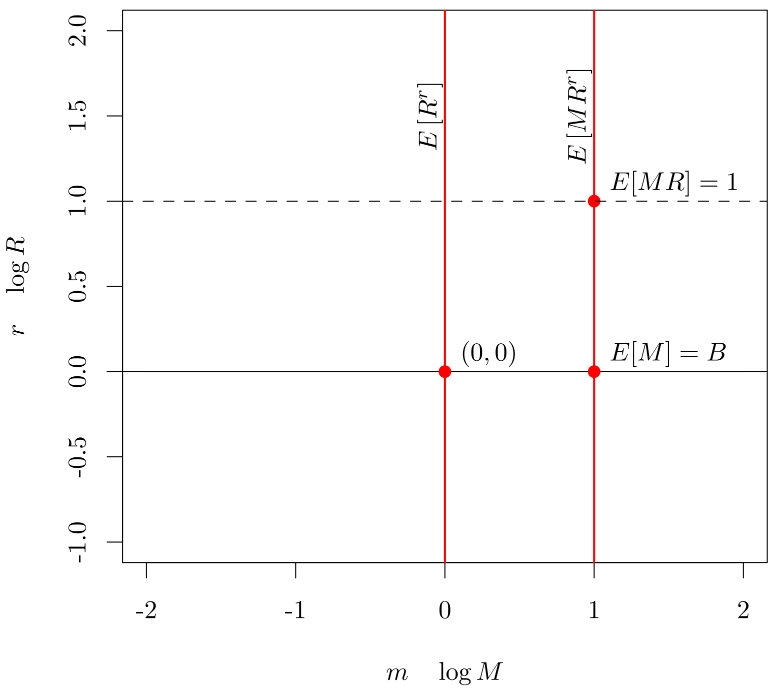
Whenever the long end of the yield curve is observable, the pricing kernel decomposition $M^P M^T$ into transient and persistent components produces additional observable constraints. Following Alvarez and Jermann (2005), $M^P$ does not affect the price of infinitely long maturity zero coupon bonds and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

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

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 $\mathcal{K}_{MR}(1,0,0) = \log E\left[R_\infty^{-1}\right]$, are marked as violet squares ($\square$).




(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 $(m, r) \in \mathcal{O}_{\mathcal{K}_{MR}} \subset \mathbb{R}^2$ 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 $R^r$.

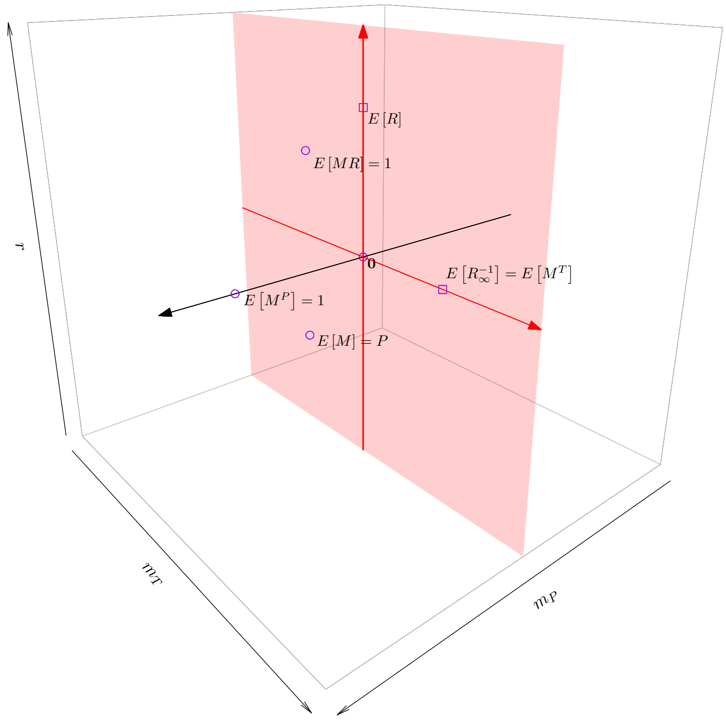



Figure 2: Observable points $\mathcal{O}_{\mathcal{K}_{MR}}$ of the joint CGF in the single risky asset case with persistent and transient pricing kernel components. The red surface and the purple points represent the tuples $(m, r) \in \mathcal{O}_{\mathcal{K}_{MR}} \subset \mathbb{R}^3$ where the joint 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.

### 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 $\{M_i : i = 1, \ldots, d\}$ ($\{R_i : i = 1, \ldots, d\}$) 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 $i = 1, \ldots, d$ are observed, then arbitrage-free CGF yields:

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

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,

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

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.[6]

### 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.

[6]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.

#### 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:*

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

*(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 $\mathcal{J}_\pi(R) := E_\pi[\mathcal{K}_R(r)] - \mathcal{K}_R(E_\pi[r])$.*

$\mathcal{J}_\pi(M,R)$ ($\mathcal{J}_\pi(M)$ 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:*

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

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

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

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

5. *$\mathcal{J}_\pi(M,R)$ is positively homogenous of degree 0, in the sense that for any $\lambda_M := (\lambda_{M1}, \ldots, \lambda_{Md_1})$*

*and* $\lambda_R := (\lambda_{M1}, \ldots, \lambda_{Md_2})$*:* $\mathcal{J}_\pi(\lambda_M \circ M, \lambda_R \circ R) = \mathcal{J}_\pi(M, R)$.[7]

6. *For any prior* $\pi$ *on* $\mathbb{R}^{d_1+d_2}$*, having prior covariance matrix* $Var_\pi(m, r)$*:*

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

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

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 measures 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*

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

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

[7] $\circ$ 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 $(A \circ B)_{ij} = A_{ij} B_{ij}$.

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

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

where $\mathcal{K}''_{MR}(\cdot,\cdot)$ is the Hessian of $\mathcal{K}$.[8] 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.[9] 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.

#### 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:

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

i.e., the $\alpha-$Rényi (1960) entropy of the stochastic discount factor. $\mathcal{E}_\alpha(M)$ is a standardized Jensen's gap of dimension $d_\pi = 1$. It's geometric interpretation based on dispersion measure $\mathcal{D}_{\pi_\alpha}(M,R)$ is illustrated in Figure 3.

As $\alpha \to 0$, we obtain $\mathcal{E}_0(M) = E[-\log(M/E(M))]$, i.e., the stochastic discount factor entropy in

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

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

with the nonzero eigenvalues $\lambda_{i\pi}$ and the corresponding orthonormal eigenvectors $q_{i\pi}$ of $Var_\pi(m,r)$, 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 $\mathcal{K}''_{MR}(E_\pi[m],E_\pi[r])$.

[9] 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.

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

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

which for $\alpha \to 0$ 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 $\mathcal{C}_\alpha(M,R) = Cov(\log M, \log R)$.

*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

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

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 $\alpha \to 0$ 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.[10]

*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:

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

$\mathcal{E}_{\alpha,\beta}(M,R)$ 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

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

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 $Cov(\log M, \log R)$

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

when returns and pricing kernel are jointly log normal, then naturally follows:

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

$\mathcal{C}_{\alpha,\beta}(M,R)$ 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.[11]

#### 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

$$\mathcal{CC}_{\alpha^*}(M_d,M_f) := \min_{\alpha\in(0,1)} \mathcal{CC}_{\alpha}(M_d,M_f) , \tag{22}$$

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

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

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.

[11]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.

## 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*

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

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

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

*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.[12] 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 mispecifications, 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.

[12] 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 depedence 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:

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

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 $M = (M_1, \ldots, M_n)$ 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

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

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:[13]

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

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 $\mathcal{K}^{\mathbb{M}}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}}$ 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 $\mathcal{K}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}} - \mathcal{K}^{\mathbb{M}}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}}$ is linear, i.e., there exists a vector $e \in \mathbb{R}^{d_1+d_2}$ such that $\mathcal{K}_{MR}(o) = \mathcal{K}^{\mathbb{M}}_{MR}(o) + e' \cdot o$ 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 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 $\mathcal{K}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}} - \mathcal{K}^{\mathbb{M}}_{MR}|_{\mathcal{O}_{\mathcal{K}_{MR}}}$ 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

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

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:

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

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 $(m_\star, r_\star) \in \overline{\mathcal{O}}_{\mathcal{K}_{MR}}$ the upper bound:

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

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* $\mathcal{K}^U_{MR}$ : $\mathbb{R}^{d_1+d_2} \to \mathbb{R} \cup \{+\infty\}$*, defined for any* $(m_\star, r_\star) \in \mathbb{R}^{d_1+d_2}$ *by:*[14]

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

*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 $\overline{\mathcal{O}}_{\mathcal{K}_{MR}}$.[15] By definition, $\mathcal{K}^U_{MR}$ 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:

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

Figure 4a 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 $\mathcal{K}_{MR}(0,r) =$

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

[15]See Peters and Wakker (1987), among other, for the properties of finite convex extensions of a convex function.

$\log E[R^r]$ 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 $(m_\star, r_\star) \in \overline{\mathcal{O}}_{\mathcal{K}_{MR}} \backslash \mathcal{O}_{\mathcal{K}_{MR}}$, an arbitrage-free dispersion violation of Type (1) arises whenever $\mathcal{K}^{\mathbb{M}}_{MR}(m_\star, r_\star) > \mathcal{K}^{U}_{MR}(m_\star, r_\star)$.*

### 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:

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

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:

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

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:

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

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 $\mathcal{K}^{L}_{MR} : \mathbb{R}^{d_1+d_2} \to \mathbb{R} \cup \{-\infty\}$, defined for any $(m_\star, r_\star) \in \mathbb{R}^{d_1+d_2}$ by:*[16]

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

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

*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, $\mathcal{K}^L_{MR}$ is observable from the set $\mathcal{O}_{\mathcal{K}_{MR}}$ of empirically observable points and implies a nontrivial inequality on the arbitrage-free CGF:

$$\mathcal{K}_{MR}(m_\star, r_\star) \geq \mathcal{K}^L_{MR}(m_\star, r_\star) \; . \tag{37}$$

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 $E_\pi[(m,r)] \in \mathcal{O}_{\mathcal{K}_{MR}}$[17]. The set of points where $\mathcal{K}^L_{MR}$ is finite follows from the following identity:

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

We denote by $\underline{\mathcal{O}}_{\mathcal{K}_{MR}} \subset \mathbb{R}^{d_1+d_2}$ the domain on which $\mathcal{K}^L_{MR}$ is finite. Figure 4b 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* $(m_\star, r_\star) \in \underline{\mathcal{O}}_{\mathcal{K}_{MR}} \backslash \mathcal{O}_{\mathcal{K}_{MR}}$*, an arbitrage-free dispersion violation of Type (2) arises whenever* $\mathcal{K}^{\mathbb{M}}_{MR}(m_\star, r_\star) < \mathcal{K}^L_{MR}(m_\star, r_\star)$.

### 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 directly follows from the definitions:

$$\mathcal{D}_\pi(M,R) \;\geq\; \left( \frac{E_\pi[\mathcal{K}^L_{MR}(m,r)] - \mathcal{K}^U_{MR}(E_\pi[(m,r)])}{tr(Var_\pi(m,r))} \right)^+ =: \mathcal{D}^L_\pi(M,R) \; , \tag{39}$$

$$\mathcal{D}_{\pi'}(M,R) \;\leq\; \frac{E_{\pi'}[\mathcal{K}^U_{MR}(m,r)] - \mathcal{K}^L_{MR}(E_{\pi'}[(m,r)])}{tr(Var_{\pi'}(m,r))} =: \mathcal{D}^U_{\pi'}(M,R) \; , \tag{40}$$

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

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

(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 $E_{\pi'}[(m,r)] \in \underline{\mathcal{O}}_{\mathcal{K}_{MR}}$.[18]

### 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:

$$\mathcal{H}_0(m^\star, r^\star) \ : \ \mathcal{K}^L_{MR}(m^\star, r^\star) \leq \mathcal{K}^{\mathbb{M}}_{MR}(m^\star, r^\star) \leq \mathcal{K}^U_{MR}(m^\star, r^\star) \ , \tag{41}$$

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

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

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

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

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:

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

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

A diagnostics test of null hypothesis

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

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

[18] In general, bounds (39) and (40) are not both always binding for the same prior $\pi$: $\mathcal{D}^L_\pi(M,R) \leq \mathcal{D}_\pi(M,R) \leq \mathcal{D}^U_\pi(M,R)$. 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.

dispersion properties, which is a test of the (composite) null hypothesis:

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

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

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

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

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 disperion, 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:

$$\mathcal{D}_\pi^L(M,R) \le \mathcal{D}_\pi^{\mathbb{M}}(M,R) \le \mathcal{D}_\pi^U(M,R)\ , \tag{49}$$

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 $\mathcal{K}_{MR}^U$:

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

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 Appendix V is extremal. Therefore, our empirical analysis of the arbitrage free dispersion properties of long-run risk models in Section Appendix V 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 constrainst 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 transienst 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 $R = (R_1, \ldots, R_{d_2})$ ($(1, e_i) \in \mathcal{O}_{\mathcal{K}_{MR}}$ 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

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

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:

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

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

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

As $\alpha \to 0$, 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:

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

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:

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

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

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

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:*

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

Figure 5 illustrates for $d_1 = d_2 = 1$ the construction of the above pricing kernel bounds. For $\alpha = 1/2$, we appy 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 constaints in Proposition 3 are tight, in the sense that they are the sharpest bounds implied by the arbitrage-free constraints on returns $(R_0, R_1, \ldots, R_{d_2})$, where $R_0$ is the risk-free return.[19] 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).[20] 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 $1-\sum_{i=1}^{d_2}\lambda_i, \lambda_1, \ldots, \lambda_{d_2}$ in returns $R_0, \ldots, R_{d_2}$.

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

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

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

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

*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 $\lambda^\star$ to this problem, it follows:*

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

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

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 $R_0, \ldots, R_{d_2}$, follows from a single arbitrage-free dispersion constraint

[19] Bansal and Lehmann (1997) and Alvarez and Jermann (2005), among others, show that the tightest pricing kernel entropy bound, which is obtained for $\alpha \to 0$ 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.

[20] This follows from the equivalence of the optimization problem:

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

with the minimal divergence problem:

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

in Almeida and Garcia (2011).

for portfolio return $R_{\lambda^\star}$.[21] The bound tightness in Proposition 4 also implies closed-form upper and lower arbitrage free CGF of the pricing kernel, wich explicit domains $D_U = [0,1]$ and $D_L = [0,1]^c$ where these functions take finite values:

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

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

## 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.

### 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 $\alpha \in \mathbb{R}$:

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

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 $\gamma \in \mathbb{R}$.*[22] *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 $\beta \in \mathbb{R}$:*

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

[21] 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.

[22] This change of measure is well-defined if and only if the marginal CGF of $M^T$ in $\gamma$ is well-defined.

*This bound is equivalent to the bound*

$$\frac{\mathcal{K}_{M^TM^P}(\gamma+(1-\gamma)\beta,\beta)}{\beta(\beta-1)} \geq -\frac{\mathcal{K}_{M^TR}(\gamma,-\beta/(1-\beta))}{\beta} \ , \tag{66}$$

*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 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:[23]

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

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

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

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.[24]

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

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

$$\mathcal{K}_{M_iR_i}(1,1) = \mathcal{K}_{MR}([e_i', e_i']) = 0 \ . \tag{69}$$

[23]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 $\alpha \in (0,1)$.

[24]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 $1/F_e$.

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

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

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

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

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 $R_{i0}, \ldots R_{iN_i}$ in market $i = 1, \ldots, d$, 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.[25]

**Proposition 6** (Multivariate Bound Tightness)**.** *Consider for* $\alpha \in \mathbb{R}^d_{++}$ *such that* $||\alpha||_1 < 1$ *the pricing kernel vector* $M^\star = (M_1^\star, \ldots, M_d^\star)$ *that solves the following minimum divergence problem:*[26]

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

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

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

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

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

*where* $R_\lambda = (R_{1\lambda_1}, \ldots, R_{d\lambda_d})$ *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*

[25] The proof is collected in the Supplemental Appendix.

[26] $\mathbb{R}^d_{++}$ denotes the $d-$dimensional strictly positive cone.

