RESEARCH

Swiss Finance Institute
Research Paper Series
N°23-81

Tradable Factor Risk Premia and Oracle Tests of Asset Pricing Models

Alberto Quaini
Erasmus University

Fabio Trojani
University of Geneva, University of Turin, and Swiss Finance Institute

Ming Yuan
Columbia University

TRADABLE FACTOR RISK PREMIA AND ORACLE TESTS OF ASSET PRICING MODELS

ALBERTO QUAINI, FABIO TROJANI AND MING YUAN*

First version: June 2022. This version: September 16, 2023

Abstract

Tradable factor risk premia are defined by the negative factor covariance with the Stochastic Discount Factor projection on returns. They are robust to misspecification or weak identification in asset pricing models, and they are zero for any factor weakly correlated with returns. We propose a simple estimator of tradable factor risk premia that enjoys the Oracle Property, i.e., it performs as well as if the weak or useless factors were known. This estimator not only consistently removes such factors, but it also gives rise to reliable tests of asset pricing models. We study empirically a family of asset pricing models from the factor zoo and detect a robust subset of economically relevant and well-identified models, which are built out of factors with a nonzero tradable risk premium. Well-identified models feature a relatively low factor space dimension and some degree of misspecification, which harms the interpretation of other established notions of a factor risk premium in the literature.

Keywords: Testing of asset pricing models, factor risk premia, useless and weak factors, factor selection, model misspecification, Oracle estimation and inference.

JEL classication: G12, C12, C13, C51, C52, C58.

---

*Alberto Quaini (email: quaini@ese.eur.nl) is with Erasmus University of Rotterdam, Erasmus school of Economics. Fabio Trojani (email: Fabio.Trojani@unige.ch) is with the University of Geneva, the AXA Chair in socio-economic risks at the University of Turin and a Senior Chair of the Swiss Finance Institute. Ming Yuan (email: Ming.Yuan@columbia.edu) is with Columbia University, Department of Statistics and Data Science Institute. For useful comments, we thank Cesare Robotti, Chen Zhou, Robin Lumsdaine, Sofonias Korsaye, Rasmus Lönn, Gustavo Freire, Chu Zhang, the participants of CFE 2022 and EcoSta 2023. The methods developed in this paper are implemented in the Rcpp package intrinsicFRP available on CRAN... All errors are ours.

1 Introduction

Over the last decades, a myriad of factors and low-dimensional factor models has been suggested to explain the cross-section of asset returns. Most of the empirical evidence for these models is based on various versions of the so-called two-step cross-sectional estimation and inference methodology, in which (i) factor risk premia are estimated by a cross-sectional projection of average returns on estimated asset return betas and (ii) asset return betas are estimated by a time series regression of returns on factors; see, e.g., Fama and MacBeth (1973) and Shanken (1992). However, this methodology has been shown to produce unreliable results in empirically relevant situationss where an asset pricing model may be misspecified or a candidate risk factor may be at best only weakly correlated with returns.[1]

Misspecification arises when asset expected returns are not completely spanned by the vectors of factor betas in an asset pricing model, and it has important implications for the properties of risk premia and test statistics produced by two-step cross-sectional regression approaches. First, misspecification gives rise to non zero pricing errors that impact the asymptotic distribution of estimated factor risk premia and derived statistics. Second, misspecification can lead to a definition of (pseudo) factor risk premia in two-step cross-sectional regression approaches that is to some extent arbitrary, as it depends on the weighting scheme chosen to penalize pricing errors in the second step of the methodology.

The presence in an asset pricing model of risk factors that may be at best only weakly correlated with returns gives rise to the so-called weak (or no) identification problem. This issue arises when the cross-sectional projection of expected returns on asset return betas does not uniquely identify a factor risk premium. Weak (or no) identification has as well important implications for the reliability of estimated risk premia and test statistics produced by two-step cross-sectional regression approaches, as it may spuriously favour statistically weak factors or even completely useless ones, i.e., factors that are uncorrelated with returns. These undesirable features may be further exacerbated by asset pricing models in which misspecification and weak (or no) identification coexist.

