Beyond Arbitrage: Deviations from the Risk-Return Tradeoff *

Benjamin Holcblat† Abraham Lioui‡ Michael Weber§

This version: March 3, 2024

Abstract

Differences in expected returns reflect and guide investment decisions in the economy. One challenge for economic agents and academics is to understand whether differences in risk exposure drive differences in expected returns. We propose two tests to assess whether risk alone can explain differences in expected returns. We develop the tests' equilibrium foundations, which hold within a large class of models. We study the tests' properties asymptotically and in simulations. Empirically, we find that risk cannot explain most differences in expected returns of characteristic-sorted portfolios. Our findings are consistent with a prominent role for frictions and behavioral biases.

JEL classification: G12, C58, C38, D53.

Keywords: Market frictions; Behavioral biases; Risk-return tradeoff; Strong SSD.

---

*The paper benefited from comments and discussions with Laurent Barras, Alon Brav, Svetlana Bryzgalova, Carter Davis, Amit Goyal, Steffen Grønneberg, Michael Halling, Christian Julliard, Christos Koulovationos, François Le Grand, Marco Lyrio, Thiago de Oliveira Souza, Stefan Nagel, Andy Neuhierl, Markus Pelger, Julien Pénasse, Seth Pruit, Mirco Rubin, Shri Santosh, Roberto Steri, Hongjun Yan, Raman Uppal, Irina Zviadadze, and seminar/conference participants at CFE London, EDHEC, French Inter Business Schools seminar in Finance, University of Luxembourg, 3rd Frontiers of Factor Investing Conference, FRA Conference, Paris Hedge Fund Conference, Paris Dauphine, FutFinInfo (HEC Paris), SoFiE, CEPR/ESSFM (Gerzensee), and EFA (Amsterdam). 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

1 Introduction

Differences in expected returns reflect and guide investment decisions in the economy, and hence, they are closely related to firm behavior and aggregate outcomes such as unemployment and economic growth. In financial economics, we typically distinguish two main approaches to explain asset returns: *relative* valuation and *absolute* valuation.

Relative valuation determines expected asset returns as a function of the expected returns of other assets. The reasoning is that similar assets should have similar returns. Unlike relative valuation, absolute valuation is concerned with the fundamental economic determinants of the expected return of an asset. Absolute valuation deduces asset returns from an equilibrium outcome determined by individuals' preferences, production technology, and other fundamental economic determinants. Absolute valuation relies on a specific model of the economy. Relative valuation relies on fewer assumptions —mainly no arbitrage—, but it is often not applicable in its pure form and does not allow the investigation of the overall determinants of asset returns. In blunt terms, relative valuation tells us that "two-quart bottles of ketchup invariably sell for twice as much as one-quart bottles of ketchup" (Summers, 1985), but it is silent about the origin of the market value of one-quart bottles of ketchup in the first place. Moreover, it ignores the impossibility of replicating one-third bottle of ketchup with one-quart bottles of ketchup, so relative valuation in its pure form is often impossible due to replication cost and error.

In the present paper, we propose two tests to investigate differences in expected returns —excess returns— without assuming a specific model of the economy. We inject *just a bit* of absolute valuation into relative valuation. If individuals prefer more to less, nobody leaves money on the table. In other words, preference monotonicity implies no arbitrage, which is the main foundation for relative valuation. In the present paper, we add risk aversion to preference monotonicity. We assume that individuals prefer higher returns *and* more stable returns. The combination of monotonicity and risk aversion leads to the risk-return tradeoff, which is the bedrock of frictionless financial economics with rational individuals.

We propose two tests to detect profitable deviations from the risk-return tradeoff. The two tests assess whether risk can explain expected excess returns, that is, the difference between the higher expected return asset and the lower expected return asset. For this purpose, we need a definition of risk. We go back to the basics and define risk as any variability risk-averse individuals dislike — anything discounted by increasing and concave utility functions. Then, the core idea behind our two tests is to assess whether every possible risk-averse individual strictly prefers the higher expected return asset over the lower expected return asset. If this preference does not hold for all individuals, at least one possible risk-averse individual prefers to forgo the higher expected return asset in exchange for the lower but less risky expected return asset. Then, risk can explain the difference in expected returns between the two assets. More precisely, the expected excess return is a possible compensation for the higher risk of the higher expected return asset.

We apply the two tests to differences in expected returns of characteristic-sorted portfolios. Researchers and practitioners sort stocks according to the value of a characteristic, such as firms' market capitalization, divide the sorted stocks into groups according to some percentiles (e.g., deciles), and then form portfolios based on the groups. The resulting portfolios are the characteristic-sorted portfolios. If the expected returns appear monotonic in the characteristic, researchers form a long-minus-short strategy by subtracting low-return portfolio returns from high-return portfolio returns. Long-minus-short strategies based on multivariate sorting similarly have a long leg with a high expected return and a short leg with a low expected return. The risk-return tradeoff stipulates that the higher expected returns of the long leg should compensate for higher risk. Thus, if risk alone cannot explain the spread in expected returns between the two legs of the strategy, we call it a deviation from the risk-return tradeoff. Regressions on these long-minus-short strategies are a standard tool to determine firms' cost of capital and portfolios' required expected return. In line with the relative valuation principle, firms' cost of capital should be equal to stocks' expected return with similar characteristics. The regression coefficients of the long-minus-short strategies are deemed to capture the part of a cost of capital or the part of the required expected return explained by the corresponding characteristic.

The null hypothesis of the first test corresponds to unconditional strict preferences of risk-averse individuals for the long leg. 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 risk). To more deeply tie the tests to financial economics, we investigate the economic content of the two tests beyond a pairwise comparison of assets. In an economy with diversification benefits, differences in expected returns between two tradable assets should compensate for *undiversified* risk. We show that if the tests' null hypotheses hold, undiversified risk alone is unlikely to explain the expected excess return;

that is, the latter should exceed risk compensations required by individuals. The intuition behind the result is that *undiversified* risk is unlikely to explain an expected returns difference if the *total* risk cannot explain the expected returns difference in the first place. The null hypotheses of the tests correspond to what we call *strong* second-order stochastic dominance (SSD), which is the standard SSD condition with strict inequality instead of weak inequality. A strict inequality is necessary for a difference in expected returns to exist, and thus, it is needed to derive the equilibrium foundations of the tests.

We show that the equilibrium foundations for both tests are valid in one-period and multi-period settings. The equilibrium foundations of the test rely on a local cross-sectional optimality condition implied by Euler equations. We also demonstrate that the equilibrium foundations hold independently of the economy's structure (e.g., whether or not individuals optimally diversify risk, whether or not markets are complete, whether or not a representative agent exists, etc.). By construction, an excess return is a costless portfolio because it consists of buying $1 of the long leg and selling $1 of the short leg, so optimality conditions for excess returns do not depend on budget constraints. Thus, the theoretical foundations of the proposed tests hold within a large class of models.

To assess the performance of the tests, we investigate their properties mathematically, numerically, and empirically. First, building on the statistics and econometrics literature on SSD (McFadden, 1989), we show that the tests have good asymptotic properties; that is, they are valid and consistent. Second, we investigate their finite-sample properties through Monte-Carlo simulations, confirming the asymptotic properties of the tests. Finally, as a proof of concept, we apply the unconditional test to the difference 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 indicate that risk can explain the US stocks' excess returns over US Treasury bills returns.

In addition to the above mentioned properties, the tests possess several note-worthy properties. First, the tests are *comprehensive.* The tests do not rely on a specific measure of risk (e.g., variance) or utility function (e.g., constant relative risk-aversion utility function) because they test the strict preference for the long leg, accounting for all types of risks disliked by risk-averse individuals, including high-order moments and tail risks.

Second, the unconditional test is immune 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). Therefore, for samples of sufficient size, we are unlikely to incorrectly classify by luck an expected excess return as a possible compensation for risk contrary to standard tests. By construction, a test for significance at the 5% level classifies an insignificant return spread as significant 5% of the time, even asymptotically, giving rise to the issues of multiple hypothesis testing and pretesting. Testing strong SSD (instead of the usual SSD) is critical to obtaining zero asymptotic type I error.

Despite the aforementioned properties, we do not claim that the proposed tests are without limitations. A possible limitation concerns the equilibrium foundations of the tests. Beyond the pairwise comparison of assets, the equilibrium foundations of the tests rely on Taylor expansions, so they are valid up to approximation errors. Taylor expansions are ubiquitous in economics and finance (e.g., log linearizations such as the Campbell-Shiller decomposition) and empirical work (e.g., inference based on asymptotic approximations) and have been found useful. In the present paper, approximation errors are unlikely to affect the empirical results because we can arbitrarily recenter the Taylor expansions to shrink approximation error terms. Nevertheless, one should take results based on Taylor approximations with a grain of salt because of the very nature of approximations. In summary, we do not claim that the present paper exhausts the question of the economic content of differences in expected returns. We hope it sheds new light on whether risk alone can explain them.

A long debate about the origin of differences in expected returns exists with three main competing explanations. The first strand of the literature asserts that risk explains differences in expected return. It typically assumes rational individuals and a frictionless financial market. A second strand of literature emphasizes the impact of market frictions, such as information asymmetry and leverage constraint, on expected returns (e.g., Merton, 1987; Shleifer and Vishny, 1997; Gromb and Vayanos, 2010; He and Krishnamurthy, 2018; Korsaye et al., 2021; Sandulescu, 2022). A third strand of literature introduces behavioral biases to explains differences in expected return (e.g., Bondt and Thaler, 1985; Jegadeesh and Titman, 1993; Hong and Stein, 1999; Bouchaud et al., 2019; Bordalo et al., 2019; Barberis et al., 2021). Thanks to empirical evidence and methodological developments (e.g., Coibon and Gorodnichenko, 2012, 2015; Korsaye, 2024), frictions and behavioral biases have become compelling alternatives to the dominant risk-based explanation, but we lack general tools to assess their importance for differences in expected returns. The present paper serves that purpose. Our empirical results indicate that frictions and/or behavioral biases are prominent drivers of differences in expected returns.

Understanding the relation between risk and returns is more than mere academic curiosity. In many situations, the practical implications of an expected excess return depend on whether risk can explain it. If risk cannot explain a difference in expected excess returns, a public authority —e.g., the US Securities and Exchange Commission— may design policies to eliminate the friction or the behavioral bias behind it to improve market efficiency. In the same situation, an individual would likely want to invest in the corresponding long-minus-short strategy —if possible— and thus earn higher expected returns. For investment decisions, if risk cannot explain a long-minus-short strategy based on characteristic-sorted portfolios, investors might avoid using it to risk-adjust future cash flows.

Related literature

To our knowledge, our paper is the first to propose general tests to detect deviations from the risk-return tradeoff. Nevertheless, in addition to the already mentioned papers, we build on and contribute to several strands of the literature.

In terms of ideas, the present paper follows a large tradition in financial economics of "doing something without having to do everything" (Hansen, 2013). This tradition investigates key equilibrium conditions without imposing a specific model of the economy so results become "immune to mistakes in how one might fill out the complete specification of the underlying economic model" (Hansen, 2013). Hansen and Singleton (1982) and Hansen and Jagannathan (1991) are prototypical examples of this tradition. Within this tradition, in terms of broad goals, the present paper is closer to Bernardo and Ledoit (2000) and Cochrane and Saa-Requejo (2000), who also strengthen no-arbitrage to get sharper pricing implications. Nevertheless, we differ in methodology and the specific application. First, we propose statistical tests instead of pricing bounds. Second, our tests rely on strong stochastic dominance instead of a "gain-risk" ratio. Third, we apply the tests to characteristic-sorted portfolios instead of options. Finally, our tests do not require the calibration of a threshold parameter that defines "good deals."

In terms of methodology, we build on a large econometric literature on tests of stochastic dominance (e.g. Davidson and Duclos, 2000; Linton et al., 2005), which essentially originates with McFadden (1989). Our unconditional test is a subsampling implementation of a modified McFadden (1989) test. Two features differentiate our unconditional test from the existing literature. First, our null hypothesis corresponds to strong SSD instead of the usual SSD. Second, guided by the Morse-Sard theorem (Morse, 1939; Sard, 1942), we deduce zero type I error from strong SSD under general assumptions. 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 long literature in mathematics on SSD, which goes back to Hardy et al. (1929). The SSD literature in finance has mainly focused on portfolio allocation or general equilibrium implications of stochastic dominance (e.g., Post, 2003; Hodder et al., 2015). Recently, Chalamandaris et al. (2021) and Arvanitis et al. (2022), building on Arvanitis et al. (2019) and Scaillet and Topaloglou (2010), propose a method to assess whether adding an asset to a given set of assets is beneficial for every risk-averse investor and for every investor with a prospect-theory utility, respectively. These are spanning tests for factor investing but do not allow to detect deviation from the risk-return tradeoff. We also contribute to this literature by introducing the concept of *strong* SSD, that is, the replacement of weak inequalities by strict inequalities in the different characterizations of SSD.[1] This modification is crucial for the equilibrium foundations of the null hypotheses of our tests: If we allowed for equality, some individuals could be indifferent between the long and the short leg so that both legs could coexist in equilibrium, and hence, no deviation from the risk-return tradeoff would exist.

2 Motivation and equilibrium foundations

We now discuss the motivation for the tests, explain their null hypotheses, and the equilibrium foundations. For simplicity, we focus on the unconditional test. Section 4 shows the logic behind the conditional test is similar to the unconditional test.

[1] *Strict* SSD qualifies the situation in which all possible *strictly* risk-averse individuals (i.e., individuals with a strictly concave von Neumann-Morgenstern utility function) 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 Excess returns are not necessarily risk compensations

2.1.1 Simple case

The primary motivation for the tests is the possible existence of expected excess returns unexplained by risk alone. To support the motivation, we now provide a simple model economy, in which an expected excess return is not a risk compensation, although factors span it. By "factor," we mean a variable that predicts differences in expected returns. In the model, the expected excess return does not compensate investors for risk, but rather arises due to a friction. The model is in the spirit of existing models that introduce a friction to explain the existence of differences in expected return (e.g., Merton, 1987; Gromb and Vayanos, 2010; Frazzini and Pedersen, 2014), but it is more parsimonious. Moreover, unlike existing models, it shows that even expected excess returns spanned by factors do not need to be risk compensations: We derive a factor model representation (equation (3) below) that explicitly incorporates the friction in the form of a factor. In empirical financial economics, the dominant risk-based approach assumes that expected excess returns spanned by factors are risk compensations. We motivate the tests with a friction-driven expected excess return, but behavioral biases can also drive expected excess return and thus motivate our tests (e.g., Hong and Stein, 1999; Bouchaud et al., 2019; Bordalo et al., 2019).

