Anomaly or Possible Risk Factor? Simple-To-Use Tests*

Benjamin Holcblat† Abraham Lioui‡ Michael Weber§

This version: April 15, 2022

*The paper benefitted from discussions with Laurent Barras, Svetlana Bryzgalova, Steffen Grønneberg, Michael Halling, Christian Julliard, Christos Koulovationos, François Le Grand, Marco Lyrio, Thiago de Oliveira Souza, Markus Pelger, Julien Pénasse, Roberto Steri, Raman Uppal, Irina Zviadadze, and seminar/conference participants at CFE London, EDHEC, French Inter Business Schools seminar in Finance (EMLyon) and University of Luxembourg. We also thank Juan Carlos Escanciano for sharing Matlab programs to compute least concave majorants, and Kenneth French, Andrew Y. Chen and Tom Zimmermann for posting their data online. Any errors are our own. Weber also gratefully acknowledges financial support from the University of Chicago Booth School of Business, the Fama Research Fund at the University of Chicago Booth School of Business, and the Fama-Miller Center.

†University of Luxembourg. benjamin.holcblat AT uni.lu

‡EDHEC Business School. abraham.lioui AT edhec.edu

§Booth School of Business, University of Chicago, CEPR, and NBER. michael.weber AT chicagobooth.edu

Anomaly or Possible Risk Factor? Simple-To-Use Tests

Abstract

Basic asset pricing theory predicts high expected returns are a compensation for risk. However, high expected returns might also constitute anomalies due to frictions or behavioral biases. We propose two complementary simple-to-use tests to assess whether risk can explain differences in expected returns. We provide general theoretical equilibrium foundations for the tests and show their properties in simulations. The tests take into account risks disliked by risk-averse individuals, including high-order moments and tail risks. None of the tests rely on the validity of a factor model nor other parametric statistical models. Empirically, we find risk cannot explain a large majority of variables predicting differences in expected returns. In particular, value, momentum, operating profitability, and investment appear to be anomalies.

JEL classification: G12, C58, C38, D53.

Keywords: Cross-section of Returns; Factor Pricing; Strong SSD; Abnormal returns; Market frictions.

1 Introduction

Expected returns reflect and guide investment decisions in the economy (e.g., Cochrane 1996), and hence they are closely related to firms' behavior and aggregate outcomes such as unemployment (Hall 2017, Borovicka and Borovicková 2018). Over the last decades, the literature has identified hundreds of factors predicting cross-sectional returns (Harvey et al. 2016).[1] Kozak et al. (2018), among others, argue that factors' returns might be a compensation for risk (e.g., Berk et al. 1999, Cooper 2006), but may also occur because of behavioral biases (e.g., Bondt and Thaler 1985, Jegadeesh and Titman 1993), institutional frictions (e.g., Gromb and Vayanos 2010, and references therein), informational frictions (e.g., Seyhun 1988, Cohen et al. 2012) and many other frictions.

We propose simple-to-use tests to shed light on the economic content of factors and assess whether risk alone can explain the difference in expected returns generated by a given factor. Researchers and practioners typically build a factor through portfolio sorts based on a given characteristic. They sort stocks according to the value of a characteristic, divide the sorted stocks into groups according to some quantiles (e.g., bottom 30%, middle 40%, top 30%), and then form portfolios based on the groups. If the average returns of the portfolios appear to be monotonic in the characteristic, researchers form a factor by subtracting low-return portfolios from high-return portfolios, so it mimics a long-short strategy. Factors based on multivariate sorting similarly have a long leg with high expected-returns and a short leg with low expected-returns. Basic asset pricing theory stipulates that the higher expected returns of the long leg should correspond to higher risk. Thus, similarly to Kelly et al. (2019), if risk alone cannot explain the spread in expected returns between the two legs of the factor, we call the latter an "anomaly," otherwise we call it a "possible risk factor." In the present paper, we do not use the term "factor" as a shorthand for "risk factor:" A factor can be an anomaly, or a return spread that risk can explain.

Distinguishing between risk factors and anomalies requires a definition of risk. For this purpose, we go back to basic microeconomics and define risk as anything a risk-averse individual dislikes, (i.e., individuals with an increasing and concave von Neumann-Morgenstern utility function). The basic idea behind our two tests is to check whether every possible risk-averse individual strictly prefers the long-leg returns over the short-leg returns. If this is *not* the case, at least one possible risk-averse individual prefers to forego the higher return of the long leg in exchange for the lower, but less risky, return of the short leg. Then, risk can possibly explain the factor's expected return, i.e., the difference in expected returns between the long and the short leg. More precisely, the factor's expected return is a possible compensation for the higher risk of the long leg with respect to the short leg.

The main empirical results of the paper indicate that a majority of factors are anomalies rather than possible risk factors. Regarding the Fama and French (2015) five factors and the

[1] In the following, we use characteristics and factors interchangeably. When we do so, we have variables in mind that help predict returns in the cross section without taking a stance on the validity of a factor model.

momentum factor (Jegadeesh and Titman 1993, Carhart 1997), our tests indicate that value, momentum, operating profitability, and investment are anomalies rather than risk factors. Evidence are mixed regarding size: The null hypothesis is rejected, but it is unclear whether the rejection is due to risk or a lack of a significant factor return. Application of the tests to a standard data set of more than 200 potential factors shows that more than 70% of factors are anomalies, and thus indicate that the main empirical finding holds beyond the widely-used Fama and French (2015) five factors and the momentum factor.

The null hypothesis of the first test corresponds to unconditional strict preferences for the long leg, while the null hypothesis of the second test corresponds to strict preferences for the long leg *conditional* on the market (i.e., after controlling for exposure to market). Because both tests check the strict preference for the long leg for every possible risk-averse individual, the tests do not rely on a specific measure of risk (e.g., variance), nor utility function (e.g., constant relative risk-aversion utility function). In this way, the tests are *comprehensive*, that is, they account for all risks disliked by risk-averse individuals, including high-order moments and tail risks. The tests are also *model-free*, in the sense that they do not assume a parametric model of returns. The large literature has assumed a linear factor model with a specific dependence structure for the errors (e.g., Ross (1976)'s Arbitrage Pricing Theory and its extensions). In particular, our proposed tests do not require us to assume a specific factor model, unlike the literature, which often equates anomalies (or mispricing) and non-zero alphas of regressions of a novel long-short strategy on a specific factor model. Thus, we can define an anomaly as a difference in expected returns that cannot be explained by risk alone, and not as a deviation from a specific factor model that is assumed to capture risk. Another advantage of the unconditional test is the immunity to the multiple hypotheses and pretesting problems: the test does not yield any type I (nor type II error) asymptotically. In other words, as the sample size increases, it is not only impossible to fail to reject a false null hypothesis (type II error), but it is also impossible to wrongly reject a true null hypothesis (type I error).

To formally tie the tests with asset-pricing theory, we also investigate the meaning of their null hypotheses beyond a pairwise comparison of factor legs. The null hypotheses correspond to what we call *strong* Second-order Stochastic Dominance (SSD), which corresponds to SSD with strict inequalities instead of weak inequalities. While the use of strict inequalities should be a mild change in practice, it is key to derive the equilibrium foundations of the tests. In an economy with diversification benefits, spreads in expected returns between two tradable assets should compensate for *non-diversified* risk. We show that if every possible risk-averse individual strictly prefers the returns of the long leg to the returns of the short leg, then non-diversified risk alone is unlikely to explain the factor's expected return, that is, the return spread should exceed any risk compensation individuals require. In line with most of the literature on factor models, for simplicity, we focus on a one-period setting. Nevertheless, we show the equilibrium foundations for both tests remain valid in multiperiod settings. We also demonstrate the equilibrium foundations hold independently of the structure of the economy (e.g., whether or not individuals optimally diversify risk, whether or not markets are complete, whether or not a representative agent exists, etc.). Thus, the theoretical foundations of the proposed tests hold under fairly general assumptions.

To assess the performance of the tests, we investigate their properties mathematically, numerically and empirically. First, building on the statistical and econometric literature on SSD, which goes back at least to McFadden (1989), we show the tests have good asymptotic properties, i.e., they are valid and consistent. Second, we investigate their finite-sample properties through Monte-Carlo simulations. Our simulation results confirm the asymptotic properties of the tests. Finally, as a proof of concept, we apply the unconditional test to the market factor, that is, the spread in expected returns between US stock returns and one-month US Treasury bill returns. Overwhelming empirical evidence exists documenting that US stocks have higher expected returns than Treasury bills, but are riskier. In line with the evidence, the tests clearly indicate that risk can explain the spread, so the market factor clearly appears as a possible risk factor unlike the majority of other factors.

The question of how to interpret factors is not a mere academic curiosity. In many situations, the practical implications of a factor discovery depend on whether it is a risk factor or an anomaly. If a factor corresponds to risk, an individual would likely try to limit her exposure to this factor. Conversely, if a factor corresponds to an anomaly, an individual would likely want to load on it —if possible— and thus earn a higher expected return. Likewise, for investment decisions, firms would likely account for a risk factor to value investment projects, but not necessarily for an anomaly. More generally, unlike an anomaly, a risk factor can typically be used for discounting, which is key both in asset pricing and for real investment decisions. Thus the difference between anomalies and risk factors is also of interest to public authorities in charge of financial markets efficiency, such as the U. S. Securities and Exchange Commission. A public authority is unlikely to want to eliminate a risk factor spread that is a compensation for a fundamental risk, but it would likely want to design policies to eliminate anomalies. For example, targeted advancement of financial literacy and targeted information-disclosure regulations can alleviate a behavioral bias and an informational friction, respectively.

Related literature

To the best of our knowledge, our paper is the first to propose model-free and comprehensive tests to distinguish anomalies from possible risk factors. Nevertheless, it builds on several strands of the literature.

The literature on factor models for the cross-section of stock returns goes back, at least, to the CAPM (Sharpe 1964, Lintner 1965, Mossin 1966), in which differences in exposure to the market return determine differences in expected returns. After some mixed evidence using individual stock returns as test assets (Miller and Scholes 1972), Black et al. (1972), Fama and MacBeth (1973) and others group stocks into portfolios to decrease the idiosyncratic noise, and provide empirical evidence in favor of the CAPM.

However, theoretically, Merton (1973) shows that the market factor does not need to be the only risk factor, and Dybvig and Ingersoll (1982) even show that a CAPM equilibrium can imply the existence of arbitrage opportunities. Empirically, starting with Basu (1977) and Banz (1981), the literature has developed several factor models that attribute important roles to risk factors other than the market factor. Fama and French (1992, 1993)'s three factors plus momentum (Jegadeesh and Titman 1993, Carhart 1997) partly synthesize these early findings.

Since then, exponential growth describes the number of newly discovered factors (Harvey et al. 2016), partially spurred by the availability of better computing power, data mining, and trial and error,[2] econometric advances,[3] and the incorporation of no-arbitrage and equilibrium constraints in statistical linear factor models.[4] Most of the literature focuses on observable factors rather than latent and unobservable factors, a feature our paper shares.

Recent attempts try to "tame" the factor "zoo" (Cochrane 2011) by using novel econometric methods. A first strand of literature proposes to reduce the dimensions of the "zoo" through the extraction of a small number of *unobservable* factors from static or dynamic PCAs.[5] A second strand proposes techniques to infer a parsimonious set of *observable* factors. Barillas and Shanken (2018) and Bryzgalova et al. (2020) develop Bayesian model-selection approaches to select factors. Freyberger et al. (2020) and Feng et al. (2020) adapt LASSO-type of techniques to shrink the number of factors. A third and small strand of literature tries to distinguish risk factors from anomalies. Charoenrook and Conrad (2008) propose conditions for a factor to be a risk factor, and assess them empirically. Pukthuanthong et al. (2018) propose to classify priced factors related to the covariance matrix as risk factors. Kelly et al. (2019) classify factors that corresponds to the exposure to some latent factors as risk factor.

The present paper is closest to this last strand of the literature. The main differences with respect to the latter are the following. (i) Our approach does not specify a specific linear statistical model of returns, which does not necessarily imply no-arbitrage for the set of traded assets.[6] (ii) It detects anomalies instead of risk factors —The rejection of the null hypotheses of our tests indicate a *possible* risk factor. (iii) It evades the Hansen and Richard (1987) critique, i.e., it does not require that conditioning on econometricians' information set and conditioning on individuals' information set coincide.

We also build on a large econometric literature on tests of stochastic dominance. The liter-

[2]See, e.g., McLean and Pontiff (2016), Harvey et al. (2016), Chinco et al. (2021), Chen and Zimmermann (2020b).

[3]See, e.g., Gibbons et al. (1989), Jagannathan and Wang (1998), Sentana and Fiorentini (2001), Stock and Watson (2002), Bai (2003), Todorov and Bollerslev (2010), Doz et al. (2011, 2012), Connor et al. (2012), Kan et al. (2013), Gagliardini et al. (2016, 2019), Forni et al. (2017), Kim and Skoulakis (2018), Raponi et al. (2020), Giglio and Xiu (2020), Uppal et al. (2018), Pelger (2019), Ando and Bai (2020), Lettau and Pelger (2020), Cattaneo et al. (2020).

[4]See, e.g., Ross (1976), Chamberlain and Rothschild (1983), Connor (1984), Milne (1988), Reisman (1988), Al-Najjar (1998), Forni and Lippi (2001), Raponi et al. (2018), Renault et al. (2019)

[5]See, e.g., Connor and Korajczyk (1993), Bai and Ng (2002), Hallin and Liška (2007), Amengual and Watson (2007), Hallin and Liška (2007), Onatski (2009, 2010), Ahn and Horenstein (2013)

[6]Linear factor models do not rule out arbitrage opportunities for observable traded assets, which are necessarily finite (Al-Najjar 1998), see also Dybvig and Ingersoll (1982)

ature mainly builds on McFadden (1989), and includes notable contributions by Davidson and Duclos (2000), Barrett and Donald (2003), Delgado and Escanciano (2012), and Donald and Hsu (2016) among others. Our unconditional test is a subsampling implementation of a modified McFadden (1989) test of SSD. From a technical point of view, it is closest to Linton et al. (2005), although the null hypotheses are different: Our null hypothesis is "the long leg strongly dominates the short leg," whereas applying Linton et al. (2005) to our setting would imply the null hypothesis "the long leg dominates the short leg or the short leg dominates the long leg." Our conditional test is a test of conditional strong SSD. It follows from an application of Durot (2003)'s approach, along the lines of Delgado and Escanciano (2013), and thus adapts the latter to strong SSD. Our block subsampling implementations of the unconditional and conditional tests allow for time-series and cross-sectional dependence.

We also build on a large literature in mathematics on SSD, which goes back to Hardy et al. (1929). Hadar and Russell (1969), Hanoch and Levy (1969), and Rothschild and Stiglitz (1970) introduce and develop SSD methods in economics and finance. Since then, the SSD literature in finance has mainly focused on portfolio allocation or general equilibrium implications of stochastic dominance with recent contributions including Post (2003), Post and Levy (2005), Carlier et al. (2012), Post and Kopa (2017). Recently, Chalamandaris et al. (2019) and Arvanitis et al. (2021), building on Arvanitis et al. (2019) and Scaillet and Topaloglou (2010), propose a method to assess whether adding a factor to a given set of factors is beneficial for every risk-averse investor, and for every prospect investor, respectively. These are spanning tests for factor investing, in the spirit of the previously mentioned strand of literature that tries to infer a parsimonious set of factors. However, they do not allow to distinguish anomalies from possible risk factors. If a given set of factors contains anomalies, then any added factor that is spanned by these anomalies should results in a rejection of their null hypothesis. We contribute to this literature by introducing the concept of *strong* SSD, i.e., the replacement of weak inequalities by strict inequalities in the different characterizations of SSD.[7] As previously mentioned, while it should be a mild modification in practice, the modification is crucial for the equilibrium foundations of the null hypotheses of our tests.

2 Unconditional test

We now develop the unconditional test as well as its equilibrium foundations. For simplicity, we focus on a one-period equilibrium framework and discuss multi-period extensions in Section 4.6.

[7] *Strict* SSD is used to qualify the situation in which all possible risk averse individuals weakly prefer a lottery to another lottery, with a strict preference for some individuals, or equivalently, in which strictly risk averse individuals strictly prefer a lottery to another lottery (Dana 2004, Definition 1). For this reason, we use the term *strong* SSD instead of *strict* SSD.

2.1 Unconditional null hypothesis

A factor typically corresponds to a long-minus-short trading strategy, in which the long leg is a high-expected-returns portfolio and the short leg corresponds to a low-expected-returns portfolio. Thus, the basic idea is to test, for each factor, whether every risk-averse individual would strictly prefer the lottery representing the long leg to the lottery representing the short leg. Accordingly, the null hypothesis of the unconditional test is

$$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)] \tag{1}$$

where $\mathbf{U}_2$ denotes a class of concave and increasing functions, and r_S and r_L denote the returns of the long leg and the short leg, respectively. If the null hypothesis (1) is rejected, then at least one possible risk-averse individual weakly prefers the short leg to the long leg, so risk can possibly explain the spread in expected returns captured by the factor. In other words, an individual who prefers more to less still prefers the short leg because it is less risky than the long leg. Testing for all possible utility functions in $\mathbf{U}_2$ allows us to sidestep the choice of specifying a specific measure of risk, that is, the choice of a specific utility function u.

The null hypothesis (1) is similar to the well-known SSD. The difference comes from the use of strict inequalities instead of weak inequalities, that is, the null hypothesis (1) rules out the possibility of risk-averse individuals who are indifferent between the long and the short leg. Hereafter, when the null hypothesis (1) holds, we say that r_L *strongly* SSD dominates r_S. While the replacement of weak inequalities by strict inequalities is a zero-Lebesgue measure modification, it is key from an economic point of view. SSD is not a sufficient condition for an anomaly for at least two reasons. First, it does not guarantee a strictly positive expected factor return $\mathbb{E}(r_L - r_S)$, which is a necessary condition for the existence of a factor. Second, the modification is key to derive the equilibrium foundations of the test in Section 2.3. If some individuals are indifferent between the long and the short leg, then both legs can coexist in equilibrium, hence no anomaly exists. In fact, any portfolio SSD dominates itself, although it necessarily coexists with itself. In contrast, no portfolio *strongly* SSD dominates itself, because strong SSD is not a reflexive binary relation.

Another way to obtain strict inequalities instead of weak inequalities is to rule out affine utility functions from the class $\mathbf{U}_2$, and rely on *strict* SSD. The latter corresponds to the situation in which all possible strictly risk-averse individuals strictly prefer the dominant lottery (Dana 2004, Definition 1 and strict Jensen's inequality). We do not pursue this path because (i) Risk neutrality (i.e., affine utility functions) is a regular benchmark in finance and economics; (ii) The existence of a strictly positive expected factor return $\mathbb{E}(r_L - r_S)$ is a necessary condition for the existence of an anomaly, so it needs to be part of the null hypothesis.

To derive the testable implications of the null hypothesis (1), the following lemma provides a characterization of strong SSD in terms of cumulative distribution functions (CDFs).

Lemma 1 (Characterizations of strong SSD in terms of CDF)**.** *Assume the support of the random* *variables r_L and r_S is a subset of the interval $[\underline{u}, \overline{u}] \subset \mathbf{R}$ with $\underline{u} \neq \overline{u}$. Denote the left derivative and right derivative of a function $u(.)$ at x with $u'_-(x)$ and $u'_+(x)$, respectively. Define the class $\mathbf{U}_2$ of concave and increasing functions $u : [\underline{u}, \overline{u}] \to \mathbf{R}$ such that (s.t.) there exist $u'_+(\underline{u}) \in \mathbf{R}$ and $u'_-(\check{u}) \in \mathbf{R} \setminus \{0\}$, where $\check{u} \neq \underline{u}$ and $\check{u} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u(x) = 0\}\right\}$.*[8] *Then the following statements are equivalent.*

(i) For all $u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)]$.

(ii) For all $z \in]\underline{u}, \infty[$, $F_L^{(2)}(z) < F_S^{(2)}(z)$, where, $\forall i \in \{H, L\}$, $F_i^{(2)}(z) := \int_{\underline{u}}^{z} (z - x) \mathrm{d}F_i(x)$ denotes the integrated CDF of r_i, with $F_i(.)$ the CDF of r_i.

Proof. See Appendix A.1.1.

□

Well-known estimators of CDFs and functionals thereof exist, so Lemma 1 provides a way to test the null hypothesis (1). Lemma 1 is the strong counterpart of the well-known Hardy-Littlewood et. al. theorem,[9] which has been popularized in economics by Rothschild and Stiglitz (1970). In the present paper, Lemma 1 is mainly used for the same purpose as the Hardy-Littlewood et. al. theorem in the SSD econometric literature.

Despite the appearance, it is not sufficient to replace the weak inequalities in the available proofs of the Hardy-Littlewood et. al. theorem by strict inequalities to prove Lemma 1. The key new ingredient of the proof is the quantity $\check{u}$, which enters in the definition of the class $\mathbf{U}_2$ of concave increasing functions. The restrictions on $\check{u}$ rules out constant functions from the class $\mathbf{U}_2$ —they would imply an equality and thus necessarily violate (1)—, while they allow short-put-payoff-type functions, whose expectations are equal to the integrated CDF. Despite these restrictions, the class $\mathbf{U}_2$ contains all strictly increasing, differentiable, and concave functions on $\mathbf{R}$. In words, the class $\mathbf{U}_2$ is the class of concave, increasing functions differentiable at the minimum $\underline{u}$ of the support and with non-zero left-derivative at the minimum between "absorbing" zeros and the maximum $\overline{u}$ of the support.

A direct consequence of Lemma 1 is the invariance of the null hypothesis (1) under strictly positive affine transformations of lotteries. This implies the formulations of the null hypothesis (1) in terms of terminal wealth, capital gain, gross returns or any other strictly positive affine transformation thereof are all mathematically equivalent, i.e., $\forall u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)] \Leftrightarrow \forall u \in \mathbf{U}_2$, $\mathbb{E}[u(W_0 r_S)] < \mathbb{E}[u(W_0 r_L)] \Leftrightarrow \forall u \in \mathbf{U}_2$, $\mathbb{E}[u(W_0(1 + r_S))] < \mathbb{E}[u(W_0(1 + r_L))]$, where $W_0 > 0$ is the initial wealth of the risk-averse individual.

In addition to Lemma 1, we require the following assumption to obtain a test statistic for the null hypothesis (1).

[8] Concavity only ensures left and right differentiability in the interior $]\underline{u}, \overline{u}[$ (e.g., Aliprantis and Border 2006/1994, Theorem 7.22), so the assumptions of right differentiability at $\underline{u}$ is not subsumed by the concavity assumption.

[9] See, e.g., Hardy et al. (1929, 1934), Blackwell (1951), Sherman (1951), Cartier et al. (1964), Strassen (1965).

Assumption 1. **(a)** *(Common bounded support) The support of the random variables r_L and r_S is $[\underline{u}_r, \overline{u}_r] \subset [\underline{u}, \overline{u}]$, where $\underline{u} = \underline{u}_r$ and $\underline{u} \neq \overline{u}$.* **(b)***(No touching without crossing) If there exists $\dot{z} \in]\underline{u}, \overline{u}]$ s.t. $F_L^{(2)}(\dot{z}) = F_S^{(2)}(\dot{z})$, then there exists $\ddot{z} \in]\underline{u}, \overline{u}]$ s.t. $F_S^{(2)}(\ddot{z}) < F_L^{(2)}(\ddot{z})$.*

Assumption 1(a) is a standard assumption in the econometrics and economic SSD literature and should be "harmless" in practice (McFadden 1989). All observable quantities are necessarily finite because computer memory is finite. Assumption 1(b) "no touching without crossing" should also be harmless in practice. A sufficient condition for the assumption is that zero is not a critical value, that is, the derivative of the function $z \mapsto F_S^{(2)}(z) - F_L^{(2)}(z)$ is non-zero in the level set of 0. The set of critical values of the function $z \mapsto F_S^{(2)}(z) - F_L^{(2)}(z)$ has zero Lebesgue measure following Sard's theorem. Thus, Assumption 1(b) is harmless in practice, although it is crucial for the present paper. Thanks to Assumption 1(b), the null hypothesis (1) does not hold if, and only if, there exists $z \in [\underline{u}, \overline{u}]$ s.t. $F_S^{(2)}(z) < F_L^{(2)}(z)$.

2.2 Unconditional test statistic

We now discuss the asymptotic properties of the unconditional test, study its properties in simulations, and discuss the issues of multiple hypothesis testing and pretesting.

2.2.1 Asymptotic properties

In most statistical tests, the idea is to reject a null hypothesis if the difference between an (unconstrained) estimator and an estimator constrained by the null hypothesis is too large. For example, given a sample $(X_t)_{t=1}^T$ of size T with independent and identically distributed data, the idea behind a t-test with null hypothesis "$\mathrm{H}_0 : \mathbb{E}X_1 = 0$" is to assess whether the difference between the average $\bar{X}_T$ and zero normalized by the standard error $\hat{\sigma}/\sqrt{T}$ (i.e., $\sqrt{T}|\bar{X}_T - 0|/\hat{\sigma}$) is large. If the normalized difference between the (unconstrained) estimator $\bar{X}_T$ and the constrained estimator 0 is beyond a plausible threshold, the null hypothesis "$\mathrm{H}_0 : \mathbb{E}X_1 = 0$" is rejected. In the present paper, both tests follow the same logic.

By Lemma 1, the null hypothesis (1) is equivalent to the null hypothesis

$$\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[, \; F_L^{(2)}(z) - F_S^{(2)}(z) < 0, \tag{2}$$

where $F_L^{(2)}(z)$ and $F_S^{(2)}(z)$ denote the integrated CDF of r_L and r_S, respectively. Moreover, the standard estimator for a CDF is the empirical CDF, so a standard estimator of the integrated CDF $F_L^{(2)}$ is the integrated empirical CDF $\hat{F}_L^{(2)}(z) := \frac{1}{T}\sum_{t=1}^T \mathbb{1}\{r_{L,t} \leqslant z\}(z - r_{L,t})$. Thus, the statistic of the unconditional test is the difference between the *unconstrained* estimator $\hat{F}_L^{(2)}(.) - \hat{F}_S^{(2)}(.)$ and the *constrained* estimator $\min\{\hat{F}_L^{(2)}(.) - \hat{F}_S^{(2)}(.), 0\}$, that is,

$$\begin{aligned}\sqrt{T}\mathrm{KS}_T^* :&= \sqrt{T} \sup_{z \in \mathbf{I}_T} \left| \hat{F}_L^{(2)}(z) - \hat{F}_S^{(2)}(z) - \min\{\hat{F}_L^{(2)}(z) - \hat{F}_S^{(2)}(z), 0\} \right| \\ &= \sqrt{T} \sup_{z \in \mathbf{I}_T} \left| \hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z) \right|, \end{aligned} \tag{3}$$

where $\mathbf{I}_T := [c_T, \overline{u}]$, with $c_T \downarrow \underline{u}$, and where $\hat{F}^{(2)}_{L\wedge S}(z)$ denotes the minimum of the integrated empirical CDF (that is, $\hat{F}^{(2)}_{L\wedge S}(z) = \min\{\hat{F}^{(2)}_L(z), \hat{F}^{(2)}_S(z)\}$).[10] The estimator $\min\{\hat{F}^{(2)}_L(.) - \hat{F}^{(2)}_S(.), 0\}$ is a constrained estimator of $F^{(2)}_L(.) - F^{(2)}_S(.)$, because it satisfies the null hypothesis (2) by construction. It can be shown that the test statistic (3) is related to the one-sided Kolmogorov-Smirnov (KS) type statistics, which has been used in the SSD literature since McFadden (1989).

The following proposition shows the KS^*_T test statistic (3) defines a valid and consistent test of the null hypothesis (1).

Proposition 1 (No type I error and No type II error)**.** *Under Assumption 1 and the Assumptions of Appendix A.2, for any level of the test* $\alpha \in]0,1]$,

(i) if the null hypothesis (1) *holds, then*

$$\lim_{T\to\infty} \mathbb{P}\left(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\right) = 0;$$

(ii) if the null hypothesis (1) *does not hold, then*

$$\lim_{T\to\infty} \mathbb{P}\left(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\right) = 1;$$

where $\hat{c}_{1-\alpha}$ *is the* $1-\alpha$ *quantile of a (centered) block-subsampling approximation of the asymptotic distribution of* $\sqrt{T}\mathrm{KS}^*_T$ *with a block size* b_T *s.t.* $\lim_{T\to\infty} b_T = \infty$ *and* $\lim_{T\to\infty} \frac{b_T}{T} = 0$.