*of returns (with constraint $R_{i\lambda_i} > 0$) denominated in the $i-th$ domestic currency. It then follows:*

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

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

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

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 $R_{\lambda^\star} = (R_{1\lambda_1^\star}, \ldots, R_{d\lambda_d^\star})$. By construction, the tightness result in Proposition 6 also identifies in closed-form the upper arbitrage-free CGF on domain $D := \{\alpha \in \mathbb{R}_{++}^d : ||\alpha||_1 < 1\}$:

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

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

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

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 $R_{i1}, \ldots, R_{iN_i}$ 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 correponsing 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:

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

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:

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

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 $\alpha \to 0$ 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,

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

where $C_t$ is consumption at time $t$, $0 < \delta < 1$ the time preference rate, $\gamma$ the parameter of relative risk aversion and $\theta := \frac{1-\gamma}{1-1/\psi}$, with $\psi$ is the elasticity of inter-temporal substitution (IES). Utility maximization is subjec to 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:

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

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

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

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.[27] 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.[28] In order to factorize the pricing kernel in the LRR model into transient and persistent conponents, 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 $s = 1, \ldots, 12$ we simulate $N = 10^6$ annual paths of state dynamics (A.152) 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 $R_{\infty,t+1}(u)$, $u = 1, \ldots, N$.[29] 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,

[27] This specification is not rejected by the overidentification tests in Bansal, Kiku, and Yaron (2012).

[28] 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.

[29] 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.

for powers $(t, p)$ in domain $D = (0,1)^2$, we compute the Monte Carlo CGF estimate as:

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

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 6. It features a pronounced convexity along the $p$-axis and a much 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 $O_{\mathcal{K}_{MR_\lambda}}$.

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

(1) Restriction at the origin: $(0,0,0) \in O_{\mathcal{K}_{MR_\lambda}}$.

(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:

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

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).[30] 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 respected to size and book-to-market (Set C = S-G + L-G + S-V + L-V + Bond).[31] 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 $\mathcal{K}^U_{MR_\lambda}$ 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 $\mathcal{K}^U_{MR_{\lambda^\star}} = \inf_\lambda \mathcal{K}^U_{MR_\lambda}$ 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 Omnibus Diagnostic Tests of Null Hypothesis $\mathcal{H}_0(\overline{\mathcal{O}}_{\mathcal{K}_{MR_{\lambda^\star}}})$

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 $D \subset \overline{\mathcal{O}}_{\mathcal{K}_{MR_{\lambda^\star}}}$:

$$\mathcal{H}_0(\overline{D}) \ : \ \mathcal{K}^{\mathbb{M}}_M|_D \leq \mathcal{K}^U_{MR_{\lambda^\star}}|_D \ . \tag{86}$$

The estimated upper arbitrage free CGF $\hat{\mathcal{K}}^U_{MR_{\hat{\lambda}^\star}}|_D$ for data sets A and C is presented in the middle and the right panel of Figure 6, 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.

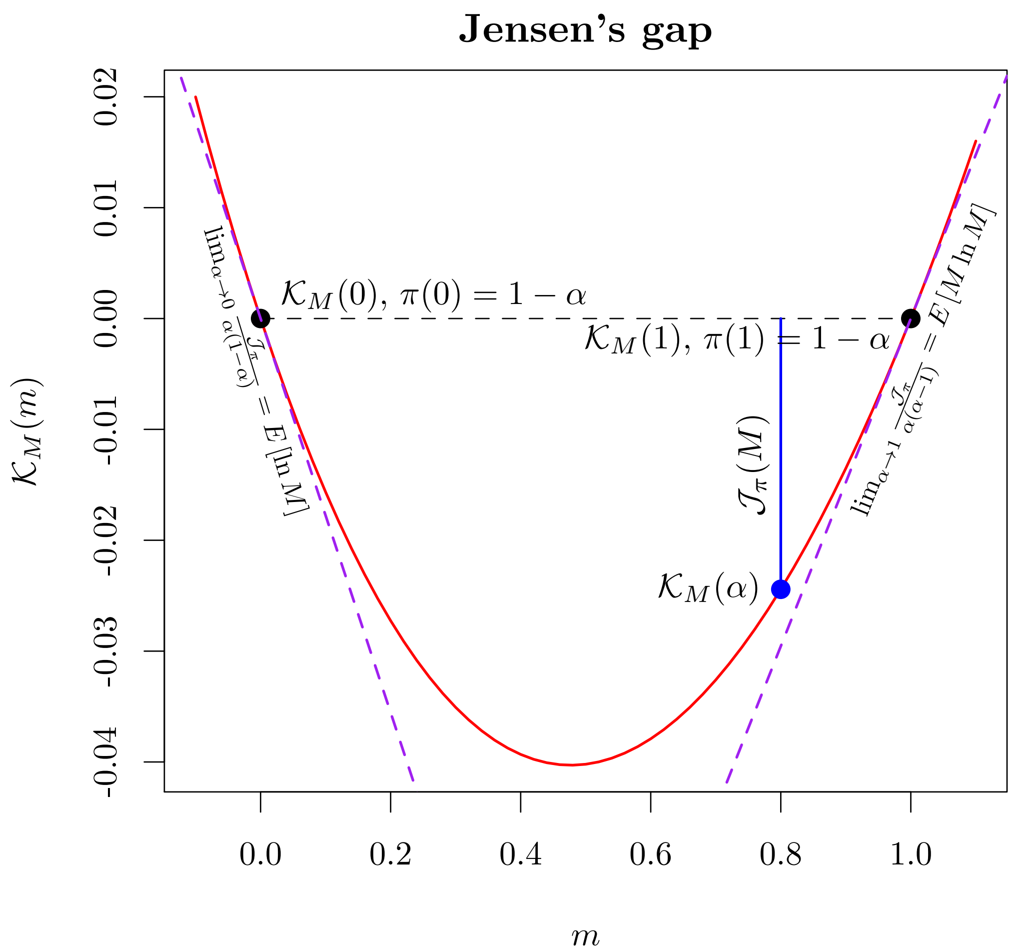
In order to rely on an accurate finite-sample inference on the LRR model, we develop a suitable bootstrap procedure for estimating bootstrap confidence intervals about point estimate $\widehat{\mathcal{K}}^U_{MR_{\hat{\lambda}^\star}}(p,t,0)$ for each $(p,t) \in D$.[32] Conservative bootstrap $p$-values for the test of null hypothesis (A.156) are presented in Figure 13. Using dataset A, we obtain no significant violation of this null hypothesis over the vast majority of domain $D$. In constrast, using datasets $B$ and $C$, we progressively obtain wider regions of rejection of null hypothesis (A.156), 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 (A.156) 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 (A.156) 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 adress the separation of model

[30] For brevity, S (L) stays for **S**mall (**L**arge) stocks and G (V) for **G**rowth (**V**alue) stocks.

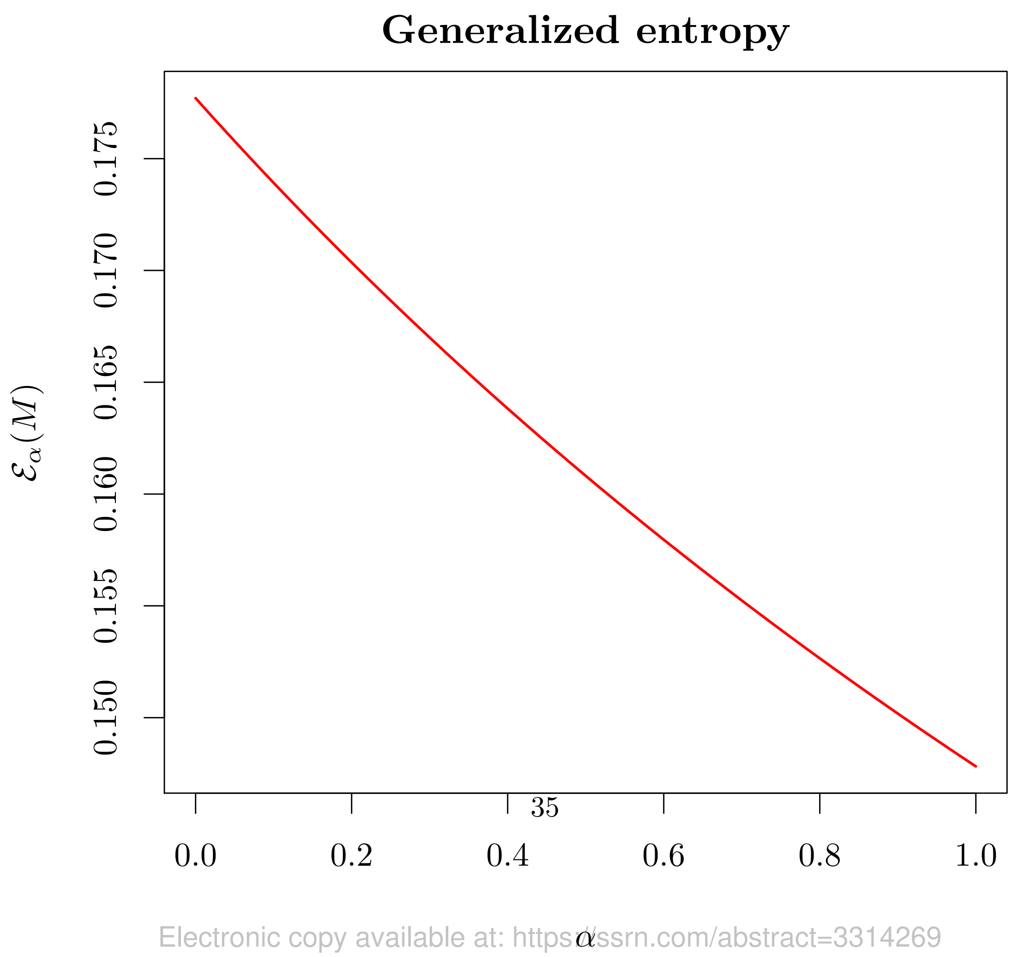
[31] We obtain these stock return data from Kenneth French's website. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

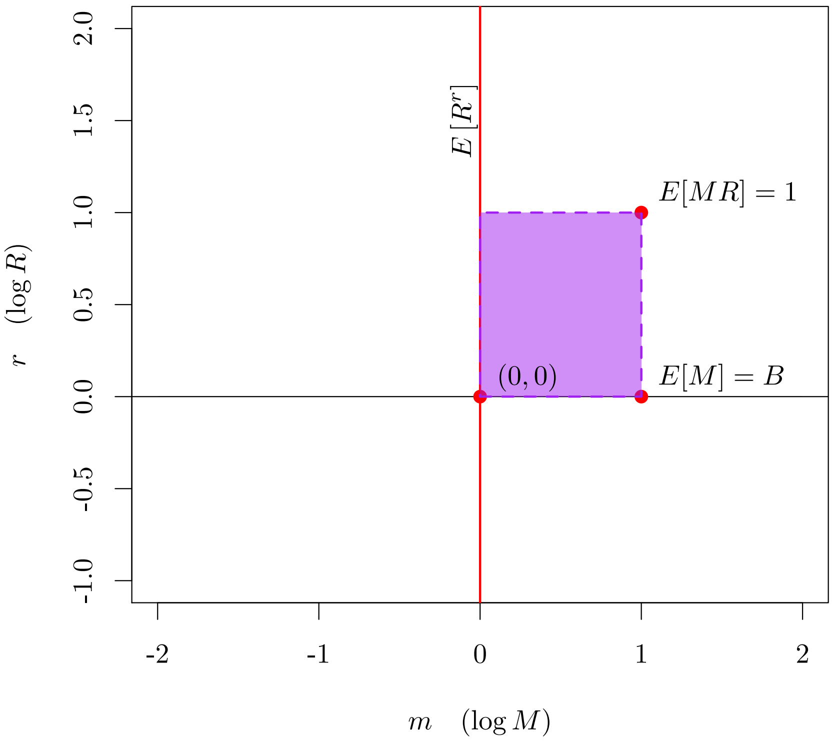
[32] Details on this bootstrap procedure are provided in the Supplemental Appendix, available from the authors upon request.

violations due to inappropriate dispersion from those due to a scaling discrepancy in the next sections.

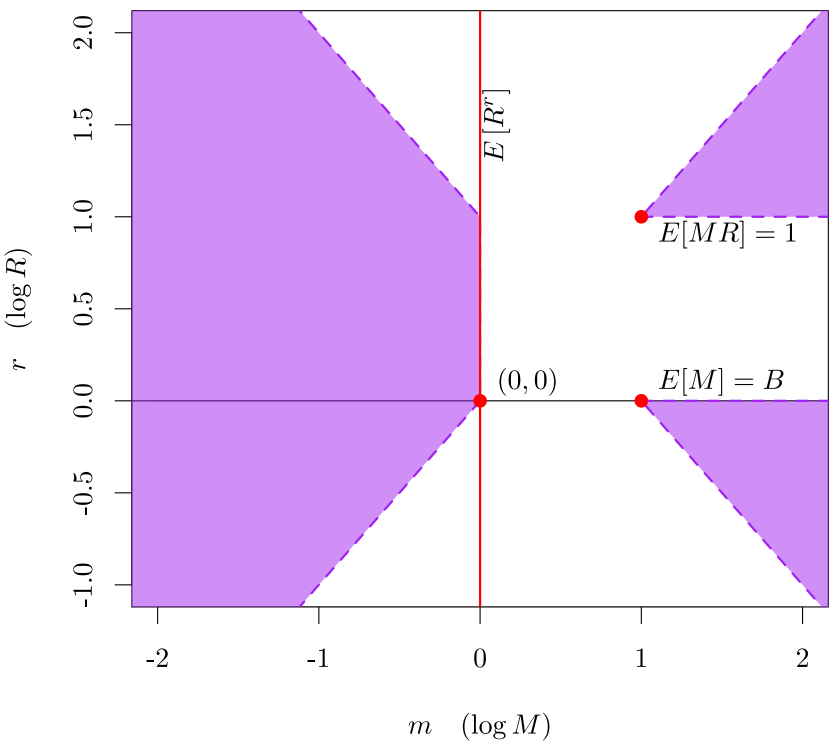



(a) Illustration of Jensen's gap $\mathcal{J}_\pi$ for stochastic discount factor $M$. $\mathcal{J}_\pi$ at $\alpha$ is calculated with prior $\pi$ such that $\pi(0) = 1-\alpha$ and $\pi(1) = \alpha$. The vertical blue line at $\alpha = 0.8$ is the gap. At $\alpha = 0$ ($\alpha = 1$) we take limits of the dispersion measure in panel 3b obtained by standardising $\mathcal{J}_\pi$ by $\alpha(1-\alpha)$ and illustrate them as the negative slope (slope) of the CGF at 0 (1).






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

Figure 4: 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 $B$.

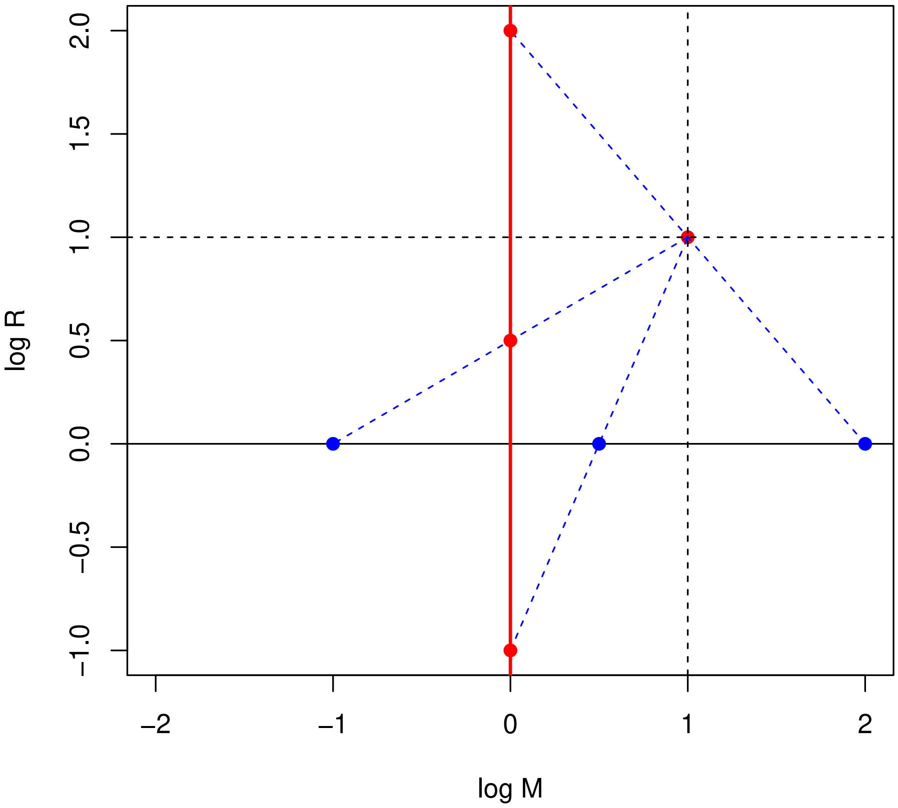



Figure 5: Formation of dispersion constraints of Type (1) and (2) in a setting with $d_1 = d_2 = 1$. A Type (1) constraint is used to bound the CGF value from above at unobservable point $(1/2, 0)$ with the use of values at observable points $(1, 1)$ and $(0, -1)$. Type (2) constraints are used to bound 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.