This paper addresses the testing of asset pricing models in the joint presence of misspecification and weak (or no) identification, based on a well-established, economically motivated, notion of a tradable factor risk

[1] See Kan and Zhang (1999), Shanken and Zhou (2007), Kleibergen (2009), Kan et al. (2013), Gospodinov et al. (2014), and Kleibergen and Zhan (2015), among many others.

premium. A tradable factor risk premium is defined by the negative factor covariance with the projection of any stochastic discount factor (SDF) on the space of asset returns. Importantly, tradable factor risk premia are well-defined whenever the standard no arbitrage condition of a finite maximum asset Sharpe ratio holds, and they are zero by definition for any factor that is only weakly correlated with returns. Therefore, they are naturally suited objects to develop inference procedures in potentially misspecified or weakly identified asset pricing models.

We show that a consistent finite sample selection of factors that are at best only weakly correlated with returns is easily achievable using tradable factor risk premia, together with a reliable and efficient inference on the risk premia of all other factors in an asset pricing model. To achieve this, we exploit the fact that a standard sample version of tradable factor risk premia follows an asymptotically Gaussian distribution under quite general assumptions, which cover models potentially including factors weakly correlated with returns. Crucially, this distribution may imply a distortion of estimated risk premia in weakly-identified models, but only for the risk premia of factors that are weakly correlated with returns.

We exploit these key properties to build a convenient Oracle estimator of tradable factor risk premia, which overcomes the potential distortions generated by weak factors in an asset pricing model. This estimator is given in closed-form as a penalized minimum distance correction of sample tradable risk premia and improves on the latter in two main directions. First, it consistently identifies the set of factors that are weakly correlated with returns. Second, it yields an (asymptotically) efficient estimate of tradable risk premia for all other factors in an asset pricing model as if these factors were known. By design, our Oracle estimator directly produces a valid inference for the tradable risk premia of factors that are not weakly correlated with returns. Alternatively, it can be used to perform a consistent preliminary screening of factors weakly correlated with returns, in order to reliably and efficiently apply in a second step standard cross-sectional inference methodologies to the risk premia of all other factors in an asset pricing model.

We make use of our Oracle estimation and inference methodology for tradable factor risk premia, in order to build a coherent and easily applicable framework for studying the asset pricing properties of a broad class of factor models from the factor zoo. For this purpose, we first form randomized families of candidate models of largest factor space dimension, in order to single out a set of well-identified submodels consisting exclusively of factors that are neither useless nor weak. Based on this set, we then pin down the properties of the resulting distribution of factors and factor risk premia across well-identified models, for various benchmark choices of test assets in the literature.

Our empirical analysis detects a quite robust set of economically relevant, well-identified models from the factor zoo, which are all built out of a relatively small set of factors having a nonzero tradable risk premium. Such factors include market, size, intermediaries capital ratio, market with a hedged unpriced component, a mispricing factor, a long-term behavioral factor, a liquidity factor, a conservative minus aggressive factor and a composite equity issuance factor. At the same time, we find that the relatively low factor space dimension of well-identified models is associated with some degree of misspecification. This feature harms the interpretation of other established definitions of a factor risk premium in the literature, which may give rise to excessively model-dependent risk premia and testing procedures in presence of untradable factor risks. For instance, for factors intermediaries capital ratio and market with a hedged unpriced component, we find that the risk premium implied by two-step cross-sectional regression approaches, across well-identified models and test assets, exhibits both an economically implausible variability and weak statistical significance.

The rest of the paper is structured as follows. After a review of the related literature, Section 2 defines tradable factor risk premia and clarifies their main properties. In Section 3, we introduce our Oracle estimator of tradable factor risk premia. We then derive its asymptotic properties under varying assumption on the (potentially vanishing) correlation between factors and returns. Finally, we explain our factor selection and inference procedure based on tradable risk premia. Section 4 provides extensive Monte Carlo simulation evidence on the finite sample accuracy of our methodology, in comparison to alternative approaches in the literature. Section 5 presents our empirical study, while Section 6 concludes and provides directions for future research.