Proof. See Appendix A.2. □

Proposition 1 (i) shows the null hypothesis is asymptotically never rejected when it is true, i.e., no type I error exists, asymptotically. Proposition 1 (i) a fortiori also means the test is valid, that is, the probability of wrongly rejecting a true hypothesis is asymptotically smaller than any level $\alpha \in]0,1]$. Proposition 1 (ii) shows the null hypothesis is rejected with probability one when it is wrong, that is, no type II error exists, asymptotically. In the present paper, we rely on centered and uncentered block subsampling to approximate the distribution of test statistics. Block subsampling implies to draw without replacement matrices $(r_{i,t+1}\ r_{i,t+2}\ \cdots\ r_{i,t+b_T})_{i\in\{L,S\}}$ of b_T consecutive observations of contemporaneous r_L and r_S, instead of any matrix $(r_{i,t_1}\ r_{i,t_2}\ \cdots\ r_{i,t_{b_T}})_{i\in\{L,S\}}$ of b_T observations of r_L and r_S. In this way, block subsampling accounts for potential time-dependence and cross-sectional dependence.

The mathematics behind Proposition 1 are standard. We just need (i) the test statistic (3) to go to zero under the null hypothesis and (ii) the test statistic to diverge under the alternative hypothesis. The crux of the mathematics is the following. Denote with $\mathbf{A}$ the subset of $\mathbf{R}$, in

[10] The absolute value is superfluous in the Kolmogorov-Smirnov (KS) test statistic (3) because, for all $z \in \mathbf{R}$, $0 \leqslant \hat{F}^{(2)}_L(z) - \hat{F}^{(2)}_{L\wedge S}(z)$ by the definition of $\hat{F}^{(2)}_{L\wedge S}(z)$. However, we keep the absolute value to make clear that the KS test statistic (3) measures the distance between the unconstrained estimator $\hat{F}^{(2)}_L$ and the constrained estimator $\hat{F}^{(2)}_{L\wedge S}(z)$.

which the null hypothesis (2) does not hold, that is,

$$\mathbf{A} := \{z \in \mathbf{R} : F_S^{(2)}(z) < F_L^{(2)}(z)\}.$$

Then, addition and subtraction of $F_L^{(2)}(z)$ and $F_{L\wedge S}^{(2)}(z)$ to the quantity maximized by the KS_T^* test statistic (3) yields

$$\begin{aligned}
\sqrt{T}\mathrm{KS}_T(z) &:= \sqrt{T}\{\hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z)\} \\
&= \sqrt{T}\left\{\hat{F}_L^{(2)}(z) - F_L^{(2)}(z) - [\hat{F}_{L\wedge S}^{(2)}(z) - F_{L\wedge S}^{(2)}(z)] + F_L^{(2)}(z) - F_{L\wedge S}^{(2)}(z)\right\} \\
&= \sqrt{T}[\hat{F}_L^{(2)}(z) - F_L^{(2)}(z)] - \sqrt{T}[\hat{F}_{L\wedge S}^{(2)}(z) - F_{L\wedge S}^{(2)}(z)] \\
&\qquad + \sqrt{T}[F_L^{(2)}(z) - F_S^{(2)}(z)]\mathbb{1}_{\mathbf{A}}(z), \qquad (4)
\end{aligned}$$

because, for all $z \notin \mathbf{A}$, $F_L^{(2)}(z) - F_{L\wedge S}^{(2)}(z) = F_L^{(2)}(z) - F_L^{(2)}(z) = 0$.

Under the null hypothesis (2), by the definition of $\mathbf{A}$, $\mathbb{1}_{\mathbf{A}}(z) = 0$, for all $z \in \mathbf{R}$. Thus, for T big enough, with probability one,

$$\begin{aligned}
\sqrt{T}\mathrm{KS}_T(z) &= \sqrt{T}[\hat{F}_L^{(2)}(z) - F_L^{(2)}(z)] - \sqrt{T}[\hat{F}_{L\wedge S}^{(2)}(z) - F_{L\wedge S}^{(2)}(z)] \\
&= \sqrt{T}[\hat{F}_L^{(2)}(z) - F_L^{(2)}(z)] - \sqrt{T}[\hat{F}_L^{(2)}(z) - F_L^{(2)}(z)] = 0,
\end{aligned}$$

because $F_{L\wedge S}^{(2)}(.) = F_L^{(2)}(.)$, and a Law of Large Numbers (LLN) implies the uniform convergence of $\hat{F}_L^{(2)}(z) := \frac{1}{T}\sum_{t=1}^{T}\mathbb{1}\{r_{L,t} \leqslant z\}(z - r_{L,t})$ and $\hat{F}_S^{(2)}(z) := \frac{1}{T}\sum_{t=1}^{T}\mathbb{1}\{r_{S,t} \leqslant z\}(z - r_{S,t})$ to $F_L^{(2)}(z) := \mathbb{E}[\mathbb{1}\{r_{L,t} \leqslant z\}(z - r_{L,t})$ and $F_S^{(2)}(z) := \mathbb{E}[\mathbb{1}\{r_{S,t} \leqslant z\}(z - r_{S,t})$, so $\hat{F}_{L\wedge S}^{(2)}(z) = \hat{F}_L^{(2)}(z)$ for T big enough. Thus, $\sqrt{T}\mathrm{KS}_T^*$ is asymptotically smaller than any positive quantity, so $\mathbb{P}\big(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}_T^*\big)$ goes to zero, as $T \to \infty$. If the null hypothesis (2) does not hold, $\sqrt{T}[\hat{F}_L^{(2)}(z) - F_L^{(2)}(z)] = \sqrt{T}\big[\frac{1}{T}\sum_{t=1}^{T}\mathbb{1}\{r_{L,t} \leqslant z\}(z - r_{L,t}) - \mathbb{E}[\mathbb{1}\{r_{L,t} \leqslant z\}(z - r_{L,t})]\big]$, which, by a Central Limit Theorem (CLT), converges to a tight limit after multiplication by $\sqrt{T}$. Similarly, by the continuous mapping theorem $\sqrt{T}[\hat{F}_{L\wedge S}^{(2)}(z) - F_{L\wedge S}^{(2)}(z)] = O_P(1)$. However, for all $z \in \mathbf{A}$, $\sqrt{T}[F_L^{(2)}(z) - F_S^{(2)}(z)]\mathbb{1}_{\mathbf{A}}(z) \to \infty$, as $T \to \infty$. Therefore, under the alternative hypothesis, as $T \to \infty$, the KS_T^* test statistic (3), which maximizes (4), goes to infinity, and thus becomes bigger than any threshold $\hat{c}_{1-\alpha}$.

2.2.2 Monte-Carlo Simulations

We find in Monte-Carlo simulations in Table 1 that the finite-sample properties of the test statistic KS_T^* are in line with Proposition 1. For all data-generating processes (DGP), p-values goes to zero when the null hypothesis (2) is wrong. Also, in line with the asymptotic theory, a large and growing proportion of p-values equals one, when the null hypothesis (1) holds, because of the absence of type I error, asymptotically. The first two DGPs are Gaussian distributions calibrated to data. More precisely, they are calibrated to two factors —size and the dividend yield— for which the null hypotheses are barely true (or false) in order to be challenging for the test. The third DGP is a stylized DGP except for the correlation between the long leg and the short leg. The latter correlation is calibrated to the average correlation of the legs of some of the most prominent factors. Further simulation results and details are available in Appendix B.

One insight from the simulations is that centered block subsampling tends to yield more rejections than uncentered block subsampling approximations. Hence, to be conservative, we use the centered subsampling approximation in our empirical implementation: Centered block subsampling should play against the main empirical result of the paper. In Section 4.2, we also investigate the finite-sample properties of the tests on actual financial data.

2.2.3 Immunity to multiple hypothesis testing and pretesting

Because of the large number of factors considered in the literature, Harvey et al. (2016) among others raise the concern of multiple hypothesis testing. The multiple hypothesis problem stems from the high probability of wrongly rejecting at least one true hypothesis, if one simultaneously tests many true hypotheses with size and level of each test exactly equal to $\alpha \in]0,1]$. E.g., by definition of the asymptotic size of a test, if one simultaneously and independently tests 100 true hypotheses at size $\alpha = 5\%$, one expects to wrongly reject 5 true hypotheses, asymptotically. The following Proposition 2 shows the unconditional test is immune to the multiple hypothesis problem.

Proposition 2 (Immunity to multiple hypothesis testing)**.** *Define a family* $(\mathrm{H}_{0,k})_{k=1}^{K}$ *of null hypotheses s.t.* $\mathrm{H}_{0,k} : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(r_{k,S})] < \mathbb{E}[u(r_{k,L})]$*, where* $r_{k,S}$ *and* $r_{k,L}$ *denote the return of the short leg and the long leg of the factor* k*. Define the set* $\mathbf{J} \subset [\![1,K]\!]$ *of true hypotheses. Under the assumptions of Proposition 1, the asymptotic family-wise error rate (FWER) is zero, i.e.,*

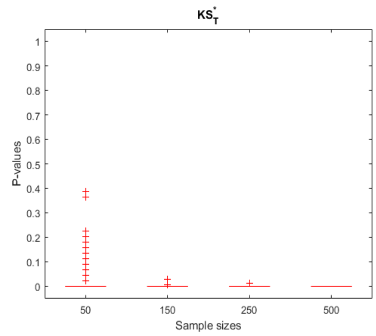
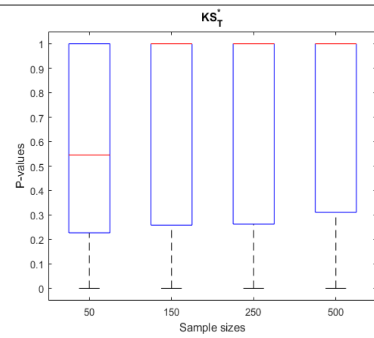
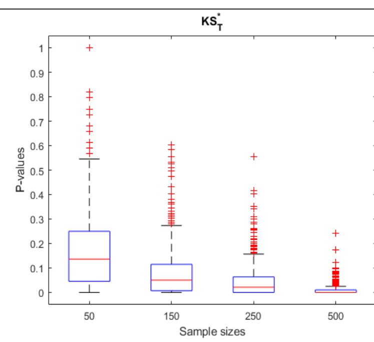
$$\lim_{T\to\infty} \mathbb{P}\left\{\exists j \in \mathbf{J} \ s.t.\ \hat{c}_{j,1-\alpha} < \sqrt{T}\mathrm{KS}^*_{j,T}\right\} = 0,$$

where $\mathrm{KS}^*_{j,T}$ *is the unconditional test statistic* (3) *that corresponds to the null hypothesis* $\mathrm{H}_{0,j}$ *and* $\hat{c}_{j,1-\alpha}$ *the* $1-\alpha$ *quantile of a (centered) block-subsampling approximation of the asymptotic distribution of* $\sqrt{T}\mathrm{KS}^*_{j,T}$ *with a block size* b_T *s.t.* $\lim_{T\to\infty} b_T = \infty$ *and* $\lim_{T\to\infty} \frac{b_T}{T} = 0$.

Proof. By positivity and additivity of probability measures, $0 \leqslant \mathbb{P}\{\exists j \in \mathbf{J} \text{ s.t. } \hat{c}_{j,1-\alpha} < \sqrt{T}\mathrm{KS}^*_{j,T}\}$ $= \mathbb{P}\left\{\bigcup_{j\in\mathbf{J}}\{\hat{c}_{j,1-\alpha} < \sqrt{T}\mathrm{KS}^*_{j,T}\}\right\} \leqslant \sum_{j\in\mathbf{J}} \mathbb{P}\{\hat{c}_{j,1-\alpha} < \sqrt{T}\mathrm{KS}^*_{j,T}\}$. Now, by Proposition 1i, we know $\lim_{T\to\infty}\sum_{j\in\mathbf{J}} \mathbb{P}\{\hat{c}_{j,1-\alpha} < \sqrt{T}\mathrm{KS}^*_{j,T}\} = 0$, so the result follows from the squeeze theorem. □

Proposition 2 stipulates that the probability of wrongly rejecting at least one true hypothesis (that is, the FWER) is close to zero for a sufficiently large sample size. As the proof shows, Proposition 2 is an immediate consequence of Proposition 1(i), which implies a zero probability of rejecting a true hypothesis, asymptotically. Proposition 2 shows the unconditional test satisfies stronger properties than asymptotic t-tests corrected for multiple hypothesis testing: Usual

Table 1: Performance of unconditional test in Monte-Carlo simulations

<table>
<tr><th>H_0</th><th>DGP</th><th>Boxplots of p-values</th></tr>
<tr><td>False</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} .015 \\ .0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>True</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} .011 \\ .010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>False</td><td>$\begin{cases} r_L \overset{IID}{\hookrightarrow} \mathrm{t}(4) \\ r_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\mathrm{or}(r_S, r_L) = .7 \end{cases}$</td><td>
</td></tr>
</table>

Note: The first two data-generating processes (DGP) correspond to Gaussian distributions calibrated to factors for which H_0 appears barely true (or false). The third DGP is a stylized DGP except for the correlation that is calibrated to data. The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of KS^*_T is approximated through centered block subsampling with block size $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

multiple hypothesis procedures for t-tests bound from above the false discovery rate (FDR), which is a less stringent criterion than FWER (e.g., Lehmann and Romano 2006).

While Proposition 2 is stronger than the property of usual multiple hypothesis testing techniques, it does not address the deeper problem of pretesting. In the context of t-tests, the pretesting problem is the following. The classical theoretical justification of an asymptotic t-test of size α is the t-statistic has a probability $1-\alpha$, asymptotically, to be between the $\alpha/2$ and $1-\alpha/2$ quantiles of a standard Gaussian distribution under the test hypothesis. However, once computed, the t-statistic is in the non-rejection region with probability 0 or 1, that is, it either *is* or it is *not* in the non-rejection region. Thus, if the result of this first test leads an econometrician to implement a second t-test of size α, the corresponding t-statistic does not typically have a probability of $1-\alpha$ asymptotically to be between the $\alpha/2$ and $1-\alpha/2$ quantiles of a standard Gaussian distribution under the test hypothesis. The observation of the first t-statistic has removed a part of the randomness of the second t-statistic. Except in specific cases, statistics based on the same data set are not independent. Hence, the classical theoretical justification does not hold for the second t-test. In fact, the econometrician would need to use the asymptotic distribution of the second t-statistic conditional on the result of the first t-statistic, and it is generally a difficult task to derive such a distribution. The pretesting problem is even more difficult because the econometrician would not only need to condition on the result of the last t-test but on all previous knowledge about the data (e.g., plots of the data, descriptive statistics, prior model selections etc.). Because of a lack of a general solutions to the pretesting problem, it is typically ignored, that is, the econometrician typically proceeds as if they had chosen the test to be implemented before any examination of the data. Multiple hypothesis testing techniques do not tackle the pretesting problem because they assume that the list of all statistics to be potentially computed is determined before any examination of the data. The latter assumption is difficult to defend in the case of factor discovery: The evolution of cross-sectional asset pricing is a hard-to-predict dialog between theory and many empirical studies. The following Proposition 3 shows the unconditional test is immune to the pretesting problem.

Proposition 3 (Immunity to pretesting)**.** *Under the assumptions of Proposition 1, for any sequence of events* $\{F_T\}_{T\in\mathbf{N}}$*,*

$$\lim_{T\to\infty}\mathbb{P}\left(\{\hat{c}_{1-\alpha}<\sqrt{T}\mathrm{KS}_T^*\}\cap F_T\right)=\lim_{T\to\infty}\mathbb{P}(\hat{c}_{1-\alpha}<\sqrt{T}\mathrm{KS}_T^*)\mathbb{P}\left(F_T\right).$$

Proof. See Appendix A.3. □

Proposition 3 shows the unconditional test is independent of any sequence of events $\{F_T\}_{T\in\mathbf{N}}$ as the sample size increases. Thus, conditioning on prior knowledge of the data is irrelevant for a sufficiently large sample size. It also means that conditioning on the result of the unconditional test is also irrelevant for further inference. To the best of our knowledge, only a few known inference procedures with such property exist (e.g., Hannan and Quinn 1979). Like Proposition 2, Proposition 3 is a direct consequence of Proposition 1.

2.3 Equilibrium Foundations for Unconditional Test

In the present section, we show that, under general assumptions, the null hypothesis (1) should be a sufficient condition for an anomaly.

2.3.1 Equilibrium Foundations without Diversification Benefits

In the absence of diversification benefits, the equilibrium implication of the null hypothesis (1) is immediate. Assume individuals have to invest all wealth either in the short leg, or in the long leg —exclusive or— so no diversification benefits exist. Assume all possible individuals have strictly increasing von Neumann-Morgenstern utility functions in $\mathbf{U}_2$. If the returns of the long leg are strictly preferred by all possible individuals to the returns of the short leg, then by the invariance of the null hypothesis under strictly positive affine transformations of lotteries (Lemma 1)

$$\begin{aligned} & \mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)] \\ \Leftrightarrow\ & \mathbb{E}[u(1+r_S)] < \mathbb{E}[u(1+r_L)] \\ \Leftrightarrow\ & u^{-1}\Big(\mathbb{E}[u(1+r_S)]\Big) < u^{-1}\Big(\mathbb{E}[u(1+r_L)]\Big), \end{aligned}$$

where $u^{-1}\Big(\mathbb{E}[u(1+r_S)]\Big)$ and $u^{-1}\Big(\mathbb{E}[u(1+r_L)]\Big)$ are private values —the certainty equivalents— of the short and long leg gross returns, respectively. In words, all possible risk averse individuals value the long leg gross returns strictly higher than the short leg gross returns. Now, by definition for gross returns, the market price of both the gross short leg $(1 + r_S)$ and of the gross long leg $(1 + r_S)$ is 1\$. Thus, every individual tries to buy the long leg so the relative price of the long leg relative to the short leg increases and its returns decrease up to a point where some individuals are indifferent between the two. At the equilibrium, the long leg cannot be strictly preferred by all individuals. Therefore, if the null hypothesis (1), —or equivalently "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2$, $\mathbb{E}[u(1 + r_S)] < \mathbb{E}[u(1 + r_L)]$"— holds, an anomaly exists.

2.3.2 Equilibrium Foundations with Diversification Benefits

In an economy with several assets, the aforementioned equilibrium implication does not necessarily hold because individuals do not have to choose one among two assets. Individuals can combine assets into portfolios, so the idiosyncratic risk of different assets can cancel out through diversification. Then, the remaining non-diversified risk corresponds to the movement of individuals' wealth, so the priced risk corresponds to the positive comovements of the factor return with individuals' wealth.

Nevertheless, the present section shows the null hypothesis (1) "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)]$" should still be a sufficient condition for an anomaly in the presence of diversification benefits. More precisely, we show the null hypothesis (1) implies that, up to a first order, the expected return of the factor cannot be explained by risk alone, that is, it exceeds the risk compensations required by risk-averse individuals. For this purpose, we first need to derive the factor risk compensations under general assumptions. The assumptions should be as general as possible to the extent they do not allow for behavioral biases nor frictions affecting the expected return of the factor. We want risk compensations and not compensations for frictions or behavioral biases. Thus, the question is to identify a parsimonious combination of ingredients that are sufficient to derive the factor risk compensations. The following simple derivation shows that it is sufficient to consider a situation in which individuals optimally and freely trade the factor in a neighborhood of their locally optimal terminal wealth. Importantly, we do not need to specify a model, that is, we can do "something without having to do everything." (Hansen 2013).

Derivation of Risk Compensation

By construction, a factor $r_L - r_S = (1 + r_L) - (1 + r_S)$ is a costless portfolio, because it implies buying 1$ of the long leg and selling 1$ of the short leg. Thus, for any individual, irrespective of budget constraints, as long as the factor freely trades in a neighborhood of the locally optimal terminal wealth of the individual, the expected marginal value of the factor is zero, that is,

$$\mathbb{E}[u'(W_1)(r_L - r_S)] = 0, \tag{5}$$

where $u(.)$ and W_1 denote, respectively, individual's utility function and terminal wealth. The logic behind the standard optimality condition (5) is the following. If $\mathbb{E}[u'(W_1)(r_L - r_S)] > 0$ (respectively $0 > \mathbb{E}[u'(W_1)(r_L - r_S)]$), one more (respectively less) marginal unit of the costless portfolio $r_L - r_S$ would increase individual's utility. See Appendix A.4 for a formal proof under general assumptions.

By the optimality condition (5), $\mathbb{C}\text{ov}(u'(W_1), r_L - r_S) + \mathbb{E}[u'(W_1)]\mathbb{E}(r_L - r_S) = 0$, so the expected return of the factor explained by risk alone is

$$\mathbb{E}(r_L - r_S) = -\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\text{ov}(u'(W_1), r_L - r_S). \tag{6}$$

In words, the expected return of the factor $\mathbb{E}(r_L - r_S)$ should be the opposite of its covariance with individuals' marginal utility normalized by individuals' expected marginal utility, that is, $-\frac{1}{\mathbb{E}[u'(W_1)]}$ $\times\, \mathbb{C}\text{ov}(u'(W_1), r_L - r_S)$. Hence, the expected return of the factor $\mathbb{E}(r_L - r_S)$ should exactly compensate for its normalized negative comovements with the marginal utility of terminal wealth W_1, and thus for its normalized positive comovements with terminal wealth W_1 —the marginal utility function $u'(.)$ is decreasing due to concavity.

Our derivation of equation (6) does not require us to specify an equilibrium model. As previously mentioned, the optimality condition (5), and thus equation (6) holds as long as individuals can freely trade the costless portfolio $r_L - r_S$ in a neighborhood around their locally optimal terminal wealth W_1. Thus, the quantity $-\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\text{ov}(u'(W_1), r_L - r_S)$ should be the risk compensation for any one-period equilibrium model. In other words, in any equilibrium model, whether partial equilibrium or general equilibrium, whether with production or not, whether with complete or incomplete financial markets etc., the risk compensation is given by the right-hand side of equation (6). If a wedge exists between the expected return of the factor $\mathbb{E}(r_L - r_S)$ and the risk compensation $-\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\text{ov}(u'(W_1), r_L - r_S)$, an explanation other than risk is needed to account for the expected return of the factor $\mathbb{E}(r_L - r_S)$. Moreover, the derivation of equation (6) indicates that the other explanation should be a friction or a behavioral bias that induces a violation of the optimality condition (5). Hence, an informational friction or a trading friction on the factor can be an explanation, but a friction on production or even a short-sale constraint on a asset that is not part of the factor cannot be an explanation.

The Null hypothesis (1) and Risk Compensations

The following proposition shows that if the null hypothesis (1) holds, then the expected return of the factor $\mathbb{E}(r_L - r_S)$ should exceed the risk compensation $-\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\text{ov}(u'(W_1), r_L - r_S)$ for a large class of increasing and concave utility functions.

Proposition 4 (Equilibrium foundation for unconditional test)**.** *For all $u \in \mathbf{U}_2$ s.t. u is twice continuously differentiable on $[\underline{u}, \overline{u}]$, which includes the support of W_1 and of the returns r_S and r_L, then, up to a first order, the null hypothesis "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)]$" implies the expected return of the factor exceeds its risk compensation, i.e.,*

$$-\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\text{ov}(u'(W_1), r_L - r_S) < \mathbb{E}(r_L - r_S).$$

Proposition 4 provides sufficient assumptions under which strict preference for the long leg implies that risk alone cannot explain the factor's expected return $\mathbb{E}(r_L - r_S)$, up to a first order. If risk alone cannot explain the factor's expected return, other explanations, such as behavioral biases or institutional frictions, are necessary to explain the factor's expected return, and thus we call the factor an anomaly. Assumptions underlying Proposition 4 are mild. They hold for any twice continuously differentiable strictly increasing and concave utility function on $[\underline{u}, \overline{u}]$. The assumption $\mathbb{P}(u'(W_1) > 0) > 0$, which necessarily holds for strictly increasing differentiable utility functions, ensures that $\mathbb{E}[u'(W_1)] > 0$. As previously explained for equation (6), the assumptions do not require us to specify a DGP for returns, nor an economy. If we were to specify the latter, it would need to generate the exact same returns as the observed returns and thus it would not matter for the test. The proof of Proposition 4 essentially only requires Taylor expansions. Because of the simplicity of the proof, we provide it in the main text below.

Proof of Proposition 4. Two first-order Taylor expansions of $u(.)$ around W_1 yield[11]

$$
\begin{aligned}
& \mathbb{E}[u(r_L + \mathbb{E}W_1) - u(r_S + \mathbb{E}W_1)] \\
= & \ \mathbb{E}\left[u(W_1) + u'(W_1)\big(r_L + \mathbb{E}(W_1) - W_1\big) + o(\epsilon_L)\right. \\
& \qquad \left. - u(W_1) - u'(W_1)\big(r_S + \mathbb{E}(W_1) - W_1\big) + o(\epsilon_S)\right] \text{ where } \epsilon_i := r_i + \mathbb{E}(W_1) - W_1,\ \forall i \in \{L, S\}; \\
= & \ \mathbb{E}\left[u'(W_1)(r_L - r_S) + o(\epsilon_L) + o(\epsilon_S)\right], \qquad (7)
\end{aligned}
$$

where the invariance of the null hypothesis (1) under strictly positive affine transformations of lotteries (Lemma 1) implies $0 < \mathbb{E}[u(r_L + \mathbb{E}W_1) - u(r_S + \mathbb{E}W_1)]$.

Thus, up to a first order,

$$
\begin{aligned}
& 0 < \mathbb{E}\left[u'(W_1)(r_L - r_S)\right] = \mathbb{C}\mathrm{ov}(u'(W_1), r_L - r_S) + \mathbb{E}[u'(W_1)]\mathbb{E}(r_L - r_S) \\
\Leftrightarrow & -\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\mathrm{ov}(u'(W_1), r_L - r_S) < \mathbb{E}(r_L - r_S).
\end{aligned}
$$

□

3 Test Conditional on the Market

As its name indicates, the unconditional test relies on the unconditional distribution of returns. However, practitioners—probably inspired by the CAPM—usually analyze returns after controlling for exposure to market risk. For this reason, we propose a test conditional on the market. The present section follows the same structure as the previous section. We first present the test conditional on the market returns and then its equilibrium foundations.

3.1 Null Hypothesis Conditional on the Market

The null hypothesis of the test conditional on the market is the same as for the unconditional test, except that it controls for the market return r_M. The idea is to test, for each factor, whether every possible risk-averse individual would strictly prefer the long-leg lottery to the short-leg lottery conditional on the market, that is,

$$
\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(r_S)|r_M] < \mathbb{E}[u(r_L)|r_M], \qquad (8)
$$

where r_M denotes the market return.

As previously mentioned, the main motivation for the null hypothesis (8) relative to the null hypothesis (1) of the unconditional test is practitioners' routine of controlling for the market through a regression with the market (excess) returns as an explanatory variable. In this way,

[11] Although the proof is based on Taylor expansions, preferences are *not* implicitly assumed risk neutral nor mean-variance because (i) The Taylor expansions are made around the terminal wealth W_1, which is random, instead of around expected quantities; (ii) The first-order term $u'(W_1)(r_L - r_S)$ exactly corresponds to the non-diversified risk as the derivation of equation (6) shows.

practitioners control for affine functions of the market returns. The test conditional on market does not only control for affine functions of market returns, but for all measurable functions of market returns. Moreover, it should not matter whether we use market returns, or the latter in excess of the risk-free rate: Conditioning on r_M, or conditioning on $r_M - r_f$ does not matter because they generate the same σ-algebra.

As for the unconditional test, a characterization of strong conditional SSD in terms of CDFs is necessary to bring the null hypothesis (8) to the data.

Lemma 2 (Characterization of conditional strong SSD in terms of CDF)**.** *Assume a complete probability space. Under Assumption 1(a), the following statements are equivalent.*

(i) *For all* $u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)|r_M] < \mathbb{E}[u(r_L)|r_M]$ *almost surely (a.s.).*

(ii) *For all* $z \in]\underline{u}, \infty[$, $F^{(2)}_{L|M}(z|r_M) < F^{(2)}_{S|M}(z|r_M)$ *a.s., where* $F^{(2)}_{L|M}(z|r_M) := \int_{\underline{u}}^{z} F_{L|M}(y|r_M)\mathrm{d}y$ *a.s.*

Proof. See Appendix A.1.2. □

Lemma 2 is the conditional counterpart of Lemma 1. Similarly to Lemma 1 for the null hypothesis (1), Lemma 2 implies the invariance of the null hypothesis (8) under strictly positive affine transformations of lotteries. In particular, the lemma implies that it does not matter whether we consider the leg's returns, or —if inspired by the CAPM— we consider the latter in excess of the risk-free rate, i.e., $\forall u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)|r_M] < \mathbb{E}[u(r_L)|r_M] \Leftrightarrow \forall u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S - r_f)|r_M] < \mathbb{E}[u(r_L - r_f)|r_M]$. As for the unconditional test, a conditional counterpart of the assumption "no touching without crossing" is necessary to bring the null hypothesis (8) to the data.