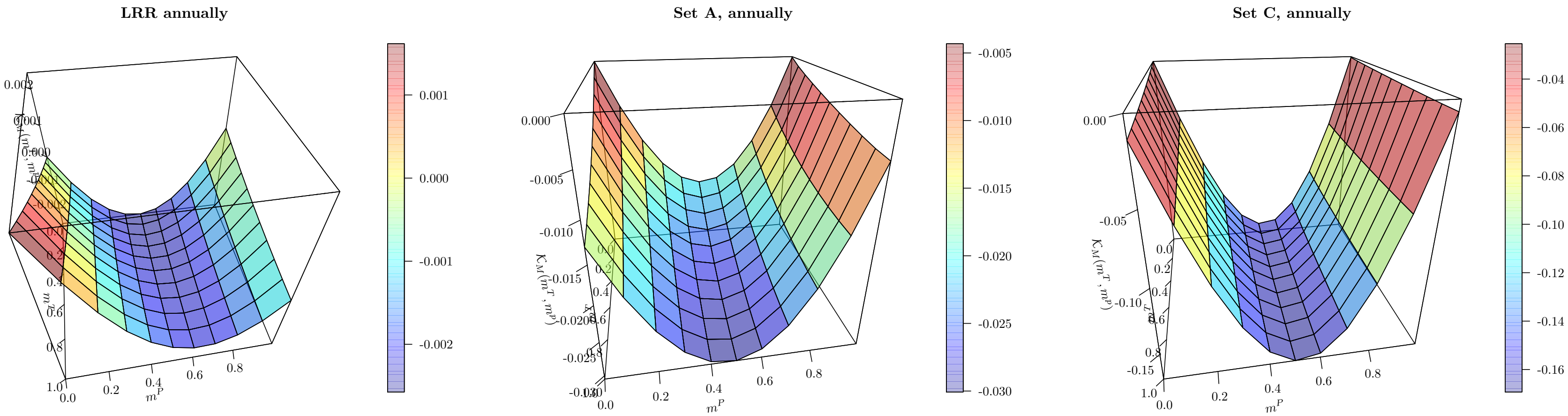



Figure 6: Model-implied marginal CGF of the SDF components (Bansal, Kiku, and Yaron (2012), leftmost panels), and estimates of $\mathcal{K}^U_{M^T M^P}(m^T, m^P)$, $(m^T, m^P) \in (0,1) \times (0,1)$, obtained as in Proposition 4. The middle panel presents results where the portfolio of assets contains 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 Appendix V.1.

### 6.5 Diagnostic 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.

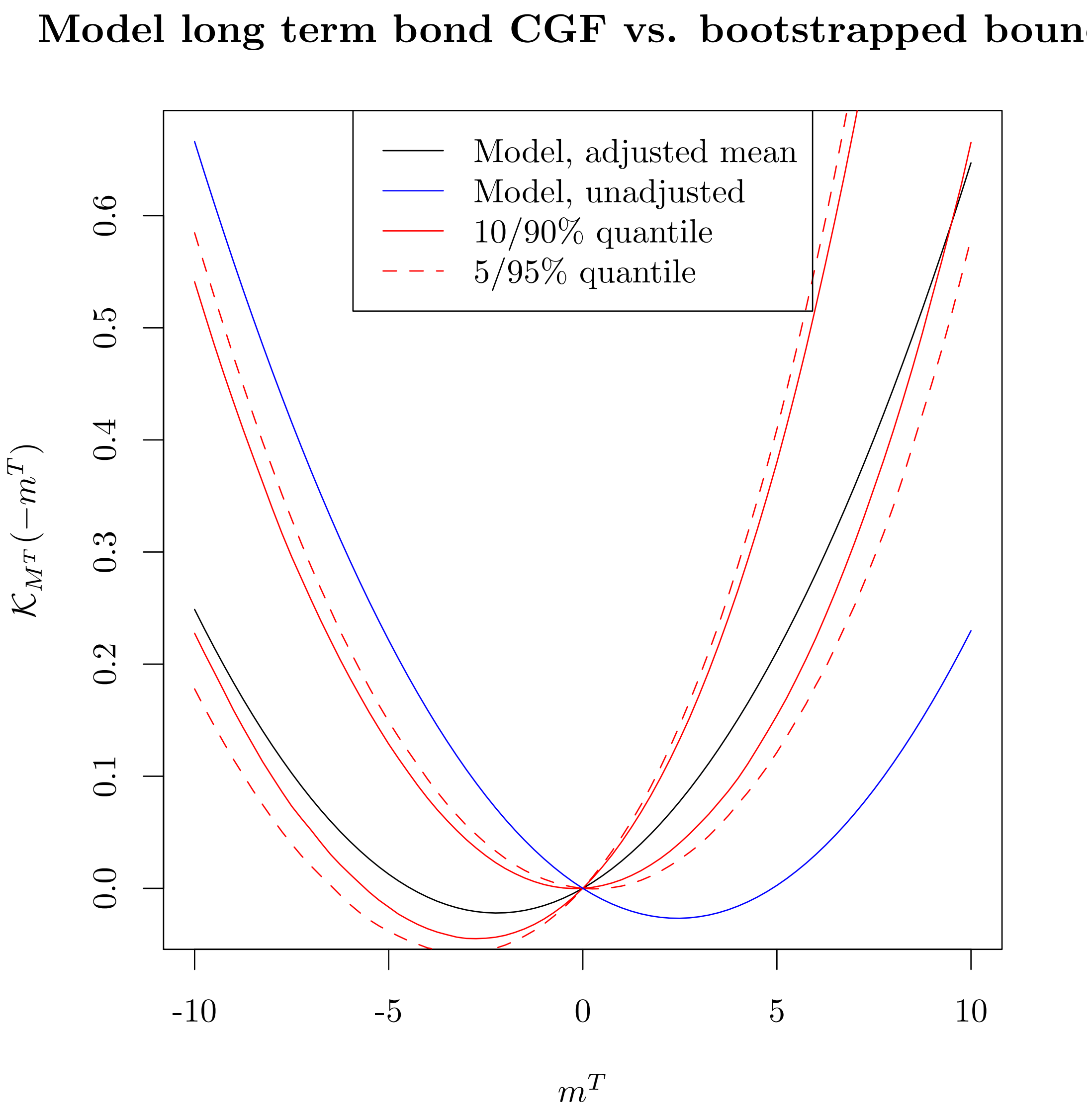
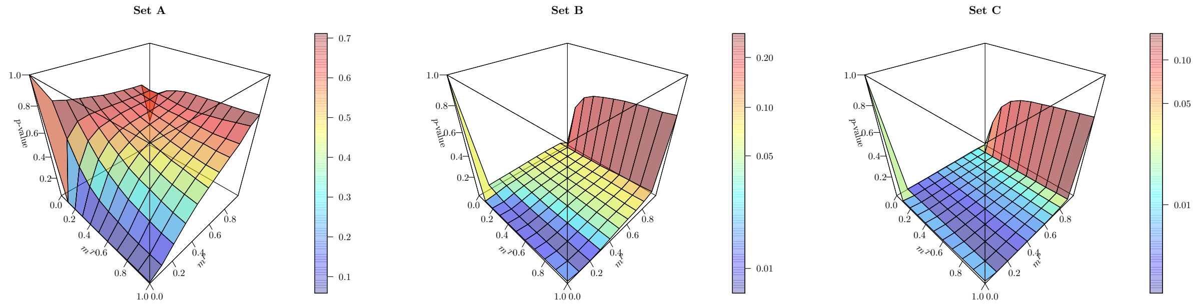
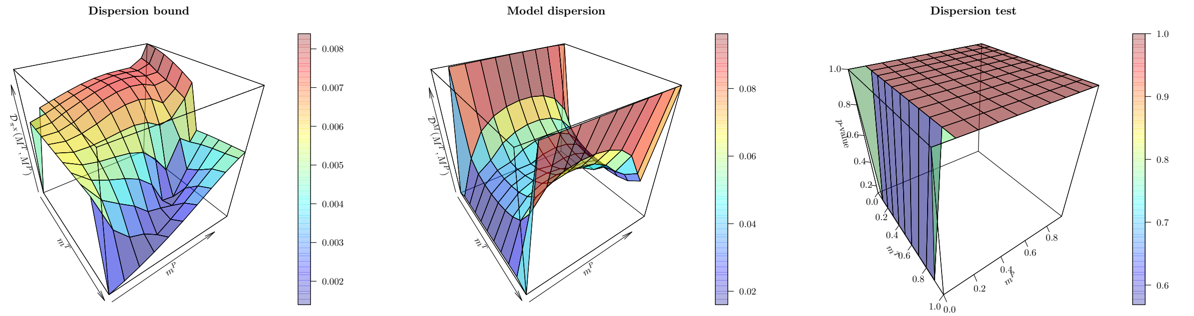



Figure 7: Model-based and estimated CGF confidence bounds of the transitory component of the stochastic discount factor, with $M^T$ estimated with real returns on long-maturity bonds. Bootstrapped confidence bounds are denoted by red solid and dashed lines. The LRR model-implied CGF is given by the blue 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) The omnibus test (see Section Appendix V.3.1): $p$-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 Appendix V.1.




(b) Dispersion test in the marginal SDF space (see Section Appendix V.3.3): $p$-values, for $(m^T, m^P) \in (0,1) \times (0,1)$, of the test of the lower dispersion bound (A.159) using only observable SDF component information. Annual frequency. Data set description is available in Section Appendix V.1.

Figure 8: Omnibus and marginal SDF space dispersion tests.

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 $E^{\mathbb{M}}[\widetilde{M}^T] = E[1/R_{LT}]$ produces a marginal model-implied CGF well-inside the bootstrap confidence intervals.[33] 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 $\pi_\alpha(0,0,0) = 1-\alpha$:

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

which is observable under Assumption 1 (5). Figure 9 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 $1/R_{LT}$.

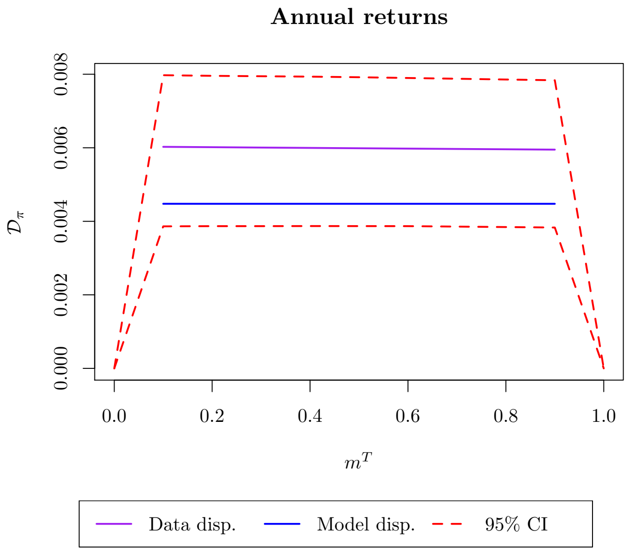



Figure 9: Observable dispersion of the transitory part of the SDF (A.157), calculated under the assumption that real returns on nominal long-maturity bonds are a proxy for real returns on infinite-maturity real bonds, which in turn are the inverse of the transitory component of the SDF, i.e. $R_\infty = \left(M^T\right)^{-1}$.

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

### 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:

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

where we used Assumption 1 (5) to write $M^T = 1/R_\infty$. Similar to the observable covariance measure $\mathrm{cov}(M^T, M^P) = B - E[1/R_\infty]$ 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 observable excess dispersion in Example 1:

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

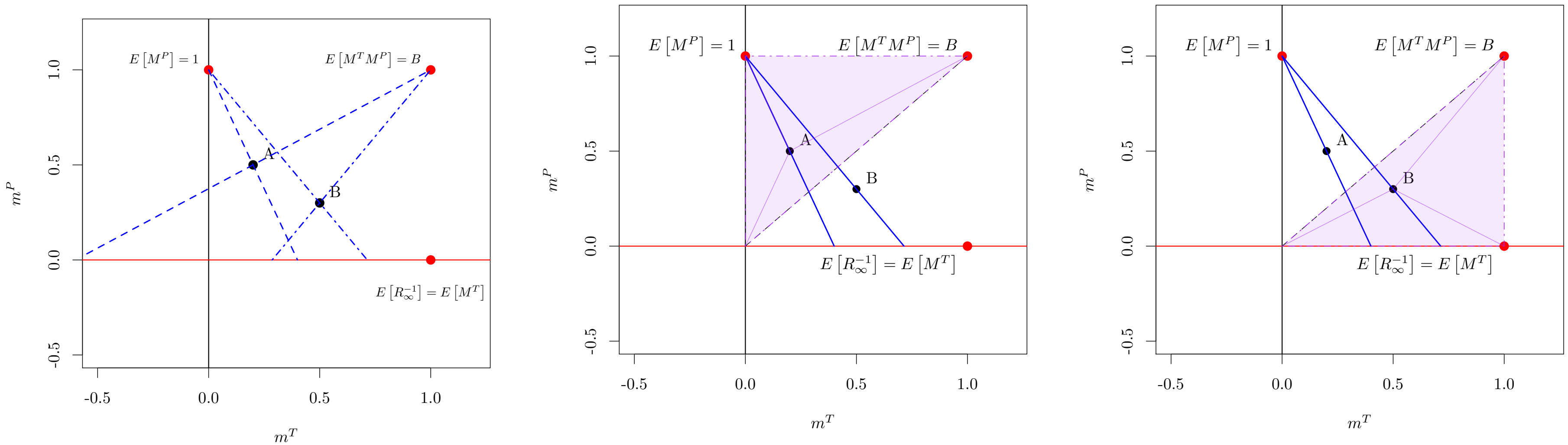
Table 1 presents the results of a direct test of null hypothesis $\mathcal{H}_0(\Delta\mathcal{J}) : \Delta\mathcal{J}_{\pi_1,\pi_2}(R_{LT}) = \Delta\mathcal{J}^{\mathbb{M}}_{\pi_1,\pi_2}(R_\infty)$ based on excess dispersion (89). We find that the model-implied excess dispersion is negative, while its 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 10, 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:

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

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




(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 10: 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 10a 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 10c and 10b the purple triangles provide two more ways of constructing Type (1) dispersion bounds. In order to construct $\mathcal{K}^U_{MR}$ one has to pick the lowest available bound value.

Consistently with the graphical description in Figure 10, 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 10a). 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 10c (Figure 10b).

The left and the middle panels of Figure 14 plot for any $(t_0, p_0) \in D$ the estimated lower bound $\widehat{\mathcal{D}}^L_{\pi_0, \Pi(t_0,p_0)}(M)$ and the model-implied dispersion $\mathcal{D}^{\mathbb{M}}_{\pi_0}(M)$, 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 (A.159). 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 $\mathcal{D}^{\mathbb{M}}_{\pi_0}(M) \geq \mathcal{D}^L_{\pi_0, \Pi(t_0,p_0)}(M)$, in the right panel of Figure 14, 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 (A.159) 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:

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

where $\mathcal{D}^L_{\pi_1}(M, R_\lambda)$ is observable empirically. By construction, $\mathcal{K}^U_{MR_\lambda}(t_0, p_0, 0)$ is finite and has the following explicit expression from Proposition 5: [34]

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

[34]Finiteness of $\mathcal{K}^U_{MR_\lambda}(t_0,p_0,0)$ 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 $\mathcal{O}_{\mathcal{K}_{MR_\lambda}}$.