2 Factor risk premia

Let $\boldsymbol{R}_t = (R_{1t}, \dots, R_{Nt})'$ and $\boldsymbol{F}_t := (F_{1t}, \dots, F_{Kt})'$ be a vector of excess returns and a vector of candidate asset pricing factors, where $K < N$, observed at times $t = 1, \dots, T$. The joint vector $\boldsymbol{Y}_t := (\boldsymbol{R}_t', \boldsymbol{F}_t')'$ has moments partitioned as:

$$\mathbb{E}\left[\boldsymbol{Y}_t\right] = \begin{pmatrix} \boldsymbol{\mu}_R \\ \boldsymbol{\mu}_F \end{pmatrix} \quad ; \quad \mathbb{C}\mathrm{ov}\left[\boldsymbol{Y}_t, \boldsymbol{Y}_t\right] = \begin{bmatrix} \boldsymbol{V}_R & \boldsymbol{V}_{RF} \\ \boldsymbol{V}_{FR} & \boldsymbol{V}_F \end{bmatrix} . \tag{1}$$

Given a set of test asset returns and candidate asset pricing factors, the main question studied by empirical asset pricers is whether some factors may reflect sources of systematic risk that are priced in the cross-section of asset returns. In order to answer this question empirically, one crucially needs both an appropriate factor risk premium definition and reliable inference techniques for testing hypotheses about factor risk premia.

Existing notions of factor risk premia all rely on the basic asset pricing equation implied by the choice of a corresponding Stochastic Discount Factor (SDF) projection M_t:

$$\boldsymbol{\lambda}(M) = -\mathbb{C}ov[\boldsymbol{F}_t, M_t] . \tag{2}$$

In all cases, SDF projection M_t is defined as some suitable approximation of a corresponding unknown SDF in an arbitrage-free market. Such SDF projections can be quite different with respect to several important features, including, e.g., the type of risks they span and the optimality criteria used to build them. More importantly for our purposes, most of these projections directly depend on the properties of an asset pricing model and are not well-defined in not or weakly identified models.

2.1 Tradable factor risk premia

In order to define tradable factor risk premia, the only assumption needed is the nonredundancy of asset returns.

Assumption 1. Covariance matrix $\boldsymbol{V}_R$ is positive definite.

Assumption 1 is standard and ensures a finite upper bound on the largest Sharpe ratio of any excess return portfolio. This bound is attained by the maximum Sharpe ratio portfolio having portfolio weights:

$$\boldsymbol{w} = \boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R . \tag{3}$$

Portfolio (3) identifies the tradable minimum variance projection of any stochastic discount factor (SDF) onto the space spanned by excess returns and constant payoffs:[2]

$$M_t^i := 1 - \boldsymbol{\mu}_R'\boldsymbol{V}_R^{-1}(\boldsymbol{R}_t - \boldsymbol{\mu}_R) . \tag{4}$$

[2] The normalization $\mathbb{E}[M_t^i] = 1$ is without loss of generality since we work with excess returns.

By construction, this projection satisfies the basic risk premium identity:

$$\mathbb{E}[\boldsymbol{R}_t] = -\,\mathbb{C}\text{ov}[\boldsymbol{R}_t, M_t^i]\ . \tag{5}$$

Consistently with these pricing properties, a tradable factor risk premium is the risk premium assigned to a candidate asset pricing factor by SDF projection (4).

Definition 1. Given Assumption 1, the tradable risk premium of factor vector $\boldsymbol{F}_t$ is defined by:

$$\boldsymbol{\lambda}^i := -\,\mathbb{C}\text{ov}[\boldsymbol{F}_t, M_t^i] = \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R\ . \tag{6}$$

The tradable factor risk premium in Definition 1 is a special case of general risk premium definition (2) with $M_t := M_t^i$ and it satisfies a number of desirable properties listed below.