3.2 Test Statistic Conditional on the Market

By Lemma 2, the hypothesis (8) is equivalent to the null hypothesis

$$\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[,\ F^{(2)}_{L|M}(z|.) - F^{(2)}_{S|M}(z|.) < 0, \tag{9}$$

where $F^{(2)}_{L|M}(z|x)$ and $F^{(2)}_{S|M}(z)$ denote the integrated CDF of r_L and r_S conditional on r_M, respectively. We cannot follow the same approach as for the unconditional test in Section 2, because conditional empirical CDFs do not follow functional CLTs. Thus, we follow Durot (2003)'s approach along the lines of Delgado and Escanciano (2013), and adapt the latter to strong SSD. The key idea is to express the null hypothesis (9) in terms of the concavity of the second-order antiderivative of the difference of integrated conditional CDF.

Under standard regularity conditions, a function is strictly negative if, and only if, its first-order antiderivative is strictly decreasing, and if, and only if, its second-order antiderivative (i.e., the antiderivative of the antiderivative of the function) is strictly concave. Thus, the null hypothesis (9) is equivalent to the null hypotheses

$$\begin{aligned}
&\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[, \int_{-\infty}^{\cdot} [F^{(2)}_{L|M}(z|\dot{x}) - F^{(2)}_{S|M}(z|\dot{x})] f_X(\dot{x}) \mathrm{d}\dot{x} = F^{(2)}_{L,M}(z,.) - F^{(2)}_{S,M}(z,.) \text{ strictly decreasing} \\
&\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[, C^{(2)}(z,.) \text{ is strictly concave,} \qquad (10)
\end{aligned}$$

where, for all $z \in \mathbf{R}$, $C^{(2)}(z,.)$ denotes a normalized antiderivative of $F^{(2)}_{L,M}(z,x) - F^{(2)}_{S,M}(z,.)$. An unconstrained estimator of $C^{(2)}(z,.)$ is the antiderivative $\hat{C}^{(2)}(z,.)$ of the integrated empirical CDF. A constrained estimator of $C^{(2)}(z,.)$ is the smallest concave majorant $\mathcal{T}\hat{C}^{(2)}(z,.)$ of $\hat{C}^{(2)}(z,.)$ because the smallest concave majorant (also called least-concave majorant) of a concave function is the concave function itself. Therefore the test statistic is

$$\sqrt{T}\mathrm{C}^*_T := \sqrt{T} \sup_{(z,u)\in]\underline{u},\infty[\times \hat{F}_M([\underline{u}_M,\overline{u}_M])} |\mathcal{T}\hat{C}^{(2)}(z,u) - \hat{C}^{(2)}(z,u)|,$$

where $[\underline{u}_M, \overline{u}_M]$ denotes the support of r_M. The following proposition shows the C^*_T test statistic defines a valid and consistent test.

Proposition 5 (Validity and consistency). *Under the Assumption 1 and the assumptions of Appendix A.6,*

(i) if the null hypothesis (8) holds, then

$$\lim_{T\to\infty} \sup \mathbb{P}\left(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{C}^*_T\right) \leqslant \alpha;$$

(ii) if the null hypothesis (8) does not hold, then

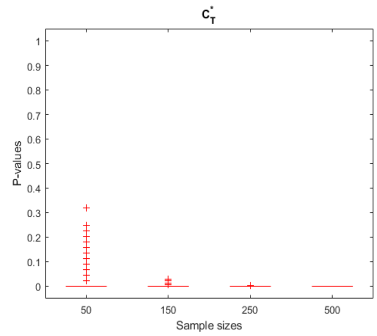
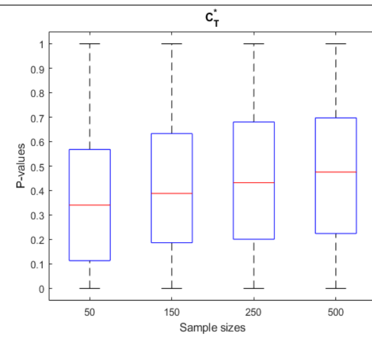
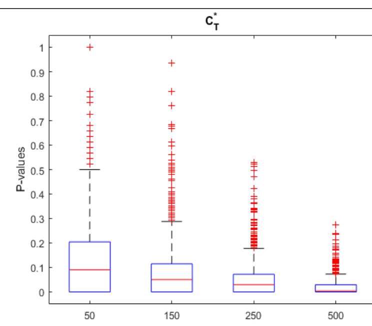
$$\lim_{T\to\infty} \mathbb{P}\left(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{C}^*_T\right) = 1;$$

where $\hat{c}_{1-\alpha}$ is the $1-\alpha$ quantile of a (centered) block-subsampling approximation of the asymptotic distribution of $\sqrt{T}\mathrm{C}^*_T$ with a block size b_T s.t. $\lim_{T\to\infty} b_T = \infty$ and $\lim_{T\to\infty} \frac{b_T}{T} = 0$.

Proof. See Appendix A.6. □

Proposition 5 shows the test conditional on the market is valid and consistent. Results from a Monte-Carlo simulation in Table 2 support Proposition 5. When the null hypothesis (8) is wrong, p-values converge to zero as the sample size increases. When the null hypothesis (8) is true, a large proportion of p-values is away from zero. For ease of comparison, the DGPs are the same as in Table 1 for the unconditional tests except for the common component x.

Table 2: Performance of conditional test in Monte-Carlo simulations

<table>
<tr><th>H_0</th><th>DGP</th><th>Boxplots of p-values</th></tr>
<tr><td>False</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} .015 \\ .0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>True</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} .011 \\ .010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>False</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \begin{bmatrix} z_L \\ z_S \end{bmatrix}$ where $\begin{cases} z_L \overset{IID}{\hookrightarrow} \mathrm{t}(4) \\ z_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\mathrm{or}(z_S, z_L) = .7 \end{cases}$</td><td>
</td></tr>
</table>

Note: The first two data-generating processes (DGP) are calibrated to data. In particular $x \overset{IID}{\hookrightarrow} \mathcal{N}(0, \sigma_x)$, where $\sigma_x = .04$ is the estimated standard deviation of monthly market returns. The third DGP is a stylized DGP except for the correlation that is calibrated to data. The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of C_T^* is approximated through centered block subsampling with block size $b_T = \sqrt{T}$. The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

3.3 Equilibrium foundations for the test conditional on the market

In the absence of diversification benefits, the equilibrium foundations of the conditional test is similar to the ones of the unconditional test. The reasoning is the same, except that investors control for the conditioning variable, that is, investors' preferences correspond to an expected utility under the distribution conditional on the market.

In the presence of diversification benefits, the following proposition formalizes the one-period equilibrium foundations for the test conditional on market.

Proposition 6 (Equilibrium foundation for test conditional on market). *Let r_W and $[\underline{u}_{W_1}, \overline{u}_{W_1}]$, respectively, denote the return on wealth (that is, $r_W := \frac{W_1}{W_0} - 1$, where W_0 denotes the initial wealth) and the support of W_1. Under Assumptions 1, for all $u \in \mathbf{U}_2$ s.t. u is twice continuously differentiable on $[\underline{u}, \overline{u}]$, which includes the support of W_1 and of the returns r_S and r_L, then, up to a first order, the null hypothesis "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2, \mathbb{E}[u(r_S)|r_W] < \mathbb{E}[u(r_L)|r_W]$" implies that the expected return of the factor exceeds its risk compensation, that is,*

$$-\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\text{ov}(u'(W_1), r_L - r_S) < \mathbb{E}(r_L - r_S).$$

Proof. Under Assumption 1, by iterated conditioning, the Hardy et. al. theorem, and Assumption 1(b) (no touching without crossing), if, $\forall u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)|r_W] < \mathbb{E}[u(r_L)|r_W]$, then, $\forall u \in \mathbf{U}_2$, $\mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)]$. Then the result follows immediately from Proposition 4. □

Proposition 6 shows that, up to a first order, strict preference for the long leg conditional on the market is a sufficient condition for an anomaly. The assumptions of Proposition 6 are similar to the assumptions of Proposition 4.

4 Empirical Results

We now apply our tests to actual data. We start by describing the dataset and, as a proof of concept, we apply the test to the market factor MKT. Then, we apply the tests to the widely-used FF5+MOM factors. Finally, we provide an overview of the test results for a standard dataset of more than 200 potential risk factors.

4.1 Data

Data for the five Fama and French factors and momentum, FF5+MOM, are from Kenneth French website. The factors are built by double sorting stocks on size and four characteristics, that is, book to market (BM), operating profitability (OP), investment (INV) and momentum (MOM). For each characteristic, stocks are double sorted into Small and Big stocks as well as tertiles of Low, Medium and High characteristics stocks. For each characteristic, the long leg of the corresponding factor is the equally weighted portfolio of two portfolios of Small and Big stocks in the highest tertiles (lowest for INV) and equivalently for the short leg. For each characteristic, the long leg of the corresponding Size factor is the equally weighted portfolio of three portfolios of Small stocks (Low, Medium and High), while the short leg is the equally weighted portfolio of three portfolios of Big stocks. Following Fama and French (2015), we built a Size factor by averaging the long and short legs across the Size factors related to BM, OP and INV. We also use as the aggregate market the CRSP value-weighted index as well as the one-month Treasury Bill for the risk-free rate.

For BM and MOM a long sample of data is available, starting from July 1926 (BM) or January 1927 (MOM). For the market and the Treasury bill yield, data are also available starting from July 1926. For OP and INV, data start only from July 1963. For this reason, we report for BM, MOM and the market MKT the findings for the full sample period as well as for a restricted period starting in July 1963. The samples for FF5+MOM factors end in October 2021.

We use data for 205 potential risk factors from Chen and Zimmermann (2020a). Stocks are sorted into quantile portfolios, where the number of quantiles depend upon data availability for the characteristic. We use the lowest and highest quantiles and retain as the short leg the quantile with the lowest sample average return over the sample period. We discuss evidence for the original samples of the published papers as well as for the post-publication samples and the full samples. The samples end in December 2020.

4.2 Proof of Concept

Propositions 1 and 5 show the unconditional test and the conditional test have good asymptotic properties. Monte-Carlo simulations (Tables 1-2 in previous sections and Appendix B) indicate that the finite sample performance of the tests are in line with the asymptotic properties. In the present section, we apply the unconditional test to the market factor MKT as a proof of concept on actual financial data.

Overwhelming empirical evidence show that US stocks have higher expected returns than Treasury bill returns, but that they are riskier. Thus, we test the following null hypothesis

$$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\, \mathbb{E}[u(r_f)] < \mathbb{E}[u(r_M)],$$

where r_f is the one-month Treasury bill return and r_M is the CRSP market return. We report results in Table 3.

Table 3: Unconditional test applied to the equity premium (i.e., market factor MKT)

	Long	**Short**	t_{NW}^{L-S}	**P-value**
1926 - 2021	0.96	0.27	4.01	0.00
1963 - 2021	0.96	0.37	3.18	0.00

Note: The columns "Long," "Short," "t_{NW}^{L-S}" and "P-value," respectively, correspond to the average return of the long leg, the average return of the short leg, the t-statistic for the null hypothesis "$\mathrm{H}_0 : \mathbb{E}(r_S) = \mathbb{E}(r_L)$," and the p-value of the unconditional test. We use Newey-West standard errors to calculate t_{NW}^{L-S}. The frequency of the data is monthly.

We clearly reject the null hypothesis, so, in line with the empirical evidence, the market factor MKT appear as a possible risk factor. In other words, levels of risk aversion exist s.t. US Treasury bills are preferred to US stocks. The results are robust to subsample analysis. While the results are a proof of concept for the unconditional test, they also indicate the tests set a high threshold to classify a factor as an anomaly, in the sense that they allow for any arbitrarily high level of risk aversion. By construction, the tests do not require the level of risk aversion (i.e., the concavity of the von Neumann-Morgenstern utility) to be plausible for actual agents in the economy. Mehra and Prescott (1985) also show a sufficiently high level of risk aversion can make individuals prefer US Treasury bills over US stocks, but they regard it as implausibly high, so they classify the market factor MKT as an anomaly, which they call the "equity premium puzzle."

4.3 Unconditional Test Applied to FF5+MOM Factors

The FF5+MOM factors are widely assumed to be risk factors and thus used to adjust for risk both in practice and academia. We apply our unconditional test to these factors to assess whether they are anomalies or possible risk factors. The results are reported in the following table.

Table 4: Unconditional test applied to FF5+MOM factors

	Long	**Short**	t_{NW}^{L-S}	**P-value**
Size 1963 - 2021	1.21	0.97	1.85	0.00
BM 1926 - 2021	1.32	0.99	2.80	0.15
BM 1963 - 2021	1.24	0.97	1.98	0.40
OP 1963 - 2021	1.18	0.92	2.71	1.00
INV 1963 - 2021	1.22	0.96	2.91	1.00
MOM 1926 - 2021	1.42	0.78	4.40	1.00
MOM 1963 - 2021	1.38	0.76	3.60	0.54
MKT 1926 - 2021	0.96	0.27	4.01	0.00
MKT 1963 - 2021	0.96	0.37	3.18	0.00

Note: The columns "Long," "Short," "t_{NW}^{L-S}" and "P-value," respectively, correspond to the average return of the long leg, the average return of the short leg, the t-statistic for the null hypothesis "$\mathrm{H}_0 : \mathbb{E}(r_S) = \mathbb{E}(r_L)$," and the p-value of the unconditional test. We use Newey-West standard errors to calculate t_{NW}^{L-S}. The frequency of the data is monthly. BM stands for book-to-market, OP for Operating Profitability, INV for Investment and MOM for Momentum.

Market factor MKT set aside, only Size has a p-value below standard thresholds. The result is robust to the different methods for constructing Size. A first potential explanation is the lack of significance of the factor's expected return: The t-statistic of the long-short portfolio t_{NW}^{L-S} is slightly below 1.96, suggesting Size might not be a factor after all, and thus neither an anomaly nor a risk factor. A second potential explanation is that Size can be explained by risk alone. This second explanation seems more plausible because a t-statistic t_{NW}^{L-S}, which is slightly below 1.96 and thus significant at 10%, is unlikely to explain a p-value of zero for the unconditional test. Moreover, in the original sample (Online Appendix) and for other constructions of the Size factor, the p-value is still zero even when the expected return is highly significant. This second, more plausible explanation lends support to Berk (1995), which explains why Size should not be regarded as an anomaly, but rather as a compensation for risk.

Regarding the factors BM, INV, OP and MOM there is strong evidence for the null hypothesis for the sub-sample period starting in July 1963. Similar results hold even if we exclude 2020 and 2021. For the MOM factor, the spread between the short and the long legs is greater than 7% on a yearly basis and hence close to the equity premium. While a high risk aversion could explain the equity premium, it cannot explain the MOM factor. The p-values are also large for the newly discovered OP and INV factors even though their expected returns are less than half the MOM factor's expected return. The findings indicate OP and INV are anomalies through the lens of our test.

The evidence for the BM factor is weaker, especially for the longest sample period. The findings complement the debate around the BM factor in Ang and Chen (2007) and Fama and French (2006) as well as to the recent value trap. A necessary condition for strong SSD is a strictly positive factor expected return. Ang and Chen (2007) document the value premium is absent pre-1963 explaining why in the longer sample the p-value of the unconditional test is much lower than in the post-1963 sample. In the latter sub-sample, the p-value of 40% strongly indicates that BM is not a risk factor. Note the sample period includes the 2010-2020 decade during which value stocks underperformed relative to growth stocks.

4.4 Test Conditional on Market applied to FF5+MOM Factors

The test conditional on the market has the main advantage relative to the unconditional test to control for exposure to market risk including nonlinear dependence. We report the results of the test conditional on the market in Table 5.

Table 5: Test conditional on market applied to FF5+MOM factors

	Long	**Short**	t_{NW}^{L-S}	**P-value**
Size 1963 - 2021	1.21	0.97	1.85	0.00
BM 1926 - 2021	1.32	0.99	2.80	0.37
BM 1963 - 2021	1.24	0.97	1.98	0.25
OP 1963 - 2021	1.18	0.92	2.71	0.40
INV 1963 - 2021	1.22	0.96	2.91	0.09
MOM 1926 - 2021	1.42	0.78	4.40	0.60
MOM 1963 - 2021	1.38	0.76	3.60	0.43

Note: The columns "Long," "Short," "t_{NW}^{L-S}" and "P-value," respectively, correspond to the average return of the long leg, the average return of the short leg, the t-statistic for the null hypothesis "$\text{H}_0 : \mathbb{E}(r_S) = \mathbb{E}(r_L)$," and the p-value of the unconditional test. We use Newey-West standard errors to calculate t_{NW}^{L-S}. The frequency of the data is monthly. BM stands for book-to-market, OP for Operating Profitability, INV for Investment and MOM for Momentum.

We still reject the null that Size is an anomaly. While the p-values drop for the other characteristics, the factors BM, OP and MOM still appear as anomalies. In the case of INV, the p-value is now only 9%, which is above the standard 5% threshold, but slightly below 10%. Again, the findings are robust to alternative construction methods of the Size factor as well as looking at recent data only.

One possible explanation for the drop in p-values relative to the unconditional test is the unusual absence of type I error for the latter, asymptotically (Proposition 1i vs Proposition 5ii). A second possible explanation is the important commonality between the market and the legs of different factors.

4.5 A Bird View on the Factor Zoo

Beyond the widely-used FF5+MOM factors studied above, hundreds of other factors —the factor "zoo"— have been discovered. In order to have a broader assessment, we also apply the two tests to a standard dataset of more than 200 potential factors. We report the detailed results in the Appendix. In the present section, we only provide an overview of the main results. We use 5% as the threshold above which we cannot reject the null hypothesis. We report the proportions of potential factors that appear as anomalies in the table below.

Table 6: Proportion of p-values above 5%

	Unconditional	Conditional on Market
Original Sample	0.92	0.87
Post-Pub. Sample	0.35	0.34
Full Sample	0.88	0.77

Note: The data base correspond to Chen and Zimmermann (2020a) data base of 205 potential factors. The frequency of the data is monthly.

A first result is that a majority of the 205 potential factors appear as anomalies in the original sample of the published papers and the full sample. For both tests, we find more than 70% appear as anomalies in the original sample and the full sample. Because the existence of a factor is necessary condition for an anomaly, this result lends support to Chen and Zimmermann (2020b), Chen (2021a,b), Jensen et al. (2021) among others, who find that most factors are replicable in the original sample. Remember the unconditional test is immune to multiple hypothesis problem and the pretesting problem (Propositions 2-3) and hence makes the results of this literature even stronger.

A second result is the dramatic drop in the proportion of anomalies from the original sample to the post publication sample: The proportion drops from about 90% to about 35% for both tests. Two potential explanations exist for this drop: (i) Many anomalies became risk factors after publication; or (ii) The phenomenon of "Anomalies elimination" occurred, that is, many anomalies disappeared because their expected returns shrank to zero. Table 7 supports the second explanation. Table 7 displays the proportion of apparent anomalies among the significant factors, that is, the proportion of p-values above 5% for the potential factors with expected returns significantly positive at the 5% level. The table shows the proportion of apparent anomalies among (significant) factors is above 80%, and often close to 90%, in line with "anomaly elimination," which has been documented (e.g., Hanson and Sunderam 2014, McLean and Pontiff 2016): Following the publication of an anomaly, some investors trade on it, so its expected return decreases after a temporary increase (Pénasse 2020).

Table 7: Proportion of p-values above 5% for significant factors

	Unconditional	Conditional on Market
Original Sample	0.93	0.89
Post-Pub. Sample	0.95	0.93
Full Sample	0.91	0.81

Note: We compute the displayed proportions as follows. (i) We keep from the Chen and Zimmermann (2020a) data base of 205 potential factors, the ones that have a t-statistics bigger than the 95% quantile of standard normal distribution. (ii) We compute the proportion of p-value above 5% among the kept factors. For simplicity, potential pretesting problems are ignored. The frequency of the data is monthly.

The third and main result is a clear majority of factors appear as anomalies in all samples. Overall, more than 80% of factors appear as anomalies in the original sample, the post-publication sample and the full sample (see Table 7). In Table 6, the proportions are lower because some potential factors do not have significantly positive expected returns and thus are not factors per se. This third result generalizes the results for the FF5+MOM factors to most of the factors documented in the literature. This generalization is not surprising given that theory and empirical evidence indicate a strong commonality between factors (e.g., Reisman 1992, Lewellen et al. 2010, Bryzgalova et al. 2020, Arvanitis et al. 2021).

4.6 Multiperiod Considerations

In line with a large part of the literature on factor models, for simplicity, we previously focused on one-period equilibrium foundations for the proposed tests. In the present section, we provide multiperiod equilibrium foundations for the unconditional test and a modified conditional test. For this purpose, as in the one-period case, we first derive the factor risk compensation required by risk-averse individuals.

Consider individuals who maximize time additive utility functions $U(C_{0:T}) := \sum_{t=0}^{T} \beta^t \mathbb{E}[u(C_t)]$, where $\beta \in]0,1[$ denotes a time discount factor, $u(.)$ an increasing and concave von Neuman-Morgenstern utility function, and $C_{0:T} := (C_0, C_1 \ldots, C_T)$ a consumption process.[12] A generalization of the one-period reasoning of Section 2.3.2 implies that, for any time period $t \in [\![1, T]\!]$ at which the factor $r_{L,t} - r_{S,t}$ is freely tradable, the following optimality condition holds

$$\mathbb{E}[u'(C_t)(r_{L,t} - r_{S,t})] = 0 \ ,$$

so the factor expected return explained by risk alone is

$$\mathbb{E}(r_{L,t} - r_{S,t}) = -\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\text{ov}(u'(C_t), r_{L,t} - r_{S,t}).$$

See Proposition A.1 in Appendix A.4 for a formal proof. Therefore, indexing returns with t, the equilibrium foundations provided by Propositions 4 and 6 still hold with C_t in lieu of W_t. The multiperiod version of Propositions 4 shows Tables 3 and 4 have multiperiod equilibrium foundations.

[12] Our tests cannot be extended to Epstein-Zin-Weil utility functions (Epstein and Zin 1989, PhilippeWeil 1989). One of the reasons is that Epstein-Zin-Weil utility functions violate first-order stochastic dominance, and thus, a fortiori, SSD. Individuals with Epstein-Zin-Weil utility functions do not always prefer more to less. More precisely, Epstein-Zin-Weil utility functions violate the monotonicity axiom according to which an agent does not choose a lottery if another available lottery is preferable in every state of the world. See Bommier et al. 2017 for a thorough analysis of this violation.

5 Summary and Discussion

Over the last decades, hundreds of factors predicting cross-sectional returns have been discovered. The present paper (i) introduces the concept of strong SSD; (ii) provides a general, but simple, derivation of the risk compensations required by risk-averse individuals to hold a factor; (iii) shows that if the long leg of a factor strongly SSD dominates its short leg, the factor's expected return should exceed its possible risk compensations in equilibrium; (iv) proposes two tests of strong SSD; (v) verifies the performance of the tests numerically,mathematically and empirically; and (vi) applies the two tests to more than 200 factors.

We propose and use two tests because they rely on slightly different assumptions and data. Despite their differences, both tests classify a majority of factors —including a majority of the widely used FF5+MOM factors— as anomalies. Thus, the factors "zoo" appears to be mainly an anomalies "zoo." This result might appear unexpected, because strong SSD sets a high threshold for anomalies. Strong SSD requires strict preference even for implausibly high level of risk aversion.

The proposed tests do not only help to detect anomalies. They also provide some guidance on which types of models can explain the anomalies. The tests and their theoretical foundations barely impose any restriction on distributions of returns nor on production, etc. Thus, explanations of the anomalies "zoo" call for models in which risk-averse individuals do not buy factors that they value higher than their market price. In particular, trading frictions on factors (e.g., He and Modest (1995), Luttmer (1996), intermediary asset pricing as in He and Krishnamurthy (2018)), or behavioral biases (e.g., Barberis et al. 2021) are possible explanations for the detected anomalies, while frictions on production are unlikely explanations.

Beyond the question of the factors "zoo," the present paper is a step toward a solution to Fama's joint hypothesis problem (Fama 1970, Roll 1977, Fama 2013), in the sense that it proposes model-free tests to detect abnormal excess returns. In its modern formulation, the joint hypothesis problem states that asset pricing tests always jointly test the existence of abnormal returns and a model of market equilibrium (e.g., CAPM). Hence, it is impossible to distinguish abnormal returns from using the wrong model of market equilibrium. In contrast, the two tests of the present paper can help detect abnormal excess returns without assuming a specific model of market equilibrium.[13] To the best of our knowledge, they are the first tests with this property. Therefore, the proposed tests should be useful to detect abnormal excess returns in many situations, especially given that the current dominating methodology assimilates abnormal returns to the alphas of regressions on a preferred factor model. In this way, both tests can provide guidance for better investment decisions and capital allocation.

[13]While our tests are a step toward a solution to the modern formulation of Fama's joint hypothesis problem, they do not address its original formulation in terms of information. Our tests do not assess whether assets prices reflect all available information. The latter remains an open issue.

References

Ahn, S. C. and Horenstein, A. R.: 2013, Eigenvalue ratio test for the number of factors, *Econometrica* **81**(3), 1203–1227.

Al-Najjar, N. I.: 1998, Factor analysis and arbitrage pricing in large asset economies, *Journal of Economic Theory* **78**(2), 231–262.

Aliprantis, C. D. and Border, K. C.: 2006/1994, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd edn, Springer.

Amengual, D. and Watson, M. W.: 2007, Consistent estimation of the number of dynamic factors in a large n and t panel, *Journal of Business & Economic Statistics* **25**(1), 91–96.

Ando, T. and Bai, J.: 2020, Quantile co-movement in financial markets: A panel quantile model with unobserved heterogeneity, *Journal of the American Statistical Association* **115**(529), 266–279.

Ang, A. and Chen, J.: 2007, Capm over the long run: 1926–2001, *Journal of Empirical Finance* **14**(1), 1–40.

Arvanitis, S., Hallam, M., Post, T. and Topaloglou, N.: 2019, Stochastic spanning, *Journal of Business & Economic Statistics* **37**(4), 573–585.

Arvanitis, S., Scaillet, O. and Topaloglou, N.: 2021, Spanning analysis of stock market anomalies under prospect stochastic dominance.

Bai, J.: 2003, Inferential theory for factor models of large dimensions, *Econometrica* **71**(1), 135–171.

Bai, J. and Ng, S.: 2002, Determining the number of factors in approximate factor models, *Econometrica* **70**(1), 191–221.

Banz, R. W.: 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* **9**(1), 3–18.

Barberis, N., Jin, L. J. and Wang, B.: 2021, Prospect theory and stock market anomalies, *The Journal of Finance* **76**(5), 2639–2687.

Barillas, F. and Shanken, J.: 2018, Comparing asset pricing models, *The Journal of Finance* **73**(2), 715–754.

Barrett, G. F. and Donald, S. G.: 2003, Consistent tests for stochastic dominance, *Econometrica* **71**(1), 71–104.

Basu, S.: 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *The Journal of Finance* **32**(3), 663–682.

Beran, J.: 1984, Bootstrap methods in statistics, *Jahresberichte des Deutschen Mathematichen Vereins* **86**, 14–30.

Berk, J. B.: 1995, A critique of size-related anomalies, *The Review of Financial Studies* **8**(2), 275–286.

Berk, J. B., Green, R. C. and Naik, V.: 1999, Optimal investment, growth options, and security returns, *The Journal of Finance* **54**, 1553–1607.

Black, F., Jensen, M. C. and Schole, M.: 1972, *Studies in the Theory of Capital Markets*, Praeger Publishers Inc., chapter The Capital Asset Pricing Model: Some Empirical Tests. Available at http://papers.ssrn.com/abstract=908569.

Blackwell, D.: 1951, Comparison of experiments, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, pp. 93–102.

Bommier, A., Kochov, A. and Legrand, F.: 2017, On monotone recursive preferences, *Econometrica* **85**(5), 1433–1466.

Bondt, W. F. M. D. and Thaler, R.: 1985, Does the stock market overreact?, *The Journal of Finance* **40**(3), 793–805.

Borovicka, J. and Borovicková, K.: 2018, Risk premia and unemployment fluctuations, *Manuscript, New York University* .

Bryzgalova, S., Huang, J. and Julliard, C.: 2020, Bayesian solutions for the factor zoo: We just ran two quadrillion models. Available at SSRN: https://ssrn.com/abstract=3481736.

Carhart, M. M.: 1997, On persistence in mutual fund performance, *The Journal of Finance* **52**(1), 57–82.

Cario, M. C. and Nelson., B. L.: 1997, Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University.