Consider a representative investor who maximizes her expected utility subject to constraints on long positions. More specifically, the representative investor solves the following mean-variance problem

$$\begin{cases} \max_{w \in \mathbf{R}^K} w'(\mu - R_0 \mathbf{1}) - \frac{\lambda}{2} w' \Sigma w \\ w_k \leqslant \bar{M}_k, \quad \text{for } k = 1, 2, \ldots, K, \end{cases} \tag{1}$$

where vector $R := (R_1 \ R_2 \ \ldots \ R_K)'$ denotes the vector of gross returns of risky assets, $\mu := \mathbb{E}(R)$ the expected gross return of risky assets, $\Sigma := \mathbb{V}(R)$ the variance-covariance matrix of risky assets' gross returns, R_0 the gross return of the risk-free rate, w_k the fraction of initial wealth invested in the asset k, $w := (w_1 \ w_2 \ \ldots w_K)'$, $\mathbf{1} := (1 \ 1 \ \ldots 1)'$ a $K \times 1$ vector of ones, $\bar{M} := (\bar{M}_1 \ \bar{M}_2 \ \ldots \ \bar{M}_K)'$ the vector of upper bounds on long positions, and $\lambda > 0$ captures risk aversion. The constraint on long position $\bar{M}_k$, for example, can be due to regulation (e.g., risk management). The existence of a solution to the mean-variance problem (1) is a sufficient condition for the existence of general equilibrium economy with a representative investor maximizing the mean-variance problem (1). Luttmer (1996) and He and Modest (1995) are prominent examples of representative agents in economies with frictions, and Luttmer (1992) provides aggregation results.

Solving (1) is equivalent to maximizing the Lagrangian

$$\max_{w \in \mathbf{R}^K} w'(\mu - R_0\mathbf{1}) - \frac{\lambda}{2} w'\Sigma w - \delta'(w - \bar{M}),$$

where δ is the vector of Lagrange multipliers for the constraints on long-position. Thus, the first order condition is $\mu - R_0\mathbf{1} - \lambda\Sigma w^* - \delta = 0$, which is also equivalent to

$$\mu - R_0\mathbf{1} = \lambda\Sigma w^* + \delta = \lambda\Sigma w^* + \lambda_\delta \Sigma w_\delta, \tag{2}$$

where $w_\delta := \frac{\Sigma^{-1}\delta}{\mathbf{1}'\Sigma^{-1}\delta}$ is Roll (1980)'s orthogonal portfolio, and $\lambda_\delta := \mathbf{1}'\Sigma^{-1}\delta$. Thus, the expected excess return of any portfolio w_p is

$$w_p'(\mu - R_0\mathbf{1}) = \lambda w_p'\Sigma w^* + \lambda_\delta w_p'\Sigma w_\delta. \tag{3}$$

The factor model (3) consists of two factors, $(w^*)'R$ and $w_\delta' R$, where $(w^*)'R$ corresponds to a market-type factor.[2] In standard factor models, the lambdas λ and λ_δ are called the prices of risk of the risk factors, $(w^*)'R$ and $w_\delta' R$, respectively. In fact, whereas $(w^*)'R$ is a risk factor, the factor $w_\delta' R$ is *not* a risk factor. It arises due to the constraints on long positions. In the absence of constraints, the Lagrange multiplier δ and λ_δ are zero, and thus the only factor is the market-type factor $(w^*)'R$, that is, the factor $w_\delta' R$ does not exist. The factor $w_\delta' R$ drives the wedge between the expected excess return $\mathbb{E}(R_k - R_0)$ and the risk compensation driven by the market-type factor $(w^*)'R$. The equivalence between the expected excess return representations (2) and (3) also shows that the decomposition of expected excess return in terms of an intercept, the so-called "alpha", and "factors" depends on the factors used: An "alpha" can be represented by a factor, and vice versa.[3]

[2]See Appendix A for a detailed derivation of an augmented-CAPM representation of the factor model (3) and its tangency portfolio.

[3]In the expected excess return representation (2), we denote the intercept with "δ" instead of "α" because we use the latter to denote tests' level in Sections 3 and 4.

2.1.2 General case

Friction-driven expected excess returns are not an artifact of the previous model nor of mean-variance preferences. Under general assumptions that allow for different types of frictions (e.g., bid-ask spreads, transaction costs, and constraints on long positions), the fundamental theorem of asset pricing and Riesz representation theorem show that no-arbitrage implies the existence of at least one strictly positive stochastic discount factor (SDF) M and a vector δ such that (s.t.)

$$\mathbb{E}[M(R - R_0\mathbf{1})] = \delta, \tag{4}$$

where δ belongs to a subset of $\mathbf{R}^K$ determined by the frictions (Jouini and Kallal, 1995; Luttmer, 1996; Korsaye et al., 2021, Proposition 1). The vector δ corresponds to the wedge due to frictions. In the standard textbook presentations of SDFs, the wedge vector $\delta = 0$ because free portfolio formation is assumed, that is, frictions are ruled out. The pricing equation (4) shows that both risk and frictions are necessary to explain differences in expected returns when free portfolio formation is not assumed. The wedge vector δ captures the part explained by frictions, whereas the SDF M captures the part explained by (undiversifiable) risk. Hereafter, without loss of generality, we impose $\mathbb{E}(M) = 1$ because we can divide both sides of the pricing equation (4) with $\mathbb{E}(M)$.

Then, by the pricing equation (4), $\mathbb{C}\mathrm{ov}(R, M) + \mathbb{E}(R - R_0\mathbf{1}) = \delta$ so $\mathbb{E}(R - R_0\mathbf{1}) = -\mathbb{C}\mathrm{ov}(R, M) + \delta$, which, in turn, implies that, the expected excess return of any portfolio w_p is

$$w_p'(\mu - R_0\mathbf{1}) = \mathbb{C}\mathrm{ov}(w_p'R, -M) + \lambda_\delta w_p'\Sigma w_\delta, \tag{5}$$

where $\mu := \mathbb{E}(R)$, $\Sigma := \mathbb{V}(R)$, $\lambda_\delta := \mathbf{1}'\Sigma^{-1}\delta$, and $w_\delta := \frac{\Sigma^{-1}\delta}{\mathbf{1}'\Sigma^{-1}\delta}$. The two-factors model (5) generalizes the simple two-factor model (3): The covariance $\mathbb{C}\mathrm{ov}(w_p'R, -M)$ corresponds to the term $\lambda w_p'\Sigma w^*$ in the simple two-factor model (3). The equivalence of the expected excess-returns representations (4) and (5) confirms that the distinction between so-called "alpha" intercept and "factors" depends on the factors used. In addition, the two-factor model (5) shows the dominant risk-based approach would wrongly classify the factor $w_\delta R$ and any expected excess return spanned by the latter as a risk compensation. This is why we propose tests to assess whether risk alone can explain the difference in expected returns.

2.2 Null hypothesis

By definition, risk is anything individuals with concave von Neumann-Morgenstern utility functions dislike. Concavity discounts departures from the mean by Jensen's inequality. The combination of preference monotonicity and concavity especially discounts negative departure from the mean, that is, downside risk. Then, a risk-averse individual has an increasing and concave von Neumann-Morgenstern utility function. We link this characterization of risk-averse individuals with excess returns. Expected excess returns correspond 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, for each expected excess return, the basic idea is to test 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

$$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\, \mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)], \tag{6}$$

where $\mathbf{U}_2$ denotes a class of concave and increasing functions, and R_S and R_L denote the gross returns of the long leg and the short leg, respectively. If the null hypothesis (6) is rejected, then at least one possible risk-averse individual weakly prefers the short leg to the long leg, so risk can explain the spread in expected returns. In other words, a possible risk-averse individual prefers to forego the higher expected return of the long leg in exchange for the lower expected return of the short leg because the latter is less risky. Then, risk can explain the expected excess return. Testing for all possible utility functions in $\mathbf{U}_2$ allows us to sidestep the choice of a specific measure of risk, that is, the choice of a specific utility function $u(.)$.

The null hypothesis (6) is similar to the well-known SSD. The difference arises from the use of *strict* inequalities instead of *weak* inequalities, that is, the null hypothesis (6) rules out the possibility of risk-averse individuals who are indifferent between the long and the short leg. Hereafter, when the null hypothesis (6) holds, we say that R_L *strongly* SSD dominates R_S.

The replacement of weak inequalities is key from an economic point of view. SSD is not a sufficient condition for a deviation from the risk-return tradeoff for at least two reasons. First, it does not guarantee a strictly positive expected excess return $\mathbb{E}(R_L - R_S)$. Second, the modification is central to the equilibrium foundations of the tests. If some individuals are indifferent between the long and the short leg, then both legs can coexist in a frictionless equilibrium, and hence no deviation from the risk-return tradeoff 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.[4]

2.3 Equilibrium foundations

In this section, we show that, under general assumptions, the null hypothesis (6) should be a sufficient condition for a deviation from the risk-return tradeoff. We label an expected excess return a deviation from the risk-return tradeoff if risk alone cannot explain the expected excess return, that is, if the expected excess return exceeds all possible risk compensations required by risk-averse individuals.

2.3.1 Equilibrium Foundations without Diversification Benefits

In the absence of diversification benefits, the equilibrium implication of the null hypothesis (6) is immediate. Assume every individual has to invest all her wealth W_0 either in the short leg or in the long leg, so no diversification benefits exist. Furthermore, assume all individuals have strictly increasing von Neumann-Morgenstern utility functions in $\mathbf{U}_2$. If all possible individuals strictly prefer the returns of the long leg to the returns of the short leg, then by the invariance of the null hypothesis under strictly positive affine

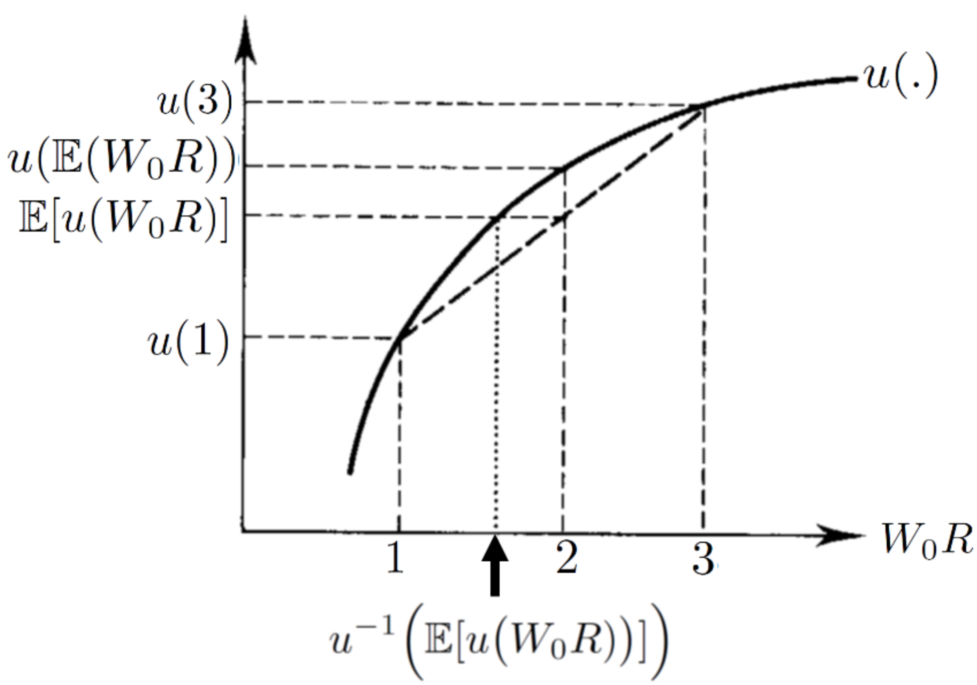
[4] Another way to obtain strict inequalities instead of weak inequalities is to rule out affine utility functions from the class $\mathbf{U}_2$ and rely on *strict* SSD. The latter corresponds to the situation in which all possible individuals with a strictly concave von Neumann-Morgenstern utility function 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 standard benchmark in finance and economics; (ii) As previously mentioned, it does not guarantee the existence of a strictly positive expected excess return $\mathbb{E}(R_L - R_S)$

transformations of lotteries (Lemma 1 on p. 22)

$$\begin{aligned} & \mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)] \\ \Leftrightarrow\ & \mathbb{E}[u(W_0 R_S)] < \mathbb{E}[u(W_0 R_L)] \\ \Leftrightarrow\ & u^{-1}\Big(\mathbb{E}[u(W_0 R_S))]\Big) < u^{-1}\Big(\mathbb{E}[u(W_0 R_L)]\Big), \end{aligned}$$

where $u^{-1}\Big(\mathbb{E}[u(W_0 R_S)]\Big)$ and $u^{-1}\Big(\mathbb{E}[u(W_0 R_L)]\Big)$ are the certainty equivalents of the investment payoffs of the short and long leg, respectively. In words, all possible risk-averse individuals value the investment payoff of the long leg strictly higher than the investment payoff of the short leg; that is, the private value of the long leg investment payoff $W_0 R_L$ is higher than that of the short leg $W_0 R_S$ for all possible risk adjustments.

Figure 1: Risk aversion and asset pricing without diversification benefits



Notes: For simplicity, we assume $\mathbb{P}(W_0 R = 1) = \mathbb{P}(W_0 R = 3) = \frac{1}{2}$ so $\mathbb{E}(W_0 R) = 2$. Risk aversion corresponds to the concavity of the von Neumann-Morgerstern utility $u(.)$. By Jensen's inequality, concavity implies $\mathbb{E}[u(W_0 R)] \leqslant u(\mathbb{E}(W_0 R))$, that is, the individual prefers the sure amount of money $\mathbb{E}(W_0 R)$ to the random payoff $W_0 R$. The certainty equivalent $u^{-1}(\mathbb{E}[u(W_0 R))])$ is the amount of money that makes an individual with von Neumann-Morgerstern utility $u(.)$ indifferent between an asset with payoff $W_0 R$ and the sure amount of money $u^{-1}(\mathbb{E}[u(W_0 R))])$. In other words, the certainty equivalent indicates how much an individual values an asset in the absence of diversificaiton benefits. Then, the risk premium is $\mathbb{E}(W_0 R) - u^{-1}(\mathbb{E}[u(W_0 R))])$.

Figure 1 illustrates how risk-averse individuals value investment payoff. Now, by the definition of gross returns, the market price of both investments is W_0. Thus, every individual tries to buy the long leg. Hence, the price of the long leg relative to the price of the short leg increases and its excess return decreases until some individuals are indifferent between the two. At the equilibrium, all individuals cannot strictly prefer the long leg. It yields the following definition of a deviation from the risk-return tradeoff.

Definition 1 (Deviation from the risk-return tradeoff in the absence of diversification benefits)**.** *In the absence of diversification benefits, an expected excess return* $\mathbb{E}[R_L - R_S]$ *is a deviation from the risk-return tradeoff if, for all von Neumann-Morgenstern utility functions* $u \in \mathbf{U}_2$*,* $\mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)]$*.*

As a mirror of Definition 1, in the absence of diversification benefits, an expected excess return $\mathbb{E}[R_L - R_S]$ is a possible risk compensation if there exists $u(.)$ in $\mathbf{U}_2$ s.t. $\mathbb{E}[u(R_L)] \leqslant \mathbb{E}[u(R_S)]$. In words, an expected excess return $\mathbb{E}[R_L - R_S]$ is a possible risk compensation if there exists a risk-averse individual who prefers to forego the higher expected return of the long leg in exchange for the lower expected but less risky return, of the short leg. In the latter case, risk alone can explain the difference in expected returns between the long and the short leg.

2.3.2 Equilibrium Foundations with Diversification Benefits

In an economy with several assets, the above 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-diversifiable risk corresponds to the movement of individuals' consumption, so the priced risk corresponds to the comovements of excess returns with individuals' consumption.