(i) It is well-defined whenever Assumption 1 holds. In particular, it remains well-defined when covariance matrix $\boldsymbol{V}_{RF}$ does not have a full column rank. This feature emerges in presence of useless or weak factors or, more generally, when covariances of different factors with returns are collinear.

(ii) It uniquely assigns a zero risk premium to any factor having a zero or a vanishing population covariance with all excess returns, i.e., any useless or weak factor.

(iii) It is independent of the misspecification properties of a factor asset pricing model, because it is directly defined via the SDF projection on the space of returns. In particular, the tradable risk premium of a factor depends exclusively on the joint distribution of that factor with test asset returns, and not on the joint distribution properties with other factors.

(iv) It is invariant to linear invertible transformations of excess returns, and thus also to a simple repackaging of the test assets.[3]

(v) It offers a straightforward economic interpretation as it coincides with the expectation of the least squares projection of factors on the span of excess returns:

$$\boldsymbol{\lambda}^i = \mathbb{E}[\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{R}_t]\ ,$$

[3] A detailed discussion of this important invariance property is given in Kandel and Stambaugh (1995).

i.e., tradable factor risk premia are interpretable as the (tradable) risk premia of factor replicating portfolios.[4]

In addition to the above properties, tradable factor risk premia naturally give rise to an exact beta representation of expected returns in correctly specified, but not necessarily identified, models.

Proposition 2.1. *Let Assumption 1 hold. Then, following statements are equivalent:*

1. *There exists a valid factor SDF of the form*
$$M_F := 1 - \boldsymbol{\gamma}'(\boldsymbol{F}_t - \boldsymbol{\mu}_F) \tag{7}$$
for some $\boldsymbol{\gamma} \in \mathbb{R}^K$.

2. *Following tradable beta-representation of expected returns holds:*
$$\boldsymbol{\mu}_R = \boldsymbol{V}_{RF}\boldsymbol{\gamma} = \boldsymbol{\beta}^i \boldsymbol{\lambda}^i \ , \tag{8}$$
for a $N \times K$ *matrix of mimicking portfolio return betas given by:*
$$\boldsymbol{\beta}^i := \boldsymbol{V}_{RF}(\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^+ \ , \tag{9}$$
where $\boldsymbol{A}^+$ *denotes the Moore–Penrose inverse of a matrix* $\boldsymbol{A}$.

Proposition 2.1 states that in correctly specified models there always exists an exact beta-representation of expected returns, based on tradable factor risk premia $\boldsymbol{\lambda}^i$ and $N \times K$ matrix $\boldsymbol{\beta}^i$ of tradable factor betas. Note that this representation is always valid, irrespective of the identification properties of an asset pricing model. Indeed, recall that $\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF}$ is the variance covariance matrix of the factor mimicking portfolio returns. Therefore, in identified models $\boldsymbol{\beta}^i$ is the uniquely given $N \times K$ matrix such that:

$$\boldsymbol{\beta}^i \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF} = \boldsymbol{V}_{RF} \ ,$$

i.e., the matrix of factor betas of mimicking portfolio returns. More generally, in unidentified models there always exists a multiplicity of $N \times K$ matrices $\tilde{\boldsymbol{\beta}}$ such that:

$$\tilde{\boldsymbol{\beta}} \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF} = \boldsymbol{V}_{RF} \ ,$$

i.e., without additional assumptions the factor betas of portfolio mimicking returns are not identified. In such settings, definition (9) identifies a unique matrix of betas of portfolio factor mimicking returns, which (i) satisfies a tradable beta representation of expected returns and (ii) implies the smallest beta exposures under the Euclidean metric.

[4] The relation between tradable factor risk premia and the risk premia of factor mimicking portfolio returns is explained in more detail in Section 2.2.