Carlier, G., Dana, R.-A. and Galichon, A.: 2012, Pareto efficiency for the concave order and multivariate comonotonicity, *Journal of Economic Theory* **147**(1), 207–229.

Cartier, P., Fell, J. M. G. and Meyer, P.-A.: 1964, Comparaison, des mesures portées par un ensemble convexe compact, *Bulletin de la Société Mathématique de France* **92**, 435–445.

Cattaneo, M. D., Crump, R. K., Farrell, M. H. and Schaumburg, E.: 2020, Characteristic-sorted portfolios: Estimation and inference, *Review of Economics and Statistics* **102**(3), 531–551.

Chalamandaris, G., Pagratis, S. and Topaloglou, N.: 2019, Are stock-market anomalies anomalous after all?

Chamberlain, G. and Rothschild, M.: 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica* **51**(5), 1281–1304.

Charoenrook, A. and Conrad, J.: 2008, Identifying risk-based factors. Older version (2005) available on ssrn at https://ssrn.com/abstract=757210.

Chen, A. Y.: 2021a, The limits of p-hacking: Some thought experiments. Forthcoming in The Journal of Finance.

Chen, A. Y.: 2021b, Most claimed statistical findings in cross-sectional return predictability are likely true. Available at SSRN: https://ssrn.com/abstract=3912915.

Chen, A. Y. and Zimmermann, T.: 2020a, Open source cross-sectional asset pricing. Available at SSRN: https://ssrn.com/abstract=3604626. Forthcoming in Critical Finance Review.

Chen, A. Y. and Zimmermann, T.: 2020b, Publication bias and the cross-section of stock returns, *The Review of Asset Pricing Studies* **10**(2), 249–289.

Chernozhukov, V. and Fernández-Val, I.: 2005, Subsampling inference on quantile regression processes, *Sankhyā: The Indian Journal of Statistics* pp. 253–276.

Chinco, A., Neuhierl, A. and Weber, M.: 2021, Estimating the anomaly base rate, *Journal of financial economics* **140**(1), 101–126.

Cochrane, J. H.: 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* **104**(3), 572–621.

Cochrane, J. H.: 2011, Presidential address: Discount rates, *The Journal of finance* **66**(4), 1047–1108.

Cohen, L., Malloy, C. and Pomorski, L.: 2012, Decoding inside information, *The Journal of Finance* **67**(3), 1009–1043.

Connor, G.: 1984, A unified beta pricing theory, *Journal of Economic Theory* **34**(1), 13–31.

Connor, G., Hagmann, M. and Linton, O.: 2012, Efficient semiparametric estimation of the fama-french model and extensions, *Econometrica* **80**(2), 713–754.

Connor, G. and Korajczyk, R. A.: 1993, A test for the number of factors in an approximate factor model, *The Journal of Finance* **48**(4), 1263–1291.

Cooper, I.: 2006, Asset pricing implications of nonconvex adjustment costs and irreversibility of investment, *The Journal of Finance* **61**(1), 139–170.

Dana, R. A.: 2004, Market behavior when preferences are generated by second-order stochastic dominance, *Journal of Mathematical Economics* **40**, 619–639.

Davidson, R. and Duclos, J.-Y.: 2000, Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica* **68**(6), 1435–1464.

Delgado, M. A. and Escanciano, J. C.: 2012, Distribution-free tests of stochastic monotonicity, *Journal of Econometrics* **170**, 68–75.

Delgado, M. A. and Escanciano, J. C.: 2013, Conditional stochastic dominance testing, *Journal of Business & Economic Statistics* **31**(1).

Donald, S. G. and Hsu, Y.-C.: 2016, Improving the power of tests of stochastic dominance, *Econometric Reviews* **35**(4), 553–585.

Doz, C., Giannone, D. and Reichlin, L.: 2011, A two-step estimator for large approximate dynamic factor models based on kalman filtering, *Journal of Econometrics* **164**(1), 188–205.

Doz, C., Giannone, D. and Reichlin, L.: 2012, A quasi-maximum likelihood approach for large, approximate dynamic factor models, *The Review of Economics and Statistics* **94**(4), 1014 –1024.

Durot, C.: 2003, A Kolmogorov-type test for monotonicity of regression, *Statistics and Probability Letters* **63**, 425–433.

Durot, C. and Tocquet, A.-S.: 2003, On the distance between the empirical process and its concave majorant in a monotone regression framework, *Annales de l'Institut Henri Poincare (B) Probability and Statistics* **39**(2).

Dybvig, P. H. and Ingersoll, J. E.: 1982, Mean-variance theory in complete markets, *The Journal of Business* **55**(2), 233–251.

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

Fama, E. F.: 1970, Efficient capital markets: A review of theory and empirical work, *The Journal of Finance,* **25**(2), 383–417.

Fama, E. F.: 2013, Two pillars of asset pricing, *Nobel Prize Lecture* .

Fama, E. F. and French, K. R.: 1992, The cross-section of expected stock returns, *The Journal of Finance* **47**(2), 427–465.

Fama, E. F. and French, K. R.: 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**, 3–56.

Fama, E. F. and French, K. R.: 2006, The value premium and the capm, *The Journal of Finance* **61**(5), 2163–2185.

Fama, E. F. and French, K. R.: 2015, A five-factor asset pricing model, *Journal of Financial Economics* **116**(1), 1–22.

Fama, E. F. and MacBeth, J. D.: 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* **81**(3), 607–636.

Feng, G., Giglio, S. and Xiu, D.: 2020, Taming the factor zoo: A test of new factors, *The Journal of Finance* **75**(3), 1327–1370.

Föllmer, H. and Schied, A.: 2011/2002, *Stochastic Finance. An Introduction in Discrete Time.*, 3rd edn, De Gruyter.

Forni, M., Hallin, M., Lippi, M. and Zaffaroni, P.: 2017, Dynamic factor models with infinite-dimensional factor space:asymptotic analysis, *Journal of Econometrics* **199**(1), 74–92.

Forni, M. and Lippi, M.: 2001, The generalized dynamic factor model: Representation theory, *Econometric Theory* **17**(6), 1113–1141.

Freyberger, J., Neuhierl, A. and Weber, M.: 2020, Dissecting characteristics nonparametrically, *The Review of Financial Studies* **33**(5), 2326–2377.

Gagliardini, P., Ossola, E. and Scaillet., O.: 2016, Time varying risk premium in large cross sectional equity data sets, *Econometrica* **84**(3), 985–1046.

Gagliardini, P., Ossola, E. and Scaillet, O.: 2019, A diagnostic criterion for approximate factor structure, *Journal of Econometrics* **212**(2), 503–521.

Gibbons, M. R., Ross, S. A. and Shanken, J.: 1989, A test of the efficiency of a given portfolio, *Econometrica* **57**(5), 1121–1152.

Giglio, S. and Xiu, D.: 2020, Asset pricing with omitted factors. Forthcoming in the Journal of Political Economy.

Gromb, D. and Vayanos, D.: 2010, Limits of arbitrage, *Annual Review of Financial Economics* **2**, 251–275.

Hadar, J. and Russell, W. R.: 1969, Rules for ordering uncertain prospects, *The American Economic Review* **59**(1), 25–34.

Hall, R. E.: 2017, High discounts and high unemployment, *American Economic Review* **107**(2), 305–30.

Hallin, M. and Liška, R.: 2007, Determining the number of factors in the general dynamic factor model, *Journal of the American Statistical Association* **102**(478), 603–617.

Hannan, E. J. and Quinn, B. G.: 1979, The determination of the order of an autoregression, *Journal of the Royal Statistical Society. Series B (Methodological)* **41**(2), 190–195.

Hanoch, G. and Levy, H.: 1969, The efficiency analysis of choices involving risk, *The Review of Economic Studies* **36**(3), 335—346.

Hansen, L. P.: 2013, Uncertainty outside and inside economic models, Nobel Prize Lecture.

Hansen, L. P. and Richard, S. F.: 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* **55**(3), 587–613.

Hanson, S. G. and Sunderam, A.: 2014, The growth and limits of arbitrage: Evidence from short interest, *The Review of Financial Studies* **27**(4), 1238–1286.

Hardy, G. H., Littlewood, J. E. and Pólya, G.: 1929, Some simple inequalities satisfied by convex functions, *Messenger of Mathematics* **58**, 145–152.

Hardy, G. H., Littlewood, J. E. and Pólya, G.: 1934, *Inequalities*, Cambridge University Press.

Harvey, C. R., Liu, Y. and Zhu, H.: 2016, . . and the cross-section of expected returns, *Review of Financial Studies* **29**(1).

He, H. and Modest, D. M.: 1995, Market frictions and consumption-based asset pricing, *Journal of Political Economy* **103**(1), 94–117.

He, Z. and Krishnamurthy, A.: 2018, Intermediary asset pricing and the financial crisis, *Annual Review of Financial Economics* **10**, 173–197.

Jagannathan, R. and Wang, Z.: 1998, An asymptotic theory for estimating beta-pricing models using cross-sectional regression, *The Journal of Finance* **53**(4), 285–1309.

Jegadeesh, N. and Titman, S.: 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *The Journal of Finance* **48**(1), 65–91.

Jensen, T. I., Kelly, B. and Pedersen, L. H.: 2021, Is there a replication crisis in finance? Available on SSRN.

Kallenberg, O.: 2002/1997, *Foundation of Modern Probability*, Probability and Its Applications, second edn, Springer.

Kan, R., Robotti, C. and Shanken, J.: 2013, Pricing model performance and the two-pass cross-sectional regression methodology, *The Journal of Finance* **68**(6), 2617–2649.

Kelly, B., Pruitt, S. and Su, Y.: 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* . Forthcoming.

Kim, S. and Skoulakis, G.: 2018, Ex-post risk premia estimation and asset pricing tests using large cross sections: The regression-calibration approach, *Journal of Econometrics* **204**(2), 159–188.

Kozak, S., Nagel, S. and Santosh, S.: 2018, Interpreting factor models, *The Journal of Finance* **73**(3).

Lehmann, E. L. and Romano, J. P.: 2006, *Testing statistical hypotheses*, Texts in statistics, Springer.

Lettau, M. and Pelger, M.: 2020, Estimating latent asset-pricing factors, *Journal of Econometrics* **218**(1), 1–31.

Lewellen, J., Nagel, S. and Shanken, J.: 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* **96**(2), 175–194.

Lintner, J.: 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *The Review of Economics and Statistics* **47**(1), 13–37.

Linton, O., Maasoumi, E. and Whang, Y.-J.: 2005, Consistent testing for stochastic dominance under general sampling schemes, *The Review of Economic Studies* **72**(3), 735–765.

Luttmer, E. G. J.: 1996, Asset pricing in economies with frictions, *Econometrica* **64**(6), 1439–1467.

McFadden, D.: 1989, *Studies in the Economics of Uncertainty*, Springer, chapter Testing for stochastic dominance, pp. 113–134.

McLean, R. D. and Pontiff, J.: 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* **71**(1), 5–32.

Mehra, R. and Prescott, E. C.: 1985, The equity premium. A puzzle, *Journal of Monetary Economics* **15**, 145–161.

Merton, R. C.: 1973, An intertemporal capital asset pricing model, *Econometrica* **41**(5), 867–887.

Miller, M. H. and Scholes, M.: 1972, *Studies in the theory of capital markets*, Praeger Publishers Inc., chapter Rates of return in relation to risk: A reexamination of some recent findings, pp. 768–783.

Milne, F.: 1988, Arbitrage and diversification in a general equilibrium asset economy, *Econometrica* **56**(4), 815–840.

Mossin, J.: 1966, Equilibrium in a capital asset market, *Econometrica* **34**(4), 768–783.

Onatski, A.: 2009, Testing hypotheses about the number of factors in large factor models, *Econometrica* **77**(5), 1447–1479.

Onatski, A.: 2010, Determining the number of factors from empirical distribution of eigenvalues, *The Review of Economics and Statistics* **92**(4), 1004–1016.

Pelger, M.: 2019, Large-dimensional factor modeling based on high-frequency observations, *Journal of Econometrics* **208**(1), 23–42.

Pénasse, J.: 2020, Understanding alpha decay. Available at http://dx.doi.org/10.2139/ssrn.2953614.

PhilippeWeil: 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* **24**(3), 401–421.

Politis, D. N., Romano, J. P. and Wolf, M.: 1999, *Subsampling*, Springer series in statistics, Springer.

Post, T.: 2003, Empirical tests for stochastic dominance efficiency, *The Journal of Finance* **58**(5), 1905–1931.

Post, T. and Kopa, M.: 2017, Portfolio choice based on third-degree stochastic dominance, *Management Science* **63**(10), 3381–3392.

Post, T. and Levy, H.: 2005, Does risk seeking drive stock prices? a stochastic dominance analysis of aggregate investor preferences and beliefs, *The Review of Financial Studies* **18**(3), 925–953.

Pukthuanthong, K., Roll, R. and Subrahmanyam., A.: 2018, A protocol for factor identification, *The Review of Financial Studies* **32**(4), 1573–1607.

Raponi, V., Robotti, C. and Zaffaroni, P.: 2020, Testing beta-pricing models using large cross-sections, *The Review of Financial Studies* **33**(6), 2796–2842.

Raponi, V., Uppal, R. and Zaffaroni, P.: 2018, Portfolio choice with model misspecification:a foundation for alpha and beta portfolios.

Reisman, H.: 1988, A general approach to the arbitrage pricing theory, *Econometrica* **56**(2), 473–476.

Reisman, H.: 1992, Reference variables, factor structure, and the approximate multibeta representation, *The Journal of Finance,* **47**(4), 1303–1314.

Renault, E., van der Heijden, T. and Werker, B. J.: 2019, Arbitrage pricing theory for idiosyncratic variance factors. Available at SSRN: https://ssrn.com/abstract=3065854.

Roll, R.: 1977, A critique of the asset pricing theory's tests part i: On past and potential testability of the theory, *Journal of Financial Economics* **4**(2), 129–176.

Ross, S. A.: 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* **13**(3), 341–360.

Rothschild, M. and Stiglitz, J. E.: 1970, Increasing risk: I. a definition, *Journal of Economic Theory* **2**(3), 225–243.

Rudin, W.: 1953, *Principles of Mathematical Analysis*, 3rd edn, McGraw-Hill.

Scaillet, O. and Topaloglou, N.: 2010, Testing for stochastic dominance efficiency, *Journal of Business & Economic Statistics* **28**(1), 169–180.

Sentana, E. and Fiorentini, G.: 2001, Identication, estimation and testing of conditionally heteroskedastic factor models, *Journal of Econometrics* **102**(2), 143–164.

Seyhun, H. N.: 1988, The information content of aggregate insider trading, *The Journal of Business* **61**(1), 1–24.

Sharpe, W. F.: 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* **19**(3), 425–442.

Sherman, S.: 1951, On a theorem of Hardy, Littlewood, Pólya, and Blackwell, *Proceedings of the National Academy of Sciences of the United States of America* **37**(12), 826.

Stock, J. H. and Watson, M. W.: 2002, Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association* **97**(460), 1167–1179.

Strassen, V.: 1965, The existence of probability measures with given marginals, *The Annals of Mathematical Statistics* **36**(2), 423–439.

Todorov, V. and Bollerslev, T.: 2010, Jumps and betas: A new framework for disentangling and estimating systematic risks, *Journal of Econometrics* **157**(2), 220–235.

Uppal, R., Zaffaroni, P. and Zviadadze, I.: 2018, Beyond the bound: Pricing assets with misspecified stochastic discount factors.

Whang, Y.-J.: 2019, *Econometric Analysis of Stochastic Dominance. Concept, Methods, Tools, and Applications*, Themes in Modern Econometrics, Cambridge University Press.

ONLINE APPENDIX TO:

Anomaly or Possible Risk Factor? Simple-To-Use Tests

Benjamin Holcblat, Abraham Lioui and Michael Weber

A Proofs

A.1 Proof of Lemma 1 and Lemma 2 (equivalent characterizations of strong SSD)

A.1.1 Unconditional strong SSD

Lemma 1 is a simplified version of the following theorem.

Theorem A.1 (Equivalent characterizations of strong SSD)**.** *Assume that the support of the random variables r_L and r_S is a subset of $[\underline{u}, \overline{u}] \subset \mathbf{R}$ with $\underline{u} \neq \overline{u}$. For a $u : [\underline{u}, \overline{u}] \to \mathbf{R}$, define $\check{u} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u(x) = 0\}\right\}$, and denote its left derivative and right derivative at x with $u'_-(x)$ and $u'_+(x)$, respectively. Then the following statements are equivalent.*

(i) For all real-valued, concave, and increasing function $u(.)$ on $[\underline{u}, \overline{u}]$ s.t. $u'_+(\underline{u}) \in \mathbf{R}$ and $u'_-(\check{u}) \in \mathbf{R} \setminus \{0\}$ with $\check{u} \neq \underline{u}$, $\mathbb{E}[u(r_S)] < \mathbb{E}[u(r_L)]$.

(ii) For all $z \in]\underline{u}, \infty[$, $\mathbb{E}[(z - r_L)^+] < \mathbb{E}[(z - r_S)^+]$.

(iii) For all $z \in]\underline{u}, \infty[$, $F_L^{(2)}(z) < F_S^{(2)}(z)$, where $F_L^{(2)}(z) := \int_{\underline{u}}^z F_L(y)\mathrm{d}y$.

Proof. Apply upcoming Theorem A.2 with $W_1 = 1$. □

A.1.2 Conditional strong SSD

Lemma 2 is a simplified version of the following Theorem. The following theorem is the conditional counterpart of Theorem A.1.

Theorem A.2 (Equivalent characterizations of conditional strong SSD)**.** *Assume that the support of the random variables r_L and r_S is a subset of $[\underline{u}, \overline{u}] \subset \mathbf{R}$ with $\underline{u} \neq \overline{u}$. Assume a complete probability space. For a function $u_{W_1} : [\underline{u}, \overline{u}] \to \mathbf{R}$ indexed by a random variable W_1, define $\check{u}_{W_1} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u_{W_1}(x) = 0\}\right\}$, and denote its left derivative and right derivative at x with $u'_{W_1,-}(x)$ and $u'_{W_1,+}(x)$, respectively. Then the following statements are equivalent.*

(i) For all real-valued, concave and increasing function $u_{W_1}(.)$ defined on $[\underline{u}, \overline{u}]$ and Borel measurable w.r.t. the index W_1 s.t. $\mathbb{E}|u_{W_1}(\underline{u})| < \infty$, $\mathbb{E}|u'_{W_1,+}(\underline{u})| < \infty$ and $\mathbb{E}|u'_{W_1,-}(\check{u}_{W_1})| < \infty$ with $u'_{W_1,-}(\check{u}_{W_1}) \neq 0$ and $\check{u}_{W_1} \neq \underline{u}$ a.s., $\mathbb{E}[u_{W_1}(r_S)|W_1] < \mathbb{E}[u_{W_1}(r_L)|W_1]$ a.s.

(ibis) *For all real-valued, concave and increasing function $u(.)$ on $[\underline{u}, \overline{u}]$ s.t. $u'_+(\underline{u}) \in \mathbf{R}$ and $u'_-(\check{u}) \in \mathbf{R} \setminus \{0\}$ with $\check{u} \neq \underline{u}$, $\mathbb{E}[u(r_S)|W_1] < \mathbb{E}[u(r_L)|W_1]$ a.s.*

(ii) *For all $z \in]\underline{u}, \infty[$, $\mathbb{E}[(z - r_L)^+|W_1] < \mathbb{E}[(z - r_S)^+|W_1]$ a.s.*

(iii) *For all $z \in]\underline{u}, \infty[$, $F^{(2)}_{L|W_1}(z|W_1) < F^{(2)}_{S|W_1}(z|W_1)$ a.s., where $F^{(2)}_{L|W_1}(z|W_1) := \int_{\underline{u}}^{z} F_{L|W_1}(y|W_1)\mathrm{d}y$ a.s.*

Before the proof of Theorem A.2, the following lemma shows that $\check{u}_{W_1}$ is well-defined and measurable.

Lemma A.1 (Existence and $\sigma(W_1)$-measurability of $\check{u}_{W_1}$)**.** *Under the assumptions of Theorem A.2, for all the members of the class of utility functions defined in the statement (i) of the latter theorem, the following statements hold.*

(i) *There exists a function $w_1 \mapsto \check{u}_{w_1}$ with values in $[\underline{u}, \overline{u}]$ s.t. $\check{u}_{w_1} := \min\big\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\}\big\}$, for all $w_1 \in \mathbf{R}$.*

(ii) *The correspondence $\varphi(w_1) := \{x \in [\underline{u}, \overline{u}] : u_{w_1}(x) = 0\}$ is closed and connected valued, and weakly measurable.*

(iii) *The correspondences $\psi_{\underline{u}}(w_1) := \begin{cases} \varphi(w_1) & \text{if } \varphi(w_1) \neq \emptyset \\ \{\underline{u}\} & \text{otherwise} \end{cases}$ is closed, connected and non-empty valued, and weakly measurable.*

(iv) *For all $w_1 \in \mathbf{R}$, $\{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\} = \emptyset$ iff $0 < d(\overline{u}, \psi_{\underline{u}}(w_1)) := \inf_{x \in \psi_{\underline{u}}(w_1)} |\overline{u} - x|$.*

(v) *The function $w_1 \mapsto \check{u}_{w_1}$ is Borel measurable.*

Proof. *(i)* For convenience, in the present proof, put $A_{w_1} := \{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\}$, where $w_1 \in \mathbf{R}$.

1st case: $\forall z \in [\underline{u}, \overline{u}]$, $\exists \dot{z} \in [z, \overline{u}]$ s.t. $u_{w_1}(\dot{z}) \neq 0$. Then, by definition, the set A_{w_1} is the empty set $\emptyset$, so its greatest lower bound is ∞ (i.e., $\inf A_{w_1} = \inf \emptyset = \infty$), which, in turn, implies that $\check{u}_{w_1} := \min\big\{\overline{u}, \inf A_{w_1}\big\} = \overline{u}$.

2nd case: $\exists z \in [\underline{u}, \overline{u}]$, s.t., $\forall \dot{z} \in [z, \overline{u}]$, $u_{w_1}(\dot{z}) = 0$. Then, A_{w_1} is not the empty set. There are two subcases. First, consider the subcase $A_{w_1} := \{\overline{u}\}$, so $\check{u}_{w_1} = \overline{u}$. Now consider the remaining subcase $A_{w_1} \neq \{\overline{u}\}$, so $\inf A_{w_1} \neq \overline{u}$. By the sequential characterization of infima, there exists a sequence $(z_n) \in A_{w_1}^{\mathbf{N}}$ s.t. $\lim_{n \to \infty} z_n = \inf A_{w_1}$. Now, A_{w_1} is a subset of the closed set $[\underline{u}, \overline{u}]$, so $(z_n) \in [\underline{u}, \overline{u}]^{\mathbf{N}}$, which, in turn, implies that $\inf A_{w_1} \in [\underline{u}, \overline{u}]$ by the sequential characterization of closed sets (e.g., Aliprantis and Border 2006/1994, Lemma 3.3.5).

(ii) Closeness, connectedness and weak measurability respectively follow from the continuity, the monotonicity of $u_{w_1}(.)$, and the measurability of correspondences defined as a level set of a Carathéodory function (e.g., Aliprantis and Border 2006/1994, Lemma 18.8.2).

(iii) We only prove the statement for $\psi_{\overline{u}}(.)$ because the proof is the same for $\psi_{\underline{u}}(.)$. By construction, the correspondence $\psi_{\overline{u}}(.)$ is closed, connected and non-empty valued by the properties of $\varphi(.)$ stated in (ii), and the properties of the singleton $\{\overline{u}\}$. Thus, it remains to show that $\psi_{\overline{u}}(.)$ is weakly measurable.

Denote the lower inverse of a correspondence $\psi : S \twoheadrightarrow X$ with $\psi^l(.)$, i.e., $\psi^l(A) = \{s \in S : \psi(s) \cap A \neq \emptyset\}$, $\forall A \subset X$ (e.g., Aliprantis and Border 2006/1994, p. 557). By definition of the lower inverse and of the correspondence $\psi_{\overline{u}}$, for all open subset O of $[\underline{u}, \overline{u}]$,

$$\begin{aligned}\psi^l_{\overline{u}}(O) &= \{w_1 \in \mathbf{R} : \varphi(w_1) \cap O \neq \emptyset\} \bigcup [\{w_1 \in \mathbf{R} : \varphi(w_1) = \emptyset\} \cap \{w_1 \in \mathbf{R} : \{\overline{u}\} \cap O \neq \emptyset\}] \\ &= \varphi^l(O) \bigcup \Big[\varphi^l(\mathbf{R})^c \cap \{w_1 \in \mathbf{R} : \overline{u} \in O\}\Big] \in \mathcal{B}(\mathbf{R})\end{aligned}$$

where the explanations for the last inclusion are the following. First, by (ii), $\varphi(.)$ is weakly measurable, so $\varphi^l(O)$ and $\varphi^l(\mathbf{R})^c$ are measurable (e.g., Aliprantis and Border 2006/1994, Definition 18.1). Second, $\{w_1 \in \mathbf{R} : \overline{u} \in O\} = \emptyset$ or $\mathbf{R}$, so it is also Borel measurable.

(iv) Fix $w_1 \in \mathbf{R}$. "$\Rightarrow$" Assume $\{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\} = \emptyset$. There are two cases.

1st case: $\psi_{\underline{u}}(w_1) = \varphi(w_1)$. By (ii), $\psi_{\underline{u}}(w_1) = \varphi(w_1) := \{x \in [\underline{u}, \overline{u}] : u_{w_1}(x) = 0\}$ is a closed connected set, which means a closed interval (e.g., Rudin 1953, Theorem 2.47). Thus, $\{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\} = \emptyset$ (i.e., $\forall z \in [\underline{u}, \overline{u}]$, $\exists x \in [z, \overline{u}]$ s.t. $u_{w_1}(x) \neq 0$) implies that $d(\overline{u}, \psi_{\underline{u}}(w_1)) > 0$.

2nd case: $\psi_{\underline{u}}(w_1) = \{\underline{u}\}$. Then, $d(\overline{u}, \psi_{\underline{u}}(w_1)) = d(\overline{u}, \underline{u}) > 0$, because $\underline{u} \neq \overline{u}$ by assumption.

"$\Leftarrow$" If $d(\overline{u}, \psi_{\underline{u}}(w_1)) > 0$, then, for all $x \in [\overline{u} - \epsilon, \overline{u}]$ where $\epsilon := d(\overline{u}, \psi_{\underline{u}}(w_1))$, $u_{w_1}(x) \neq 0$ by definition of $\psi_{\underline{u}}(.)$. Thus, $\forall z \in [\underline{u}, \overline{u}]$, $\exists x \in [\max(z, \overline{u} - \epsilon), \overline{u}]$ s.t. $u_{w_1}(x) \neq 0$. Thus, $\{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\} = \emptyset$.

(v) By (iii), the correspondence $\psi_{\underline{u}}(.)$ is weakly measurable and nonempty-valued. Thus, the distance function $\delta : [\underline{u}, \overline{u}] \times \mathbf{R} \to \mathbf{R}$ s.t. $\delta(z, w_1) := d(z, \psi_{\underline{u}}(w_1)) := \inf_{x \in \psi_{\underline{u}}(w_1)} |z - x|$ is Carathéodory (e.g., Aliprantis and Border 2006/1994, Theorem 18.5), so, the set $B := \{w_1 \in \mathbf{R} : \delta(\overline{u}, w_1) > 0\} = \{w_1 \in \mathbf{R} : d(\overline{u}, \psi_{\underline{u}}(w_1)) > 0\}$ is Borel measurable. Moreover, by (iii), the correspondence $\psi_{\underline{u}}(.)$ is closed and nonempty valued and weakly measurable, so, by the Castaing representation theorem (e.g., Aliprantis and Border 2006/1994, Corollary 18.14.2), there exists a sequence of Borel measurable selectors $(f_n)_{n \in \mathbf{N}}$ s.t. $\psi_{\underline{u}}(w_1) = \overline{\{f_1(w_1), f_2(w_1), \ldots\}}$, for all $w_1 \in \mathbf{R}$. Then, by (iv),

$$\check{u}_{w_1} = \overline{u}\mathbb{1}_B(w_1) + \{\inf_{n \in \mathbf{N}} f_n(w_1)\}\mathbb{1}_{B^c}(w_1),$$

which is Borel measurable as the product and the addition of Borel measurable functions. $\square$

Proof of Theorem A.2. The proof —especially that *(ii)* implies *(i)*— does not follow the usual proof of the Hardy-Littlewood et. al. theorem provided in the economic and finance literature. The latter proof relies on limiting arguments (e.g., Rothschild and Stiglitz 1970) that do not go well with strict inequalities. In particular, for two real-valued sequences (u_n) and (v_n), the strict inequalities $u_n < v_n$, for all $n \in \mathbf{N}$, do not imply $\lim_{n\to\infty} u_n < \lim_{n\to\infty} v_n$. The proof follows from the introduction of the quantity $\check{u} \neq 0$, careful modifications of the proof techniques used in the mathematical literature (e.g., Föllmer and Schied 2011/2002, for a textbook presentation), and new technical lemmas.