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

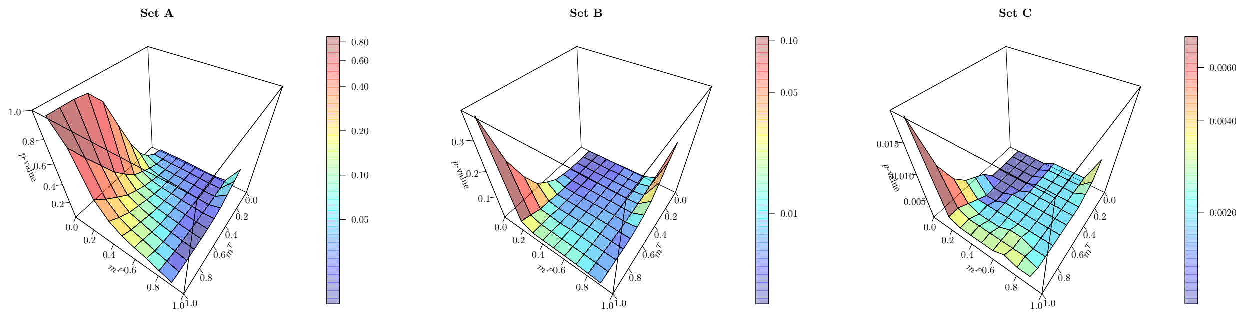
$$\mathcal{D}_{\pi_1}(M) \geq \mathcal{D}^L_{\pi_1}(M, R_{\lambda^\star}) \ , \tag{93}$$

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

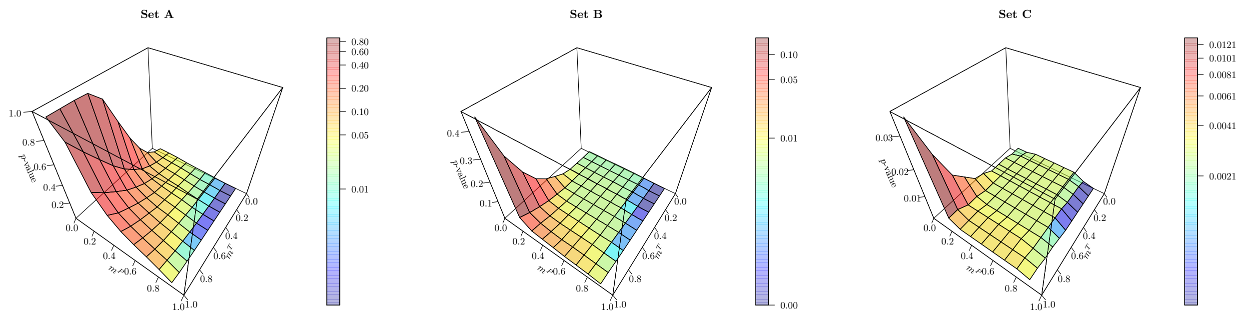
$$\mathcal{H}_0 \ : \ \mathcal{D}^{\mathbb{M}}_{\pi_1}(M) \geq \mathcal{D}^L_{\pi_1}(M, R^\star_\lambda) \ . \tag{94}$$

Figure 11 summarizes the test results of null hypothesis (A.162), 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 11a, 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 (A.162) 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 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 $p, t \to 0$, which correponds to a univariate test of the standard entropy bound (52), does not imply a model rejection at the 5% significance level.

[35]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.




(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 11: Dispersion tests in $(M, R)$ space (see Section Appendix V.3.4): $p$-values. The null hypothesis (A.162) is tested for $D_{\pi_{t,p}}$ as in (A.160) 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 Appendix V.1.

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 $R^{\star}_{\lambda}$ in null hypothesis (93) depends on the return of the long-maturity bond, we get in Figure 11b 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 helful to test multivariate pricing kernel specifications with a unifying approach. While our approach naturally incorporates existing AFD constrains 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.[36] 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.

[36]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 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).

| $1-\alpha$ | Monthly | | Quarterly | |
|---|---|---|---|---|
| data | 0.000336 | | 0.0007322 | |
| model | $-0.001673$ | | $-0.001812$ | |
| 95% | $-0.0006365$ | 0.001247 | $-0.00204$ | 0.00328 |
| 99% | $-0.001016$ | 0.001524 | $-0.002931$ | 0.004129 |

Table 2: Excess dispersion in the Bansal, Kiku, and Yaron (2012) model and in the data. Model values calculated with the use of monthly and quarterly calibration parameters in Table VI of their paper. Data values calculated from sample ranging from 1946-01-31 to 2012-12-31. Confidence intervals calculated with the use of a time-series bootstrap (basic confidence interval type).

## Appendix II Supplementary Figures

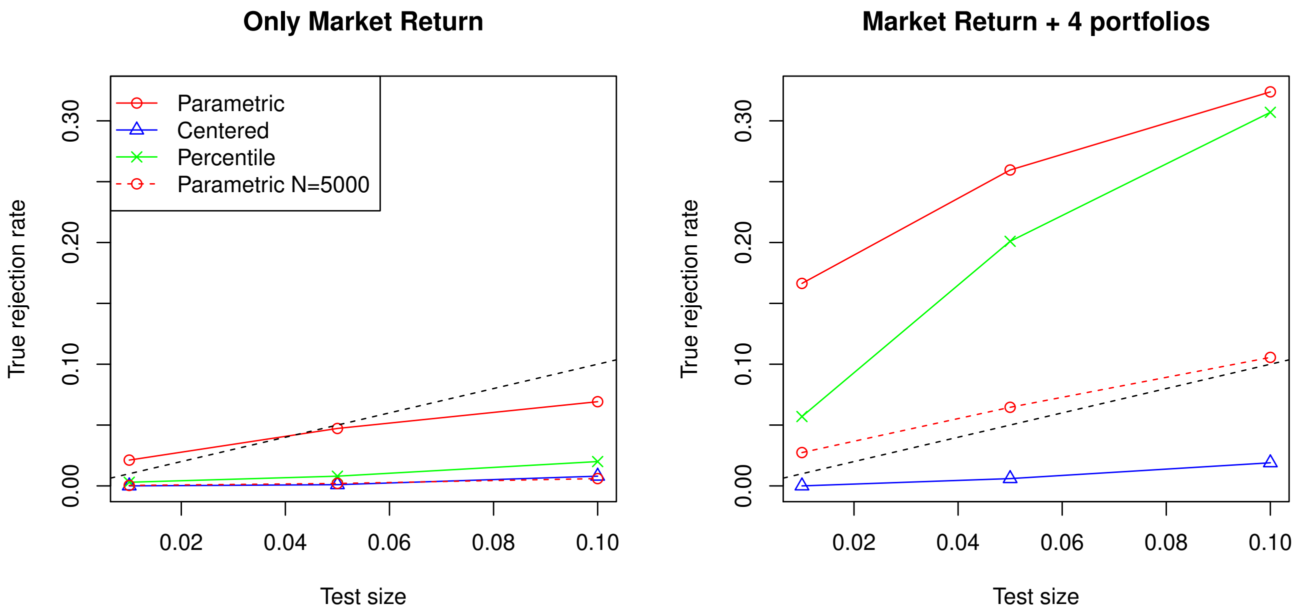
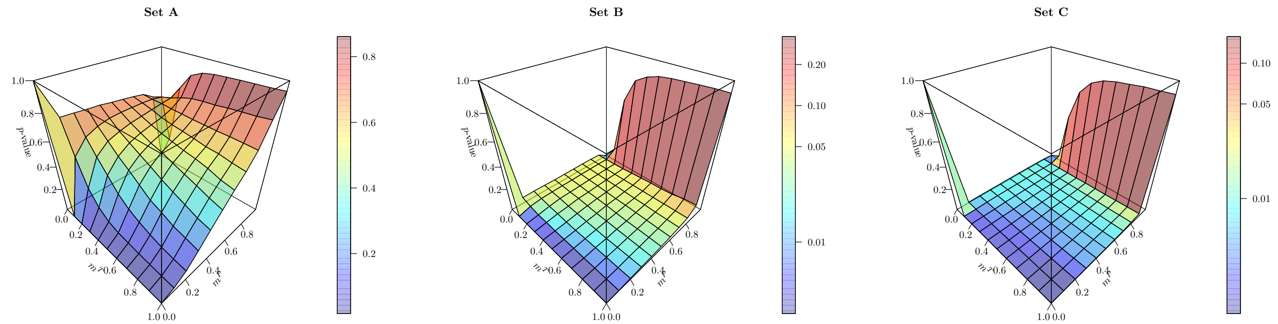
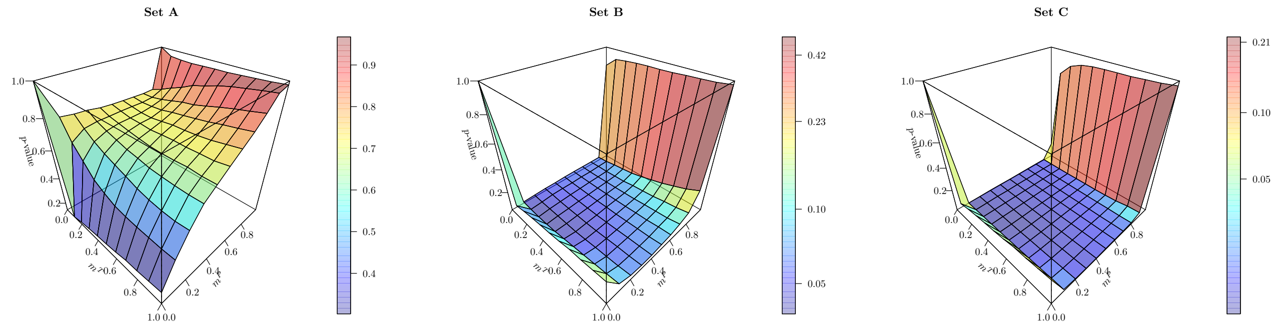



Figure 12: This figure plots the rejection rates using (a) the parametric bootstrap approach of Bakshi and Chabi-Yo (2014) (red circles), (b) a non-parametric percentile bootstrap (green cross) and (c) a non-parametric centered bootstrap (blue triangle). Points under the black dashed line constitute a conservative testing procedure, i.e. with a true size not greater than the nominal size. The $\alpha$ parameter of the dispersion measure is set to 0.5 in this graph. To calculate the coverage ratios under the LRR model, we simulated $N = 900$ monthly observations 5000 times and calculated the number of times the given bootstrap procedure would reject the true model. We also show results for the parametric bootstrap the case when we assume that a longer dataset of $N = 5000$ monthly observations is available.

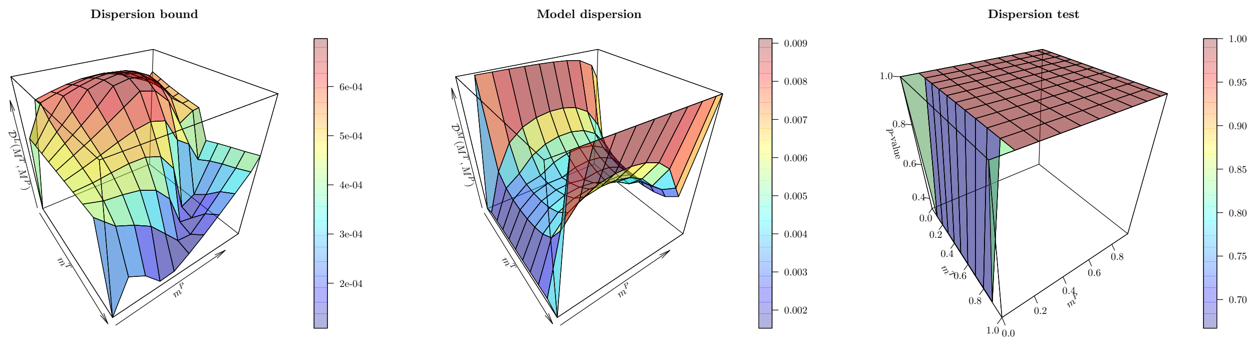



(a) $p$-values of omnibus test for $H_0 : \mathcal{K}^{\mathbb{M}}_{M^T,M^P} \leq \mathcal{K}^{\mathbb{P}}_{M^T,M^P}$ for monthly returns.

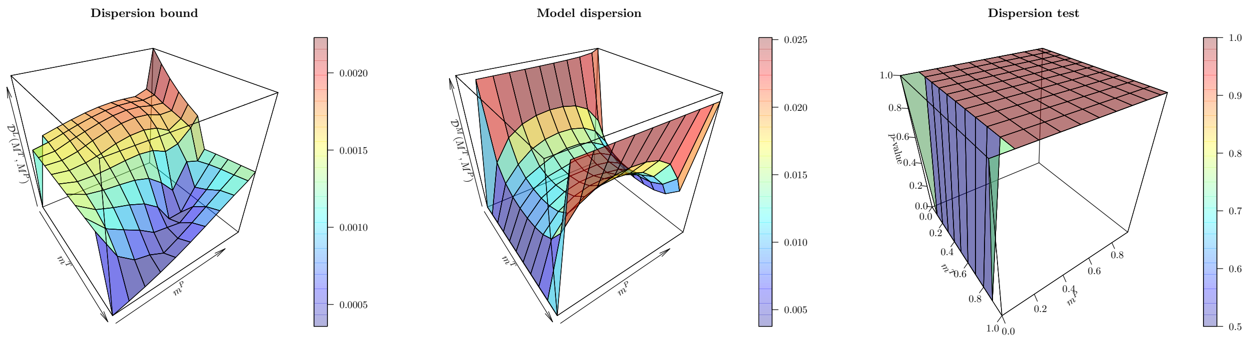



(b) $p$-values of omnibus test for $H_0 : \mathcal{K}^{\mathbb{M}}_{M^T,M^P} \leq \mathcal{K}^{\mathbb{P}}_{M^T,M^P}$ for quarterly returns.

Figure 13: The omnibus test (see Section Appendix V.3.1): $p$-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 Appendix V.1.




(a) Dispersion test in SDF space for monthly returns.




(b) Dispersion test in SDF sapce for quarterly returns.

Figure 14: Dispersion test in the marginal SDF space (see Section Appendix V.3.3): $p$-values, for $(m^T, m^P) \in (0,1) \times (0,1)$, of the test of the lower dispersion bound (A.159) using only observable SDF component information. Monthly and quarterly frequency. Data set description is available in Section Appendix V.1.

# Appendix III Proofs

*Proof of Proposition 2.* Let the dimension of the observable points be $N$ and choose $o_1, ..., o_N$ such that it formes a basis. Observing $\mathcal{K}^{\mathbb{P}}$ and $\mathcal{K}^{\mathbb{M}}$ on these points pins down the value $e$. Now choose another point $o_{N+1}$ and by contradiction assume that $\mathcal{K}^{\mathbb{P}}(o_{N+1}) - \mathcal{K}^{\mathbb{M}}(o_{N+1}) \neq e' \cdot o_{N+1}$.

(i) If $o_{N+1} \in Co(\{\mathbf{0}, o_1, ..., o_N\})$ then:

$$\begin{aligned}\mathcal{J}^{\mathbb{M}}_{\pi}(M) - \mathcal{J}^{\mathbb{P}}_{\pi}(M) &= \pi'[0, e'o_1, ..., e'o_N]' - (\mathcal{K}^{\mathbb{P}}(o_{N+1}) - \mathcal{K}^{\mathbb{M}}(o_{N+1})) \\ &\neq e'o_{N+1} - e' \cdot o_{N+1} = 0\end{aligned}$$

where $\pi$ is a prior over $\{\mathbf{0}, o_1, ..., o_N\}$ that has mean $o_{N+1}$. Thus in this case there is a dispersion mismatch, which is a contradiction.

(ii) If $o_{N+1} \notin Co(\{\mathbf{0}, o_1, ..., o_N\})$, then $(\{\mathbf{0}, o_1, ..., o_N, o_{N+1}\})$ forms an extremal set of dimension $N$. Then there exists a point $o$ that can be expressed as two distinct convex combinations of $(\{\mathbf{0}, o_1, ..., o_N, o_{N+1}\})$, $o = \alpha_i'[\mathbf{0}, o_1, ..., o_N, o_{N+1}]'$, $i = 1, 2$ and $\alpha_{1,N+2} \neq \alpha_{2,N+2}$. The excess dispersion under $\mathbb{P}$ minus the excess dispersion under $\mathbb{M}$ using the priors $\alpha_1$ and $\alpha_2$ is given by:

$$\begin{aligned}\left[\mathcal{J}^{\mathbb{P}}_{\alpha_1}(M) - \mathcal{J}^{\mathbb{P}}_{\alpha_2}(M)\right] - \left[\mathcal{J}^{\mathbb{M}}_{\alpha_1}(M) - \mathcal{J}^{\mathbb{M}}_{\alpha_2}(M)\right] &= (\alpha_1 - \alpha_2)'[0, e'o_1, .., e'o_N, \mathcal{K}^{\mathbb{P}}(o_{N+1}) - \mathcal{K}^{\mathbb{M}}(o_{N+1})]' \\ &= (\alpha_{1,N+2} - \alpha_{2,N+2})\left(\mathcal{K}^{\mathbb{P}}(o_{N+1}) - \mathcal{K}^{\mathbb{M}}(o_{N+1}) - e'o_{N+1}\right) \neq 0\end{aligned}$$

implying an excess dispersion mismatch between the model and the data.