2.2 Alternative notions of factor risk premia

Tradable factor risk premia differ from other commonly used notions of factor risk premia in the literature, because of the different SDF projection underlying them. Tradable risk premia are defined by a linear SDF projection on the return space, which correctly prices all test assets. Other available notions are defined by a SDF projections on the factor space, which implies some degree of asset mispricing in misspecified models. To clarify these differences, recall the definition of a misspecification-robust factor risk premium in Kan et al. (2013):

$$\boldsymbol{\lambda}(\boldsymbol{W}) := -\,\mathbb{C}\mathrm{ov}[\boldsymbol{F}_t, M_t(\boldsymbol{W})] \ , \tag{10}$$

for some $N \times N$ symmetric and positive-definite weighting matrix $\boldsymbol{W}$, where

$$M_t(\boldsymbol{W}) := 1 - \boldsymbol{\gamma}'(\boldsymbol{W})(\boldsymbol{F}_t - \boldsymbol{\mu}_F)$$

is a SDF projection on the factor space, with factor loadings such that:

$$\boldsymbol{\gamma}(\boldsymbol{W}) \in \underset{\boldsymbol{\gamma}\in\mathbb{R}^K}{\operatorname{argmin}} \ \mathbb{E}[M_t(\boldsymbol{W})\boldsymbol{R}_t]'\boldsymbol{W}\mathbb{E}[M_t(\boldsymbol{W})\boldsymbol{R}_t] \ . \tag{11}$$

Factor risk premium (10) is another special case of the general risk premium definition (2), which is however identified if and only if factor loadings (11) are identified, i.e., the covariance matrix $\boldsymbol{V}_{RF}$ has full column rank, in which case:

$$\boldsymbol{\gamma}(\boldsymbol{W}) = (\boldsymbol{V}_{FR}\boldsymbol{W}\boldsymbol{V}_{RF})^{-1}\boldsymbol{V}_{FR}\boldsymbol{W}\boldsymbol{\mu}_R \ . \tag{12}$$

This gives rise to following standard two-step weighted least squares expression for factor risk premia in identified models with non redundant factors:

$$\boldsymbol{\lambda}(\boldsymbol{W}) = (\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\boldsymbol{W}\boldsymbol{\mu}_R \ , \tag{13}$$

where $\boldsymbol{\beta} = \boldsymbol{V}_{RF}\boldsymbol{V}_F^{-1}$ is the $N \times K$ matrix of factor betas. These identities give rise to following direct links between tradable factor risk premia and misspecification-robust factor risk premia in identified models.

Proposition 2.2. *Let Assumption 1 hold and matrix $\boldsymbol{V}_{RF}$ have full column rank. It then follows:*

$$\boldsymbol{\lambda}(\boldsymbol{V}_R^{-1}) = \boldsymbol{V}_F(\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^{-1}\boldsymbol{\lambda}^i \ . \tag{14}$$

In particular, $\lambda_j(\boldsymbol{V}_R^{-1}) = \lambda_j^i$ for any factor F_j spanned by test asset returns. If furthermore correct specification property (7) in Proposition 2.1 holds and matrix $\boldsymbol{V}_F$ is positive definite, then for any weighting matrix $\boldsymbol{W}$ it follows:

$$\boldsymbol{\lambda}(\boldsymbol{W}) = (\boldsymbol{\beta}'\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\boldsymbol{\mu}_R = (\boldsymbol{\beta}'\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\boldsymbol{\beta}^i\boldsymbol{\lambda}^i \ , \tag{15}$$

with

$$\boldsymbol{\lambda}^i = (\boldsymbol{\beta}^{i\prime}\boldsymbol{\beta}^i)^{-1}\boldsymbol{\beta}^{i\prime}\boldsymbol{\mu}_R \ . \tag{16}$$

Proposition 2.2 shows that in identified models misspecification-robust factor risk premia implied by weighting matrix $\boldsymbol{W} = \boldsymbol{V}_R^{-1}$ are given by an invertible linear transformation of tradable factor risk premia. Since in such models $\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF}$ is a (positive-definite) variance-covariance matrix of the vector of factor mimimicking portfolio payoffs, this also shows that the factor market prices of risk induced by both definitions are the same:

$$\boldsymbol{V}_F^{-1}\boldsymbol{\lambda}(\boldsymbol{V}_R^{-1}) = (\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^{-1}\boldsymbol{\lambda}^i \ . \tag{17}$$