(i) ⇒ (ibis) If $u_{W_1}(.) = u(.)$, then $|u'_+(\underline{u})| = \mathbb{E}|u'_{W_1,+}(\underline{u})| \in \mathbf{R}$ and $|u'_-(\check{u})| = \mathbb{E}|u'_{W_1,-}(\check{u})| \in \mathbf{R}\setminus\{0\}$.

(ibis) ⇒ (ii). For any $z \in]\underline{u}, \infty[$, the function $x \mapsto -(z-x)^+$ is a real-valued, concave, increasing function on $[\underline{u}, \overline{u}]$. Moreover, $\check{u} = z$ if $z \in]\underline{u}, \overline{u}]$, and $\check{u} = \overline{u}$ otherwise, so $u'_-(\check{u}) = 1 \neq 0$ and $\check{u} \neq \underline{u}$. Moreover, for any $z \in]\underline{u}, \infty[$, if $u(x) = -(z-x)^+$, then $u'_+(\underline{u}) = 1$. Thus, putting $u(x) = -(z-x)^+$, by assumption, $-\mathbb{E}[(z-r_S)^+|W_1] < -\mathbb{E}[(z-r_L)^+|W_1]$ a.s., which is equivalent to the needed result $\mathbb{E}[(z-r_L)^+|W_1] < \mathbb{E}[(z-r_S)^+|W_1]$ a.s.

(ii) ⇒ (i). Let $u_{W_1}(.)$ be real-valued, concave, continuous, and increasing function $u_{W_1}(.)$ defined on $[\underline{u}, \overline{u}]$ and Borel measurable w.r.t. the index W_1 s.t. $\mathbb{E}|u_{W_1}(\underline{u})| < \infty$, $\mathbb{E}|u'_{W_1,+}(\underline{u})| < \infty$ and $\mathbb{E}|u'_{W_1,-}(\check{u}_{W_1})| < \infty$ with $u'_{W_1,-}(\check{u}_{W_1}) \neq 0$ and $\check{u}_{W_1} \neq \underline{u}$ a.s., Then, $h_{W_1}(.) := -u_{W_1}(.)$ is a convex function. By the fundamental theorem of calculus for convex functions (e.g., Föllmer and Schied 2011/2002, Proposition A.4), for all $x \in [\underline{u}, \overline{u}]$, a.s.,

$$
\begin{aligned}
& h_{W_1}(x) \\
= {}& h_{W_1}(\check{u}_{W_1}) + \int_{\check{u}_{W_1}}^{x} \overline{h}'_{W_1,-}(y)\mathrm{d}y \text{ where } \overline{h}'_{W_1,-}(.) := h'_{W_1,-}(.)\mathbb{1}_{]\underline{u},\overline{u}]}(.) + h'_{W_1,+}(.)\mathbb{1}_{\{\underline{u}\}}(.) \\
= {}& h_{W_1}(\check{u}_{W_1}) - \int_{x}^{\check{u}_{W_1}} \overline{h}'_{W_1,-}(y)\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} \\
& \text{because, by definition of } \overline{h}'_{W_1,-}(.) \text{ and } \check{u}_{W_1}, \forall y \in]\check{u}_{W_1}, \overline{u}],\ \overline{h}'_{W_1,-}(y) = 0; \\
\overset{(a)}{=} {}& h_{W_1}(\check{u}_{W_1}) - \int_{x}^{\check{u}_{W_1}} [\overline{h}'_{W_1,-}(y) - \overline{h}'_{W_1,-}(\check{u}_{W_1}) + \overline{h}'_{W_1,-}(\check{u}_{W_1})]\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} \\
= {}& h_{W_1}(\check{u}_{W_1}) - \int_{x}^{\check{u}_{W_1}} \overline{h}'_{W_1,-}(\check{u}_{W_1})\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} - \int_{x}^{\check{u}_{W_1}} [\overline{h}'_{W_1,-}(y) - \overline{h}'_{W_1,-}(\check{u}_{W_1})]\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} \\
\overset{(b)}{=} {}& h_{W_1}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\check{u}_{W_1})(\check{u}_{W_1} - x)\mathbb{1}\{x \leqslant \check{u}_{W_1}\} + \int_{x}^{\check{u}_{W_1}} [\overline{h}'_{W_1,-}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(y)]\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} \\
\overset{(c)}{=} {}& h_{W_1}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\check{u}_{W_1})(\check{u}_{W_1} - x)^+ + \int_{x}^{\check{u}_{W_1}} \int_{y}^{\check{u}_{W_1}} \gamma_{W_1}(\mathrm{d}z)\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} \text{ where } \gamma_{W_1} \text{ is a random} \\
& \sigma\text{-finite Borel measure on } [\underline{u}, \overline{u}[\text{ s.t., } \forall (a,b) \in [\underline{u}, \overline{u}]^2,\ \gamma_{W_1}([a,b[) = \overline{h}'_{W_1,-}(b) - \overline{h}'_{W_1,-}(a); \\
\overset{(d)}{=} {}& h_{W_1}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\check{u}_{W_1})(\check{u}_{W_1} - x)^+ + \int_{\underline{u}}^{\check{u}_{W_1}} \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\mathrm{d}y\gamma_{W_1}(\mathrm{d}z)\mathbb{1}\{x \leqslant \check{u}_{W_1}\} \\
\overset{(e)}{=} {}& h_{W_1}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\check{u}_{W_1})(\check{u}_{W_1} - x)^+ + \int_{\underline{u}}^{\check{u}_{W_1}} (z-x)^+\gamma_{W_1}(\mathrm{d}z)
\end{aligned}
\tag{A.1}
$$

(a) By assumption, $\mathbb{E}|h'_{W_1,-}(\check{u}_{W_1})| = \mathbb{E}|u'_{W_1,-}(\check{u}_{W_1})| < \infty$, so $h'_{W_1,-}(\check{u}_{W_1})$ is finite a.s.[14] Now,

[14] Concavity of $u_{W_1}(.)$ ensure the existence of $u'_{W_1,-}(\check{u}_{W_1})$ only if $\check{u}_{W_1} \in]\underline{u}, \overline{u}[$.

$\overline{h}'_{W_1,-}(.) := h'_{W_1,-}(.)\mathbb{1}_{]\underline{u},\overline{u}]}(.) + h'_{W_1,+}(.)\mathbb{1}_{\{\underline{u}\}}(.) = h'_{W_1,-}(.)$ a.s. because $\check{u}_{W_1} \neq \underline{u}$ a.s. by assumption. Thus, $\overline{h}'_{W_1,-}(\check{u}_{W_1})$ is finite a.s. *(b)* Standard algebra yields $\int_x^{\check{u}_{W_1}} \overline{h}'_{W_1,-}(\check{u}_{W_1})\mathrm{d}y = \overline{h}'_{W_1,-}(\check{u}_{W_1}) \int_x^{\check{u}_{W_1}} \mathrm{d}y = \overline{h}'_{W_1,-}(\check{u}_{W_1})(\check{u}_{W_1} - x)$. *(c)* By Lemmas A.2 and A.4 (p. OA.6 & OA.7), there exists a unique σ-finite random Borel measure γ_{W_1} on $[\underline{u}, \check{u}_{W_1}[$ s.t. $\gamma_{W_1}([a,b[) = \overline{h}'_{W_1,-}(b) - \overline{h}'_{W_1,-}(a)$, $\forall (a,b) \in [\underline{u}, \overline{u}]^2$ a.s. *(d)* $\int_x^{\check{u}_{W_1}} \int_y^{\check{u}_{W_1}} \gamma_{W_1}(\mathrm{d}z)\mathrm{d}y = \int_{\underline{u}}^{\check{u}_{W_1}} \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{1}\{y \leqslant z\}\gamma_{W_1}(\mathrm{d}z)\mathbb{1}\{x \leqslant y\}\mathrm{d}y = \int_{\underline{u}}^{\check{u}_{W_1}} \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\gamma_{W_1}(\mathrm{d}z)\mathrm{d}y = \int_{\underline{u}}^{\check{u}_{W_1}} \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\mathrm{d}y\gamma_{W_1}(\mathrm{d}z)$ where the last equality follows from Fubini-Tonelli's theorem (e.g., Kallenberg 2002/1997, Theorem 1.27) because the Lebesgue measure and γ_{W_1} are σ-finite on $[\underline{u}, \overline{u}]$. *(e)* Standard algebra yields, $\forall z \in [\underline{u}, \check{u}_{W_1}]$, $\int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\mathrm{d}y\mathbb{1}\{x \leqslant \check{u}_{W_1}\} = \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\mathrm{d}y = (z-x)\mathbb{1}\{x \leqslant z\} = (z-x)^+$.

Then, by the theorem of disintegration of measures (e.g., Kallenberg 2002/1997, Theorem 6.3-6.4 with equation (6)) and Lemma A.1v on p. OA.2,, a.s.,

$$
\begin{aligned}
&-\mathbb{E}[u_{W_1}(r_L)|W_1] = \mathbb{E}[h_{W_1}(r_L)|W_1] = \int_{\underline{u}}^{\overline{u}} h_{W_1}(x)\mathrm{d}F_{L|W_1}(x|W_1) \\
\overset{(a)}{=}\ & h_{W_1}(\check{u}_{W_1}) \int_{\underline{u}}^{\overline{u}} \mathrm{d}F_{L|W_1}(x|W_1) - \overline{h}'_{W_1,-}(\check{u}_{W_1}) \int_{\underline{u}}^{\overline{u}} (\check{u}_{W_1} - x)^+ \mathrm{d}F_{L|W_1}(x|W_1) \\
&+ \int_{\underline{u}}^{\overline{u}} \int_{\underline{u}}^{\check{u}_{W_1}} (z-x)^+ \gamma_{W_1}(\mathrm{d}z)\mathrm{d}F_{L|W_1}(x|W_1) \\
\overset{(b)}{=}\ & h_{W_1}(\check{u}_{W_1})[F_{L|W_1}(\overline{u}|W_1) - F_{L|W_1}(\underline{u}|W_1)] - \overline{h}'_{W_1,-}(\check{u}_{W_1})\mathbb{E}[(\check{u}_{W_1} - r_L)^+|W_1] \\
&+ \int_{\underline{u}}^{\check{u}_{W_1}} \int_{\underline{u}}^{\overline{u}} (z-x)^+ \mathrm{d}F_{L|W_1}(x|W_1)\gamma_{W_1}(\mathrm{d}z) \\
\overset{(c)}{=}\ & h_{W_1}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\check{u}_{W_1})\mathbb{E}[(\check{u}_{W_1} - r_L)^+|W_1] + \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{E}[(z - r_L)^+|W_1]\gamma_{W_1}(\mathrm{d}z) \\
\overset{(d)}{<}\ & h_{W_1}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\check{u}_{W_1})\mathbb{E}[(\check{u}_{W_1} - r_S)^+|W_1] + \int_{\underline{u}}^{\check{u}_{W_1}} \mathbb{E}[(z - r_S)^+|W_1]\gamma_{W_1}(\mathrm{d}z) \\
=\ & \mathbb{E}[h_{W_1}(r_S)|W_1] = -\mathbb{E}[u_{W_1}(r_S)|W_1]
\end{aligned}
$$

(a) Show the three terms of equation (A.1) have a finite expectation so their conditional expectation are well-defined (e.g., Kallenberg 2002/1997, Theorem 6.1.i&iii), which, in turn, implies that the integral of the sum is the sum of the integrals. Firstly, by definition, the support of $\check{u}_{W_1}$ is in $[\underline{u}, \overline{u}]$, so $\mathbb{E}|h_{W_1}(\check{u}_{W_1})| < \infty$ by Lemma A.5 on p. OA.7. Secondly, by the triangle inequality, provided that $\check{u}_{W_1}$ and r_L take values in $[\underline{u}, \overline{u}]$, $\mathbb{E}|\overline{h}'_{W_1,-}(\check{u}_{W_1})(\check{u}_{W_1} - r_L)^+| \leqslant \mathbb{E}|\overline{h}'_{W_1,-}(\check{u}_{W_1})||\overline{u} - \underline{u}| = |\overline{u} - \underline{u}|\mathbb{E}|h'_{W_1,-}(\check{u}_{W_1})| = |\overline{u} - \underline{u}|\mathbb{E}|u'_{W_1,-}(\check{u}_{W_1})| < \infty$ by assumption, the definition of $\overline{h}'_{W_1,-}(.)$, and the assumption $\check{u}_{W_1} \neq \underline{u}$. Thirdly, by the triangle inequality and the monotonicity of the Lebesgue integral (e.g., Aliprantis and Border 2006/1994, Theorem 11.13.3), $\mathbb{E}|\int_{\underline{u}}^{\check{u}_{W_1}} (z - r_L)^+ \gamma_{W_1}(\mathrm{d}z)| \leqslant \mathbb{E}\int_{\underline{u}}^{\check{u}_{W_1}} |\overline{u} - \underline{u}|\gamma_{W_1}(\mathrm{d}z) = |\overline{u} - \underline{u}|\mathbb{E}|\overline{h}'_{W_1,-}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\underline{u})| \leqslant |\overline{u} - \underline{u}|[\mathbb{E}|\overline{h}'_{W_1,-}(\check{u}_{W_1})| + \mathbb{E}|\overline{h}'_{W_1,-}(\underline{u})|] = |\overline{u} - \underline{u}|[\mathbb{E}|h'_{W_1,-}(\check{u}_{W_1})| + \mathbb{E}|h'_{W_1,+}(\underline{u})|] < \infty$ by assumption, and where the last equality follows from the definition of the extended

derivative $\overline{h}'_{W_1,-}(.)$, which is a.s. equal to $h'_{W_1,-}(.)\mathbb{1}_{]\underline{u},\overline{u}]}(.) + h'_{W_1,+}(.)\mathbb{1}_{\{\underline{u}\}}(.)$, and the assumption $\check{u}_{W_1} \neq \underline{u}$. *(b)* First, by definition, the probability measure corresponding to the c.d.f. $F_{L|W_1}$ is finite, and thus σ-finite. Second, by Lemma A.2, the random measure $\gamma_{W_1}(.)$ is σ-finite. Thus, by Fubini-Tonelli's theorem (e.g., Kallenberg 2002/1997, Theorem 1.27), $\int_{\underline{u}}^{\overline{u}}\int_{\underline{u}}^{\overline{u}}(z-x)^+\gamma_{W_1}(\mathrm{d}z)\mathrm{d}F_{L|W_1}(x|W_1) = \int_{\underline{u}}^{\overline{u}}\int_{\underline{u}}^{\overline{u}}(z-x)^+\mathrm{d}F_{L|W_1}(x|W_1)\gamma_{W_1}(\mathrm{d}z)$. *(c)* By definition of c.d.f. with support $[\underline{u},\overline{u}]$, $F_{L|W_1}(\overline{u}|W_1) = 1$ and $F_{L|W_1}(\underline{u}|W_1) = 0$, so $F_{L|W_1}(\overline{u}|W_1) - F_{L|W_1}(\underline{u}|W_1) = 1$. *(d)* Firstly, by assumption, $\forall z \in]\underline{u},\overline{u}]$, $\mathbb{E}[(z-r_L)^+|W_1] < \mathbb{E}[(z-r_S)^+|W_1]$, and $\check{u}_{W_1} \neq \underline{u}$, so $-\overline{h}'_{W_1,-}(\check{u}_{W_1})\mathbb{E}[(\check{u}_{W_1}-r_L)^+|W_1] < -\overline{h}'_{W_1,-}(\check{u}_{W_1})\mathbb{E}[(\check{u}_{W_1}-r_S)^+|W_1]$ by Lemma A.3 on p. OA.6. Secondly, by assumption, $\forall z \in]\underline{u},\overline{u}]$, $\mathbb{E}[(z-r_L)^+|W_1] < \mathbb{E}[(z-r_S)^+|W_1]$ a.s., so $\int_{\underline{u}}^{\overline{u}}\mathbb{E}[(z-r_L)^+|W_1]\gamma_{W_1}(\mathrm{d}z) \leqslant \int_{\underline{u}}^{\overline{u}}\mathbb{E}[(z-r_S)^+|W_1]\gamma_{W_1}(\mathrm{d}z)$ by the monotonicity of the Lebesgue integral (e.g., Kallenberg 2002/1997, Lemma 1.18). Moreover, as previously noticed in the explanation for (a), $\mathbb{E}|\int_{\underline{u}}^{\check{u}_{W_1}}(z-x)^+\gamma_{W_1}(\mathrm{d}z)| \leqslant \mathbb{E}\int_{\underline{u}}^{\check{u}_{W_1}}|\overline{u}-\underline{u}|\gamma_{W_1}(\mathrm{d}z) = |\overline{u}-\underline{u}|\mathbb{E}|\overline{h}'_{W_1,-}(\check{u}_{W_1}) - \overline{h}'_{W_1,-}(\underline{u})| \leqslant |\overline{u}-\underline{u}|\left[\mathbb{E}|\overline{h}'_{W_1,-}(\check{u}_{W_1})| + \mathbb{E}|\overline{h}'_{W_1,-}(\underline{u})|\right] = |\overline{u}-\underline{u}|[\mathbb{E}|h'_{W_1,-}(\check{u}_{W_1})| + \mathbb{E}|h'_{W_1,+}(\underline{u})|] < \infty$, so $\mathbb{E}|\mathbb{E}[\int_{\underline{u}}^{\check{u}_{W_1}}(z-r_L)^+\gamma_{W_1}(\mathrm{d}z)|W_1]| = \mathbb{E}|\int_{\underline{u}}^{\check{u}_{W_1}}\mathbb{E}[(z-r_L)^+|W_1]\gamma_{W_1}(\mathrm{d}z)| < \infty$, which implies that $\int_{\underline{u}}^{\check{u}_{W_1}}\mathbb{E}[(z-r_L)^+|W_1]\gamma_{W_1}(\mathrm{d}z)$ is finite a.s.

(ii) $\Leftrightarrow$ *(iii)*. By the theorem of disintegration of measures, we can follow the standard mathematical proof based on Fubini-Tonelli's theorem. □

Lemma A.2. *Under the assumptions of Theorem A.2, for all the members of the class of utility functions defined in the statement (i) of the latter theorem, there exists a unique random σ-finite measure $\gamma_{W_1}(.)$ on $[\underline{u},\overline{u}]$ s.t. $\gamma_{W_1}([a,b[) = \overline{h}'_{W_1,-}(b) - \overline{h}'_{W_1,-}(a)$ a.s., where $\overline{h}'_{W_1,-}(.) := h'_{W_1,-}(.)\mathbb{1}_{]\underline{u},\overline{u}]}(.) + h'_{W_1,+}(.)\mathbb{1}_{\{\underline{u}\}}(.)$ a.s. with $h(.) := -u(.)$.*

Proof. By Lemma A.3 and A.4 on p. OA.6, the extended left-derivative $\overline{h}'_{W_1,-}(.)$ is increasing and left continuous. Therefore, by a standard result for Lebesgue-Stieltjes integrals (e.g., Aliprantis and Border 2006/1994, Theorem 10.48 and comment just below), there exists a unique σ-finite Borel measure γ_{W_1} on $[\underline{u},\overline{u}]$ s.t. $\gamma_{w_1}([a,b[) = \overline{h}'_{-,W_1}(b) - \overline{h}'_{-,W_1}(a)$, $\forall (a,b) \in [\underline{u},\overline{u}]^2$ a.s.. In fact, the measure γ_{W_1} is finite a.s., because, $\forall A \in \mathcal{B}([\underline{u},\overline{u}])$, $\gamma_{W_1}(A) \leqslant \overline{h}'_{-,W_1}(\overline{u}) - \overline{h}'_{-,W_1}(\underline{u}) = h'_{-,W_1}(\overline{u}) - h'_{+,W_1}(\underline{u}) < \infty$ a.s. where the last inequality follows from Lemma A.4 on p. OA.7. Now, $\{[a,b[\colon (a,b) \in [\underline{u},\overline{u}]^2\}$ is a π-system that generates the Borel σ-algebra $\mathcal{B}([\underline{u},\overline{u}])$ (e.g., Aliprantis and Border 2006/1994, Lemma 4.19-4.20), and, for all $(a,b) \in [\underline{u},\overline{u}]^2$, $w_1 \mapsto \overline{h}'_{-,w_1}(b) - \overline{h}'_{-,W_1}(a)$ is Borel measurable because, for all $x \in [\underline{u},\overline{u}]$, the left derivative $w_1 \mapsto h'_{-,w_1}(x)$ inherits the measurability of $w_1 \mapsto h_{w_1}(a) := -u_{w_1}(x)$ by stability of measurability under limits (e.g., Aliprantis and Border 2006/1994, Theorem 4.27). Thus, by a standard result about random finite measures (e.g., Kallenberg 2002/1997, Lemma 1.40, which immediately extends to finite measures), the result follows. □

Lemma A.3 (Extended conditional left-derivative)**.** *Let $h_{W_1} : [\underline{u},\overline{u}] \to \mathbf{R}$ be a convex decreasing function indexed by a random variable W_1. Then, if $\mathbb{E}|h'_{W_1,+}(\underline{u})| < \infty$ and $\mathbb{E}|h'_{W_1,-}(\overline{u})| < \infty$,* *there exists a.s. a finite extended left-derivative on* $[\underline{u}, \overline{u}]$,

$$\overline{h}'_{W_1,-}(x) := \begin{cases} h'_{W_1,-}(x) & \forall x \in]\underline{u}, \overline{u}] \\ h'_{W_1,+}(x) & \text{for } x = \underline{u} \end{cases}$$

which is

(i) left-continuous,

(ii) increasing, and

(iii) negative.

Proof. It follows from the convexity of $h(.)$. □

Lemma A.4. *Let* $h_{W_1} : [\underline{u}, \overline{u}] \to \mathbf{R}$ *be a convex decreasing function indexed by a random variable* W_1*. Let* $\check{u}_{W_1}$ *be a random variable s.t.* $\check{u}_{W_1} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u_{W_1}(x) = 0\}\right\}$*, where* $u_{W_1}(.) := -h_{W_1}(.)$*. Then* $\mathbb{E}|h'_{W_1,+}(\underline{u})| < \infty$ *and* $\mathbb{E}|h'_{W_1,-}(\overline{u})| < \infty$*, iff,* $\mathbb{E}|h'_{W_1,+}(\underline{u})| < \infty$ *and* $\mathbb{E}|h'_{W_1,-}(\check{u}_{W_1})| < \infty$*.*

Proof. It follows from the increasing slope criterion for convex functions and the definition of $\check{u}_{W_1}$. □

Lemma A.5. *Let* $h_{W_1} : [\underline{u}, \overline{u}] \to \mathbf{R}$ *be a convex function indexed by a random variable* W_1 *s.t.* $\mathbb{E}|h_{W_1}(\underline{u})| < \infty$, $\mathbb{E}|h'_{W_1,+}(\underline{u})| < \infty$ *and* $\mathbb{E}|h'_{W_1,-}(\overline{u})| < \infty$*. If* X *is a random variable with its support in* $[\underline{u}, \overline{u}]$, $\mathbb{E}|h_{W_1}(X)| < \infty$*.*

Proof. By the increasing slope criterion for convex functions and its corollaries (e.g., Aliprantis and Border 2006/1994, Theorem 7.21-7.22), for all $x \in]\underline{u}, \overline{u}]$,

$$\begin{aligned} h'_{W_1,+}(\underline{u}) \leqslant \frac{h_{W_1}(x) - h_{W_1}(\underline{u})}{x - \underline{u}} \leqslant h'_{W_1,-}(\overline{u}) \\ \Rightarrow h_{W_1}(\underline{u}) + h'_{W_1,+}(\underline{u})(x - \underline{u}) \leqslant h_{W_1}(x) \leqslant h_{W_1}(\underline{u}) + h'_{W_1,-}(\overline{u})(x - \underline{u}) \end{aligned}$$

Moreover, the latter equality is also true if $x = \underline{u}$. Now, on one hand, if $0 \leqslant h_{W_1}(x)$, then $|h_{W_1}(X)| \leqslant |h_{W_1}(\underline{u}) + h'_{W_1,-}(\overline{u})(X - \underline{u})|$, and, on the other hand, if $h_{W_1}(x) \leqslant 0$, then $|h_{W_1}(X)| \leqslant |h_{W_1}(\underline{u}) + h'_{W_1,+}(\underline{u})(X - \underline{u})|$. Thus, for any random variable X with support in $[\underline{u}, \overline{u}]$,

$$\begin{aligned} |h_{W_1}(X)| &\leqslant |h_{W_1}(\underline{u}) + h'_{W_1,-}(\overline{u})(X - \underline{u})| + |h_{W_1}(\underline{u}) + h'_{W_1,+}(\underline{u})(X - \underline{u})| \\ &\overset{(a)}{\leqslant} 2|h_{W_1}(\underline{u})| + |h'_{W_1,-}(\overline{u})||X - \underline{u}| + |h'_{W_1,+}(\underline{u})||X - \underline{u}| \\ &\overset{(b)}{\leqslant} 2|h_{W_1}(\underline{u})| + |h'_{W_1,-}(\overline{u})||\overline{u} - \underline{u}| + |h'_{W_1,+}(\underline{u})||\overline{u} - \underline{u}| \\ \overset{(c)}{\Rightarrow} \mathbb{E}|h_{W_1}(X)| &\leqslant 2\mathbb{E}|h_{W_1}(\underline{u})| + \mathbb{E}|h'_{W_1,-}(\overline{u})||\overline{u} - \underline{u}| + \mathbb{E}|h'_{W_1,+}(\underline{u})||\overline{u} - \underline{u}| \overset{(d)}{<} \infty \end{aligned}$$

(a) Apply triangle inequality, and note that the absolute value of a product is equal to the product of the absolute values. *(b)* By assumption, $\underline{u} \leqslant X \leqslant \overline{u}$. *(c)* Monotonicity and linearity of integrals (e.g., Aliprantis and Border 2006/1994, Theorem 11.13). *(d)* By assumption, $\mathbb{E}|h_{W_1}(\underline{u})| < \infty$, $\mathbb{E}|h'_{W_1,+}(\underline{u})| < \infty$ and $\mathbb{E}|h'_{W_1,-}(\overline{u})| < \infty$. □

A.2 Proposition 1

Assumption 2 (Weak convergence of normalized integrated CDF& c_T)**.** *Denote the weak convergence with "$\rightsquigarrow$." As $T \to \infty$,*

$$\sqrt{T}\begin{pmatrix} \hat{F}_S^{(2)} - F_S^{(2)} \\ \hat{F}_L^{(2)} - F_L^{(2)} \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{H}_S \\ \mathbb{H}_L \end{pmatrix}$$

where the process $\{\mathbb{H}(z)\}_{z\in[\underline{u},\overline{u}]} := \{(\mathbb{H}_S(z)\ \mathbb{H}_L(z))'\}_{z\in[\underline{u},\overline{u}]}$ has a tight measurable Borel measurable version that lies in the space $UC([\underline{u},\overline{u}],\rho)$ of (uniformly) continuous functions on $[\underline{u},\overline{u}]$ endowed with the supremum norm ρ. Moreover, c_T converges sufficiently slowly to $\underline{u}$ from above.

Assumption 3 (Strict stationarity with strong mixing)**.** *The bivariate process $(\underline{r}_t)_{t=1}^T := (r_{S,t}\ r_{L,t})_{t=1}^T$ is strictly stationary and α-mixing.*

Assumption 3 is often required to check Assumption 2, so the former is not really more restrictive than the latter.

Lemma A.6 (Asymptotic limit of KS_T^*)**.** *Under Assumptions 1 and 2,*

(i) if H_0 holds, then, for T big enough, $\sup_{z\in\mathbf{I}_T}\left|\hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z)\right| = 0$ with probability one (w.p.1.).

(ii) if H_0 does not hold, then as $T \to \infty$, $\mathrm{KS}_T^* = \sup_{z\in\mathbf{I}_T}\left|\hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z)\right|$ converges to a non-zero positive constant $\overline{\mathrm{KS}}^*$ w.p.1.