In both cases we reached contradiction, which concludes the proof. □

*Proof of Proposition 4.* We study for $\alpha \in \mathbb{R}$ the minimum divergence problem:

$$\inf_{M} \left\{ \frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha - 1)} \right\} , \tag{A.95}$$

subject to $M > 0$ and $E[MR_k] = 1$ for each $k = 0, \dots, d_2$, where $R_0$ is the risk-free return. This is equivalent to the minimum divergence problem in Almeida and Garcia (2011):

$$\inf_M \left\{ \frac{E\left[M^\alpha\right]}{\alpha(\alpha-1)} \right\} , \tag{A.96}$$

subject to $M > 0$, $E[M(R_k - R_0)] = 0$ and $E[MR_0] = 1$. The Lagrange function for this problem is:

$$\mathcal{L}(M, \mu, \nu) = \frac{E\left[M^\alpha\right]}{\alpha(\alpha-1)} - \mu_0 E[MR_0 - 1] - \sum_{k=1}^{d_2} \mu_k E[M(R_k - R_0)] - \nu M ,$$

with the multiplier vector $\mu \in \mathbb{R}^{d_2+1}$ for the pricing constraints and the random multiplier $\nu$ for the positivity constraint. As the optimal pricing kernel needs to be strictly positive, multiplier $\nu$ vanishes almost surely and the first oder conditions for an optimum yield:

$$\frac{M^\alpha}{\alpha - 1} = M(\mu_0 R_0 + \mu_k(R_k - R_0)) =: M\mu_0 R_\mu . \tag{A.97}$$

Taking expectations on the RHS and the LHS, the pricing constraints give:

$$E[M^\alpha] = (\alpha - 1)\mu_0 . \tag{A.98}$$

Inserting this last expression in the first and the second first-order condition, it follows:

$$M_1 = \left[(\alpha - 1)\mu_0 R_\mu\right]^{1/(\alpha-1)} . \tag{A.99}$$

From this condition, the optimal pricing kernels $M^\star$ and optimal return $R_{\mu^\star}$ are such that:

$$E[(M^\star)^\alpha] = E\left[R_{\mu^\star}^{\alpha/(1-\alpha)}\right]^{1-\alpha} . \tag{A.100}$$

For any stochastic discount factor $M$ pricing returns, these findings yield the following inequalities:

$$\begin{aligned}
\frac{\mathcal{K}_M(\alpha)}{\alpha(\alpha-1)} &= \frac{\log E[M^\alpha]}{\alpha(\alpha-1)} \\
&\geq \frac{\log E[(M^\star)^\alpha]}{\alpha(\alpha-1)} \\
&= -\frac{\log E\left[R_{\mu^\star}^{\alpha/(\alpha-1)}\right]}{\alpha} \\
&= -\frac{\mathcal{K}_{R_{\mu^\star}}(\alpha/(\alpha-1))}{\alpha} \ ,
\end{aligned} \tag{A.101}$$

showing that the convexity bound implied by portfolio return $R_{\mu^\star}$ is tight. The resulting optimal stochastic discount factor is given by:

$$M_\star = R_{\mu^\star}^{1/(\alpha-1)} / E\left[R_{\mu^\star}^{\alpha/(\alpha-1)}\right] \ , \tag{A.102}$$

where optimal return

$$R_{\mu^\star} = R_0 + \sum_{k=1}^{d_2} \mu_k^\star (R_k - R_0) > 0 \ , \tag{A.103}$$

is such that for any $k = 1, \ldots, d_2$:

$$E\left[R_{\mu^\star}^{1/(\alpha-1)} (R_k - R_0)\right] = 0 \ . \tag{A.104}$$

The sharp convexity bound on the RHS of inequality (A.101) is obtained directly, by solving the dual maximization problem:

$$\sup_\mu \left\{ -\frac{\mathcal{K}_{R_\mu}(\alpha/(\alpha-1))}{\alpha} \right\} \ , \tag{A.105}$$

over portfolio weight vectors $\mu$ such that:

$$R_\mu = R_0 + \sum_{k=1}^{d_2} \mu_k (R_k - R_0) > 0 \ . \tag{A.106}$$

Indeed, this optimization problem is globally strictly concave and the first-order conditions of the unique maximum are for $k = 1, \ldots, d_2$:

$$0 = \left. \frac{\partial \mathcal{K}_{R_\mu}(\alpha/(\alpha-1))}{\partial \mu} \right|_{\mu=\mu^\star} = E\left[ R_{\mu^\star}^{1/(\alpha-1)} (R_k - R_0) \right] \ . \tag{A.107}$$

As these first-order conditions are identical to the pricing constraints for multiplier $\mu$ in the primal minimum divergence problem, the sharpness of the convexity bound resulting from maximization problem (A.105) is shown. This concludes the proof.

□

*Proof of Proposition 5.* $M^{\mathbb{T}} := M(M^T)^{-\gamma} E[(M^T)^\gamma]$ is strictly positive and for any marketed gross return $R$ we obtain:

$$1 = E[MR] = E[M^{\mathbb{T}} R (M^T)^\gamma / E[(M^T)^\gamma]] = E^{\mathbb{T}}[M^{\mathbb{T}} R] \ . \tag{A.108}$$

Therefore, for any $\beta \in \mathbb{R}$:

$$\frac{\mathcal{K}^{\mathbb{T}}_{M^{\mathbb{T}}}(\beta)}{\beta(\beta-1)} \geq -\frac{\mathcal{K}^{\mathbb{T}}_R(-\beta/(1-\beta))}{\beta} \ , \tag{A.109}$$

and this bound is sharp. Moreover,

$$\mathcal{K}^{\mathbb{T}}_R(-\beta/(1-\beta)) = -\mathcal{K}_{M^T}(\gamma) + \mathcal{K}_{M^T R}(\gamma, -\beta/(1-\beta)) \ . \tag{A.110}$$

Similarly,

$$\mathcal{K}^{\mathbb{T}}_{M^{\mathbb{T}}}(\beta) = \mathcal{K}_{M^T M^P}(\gamma + (1-\gamma)\beta, \beta) + (\beta - 1)\mathcal{K}_{M^T}(\gamma) \ . \tag{A.111}$$

This concludes the proof.

□

*Proof of Proposition 6.* We first study for given $\alpha \in \mathbb{R}^d_{++}$ such that $||\alpha||_1 < 1$ the solution of the multivariate minimum divergence problem:

$$\inf_M \left\{ -\frac{\mathcal{K}_M(\alpha)}{\prod_{i=1}^d \alpha_i} \right\} \ , \tag{A.112}$$

subject to $M_i > 0$ and $E[M_i R_{ik_i}] = 1$ for each $i = 1, \ldots, d$ and $k_i = 0, \ldots, N_i$, where $R_{i0}$ is the risk-free return in the $i-$th domestic market. This is equivalent to studying the problem:

$$\inf_M \left\{ -\frac{E\left[\prod_{i=1}^d M_i^{\alpha_i}\right]}{\prod_{i=1}^d \alpha_i} \right\} \ , \tag{A.113}$$

subject to $M_i > 0$, $E[M_i(R_{ik_i} - R_{i0})] = 0$ and $E[M_i R_{i0}] = 1$ for each $i = 1, \ldots, d$ and $k_i = 1, \ldots, N_i$. For brevity, we first prove the result in detail for a bivariate vector of pricing kernel components $M = (M_1, M_2)$ and for bivariate domestic and foreign markets with risky return vectors $R_1 = (R_{11}, R_{12})$, $R_2 = (R_{21}, R_{22})$ and riskless returns $R_{10}$, $R_{20}$. From these findings, we then provide the solution for the general case, without detailing all the intermediate steps. This optimization problem is a well-posed convex problem. The Lagrange function for this problem is:

$$\mathcal{L}(M, \mu, \nu) = -\frac{E\left[\prod_{i=1}^d M_i^{\alpha_i}\right]}{\prod_{i=1}^d \alpha_i} - \sum_{i=1}^d \mu_{i0} E[M_i R_{i0} - 1] - \sum_{i=1}^d \sum_{j=0}^d \mu_{ij} E[M_i(R_{ij} - R_{i0})] - \sum_{i=1}^d \nu_i M_i \ ,$$

with the multiplier matrix $\mu$ for the pricing constraints and the multiplier random vector $\nu$ for the positivity constraints. As the optimal pricing kernels need to be strictly positive, multiplier $\nu$ vanishes and the first oder conditions for an optimum are:

$$
\begin{aligned}
\frac{M_1^{\alpha_1} M_2^{\alpha_2}}{\alpha_2} &= M_1(\mu_{10}R_{10} + \mu_{11}(R_{11} - R_{10}) + \mu_{12}(R_{12} - R_{10})) =: M_1\mu_{10}R_{\mu_{1.}} \ , && \text{(A.114)} \\
\frac{M_1^{\alpha_1} M_2^{\alpha_2}}{\alpha_1} &= M_2(\mu_{20}R_{20} + \mu_{21}(R_{21} - R_{20}) + \mu_{22}(R_{22} - R_{20})) =: M_2\mu_{20}R_{\mu_{2.}} \ . && \text{(A.115)}
\end{aligned}
$$

Taking expectations on the RHS and the LHS, the pricing constraints give:

$$
E[M_1^{\alpha_1} M_2^{\alpha_2}] = \alpha_2\mu_{10} = \alpha_1\mu_{20} \ , \tag{A.116}
$$

$\alpha_2/\alpha_1 = \mu_{20}/\mu_{10}$ and:

$$
M_2 = M_1 \cdot \frac{R_{\mu_{1.}}}{R_{\mu_{2.}}} \ . \tag{A.117}
$$

Inserting this last expression in the first and the second first-order condition, it follows:

$$
\begin{aligned}
M_1 &= \left[\alpha_2\mu_{10}R_{\mu_{2.}}^{\alpha_2} R_{\mu_{1.}}^{1-\alpha_2}\right]^{1/(\alpha_1+\alpha_2-1)} \ , && \text{(A.118)} \\
M_2 &= \left[\alpha_1\mu_{20}R_{\mu_{2.}}^{1-\alpha_1} R_{\mu_{1.}}^{\alpha_1}\right]^{1/(\alpha_1+\alpha_2-1)} \ . && \text{(A.119)}
\end{aligned}
$$

From these conditions, the optimal pricing kernels $M_1^\star$, $M_2^\star$ and optimal returns $R_{\mu_{1.}^\star}$, $R_{\mu_{2.}^\star}$ are such that:

$$
(M_1^\star)^{\alpha_1}(M_2^\star)^{\alpha_2} = (\alpha_2\mu_{10})^{(\alpha_1+\alpha_2)/(\alpha_1+\alpha_2-1)} \left(R_{\mu_{1.}^\star}^{\alpha_1} R_{\mu_{2.}^\star}^{\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)} \ . \tag{A.120}
$$

Therefore, taking expectations and recalling equation (A.116), the optimal stochastic discount factors are such that:

$$
E[(M_1^\star)^{\alpha_1}(M_2^\star)^{\alpha_2}] = E\left[\left(R_{\mu_{1.}^\star}^{\alpha_1} R_{\mu_{2.}^\star}^{\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)}\right]^{1-\alpha_1-\alpha_2} \ . \tag{A.121}
$$

For any two stochastic discount factors $M_1$, $M_2$ pricing returns, the above findings yield the following inequalities:

$$\begin{aligned} -\mathcal{K}_M(\alpha) &= -\log E[M_1^{\alpha_1} M_2^{\alpha_2}] \\ &\geq -\log E[(M_1^\star)^{\alpha_1} (M_2^\star)^{\alpha_2}] \\ &= (\alpha_1+\alpha_2-1)\log E\left[\left(R_{\mu_{1\cdot}^\star}^{\alpha_1} R_{\mu_{2\cdot}^\star}^{\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)}\right] \\ &= (\alpha_1+\alpha_2-1)\mathcal{K}_{R_{\mu_{1\cdot}^\star} R_{\mu_{2\cdot}^\star}}(\alpha/(\alpha_1+\alpha_2-1)) \ , \end{aligned} \tag{A.122}$$

showing that the convexity bound implied by portfolio returns $R_{\mu_{1\cdot}^\star}$ and $R_{\mu_{2\cdot}^\star}$ is tight. The resulting optimal stochastic discount factors are given by:

$$M_1^\star = \left[R_{\mu_{2\cdot}^\star}^{\alpha_2} R_{\mu_{1\cdot}^\star}^{1-\alpha_2}\right]^{1/(\alpha_1+\alpha_2-1)} / E\left[\left(R_{\mu_{1\cdot}^\star}^{\alpha_1} R_{\mu_{2\cdot}^\star}^{\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)}\right] \ , \tag{A.123}$$

$$M_2^\star = \left[R_{\mu_{2\cdot}^\star}^{1-\alpha_1} R_{\mu_{1\cdot}^\star}^{\alpha_1}\right]^{1/(\alpha_1+\alpha_2-1)} / E\left[\left(R_{\mu_{1\cdot}^\star}^{\alpha_1} R_{\mu_{2\cdot}^\star}^{\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)}\right] \ , \tag{A.124}$$

where optimal returns

$$R_{\mu_{1\cdot}^\star} := R_{10} + \sum_{k_1=1}^{N_1} \mu_{1k_1}^\star (R_{1k_1} - R_{10}) > 0 \ , \tag{A.125}$$

$$R_{\mu_{2\cdot}^\star} := R_{20} + \sum_{k_2=1}^{N_2} \mu_{2k_2}^\star (R_{2k_2} - R_{20}) > 0 \ , \tag{A.126}$$

are such that for any $k_1 = 1, \ldots, N_1$ and $k_2 = 1 \ldots, N_2$:

$$E\left[\left(R_{\mu_{2\cdot}^\star}^{\alpha_2} R_{\mu_{1\cdot}^\star}^{1-\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)} (R_{1k_1} - R_{10})\right] = 0 = E\left[\left(R_{\mu_{2\cdot}^\star}^{1-\alpha_1} R_{\mu_{1\cdot}^\star}^{\alpha_1}\right)^{1/(\alpha_1+\alpha_2-1)} (R_{2k_2} - R_{20})\right] \tag{A.127}$$

The sharp convexity bound on the RHS of inequality (A.122) is obtained directly, by solving the maximization problem:

$$\sup_{\mu_{1\cdot}\mu_{2\cdot}} \left\{-\mathcal{K}_{R_{\mu_{1\cdot}} R_{\mu_{2\cdot}}}(\alpha/(\alpha_1+\alpha_2-1))\right\} \ , \tag{A.128}$$

over portfolio weight vectors $\mu_{1\cdot}$, $\mu_{2\cdot}$ such that:

$$R_{\mu_{1\cdot}} = R_{10} + \sum_{k_1=1}^{N_1} \mu_{1k_1}(R_{1k_1} - R_{10}) > 0 \ , \tag{A.129}$$

$$R_{\mu_{2\cdot}} = R_{20} + \sum_{k_2=1}^{N_2} \mu_{2k_2}(R_{2k_2} - R_{20}) > 0 \ . \tag{A.130}$$

Indeed, this optimization problem is globally strictly concave and the first-order conditions for a maximum read for $k_1 = 1, \ldots, N_1$ and $k_2 = 1, \ldots, N_2$:

$$0 = \left.\frac{\partial \mathcal{K}_{R_{\mu_{1\cdot}} R_{\mu_{2\cdot}}}(\alpha/(\alpha_1+\alpha_2-1))}{\partial \mu_{1k_1}}\right|_{\mu_{1\cdot}=\mu_{1\cdot}^\star} = E\left[\left(R_{\mu_{2\cdot}^\star}^{\alpha_2} R_{\mu_{1\cdot}^\star}^{1-\alpha_2}\right)^{1/(\alpha_1+\alpha_2-1)}(R_{1k_1} - R_{10})\right] \tag{A.131}$$

$$0 = \left.\frac{\partial \mathcal{K}_{R_{\mu_{1\cdot}} R_{\mu_{2\cdot}}}(\alpha/(\alpha_1+\alpha_2-1))}{\partial \mu_{2k_2}}\right|_{\mu_{2\cdot}=\mu_{2\cdot}^\star} = E\left[\left(R_{\mu_{2\cdot}^\star}^{1-\alpha_1} R_{\mu_{1\cdot}^\star}^{\alpha_1}\right)^{1/(\alpha_1+\alpha_2-1)}(R_{2k_2} - R_{20})\right] \tag{A.132}$$