Moreover, when a factor in an asset pricing model is tradable, i.e., spanned by test asset returns, then its misspecification-robust and tradable factor risk premia are identical. Therefore, misspecification-robust risk premia differ from tradable risk premia by the fact that they assign to not tradable factors a rescaled risk premium, which is adjusted to the larger variance covariance matrix of not tradable factors relative to the variance-covariance matrix of factor mimicking portfolio payoffs.

In models that are additionally correctly specified, equation (16) shows that tradable factor risk premia are equivalently given by the coefficients of a two-step cross-sectional ordinary least squares regression of expected returns on factor mimicking portfolio return betas. Such a definition of tradable factor risk premia is akin to a well-established approach in the literature, which aims to identify with a two-step cross-sectional least squares approach the risk premia of factor mimicking portfolios, as opposed to the risk premia of the original factors.[5] However, this equivalent definition implicitly assumes the nondegeneracy of the covariance matrix of factor mimicking portfolio returns, which is instead singular in not or weakly identified models where matrix $\boldsymbol{V}_{RF}$ is reduced rank.[6] Finally, equations (15)–(16) together show that the factor risk premia of correctly specified and identified models can always be recovered from tradable risk premia, by first projecting expected returns on tradable factor betas and then regressing that projection on factor betas.

Having clarified the desirable properties of tradable factor risk premia when working with potentially unidentified models, relative to other common notions in the literature, we study in the next section the asymptotic

[5]This approach dates back at least to Huberman et al. (1987), Grinblatt and Titman (1987) and Lehmann and Modest (1988).

[6]In identified but potentially misspecified models, the two-step definition of factor mimicking portfolio risk premia is not robust with respect to the weighting matrix $\boldsymbol{W}$ used to weight pricing errors. In contrast, Definition 1 is universally valid and robust.

properties of a convenient class of tradable factor risk premium estimators, under varying assumptions on the data generating process for returns and potentially useless or weak asset pricing factors. Building on these estimators, we then set up our Oracle tests of asset pricing models.

3 Estimation of tradable factor risk premia

Definition 1 naturally motivates a direct estimator of tradable factor risk premia, which is given by the sample version of equation (1):

$$\hat{\boldsymbol{\lambda}} := \hat{\boldsymbol{V}}_{FR}\hat{\boldsymbol{V}}_R^{-1}\hat{\boldsymbol{\mu}}_R \ , \tag{18}$$

with the sample mean and sample covariance matrix estimators given by:

$$\hat{\boldsymbol{\mu}} := (\hat{\boldsymbol{\mu}}_R', \hat{\boldsymbol{\mu}}_F')' := \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{Y}_t \ , \tag{19}$$

and

$$\hat{\boldsymbol{V}} := \begin{bmatrix} \hat{\boldsymbol{V}}_R & \hat{\boldsymbol{V}}_{RF} \\ \hat{\boldsymbol{V}}_{FR} & \hat{\boldsymbol{V}}_F \end{bmatrix} := \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{Y}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{Y}_t - \hat{\boldsymbol{\mu}})' \ . \tag{20}$$

In the sequel, we first show that estimator (18) has a number of useful properties, which make it a convenient initial estimator for building tests of potentially misspecified and not necessarily identified asset pricing models. To this end, we prove in Section 3.1 that estimator (18) always has a standard asymptotically Gaussian distribution, even in the presence of useless factors uncorrelated with returns, or, more generally, weak factors having a correlation with returns that vanishes sufficiently fast with the sample size.[7] Furthermore, we show in Section 3.2 that in presence of weak factors having return correlations vanishing at a slow rate,[8] the asymptotic distribution of estimator (18) is still Gaussian, but dependent on an asymptotic bias term. Crucially, this bias only affects the estimated risk premia of the weak factors with slowly vanishing return correlations. Moreover, it does not affect the estimator's asymptotic mean square error.