Proof. It follows from a reasoning along the lines of the mathematical arguments after Proposition 1 on p. 11. □

Lemma A.7 (Subsampling CDF of $\mathrm{KS}_{T,i}^*$)**.** *Assume $(b_T) \in [\![1,\infty[\![^{\mathbf{N}}$ s.t. $\lim_{T\to\infty} b_T = \infty$ and $\lim_{T\to\infty}\frac{b_T}{T} = 0$. Under Assumptions 1 , 2, and 3, if H_0 does not hold,*

(i) for all $x \in \mathbf{R}\setminus\{\overline{\mathrm{KS}}^*\}$, with probability one, as $T \to \infty$, $\hat{G}_{T,b_T}^0(x) \to \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)$ where $\hat{G}_{T,b_T}^0(x) := \frac{1}{T-b_T+1}\sum_{i=1}^{T-b_T+1}\mathbb{1}(\mathrm{KS}_{T,i}^* \leqslant x)$; and

(ii) for all $\alpha \in [0,1[$, as $T \to \infty$, $g_{T,b_T,1-\alpha}^0 \to \overline{\mathrm{KS}}^*$ with probability one, where $g_{T,b_T,1-\alpha}^0 := \inf\{y : 1-\alpha \leqslant \hat{G}_{T,b_T}^0(y)\}$

Proof. *(i)* By triangle inequality for the L_2 norm $|.|_2$,

$$
\begin{aligned}
|\hat{G}^0_{T,b_T}(x) - \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)|_2 &\leqslant |\hat{G}^0_{T,b_T}(x) - \mathbb{E}[\hat{G}^0_{T,b_T}(x)]|_2 + |\mathbb{E}[\hat{G}^0_{T,b_T}(x)] - \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)|_2 \\
&= \sqrt{\mathbb{V}[\hat{G}^0_{T,b_T}(x)]} + |\mathbb{P}(\mathrm{KS}^*_{T,1} \leqslant x) - \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)|_2
\end{aligned}
$$

because $\mathbb{E}[\hat{G}^0_{T,b_T}(x)] = \mathbb{E}[\frac{1}{T-b_T+1}\sum_{i=1}^{T-b_T+1}\mathbb{1}(\mathrm{KS}^*_{T,i} \leqslant x)] = \mathbb{E}[\mathbb{1}(\mathrm{KS}^*_{T,1} \leqslant x)] = \mathbb{P}(\mathrm{KS}^*_{T,1} \leqslant x)$ where the second equality comes from strict stationarity (i.e., Assumption 3). Now, for all $x \in \mathbf{R} \setminus \{\overline{\mathrm{KS}}^*\}$, as $T \to \infty$, $|\mathbb{P}(\mathrm{KS}^*_{T,1} \leqslant x) - \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)|_2)) = |\mathbb{P}(\mathrm{KS}^*_{T,1} \leqslant x) - \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)| \to 0$ w.p.1 because $\mathrm{KS}^*_{T,1} = \mathrm{KS}^*_{b_T}$, which converges in distribution to the non-zero positive constant $\overline{\mathrm{KS}}^*$ by Lemma A.6ii. Thus, it is sufficient to prove that $\mathbb{V}[\hat{G}^0_{T,b_T}(x)] \to 0$, as $T \to \infty$ w.p.1. using strong mixing.

(ii) Let $\eta > 0$ and $\epsilon > 0$ s.t. $1 - \alpha < 1 - \epsilon$ & $\epsilon < 1 - \alpha$, i.e., $\epsilon \in]0, \min\{\alpha, 1-\alpha\}[$. By (i), w.p.1, there exists $\bar{T} \in [\![1, \infty[\![$ s.t. $T \geqslant \bar{T}$ implies

$$
\begin{aligned}
&\begin{cases} 1 - \hat{G}^0_{T,b_T}(\overline{\mathrm{KS}}^* + \eta) < \epsilon \\ \hat{G}^0_{T,b_T}(\overline{\mathrm{KS}}^* - \eta) - 0 < \epsilon \end{cases} \\
\Leftrightarrow &\begin{cases} 1 - \epsilon < \hat{G}^0_{T,b_T}(\overline{\mathrm{KS}}^* + \eta) \\ \hat{G}^0_{T,b_T}(\overline{\mathrm{KS}}^* - \eta) < \epsilon \end{cases} \\
\Rightarrow &\begin{cases} 1 - \alpha < \hat{G}^0_{T,b_T}(\overline{\mathrm{KS}}^* + \eta) \\ \hat{G}^0_{T,b_T}(\overline{\mathrm{KS}}^* - \eta) < 1 - \alpha \end{cases}
\end{aligned}
$$

because $\epsilon > 0$ s.t. $1 - \alpha < 1 - \epsilon$ & $\epsilon < 1 - \alpha$. Now, $g^0_{T,b_T,1-\alpha} := \inf\{y : 1 - \alpha \leqslant \hat{G}^0_{T,b_T}(y)\}$, where $\hat{G}^0_{T,b_T}(.)$ is an increasing function. Thus, w.p.1, $\forall T \geqslant \bar{T}$, $\overline{\mathrm{KS}}^* - \eta < g_{T,b_T,1-\alpha} \leqslant \overline{\mathrm{KS}}^* + \eta$. □

Lemma A.8 (Centered Subsampling CDF of $\mathrm{KS}^*_{T,i}$)**.** *Assume* $(b_T) \in [\![1, \infty[\![^{\mathbf{N}}$ *s.t.* $\lim_{T\to\infty} b_T = \infty$ *and* $\lim_{T\to\infty} \frac{b_T}{T} = 0$. *Under Assumptions 1 , 2, and 3, if* H_0 *does not hold,*

(i) for all $x \in \mathbf{R} \setminus \{\overline{\mathrm{KS}}^*\}$, *w.p.1, as* $T \to \infty$, $\check{G}^0_{T,b_T}(x) \to \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)$ *where* $\check{G}^0_{T,b_T}(x) := \frac{1}{T-b_T+1}\sum_{i=1}^{T-b_T+1}\mathbb{1}(\mathrm{KS}^*_{T,i} - \mathrm{KS}^*_T \leqslant x)$*; and*

(ii) for all $\alpha \in [0, 1[$, *as* $T \to \infty$, $\check{g}^0_{T,b_T,1-\alpha} \to \overline{\mathrm{KS}}^*$ *w.p.1, where* $\check{g}^0_{T,b_T,1-\alpha} := \inf\{y : 1 - \alpha \leqslant \check{G}^0_{T,b_T}(y)\}$

Proof. Adapt the proof of Lemma A.7. □

Proof of Proposition 1. Case 1.1: H_0 *holds. Uncentered subsampling.* By definition of $\hat{F}^{(2)}_{L\wedge S,b_T,i}(.)$, $0 \leqslant \sqrt{b_T}\mathrm{KS}^*_{b_T,i} := \sqrt{b_T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L,b_T,i}(z) - \hat{F}^{(2)}_{L\wedge S,b_T,i}(z)|$. Thus, under Assumptions 1 and 2, by Lemma A.6i, for T big enough, w.p.1, $\sqrt{T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_L(z) - \hat{F}^{(2)}_{L\wedge S}(z)| = 0 \leqslant \sqrt{b_T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L,b_T,i}(z) - \hat{F}^{(2)}_{L\wedge S,b_T,i}(z)|$, $\forall i \in [\![1, T - b_T + 1]\!]$. Therefore, $\sqrt{T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_L(z) - \hat{F}^{(2)}_{L\wedge S}(z)|$ is smaller than any quantile of the distribution of the $\sqrt{b_T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L,b_T,i}(z) - \hat{F}^{(2)}_{L\wedge S,b_T,i}(z)|$.

Case 1.2: H_0 *holds. Centered subsampling.* Under Assumptions 1 and 2, by Lemma A.6i, for T big enough, w.p.1, $\sqrt{T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}_L^{(2)}(z)-\hat{F}_{L\wedge S}^{(2)}(z)|=0$. Thus,for T big enough, w.p.1, the centered subsampled statistics $\sqrt{b_T}\dot{\mathrm{KS}}^*_{T,i}$ are equal to the uncentered susbsampled test statistic $\sqrt{b_T}\mathrm{KS}^*_{T,i}$, i.e., $\sqrt{b_T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}_{L,b_T,i}^{(2)}(z)-\hat{F}_{L\wedge S,b_T,i}^{(2)}(z)|=\sqrt{b_T}[\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}_{L,b_T,i}^{(2)}(z)-\hat{F}_{L\wedge S,b_T,i}^{(2)}(z)|-\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}_L^{(2)}(z)-\hat{F}_{L\wedge S}^{(2)}(z)|]$. Thus, the same proof as in the uncentered case applies.

Case 2.1: H_0 *does not holds. Uncentered subsampling, i.e.,* $\hat{c}_{1-\alpha}:=\inf\{x:1-\alpha\leqslant\hat{G}_{T,b_T}(x)\}$ *where* $\hat{G}_{T,b_T}(x):=\frac{1}{T-b_T+1}\sum_{i=1}^{T-b_T+1}\mathbb{1}(\sqrt{b_T}\mathrm{KS}^*_{T,i}\leqslant x)$.

By definition of $g_{T,b_T,1-\alpha}$,

$$\begin{aligned}
&\left\{g_{T,b_T,1-\alpha}<\sqrt{T}\mathrm{KS}^*_T\right\}\\
=&\left\{\inf\{x:1-\alpha\leqslant\hat{G}_{T,b_T}(x)\}<\sqrt{T}\mathrm{KS}^*_T\right\}\\
=&\left\{\inf\{\frac{x}{\sqrt{b_T}}:1-\alpha\leqslant\hat{G}_{T,b_T}(x)\}<\sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T\right\}\\
\overset{(a)}{=}&\left\{\inf\{y:1-\alpha\leqslant\hat{G}_{T,b_T}(\sqrt{b_T}y)\}<\sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T\right\}\\
\overset{(b)}{=}&\left\{\inf\{y:1-\alpha\leqslant\hat{G}^0_{T,b_T}(y)\}<\sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T\right\}\\
=&\left\{g^0_{T,b_T,1-\alpha}<\sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T\right\}
\end{aligned}$$

(a) Put $y=x/b_T$. *(b)* $\hat{G}^0_{T,b_T}(y)=\frac{1}{T-b_T+1}\sum_{t=1}^{T-b_T+1}\mathbb{1}(\mathrm{KS}^*_{T,i}\leqslant y)=\frac{1}{T-b_T+1}\sum_{t=1}^{T-b_T+1}\mathbb{1}(\sqrt{b_T}\mathrm{KS}^*_{T,i}\leqslant\sqrt{b_T}y)=\hat{G}_{T,b_T}(\sqrt{b_T}y)$

Now, under Assumptions 1 , 2, and 3, , $\lim_{T\to\infty}\mathbb{P}\Big\{g^0_{T,b_T,1-\alpha}<\sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T\Big\}=1$ because $\lim_{T\to\infty}g^0_{T,b_T,1-\alpha}=\overline{\mathrm{KS}}^*\leqslant\lim_{T\to\infty}\sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T=\lim_{T\to\infty}\sqrt{\frac{T}{b_T}}\overline{\mathrm{KS}}^*=\infty$ w.p.1. by Lemma A.7ii and $\lim_{T\to\infty}\frac{b_T}{T}=0$ by assumption.

Case 2.2: H_0 *does not holds. Centered subsampling, i.e.,* $\hat{c}_{1-\alpha}:=\inf\{x:1-\alpha\leqslant\hat{G}_{T,b_T}(x)\}$ *where* $\hat{G}_{T,b_T}(x):=\frac{1}{T-b_T+1}\sum_{i=1}^{T-b_T+1}\mathbb{1}(\sqrt{b_T}(\mathrm{KS}^*_{T,i}-\mathrm{KS}^*_T)\leqslant x)$. Follow the same reasoning as in the case 2.1. □

A.3 Proof of Proposition 3

Proof. 1st case: H_0 *is true.* By positivity and monotonicity of probability measures, $0\leqslant\mathbb{P}\left(\{\hat{c}_{1-\alpha}<\sqrt{T}\mathrm{KS}^*_T\}\cap F_T\right)\leqslant\mathbb{P}(\hat{c}_{1-\alpha}<\sqrt{T}\mathrm{KS}^*_T)$. Now, if H_0 is true, $\lim_{T\to\infty}\mathbb{P}(\hat{c}_{1-\alpha}<\sqrt{T}\mathrm{KS}^*_T)=0$. Thus, the result follows from the squeeze theorem because $\lim_{T\to\infty}\mathbb{P}\left(\{\hat{c}_{1-\alpha}<\sqrt{T}\mathrm{KS}^*_T\}\right)\times\mathbb{P}\left(F_T\right)=0$

2st case: H_0 *is wrong.* On one hand, by additivity of probability measures, for all $T \in [\![1, \infty[\![$,

$$
\begin{aligned}
& \mathbb{P}(F_T) = \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}) + \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}^c) \\
\Rightarrow\ & \mathbb{P}(F_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}) = \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}^c) \\
\overset{(a)}{\Rightarrow}\ & \mathbb{P}(F_T)\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}) \leqslant \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}^c) \\
\overset{(b)}{\Rightarrow}\ & \mathbb{P}(F_T)\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}) \leqslant 1 - \mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T)
\end{aligned}
$$

(a) $\mathbb{P}(F_T)\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}) \leqslant \mathbb{P}(F_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\})$ because $\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) \in [0, 1]$ by definition of probability. *(b)* By monotonicity of probability measures, $\mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}^c) \leqslant \mathbb{P}(\{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}^c) = 1 - \mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T)$.

On the other hand, for all $T \in [\![1, \infty[\![$,

$$
\begin{aligned}
& \mathbb{P}(F_T)\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - \mathbb{P}(F_T) \leqslant \mathbb{P}(F_T)\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\}) \\
\Leftrightarrow\ & \mathbb{P}(F_T)[\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - 1] \leqslant \mathbb{P}(F_T)\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - \mathbb{P}(F_T \cap \{\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T\})
\end{aligned}
$$

Now, by Proposition 1ii (p. 11), $\lim_{T\to\infty} \mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) = 1$, so that $\lim_{T\to\infty} 1 - \mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) = 0$ and $\lim_{T\to\infty}[\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) - 1] = \lim_{T\to\infty} \mathbb{P}(F_T)[1 - \mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T)] = 0$ because $\mathbb{P}(F_T)$ is bounded. Therefore, the result follows from the squeeze theorem. □

A.4 Proof of optimality condition and risk compensation

The following Proposition A.1 establishes the optimality condition and the risk compensation for factors in the one-period case, and in the multiperiod case. The one-period case corresponds to $T = 1$ and a given C_0 because a strictly increasing utility functions implies $C_1 = W_1$ in a one-period framework.

Proposition A.1 (Optimality condition & risk compensation)**.** *Assume the factor $r_{L,t} - r_{S,t}$ is different from zero with probability one, i.e., $\mathbb{P}(r_L - r_S \neq 0) = 1$. Assume time-additive utility functions $U(C_{0:T}) := \sum_{t=0}^{T} \beta^t \mathbb{E}[u(C_t)]$ where $\beta > 0$ is the time discount factor, $T \in [\![1, \infty[\![$ the time horizon, and $u(.)$ a continuously differentiable von Neuman-Morgenstern utility function. Under Assumption 1(a), if $C_{0:T} := (C_0, C_1, \ldots, C_T)$ is a locally optimal consumption process with values in the interior of $[\underline{u}, \overline{u}]$ for an individual with utility function $U(C_{0:T}) := \sum_{t=0}^{T} \beta^t \mathbb{E}[u(C_t)]$, then, for any time period $\dot{t} \in [\![1, T]\!]$ at which the factor $r_{L,\dot{t}} - r_{S,\dot{t}}$ is freely tradable in a neighborhood of $C_{\dot{t}}$,*

- *(i) [Optimality condition] $\mathbb{E}[u'(C_{\dot{t}})(r_{L,\dot{t}} - r_{S,\dot{t}})] = 0$; and*
- *(ii) [Risk compensation] under the additional assumption that $\mathbb{E}[u'(C_{\dot{t}})] \neq 0$, $\mathbb{E}(r_{L,\dot{t}} - r_{S,\dot{t}}) = -\dfrac{1}{\mathbb{E}[u'(C_{\dot{t}})]}\mathbb{C}\mathrm{ov}(u'(C_{\dot{t}}), r_{L,\dot{t}} - r_{S,\dot{t}})$.*

Proof. *(i)* For any $\dot{t} \in [\![1, T]\!]$, define the consumption process $\tilde{C}_{0:T} := (\tilde{C}_0, \tilde{C}_1, \ldots, \tilde{C}_T)$ s.t., $\forall k \in [\![1, T]\!] \setminus \{\dot{t}\}$, $\tilde{C}_k = C_k$ and $\tilde{C}_{\dot{t}} = C_{\dot{t}} + \epsilon(R_{L,\dot{t}} - R_{S,\dot{t}})$ where $\epsilon > 0$. Then, on one hand, by

Assumption 1(a), for ϵ small enough, $C_i + \epsilon(R_{L,i} - R_{S,i})$ is in any arbitrary small neighborhood of C_i so the local optimality of $C_{0:T}$ implies

$$\begin{aligned} & 0 \leqslant U(C_{0:T}) - U(\tilde{C}_{0:T}) = \beta\mathbb{E}[u(C_i)] - \beta\mathbb{E}[u(C_i + \epsilon(R_{L,i} - R_{S,i}))] \\ \overset{(a)}{\Leftrightarrow} & 0 \leqslant \mathbb{E}\left[\frac{[u(C_i) - u(C_i + \epsilon(R_{L,i} - R_{S,i}))]}{\epsilon(R_{L,i} - R_{S,i})}(R_{L,i} - R_{S,i})\right] \overset{(b)}{\to} \mathbb{E}[u'(C_i)(R_{L,i} - R_{S,i})] \text{ , as } \epsilon \downarrow 0. \end{aligned}$$

(a) Divide both sides by $1/(\beta\epsilon)$, and multiply the numerator and the denominator of the fraction with $(R_{L,i} - R_{S,i})$. *(b)* By Assumption 1(a), for ϵ small enough $C_i + \epsilon(R_{L,i} - R_{S,i})$ is in the interior of $[\underline{u}, \overline{u}]$ with probability one. Now, by the mean-value theorem and the continuity of the derivative on $[\underline{u}, \overline{u}]$, $\epsilon \mapsto \frac{[u(C_i) - u(C_i + \epsilon(R_{L,i} - R_{S,i}))]}{\epsilon(R_{L,i} - R_{S,i})}$ is bounded for ϵ small enough. Thus, by the definition of derivatives, Lebesgue's dominated convergence theorem yields the result.

On the other hand, following a similar reasoning with $\tilde{C}_i = C_i - \epsilon(R_{L,i} - R_{S,i})$ implies $\mathbb{E}[u'(C_i)(R_{L,i} - R_{S,i})] \leqslant 0$. Thus, the result follows.

(ii) Standard calculations yield

$$\begin{aligned} & \mathbb{E}[u'(C_i)(r_{L,i} - r_{S,i})] = 0 \\ \Leftrightarrow\ & \mathbb{C}\mathrm{ov}(u'(C_i), r_{L,i} - r_{S,i}) + \mathbb{E}[u'(C_i)]\mathbb{E}(r_{L,i} - r_{S,i}) = 0 \\ \Leftrightarrow\ & \mathbb{E}(r_{L,i} - r_{S,i}) = -\frac{\mathbb{C}\mathrm{ov}(u'(C_i), r_{L,i} - r_{S,i})}{\mathbb{E}[u'(C_i)]} \end{aligned}$$

□

Remark 1 (Infinite horizon). Inspection of the proof shows Proposition A.1 can be extended to infinite horizon under the additional assumption that $\sum_{t=0}^{\infty} |\beta^t \mathbb{E}[u(C_t)]| < \infty$. ◇

Remark 2. Another way to derive the optimality condition is to go through standard Euler equations. We do not follows this other way because it would require more assumptions: It would at least require each leg of the factor to be freely tradable, separately. ◇

A.5 Supplementary results

The following result seems to be known, although no proofs or statements is available in the literature to the best of our knowledge.

Theorem A.3 (Equivalent characterizations of conditional SSD). *Assume that the support of the random variables r_L and r_S is a subset of $[\underline{u}, \overline{u}] \subset \mathbf{R}$ with $\underline{u} \neq \overline{u}$. Then the following statements are equivalent.*

(i) For all real-valued, concave and increasing function $u_{W_1}(.)$ defined on $[\underline{u}, \overline{u}]$ and Borel measurable w.r.t. the index W_1 s.t. $\mathbb{E}|u_{W_1}(\underline{u})| < \infty$, $\mathbb{E}|u'_{W_1,+}(\underline{u})| < \infty$ and $\mathbb{E}|u'_{W_1,-}(\overline{u})| < \infty$, the following inequality holds $\mathbb{E}[u_{W_1}(r_S)|W_1] \leqslant \mathbb{E}[u_{W_1}(r_L)|W_1]$ a.s.

(ibis) *For all real-valued, concave and increasing function* $u(.)$ *on* $[\underline{u}, \overline{u}]$ *s.t.* $u'_+(\underline{u}) \in \mathbf{R}$ *and* $u'_-(\overline{u}) \in \mathbf{R}$*, the following inequality holds* $\mathbb{E}[u(r_S)|W_1] \leqslant \mathbb{E}[u(r_L)|W_1]$ *a.s.*

(ii) *For all* $z \in \mathbf{R}$, $\mathbb{E}[(z - r_L)^+|W_1] \leqslant \mathbb{E}[(z - r_S)^+|W_1]$ *a.s.*

(iii) *For all* $z \in \mathbf{R}$, $F^{(2)}_{L|W_1}(z|W_1) \leqslant F^{(2)}_{S|W_1}(z|W_1)$ *a.s., where* $F^{(2)}_{L|W_1}(z|W_1) := \int_{\underline{u}}^{z} F_{L|W_1}(y|W_1)\mathrm{d}y$ *a.s.*

Proof of Theorem A.3. Repeat the proof of Theorem A.2 with $\overline{u}$ in lieu of $\breve{u}_{W_1}$. □

A.6 Proposition 5

Assumption 4 (Conditional no touching without crossing)**.** *If there exists* $\dot{z} \in]\underline{u}, \overline{u}]$ *s.t.* $F^{(2)}_{L|M}(\dot{z}) = F^{(2)}_{S|M}(\dot{z})$*, then there exists* $\ddot{z} \in]\underline{u}, \overline{u}]$ *s.t.* $F^{(2)}_{S|M}(\ddot{z}) < F^{(2)}_{L|M}(\ddot{z})$.

Assumption 5 (Weak convergence)**.** **(a)** *If* H_0 *holds,* $\sqrt{T}\mathrm{C}^*_T$ *converges weakly to a limiting law, as* $T \to \infty$. **(b)** *As* $T \to \infty$, $\sqrt{T}(\hat{C}^{(2)} - C^{(2)}) \rightsquigarrow \mathbb{H}_C$*, where* $\mathbb{H}_C$ *has a tight measurable Borel measurable version that lies in the space of uniformly continuous functions endowed with the supremum norm* ρ.

Assumption 6 (Strict stationarity with strong mixing)**.** *The process* $(r_{S,t}\ r_{L,t}\ r_{M,t})_{t=1}^{T}$ *is strictly stationary and* α*-mixing.*

Proof of Proposition 5. (i) Use properties of least concave majorant (Durot and Tocquet 2003, Sec. 2), and adapt the proof of Beran (1984, Theorem 1) along the lines of Politis et al. (1999, Theorem 3.2.1).

(ii) It follows from the same logic as the proof of Proposition 1(ii). □

B Monte-Carlo simulations

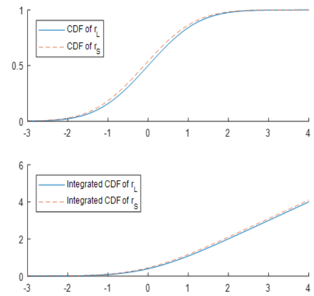
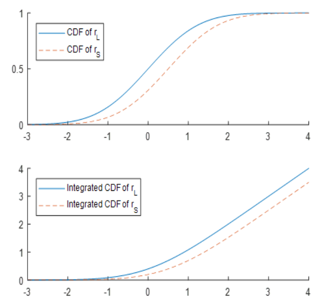
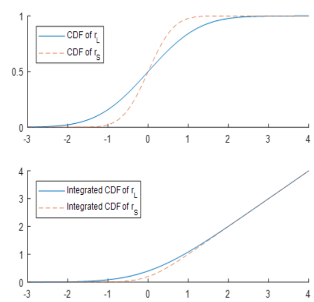
The objective of this section is to (i) explore the finite-sample behaviour of the tests; (ii) compare them with alternative implementations.

B.1 DGPs

B.1.1 Stylized DGPs

The stylized DGPs, which are taken from Whang (2019, p. 225–227) and displayed in Table A.1 (p. OA.14), allow to assess the performance of the tests in well-understood situations. A Gaussian distribution is strictly preferred by all risk-averse agents to another Gaussian distribution if its mean and variance are smaller.

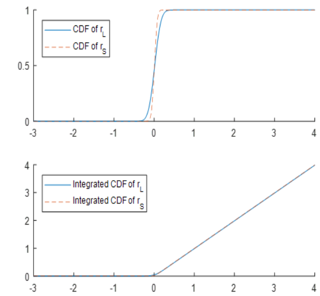
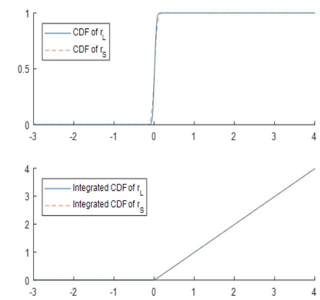
Table A.1: Stylized DGPs

H_0	**DGP**	**Plots of CDF & Integrated CDF**
True	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ -.1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$	
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ .5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$	
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & .5^2 \end{bmatrix} \right)$	

B.1.2 DGPs calibrated to data

In Table A.2 (p. OA.15), the DGPs are calibrated to data. They allow to assess the finite-sample performance of the test in situations that mimick the data. For this purpose, we calibrate Gaussian distributions to factors for which the null hypotheses are barely true (or false). More precisely, the mean and the variance are calibrated to the average and the empirical variance of the legs of the factor SIZE and the factor DY in original sample.

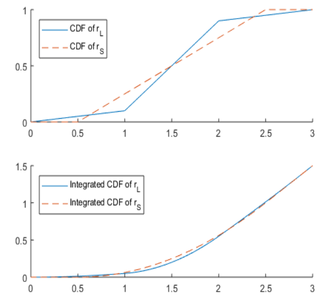
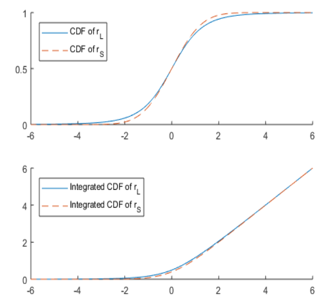
Table A.2: DGPs calibrated to data

H_0	**DGP**	**Plots of CDF & Integrated CDF**
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} .015 \\ .0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix}\right)$	
True	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} .011 \\ .010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix}\right)$	

B.1.3 Non-Gaussian DGPs with correlation calibrated to data

The non-Gaussian DGPs with correlation calibrated from data, which are displayed in Table A.6 (p. OA.20), correspond to examples of distributions mentioned in the stochastic dominance literature. The correlation is calibrated to the average correlation between the short and the long legs of factors in the original sample, that is .7. We rely on the NORTA algorithm (Cario and Nelson. 1997) to generate the data with the desired correlation and marginal distributions. The first DGP, which is adapted from Whang (2019, p. 10) and Rothschild and Stiglitz (1970, Sec. IV) is known to be a challenging DGP. The second DGP allows to assess the performance of the tests in the present of fat tails: Students distributions are leptokurtic.

Table A.3: Non-Gaussian DGPs with correlation calibrated to data

H_0	**DGP**	**Plots of CDF & Integrated CDF**
False	$\begin{cases} r_L \hookrightarrow .3\mathcal{U}_{[0,3]} + .7\mathcal{U}_{[1,2]} \\ r_S \hookrightarrow \mathcal{U}_{[.5,2.5]} \\ \mathbb{C}\mathrm{or}(r_S, r_L) = .7 \end{cases}$	
False	$\begin{cases} r_L \overset{IID}{\hookrightarrow} \mathrm{t}(4) \\ r_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\mathrm{or}(r_S, r_L) = .7 \end{cases}$	

B.2 Unconditional Test

B.2.1 Number of grid points and subsample size b_T

Like other tests of stochastic dominance à la McFadden (1989), our test requires to choose the number of gridpoints used to approximate the supremum in the test statistic. In the literature, the usual number of gridpoints seems to be 100 or less (e.g., Barrett and Donald 2003, Whang 2019). For caution, we use 200, and we have checked that our simulation results are not affected up to two decimals after the dot if we double the number of nodes to 400.