As these first-order conditions are identical to the pricing constraints (A.127) in the primal multivariate minimum divergence problem, the sharpness of the convexity bound resulting from maximization problem (A.128) is shown. In the general case with $i = 1, \ldots, d$ domestic markets, the first-order conditions for a minimum in the primal problem read:

$$\frac{\prod_{j=1}^d M_j^{\alpha_j}}{\prod_{j\neq i} \alpha_j} = M_i \mu_{i0} R_{\mu_{i\cdot}} \ , \tag{A.133}$$

where

$$E\left[\prod_{j=1}^d M_j^{\alpha_j}\right] = \mu_{i0} \prod_{j\neq i} \alpha_j \ , \tag{A.134}$$

and for any $j \neq i$:

$$M_i = M_j \frac{R_{\mu_{i\cdot}}}{R_{\mu_{j\cdot}}} \ . \tag{A.135}$$

Therefore,

$$M_i = \left[ \left( \mu_{i0} \prod_{j \neq i} \alpha_j \right) R_{\mu_{i\cdot}}^{1-\sum_{j\neq i} \alpha_j} \prod_{j \neq i} R_{\mu_{j\cdot}}^{\alpha_j} \right]^{1/(\sum_{j=1}^{d} \alpha_j - 1)} . \tag{A.136}$$

The optimal vector of pricing kernel components $M^\star = (M_1^\star, \ldots, M_d^\star)$ follows as:

$$M_i^\star = \frac{\left[ R_{\mu_{i\cdot}^\star}^{1-\sum_{j\neq i} \alpha_j} \prod_{j\neq i}^{d} R_{\mu_{j\cdot}^\star}^{\alpha_j} \right]^{1/(\sum_{j=1}^{d} \alpha_j - 1)}}{E\left[ \prod_{j=1}^{d} R_{\mu_{j\cdot}^\star}^{\alpha_j} \right]} , \tag{A.137}$$

with the optimal returns

$$R_{\mu_{j\cdot}^\star} = R_{j0} + \sum_{k_j=1}^{N_j} \mu_{jk}^\star (R_{jk} - R_{j0}) > 0 , \tag{A.138}$$

and multipliers $\mu_{j1}, \ldots, \mu_{jN_j}$ that satisfy the pricing constraints in the $j-$th $(j = 1, \ldots, d)$ domestic market. Overall, we obtain that the optimal vector of pricing kernel components is such that

$$E\left[ \prod_{i=1}^{d} (M_i^\star)^{\alpha_i} \right] = E\left[ \prod_{i=1}^{d} R_{\mu_{i\cdot}^\star}^{\alpha_i} \right]^{1-\sum_{i=1}^{d} \alpha_i} , \tag{A.139}$$

implying

$$\begin{aligned} -\mathcal{K}_M(\alpha) &\geq -\log E\left[ \prod_{i=1}^{d} (M_i^\star)^{\alpha_i} \right] \\ &= \left( \sum_{i=1}^{d} \alpha_i - 1 \right) \mathcal{K}_{R_{\mu_\cdot^\star}} \left( \alpha / \left( \sum_{i=1}^{d} \alpha_i - 1 \right) \right) , \end{aligned} \tag{A.140}$$

with the optimal vector of returns $R_{\mu_\cdot^\star} := (R_{\mu_{1\cdot}^\star}, \ldots, R_{\mu_{d\cdot}^\star})$. This shows the tightness of the convexity bound implied by return vector $R_{\mu_\cdot^\star}$. Following the same arguments as in the bivariate case, it is straightforward to see that this bound is attained, by solving the maximization problem:

$$\sup_{\mu_{1\cdot}, \mu_{2\cdot}, \ldots, \mu_{d\cdot}} \left\{ -\mathcal{K}_{R_\mu} (\alpha / (\sum_{i=1}^{d} \alpha_i - 1)) \right\} , \tag{A.141}$$

over portfolio weight vectors $\mu_{1\cdot}, \mu_{2\cdot}, \ldots, \mu_{d\cdot}$ such that:

$$R_{\mu_{i\cdot}} = R_{i0} + \sum_{k_i=1}^{N_i} \mu_{ik_i}(R_{ik_i} - R_{i0}) > 0 \ , \tag{A.142}$$

for all $i = 1, \ldots, d$. This concludes the proof. □

*Proof.* We study for given $\alpha \in \mathbb{R}^d_{++}$ such that $||\alpha||_1 < 1$ the solution of the multivariate minimum divergence problem:

$$\inf_M \left\{ -\frac{\mathcal{K}_M(\alpha)}{\prod_{i=1}^d \alpha_i} \right\} \ , \tag{A.143}$$

subject to $M_i > 0$ and $E[M_i R_d/R_{i-1}] = E[M_i R_{0d}/R_{0(i-1)}] = 1$ for each $i = 1, \ldots, d$, where $R_{00} := 1$. This is equivalent to studying the problem:

$$\inf_M \left\{ -\frac{E\left[\prod_{i=1}^d M_i^{\alpha_i}\right]}{\prod_{i=1}^d \alpha_i} \right\} \ , \tag{A.144}$$

subject to $M_i > 0$, $E[M_i R_{0d}/R_{0(i-1)}] = 1$ and $E[M_i(R_d/R_{i-1} - R_{0d}/R_{0(i-1)})] = 0$ for each $i = 1, \ldots, d$. □

The first order conditions for this problem are:

$$\prod_{i=1}^d M_i^{\alpha_i} = \prod_{j \neq i} \alpha_j M_i(\mu_{i0} R_{0d}/R_{0(i-1)} + \mu_{i1}(R_d/R_{i-1} - R_{0d}/R_{0(i-1)})) =: \mu_{i0} \prod_{j \neq i} \alpha_j M_i R_{\mu_{i\cdot}} \ . \tag{A.145}$$

# Appendix IV Bootstrapping convexity bounds

## Appendix IV.1 Bootstrap testing

In the above we treated the case when $\mathbb{P}$ is perfectly observable (known). With $\mathbb{P}$ left to be estimated based upon a finite number of observations, we incorporate statistical uncertainty into our testing procedure.

The separation of potential model violations across observable points and pricing implied dispersion inequalities is also useful when considering statistical testing. For the case of observable points the statistical testing is relatively straightforward, since the cumulant generating function is a non-linear transformation of a sample mean. Therefore confidence bounds for the population quantities can be derived using standard boostrapping techniques for the sample mean, taking into account the potential timeseries dependence. As in the above, one can derive confidence bounds for observable dispersions, excess dispersions and $N$ linearly independent observable points, which together test all the observable point restrictions. If there is any violation occurring, then this can immediately be attributed to the appropriate model feature mismatch (dispersion / excess dispersion / scale violation).

On the other hand, calculating bootstrapped confidence intervals for convexity bounds (coming from pricing restrictions) with valid coverage ratios is a more challenging task. In this section we outline a valid bootstrap approach to test models by comparing data implied convexity bounds to model implied quantities. Given that we are working with inequalities, we can only hope to develop testing procedures that have a *conservative* rejection rate, i.e. we can make sure that we do not reject the correct model with more than $q\%$ probability for some small $q$. If with such a conservative procedure we can still reject a model, then we can be sure with probability atleast $1-q$ that the model is not a good description of reality.

We now illustrate our bootstrap approach in the context of the univariate bound (52). For simplicity we assume in the following that the risk-free return is known to be 0, this is purely for expositional purposes. To develop a valid bootstrap approach we can write the convexity bound as a *functional* of the joint distribution of the returns. Given a distribution $\mathbb{D}$ of returns, the optimal convexity bound functional $\mathcal{B}$ is given by:

$$\mathcal{B}(\mathbb{D}) = \max_{\lambda} \left\{ -\frac{\mathcal{K}^{\mathbb{D}}_{R_\lambda}(-\alpha/(1-\alpha))}{\alpha} \right\} \tag{A.146}$$

where $R_\lambda$ is the portfolio return with weights $\lambda$ and the cumulant generating function $\mathcal{K}$ is evaluated under the measure $\mathbb{D}$.

Then under the true statistical distribution $\mathbb{P}$, the following inequality holds:

$$\mathcal{B}(\mathbb{P}) \leq \mathcal{E}^{\mathbb{P}}_{\alpha}(M) \tag{A.147}$$

Our goal is to provide a consistent one-sided confidence interval on $\mathcal{B}(\mathbb{P})$, given a finite sample drawn from this distribution $\xi_1, ..., \xi_n$, and the associated discrete approximating distribution: $\mathbb{P}_n$. More formally we are looking for a functional of the finite sample distribution $c_n(q) = c(\xi_1, ..., \xi_n)$ that asymptotically satisfies:

$$\lim_{n\to\infty} P\left(\mathcal{B}(\mathbb{P}) \leq c_n(q)\right) = q \tag{A.148}$$

Given inequality (A.147), a consistent one-sided confidence interval on $\mathcal{B}(\mathbb{P})$ will then provide an asymptotically *conservative* confidence interval on $\mathcal{E}^{\mathbb{P}}_{\alpha}(M)$, i.e.:

$$\lim_{n\to\infty} P\left(\mathcal{E}^{\mathbb{P}}_{\alpha}(M) \leq c_n(q)\right) \leq q \tag{A.149}$$

Under regularity conditions on the functional $\mathcal{B}$,[37] Politis and Romano (1994) establish that the stationary bootstrap resampling provides a consistent confidence interval for $\mathcal{B}(\mathbb{P})$. Let the discrete distributions $\mathbb{P}_n^*$ be generated by applying a stationary resampling scheme to $\xi_1, ..., \xi_n$ and denote by $L_n^*$ the distribution function of the random variable $\sqrt{n}(\mathcal{B}(\mathbb{P}_n^*) - \mathcal{B}(\mathbb{P}_n))$ (conditional on $\xi_1, ..., \xi_n$), whilst let the distribution function $L_n$ denote the distribution function of $\sqrt{n}(\mathcal{B}(\mathbb{P}_n) - \mathcal{B}(\mathbb{P}))$. Then $\rho(L_n, L_n^*) \to 0$ in probability, for any $\rho$ that metricizes weak convergence of distribution functions.

This implies that if $u_n(1-q)$ is the $1-q$ quantile of $L_n^*$, then $c_n(q) = \mathcal{B}(\mathbb{P}_n) - u_n(1-q)$ provides a consistent one-sided confidence interval for $\mathcal{B}(\mathbb{P})$. Since the limiting distribution of $L_n$ is symmetric Gaussian, $c_n = \mathcal{B}(\mathbb{P}_n) + u_n(q)$ also provides a consistent confidence interval (percentile bootstrap). However in finite samples the distribution of $L_n^*$ might be heavily skewed, especially if the dimension of $\lambda$ is large, implying many assets over which we optimize.

We now contrast our approach with the one taken in Bakshi and Chabi-Yo (2014) and we show that their parametric approach can result in overly many rejections of a model and simple non-parametric bootstrap stragies such as the percentile bootstrap also suffer from over-rejection. We show through a simulation exercise that a basic centered bootstrap can provide reasonable rejection ratios even if multiple asset returns are used to derive bounds.

### Appendix IV.2 Sample versions of the bounds

It important to distinguish between the sample version of a convexity bound such as (2) and a bound on the sample realization of $M$ (which is not observed). To understand the difference, assume that we have a sample of i.i.d. realizations $(M_i, R_i)$, $i = 1..N$ of the bivariate random variable $(M, R)$ which satisfies $\mathbb{E}(MR) = 1$[38]. In this case due to the *exact equality* $\mathbb{E}(MR) = 1$ we have inequality (2) holding for the *population* moments of $M, R$, in particular we know that:

$$\mathcal{K}(\alpha, 0) \leq (1-\alpha)\mathcal{K}\left(0, -\alpha/(1-\alpha)\right)$$

where $\mathcal{K}$ is the *population* cumulant generating function of $(M, R)$ .

Whilst in population the equality $E(MR) = 1$ holds, this is not necessarily true in sample, i.e. in general:

$$\frac{\sum_{i=1}^N M_i R_i}{N} \neq 1 \to \widehat{\mathcal{K}}(\alpha, 0) \nleq (1-\alpha)\widehat{\mathcal{K}}\left(0, -\alpha/(1-\alpha)\right)$$

where $\widehat{\mathcal{K}}(a, b) = \log\left(\frac{\sum_{i=1}^N \exp(a \log M_i + b \log R_i)}{N}\right)$ is the sample version of the joint cumulant generating function. In other words, the *sample version* of the bound does not provide a bound for the *sample cumulant* of the SDF.

This shows why the approach put forward in Bakshi and Chabi-Yo (2014) does not provide conservative rejection rates. Their approach relies on calculating the sampling distribution (given a

[37]In section 4.3 Politis and Romano require that $\mathcal{B}$ is Fréchet differentiable for some influence function $h$, and that for some $d > 0$ $\mathbb{E}\left[h(X_i)^{d+2}\right] < \infty$ and the data generating process is sufficiently strong mixing.

[38]Recall that we only observe the second marginal of the sample, i.e. the realizations $R_i$.

parametric model) of the sample SDF cumulant (parametric bootstrap) with simulated sample size equal to the available real sample size (generally around 900 monthly observations in US data). More formally, given model $M$ they approximate the distribution of the *model implied values* of the $\alpha$-entropy $\widehat{\mathcal{E}_\alpha} = \frac{\widehat{\mathcal{K}}(\alpha,0)-\alpha\widehat{\mathcal{K}}_i(1,0)}{\alpha(\alpha-1)}$, by drawing independent trajectories of length $N = 900$ and calculating $\mathcal{E}_i = \frac{\widehat{\mathcal{K}_i}(\alpha,0)-\alpha\widehat{\mathcal{K}}_i(1,0)}{\alpha(\alpha-1)}$ for each trajectory $i = 1, ..., 50000$.

They then compare the quantiles of the sampling distribution of $\mathcal{E}_i$ with the sample version of the entropy bound $\mathcal{B}_{sample} = \frac{(1-\alpha)\widehat{\mathcal{K}}(0,-\alpha/(1-\alpha))-\alpha\widehat{\mathcal{K}}(1,0)}{\alpha(\alpha-1)}$, where the latter is calculated on real data. However, since the sample bound $(1-\alpha)\widehat{\mathcal{K}}\,(0, -\alpha/(1-\alpha))$ calculated on real data has no general relation to the real sample SDF cumulant $\widehat{\mathcal{K}}(\alpha, 0)$ (also associated with the real data), we cannot conclude that the model violates SDF bounds if the $\alpha$ quantile of $\mathcal{E}_i$ is greater than the sample upper bound $\mathcal{B}_{sample}$.

In general we could expect that if the sample is long then $\frac{\sum_{i=1}^N M_i R_i}{N} \approx 1$ and the issue is not very serious. Below we show however that if multiple assets are considered and the sample is not very long (in the order of 900 months), then this issue can result in overly many rejections. This is due to the fact that the more assets we consider, the more likely there will be one for which the *sample* pricing relation does not hold very well. The small sample behaviour of the parametric approach therefore might be quite far from being conservative whenever the ratio of monthly observations $N$ to number of test assets $K$ is small, which we verify in a simulation exercise below. Since Bakshi and Chabi-Yo (2014) use as many as 25 portfolios in their tests, we believe that there is no guarantee that their testing approach is conservative. On Figure 12 we can see that in our simulation exercise, even with the market return and only 4 other portfolios the true rejection rate is above 15% for a "conservative test" of size 1%.

### Appendix IV.3 Empirical test

To illustrate the performance of different approaches to testing asset pricing models, we take as our parametric model the LRR model as parametrized in Bansal, Kiku, and Yaron (2012). In their paper they also provide dynamics for dividends for different assets thereby allowing for the pricing of different stock-portfolios. In particular they provide estimates for the aggregate stock-market and four other portfolios distinguishing between small / large/ value / growth type portfolios. The general log dividend dynamics for asset $j$ is given by:

$$\Delta d_{j,t+1} = \mu_j + \phi_j x_t + \psi_j \sigma_t u_{j,t+1} \tag{A.150}$$

where $\phi_j$ is the loading on the long-run consumption growth factor and $u_{j,t+1}$ is a shock that is correlated with instantaneous consumption risk, whilst $\psi_j$ describes the sensitivity of the asset to long run volatility fluctuations. The estimates for the aggregate market and the four portfolios are given in Table VIII of their paper. In particular they show that the model correctly captures the size and value premia.

For our simulation exercise we draw 1000 independent trajectories from this model, where the state variables are initialized at their stationary distributions. Each trajectory is of length 900 which approximately mimicks the length of the available US data. For each trajectory we calculate the time-series of 1 month real bond prices and returns on the aggregate market and each of the four portfolios. Then for each sample we apply each of the three mentioned testing approaches, namely the parametric approach of Bakshi and Chabi-Yo (2014), the percentile and the centered bootstrap. For the non-parametric methods we apply a block-bootstrap approach with $l = 15$ lags and we approximate the distribution of $D^*$ using 1000 resamples.

Figure 12 demonstrates the rejection rates for three different test sizes $\alpha = 1\%, 5\%, 10\%$. The first panel shows the results if only the aggregate market (and the risk-free bond) is required to be correctly priced. The second panel enforces in addition the correct pricing of the 4 additional portfolios.