We now show the null hypothesis (6) "$H_0 : \forall u \in \mathbf{U}_2, \mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)]$" should still be a sufficient condition for a deviation from the risk-return tradeoff in the presence of diversification benefits. More precisely, we show the null hypothesis (6) implies risk alone is unlikely to explain the expected excess return; that is, the latter exceeds the risk compensations required by risk-averse individuals.

For this purpose, we first derive the possible risk compensations for expected excess returns under general assumptions. The assumptions should be as general as possible but not allow for behavioral biases or frictions affecting the expected excess return: We want risk compensations, not compensations for frictions or behavioral biases. The following derivation shows it is sufficient to consider a situation in which such individuals optimally and freely trade the excess return in a neighborhood of their locally optimal consumption. Importantly, we do not need to assume a specific equilibrium model.

Derivation of Risk Compensation

By construction, an excess return $R_L - R_S$ is a costless portfolio because it consists of buying \$1 of the long leg and selling \$1 of the short leg. Thus, at any date t, for any individual, irrespective of budget constraints, as long as an excess return freely trades in a neighborhood of the locally optimal consumption C_t, the expected marginal value of an excess return is zero, that is,

$$\mathbb{E}[u'(C_t)(R_{L,t} - R_{S,t})] = 0, \tag{7}$$

where $u(.)$, C_t, $R_{L,t}$ and $R_{S,t}$ denote, respectively, the individual's utility function, consumption at time t, the gross return of the long leg and short leg at time t. The mathematics behind the standard optimality condition (7) corresponds to the following Taylor approximations around C_t, that state, up to approximation errors,

$$\mathbb{E}[u(C_t + (R_{L,t} - R_{S,t}))] - \mathbb{E}[u(C_t)] = \mathbb{E}[u'(C_t)(R_{L,t} - R_{S,t})] \tag{8}$$

$$\mathbb{E}[u(C_t - (R_{L,t} - R_{S,t}))] - \mathbb{E}[u(C_t)] = -\mathbb{E}[u'(C_t)(R_{L,t} - R_{S,t})] \tag{9}$$

By the first Taylor approximation (8), if $\mathbb{E}[u'(C_t)(R_{L,t} - R_{S,t})] > 0$, one more unit of the costless portfolio $R_{L,t} - R_{S,t}$ would increase individual's utility so C_t would not be locally optimal. Similarly, by the second Taylor approximation (9), if $\mathbb{E}[u'(C_t)(R_{L,t} - R_{S,t})] < 0$, one less unit of the costless portfolio $R_{L,t} - R_{S,t}$ would increase individual's utility so C_t would not be locally optimal.[5]

By the optimality condition (7), $\mathbb{C}\text{ov}(u'(C_t), R_{L,t} - R_{S,t}) + \mathbb{E}[u'(C_t)]\mathbb{E}(R_{L,t} - R_{S,t}) = 0$, so an expected excess return explained solely by risk is s.t.

$$\mathbb{E}(R_{L,t} - R_{S,t}) = -\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\text{ov}(u'(C_t), R_{L,t} - R_{S,t}). \tag{10}$$

[5]See Appendix B.2 for a complete proof under general assumptions. Under more restrictive assumptions, the optimality condition (7) is also a direct implication of the standard Euler equations.

The expected excess return $\mathbb{E}(R_{L,t}-R_{S,t})$ should be the negative of its covariance with individuals' marginal utility normalized by individuals' expected marginal utility. Hence, the expected excess return should precisely compensate for its normalized negative comovements with the marginal utility of consumption C_t, and thus, for its normalized positive comovements with consumption C_t —the marginal utility function $u'(.)$ is decreasing due to concavity. If the expected excess return exceeds risk compensations required by all possible risk averse individuals, we call it a deviation from the risk-return tradeoff. More precisely, an expected excess return is a deviation from the risk-return tradeoff if marginal utilities of all possible risk-averse individuals cannot explain the expected excess return.

Definition 2 (Deviation from the risk-return tradeoff in the presence of diversification benefits)**.** *In the presence of diversification benefits, an excess return $R_{L,t} - R_{S,t}$ is a deviation from the risk-return tradeoff if, for all von Neumann-Morgenstern utility functions $u \in \mathbf{U}_2$,*

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

Definition 2 does not require us to assume a specific equilibrium model. The optimality condition (7), and thus equation (10), holds as long as individuals can freely trade the costless portfolio $R_{L,t} - R_{S,t}$ in a neighborhood around their locally optimal consumption C_t (see Appendix B.2). Thus, the quantity $-\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\text{ov}(u'(C_t), R_{L,t} - R_{S,t})$ should be the risk compensation for a large class of equilibrium models. In particular, Definition 2 accounts for so-called "background risk" because the latter affects the consumption C_t, which is all an individual cares about. Any source of income, such as human capital and real estate, ultimately funds consumption, which yields utility to individuals. For this reason, Definition 2 also covers the ICAPM, which is a particular case of the consumption CAPM by the envelope condition (e.g., Cochrane, 2005, Sec. 9.3). In a large class of equilibrium models, whether partial equilibrium or general equilibrium, whether with production or not, whether with complete or incomplete financial markets, etc., the right-hand side of equation (10) delivers the required risk compensation. If a wedge exists between the expected excess return $\mathbb{E}(R_{L,t} - R_{S,t})$ and the required risk compensation $-\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\text{ov}(u'(C_t), R_{L,t} - R_{S,t})$, an explanation other than risk is needed to account for the expected return of the factor $\mathbb{E}(R_{L,t} - R_{S,t})$. In the simple economy of Section 2.1, for example, the risk compensation is $-\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\text{ov}(u'(C_t), R_{L,t} - R_{S,t}) = \lambda w'\Sigma w^*$

for the representative agent, so the wedge is $\lambda_\delta w'\Sigma w_\delta$. By avoiding the specification of a particular equilibrium model, the equation (10) is "immune to mistakes in how one might fill out the complete specification of the underlying economic model" (Hansen, 2013).

Moreover, the derivation of equation (10) indicates alternative explanations should be based on frictions or behavioral biases that induce a violation of the optimality condition (7). Hence, an explanation can be an informational or trading friction on the long-minus-short strategy $R_{L,t} - R_{S,t}$. However, frictions on production or even a short-sale constraint on an asset that is not part of the factor cannot be an explanation. Note also that if a wedge exists for all concave increasing utility functions, the sole presence of "irrational" individuals cannot be an explanation as long as "rational" unconstrained individuals are present because equation (10) would need to hold for the "rational" individuals.

The Null Hypothesis (6) and Risk Compensation

The following proposition shows if the null hypothesis (6) holds, then the expected excess return $\mathbb{E}(R_{L,t} - R_{S,t})$ should exceed the risk compensation $-\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\text{ov}(u'(C_t), R_{L,t} - R_{S,t})$ for a large class of increasing and concave utility functions.

Proposition 1 (Equilibrium foundation for unconditional test)**.** *For any twice continuously differentiable strictly increasing and concave utility function u on $[\underline{u}, \overline{u}]$, which includes the support of C_t and the returns $R_{S,t}$ and $R_{L,t}$, up to approximation errors, the null hypothesis "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(R_{S,t})] < \mathbb{E}[u(R_{L,t})]$"[6] implies the expected excess return exceeds its risk compensation, i.e.,*

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

Proposition 1 provides sufficient assumptions under which strict preference for the long leg implies the existence of a deviation from the risk-return tradeoff up to approximation errors. If risk alone cannot explain the expected excess return $\mathbb{E}(R_{L,t} - R_{S,t})$, other explanations, such as behavioral biases or institutional frictions, are necessary to explain the expected excess return; so, we call the excess return $R_{L,t} - R_{S,t}$ a deviation from the risk-return tradeoff. The intuition behind Proposition 1 is the following. If total

[6] By stationarity (Assumption 3 in Appendix B.4), the time index in the statement of the null hypothesis is innocuous, i.e., the null hypothesis "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(R_{S,t})] < \mathbb{E}[u(R_{L,t})]$" is equivalent to the null hypothesis "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)]$."

risk cannot explain expected excess returns $\mathbb{E}(R_{L,t} - R_{S,t})$, undiversified risk is unlikely to explain it. In mathematical terms, if the excess returns' departures from their mean cannot explain expected excess returns $\mathbb{E}(R_{L,t} - R_{S,t})$, the part of the departures spanned by the marginal utility $u'(C_t)$ is also unlikely to explain it. The proof of Proposition 1 is based on Taylor expansions similar to (8) and (9). In the proof, it is key that Taylor expansions are around the random consumption C_t, so the random changes in C_t can account for the curvature of the utility function $u(.)$. In particular, approximating around C_t allows accounting for the concavity of the utility function, that is, risk aversion. In contrast, if the Taylor approximations were around the fixed value $\mathbb{E}(C_t)$, it would not be possible to account for the curvature of $u(.)$, thus neutralizing risk aversion.[7] Note the assumptions underlying Proposition 1 are mild. The assumptions do not require us to specify a data-generating process (DGP) for returns nor the primitives of an economy.

The presence of a deviation from the risk-return tradeoff, or more generally, the violation of the "frictionless" optimality condition (7), does not imply the existence of arbitrage opportunities in the economy. For example, in the simple economy of Section 2.1, the constraint on long positions implies the violation of the "frictionless" optimality condition (7) and the existence of a deviation from the risk-return tradeoff, but no arbitrage opportunity exists. If there were arbitrage opportunities, no (finite) solution to the portfolio choice problem (1) of the representative individual would exist. In fact, the second part of the fundamental theorem of asset pricing, that is, the equivalence between the absence of arbitrage and the existence of a solution to a portfolio choice problem, has been generalized to an economy with frictions (Jouini and Kallal, 1999).

2.4 Extension to exotic preferences

Several "exotic preferences", such as Epstein-Zin-Weil recursive preferences and preferences for environmental, social, and corporate governance (ESG) —non-pecuniary benefits from investing in stocks with high ESG scores—, have become common in financial economics.[8] This section shows that the null hypothesis (6) should be a sufficient condition for a deviation from the risk-return tradeoff even for these common exotic preferences. The null hypothesis (6) has implications beyond standard time-additive von Neumann-Morgenstern

[7] See Appendix B.3 for more details.

[8] Following Backus et al. (2004), by "exotic preferences," we mean preferences different from time additive von Neumann-Morgenstern utility.

utility because (i) it holds for a large class of increasing and concave utility functions and (ii) most exotic preferences yield optimality conditions similar to the optimality condition (10). For brevity, the goal of this section is not to be exhaustive but to focus on some of the most prominent examples of exotic preferences.

2.4.1 External habit and Epstein-Zin-Weil preferences

In the standard constant relative risk aversion (CRRA) consumption-based asset pricing model, the optimality condition (10) reduces to

$$\mathbb{E}[C_t^{-\gamma}(R_{L,t} - R_{S,t})] = 0, \tag{11}$$

where γ is the coefficient of relative risk aversion. This equation is a direct implication of the standard Euler equations $C_{t-1}^{-\gamma} = \mathbb{E}_{t-1}[\beta C_t^{-\gamma} R_i]$, for all $i \in \{L, S\}$. Rare disaster models a la Rietz (1988) and Barro (2006) imply the same optimality condition (11). In the case of exotic preferences, the optimality condition (11) becomes

$$\mathbb{E}[C_t^{-\gamma} Z_t(R_{L,t} - R_{S,t})] = 0, \tag{12}$$

where $Z_t := S_t^{-\gamma}$ for the Campbell and Cochrane (1999) external habit formation model with S_t being the consumption surplus ratio, and $Z_t := U_t^{-(\gamma-\frac{1}{\psi})}$ for Epstein-Zin-Weil preferences (Weil, 1989; Epstein and Zin, 1989) with U_t being the continuation utility, and ψ the intertemporal elasticity of substitution (IES). The Bansal and Yaron (2004) long run-risk model, which features Epstein-Zin-Weil preferences, implies the optimality condition (12). Thanks to the similarity between the optimality conditions (10) and (12), following a similar proof, we can obtain the counterpart of Proposition 1 for Campbell and Cochrane (1999) external habit and Epstein-Zin-Weil preferences. Thus, the null hypothesis (6) should also be a sufficient condition for a deviation from the risk-return tradeoff for these exotic preferences.

2.4.2 Non-pecunary ESG preferences

The tests should also be useful to economies in which individuals feature non-pecuniary ESG preferences. Pastor et al. (2021), Avramov et al. (2022), Baker et al. (2022) and Zerbib (2022) introduce individuals with a non-pecuniary utility for holding stocks which is proportional to the ESG score of the stock. Avramov et al. (2023) generalizes these papers to a multiperiod setting with Epstein-Zin-Weil preferences. They show that, in equilibrium, the Euler equation for the gross return on any asset k with an ESG score equal to $G_{k,t-1}$ is given by

$$\mathbb{E}_{t-1}\left[M_t R_{k,t}\right] = 1 - \eta_{t-1} G_{k,t-1}, \tag{13}$$

where η_{t-1} captures an ESG preference, and $M_{t+1} := \beta^{\theta}\left(\frac{C_t}{C_{t-1}}\right)^{-\frac{\theta}{\psi}} \tilde{R}_{W,t}^{\theta-1}$ is the SDF with $\tilde{R}_{W,t} := \frac{R_{W,t}}{1-\eta_{t-1}G_{W,t-1}}$ the ESG-adjusted gross return on the consumption asset, β is the time discount factor, ψ is the IES, γ the coefficient of relative risk, $\theta := \frac{1-\gamma}{1-\frac{1}{\psi}}$, and $C_{t-1} := \dot{C}_{t-1} + \eta_{t-1}G_{W,t-1}\left(W_{t-1} - \dot{C}_{t-1}\right)$ a consumption bundle consisting the consumption good $\dot{C}_{t-1}$ and the financial ESG consumption scaled by the ESG score $G_{W,t-1} := \sum_{k=1}^{K} w_{k,t-1}G_{k,t-1}$ of the wealth portfolio $(w_{1,t-1}\, w_{2,t-1}\, \ldots\, w_{K,t-1})$. In the Euler equations (13), the SDF M_t accounts for the ESG consumption risk, whereas $\eta_{t-1}G_{k,t-1}$ accounts for the ESG bias. The Euler equation (13) is similar to the Euler equation in Krishnamurthy and Vissing-Jorgensen (2012), where individuals face liquidity costs. The Euler equation (13) implies the following optimality condition

$$\mathbb{E}[C_t^{-\frac{\theta}{\psi}} Z_t (R_{L,t} - R_{S,t})] = \dot{\delta}, \tag{14}$$

where $\dot{\delta} := \mathbb{E}[\eta_{t-1}(G_{S,t-1} - G_{L,t-1})]$ embodies the ESG bias, and $Z_t := \beta^{\theta}\left(\frac{1}{C_{t-1}}\right)^{-\frac{\theta}{\psi}} \tilde{R}_{W,t}^{\theta-1}$. The optimality condition (14) is the counterpart of equation (4). Thus, similarly to Section 2.1.2, we can derive a two-factors model with risk and bias-driven factors. Thanks to the similarity between the optimality conditions (10) and (14), a counterpart of Proposition 1 for non pecuniary ESG preferences exists, so the tests should help disentangle between a consumption-risk factor and an ESG bias factor.