Regarding the subsample size b_T, asymptotic theory requires $\lim_{T\to\infty} b_T = \infty$ and $\lim_{T\to\infty} \frac{b_T}{T} = 0$ (Propositions 1 and 5 on p. 11 & 21). This leaves a wide choice of subsample sizes. The trade off is the following. If b_T is too big (i.e., too close to the sample size T), the subsample statistics are too close to each other, so the subsampling distribution is too tight. Conversely, if b_T is too small (e.g., $b_T = 1$), the subsample statistics are too far from each other, so the subsampling distribution is too wide. While some automatic data-dependent methods have been to proposed to choose the subsample size b_T (e.g., Linton et al. 2005, Politis et al. 1999, Chap. 9), there is no consensus about which data-dependent methods to choose. Now, by the CLT, under general assumptions, the rate of convergence of estimators (i.e., the rate of accumulation of information) is $\sqrt{T}$, so we choose subsample size $b_T = \lfloor \sqrt{T} \rfloor$ where $\lfloor a \rfloor := \max\{n \in \mathbf{N} : n \leqslant a\}$. For robustness, we also tried $b_T = \lfloor m + \sqrt{T} \rfloor$ with $m \in \{5, 10, 20\}$, and $b_T = \left\lceil \frac{\eta T}{\log[\log(\mathrm{e}^{\mathrm{e}} + T)]} \right\rceil$ with $\eta \in \{.25, .5\}$

and where $\lceil a \rceil := \min\{n \in \mathbf{N} : a \leqslant n\}$ for all $a \in \mathbf{R}$.[15] Monte-Carlo simulations, which are available upon request, indicate that none of this alternatives work better than $b_T = \lfloor \sqrt{T} \rfloor$. Moreover, our empirical results appear qualitatively robust to these different subsample sizes. Thus, we stick to $b_T = \lfloor \sqrt{T} \rfloor$.

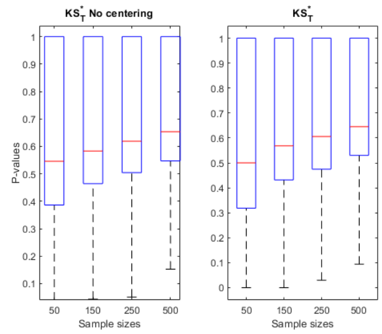
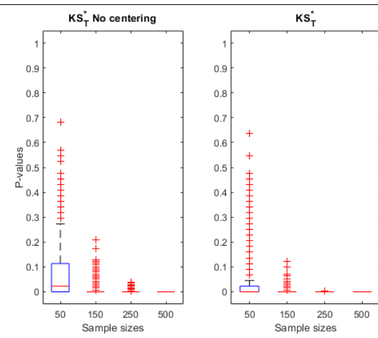
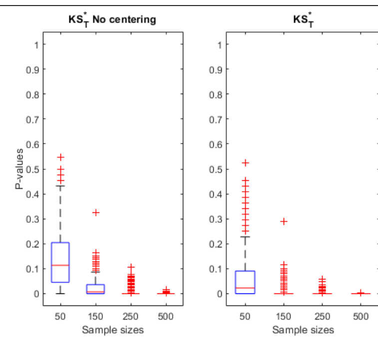
B.2.2 Results

We compare uncentered and centered block subsampling. In some situations, it has been found that centered subsampling outperforms the original uncentered subsampling in small sample (e.g., Chernozhukov and Fernández-Val 2005). Our analysis focuses on the boxplots of the p-values.

Overall, the different implementations of the tests appear to have a satisfactory finite-sample behaviour, i.e., the p-values are usually high under the null hypothesis, while the distribution of the p-values tends to converge to a point mass at zero under the alternative. Nevertheless, some patterns indicate some systematically different finite-sample behaviors. In particular, centered block subsampling implementation performs similarly to our uncentered, except that the p-values are generally smaller. Thus, for caution, in the empirical section of the main text, we only report results from our centered subsampling implementation so it goes against our main result. For the DGPs calibrated to data and the Non-Gaussian DGPs with correlation calibrated to data, the good finite-sample performance of the tests is partly due to the correlation between the short and the long legs : The higher the correlation, the less probable are crossing of the integrated empirical CDFs under the null hypothesis, and the more probable are crossing under the alternative hypothesis.

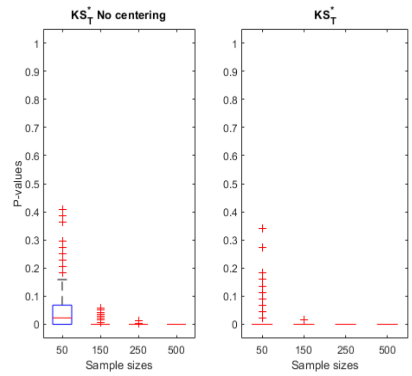
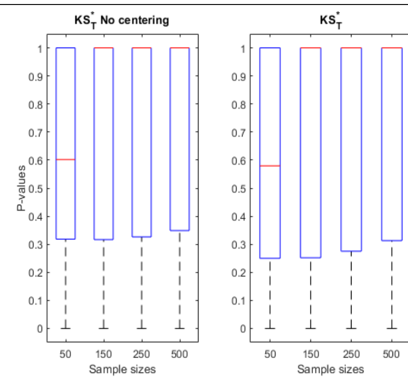
[15] The term e^{e} guarantees that the denominator is bigger than one, so the subsample size cannot be negative nor bigger than the sample size.

Table A.4: Monte-Carlo simulations of $\mathbb{KS}_T^*$: Stylized DGPs

H_0	**DGP**	**Boxplots of p-values**
True	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ -.1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$	
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ .5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$	
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & .5^2 \end{bmatrix}\right)$	

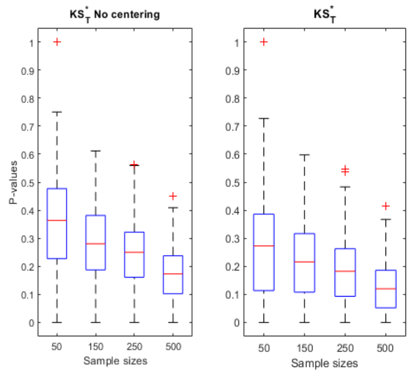
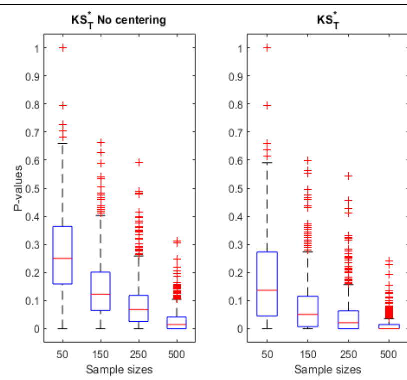
Note: The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of $\mathbb{KS}_T^*$ is approximated through block subsampling for "$\mathbb{KS}_T^*$ No centering," and centered block subsampling for "$\mathbb{KS}_T^*$." The block size is $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

Table A.5: Monte-Carlo simulations of $\mathbb{KS}_T^*$: Calibrated DGPs

<table>
<tr><th>H_0</th><th>DGP</th><th>Boxplots of p-values</th></tr>
<tr><td>False</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} .015 \\ .0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>True</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} .011 \\ .010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix}\right)$</td><td>
</td></tr>
</table>

Note: The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of $\mathbb{KS}_T^*$ is approximated through block subsampling for "$\mathbb{KS}_T^*$ No centering," and centered block subsampling for "$\mathbb{KS}_T^*$." The block size is $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

Table A.6: Monte-Carlo simulations of $\mathbb{KS}_T^*$:Non-Gaussian DGPs with correlation calibrated to data

H_0	**DGP**	**Boxplots of p-values**
False	$\begin{cases} r_L \overset{IID}{\hookrightarrow} .3\mathcal{U}_{[0,3]} + .7\mathcal{U}_{[1,2]} \\ r_S \overset{IID}{\hookrightarrow} \mathcal{U}_{[.5,2.5]} \\ \mathbb{C}\text{or}(r_S, r_L) = .7 \end{cases}$	
False	$\begin{cases} r_L \overset{IID}{\hookrightarrow} \mathrm{t}(4) \\ r_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\text{or}(r_S, r_L) = .7 \end{cases}$	

Note:The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of $\mathbb{KS}_T^*$ is approximated through block subsampling for "$\mathbb{KS}_T^*$ No centering," and centered block subsampling for "$\mathbb{KS}_T^*$." The block size is $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

B.3 Conditional tests

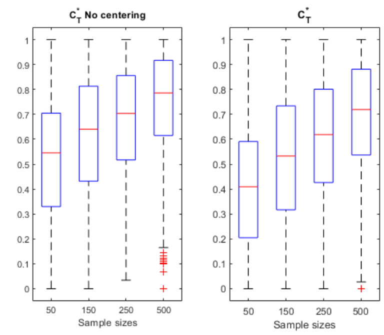
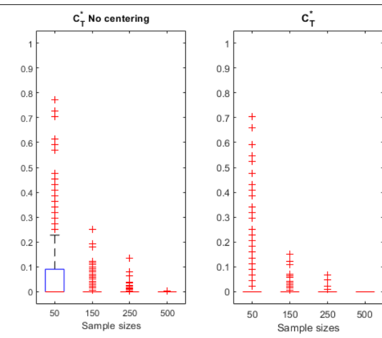
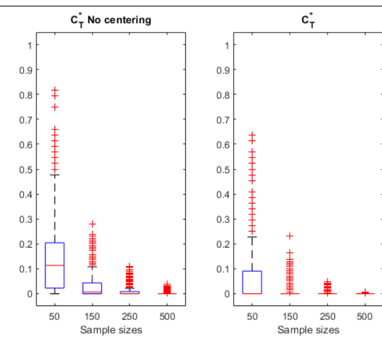
For ease of comparison, the parameterization and the DGPs are similar to the ones for the unconditional tests, except for a new common component. More precisely, we add a common independent Gaussian component $x \hookrightarrow \mathcal{N}(0, \sigma_x^2)$ to each of the DGPs. E.g., the first DGP is

$$\begin{bmatrix} r_L \\ r_S \end{bmatrix} = x + \begin{bmatrix} z_L \\ z_S \end{bmatrix}$$

where $x \overset{IID}{\hookrightarrow} \mathcal{N}(0, \sigma_x^2)$, $\begin{bmatrix} z_L \\ z_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$, and x is independent of $[z_L \quad z_S]'$. The parameter σ_x is calibrated to correspond to an estimate of the standard deviation of the monthly market returns, i.e., $\sigma_x = 4\%$. Regarding the parameterization, as in the unconditional test and for the same reasons, we keep the subsample size $b_T = \sqrt{T}$ and the number of nodes to 200.

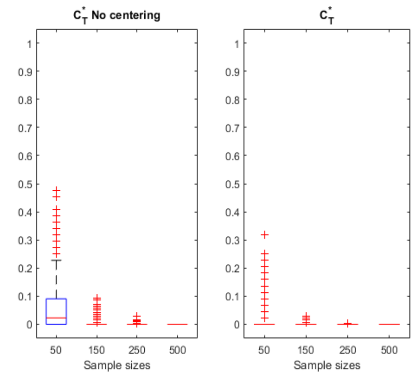
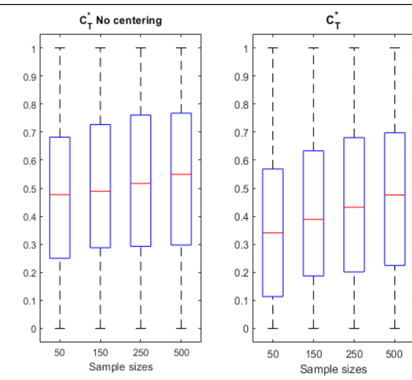
The patterns of the p-value distributions appear similar to the ones of the unconditional tests, namely smaller p-values for centered subsampling, better performance when the correlation between boths legs is higher.

Table A.7: Monte-Carlo simulations of C_T^*: Stylized DGPs

<table>
<tr><th>H_0</th><th>DGP</th><th>Boxplots of p-values</th></tr>
<tr><td>True</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} 0 \\ -.1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>False</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} 0 \\ .5 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$</td><td>
</td></tr>
<tr><td>False</td><td>$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & .5^2 \end{bmatrix}\right)$</td><td>
</td></tr>
</table>

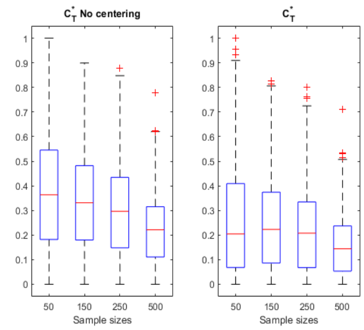
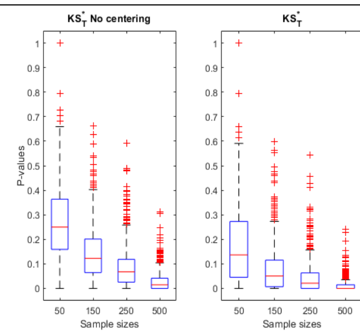
Note:The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of C_T^* is approximated through block subsampling for "C_T^* No centering," and centered block subsampling for "C_T^*." The block size is $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

Table A.8: Monte-Carlo simulations of C_T^*: Calibrated DGPs

H_0	**DGP**	**Boxplots of p-values**
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} .015 \\ .0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix} \right)$	
True	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} .011 \\ .010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix} \right)$	

Note: The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of C_T^* is approximated through block subsampling for "C_T^* No centering," and centered block subsampling for "C_T^*." The block size is $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

Table A.9: Monte-Carlo simulations of C_T^*: Non-Gaussian DGPs

H_0	**DGP**	**Boxplots of p-values**
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \begin{bmatrix} z_L \\ z_S \end{bmatrix}$ where $\begin{cases} z_L \overset{IID}{\hookrightarrow} .3\mathcal{U}_{[0,3]} + .7\mathcal{U}_{[1,2]} \\ z_S \overset{IID}{\hookrightarrow} \mathcal{U}_{[.5,2.5]} \\ \mathbb{C}\mathrm{or}(z_S, z_L) = .7 \end{cases}$	
False	$\begin{bmatrix} r_L \\ r_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \begin{bmatrix} z_L \\ z_S \end{bmatrix}$ where $\begin{cases} z_L \overset{IID}{\hookrightarrow} \mathrm{t}(4) \\ z_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\mathrm{or}(z_S, z_L) = .7 \end{cases}$	

Note: The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of C_T^* is approximated through block subsampling for "C_T^* No centering," and centered block subsampling for "C_T^*." The block size is $b_T = \sqrt{T}$.The tops and bottoms of each "box" are the 25th and 75th percentiles of the p-values, respectively. The line in the middle of each box is the median. Crosses beyond the whisker length indicate outliers. An outlier is a value that is more than 1.5 times the interquartile range away from the corresponding end of the interquartile ranges. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length.

C Additional empirical evidence

Table A.10: Acronym and Description of the 205 Characteristics

This Table provides a short description of each of the 205 characteristics used.

	Description
AM	Total assets to market
AOP	Analyst Optimism
AbnormalAccruals	Abnormal Accruals
Accruals	Accruals
AccrualsBM	Book-to-market and accruals
Activism1	Takeover vulnerability
Activism2	Active shareholders
AdExp	Advertising Expense
AgeIPO	IPO and age
AnalystRevision	EPS forecast revision
AnalystValue	Analyst Value
AnnouncementReturn	Earnings announcement return
AssetGrowth	Asset growth
BM	Book to market using most recent ME
BMdec	Book to market using December ME
BPEBM	Leverage component of BM
Beta	CAPM beta
BetaFP	Frazzini-Pedersen Beta
BetaLiquidityPS	Pastor-Stambaugh liquidity beta
BetaTailRisk	Tail risk beta
BidAskSpread	Bid-ask spread
BookLeverage	Book leverage (annual)
BrandInvest	Brand capital investment
CBOperProf	Cash-based operating profitability
CF	Cash flow to market
Cash	Cash to assets
CashProd	Cash Productivity
ChAssetTurnover	Change in Asset Turnover
ChEQ	Growth in book equity
ChForecastAccrual	Change in Forecast and Accrual
ChInv	Inventory Growth
ChInvIA	Change in capital inv (ind adj)
ChNAnalyst	Decline in Analyst Coverage
ChNNCOA	Change in Net Noncurrent Op Assets
ChNWC	Change in Net Working Capital
ChTax	Change in Taxes
ChangeInRecommendation	Change in recommendation
CitationsRD	Citations to RD expenses
CompEquIss	Composite equity issuance
CompositeDebtIssuance	Composite debt issuance
ConsRecomm	Consensus Recommendation
ConvDebt	Convertible debt indicator
CoskewACX	Coskewness using daily returns
Coskewness	Coskewness
CredRatDG	Credit Rating Downgrade
CustomerMomentum	Customer momentum
DebtIssuance	Debt Issuance
DelBreadth	Breadth of ownership
DelCOA	Change in current operating assets
DelCOL	Change in current operating liabilities

Table A.10 (continued)

	Description
DelDRC	Deferred Revenue
DelEqu	Change in equity to assets
DelFINL	Change in financial liabilities
DelLTI	Change in long-term investment
DelNetFin	Change in net financial assets
DivInit	Dividend Initiation
DivOmit	Dividend Omission
DivSeason	Dividend seasonality
DivYieldST	Predicted div yield next month
DolVol	Past trading volume
DownRecomm	Down forecast EPS
EBM	Enterprise component of BM
EP	Earnings-to-Price Ratio
EarnSupBig	Earnings surprise of big firms
EarningsConsistency	Earnings consistency
EarningsForecastDisparity	Long-vs-short EPS forecasts
EarningsStreak	Earnings surprise streak
EarningsSurprise	Earnings Surprise
EntMult	Enterprise Multiple
EquityDuration	Equity Duration
ExchSwitch	Exchange Switch
ExclExp	Excluded Expenses
FEPS	Analyst earnings per share
FR	Pension Funding Status
FirmAge	Firm age based on CRSP
FirmAgeMom	Firm Age - Momentum
ForecastDispersion	EPS Forecast Dispersion
Frontier	Efficient frontier index
GP	gross profits / total assets
Governance	Governance Index
GrAdExp	Growth in advertising expenses
GrLTNOA	Growth in long term operating assets
GrSaleToGrInv	Sales growth over inventory growth
GrSaleToGrOverhead	Sales growth over overhead growth
Herf	Industry concentration (sales)
HerfAsset	Industry concentration (assets)
HerfBE	Industry concentration (equity)
High52	52 week high
IO_ShortInterest	Inst own among high short interest
IdioRisk	Idiosyncratic risk
IdioVol3F	Idiosyncratic risk (3 factor)
IdioVolAHT	Idiosyncratic risk (AHT)
Illiquidity	Amihud's illiquidity
IndIPO	Initial Public Offerings
IndMom	Industry Momentum
IndRetBig	Industry return of big firms
IntMom	Intermediate Momentum
IntanBM	Intangible return using BM
IntanCFP	Intangible return using CFtoP
IntanEP	Intangible return using EP
IntanSP	Intangible return using Sale2P
InvGrowth	Inventory Growth

Table A.10 (continued)

	Description
InvestPPEInv	change in ppe and inv/assets
Investment	Investment to revenue
LRreversal	Long-run reversal
Leverage	Market leverage
MRreversal	Medium-run reversal
MS	Mohanram G-score
MaxRet	Maximum return over month
MeanRankRevGrowth	Revenue Growth Rank
Mom12m	Momentum (12 month)
Mom12mOffSeason	Momentum without the seasonal part
Mom6m	Momentum (6 month)
Mom6mJunk	Junk Stock Momentum
MomOffSeason	Off season long-term reversal
MomOffSeason06YrPlus	Off season reversal years 6 to 10
MomOffSeason11YrPlus	Off season reversal years 11 to 15
MomOffSeason16YrPlus	Off season reversal years 16 to 20
MomRev	Momentum and LT Reversal
MomSeason	Return seasonality years 2 to 5
MomSeason06YrPlus	Return seasonality years 6 to 10
MomSeason11YrPlus	Return seasonality years 11 to 15
MomSeason16YrPlus	Return seasonality years 16 to 20
MomSeasonShort	Return seasonality last year
MomVol	Momentum in high volume stocks
NOA	Net Operating Assets
NetDebtFinance	Net debt financing
NetDebtPrice	Net debt to price
NetEquityFinance	Net equity financing
NetPayoutYield	Net Payout Yield
NumEarnIncrease	Earnings streak length
OPLeverage	Operating leverage
OScore	O Score
OperProf	operating profits / book equity
OperProfRD	Operating profitability R&D adjusted
OptionVolume1	Option to stock volume
OptionVolume2	Option volume to average
OrderBacklog	Order backlog
OrderBacklogChg	Change in order backlog
OrgCap	Organizational capital
PS	Piotroski F-score
PatentsRD	Patents to R&D expenses
PayoutYield	Payout Yield
PctAcc	Percent Operating Accruals
PctTotAcc	Percent Total Accruals
PredictedFE	Predicted Analyst forecast error
Price	Price
PriceDelayRsq	Price delay r square
PriceDelaySlope	Price delay coeff
PriceDelayTstat	Price delay SE adjusted
ProbInformedTrading	Probability of Informed Trading
RD	R&D over market cap
RDAbility	R&D ability
RDIPO	IPO and no R&D spending
RDS	Real dirty surplus

Table A.10 (continued)

	Description
RDcap	R&D capital-to-assets
REV6	Earnings forecast revisions
RIO_Disp	Inst Own and Forecast Dispersion
RIO_MB	Inst Own and Market to Book
RIO_Turnover	Inst Own and Turnover
RIO_Volatility	Inst Own and Idio Vol
ResidualMomentum	Momentum based on FF3 residuals
ReturnSkew	Return skewness
ReturnSkew3F	Idiosyncratic skewness (3F model)
RevenueSurprise	Revenue Surprise
RoE	net income / book equity
SP	Sales-to-price
STreversal	Short term reversal
ShareIss1Y	Share issuance (1 year)
ShareIss5Y	Share issuance (5 year)
ShareRepurchase	Share repurchases
ShareVol	Share Volume
ShortInterest	Short Interest
Size	Size
SmileSlope	Put volatility minus call volatility
Spinoff	Spinoffs
SurpriseRD	Unexpected R&D increase
Tax	Taxable income to income
TotalAccruals	Total accruals
UpRecomm	Up Forecast
VarCF	Cash-flow to price variance
VolMkt	Volume to market equity
VolSD	Volume Variance
VolumeTrend	Volume Trend
XFIN	Net external financing
betaVIX	Systematic volatility
cfp	Operating Cash flows to price
dNoa	change in net operating assets
fgr5yrLag	Long-term EPS forecast
grcapx	Change in capex (two years)
grcapx3y	Change in capex (three years)
hire	Employment growth
iomom_cust	Customers momentum
iomom_supp	Suppliers momentum
realestate	Real estate holdings
retConglomerate	Conglomerate return
roaq	Return on assets (qtrly)
sfe	Earnings Forecast to price
sinAlgo	Sin Stock (selection criteria)
skew1	Volatility smirk near the money
std_turn	Share turnover volatility
tang	Tangibility
zerotrade	Days with zero trades
zerotradeAlt1	Days with zero trades
zerotradeAlt12	Days with zero trades

Table A.11: Unconditional and Conditional Tests on the Market for 205 Characteristic Sorted Portfolios

This table presents results for the unconditional and the conditional tests applied to 205 characteristics. For each characteristic, stocks are sorted into deciles, quintiles or median portfolios. We retain the portfolios in the lowest and the highest of these sorting. For the return spread between the Low and High legs we report the Newey-West t statistics with an optimal choice of lags. For each test are reported the p-value for the null corresponding to the portfolio with the highest mean returns dominates the portfolio with the lowest mean return. For each characteristic we retain three samples: the original one, the post publication one and the full sample (ending in December 2020).

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.
AM	0.80	1.43	2.93	1.00	0.49	0.93	1.28	1.04	0.42	0.06	0.87	1.35	2.36	1.00	0.09
AOP	1.10	1.46	1.61	1.00	0.12	0.97	1.01	0.24	1.00	0.01	1.02	1.18	1.28	1.00	0.00
AbnormalAccruals	0.88	1.43	4.21	0.55	0.64	1.04	0.93	0.77	1.00	0.13	0.97	1.14	1.79	0.03	0.04
Accruals	0.83	1.40	3.86	1.00	0.10	0.74	1.01	3.10	1.00	0.85	0.79	1.21	4.77	1.00	0.15
AccrualsBM	0.63	2.07	3.64	1.00	0.19	0.94	2.07	2.80	0.60	0.18	0.80	2.07	4.33	1.00	0.24
Activism1	1.39	1.63	1.15	0.13	0.25	0.94	0.86	0.35	1.00	0.04	1.25	1.39	0.91	0.05	0.08
Activism2	1.36	1.79	0.91	0.56	0.41	0.37	1.30	1.85	0.38	0.18	1.05	1.63	1.63	0.48	0.51
AdExp	1.35	2.00	2.49	1.00	0.24	0.90	1.27	1.39	0.45	0.09	1.11	1.62	2.67	0.51	0.30
AgeIPO	-0.96	0.45	1.98	1.00	0.56	0.37	1.04	2.26	1.00	0.42	0.23	0.98	2.69	1.00	0.44
AnalystRevision	1.28	2.20	2.71	0.50	0.50	0.69	1.32	5.50	1.00	0.92	0.75	1.42	5.99	1.00	0.96
AnalystValue	1.08	1.35	1.33	0.37	0.56	0.87	0.99	0.36	0.46	0.07	0.95	1.13	0.80	0.50	0.06
AnnouncementReturn	0.86	2.06	5.51	0.19	0.74	0.61	1.70	6.08	1.00	0.86	0.70	1.83	7.91	1.00	0.86
AssetGrowth	0.38	1.89	5.27	1.00	0.18	0.57	0.85	1.08	0.01	0.00	0.45	1.56	5.05	1.00	0.12
BM	0.78	2.38	3.08	1.00	0.26	0.72	1.70	3.27	1.00	0.32	0.74	1.87	4.34	1.00	0.32
BMdec	0.69	1.66	4.21	1.00	0.49	1.02	1.52	2.33	0.38	0.19	0.86	1.59	4.52	1.00	0.33
BPEBM	1.13	1.36	2.40	0.33	0.67	0.89	0.94	0.49	0.00	0.00	1.05	1.22	2.29	0.05	0.08
Beta	1.10	1.77	1.70	0.00	0.00	0.91	0.97	0.18	0.00	0.00	0.99	1.31	1.35	0.00	0.00
BetaFP	1.15	1.18	0.08	0.00	0.00	0.68	0.56	0.16	1.00	0.00	1.11	1.12	0.05	0.00	0.00
BetaLiquidityPS	1.05	1.40	1.78	1.00	0.39	0.32	0.61	1.39	0.23	0.49	0.77	1.10	2.25	0.61	0.45
BetaTailRisk	0.92	1.38	2.82	0.00	0.00	1.07	0.99	0.26	1.00	0.00	0.95	1.31	2.48	0.00	0.00
BidAskSpread	0.98	1.69	1.55	0.00	0.00	0.97	0.93	0.11	1.00	0.06	0.98	1.19	0.77	0.00	0.00
BookLeverage	0.95	1.23	2.72	0.56	0.51	1.11	1.26	0.55	0.06	0.07	1.03	1.25	1.38	0.09	0.09
BrandInvest	1.29	1.85	1.82	0.05	0.05	1.10	1.09	0.03	1.00	0.34	1.25	1.68	1.72	0.04	0.01

Table A.11 (continued)