The results clearly indicate that while for the single asset case all three aproaches are approximately conservative (i.e. they mostly fall below the black dashed line) as they should be, the size distortions become very considerable once more assets are considered. In particular both the parametric bootstrap and the percentile bootstrap imply rejection rates well over the intended probability level. In contrast, the centered (basic) bootstrap correctly maintains a *conservative* size, i.e. the rejection rates are below the predefined test sizes.

The behaviour of the parametric bootstrap (and also of the percentile bootstrap) depends strongly on the ratio of available return observations and number of test assets. We see on Figure 12 that if we would increase the sample size to 5000 monthly observations, then the size of the parametric bootstrap test also improves. However we see that even for 5 test assets we require a sample size that is unavailable even when using US historical data.

All in all we conclude that when multiple assets are used, the bootstrapping technique becomes very relevant and from the three considered approaches only the centered bootstrap provides properly conservative rejection rates of the true model. Therefore in our empirical tests we resort to using this approach to test the models considered.

# Appendix V The Anatomy of Arbitrage Free Dispersion in Long Run Risk Models – Monthly and Quarterly estimates

We make use of the testing framework developed in Section 4.6 in order to systematically characterize the arbitrage free dispersion properties of Bansal and Yaron (2004)-type long-run risk (LRR) models. This model is based on a stochastic discount factor derived from a representative agent with recursive utility $V_t$ given by:

$$V_t = \left[ (1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left( E_t[V_{t+1}^{1-\gamma}] \right)^{1/\theta} \right]^{\frac{\theta}{1-\gamma}}, \tag{A.151}$$

where $C_t$ is consumption at time $t$, $0 < \delta < 1$ the time discounting factor and $\theta = \frac{1-\gamma}{1-1/\psi}$, with $\psi$ the elasticity of intertemporal substitution and $\gamma$ the relative risk aversion parameter. Consumption growth follows the the dynamics:

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \ , \\ x_{t+1} &= \rho x_t + \psi_e \sigma_t e_{t+1} \ , \\ \sigma^2_{t+1} &= \sigma^2_0 + \nu(\sigma^2_t - \sigma^2_0) + \sigma_w w_{t+1} \ . \end{aligned} \tag{A.152}$$

where $(\eta_{t+1}, e_{t+1}, w_{t+1}) \sim IIN(0, I_3)$. We focus on parameterizations implied by recent model calibrations in Bansal, Kiku, and Yaron (2012), according to monthly and quarterly parameterv estimates in Table VI of their paper, and study systematically:

1. Observable dispersion and excess dispersion properties.
2. A broad set of arbitrage free dispersion constraints of Type (1), giving rise to a corresponding upper arbitrage free CGF.
3. A wide family of arbitrage free dispersion bounds.

## Appendix V.1 Dataset

We follow Bansal, Kiku, and Yaron (2012) and rely on US data from January 1946 to December 2012, implying $T = 804$ monthly return observations and $T = 802$ overlapping quarterly observations. Average risk-free returns $R_0$ for monthly and quarterly maturity bonds are computed from the average holding period return of T-Bills having the shortest maturity above one and three months, respectively. In all other cases, quarterly returns are aggregated from monthly returns.

We consider different sets of risky asset returns, based on value-weighted market index returns and four double-sorted Size and Book/Market portfolio returns from Kenneth French's website.[39] We proxy for the return of the real infinite maturity bond using the holding period returns of 30-year nominal bonds in the Fixed Term Indices data set of CRSP. We obtain real risky return series by deflating nominal monthly and quarterly returns with monthly and quarterly realized CPI inflation, respectively. In contrast, we compute the real return on one month, three month and 30-years bonds by deflating nominal returns with data driven proxies of inflation expectations, as in Bansal, Kiku, and Yaron (2012).[40]

We consider three sets of test assets for generating the family of observable points in the joint arbitrage free CGF of pricing kernel components and returns. First, we use the return vector $R := R_A := (R_0, R_M)$ consisting of the T-Bill real return $R_0$ and the risky real return $R_M$ on the value weighted market index (Set A = Mkt + Bond). Second, we consider the return vector $R := R_B := (R_0, R_1, R_2)$, where $R_1$ and $R_2$ are the real returns of two size-sorted portfolios (**S**mall/**L**arge) of returns in the upper book-to-market half (**G**rowth stocks) of the CRSP universe (Set B = S-G + L-G + Bond). The last set of asset returns (Set C = S-G + L-G + S-V + L-V + Bond) consists

[39] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[40] The results using realized rather than expected inflation are identical.

of a return vector $R := R_C := (R_0, R_{11}, R_{12}, R_{21}, R_{22})$, including double-sorted real risky returns $R_{ij}$ $(1 \leq i, j \leq 2)$ with respect to book-to-market (**G**rowth/**V**alue stocks) and size (**S**mall/**L**arge). For portfolio weight vectors $\lambda_A \in \mathbb{R}$, $\lambda_B \in \mathbb{R}^2$ and $\lambda_C \in \mathbb{R}^4$, we then construct portfolio returns $R_{\lambda_A} := R_0 + \lambda_A(R - R_0)$, $R_{\lambda_B} := R_0 + \lambda_{B1}(R_1 - R_0) + \lambda_{B2}(R_2 - R_0)$ and $R_{\lambda_C} := R_0 + \sum_{i,j=1}^{2} \lambda_{Cij}(R_{ij} - R_0)$. These portfolio returns correspond to joint CGFs $\mathcal{K}_{MR_A}$, $\mathcal{K}_{MR_B}$ and $\mathcal{K}_{MR_C}$. They are used to compute on domain $D = (0, 1) \times (0, 1)$ three distinct upper marginal CGFs $\mathcal{K}^U_{M_A}$, $\mathcal{K}^U_{M_B}$ and $\mathcal{K}^U_{M_C}$ of pricing kernel components, by optimizing with respect to $\lambda_A$, $\lambda_B$ and $\lambda_C$ the corresponding convexity bounds in Section 4.6.

## Appendix V.2 Model-Implied Joint CGF of Pricing Kernel Components

We factorize the pricing kernel in the LRR model as $M_{t+1} = M^T_{t+1} M^P_{t+1}$, where permanent component $M^P$ has unit conditional expectation. We follow Alvarez and Jermann (2005) and identify the transient component $M^T$ using the return of infinite maturity bonds: $M^T_{t+1} = 1/R_{\infty t+1}$. The unconditional monthly (quarterly) arbitrage free CGF of $M = (M^T, M^P)$ is our focus and we calculate it by Monte Carlo simulation, by simulating the stationary distribution of state process $\{(c_t, x_t, \sigma^2_t) : t \in \mathbb{N}\}$ in dynamics (A.152), under the monthly (quarterly) parameter estimates in Bansal, Kiku, and Yaron (2012). Along each trajectory, we compute the time series $\{M_{t+1} : t \in \mathbb{N}\}$ based on the model-implied expressions and the returns $\{R_{\infty t+1} : t \in \mathbb{N}\}$ on long term bonds for a sufficiently large maturity. In this way, we obtain the corresponding simulated trajectories for $M^T_{t+1} = 1/R_{\infty t+1}$ and $M^P_{t+1} = M_{t+1}/M^T_{t+1}$.[41] For power arguments $(t, p)$ in domain $D = (0, 1) \times (0, 1)$, we finally calculate by Monte Carlo simulation the CGF values:

$$\mathcal{K}^{\mathbb{M}}_M(p, t) = \log E^{\mathbb{M}} \left[ \left(M^T_{t+1}\right)^t \left(M^P_{t+1}\right)^p \right] , \tag{A.153}$$

where $\mathbb{M}$ emphasizes the model-implied character of this CGF.

Panel "LRR monthly" and Panel "LRR quarterly" of Figure 6 plot the resulting model-implied arbitrage free CGFs for monthly and quarterly parameter estimates, respectively. In each panel, the model-implied CGF features a pronounced convexity along the $p-$axis and a much flatter profile along the $t-$ axis, where the convexity along the $p-$axis is stronger for the quarterly CGF. Consistent with intuition, these CGF convexity properties induce a significant dispersion in the permanent pricing kernel component and a much lower dispersion in the transient pricing kernel component of LRR models. Such model-implied dispersion features have to be consistent with (i) observable dispersion and excess dispersion properties, (ii) arbitrage free dispersion constraints of Type (1) or (2) and (iii) arbitrage free dispersion bounds implied by the arbitrage free CGF in the data.

[41] The yields of discount real bonds are affine functions of the state variables after a log-linearization. These affine functions are calculated iteratively as bond maturity increases. In this way, the return on infinite maturity bonds is easily obtained numerically, avoiding solving 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.

### Appendix V.3 Testing Framework and Empirical Results

We generate observable sets $\mathcal{O}_{\mathcal{K}_{MR_u}}$, $u = A, B, C$, using the following CGF information:

(1) Restriction at the origin: $\mathcal{K}_{MR}(0_{d_1+d_2}) = 0$.

(2) Martingale normalization of $M^P$: $\mathcal{K}_{MR}(e_2) = 0$.

(3) Pricing of short-term bond: $\mathcal{K}_{MR}(e_1 + e_2) = -\log R_0$.

(4) Pricing of risky returns: $\mathcal{K}_{MR}(e_1 + e_2 + e_{d_2+i}) = 0$ for $i = 1, \ldots, d_2$.

(5) Physical risky return observation:

$$\mathcal{K}_{MR}\left(\sum_{i=1}^{d_2} r_i e_{d_1+i}\right) = \log E\left[\prod_{i=1}^{d_2} R_i^{r_i}\right] \ , \tag{A.154}$$

for $r_1, \ldots, r_{d_2} \in \mathbb{R}$.

(5') Physical risky return and long term bond return observation:

$$\mathcal{K}_{MR}\left(te_1 + \sum_{i=1}^{d_2} r_i e_{d_1+i}\right) = \log E\left[R_\infty^{-t}\prod_{i=1}^{d_2} R_i^{r_i}\right] \ , \tag{A.155}$$

using the identity $M^T = 1/R_\infty$, for $t, r_1, \ldots, r_{d_2}, \in \mathbb{R}$.

The difference between assumption (5) and (5') relies in the physical observability of the infinite maturity bond return. When return $R_\infty$ can be assumed to be well measured by the observable real return of long term bonds in the data, then assumption (5') is a convenient one. Alternatively, the weaker assumption (5) can be used. Figure 2 illustrates the observable set $\mathcal{O}_{\mathcal{K}_{MR_A}}$ generated by assumptions (1)-(4), (5') for the tests asset returns $R_A = (R_0, R_M)$.

We study restrictions on the marginal arbitrage free CGF of $(M^T, M^P)$ on domain $D = (0,1) \times (0,1)$. As $D \subset \overline{\mathcal{O}}_{\mathcal{K}_{MR_u}}$ for each $u = A, B, C$, the marginal upper CGFs $\mathcal{K}^U_{M_A}$, $\mathcal{K}^U_{M_B}$ and $\mathcal{K}^U_{M_C}$ are finite on $D$ ad provide useful information about the admissible values of $\mathcal{K}_M$ on this domain.

#### Appendix V.3.1 Omnibus test and dispersion constraints of Type (1)

Following the Type (1) dispersion constraints in the LHS of equation (41), we test the null hypothesis:

$$\mathcal{H}_0 \ : \ \mathcal{K}^{\mathbb{M}}_{M_u}|_D \leq \mathcal{K}^U_{M_u}|_D \ \ , \tag{A.156}$$

for the sets of test asset A, B and C. In order to account for estimation uncertainty and develop an accurate inference, we apply a suitable bootstrap procedure for estimating the corresponding bootstrap confidence intervals about point estimate $\widehat{\mathcal{K}}^U_{M_u}(p,t)$, for each $(p,t) \in D$. Details on this bootstrap procedure are provided in Appendix IV.1.

The estimated upper arbitrage free CGFs for data sets A and C are presented in Figure 6, based on the observability assumptions (1)-(4), (5'). Similar to the model-implied arbitrage free CGFs, both upper CGFs imply a pronounced convexity along the $p-$axis and a flatter profile along the $t-$axis, suggesting a dominating dispersions of permanent relative to transient pricing kernel components. However, shapes and levels of upper CGFs and model-implied CGFs are different in a number of cases, to the point that null hypothesis (A.156) is rejected in some regions of domain $D$.

Bootstrap $p-$values for the test of null hypothesis (A.156) are presented in Panel 13a (13b) of Figure 13 for monthly (quarterly) returns. With set A, the upper CGF is well above the model-implied CGF in virtually all its domain. Test violations arise only in a neighborhood of the $t-$axis ($p = 0$) for monthly returns, indicating that on a monthly frequency violations might arise because the return of infinite maturity bonds might not be well approximated by the return of 30 year bonds. Using set B, we obtain wider regions of rejection at a significance level of 0.05, both for monthly and quarterly data, also in the interior of domain $D$. Such interior violations are dispersion violations generated by an inapproriate convexity of the joint model-implied CGF. For monthly returns, the interior violations are concentrated below the main diagonal in $(t,p)-$coordinates. For quarterly returns, they are mostly concentrated below the antidiagonal in $(t,p)-$coordinates. Finally, based on set C null hypothesis (A.156) is rejected virtually over the entire domain $D$, with only a few exceptions along the $t-$axis for quarterly data.

In the sequel, we study more systematically the anatomy of the violations of dispersion constraints of Type (1) in the LRR model, by studying consistently with the concepts developed in Section 4 the observable dispersion and excess dispersion properties, as well as direct dispersion bounds when they become available.

### Appendix V.3.2 Observable pricing kernel dispersion and excess dispersion

Violations or non violations of Type (1) dispersion constrains in the LRR model are obviously related also to the observable dispersion properties of transient pricing kernel component $M^T$, which is assumed observable in the above tests through the return of the infinite maturity bond. For instance, for any prior $\pi_\alpha$ with support in $(0,1)$ and such that $\pi_\alpha(1) = \alpha \in (0,1)$, the marginal dispersion measure

$$\mathcal{D}_{\pi_\alpha}(M^T) = \frac{E_{\pi_\alpha}\left[\mathcal{K}_{M^T}(t)\right] - \mathcal{K}_{M^T}\left(E_{\pi_\alpha}[t]\right)}{\alpha(1-\alpha)} = \frac{\alpha\mathcal{K}_{M^T}(1) - \mathcal{K}_{M^T}(\alpha)}{\alpha(1-\alpha)} , \tag{A.157}$$

is observable, as $\mathcal{K}_{M^T}(1) = \log E[M^T] = \log E[1/R_\infty]$ is observable under observability assumption (5'). While a geometric interpretation of this dispersion measure is produced in Figure 3, Figure 9 presents results of a direct test of the null hypothesis $\mathcal{D}_{\pi_\alpha}(M^T) = \mathcal{D}^{\mathbb{M}}_{\pi_\alpha}(M^T)$ for different priors $\pi_\alpha$. We find that while the model-implied dispersion for quarterly investment horizons is systematically too low, the model-implied dispersion for monthly horizons is well inside the 95% confidence interval.

According to Proposition 2, the latter marginal dispersion properties are compatible with violations of dispersion constraints of Type (1) along the $t-$axis that derive from an inappropriate scaling of the transient pricing kernel component. We investigate this aspect further in Figure 7, where we plot the monthly marginal arbitrage free CGF in the data and in the model, together with bootstrapped 95% confidence intervals. We find that while the model-implied CGF clearly does not fit within the confidence bounds, a simple (dispersion-invariant) rescaling produces a rescaled CGF well within the bounds.[42] Therefore, the monthly paramterization captures well properties of long-term bond returns after an adjustment of the first moment.

Additional insight into the properties of violations or non violations of Type (1) dispersion constraints is gained by studying the joint observable excess dispersion of transient and persistent pricing kernel components. According to Definition 6, excess dispersion $\Delta\mathcal{J}_{\pi_1,\pi_2}$ is observable when two priors with identical mean and with support in observable set $\mathcal{O}_{\mathcal{K}_{MR}}$ exist. In the LRR model, we obtain a family of observable excess dispersions using priors $\pi_1, \pi_2$ such that:

(1) $\pi_1$ has support in $\{(0,1),(t/(1-p),0)\}$.

(2) $\pi_2$ has support in $\{(1,1),((t-p)/(1-p),0)\}$.

(2) $\pi_1(0,1)=\pi_2(1,1)=p\in(0,1)$.