3 Unconditional Test

We now derive a testable version of the unconditional test and discuss its statistical properties.

3.1 Unconditional Null Hypothesis in a Testable Form

To derive the testable implications of the null hypothesis (6), 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 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 $u : [\underline{u}, \overline{u}] \to \mathbf{R}$ 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 $\check{u} := \min\big\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}]$ s.t., $\forall x \in [z, \overline{u}],\ u(x) = 0\}\big\}$. 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 B.1.1.

□

Well-known estimators of CDFs and functionals thereof exist, so Lemma 1 provides a way to test the null hypothesis (6). Lemma 1 is the strong counterpart of the well-known Hardy-Littlewood et. al. theorem for SSD.

Note, it is not sufficient to replace the weak inequalities in standard 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 (6)—, 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 other 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 (6) under strictly positive affine transformations of lotteries. This result implies the formulations of the null hypothesis (6) in terms of terminal wealth, capital gain, gross returns or any other strictly positive affine transformation thereof, are all mathematically equivalent, that is, $\forall u \in \mathbf{U}_2$, $\mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)] \Leftrightarrow \forall u \in \mathbf{U}_2$, $\mathbb{E}[u(W_0 R_S)] < \mathbb{E}[u(W_0 R_L)]$, 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 (6).

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). We can relax Assumption 1(a) at the cost of notational and mathematical complications. In particular, it saves us a list of technical assumptions ensuring the existence of the expectations. 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 the Morse-Sard theorem (Morse, 1939; Sard, 1942). Thus, Assumption 1(b) is harmless in practice, although it is crucial for the present paper. Thanks to Assumption 1(b), the null hypothesis (6) does not hold if, and only if, there exists $z \in (\underline{u}, \overline{u}]$ s.t. $F_S^{(2)}(z) < F_L^{(2)}(z)$.

3.2 Unconditional Test Statistic

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

3.2.1 Asymptotic properties

In many 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 (6) is equivalent to the null hypothesis

$$\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[,\ F_L^{(2)}(z) - F_S^{(2)}(z) < 0, \tag{15}$$

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 $F_i^{(2)}$ is the integrated empirical CDF $\hat{F}_i^{(2)}(z) := \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\{R_{i,t} \leqslant z\}(z - R_{i,t})$, for $i \in \{L, S\}$. 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,

$$\begin{aligned}
\sqrt{T}\mathrm{KS}_T^* :&= \sqrt{T} \sup_{z \in \mathbf{I}_T} \left| \hat{F}_L^{(2)}(z) - \hat{F}_S^{(2)}(z) - \min\{\hat{F}_L^{(2)}(z) - \hat{F}_S^{(2)}(z), 0\} \right| \\
&= \sqrt{T} \sup_{z \in \mathbf{I}_T} \left| \hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z) \right|, \qquad (16)
\end{aligned}$$

where $\mathbf{I}_T := [c_T, \overline{u}]$, with $c_T \downarrow \underline{u}$, and $\hat{F}_{L\wedge S}^{(2)}(z)$ denotes the minimum of the integrated empirical CDF (that is, $\hat{F}_{L\wedge S}^{(2)}(z) = \min\{\hat{F}_L^{(2)}(z), \hat{F}_S^{(2)}(z)\}$).[9] The estimator $\min\{\hat{F}_L^{(2)}(.) - \hat{F}_S^{(2)}(.), 0\}$ is a constrained estimator of $F_L^{(2)}(.) - F_S^{(2)}(.)$, because it satisfies the null hypothesis (15) by construction.

The following proposition shows the KS_T^* test statistic (16) defines a valid and consistent test of the null hypothesis (6).

Proposition 2 (No type I error and No type II error)**.** *Under Assumption 1 and the assumptions of Appendix B.4, for any level of the test* $\alpha \in]0, 1]$,

[9] The absolute value is superfluous in the Kolmogorov-Smirnov (KS) test statistic (16) because, for all $z \in \mathbf{R}$, $0 \leqslant \hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z)$ by the definition of $\hat{F}_{L\wedge S}^{(2)}(z)$. However, we keep the absolute value to emphasize that the KS test statistic (16) measures the distance between the unconstrained estimator $\hat{F}_L^{(2)}$ and the constrained estimator $\hat{F}_{L\wedge S}^{(2)}(z)$.

(i) if the null hypothesis (6) *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 (6) *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 B.4. □

Proposition 2 (i) shows the null hypothesis is asymptotically never rejected when it is true, i.e., no type I error exists asymptotically. We obtain this noteworthy property under general assumptions thanks to the Morse-Sard theorem (Morse, 1939; Sard, 1942) and the strict inequalities of the null hypothesis. The latter ensures that the test statistic $\sqrt{T}\mathrm{KS}_T^*$ equals 0 with probability 1 for T big enough if the null hypothesis is true. Proposition 2 (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 2 (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 drawing 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 returns 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 manner, block subsampling accounts for potential time- and cross-sectional dependence.

3.2.2 Monte-Carlo Simulations

In Table 1, Monte-Carlo simulation results show that the test statistic KS_T^*'s finite-sample properties align with Proposition 2. For all DGPs, p-values go to zero when the null hypothesis (15) is wrong. Also, in line with the asymptotic theory, a large and growing proportion of p-values equals one when the null hypothesis (6) holds because of the absence of type I error, asymptotically. The first two DGPs are Gaussian distributions calibrated to data. More precisely, we calibrate the DGPs to two long-minus-short strategies based on characteristic-sorted portfolios —size and the dividend yield— for which the null hypotheses are barely true (or false). This calibration should be challenging for the test. The third DGP is a stylized DGP except for the correlation between the long and short legs. We calibrate the latter correlation to the average correlation of the legs of some of the most prominent long-minus-short strategies based on characteristic-sorted portfolios. Further simulation results and details are available in Appendix C.

One insight from the simulations is that centered block subsampling tends to yield more rejections than uncentered block subsampling approximations. Hence, we use the centered subsampling approximation in our empirical implementation to be conservative. In Section 3.3, we also investigate the finite-sample properties of the tests on actual financial data.

3.2.3 Immunity to Multiple Hypothesis Testing and Pretesting

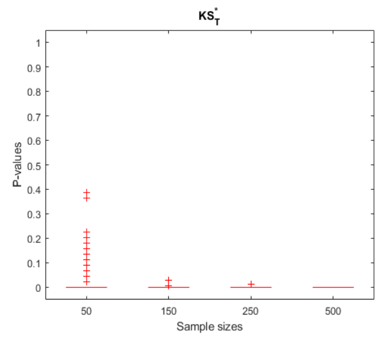
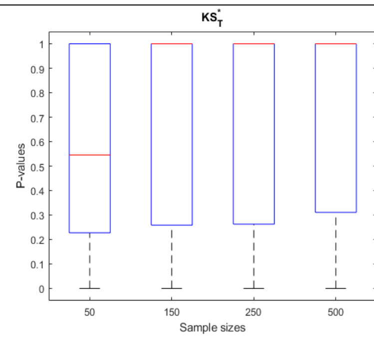
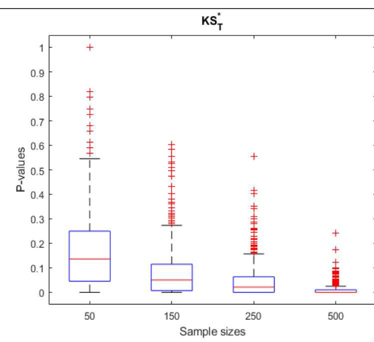
Because of the large number of excess returns often considered in the literature, multiple hypothesis testing is a concern. Harvey et al. (2016) highlights the concern in the case of long-minus-short strategies based on characteristic-sorted portfolios. The multiple hypothesis problem originates from the probability of wrongly rejecting at least one true hypothesis if one simultaneously tests many true hypotheses with the size and level of each test precisely equal to $\alpha \in (0, 1]$. 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 five true hypotheses asymptotically. The following Proposition 3 shows the unconditional test is immune to the multiple hypothesis problem.

Proposition 3 (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 and the long leg of the excess return* k*. Define the set* $\mathbf{J} \subset [\![1, K]\!]$ *of true hypotheses. Under the assumptions of Proposition 2, 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* (16) *that corresponds to the null hypothesis*

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

H_0	DGP	Boxplots of p-values
False	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 1.015 \\ 1.0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix}\right)$	
True	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 1.011 \\ 1.010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix}\right)$	
False	$\begin{cases} R_L \overset{IID}{\hookrightarrow} 1 + \mathrm{t}(4) \\ R_S \overset{IID}{\hookrightarrow} \mathcal{N}(1,1) \\ \mathbb{C}\mathrm{or}(R_S, R_L) = .7 \end{cases}$	

Notes: The first two data-generating processes (DGP) correspond to Gaussian distributions calibrated to factors for which H_0 are barely true (or false). The third DGP is a stylized DGP except for the correlation 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 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.

$\mathrm{H}_{0,j}$ *and* $\hat{c}_{j,1-\alpha}$ *is 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 2i, 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. □

Usual 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 3 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 that 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 difficult 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 solution 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 examining the data. The latter assumption is difficult to defend in the case of excess returns: The evolution of asset pricing is a hard-to-predict dialog between theory and many empirical studies. The following Proposition 4 shows that the unconditional test is immune to the pretesting problem.

Proposition 4 (Immunity to pretesting). *Under the assumptions of Proposition 2, for any sequence of events* $\{F_T\}_{T\in\mathbf{N}}$,

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

Proof. See Appendix B.5. □

Proposition 4 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 unconditional test's result is irrelevant for further inference. To the best of our knowledge, only a few known inference procedures with such a property exist (e.g., Hannan and Quinn, 1979; Gagliardini et al., 2019). Like Proposition 3, Proposition 4 is a direct consequence of Proposition 2.

3.3 Proof of Concept

Proposition 2 shows the unconditional test has good asymptotic properties. Monte-Carlo simulations (Table 1 and Appendix C) indicate that the finite sample performance of the test is in line with the asymptotic properties. In the present section, we apply the unconditional test to the excess return of the market over the risk-free asset as a proof of concept on actual financial data. The difference in expected returns between US stock returns and one-month US Treasury bill returns corresponds to the equity premium.

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

$$\text{H}_0 : \forall u \in \mathbf{U}_2,\ \mathbb{E}[u(R_f)] < \mathbb{E}[u(R_M)],$$

where R_f is the one-month Treasury bill gross return and R_M is the CRSP value-weighted market gross return. The excess return $R_M - R_f$ is known as the market factor MKT.

We report results in Table 2.

Table 2: 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

Notes: The columns "Long," "Short," "t_{NW}^{L-S}" and "P-value" correspond to the average return of the long leg, the average return of the short leg, the t-statistic for the null hypothesis "$H_0 : \mathbb{E}(R_S) = \mathbb{E}(R_L)$," and the p-value of the unconditional test, respectively. 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 is not a deviation from the risk-return tradeoff, but 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 an excess return as a deviation from the risk-return tradeoff, 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 function) to be plausible for actual agents in the economy. Mehra and Prescott (1985) also show that a sufficiently high level of risk aversion can make individuals prefer US Treasury bills over US stocks, but they regard it as implausibly high and coin the term "equity premium puzzle."

4 Test Conditional on the Market

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.

4.1 Null Hypothesis Conditional on the Market

The test's null hypothesis 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 excess return, whether every possible risk-averse individual would strictly prefer the long-leg lottery to the short-leg lottery conditional on the market, that is,

$$\mathrm{H}_0 : \forall u \in \mathbf{U}_2,\, \mathbb{E}[u(R_S)|R_M] < \mathbb{E}[u(R_L)|R_M], \tag{17}$$

where R_M denotes the market return.

The primary motivation for the null hypothesis (17) relative to the null hypothesis (6) of the unconditional test is the practice of controlling for the market through a regression with the market (excess) returns as an explanatory variable. In this way, practitioners control for affine functions of the market return. The test conditional on the market does not only control for affine functions of market returns but for all measurable functions of market returns because Lettau et al. (2014), among others, highlight the importance of nonlinearities. Moreover, whether we use market returns or excess returns should be immaterial: 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 (17) 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 B.1.2. □

Lemma 2 is the conditional counterpart of Lemma 1. Similarly to Lemma 1 for the null hypothesis (6), Lemma 2 implies the invariance of the null hypothesis (17) 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 (17) to the data.

4.2 Test Statistic Conditional on the Market

By Lemma 2, the hypothesis (17) is equivalent to the null hypothesis

$$\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[,\ F^{(2)}_{L|M}(z|.) - F^{(2)}_{S|M}(z|.) < 0, \tag{18}$$

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 3, 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 (18) in terms of the concavity of the second-order antiderivative of the difference of integrated conditional CDFs.

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 (18) is equivalent to the null hypotheses

$$\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[, \int_{-\infty}^{.} [F^{(2)}_{L|M}(z|\dot{x}) - F^{(2)}_{S|M}(z|\dot{x})] f_X(\dot{x}) \mathrm{d}\dot{x} = F^{(2)}_{L,M}(z,.) - F^{(2)}_{S,M}(z,.) \text{ strictly decreasing}$$
$$\mathrm{H}_0 : \forall z \in]\underline{u}, \infty[, C^{(2)}(z,.) \text{ is strictly concave}, \tag{19}$$

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 B.7,*

(i) if the null hypothesis (17) *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 (17) *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* $\lim_{T\to\infty} \frac{b_T}{T} = 0$.

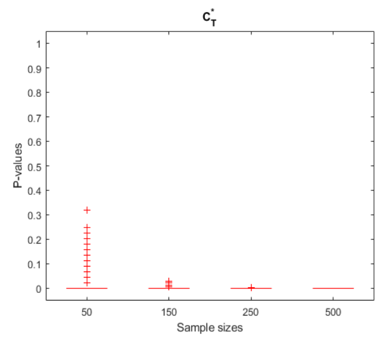
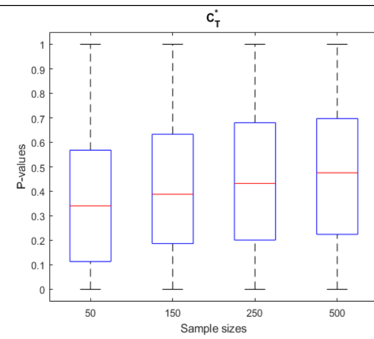
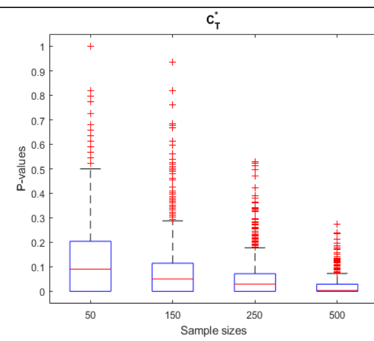
Proof. See Appendix B.7. □

Proposition 5 shows the test conditional on the market is valid and consistent. In Table 3, Monte-Carlo simulation results support Proposition 5. When the null hypothesis (17) is wrong, p-values converge to zero as the sample size increases. When the null hypothesis (17) 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.