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**	**Low**	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**	**Low**	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**
CBOperProf	0.59	1.05	2.70	1.00	0.07	0.88	1.56	1.28	1.00	0.95	0.62	1.11	2.97	1.00	0.09
CF	0.52	1.35	3.34	1.00	0.90	1.09	1.24	0.52	0.38	0.27	0.84	1.29	2.26	0.37	0.25
Cash	0.83	1.53	2.36	0.05	0.06	0.92	1.33	1.04	0.02	0.00	0.85	1.49	2.57	0.04	0.03
CashProd	0.88	1.44	2.82	1.00	0.22	0.87	0.70	0.73	1.00	0.49	0.88	1.22	2.15	1.00	0.17
ChAssetTurnover	0.86	1.16	3.47	1.00	0.87	1.11	1.10	0.06	1.00	0.27	0.98	1.13	2.61	1.00	0.94
ChEQ	0.95	1.51	3.51	1.00	0.08	0.79	1.04	1.23	0.03	0.04	0.91	1.40	3.64	0.64	0.03
ChForecastAccrual	1.03	1.39	3.26	0.54	0.99	0.86	0.98	1.62	0.54	0.50	0.93	1.14	3.43	0.40	0.61
ChInv	0.87	1.64	4.60	1.00	0.38	0.93	1.36	2.36	0.72	0.21	0.90	1.52	5.01	1.00	0.46
ChInvIA	1.44	1.94	4.28	1.00	0.52	0.96	1.30	2.81	0.26	0.17	1.11	1.50	4.33	0.31	0.29
ChNAnalyst	0.14	0.55	0.65	0.28	0.06	-2.65	-0.67	0.96	1.00	0.23	-0.50	0.27	1.17	1.00	0.09
ChNNCOA	0.74	1.09	3.54	1.00	0.81	1.15	1.19	0.57	0.34	0.08	0.94	1.14	3.20	1.00	0.62
ChNWC	0.86	1.02	2.49	0.28	0.79	1.01	0.97	0.59	0.50	0.35	0.93	1.00	1.40	0.33	0.72
ChTax	0.85	1.94	5.71	0.42	0.85	0.75	1.06	1.85	0.60	0.66	0.81	1.66	5.99	0.58	0.94
ChangeInRecommendation	0.78	1.82	3.48	0.29	0.65	0.70	1.16	4.36	1.00	0.97	0.71	1.28	5.04	1.00	0.95
CitationsRD	1.17	1.19	0.04	0.63	0.02	1.67	3.18	0.62	1.00	0.00	1.21	1.36	0.27	0.62	0.01
CompEquIss	0.97	1.23	2.15	1.00	0.84	0.67	1.11	2.72	0.20	0.61	0.87	1.19	3.22	1.00	0.98
CompositeDebtIssuance	1.24	1.55	4.10	1.00	0.28	0.79	1.00	2.19	0.37	0.39	1.10	1.39	4.64	1.00	0.27
ConsRecomm	1.35	1.89	1.31	1.00	0.89	0.31	0.78	1.70	1.00	0.66	0.48	0.95	1.90	1.00	0.61
ConvDebt	0.76	1.14	3.46	1.00	0.09	0.83	1.14	1.75	1.00	0.33	0.77	1.14	3.83	1.00	0.09
CoskewACX	1.09	1.38	2.58	0.35	0.26	0.87	1.40	2.28	0.27	0.69	1.01	1.39	3.44	0.36	0.57
Coskewness	0.87	1.14	1.88	0.09	0.14	0.76	0.96	1.70	0.36	0.26	0.82	1.05	2.57	0.08	0.16
CredRatDG	0.38	1.11	2.38	1.00	0.79	0.41	1.07	1.83	1.00	0.19	0.40	1.08	2.74	1.00	0.31
CustomerMomentum	0.30	1.46	2.83	0.24	0.49	1.20	1.01	0.41	0.03	0.21	0.65	1.28	2.05	0.27	0.69
DebtIssuance	1.78	1.95	2.46	1.00	0.44	0.98	1.35	3.77	1.00	0.86	1.24	1.54	4.34	1.00	0.86
DelBreadth	0.96	1.65	3.39	0.56	0.88	0.59	1.05	1.44	1.00	0.84	0.77	1.33	2.89	0.59	0.95
DelCOA	0.95	1.49	4.63	1.00	0.16	0.97	1.14	1.19	0.50	0.09	0.96	1.37	4.48	1.00	0.37
DelCOL	1.08	1.43	3.79	1.00	0.06	0.94	1.06	0.86	0.70	0.08	1.03	1.31	3.56	1.00	0.08
DelDRC	0.59	1.30	1.56	0.18	0.68	1.08	1.18	0.61	1.00	0.41	0.93	1.22	1.64	0.33	0.61
DelEqu	1.03	1.49	2.91	1.00	0.05	0.84	1.23	1.63	0.04	0.00	0.97	1.41	3.22	0.73	0.01
DelFINL	0.84	1.57	7.03	1.00	0.97	0.83	1.10	2.78	1.00	0.76	0.84	1.42	7.15	1.00	0.96
DelLTI	1.17	1.34	2.34	0.22	0.12	0.97	1.10	1.67	0.17	0.08	1.11	1.26	2.82	0.19	0.07

Table A.11 (continued)

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.
DelNetFin	0.94	1.49	6.28	1.00	0.97	0.99	1.03	0.32	1.00	0.18	0.96	1.34	5.28	1.00	0.77
DivInit	1.26	1.84	4.13	1.00	0.80	1.11	1.31	1.24	0.22	0.09	1.18	1.54	3.35	0.28	0.11
DivOmit	0.76	1.28	2.01	0.69	0.00	0.45	1.11	1.92	1.00	0.99	0.59	1.18	2.67	1.00	0.44
DivSeason	1.02	1.35	8.08	0.47	0.81	1.10	1.17	1.37	0.55	0.51	1.02	1.33	8.19	0.50	0.85
DivYieldST	1.00	1.42	3.34	1.00	0.13	1.14	1.75	5.81	0.23	0.25	1.07	1.59	6.07	0.24	0.14
DolVol	0.93	1.69	2.75	0.49	0.00	0.84	1.29	2.21	1.00	0.00	0.89	1.50	3.50	1.00	0.00
DownRecomm	1.07	1.70	2.74	0.26	0.42	0.69	1.00	4.20	1.00	0.76	0.75	1.11	4.68	1.00	0.88
EBM	1.06	1.36	3.24	1.00	0.51	0.89	0.93	0.31	0.08	0.04	1.00	1.22	2.92	0.48	0.46
EP	0.99	1.38	2.17	1.00	0.38	1.03	1.26	1.72	0.34	0.28	1.02	1.29	2.39	0.43	0.28
EarnSupBig	1.10	1.47	2.07	0.32	0.12	0.87	1.01	0.76	0.61	0.70	1.01	1.30	2.17	0.53	0.31
EarningsConsistency	1.04	1.25	2.28	0.71	0.92	1.00	1.24	1.40	1.00	0.23	1.03	1.25	2.59	1.00	0.75
EarningsForecastDisparity	0.68	1.33	3.37	0.52	0.61	0.58	0.80	1.01	0.34	0.85	0.64	1.14	3.41	0.37	0.92
EarningsStreak	0.46	1.55	5.51	1.00	0.84	0.81	1.21	3.33	1.00	0.83	0.58	1.44	6.22	1.00	0.99
EarningsSurprise	1.20	2.35	3.58	0.47	0.65	0.89	1.34	4.03	0.47	0.95	0.95	1.51	5.14	0.47	0.98
EntMult	0.85	1.70	4.23	1.00	0.17	1.16	1.06	0.33	1.00	0.75	0.91	1.58	3.80	1.00	0.20
EquityDuration	0.81	1.37	2.73	1.00	0.82	0.63	0.80	0.49	0.63	0.01	0.74	1.15	2.14	1.00	0.23
ExchSwitch	0.71	1.16	2.55	1.00	0.16	0.42	1.21	3.96	1.00	0.67	0.56	1.18	4.62	1.00	0.54
ExclExp	1.45	1.72	2.58	1.00	0.92	1.00	1.17	1.25	1.00	0.00	1.17	1.37	2.19	1.00	0.13
FEPS	0.01	1.47	2.51	1.00	0.12	0.67	0.95	0.85	1.00	0.04	0.32	1.23	2.58	1.00	0.06
FR	1.06	1.37	1.62	1.00	0.40	1.51	1.00	1.49	0.00	0.00	1.27	1.20	0.34	0.00	0.00
FirmAge	1.39	1.39	0.06	0.49	0.19	1.12	1.04	0.64	1.00	0.00	1.27	1.23	0.52	1.00	0.02
FirmAgeMom	-0.70	1.59	4.05	1.00	0.75	0.02	1.26	3.37	1.00	0.75	-0.34	1.43	5.09	1.00	0.79
ForecastDispersion	0.88	1.53	2.38	1.00	0.34	0.70	0.95	0.68	1.00	0.06	0.80	1.27	2.17	1.00	0.11
Frontier	0.61	2.70	4.67	1.00	0.18	0.87	1.68	2.01	0.34	0.08	0.71	2.28	4.93	1.00	0.15
GP	0.78	1.08	2.14	0.75	0.31	0.82	1.38	1.69	1.00	0.87	0.79	1.13	2.66	0.71	0.40
Governance	1.30	1.82	1.64	0.44	0.77	1.12	0.16	2.26	1.00	0.12	1.22	1.09	0.47	1.00	0.18
GrAdExp	0.96	1.40	3.32	1.00	0.20	1.20	1.22	0.10	0.51	0.53	1.01	1.36	3.08	1.00	0.15
GrLTNOA	0.92	1.29	2.98	1.00	0.13	0.77	0.85	0.76	0.33	0.48	0.85	1.08	2.81	1.00	0.35
GrSaleToGrInv	1.41	1.72	3.08	0.42	0.70	0.97	1.14	1.98	0.58	0.88	1.11	1.33	3.23	0.48	0.97
GrSaleToGrOverhead	1.54	1.48	0.38	0.01	0.00	1.11	1.02	1.04	0.72	0.17	1.25	1.16	1.05	0.27	0.04
Herf	1.25	1.46	1.84	0.37	0.68	1.03	1.06	0.16	0.15	0.03	1.18	1.33	1.53	0.41	0.45

Table A.11 (continued)

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.
Herf	1.32	1.51	1.36	0.11	0.32	1.08	0.99	0.50	0.39	0.64	1.24	1.34	0.86	0.20	0.21
HerfAsset	1.30	1.52	1.63	0.26	0.46	1.05	1.02	0.23	0.30	0.49	1.22	1.36	1.26	0.42	0.35
HerfBE	0.94	1.45	2.12	1.00	0.18	0.88	0.82	0.14	0.00	0.00	0.92	1.24	1.49	1.00	0.09
High52	-1.53	0.69	2.80	1.00	0.96	-2.67	1.06	3.24	1.00	0.35	-2.03	0.85	4.25	1.00	0.80
IO_ShortInterest	0.06	1.05	2.89	1.00	0.21	0.57	0.73	0.34	1.00	0.01	0.24	0.94	2.51	1.00	0.04
IdioRisk	0.10	1.06	2.75	1.00	0.21	0.59	0.70	0.23	1.00	0.03	0.27	0.94	2.33	1.00	0.04
IdioVol3F	0.44	1.34	2.06	1.00	0.39	0.66	0.70	0.05	1.00	0.01	0.56	1.01	1.22	1.00	0.02
IdioVolAHT	1.02	1.59	3.00	0.22	0.23	0.78	0.82	0.22	1.00	0.00	0.92	1.28	2.63	0.61	0.05
Illiquidity	1.04	1.70	1.99	1.00	0.62	0.88	1.15	1.50	1.00	0.07	0.92	1.30	2.37	1.00	0.09
IndIPO	1.14	1.42	1.81	0.34	0.66	0.72	1.24	2.00	0.56	0.68	0.96	1.34	2.66	0.47	0.72
IndMom	0.12	2.33	5.54	0.16	0.88	0.41	1.47	3.62	1.00	0.84	0.23	2.00	6.50	0.25	0.96
IndRetBig	0.25	1.49	5.06	1.00	0.97	0.69	1.02	0.67	1.00	0.55	0.30	1.44	5.08	1.00	0.99
IntMom	1.03	1.42	2.13	1.00	0.29	0.92	0.90	0.08	1.00	0.74	0.99	1.25	1.75	0.49	0.21
IntanBM	1.08	1.48	2.14	1.00	0.20	0.83	1.03	0.81	0.07	0.19	1.00	1.34	2.23	0.49	0.22
IntanCFP	1.07	1.41	2.20	1.00	0.11	0.84	0.93	0.44	0.17	0.33	1.00	1.26	2.08	0.57	0.11
IntanEP	1.10	1.62	2.30	0.15	0.04	0.93	1.00	0.20	0.00	0.00	1.04	1.42	1.94	0.00	0.00
IntanSP	0.73	1.60	5.20	1.00	0.81	0.95	0.96	0.03	0.30	0.20	0.78	1.48	4.85	1.00	0.66
InvGrowth	0.86	1.66	5.66	1.00	0.37	0.76	0.94	1.34	1.00	0.64	0.83	1.45	5.76	1.00	0.44
InvestPPEInv	1.00	1.26	2.05	0.22	0.29	0.91	1.03	0.54	0.10	0.07	0.96	1.14	1.51	0.05	0.06
Investment	0.99	1.78	2.88	0.14	0.10	0.93	1.39	1.55	0.00	0.00	0.97	1.62	3.20	0.00	0.00
LRreversal	1.16	1.52	2.48	0.69	0.37	0.86	1.15	1.06	1.00	0.05	0.99	1.31	1.88	1.00	0.10
Leverage	1.42	1.82	2.10	0.46	0.33	0.97	1.25	1.65	0.04	0.00	1.22	1.56	2.67	0.08	0.03
MRreversal	0.14	1.48	4.28	1.00	0.63	0.63	1.08	2.10	1.00	0.67	0.36	1.30	4.75	1.00	0.60
MS	-0.05	0.84	2.50	1.00	0.08	0.66	0.72	0.13	1.00	0.01	0.13	0.81	2.29	1.00	0.02
MaxRet	0.82	1.37	3.41	0.19	0.27	1.12	1.11	0.05	0.58	0.18	0.99	1.23	2.50	0.42	0.24
MeanRankRevGrowth	0.50	1.87	4.24	0.48	0.85	0.90	1.39	1.11	1.00	0.30	0.72	1.61	3.13	0.50	0.51
Mom12m	0.51	1.74	4.14	0.44	0.84	0.80	1.41	1.02	1.00	0.27	0.60	1.63	3.62	1.00	0.43
Mom12mOffSeason	0.53	1.57	3.49	0.51	0.86	0.83	1.45	1.69	1.00	0.44	0.69	1.50	3.32	1.00	0.51
Mom6m	0.40	1.98	3.28	1.00	0.65	0.59	0.88	0.70	1.00	0.74	0.48	1.53	3.22	1.00	0.70
Mom6mJunk	0.45	1.76	4.41	0.45	0.19	1.05	1.15	0.24	0.00	0.00	0.65	1.56	3.75	0.53	0.11
MomOffSeason	0.88	1.46	3.82	0.46	0.65	0.65	1.51	3.44	0.30	0.78	0.80	1.48	4.90	0.39	0.65

Table A.11 (continued)

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.	Low	High	t_{NW}^{Spread}	Uncond.	Cond.
MomOffSeason11YrPlus	1.14	1.38	2.00	0.79	0.80	1.19	1.32	0.67	0.34	0.55	1.16	1.36	1.99	0.73	0.73
MomOffSeason16YrPlus	1.03	1.38	2.38	0.48	0.30	1.03	1.35	1.81	1.00	0.53	1.03	1.37	2.95	0.55	0.30
MomRev	0.47	1.67	4.12	1.00	0.52	0.96	1.20	0.62	1.00	0.64	0.64	1.51	3.79	1.00	0.47
MomSeason	0.78	1.60	4.59	1.00	0.76	0.85	1.32	1.89	0.44	0.67	0.80	1.51	4.76	0.44	0.91
MomSeason06YrPlus	0.86	1.60	4.98	1.00	1.00	1.05	1.26	0.99	0.52	0.23	0.92	1.49	4.57	0.47	0.98
MomSeason11YrPlus	0.88	1.63	5.67	1.00	0.98	1.00	1.29	1.60	0.56	0.80	0.92	1.52	5.59	1.00	0.99
MomSeason16YrPlus	0.91	1.50	4.31	1.00	0.98	0.87	1.34	2.57	0.34	0.81	0.90	1.45	4.86	1.00	1.00
MomSeasonShort	0.40	1.76	6.10	1.00	0.97	1.19	1.06	0.52	0.09	0.10	0.66	1.54	4.95	1.00	0.97
MomVol	-0.41	1.18	4.04	0.45	0.88	-0.01	1.11	1.99	1.00	0.64	-0.23	1.15	4.11	1.00	0.59
NOA	0.43	1.51	5.01	1.00	0.81	0.79	1.20	1.40	0.28	0.02	0.55	1.41	4.78	1.00	0.54
NetDebtFinance	0.62	1.37	5.46	1.00	0.76	0.82	1.32	3.37	1.00	0.92	0.70	1.35	6.30	1.00	0.88
NetDebtPrice	1.31	1.86	2.82	0.50	0.47	1.08	1.65	1.60	1.00	0.89	1.24	1.79	3.15	1.00	0.89
NetEquityFinance	0.61	1.67	3.96	1.00	0.51	0.65	1.32	2.04	1.00	0.08	0.63	1.53	4.40	1.00	0.21
NetPayoutYield	0.76	1.63	2.19	1.00	0.13	0.35	1.15	2.23	1.00	0.27	0.57	1.41	3.06	1.00	0.12
NumEarnIncrease	0.76	1.27	4.53	1.00	0.89	1.06	1.24	1.63	1.00	0.54	0.86	1.26	4.78	1.00	0.86
OPLeverage	0.96	1.31	2.07	0.00	0.01	0.94	1.66	1.99	0.66	0.27	0.95	1.38	2.73	0.00	0.00
OScore	0.24	1.25	2.46	1.00	0.80	0.34	1.08	2.07	1.00	0.11	0.30	1.14	3.06	1.00	0.20
OperProf	0.67	1.39	2.40	1.00	0.18	0.78	1.12	1.90	1.00	0.16	0.71	1.28	2.90	1.00	0.16
OperProfRD	0.66	0.99	1.57	1.00	0.06	0.70	1.53	1.39	1.00	0.40	0.66	1.05	1.89	1.00	0.12
OptionVolume1	0.53	1.21	1.85	1.00	0.16	0.62	0.98	2.00	1.00	0.18	0.57	1.12	2.34	1.00	0.17
OptionVolume2	0.71	1.24	1.93	0.30	0.37	0.78	0.86	0.87	1.00	0.16	0.74	1.09	2.11	0.25	0.29
OrderBacklog	0.96	1.46	2.74	1.00	0.33	1.32	1.14	1.08	0.50	0.46	1.15	1.29	1.12	0.43	0.53
OrderBacklogChg	1.13	1.51	2.50	0.65	0.89	1.05	1.36	1.32	0.60	0.65	1.09	1.44	2.62	0.67	0.96
OrgCap	0.80	1.17	2.70	1.00	0.40	1.26	1.43	1.17	1.00	0.39	0.91	1.23	2.94	1.00	0.43
PS	1.32	2.23	2.84	1.00	0.60	0.12	1.03	1.76	1.00	0.52	0.68	1.59	2.90	1.00	0.53
PatentsRD	1.22	1.38	0.29	0.61	0.01	NaN	NaN	NaN	NaN	NaN	1.22	1.38	0.29	0.61	0.01
PayoutYield	1.04	1.47	2.42	1.00	0.08	0.93	0.93	0.00	0.51	0.44	0.99	1.22	1.70	0.56	0.25
PctAcc	0.41	0.87	3.05	0.24	0.42	1.15	1.24	0.79	0.14	0.19	0.69	1.01	3.09	0.46	0.65
PctTotAcc	0.59	1.09	4.01	1.00	0.75	1.41	1.48	0.71	0.27	0.77	0.90	1.23	3.81	0.49	0.95
PredictedFE	1.06	1.36	0.86	1.00	0.09	1.11	0.98	0.56	0.03	0.07	1.09	1.09	0.03	0.00	0.02
Price	1.09	2.51	2.57	0.00	0.00	1.06	1.41	1.19	0.00	0.00	1.07	1.91	2.80	0.00	0.00

Table A.11 (continued)

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**	**Low**	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**	**Low**	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**
PriceDelayRsq	1.07	1.55	2.81	1.00	0.00	0.73	1.04	1.22	1.00	0.02	0.95	1.38	3.03	1.00	0.00
PriceDelaySlope	1.31	1.48	2.14	1.00	0.00	0.80	0.99	1.21	1.00	0.02	1.14	1.32	2.44	1.00	0.00
PriceDelayTstat	1.21	1.36	1.66	1.00	0.00	0.83	0.85	0.14	1.00	0.01	1.08	1.19	1.48	1.00	0.00
ProbInformedTrading	0.29	1.59	3.96	1.00	0.30	-0.07	1.41	1.58	1.00	0.00	0.21	1.55	4.06	1.00	0.21
RD	1.55	2.56	3.89	0.49	0.82	1.00	2.09	2.22	0.00	0.00	1.25	2.30	3.58	0.02	0.03
RDAbility	1.18	1.45	1.43	0.70	0.58	1.40	1.28	0.62	1.00	0.37	1.24	1.40	1.10	0.56	0.62
RDIPO	0.32	1.29	2.47	1.00	0.79	0.57	1.08	2.31	1.00	0.75	0.47	1.16	3.35	1.00	0.80
RDS	1.21	1.70	3.41	0.43	0.09	0.92	0.88	0.33	1.00	0.21	1.10	1.39	2.88	0.32	0.06
RDcap	1.10	1.56	1.75	0.02	0.06	0.75	1.22	1.44	0.07	0.06	0.99	1.45	2.23	0.03	0.04
REV6	0.66	1.95	3.97	1.00	0.95	0.51	1.10	1.91	1.00	0.49	0.56	1.41	3.65	1.00	0.55
RIO_Disp	0.68	1.31	2.27	1.00	0.54	0.58	0.83	1.03	0.42	0.65	0.64	1.11	2.48	1.00	0.74
RIO_MB	0.58	1.47	3.04	0.05	0.10	0.88	1.04	0.80	0.00	0.00	0.70	1.29	3.10	0.00	0.01
RIO_Turnover	1.00	1.65	2.06	0.68	0.48	0.61	0.91	1.16	0.32	0.71	0.84	1.34	2.38	0.52	0.70
RIO_Volatility	-0.01	1.00	3.31	1.00	0.99	0.57	1.14	1.75	1.00	0.89	0.23	1.06	3.73	1.00	0.98
ResidualMomentum	0.71	1.66	6.85	1.00	0.63	0.93	1.08	0.65	1.00	0.28	0.73	1.59	6.84	1.00	0.59
ReturnSkew	0.87	1.28	4.02	1.00	0.12	0.83	0.93	0.50	0.48	0.74	0.86	1.23	3.91	1.00	0.14
ReturnSkew3F	0.93	1.22	3.73	1.00	0.19	0.93	0.90	0.19	0.02	0.00	0.93	1.18	3.49	1.00	0.14
RevenueSurprise	1.02	1.77	4.43	0.13	0.57	0.80	1.17	2.51	1.00	0.37	0.91	1.47	4.79	0.30	0.76
RoE	1.14	1.46	2.16	1.00	0.48	0.72	1.05	1.56	1.00	0.08	0.87	1.20	2.23	1.00	0.11
SP	0.89	1.60	1.98	1.00	0.40	0.75	1.50	2.35	0.43	0.32	0.79	1.53	2.99	0.42	0.41
STreversal	-0.03	2.91	7.25	1.00	0.85	0.40	2.04	4.31	0.02	0.13	0.14	2.58	8.37	0.18	0.37
ShareIss1Y	0.89	1.51	4.12	1.00	0.11	0.59	1.03	2.20	1.00	0.12	0.79	1.35	4.64	1.00	0.07
ShareIss5Y	0.99	1.51	4.03	1.00	0.12	0.77	1.02	1.92	0.40	0.75	0.92	1.35	4.30	0.36	0.16
ShareRepurchase	0.92	1.24	2.90	1.00	0.78	1.19	1.29	0.86	0.35	0.09	1.12	1.27	1.77	1.00	0.13
ShareVol	0.34	1.25	3.58	1.00	0.10	0.80	1.07	1.39	1.00	0.09	0.57	1.16	3.66	1.00	0.06
ShortInterest	0.99	1.82	4.44	1.00	0.24	0.57	1.39	4.37	1.00	0.05	0.74	1.56	6.03	1.00	0.04
Size	0.99	1.49	2.34	0.00	0.00	1.11	1.29	1.48	0.25	0.00	1.05	1.39	2.79	0.00	0.00
SmileSlope	0.06	1.84	4.15	0.25	0.62	0.15	1.03	4.57	1.00	0.90	0.11	1.36	5.60	1.00	0.95
Spinoff	0.87	1.28	2.05	0.00	0.01	0.96	1.12	0.90	0.05	0.04	0.92	1.19	1.99	0.03	0.02
SurpriseRD	1.55	1.84	2.38	0.03	0.03	1.18	1.27	0.76	0.02	0.02	1.40	1.61	2.41	0.03	0.04
Tax	0.96	1.41	2.93	0.41	0.48	0.68	1.09	3.45	1.00	0.99	0.84	1.27	4.18	0.37	0.86

Table A.11 (continued)

	Original Sample					Post Publication					Full Sample				
	Returns			p-values		Returns			p-values		Returns			p-values	
	Low	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**	**Low**	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**	**Low**	**High**	t_{NW}^{Spread}	**Uncond.**	**Cond.**
TotalAccruals	1.07	1.35	2.28	0.18	0.03	0.92	1.14	0.94	0.00	0.00	1.02	1.28	2.25	0.03	0.00
UpRecomm	1.27	1.88	2.83	1.00	0.81	0.77	1.08	3.98	1.00	0.92	0.85	1.21	4.52	1.00	0.93
VarCF	1.80	1.24	1.70	0.04	0.05	1.23	1.02	0.56	0.00	0.00	1.43	1.10	1.25	0.00	0.00
VolMkt	1.13	1.58	1.58	1.00	0.00	0.58	0.96	1.27	1.00	0.02	0.77	1.18	1.87	1.00	0.01
VolSD	0.93	1.32	2.99	1.00	0.00	0.79	0.84	0.23	1.00	0.00	0.87	1.10	1.95	1.00	0.00
VolumeTrend	1.19	1.73	2.28	1.00	0.10	0.70	1.36	4.14	1.00	0.21	0.87	1.49	4.58	1.00	0.08
XFIN	0.44	1.58	3.34	1.00	0.16	0.60	1.33	1.99	1.00	0.06	0.50	1.48	3.94	1.00	0.09
betaVIX	0.60	1.66	3.15	0.30	0.82	0.55	0.73	0.84	1.00	0.11	0.57	1.13	2.84	0.40	0.81
cfp	1.38	1.74	1.85	1.00	0.55	1.07	1.25	0.45	0.28	0.08	1.23	1.51	1.27	0.32	0.09
dNoa	0.63	1.68	6.02	1.00	0.55	0.97	1.27	1.76	0.31	0.03	0.74	1.55	5.93	1.00	0.56
fgr5yrLag	0.39	1.22	1.92	1.00	0.08	1.12	1.10	0.04	0.02	0.05	0.96	1.13	0.69	1.00	0.05
grcapx	1.30	1.80	3.93	1.00	0.30	0.88	1.07	1.44	0.72	0.20	1.10	1.46	3.84	1.00	0.34
grcapx3y	1.30	1.89	3.81	0.56	0.18	0.92	1.04	0.88	0.33	0.01	1.13	1.50	3.49	0.64	0.22
hire	0.99	1.51	4.65	1.00	0.33	0.92	0.98	0.29	0.46	0.19	0.98	1.41	4.31	1.00	0.32
iomom_cust	0.68	1.40	2.38	0.38	0.63	0.63	1.09	1.83	1.00	0.57	0.66	1.26	2.99	0.42	0.77
iomom_supp	0.81	1.41	1.82	0.41	0.46	0.33	0.90	1.84	1.00	0.49	0.60	1.19	2.55	1.00	0.53
realestate	0.88	1.17	1.90	0.67	0.78	1.06	1.30	1.36	1.00	0.52	0.93	1.21	2.31	0.57	0.90
retConglomerate	0.43	1.76	2.75	1.00	0.00	0.70	0.93	0.33	0.45	0.01	0.48	1.60	2.71	1.00	0.00
roaq	0.28	1.97	4.31	1.00	0.56	0.36	0.95	1.59	1.00	0.15	0.31	1.63	4.56	1.00	0.51
sfe	0.81	1.62	2.13	1.00	0.33	1.02	1.20	0.30	0.32	0.06	0.93	1.38	1.21	0.34	0.05
sinAlgo	1.11	1.32	1.64	1.00	0.36	0.80	1.36	1.81	0.03	0.33	1.04	1.33	2.37	0.41	0.45
skew1	0.45	0.99	2.18	0.26	0.60	0.48	0.79	2.08	0.47	0.91	0.47	0.88	3.02	0.28	0.86
std_turn	0.65	1.45	3.20	1.00	0.06	0.54	0.74	0.41	1.00	0.00	0.60	1.13	2.07	1.00	0.01
tang	1.04	1.75	2.81	0.32	0.29	1.09	1.23	0.53	0.00	0.00	1.06	1.54	2.62	0.13	0.07
zerotrade	0.77	1.26	2.87	1.00	0.00	0.68	0.89	0.57	1.00	0.00	0.74	1.15	2.61	1.00	0.00
zerotradeAlt1	0.72	1.36	3.66	1.00	0.00	0.57	0.90	0.88	1.00	0.02	0.68	1.23	3.37	1.00	0.00
zerotradeAlt12	0.90	1.29	2.96	1.00	0.00	0.82	0.83	0.04	1.00	0.00	0.87	1.16	2.23	1.00	0.00