Such excess dispersions are observable, because $M^T = 1/R_\infty$ under observability assumption (5'). They are explicitly given by:

$$\Delta\mathcal{J}_{\pi_1,\pi_2} = (1-p)\log\left(E\left[R_\infty^{-\frac{t}{1-p}}\right] / E\left[R_\infty^{-\frac{t-p}{1-p}}\right]\right) - p\log B\ . \tag{A.158}$$

For $t=p=1/2$ this excess dispersion is robust with respect to the scale of $R_\infty$ and it corresponds to the measure of co-dispersion in equation (26) of Example 1. Table 2 presents the results of a direct test of the null hypothesis $\Delta\mathcal{J}_{\pi_1,\pi_2} = \Delta\mathcal{J}^{\mathbb{M}}_{\pi_1,\pi_2}$ for this choice of the two priors. We find that while the null hypothesis of identical excess dispersion in the data and the model is not rejected at the 5% significance level for a quarterly horizon, it is rejected for a monthly horizon. As this violation is independent of the scale of $R_\infty$, the monthly parameterization is subject to observable bivariate dispersion misspecifications that are independent of the model's inability to fit the average return on long term bonds.

In summary, the violations of dispersion constraints of Type (1) in the joint marginal CGF of $M^T$ and $M^P$ cannot be corrected by a simple rescaling of the transient pricing kernel component, neither for monthly nor for quarterly horizons, as observable dispersion or excess dispersion violations arise for both settings. In order to investigate the arbitrage free dispersion properties in unobservable regions of domain $D$, we now borrow from Section 4.6.2 and consider direct arbitrage free dispersion bounds for the LRR model.

[42] Rescaling of $M^T$ is equivalent to modifying the model-implied CGF by a linear function that implies $\mathcal{K}^{\mathbb{M}}_{R^\infty}(1) = \mathcal{K}_{R^\infty}(1)$.

### Appendix V.3.3 Marginal lower bounds on pricing kernel dispersion

As illustrated grafically by Figure 10, for any point $(t_0, p_0)$ in domain $D = (0,1) \times (0,1)$ we can obtain multiple dispersion constraints of Type (1), using distinct priors $\pi$ with support in $\mathcal{O}_{\mathcal{K}_M}$ and identical mean $(E_\pi[(t,p)] = (t_0, p_0))$. By construction, these priors only depend on properties of the marginal pricing kernel components. Therefore, dispersion tests based on these priors test the specification of a model using exclusively observable information generated by permanent and transient pricing kernel components. Given the set $\Pi(t_0, p_0)$ of such priors, we then easily obtain the following binding lower bound on the marginal pricing kernel dispersion implied by a prior $\pi_0 \in \Pi(t_0, p_0)$:

$$\begin{aligned}\mathcal{D}_{\pi_0}(M) &= \frac{E_\pi[\mathcal{K}_M(t,p)] - \mathcal{K}(E_\pi[(t,p)])}{tr(Var_{\pi_0}(t,p))} \\ &\geq \frac{E_\pi[\mathcal{K}_M(t,p)] - \inf_{\pi \in \Pi(t_0,p_0)} E_\pi[\mathcal{K}_M(t,p)]}{tr(Var_{\pi_0}(t,p))} =: \mathcal{D}^L_{\pi_0,\Pi(t_0,p_0)}(M) > 0 \ ,\end{aligned} \tag{A.159}$$

where $\mathcal{D}^L_{\pi_0,\Pi(t_0,p_0)}(M)$ is empirically observable and strictly positive if the above infimum is not attained in $\pi_0$. Motivated by the graphical description in Figure 10, we construct set $\Pi(t_0, p_0)$ as follows.

1. For any $(t, p) \in D$ set $\Pi(t_0, p_0)$ includes following priors:
   - A prior $\pi_1$ with support in $\{(1,1), ((t-p)/(1-p), 0)\}$ and such that $\pi_1(1,1) = p$ (Figure 10a).
   - A prior $\pi_2$ with support in $\{(0,1), (t/(1-p), 0)\}$ and such that $\pi_1(0,1) = p$ (Figure 10a).
2. For any $0 < t \leq p < 1$ set $\Pi(t_0, p_0)$ additionally includes following prior:
   - A prior $\pi_3$ with support in $\{(0,0), (0,1), (1,1)\}$ and such that $\pi_1(0,0) = (1-t)$ and $\pi_1(0,1) = t - p$ (Figure 10c).
3. For any $1 > t > p > 1$ set $\Pi(t_0, p_0)$ additionally includes following prior:
   - A prior $\pi_3$ with support in $\{(0,0), (1,0), (1,1)\}$ and such that $\pi_1(0,0) = (1-t)$ and $\pi_1(1,0) = t - p$ (Figure 10b).

Figure 14 illustrates the properties of bound (A.159) when applied to LRR models. For any $(t_0, p_0) \in D$, we select prior $\pi_0 = \arg\sup_{\pi \in \Pi(t_0,p_0)} E_\pi[\mathcal{K}_M(t,p)]$, in order to ensure a non trivial lower dispersion bound (A.159). We then plot in the leftmost panels of Figure 14a and 14b the point estimate for $\mathcal{D}^L_{\pi_0,\Pi(t_0,p_0)}(M)$ using monthly and quarterly data, respectively. The panels in the middle of Figure 14 plot the model-implied dispersion $\mathcal{D}^{\mathbb{M}}_{\pi_0}(M)$ for monthly an quarterly data, which we find to be typically ways larger than the lower dispersion bound for most points in $D$. The large bootstrap $p-$values for a test of the null hypothesis $\mathcal{D}^{\mathbb{M}}_{\pi_0}(M) \geq \mathcal{D}^L_{\pi_0,\Pi(t_0,p_0)}(M)$ in the right panels confirm this impressions, indicating that information from marginal princing kernel components is insufficient to identify violations of dispersion bounds in regions where the arbitrage-free CGF is not fully observable.

### Appendix V.3.4 Joint lower bounds on pricing kernel dispersion

Tighter dispersion bounds than bound (A.159) can be obtained using risky return information and the correponding upper arbitrage-free CGF, when it is available. Indeed, for any point $(t_0, p_0)$ in domain $D = (0,1) \times (0,1)$ we can obtain multiple dispersion constraints of Type (1) also using the joint arbitrage-free CGF of pricing kernels and returns. More precisely, given a prior $\pi_0$ with support in $\mathcal{O}_{\mathcal{K}_{MR}}$ and such that $(t_0, p_0, 0) = E_{\pi_0}[(m,r)]$, tighter dispersion bounds are then directly available. For instance, using prior $\pi_1$ in the previous section, we obtain:

$$\begin{aligned}\mathcal{D}_{\pi_1}(M) &= \frac{E_{\pi_1}[\mathcal{K}_M(t,p)] - \mathcal{K}(t_0,p_0)}{tr(Var_{\pi_1}(t,p))} \\ &\geq \frac{E_{\pi_1}[\mathcal{K}_M(t,p)] - \mathcal{K}_M^U(t_0,p_0)}{tr(Var_{\pi_1}(t,p))} =: \mathcal{D}_{\pi_1}^L(M,R) > 0 \ , \end{aligned} \tag{A.160}$$

where $\mathcal{D}_{\pi_1}^L(M,R)$ is observable empirically. As $(t_0, p_0, 0) \in D$ is in the convex hull of $\mathcal{O}_{\mathcal{K}_{MR}} = \{(t,0,r) : t, r \in \mathbb{R}\} \cup \{(1,1,1),(1,1,0)\}$, $\mathcal{K}_M^U(t_0,p_0)$ is finite and the bound is sharp according to Proposition 5, with an explicit expression for the upper arbitrage free CGF given by:

$$\mathcal{K}_M^U(t_0,p_0) = (1-p)\mathcal{K}_{M^T R}\left(\frac{t-p}{1-p}, -\frac{p}{1-p}\right) = (1-p)\mathcal{K}_{R_\infty R}\left(\frac{p-t}{1-p}, -\frac{p}{1-p}\right) \ . \tag{A.161}$$

The applicability of lower dispersion bound (A.160) depends on whether $R_\infty$ can be assumed to be observable. For monthly data, we verified that the marginal model-implied CGF of $R_\infty$ is well approximated by the marginal CGF of long-bond returns in the data, up to a scaling transformation. Whenever the portfolio generating return $R$ does not contain a position in the long bond, null hypothesis

$$H_0 \ : \ \mathcal{D}_{\pi_1}^{\mathbb{M}}(M) \geq \mathcal{D}_{\pi_1}^L(M,R) \ , \tag{A.162}$$

is invariant to scaling of $R_\infty$ and is therefore a robust null hypothesis for the monthly parameterization of the LRR model.[43] In contrast, as we verified above, observability of $R_\infty$ is a more restrictive assumption for the quarterly parameterization. Thus, for this case we focus on priors $\pi_1$ such that $p = t = \alpha \in (0,1)$, in order to avoid requiring observability of $R_\infty$. The resulting explicit lower dispersion bound,

$$\mathcal{D}_{\pi_1}(M) = \frac{\alpha\mathcal{K}_{MR}(1,1,0) - \mathcal{K}_{MR}(\alpha,\alpha,0)}{\alpha(1-\alpha)} \geq \frac{\alpha \log B - (1-\alpha)\mathcal{K}_R\left(\frac{-\alpha}{1-\alpha}\right)}{\alpha(1-\alpha)} = \mathcal{D}_{\pi_1}^L(M,R) \ , \tag{A.163}$$

generates a testable null hypothesis (A.162) that abstracts from the observability of $R_\infty$.

Figure 11 summarizes the test results of null hypothesis (A.162) on monthly and quarterly data,

[43] This scale invariance of null hypothesis $H_0$ 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 $R_\infty$.

for each dataset A,B and C and for $(t_0, p_0) \in D$. For each dataset, we obtain in Figure 11a a substantial variability of monthly $p-$values over domain $D$. Consistent monthly rejections over the whole domain are obtained for Dataset C, suggesting a general difficulty of the LRR model to fit the value premium at monthly horizons. Dataset B generates both rejections and non rejections of the null at monthly horizons, with rejections being stronger in regions where $t < p < 0.8$, in which the permanent pricing kernel component loads more heavily on dispersion but is not dispersion-exhaustive. In these regions, it appears that the LRR model generates a too low dispersion of the permanent pricing kernel component and/or an insufficient co-dispersion of permanent and transient pricing kernel components. Finally, while the monthly $p-$values of Dataset A also vary a lot, they don't lead to enough power for violating the lower dispersion bound. This low power reflects an insufficiently tight upper CGF bound on $\mathcal{K}_M(t_0, p_0)$ using market return information. The results for quartely lower dispersion bounds in Figure 11b generate a similar evidence as for the monthly dispersion bounds, despite the fact that the quarterly test does not assume observability of return $R_\infty$. We find that in dataset C the dispersion bound is violated with significance 0.05 for $0 < \alpha < 1$. In constrast, Dataset B yiels a rejection for $\alpha = 0.9$ and Dataset A no rejection.

# Appendix VI Multivariate pricing kernel components and co-dispersion bounds

Given $\alpha \in \mathbb{R}^d_{++}$ such that $||\alpha|| < 1$, consider the arbitrage-free dispersion constraint of Type (1):

$$\frac{\mathcal{K}_M(\alpha)}{(\sum_{i=1}^d \alpha_i - 1)\prod_{i=1}^d \alpha_i} \geq -\frac{\mathcal{K}_R\left(-\alpha/(1-\sum_{i=1}^d \alpha_i)\right)}{\prod_{i=1}^d \alpha_i} \ , \tag{A.164}$$

it is easy to formulate this constraint as a constraint on the co-dispersion of the stochastic discount factor components, using the fact that under stochastic independence it follows:

$$\mathcal{K}_M(\alpha) = \sum_{i=1}^d \mathcal{K}_{M_i}(\alpha_i) \ . \tag{A.165}$$

In this way, we obtain the following multivariate co-dispersion measure:

$$\mathcal{C}_\alpha(M) := \frac{\mathcal{K}_M(\alpha) - \sum_{i=1}^d \mathcal{K}_{M_i}(\alpha_i)}{\prod_{i=1}^d \alpha_i} = \frac{1}{\prod_{i=1}^d \alpha_i} \log\left(\frac{E\left[\prod_{i=1} M_i^{\alpha_i}\right]}{\prod_{i=1} E\left[M_i^{\alpha_i}\right]}\right) \ . \tag{A.166}$$

Therefore, from the Type (1) dispersion constraint above:

$$\mathcal{C}_\alpha(M) \leq \frac{1-\sum_{i=1}^d \alpha_i}{\prod_{i=1}^d \alpha_i} \cdot \mathcal{K}_R\left(-\alpha/(1-\sum_{i=1}^d \alpha_i)\right) - \frac{1}{\prod_{i=1}^d \alpha_i}\sum_{i=1}^d \log E\left[M_i^{\alpha_i}\right] \ . \tag{A.167}$$

We can also write this inequality differently, in order to emphasize the contribution to the multivariate dispersion bound of the marginal stochastic discount factor dispersions. Note that:

$$\begin{aligned}
\frac{\mathcal{K}_M(\alpha)}{\prod_{i=1}^d \alpha_i} &= \frac{1}{\prod_{i=1}^d \alpha_i}\left(\log\left(\frac{E\left[\prod_{i=1} M_i^{\alpha_i}\right]}{\prod_{i=1} E\left[M_i^{\alpha_i}\right]}\right) + \log\left(\frac{\prod_{i=1} E[M_i^{\alpha_i}]}{\prod_{i=1} E[M_i]^{\alpha_i}}\right) + \log\left(\prod_{i=1} E[M_i]^{\alpha_i}\right)\right) \\
&= \mathcal{C}_\alpha(M) - \frac{\sum_{i=1}^n \alpha_i(1-\alpha_i)\mathcal{E}_{\alpha_i}(M_i)}{\prod_{i=1}^d \alpha_i} + \frac{\sum_{i=1}^n \alpha_i \log E[B_i]}{\prod_{i=1}^d \alpha_i} \ ,
\end{aligned} \tag{A.168}$$

with the marginal Renyi entropy $\mathcal{E}_{\alpha_i}(M_i)$ of $M_i$ and the zero-bond price $B_i$ implied by stochastic discount factor component $M_i$. Overall, this gives:

$$\begin{aligned}
\mathcal{C}_\alpha(M) - \frac{\sum_{i=1}^n \alpha_i(1-\alpha_i)\mathcal{E}_{\alpha_i}(M_i)}{\prod_{i=1}^d \alpha_i} &= \frac{\mathcal{K}_M(\alpha) - \sum_{i=1}^n \alpha_i \log E[B_i]}{\prod_{i=1}^d \alpha_i} \\
&\leq \frac{(1-\sum_{i=1}^n \alpha_i)\mathcal{K}_R\left(-\alpha/(1-\sum_{i=1}^d \alpha_i)\right) - \sum_{i=1}^n \alpha_i \log E[B_i]}{\prod_{i=1}^d \alpha_i} \ .
\end{aligned} \tag{A.169}$$

A number of special cases are interesting. First, under stationarity of pricing kernels and zero coupon bond prices, it follows:

$$\frac{\sum_{i=1}^n \alpha_i(1-\alpha_i)}{\prod_{i=1}^d \alpha_i}\mathcal{E}_{\alpha_i}(M_1) - \mathcal{C}_\alpha(M) \geq \frac{\left(\sum_{i=1}^n \alpha_i - 1\right)\mathcal{K}_R\left(-\alpha/(1-\sum_{i=1}^d \alpha_i)\right) + \sum_{i=1}^n \alpha_i \log E[B_1]}{\prod_{i=1}^d \alpha_i} \ . \tag{A.170}$$

Second, under independence of single-period returns, it follows:

$$\frac{\sum_{i=1}^n \alpha_i(1-\alpha_i)}{\prod_{i=1}^d \alpha_i}\mathcal{E}_{\alpha_i}(M_1) - \mathcal{C}_\alpha(M) \geq \sum_{i=1}^n \alpha_i \frac{\log E[B_1] - \mathcal{K}_{R_i}\left(-\alpha_i/(1-\sum_{i=1}^d \alpha_i)\right)\left(1-\sum_{i=1}^n \alpha_i\right)/\alpha_i}{\prod_{i=1}^d \alpha_i} \ . \tag{A.171}$$

## 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.

Politis, D. N., and J. P. Romano, 1994, "Limit theorems for weakly dependent Hilbert space valued random variables with applications to the stationary bootstrap," *Statistica Sinica*, 4(2), 461–476.

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. G., and F. Trojani, 2015, "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.

## Swiss Finance Institute

Swiss Finance Institute (SFI) is the national center for fundamental research, doctoral training, knowledge exchange, and continuing education in the fields of banking and finance. SFI's mission is to grow knowledge capital for the Swiss financial marketplace. Created in 2006 as a public–private partnership, SFI is a common initiative of the Swiss finance industry, leading Swiss universities, and the Swiss Confederation.

swiss:finance:institute

c/o University of Geneva, Bd. Du Pont d'Arve 42, CH-1211 Geneva 4
T +41 22 379 84 71, rps@sfi.ch, www.sfi.ch