4.3 Equilibrium Foundations for the Test Conditional on the Market

In the absence of diversification benefits, the equilibrium foundations of the conditional test are similar to the ones of the unconditional test. The only difference is that investors'

Table 3: Performance of conditional test in Monte-Carlo simulations

H_0	**DGP**	**Boxplots of p-values**
False	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} 1.015 \\ 1.0078 \end{bmatrix}, \begin{bmatrix} .12^2 & .0051 \\ & .057^2 \end{bmatrix}\right)$	
True	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix} 1.011 \\ 1.010 \end{bmatrix}, \begin{bmatrix} .039^2 & .0012 \\ & .057^2 \end{bmatrix}\right)$	
False	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} x + \begin{bmatrix} z_L \\ z_S \end{bmatrix}$ where $\begin{cases} z_L \overset{IID}{\hookrightarrow} 1 + \mathrm{t}(4) \\ z_S \overset{IID}{\hookrightarrow} \mathcal{N}(1,1) \\ \mathbb{C}\mathrm{or}(z_S, z_L) = .7 \end{cases}$	

Notes: 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.

preferences correspond to an expected utility under the distribution conditional on the market.

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

Proposition 6 (Equilibrium foundation for test conditional on market)**.** *Let $R_{M,t}$ denote the market return at time t. Under Assumptions 1, for all $u \in \mathbf{U}_2$ s.t. u is strictly increasing and twice continuously differentiable on $[\underline{u}, \overline{u}]$, which includes the support of $R_{M,t}$ and the returns $R_{S,t}$ and $R_{L,t}$, then, up to approximation errors, the null hypothesis "$\mathrm{H}_0 : \forall u \in \mathbf{U}_2, \mathbb{E}[u(R_{S,t})|R_{M,t}] < \mathbb{E}[u(R_{L,t})|R_{M,t}]$" implies the expected return of the factor exceeds its risk compensation, that is,*

$$-\frac{1}{\mathbb{E}[u'(C_t)]}\mathbb{C}\mathrm{ov}(u'(C_t), R_{L,t} - R_{S,t}) < \mathbb{E}(R_{L,t} - R_{S,t}).$$

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

Proposition 6 shows that, up to approximation errors, strict preference for the long leg conditional on the market is a sufficient condition for a deviation from the risk-return tradeoff. The assumptions of Proposition 6 are similar to the assumptions of Proposition 1.

5 Deviation from the risk-return tradeoff or possible risk factor?

In this section, we apply our tests to assess whether risk can explain long-minus-short portfolios based on characteristic-sorted portfolios. These long-minus-short portfolios correspond to the factors routinely used to compute firms' cost of capital and portfolios' required expected return. Over the last decades, the literature has identified hundreds of these factors called the factor "zoo" (Cochrane, 2011).[10] However, the economic content of

[10]In the following, we do *not* use the term "factor" as a shorthand for "risk factor." A factor can be a deviation from the risk-return tradeoff or an excess return that risk can explain. When we use the term

factors is an open question. Factor returns might be a compensation for risk as frictionless rational asset pricing theory asserts, but they may also arise because of behavioral biases, institutional, informational, and many other frictions.[11] In Section 2.1, we produced a simple example of such a friction. Hereafter, we say that a factor is a possible risk factor if risk can explain its expected return, which is the difference in expected return between the long and short legs. Otherwise, we call it a deviation from the risk-return tradeoff.

We start by describing the dataset. Then, we apply the tests to the widely-used Fama and French 4 factors plus momentum (FF4+MOM). Finally, we provide an overview of the test results for a standard dataset of more than 200 possible risk factors.

5.1 Data

Data for the Fama and French factors and momentum, FF4+MOM, are from Kenneth French's website. The frequency is monthly. 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 stocks with Low, Medium, and High characteristic value. 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). In contrast, 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 the CRSP value-weighted index for the aggregate market and the one-month Treasury Bill for the risk-free rate.

A long sample is available for BM and MOM, starting from July 1926 (BM) or January 1927 (MOM). Data for the market and the Treasury bill yield are also available starting from July 1926. For OP and INV, data start only from July 1963. For this reason, we report the findings for BM, MOM, and the market MKT for the full sample period and

"factor," we have variables in mind that help predict returns in the cross-section without taking a stance on the validity of a factor model.

[11] See e.g., Berk et al., 1999; Gomes et al., 2003; Cooper, 2006 for risk compensation, Bondt and Thaler, 1985; Jegadeesh and Titman, 1993; Bouchaud et al., 2019 for biases, Gromb and Vayanos, 2010 for institutional frictions, and Cohen et al., 2012 for informational frictions.

for a restricted period starting in July 1963. The samples for the FF4+MOM factors end in October 2021.

Moreover, we use data for 205 potential risk factors from Chen and Zimmermann (2022). Stocks are sorted into quantile portfolios, where the number of quantiles depends on data availability for the respective characteristic. We use the lowest and highest quantiles and retain the quantile with the low average return over the sample period as the short leg. We discuss evidence for the original samples of the published papers, the post-publication, and the full samples. The data end in December 2020.

5.2 Unconditional Test Applied to FF4+MOM Factors

The FF4+MOM factors are widely assumed to be risk factors and used to adjust for risk in practice and academia. We apply our unconditional test to these factors to assess whether risk can explain their expected return. We report the results in Table 4.

Table 4: Unconditional test applied to FF4+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

Notes: The columns "Long," "Short," "t_{NW}^{L-S}" and "P-value" correspond to the average return of the long leg, the average return of the short leg, the t-statistic for the null hypothesis "$H_0 : \mathbb{E}(R_S) = \mathbb{E}(R_L)$," and the p-value of the unconditional test, respectively. 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.

Setting aside the Market factor, only Size has a p-value below standard thresholds. The result is robust to different methods for constructing Size. A first potential explanation is the lack of significance of the expected return of Size: The t-statistic of the long-minus-short portfolio t_{NW}^{L-S} is slightly below 1.96, suggesting Size might not be a factor after all, and thus neither a deviation from the risk-return tradeoff nor a risk factor. A second potential explanation is that risk alone explains Size. 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 (see 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 supports Berk (1995), who put forward a risk-based explanation for Size.

Regarding BM, INV, OP, and MOM, we cannot reject the null hypothesis for the sample period starting in July 1963. Similar results hold even if we exclude 2020 and 2021. For MOM, the spread between the short and the long legs is bigger than 7% on an annual basis and, hence, close to the equity premium. Whereas a high risk aversion can explain the equity premium, it cannot explain the MOM factor. This result lends support to behavioral explanations for MOM (e.g., Jegadeesh and Titman, 1993; Hong and Stein, 1999). 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 deviations from the risk-return tradeoff through the lens of our test, in line with findings in Bouchaud et al. (2019).

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

5.3 Test Conditional on Market applied to FF4+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 FF4+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

Notes: The columns "Long," "Short," "t_{NW}^{L-S}" and "P-value" 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 conditional test, respectively. 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 a deviation from the risk-return tradeoff. Whereas the p-values drop for the other characteristics, BM, OP and MOM still appear as deviations from the risk-return tradeoff. 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 (compare Proposition 2i to Proposition 5ii). A second possible explanation is the commonality between the market and the legs of different factors.

5.4 A Bird View on the factor Zoo

Beyond the widely-used FF4+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 Online Appendix. Here, we only provide an overview of the main results. We use 5% as the threshold above which we cannot reject the null hypothesis. In the table below, we report the proportions of potential factors that appear as deviations from the risk-return tradeoff.

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

Notes: The data base correspond to Chen and Zimmermann (2022) dataset of 205 potential factors. The frequency of the data is monthly.

The first result is that a majority of the 205 potential factors appear to be deviations from the risk-return tradeoff in the original sample of the published papers and the full sample. For both tests, we find more than 70% appear as deviations from the risk-return tradeoff. Because the existence of a factor is a necessary condition for a deviation from the risk-return tradeoff, this result lends support to Chen and Zimmermann (2020); Chen (2021a,b); Jensen et al. (2022), who find that most factors can be replicated in the original sample. Remember the unconditional test is immune to the multiple hypothesis problem and the pretesting problem, making the results of this literature even stronger.

The second result is a dramatic change in the proportion of deviations from the risk-return tradeoff between the original and post-publication samples: The proportion drops from about 90% to about 35% for both tests. Two potential explanations exist for this drop: (i) Many deviations from the risk-return tradeoff became risk factors after publication; or (ii) The phenomenon of "anomaly elimination" occurred, that is, many deviations from the risk-return tradeoff disappeared because their expected returns shrank to zero. Table 7 supports the second explanation. It displays the proportion of apparent deviations from the risk-return tradeoff 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 deviations from the risk-return tradeoff 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 a factor, some investors trade on it, so its expected return decreases after a temporary increase (Pénasse, 2022).

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

Notes: We compute the displayed proportions as follows. (i) We keep from the Chen and Zimmermann (2022) dataset 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 remaining factors. For simplicity, potential pretesting problems are ignored. The frequency of the data is monthly.

The third and main result is that a clear majority of factors appears to be deviations from the risk-return tradeoff in all samples. Overall, risk cannot explain more than 80% of factors in the original sample, the post-publication sample, and the full sample (see Table 7). In Table 6, the proportions are lower than in Table 7 because some potential factors do not have significantly positive expected returns and thus are not factors in the first place. This third result generalizes the results for the FF4+MOM factors to most of the factors documented in the literature. This generalization is not surprising because theory and empirical evidence indicate strong commonality across factors (e.g., Reisman, 1992; Bryzgalova et al., 2020) and given the literature stressing the role of behavioral biases (e.g., Bouchaud et al., 2019; Barberis et al., 2021) and frictions for factors (e.g., Weber, 2018; Bowles et al., 2022; Kim et al., 2022; Muravyev et al., 2023). It suggests that the factor "zoo" is mainly an "anomaly" zoo in the sense that risk alone cannot explain it.

More generally, the equilibrium foundations of the tests indicate that this last and main empirical result is consistent with the literature highlighting the importance of market frictions and behavioral biases for differences in cross-sectional returns.[12] Recently, Korsaye et al. (2021), Dello-Preite et al. (2022) and Cong et al. (2022) find non-systematic variables are helpful to explain cross-sectional returns in line with market frictions, whereas Lopez-Lira and Roussanov (2023) find that latent common factors have limited explanatory power for stock returns. Chinco et al. (2022), instead, survey investors and find they do not make investment decisions based on the covariance between asset returns and consumption growth, making it less likely that this covariance, which captures *non-diversified* risk, explains cross-sectional returns. Our empirical results also extend a long

[12] Luttmer (1996), He and Modest (1995), Guvenen (2009) and Czellar et al. (2022) show market frictions can even explain the equity premium puzzle.

literature on "low-risk anomalies" (e.g., Haugen and Heins, 1975; Frazzini and Pedersen, 2014; Kapadia et al., 2019; Schneider et al., 2020). To our knowledge, the present paper is the first to show that risk cannot explain most of the factor "zoo" without relying on a factor model with a specific dependence structure for the errors, nor on a particular measure of risk.

6 Conclusion

No-arbitrage is the central pillar of relative valuation. No-arbitrage mainly follows from preference monotonicity: If individuals prefer more to less, nobody leaves money on the table. In the present paper, we add risk aversion to preference monotonicity. It allows us to go beyond arbitrage. It allows us to investigate deviations from the risk-return tradeoff without assuming a specific model of the economy.

In the present paper, we (i) provide as a motivation a simple model economy, in which a limit on long positions yields an expected excess return unexplained by risk alone although factors span it; (ii) define deviations from the risk-return tradeoff; (iii) introduce the concept of strong SSD; (iv) show that if the long leg of an excess return strongly SSD dominates its short leg, it should be a deviation from the risk-return tradeoff; (v) propose two tests based on strong SSD; (vi) verify the performance of the tests numerically, mathematically, and empirically; and (vii) apply the two tests to more than 200 excess returns of characteristic-sorted portfolios.

We propose and use two tests because they rely on different assumptions. Despite their differences, both tests classify most expected excess returns of characteristic-sorted portfolios as deviations from the risk-return tradeoff. This empirical result is consistent with a prominent role of frictions and behavioral biases in financial markets. It might be unexpected because strong SSD sets a high threshold for deviations from the risk-return tradeoff. Strong SSD requires strict preference even for implausibly high levels of risk aversion.

The proposed tests should be helpful to detect "abnormal" returns in many situations, especially given that the prevailing method equates "abnormal" returns to the alphas of regressions on a preferred factor model, that is, on some excess returns of characteristic-sorted portfolios. In this way, both tests can guide better investment decisions and capital allocation.

The proposed tests do not only help to detect deviations from the risk-return tradeoff. They also provide some guidance on which types of models can explain these deviations. The tests and their theoretical foundations barely impose any restriction on distributions of returns nor production, etc. Thus, explanations of the deviation from the risk-return tradeoff call for models in which risk-averse individuals do not buy factors they value higher than their market price. In particular, trading frictions, financial intermediaries, and behavioral biases are possible explanations for the detected deviations from the risk-return tradeoff.

References

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

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

Arvanitis, S., O. Scaillet, and N. Topaloglou (2022, Dec). Spanning analysis of stock market anomalies under prospect stochastic dominance. Available at SSRN: https://ssrn.com/abstract=3569755.

Avramov, D., S. Cheng, A. Lioui, and A. Tarelli (2022). Sustainable investing with esg rating uncertainty. *Journal of Financial Economics 145*(2), 642–664.

Avramov, D., A. Lioui, Y. Liu, and A. Tarelli (2023, December). Dynamic esg equilibrium. Available at SSRN: https://ssrn.com/abstract=3935174.

Backus, D. K., B. R. Routledge, and S. E. Zin (2004). Exotic preferences for macroeconomists. *NBER Macroeconomics Annual 19*, 319–390.

Baker, S. D., B. Hollifield, and E. Osambela (2022). Asset prices and portfolios with externalities. *Review of Finance 26*(6), 1433–1468.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance 59*(4), 1481–1509.

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

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics 121*(3), 823–866.

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

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

Bernardo, A. E. and O. Ledoit (2000). Gain, loss, and asset pricing. *Journal of Political Economy 108*(1), 144–172.

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

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *The Journal of Finance 74*(6), 2839–2874.

Bouchaud, J., P. Krueger, A. Landier, and D. Thesmar (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance 74*(2), 639–674.

Bowles, B., A. V. Reed, M. C. Ringgenberg, and J. R. Thornock (2022). Anomaly time. Available at SSRN: https://ssrn.com/abstract=3069026.

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

Campbell, J. Y. and J. H. Cochrane (1999, Apr.). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*(2), 205–251.

Chalamandaris, G., K. Pukthuanthong, and N. Topaloglou (2021, Feb.). Are stock-market anomalies anomalous after all? Available at SSRN: https://ssrn.com/abstract=3752177.

Chen, A. Y. (2021a). The limits of p-hacking: Some thought experiments. *The Journal of Finance 76*(5), 2447–2480.

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

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

Chen, A. Y. and T. Zimmermann (2022). Open source cross-sectional asset pricing. *Critical Finance Review 27*(2), 207–264.

Chinco, A., S. M. Hartzmark, and A. B. Sussman (2022). A new test of risk factor relevance. *The Journal of Finance 77*(4), 2183–2238.

Cochrane, J. (2005). *Asset pricing* (2nd ed.). Princeton University Press.

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

Cochrane, J. H. and J. Saa-Requejo (2000). Gain, loss, and asset pricing. *Journal of Political Economy 108*(1), 144–172.

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

Coibon, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigiditiies? *Journal of Political Economy 120*(1), 116–159.

Coibon, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(3), 2644–2678.

Cong, L. W., G. Feng, J. He, and J. Li (2022, September). Uncommon factors for Bayesian asset clusters. Available at SSRN: https://ssrn.com/abstract=4219905.

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

Czellar, V., R. Garcia, and F. Le Grand (2022, Feb). Uncovering asset market participation from household consumption and income. Available at SSRN: https://ssrn.com/abstract=3975672.

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

Delgado, M. A. and J. C. Escanciano (2013). Conditional stochastic dominance testing. *Journal of Business & Economic Statistics 31*(1), 16–28.

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze (2022, June). What is missing in asset-pricing factor models?

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

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

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

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

Frazzini, A. and L. H. Pedersen (2014). Betting against beta. *Journal of Financial Economics 111*(1), 1–25.

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

Gomes, J., L. Kogan, and L. Zhang (2003). Equilibrium cross section of returns. *Journal of Political Economy 111*(4).

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

Guvenen, F. (2009). A parsimonious macroeconomic model for asset pricing. *Econometrica 77*(6), 1711–1750.

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

Hansen, L. P. (2013, Dec). Uncertainty outside and inside economic models. Nobel Prize Lecture.

Hansen, L. P. and R. Jagannathan (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy 99*(2), 225–262.

Hansen, L. P. and K. J. Singleton (1982, Sep). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica 50*(5), 1269–1286.

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

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

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

Haugen, R. A. and A. J. Heins (1975). Risk and the rate of return on financial assets: Some old wine in new bottles. *The Journal of Financial and Quantitative Analysis 10*(5), 775–784.

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

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

Hodder, J. E., J. C. Jackwerth, and O. Kolokolova (2015). Improved portfolio choice using second-order stochastic dominance. *Review of Finance 19*(4), 1623âĂŞ1647.

Hong, H. and J. C. Stein (1999, Dec.). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance 54*(6), 2143–2184.

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

Jensen, T. I., B. Kelly, and L. H. Pedersen (2022). Is there a replication crisis in finance? Forthcoming in The Journal of Finance.

Jouini, E. and H. Kallal (1995). Martingales and arbitrage in securities markets with transaction costs. *Journal of Economic Theory 66*(1), 178–197.

Jouini, E. and H. Kallal (1999, July). Viability and equilibrium in securities markets with frictions. *Mathematical Finance 9*(3), 275–292.

Kapadia, N., B. B. Ostdiek, J. P. Weston, and M. Zekhnini (2019). Getting paid to hedge: Why don't investors pay a premium to hedge downturns? *Journal of Financial and Quantitative Analysis 54*(3), 1157–1192.

Kim, Y. H., Z. Ivkovich, and D. Muravyev (2022). Causal effect of information costs on asset pricing anomalies. Available at SSRN https://ssrn.com/abstract=3069026.

Korsaye, S. A. (2024). Investor beliefs and trading frictions. Available at SSRN: https://ssrn.com/abstract=4277979.

Korsaye, S. A., A. Quaini, and F. Trojani (2021). Smart stochastic discount factors. Available at SSRN: https://ssrn.com/abstract=3878300.

Krishnamurthy, A. and A. Vissing-Jorgensen (2012). The aggregate demand for treasury debt. *Journal of Political Economy 120*(2), 233–267.

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

Lettau, M., M. Maggiori, and M. Weber (2014). Conditional risk premia in currency markets and other asset classes. *Journal of Financial Economics 114*(2), 197–225.

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

Lopez-Lira, A. and N. L. Roussanov (2023). Do common factors really explain the cross-section of stock returns? Available at SSRN at https://ssrn.com/abstract=3628120.

Luttmer, E. G. (1992). Asset pricing in economies with frictions. Ph.D. Thesis, Department of Economics, The University of Chicago.

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

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

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

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

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance 42*(4), 483–510.

Morse, A. P. (1939). The behaviour of a function on its critical set. *Annals of Mathematics 40*(1), 62–70.

Muravyev, D., N. D. Pearson, and J. M. Pollet (2023). Anomalies and their short-sale costs. Available at SSRN: https://ssrn.com/abstract=4266059.

Pastor, L., R. F. Stambaugh, and L. A. Taylor (2021). Sustainable investing in equilibrium. *Journal of Financial Economics 142*(2), 550–571.

Pénasse, J. (2022). Understanding alpha decay. *Management Science 68*(5).

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

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

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics 22*(1), 117–131.

Roll, R. (1980). Orthogonal portfolios. *Journal of Financial and Quantitative Analysis 15*(5), 1005–1023.

Sandulescu, M. (2022). How integrated are corporate bond and stock markets? Ross School of Business Paper Forthcoming, Swiss Finance Institute Research Paper No. 20-09.

Sard, A. (1942). The measure of the critical values of differentiable maps. *Bulletin of the American Mathematical Society 48*(12), 883–890.

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

Schneider, P., C. Wagner, and J. Zechner (2020). Low-risk anomalies? *The Journal of Finance 75*(5), 2673–2718.

Shleifer, A. and R. W. Vishny (1997, Mar.). The limits of arbitrage. *The Journal of Finance 52*(1), 35–55.

Summers, L. (1985). On economics and finance. *The Journal of Finance 40*(3), 633–635.

Weber, M. (2018). Cash flow duration and the term structure of equity returns. *Journal of Financial Economics 128*(3), 486–503.

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

Zerbib, O. D. (2022). A sustainable capital asset pricing model (s-capm): Evidence from environmental integration and sin stock exclusion. *Review of Finance 26*(6), 1345–1388.

Note

The present paper supersedes the working paper called "Anomaly or Possible Risk Factor? Simple-To-Use Tests."

ONLINE APPENDIX TO:

Beyond Arbitrage: Deviations from the Risk-Return Tradeoff

Benjamin Holcblat, Abraham Lioui and Michael Weber

A Augmented-CAPM

Consider a representative investor who maximizes her expected utility subject to constraints on long positions. More specifically, the representative investor maximizes the following mean-variance problem

$$\begin{cases} \max_{w\in\mathbf{R}^K} w'(\mu - R_0\mathbf{1}) - \frac{\lambda}{2}w'\Sigma w \\ w_k \leqslant \bar{M}_k, \quad \text{for } k = 1, 2, \dots, K, \end{cases} \tag{A.1}$$

where vector $R := (R_1\ R_2\ \dots\ R_K)'$ denotes the vector of gross returns of risky assets, $\mu := \mathbb{E}(R)$ the expected gross return of risky assets, $\Sigma := \mathbb{V}(R)$ the variance-covariance matrix of risky assets' gross returns, R_0 the gross return of the risk-free rate, w_k the fraction of initial wealth invested in the asset k, $w := (w_1\ w_2\ \dots w_K)'$, $\mathbf{1} := (1\ 1\ \dots 1)'$ a $K\times 1$ vector of ones, $\bar{M} := (\bar{M}_1\ \bar{M}_2\ \dots\ \bar{M}_K)'$ the vector of upper bounds on long positions, and $\lambda > 0$ captures risk aversion. The constraint on long position $\bar{M}_k$, for example, can be due to regulation (e.g., risk management). The existence of a solution to the mean-variance problem (A.1) is a sufficient condition for the existence of general equilibrium economy with a representative investor maximizing the mean-variance problem (A.1). See Luttmer (1996), He and Modest (1995) for prominent examples of representative agents in economies with frictions, and Luttmer (1992) for aggregation results.

Solving (A.1) is equivalent to maximizing the Lagrangian

$$\max_{w\in\mathbf{R}^K} w'(\mu - R_0\mathbf{1}) - \frac{\lambda}{2}w'\Sigma w - \delta'(w - \bar{M}),$$

where δ is the vector of Lagrange multipliers for the constraints on long positions. Thus, the first order condition is

$$\mu - R_0\mathbf{1} - \lambda\Sigma w^* - \delta = 0, \tag{A.2}$$

resulting in optimal portfolio weights

$$w^* = \frac{\lambda_\tau}{\lambda} w_\tau - \frac{\lambda_\delta}{\lambda} w_\delta, \tag{A.3}$$

where $\lambda_\tau := \mathbf{1}'\Sigma^{-1}(\mu - R_0\mathbf{1})$, $\lambda_\delta := \mathbf{1}'\Sigma^{-1}\delta$, $w_\tau := \frac{\Sigma^{-1}(\mu - R_0\mathbf{1})}{\mathbf{1}'\Sigma^{-1}(\mu - R_0\mathbf{1})}$, $w_\delta := \frac{\Sigma^{-1}\delta}{\mathbf{1}'\Sigma^{-1}\delta}$. In a frictionless economy, the portfolio w_τ is the standard tangency portfolio. The portfolio w_δ reflects the distortion to the optimal demand for risky assets due to the friction.

The first order condition (A.2) is also equivalent to

$$\mu - R_0\mathbf{1} = \lambda\Sigma w^* + \delta = \lambda\Sigma w^* + \lambda_\delta \Sigma w_\delta, \tag{A.4}$$

so the expected return of the optimal portfolio is

$$(w^*)'(\mu - R_0\mathbf{1}) = \lambda (w^*)' \Sigma w^* + \lambda_\delta (w^*)' \Sigma w_\delta. \tag{A.5}$$

In a frictionless economy, the expected return of an efficient portfolio is proportional to its level of risk $(w^*)' \Sigma w^*$. With a constraint on long positions, a correction exists proportional to the covariance of the efficient portfolio with the friction portfolio w_δ.

According to the optimality condition (A.4), the expected excess return of any portfolio w_p is

$$w_p'(\mu - R_0\mathbf{1}) = \lambda w_p' \Sigma w^* + \lambda_\delta w_p' \Sigma w_\delta. \tag{A.6}$$

The factor model (A.6) consists of two factors, $(w^*)'R$ and $w_\delta' R$, where $(w^*)'R$ corresponds to a market-type factor. In the standard approach, which assumes factors are necessarily risk factors, the lambdas λ and λ_δ would be called the prices of risk of the factors, $(w^*)'R$ and $w_\delta' R$, respectively. While the factor $(w^*)'R$ is a risk factor, the factor $w_\delta' R$ is *not* a risk factor. It is due to the constraints on long positions. In the absence of constraints, the Lagrange multiplier δ and the lambda λ_δ are zero, and thus the only factor is the market-type factor $(w^*)'R$, that is, the factor $w_\delta' R$ does not exist.

We can also derive an augmented-CAPM representation of the factor model (A.6). In equilibrium, the representative individual holds the market portfolio with weigths $w_M := \frac{w^*}{\mathbf{1}'w^*}$. By the optimality condition (A.4), the expected excess returns of any portfolio w_p and of the market portfolio w_M are

$$\mu_p - R_0 = w_p'(\mu - R_0\mathbf{1}) = \bar{\lambda} w_p' \Sigma w_M + \lambda_\delta w_p' \Sigma w_\delta$$
$$\mu_M - R_0 = w_M'(\mu - R_0\mathbf{1}) = \bar{\lambda} w_M' \Sigma w_M + \lambda_\delta w_M' \Sigma w_\delta,$$

where $\bar{\lambda} := \lambda \mathbf{1}' w^*$. Combination of the last two equalities yields the augmented-CAPM representation

$$\mu_p - R_0 = \beta_{p,M} (\mu_M - R_0) + (\beta_{p,\delta} - \beta_{p,M}\beta_{M,\delta}) \bar{\lambda}_\delta, \tag{A.7}$$

where $\beta_{i,j}$ is the beta of portfolio i relative to portfolio j and $\bar{\lambda}_\delta := \lambda_\delta w_\delta' \Sigma w_\delta$. The expected excess return of any portfolio w_p has two components: A CAPM component $\beta_{p,M} (\mu_M - R_0)$ proportional to the market excess return but also a second component $(\beta_{p,\delta} - \beta_{p,M}\beta_{M,\delta}) \bar{\lambda}_\delta$ related to the exposure to the friction portfolio w_δ. The exposure to the friction portfolio accounts for the fact that in equilibrium the market will also be impacted by its exposure to w_δ. To avoid double counting, the exposure of the portfolio w_p to the friction portfolio w_δ is its beta relative to this portfolio $\beta_{p,\delta}$ net of the compensation for the presence of the second factor in the market portfolio $\beta_{p,M}\beta_{M,\delta}$.

B Proofs

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

B.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 $\breve{u} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u(x) = 0\}\right\}$, and denote its left derivative and right derivative at x with $u'_-(x)$ and $u'_+(x)$, respectively.*[13] *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 $u'_-(\breve{u}) \in \mathbf{R} \setminus \{0\}$ with $\breve{u} \neq \underline{u}$, $\mathbb{E}[u(R_S)] < \mathbb{E}[u(R_L)]$.*
- *(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 $F_L^{(2)}(z) := \int_{\underline{u}}^z F_L(y)\mathrm{d}y$.*

Theorem A.1 is the strong counterpart of the well-known Hardy-Littlewood et. al. theorem (Hardy et al., 1929, 1934; Blackwell, 1951; Sherman, 1951; Cartier et al., 1964;

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

Strassen, 1965), which has been popularized in economics by Rothschild and Stiglitz (1970),

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

B.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 $\breve{u}_{W_1} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u_{W_1}(x) = 0\}\right\}$, 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 $\mathbb{E}|u'_{W_1,-}(\breve{u}_{W_1})| < \infty$ with $u'_{W_1,-}(\breve{u}_{W_1}) \neq 0$ and $\breve{u}_{W_1} \neq \underline{u}$ 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 $u'_-(\breve{u}) \in \mathbf{R} \setminus \{0\}$ with $\breve{u} \neq \underline{u}$, $\mathbb{E}[u(R_S)|W_1] < \mathbb{E}[u(R_L)|W_1]$ 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 $\breve{u}_{W_1}$ is well-defined and measurable.

Lemma A.1 (Existence and $\sigma(W_1)$-measurability of $\breve{u}_{W_1}$)**.** *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 $w_1 \mapsto \breve{u}_{w_1}$ with values in $[\underline{u}, \overline{u}]$ s.t. $\breve{u}_{w_1} := \min\left\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\}\right\}$, 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 $w_1 \in \mathbf{R}$, $\{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\} = \emptyset$ *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 $A_{w_1} := \{z \in [\underline{u}, \overline{u}] \text{ s.t., } \forall x \in [z, \overline{u}],\ u_{w_1}(x) = 0\}$, 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 $\check{u}_{w_1} := \min\left\{\overline{u}, \inf A_{w_1}\right\} = \overline{u}$.

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, 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, 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, 1994, p. 557). By definition of the lower inverse and of the correspondence $\psi_{\overline{u}}$, for all open subset O of $[\underline{u}, \overline{u}]$,

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

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, 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, 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, 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),

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

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 $\breve{u} \neq 0$, careful modifications of the proof techniques used in the mathematical literature (e.g., Föllmer and Schied, 2002, for a textbook presentation), and new technical lemmas.

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

(ibis) $\Rightarrow$ *(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, $\breve{u} = z$ if $z \in]\underline{u}, \overline{u}]$, and $\breve{u} = \overline{u}$ otherwise, so $u'_-(\breve{u}) =$

$1 \neq 0$ and $\breve{u} \neq \underline{u}$. 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) $\Rightarrow$ *(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 $\mathbb{E}|u'_{W_1,-}(\breve{u}_{W_1})| < \infty$ with $u'_{W_1,-}(\breve{u}_{W_1}) \neq 0$ and $\breve{u}_{W_1} \neq \underline{u}$ 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, 2002, Proposition A.4), for all $x \in [\underline{u}, \overline{u}]$, a.s.,

$$
\begin{aligned}
& h_{W_1}(x) \\
&= h_{W_1}(\breve{u}_{W_1}) + \int_{\breve{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}(\breve{u}_{W_1}) - \int_{x}^{\breve{u}_{W_1}} \overline{h}'_{W_1,-}(y)\mathrm{d}y\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} \\
&\qquad \text{because, by definition of } \overline{h}'_{W_1,-}(.) \text{ and } \breve{u}_{W_1},\ \forall y \in]\breve{u}_{W_1}, \overline{u}],\ \overline{h}'_{W_1,-}(y) = 0; \\
&\overset{(a)}{=} h_{W_1}(\breve{u}_{W_1}) - \int_{x}^{\breve{u}_{W_1}} [\overline{h}'_{W_1,-}(y) - \overline{h}'_{W_1,-}(\breve{u}_{W_1}) + \overline{h}'_{W_1,-}(\breve{u}_{W_1})]\mathrm{d}y\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} \\
&= h_{W_1}(\breve{u}_{W_1}) - \int_{x}^{\breve{u}_{W_1}} \overline{h}'_{W_1,-}(\breve{u}_{W_1})\mathrm{d}y\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} - \int_{x}^{\breve{u}_{W_1}} [\overline{h}'_{W_1,-}(y) - \overline{h}'_{W_1,-}(\breve{u}_{W_1})]\mathrm{d}y\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} \\
&\overset{(b)}{=} h_{W_1}(\breve{u}_{W_1}) - \overline{h}'_{W_1,-}(\breve{u}_{W_1})(\breve{u}_{W_1} - x)\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} + \int_{x}^{\breve{u}_{W_1}} [\overline{h}'_{W_1,-}(\breve{u}_{W_1}) - \overline{h}'_{W_1,-}(y)]\mathrm{d}y\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} \\
&\overset{(c)}{=} h_{W_1}(\breve{u}_{W_1}) - \overline{h}'_{W_1,-}(\breve{u}_{W_1})(\breve{u}_{W_1} - x)^+ + \int_{x}^{\breve{u}_{W_1}} \int_{y}^{\breve{u}_{W_1}} \gamma_{W_1}(\mathrm{d}z)\mathrm{d}y\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} \text{ where } \gamma_{W_1} \text{ is a random} \\
&\qquad \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}(\breve{u}_{W_1}) - \overline{h}'_{W_1,-}(\breve{u}_{W_1})(\breve{u}_{W_1} - x)^+ + \int_{\underline{u}}^{\breve{u}_{W_1}} \int_{\underline{u}}^{\breve{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\mathrm{d}y\gamma_{W_1}(\mathrm{d}z)\mathbb{1}\{x \leqslant \breve{u}_{W_1}\} \\
&\overset{(e)}{=} h_{W_1}(\breve{u}_{W_1}) - \overline{h}'_{W_1,-}(\breve{u}_{W_1})(\breve{u}_{W_1} - x)^+ + \int_{\underline{u}}^{\breve{u}_{W_1}} (z-x)^+\gamma_{W_1}(\mathrm{d}z) \qquad \text{(A.8)}
\end{aligned}
$$

(a) By assumption, $\mathbb{E}|h'_{W_1,-}(\breve{u}_{W_1})| = \mathbb{E}|u'_{W_1,-}(\breve{u}_{W_1})| < \infty$, so $h'_{W_1,-}(\breve{u}_{W_1})$ is finite a.s.[14] Now, $\overline{h}'_{W_1,-}(.) := h'_{W_1,-}(.)\mathbb{1}_{]\underline{u},\overline{u}]}(.) + h'_{W_1,+}(.)\mathbb{1}_{\{\underline{u}\}}(.) = h'_{W_1,-}(.)$ a.s. because $\breve{u}_{W_1} \neq \underline{u}$ a.s. by assumption. Thus, $\overline{h}'_{W_1,-}(\breve{u}_{W_1})$ is finite a.s. *(b)* Standard algebra yields $\int_x^{\breve{u}_{W_1}} \overline{h}'_{W_1,-}(\breve{u}_{W_1})\mathrm{d}y = \overline{h}'_{W_1,-}(\breve{u}_{W_1}) \int_x^{\breve{u}_{W_1}} \mathrm{d}y = \overline{h}'_{W_1,-}(\breve{u}_{W_1})(\breve{u}_{W_1} - x)$. *(c)* By Lemmas A.2 and A.4 (p. OA.9 & OA.10), there exists a unique σ-finite random Borel measure γ_{W_1} on $[\underline{u}, \breve{u}_{W_1}[$ 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)* $\int_x^{\breve{u}_{W_1}} \int_y^{\breve{u}_{W_1}} \gamma_{W_1}(\mathrm{d}z)\mathrm{d}y = \int_{\underline{u}}^{\breve{u}_{W_1}} \int_{\underline{u}}^{\breve{u}_{W_1}} \mathbb{1}\{y \leqslant z\}\gamma_{W_1}(\mathrm{d}z)\mathbb{1}\{x \leqslant y\}\mathrm{d}y = \int_{\underline{u}}^{\breve{u}_{W_1}} \int_{\underline{u}}^{\breve{u}_{W_1}} \mathbb{1}\{x \leqslant y \leqslant z\}\gamma_{W_1}(\mathrm{d}z)\mathrm{d}y = \int_{\underline{u}}^{\breve{u}_{W_1}} \int_{\underline{u}}^{\breve{u}_{W_1}} \mathbb{1}\{x \leqslant$

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

$y \leqslant z\}\mathrm{d}y\gamma_{W_1}(\mathrm{d}z)$ where the last equality follows from Fubini-Tonelli's theorem (e.g., Kallenberg, 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, 1997, Theorem 6.3-6.4 with equation (6)) and Lemma A.1v on p. OA.4,, a.s.,

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

(a) Show the three terms of equation (A.8) have a finite expectation so their conditional expectation are well-defined (e.g., Kallenberg, 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.10. 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, 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, 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.9. 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, 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) ⇔ *(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.9, 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, 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.10. 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, 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, 1994, Theorem 4.27). Thus, by a standard result about random finite measures (e.g., Kallenberg, 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* $[\underline{u}, \overline{u}]$,

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

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.* $\check{u}_{W_1} := \min\big\{\overline{u}, \inf\{z \in [\underline{u}, \overline{u}]$ *s.t.*, $\forall x \in [z, \overline{u}],\ u_{W_1}(x) = 0\}\big\}$, *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, 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, 1994, Theorem 11.13). *(d)* By assumption, $\mathbb{E}|h_{W_1}(\underline{u})| < \infty$, $\mathbb{E}|h'_{W_1,+}(\underline{u})| < \infty$ and $\mathbb{E}|h'_{W_1,-}(\overline{u})| < \infty$. □

B.2 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.

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, $C_{\dot{t}} + \epsilon(R_{L,\dot{t}} - R_{S,\dot{t}})$ is in any arbitrary small neighborhood of $C_{\tilde{t}}$ so the local optimality of $C_{0:T}$ implies

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

(a) Divide both sides by $1/(\beta\epsilon)$, and multiply the numerator and the denominator of the fraction with $(R_{L,\tilde{t}} - R_{S,\tilde{t}})$. *(b)* By Assumption 1(a), for ϵ small enough $C_{\tilde{t}} + \epsilon(R_{L,\tilde{t}} - R_{S,\tilde{t}})$ 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 $[\underline{u}, \overline{u}]$, $\epsilon \mapsto \frac{[u(C_{\tilde{t}}) - u(C_{\tilde{t}} + \epsilon(R_{L,\tilde{t}} - R_{S,\tilde{t}}))]}{\epsilon(R_{L,\tilde{t}} - R_{S,\tilde{t}})}$ 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 $\tilde{C}_{\tilde{t}} = C_{\tilde{t}} - \epsilon(R_{L,\tilde{t}} - R_{S,\tilde{t}})$ implies $\mathbb{E}[u'(C_{\tilde{t}})(R_{L,\tilde{t}} - R_{S,\tilde{t}})] \leqslant 0$. Thus, the result follows.

(ii) Standard calculations yield

$$\begin{aligned} & \mathbb{E}[u'(C_{\tilde{t}})(R_{L,\tilde{t}} - R_{S,\tilde{t}})] = 0 \\ \Leftrightarrow \; & \mathbb{C}\text{ov}(u'(C_{\tilde{t}}), R_{L,\tilde{t}} - R_{S,\tilde{t}}) + \mathbb{E}[u'(C_{\tilde{t}})]\mathbb{E}(R_{L,\tilde{t}} - R_{S,\tilde{t}}) = 0 \\ \Leftrightarrow \; & \mathbb{E}(R_{L,\tilde{t}} - R_{S,\tilde{t}}) = -\frac{\mathbb{C}\text{ov}(u'(C_{\tilde{t}}), R_{L,\tilde{t}} - R_{S,\tilde{t}})}{\mathbb{E}[u'(C_{\tilde{t}})]} \end{aligned}$$

□

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 follow this other way because it would require more assumptions: It would at least require each leg of the factor to be freely tradable, separately. ◇

B.3 Proof of Proposition 1

The proof is based on Taylor expansions. The key idea is to show that the first term that does not cancel out corresponds to $\mathbb{E}[u'(W_1)(R_L - R_S)]$, which determines non-diversified risk. See the derivation of equation (10) in Section 2.3.2.

Proof of Proposition 1. Two first order Taylor expansions of $u(.)$ around W_1 yield, up to approximation error,

$$
\begin{aligned}
&\mathbb{E}[u(W_0R_L) - u(W_0R_S)] \\
&= \mathbb{E}\left[u(W_1) + u'(W_1)\big(W_0R_L - W_1\big) - u(W_1) - u'(W_1)\big(W_0R_S - W_1\big)\right] \\
&= W_0\mathbb{E}\left[u'(W_1)(R_L - R_S)\right],
\end{aligned}
\tag{A.9}
$$

where, by Lemma 1, the null hypothesis (6) implies $0 < \mathbb{E}[u(W_0R_L) - u(W_0R_S)]$.

Thus, up to approximation error, dividing both sides by W_0,

$$
\begin{aligned}
&0 < \mathbb{E}\left[u'(W_1)W_0(R_L - R_S)\right] = W_0\mathbb{C}\mathrm{ov}(u'(W_1), R_L - R_S) + W_0\mathbb{E}[u'(W_1)]\mathbb{E}(R_L - R_S) \\
\Leftrightarrow\ & -\frac{1}{\mathbb{E}[u'(W_1)]}\mathbb{C}\mathrm{ov}(u'(W_1), R_L - R_S) < \mathbb{E}(R_L - R_S).
\end{aligned}
$$

□

Remark 3. A sufficient (but not necessary) condition for the approximation errors to be negligible is $|\mathbb{E}[\int_{W_1}^{W_0R_L}(W_0R_L - x)u''(x)\mathrm{d}x - \int_{W_1}^{W_0R_S}(W_0R_s - x)u''(x)\mathrm{d}x]| < |\mathbb{E}[u(W_0R_L) - u(W_0R_S)]|$. ◇

Remark 4. A side product of the proof is to show that Roll (1977)'s critique, that is unobserved wealth, is of second order for the proposed tests: The wealth term W_0 and $u(W_1)$ cancel out in the Taylor expansions. ◇

Discussion: Taylor approximations and approximation errors

Taylor approximations have been shown to be helpful in many areas, including asset pricing theory (e.g., log linearization such as the Campbell-Shiller decomposition and solution methods to asset pricing models with Epstein-Zin preferences) and empirical works (e.g., inference based on asymptotic approximations). However, because of the potential effect of approximation errors, they should be used with caution.

In the case of the Proof of Proposition 1, there are several reasons why we can argue up to approximation errors for the purpose of the paper. First, the invariance of the null hypothesis (6) under strictly positive affine transformations of lotteries (Lemma 1) allows to arbitrarily recenter the Taylor expansions in order to reduce the magnitude of the higher error terms. For this reason, Proposition 1 can still hold even when the approximation errors of the corresponding Taylor approximation is arbitrarily big for some utility functions.[15] Second, the Taylor expansion are around the random terminal

[15]We thank Shri Santosh for providing in his discussion a simple example where the approximation error can be made arbitrarily big for a specific utility function, while Proposition 1 still holds.

wealth W_1, so the random changes of W_1 allow to account for the curvature of the utility function $u(.)$. In particular, it allows to account for its concavity, which embodies risk aversion. In contrast, if Taylor approximations were around the fixed value $\mathbb{E}(W_1)$, then risk aversion would be neutralised. Finally, note that Taylor expansions in the Proof of Proposition 1 are similar to the Taylor expansion behind the portfolio optimality condition (7): One more marginal unit of the costless portfolio $R_L - R_S$ yields a new utility $\mathbb{E}[u(W_1 + R_L - R_S)] \simeq \mathbb{E}[u(W_1)] + \mathbb{E}[u'(W_1)(R_L - R_S)]$, so the utility change $\mathbb{E}[u(W_1 + R_L - R_S)] - \mathbb{E}[u(W_1)] \simeq \mathbb{E}[u'(W_1)(R_L - R_S)]$ is zero at the optimum. This is the mathematical logic behind the portfolio optimality condition (7).

B.4 Proposition 2

B.4.1 Core of the proof

The mathematics are standard. We just need (i) the test statistic (16) 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 which the null hypothesis (15) 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 (16) yields

$$\begin{aligned}
\sqrt{T}\text{KS}_T(z) &:= \sqrt{T}\{\hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z)\} \\
&= \sqrt{T}\left\{\hat{F}_L^{(2)}(z) - F_L^{(2)}(z) - [\hat{F}_{L\wedge S}^{(2)}(z) - F_{L\wedge S}^{(2)}(z)] + F_L^{(2)}(z) - F_{L\wedge S}^{(2)}(z)\right\} \\
&= \sqrt{T}[\hat{F}_L^{(2)}(z) - F_L^{(2)}(z)] - \sqrt{T}[\hat{F}_{L\wedge S}^{(2)}(z) - F_{L\wedge S}^{(2)}(z)] \\
&\quad + \sqrt{T}[F_L^{(2)}(z) - F_S^{(2)}(z)]\mathbb{1}_{\mathbf{A}}(z), \qquad \text{(A.10)}
\end{aligned}$$

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 (15), 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,

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

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 (15) 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 (16), which maximizes (A.10), goes to infinity, and thus becomes bigger than any threshold $\hat{c}_{1-\alpha}$.

B.4.2 Assumptions and intermediary results

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.[16]

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

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

(ii) if H_0 does not hold, then as $T \to \infty$, $\mathrm{KS}_T^* = \sup_{z\in\mathbf{I}_T}\Big|\hat{F}_L^{(2)}(z) - \hat{F}_{L\wedge S}^{(2)}(z)\Big|$ 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 of the core of the proof. □

[16] As in the literature (e.g., Politis et al., 1999), we still state both assumptions to simplify the presentation.

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* $T \to \infty$, $\hat{G}^0_{T,b_T}(x) \to \mathbb{1}(\overline{\mathrm{KS}}^* \leqslant x)$ *where* $\hat{G}^0_{T,b_T}(x) := \frac{1}{T-b_T+1} \sum_{i=1}^{T-b_T+1} \mathbb{1}(\mathrm{KS}^*_{T,i} \leqslant x)$*; and*

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

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 2. 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, $\sqrt{T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L}(z) - \hat{F}^{(2)}_{L\wedge S}(z)| = 0$. 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., $\sqrt{b_T}\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L,b_T,i}(z) - \hat{F}^{(2)}_{L\wedge S,b_T,i}(z)| = \sqrt{b_T}[\sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L,b_T,i}(z) - \hat{F}^{(2)}_{L\wedge S,b_T,i}(z)| - \sup_{z\in[\underline{u},\overline{u}]}|\hat{F}^{(2)}_{L}(z) - \hat{F}^{(2)}_{L\wedge S}(z)|]$. 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, , $\lim_{T\to\infty}\mathbb{P}\left\{g^0_{T,b_T,1-\alpha} < \sqrt{\frac{T}{b_T}}\mathrm{KS}^*_T\right\} = 1$ 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. □

B.5 Proof of Proposition 4

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

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 2ii (p. 24), $\lim_{T\to\infty}\mathbb{P}(\hat{c}_{1-\alpha} < \sqrt{T}\mathrm{KS}^*_T) = 1$, 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. □

B.6 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 $\breve{u}_{W_1}$. □

B.7 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 2(ii). □

C Monte-Carlo simulations

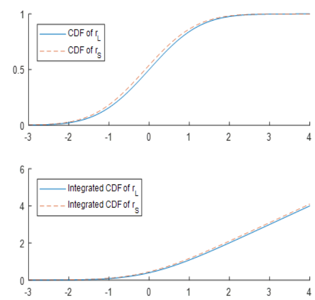
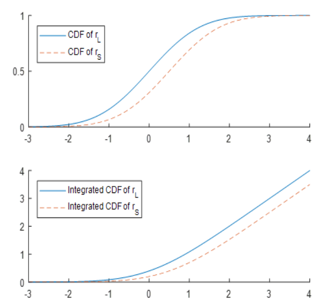
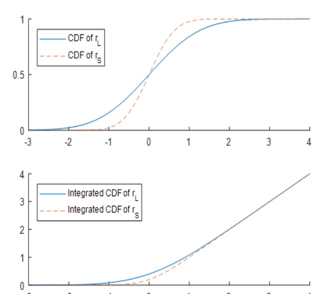
The objective of this section is to (i) explore the finite-sample behaviour of the tests; (ii) compare them with alternative implementations.

C.1 DGPs

C.1.1 Stylized DGPs

The stylized DGPs, which are taken from Whang (2019, p. 225–227) and displayed in Table A.1 (p. OA.20), 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.

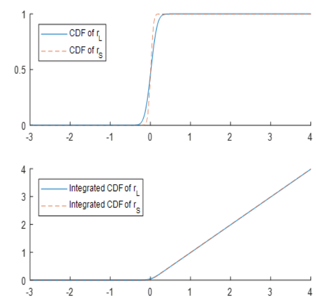
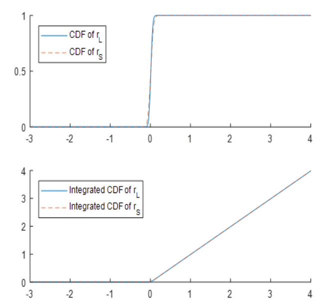
Table A.1: Stylized DGPs

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	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & .5^2 \end{bmatrix}\right)$	

C.1.2 DGPs calibrated to data

In Table A.2 (p. OA.21), 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.

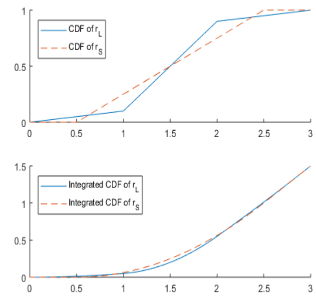
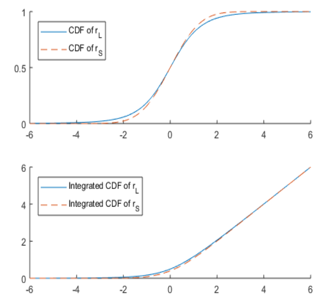
Table A.2: DGPs calibrated to data

<table>
<tr><th>H_0</th><th>DGP</th><th>Plots of CDF & Integrated CDF</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>

C.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.26), 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	$\begin{cases} R_L \hookrightarrow .3\mathcal{U}_{[0,3]} + .7\mathcal{U}_{[1,2]} \\ R_S \hookrightarrow \mathcal{U}_{[.5,2.5]} \\ \mathbb{C}\text{or}(R_S, R_L) = .7 \end{cases}$	
False	$\begin{cases} R_L \overset{IID}{\hookrightarrow} \text{t}(4) \\ R_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\text{or}(R_S, R_L) = .7 \end{cases}$	

C.2 Unconditional Test

C.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 2 and 5 on p. 24 & 33). 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 $a \in \mathbf{R}$.[17] 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 $b_T = \lfloor\sqrt{T}\rfloor$.

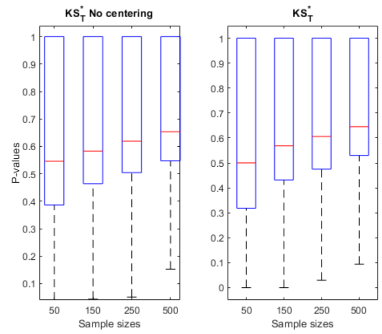
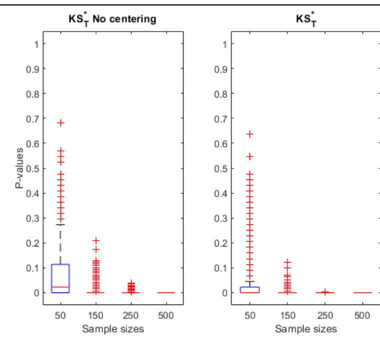
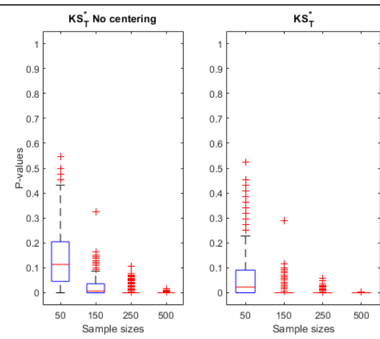
C.2.2 Results

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.

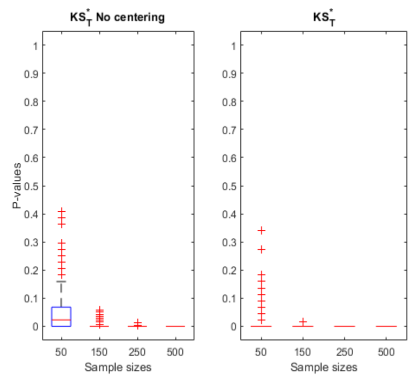
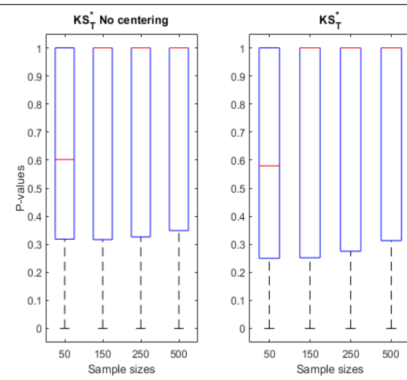
[17] The term e^{e} guarantees that the denominator is bigger than one, so the subsample size cannot be negative nor bigger than the sample size.

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	$\begin{bmatrix} R_L \\ R_S \end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & .5^2 \end{bmatrix}\right)$	

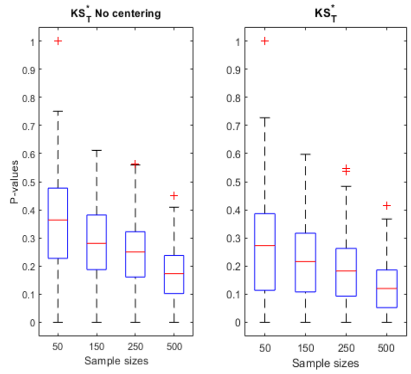
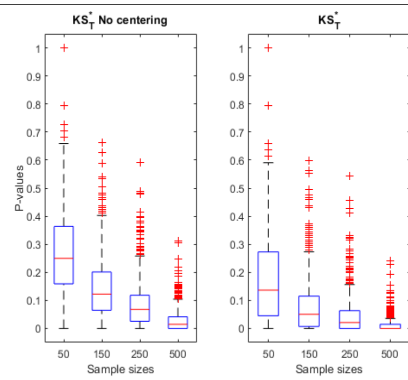
Note: The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of $\mathbb{KS}_T^*$ is approximated through block subsampling for "$\mathbb{KS}_T^*$ No centering," and centered block subsampling for "$\mathbb{KS}_T^*$." The block size is $b_T = \sqrt{T}$.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

H_0	**DGP**	**Boxplots of p-values**
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	$\begin{bmatrix}R_L\\R_S\end{bmatrix} \overset{IID}{\hookrightarrow} \mathcal{N}\left(\begin{bmatrix}.011\\.010\end{bmatrix}, \begin{bmatrix}.039^2 & .0012\\ & .057^2\end{bmatrix}\right)$	

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	$\begin{cases} R_L \overset{IID}{\hookrightarrow} .3\mathcal{U}_{[0,3]} + .7\mathcal{U}_{[1,2]} \\ R_S \overset{IID}{\hookrightarrow} \mathcal{U}_{[.5,2.5]} \\ \mathbb{C}\mathrm{or}(R_S, R_L) = .7 \end{cases}$	
False	$\begin{cases} R_L \overset{IID}{\hookrightarrow} \mathrm{t}(4) \\ R_S \overset{IID}{\hookrightarrow} \mathcal{N}(0,1) \\ \mathbb{C}\mathrm{or}(R_S, R_L) = .7 \end{cases}$	

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.

C.3 Conditional tests

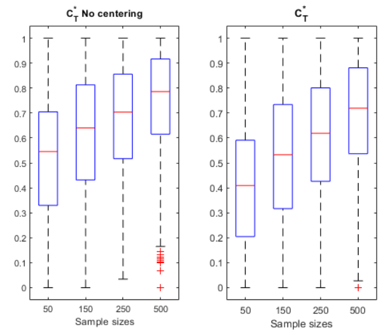
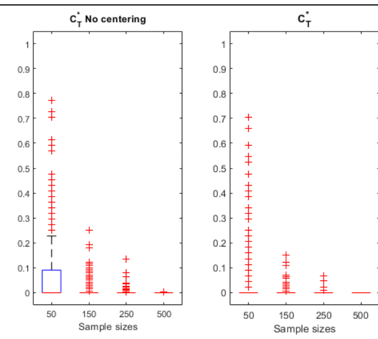
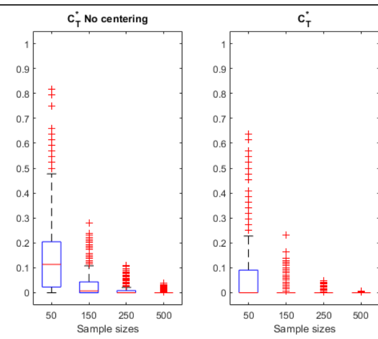
For ease of comparison, the parameterization and the DGPs are similar to the ones for the unconditional tests, except for a new common component. More precisely, we add a common independent Gaussian component $x \hookrightarrow \mathcal{N}(0, \sigma_x^2)$ to each of the DGPs. E.g., the first DGP is

$$\begin{bmatrix} R_L \\ R_S \end{bmatrix} = x + \begin{bmatrix} z_L \\ z_S \end{bmatrix}$$

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.

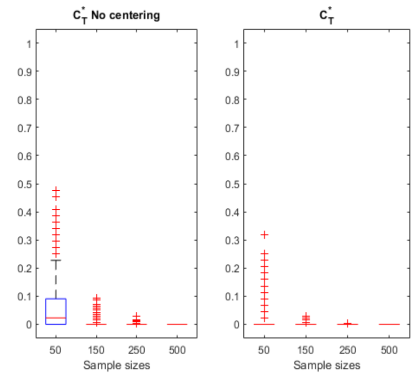
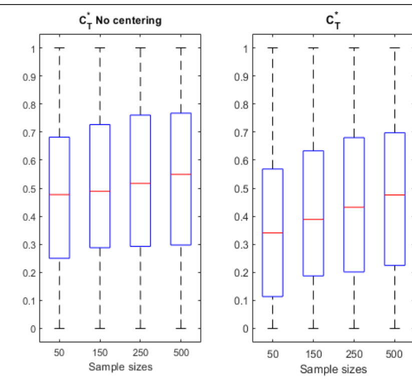
The patterns of the p-value distributions appear similar to the ones of the unconditional tests, namely smaller p-values for centered subsampling, better performance when the correlation between boths legs is higher.

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

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

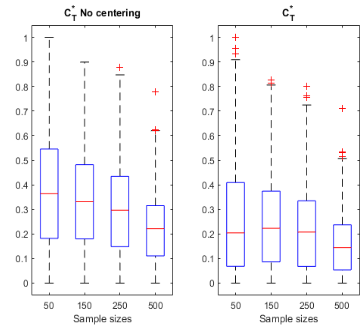
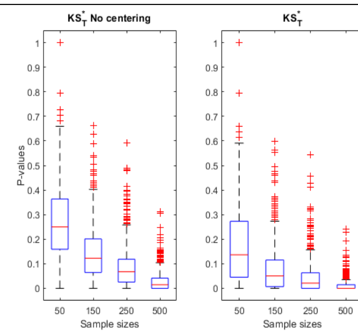
Note:The reported p-values are based on 1000 simulated samples of sample size equal to the indicated T. The distribution of C_T^* is approximated through block subsampling for "C_T^* No centering," and centered block subsampling for "C_T^*." The block size is $b_T = \sqrt{T}$.The tops and bottoms of 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	$\begin{bmatrix}R_L\\R_S\end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix}1.015\\1.0078\end{bmatrix}, \begin{bmatrix}.12^2 & .0051\\ & .057^2\end{bmatrix}\right)$	
True	$\begin{bmatrix}R_L\\R_S\end{bmatrix} \overset{IID}{\hookrightarrow} x + \mathcal{N}\left(\begin{bmatrix}1.011\\1.010\end{bmatrix}, \begin{bmatrix}.039^2 & .0012\\ & .057^2\end{bmatrix}\right)$	

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.

D 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

References

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

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

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

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

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

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

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

Durot, C. and A.-S. Tocquet (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).

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

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

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

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

Kallenberg, O. (2002/1997). *Foundation of Modern Probability* (Second ed.). Probability and Its Applications. Springer.

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

Luttmer, E. G. (1992). Asset pricing in economies with frictions. Ph.D. Thesis, Department of Economics, The University of Chicago.

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

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

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

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.

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

Rudin, W. (1953). *Principles of Mathematical Analysis* (3rd ed.). McGraw-Hill.

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.

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

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