By exploiting the asymptotic properties of sample tradable risk premia, we next build in Section 3.3 an Oracle estimator of tradable factor risk

[7] I.e., at rate at least $1/T^{\alpha}$ for some $\alpha > 1/2$.

[8] I.e., at rate $1/\sqrt{T}$.

premia. This estimator is given in closed-form by a penalized minimum distance correction of sample tradable risk premia, which corrects the potential biases generated by weak factors with slowly vanishing return correlations. We show that our Oracle estimator improves on sample tradable risk premia in two directions. First, by ensuring a consistent finite-sample selection of factors that are not weak or useless. Second, by giving rise to an efficient asymptotic distribution for the estimated risk premia of all other factors, which is independent of the potential presence of useless or weak factors in an asset pricing model. Therefore, our Oracle estimator naturally produces a valid asymptotic inference framework for testing potentially misspecified and not or only weakly identified asset pricing models.

3.1 Tradable factor risk premium estimation with useless factors

Let the vectors of excess return and factor innovations at time t be denoted by:

$$\bar{\boldsymbol{R}}_t := \boldsymbol{R}_t - \boldsymbol{\mu}_R \ , \tag{21}$$
$$\bar{\boldsymbol{F}}_t := \boldsymbol{F}_t - \boldsymbol{\mu}_F \ . \tag{22}$$

The next proposition establishes the asymptotic properties of tradable factor risk premium estimator (18), under the general assumption of factor and excess return stationarity and ergodicity.[9]

Proposition 3.1. *Let Assumption 1 hold and $\mathbf{Y}_t$ be jointly stationary and ergodic with finite fourth moments. Let further matrix:*

$$\boldsymbol{\Sigma} := \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{E}[\boldsymbol{h}_t \boldsymbol{h}_s'] \ , \tag{23}$$

where

$$\boldsymbol{h}_t := \bar{\boldsymbol{F}}_t \bar{\boldsymbol{R}}_t' \boldsymbol{V}_R^{-1} \boldsymbol{\mu}_R - \boldsymbol{V}_{FR} \boldsymbol{V}_R^{-1} \bar{\boldsymbol{R}}_t \bar{\boldsymbol{R}}_t' \boldsymbol{V}_R^{-1} \boldsymbol{\mu}_R + \boldsymbol{V}_{FR} \boldsymbol{V}_R^{-1} \bar{\boldsymbol{R}}_t \ , \tag{24}$$

be positive definite. It then follows:

$$\sqrt{T}(\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i) \to_d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \ . \tag{25}$$

[9] This assumption, together with the assumption of finite fourth moments of random vector $\boldsymbol{Y}_t$, is needed to invoke a suitable central limit theorem for jointly stationary and ergodic random vector $((\boldsymbol{Y}_t - \mathbb{E}[\boldsymbol{Y}_t])', \mathrm{vec}(\boldsymbol{Y}_t \boldsymbol{Y}_t' - \boldsymbol{V})')'$.

In Proposition 3.1, the asymptotic covariance matrix (23) depends on three sources of estimation uncertainty, which are present in the sum on the right hand side of equation (24). They capture the uncertainty generated by the estimation of population moments $\boldsymbol{V}_{RF}$, $\boldsymbol{V}_R$ and $\boldsymbol{\mu}_R$ in the definition of tradable factor risk premia.

Several useful properties of asymptotic distribution (25) are worth noticing. First, the marginal asymptotic distribution of sample tradable risk premia only depends on the joint factor distribution with asset returns, but not on the joint distribution with other factors.[10] This property does not hold in general, e.g., for factor risk premium estimators implied by two-step cross-sectional regression approaches. Second, in Proposition 3.1 the $N \times K$ matrix $\boldsymbol{V}_{RF}$ is not required to have a full column rank. Therefore, asymptotic distribution (25) is valid also in presence of useless factors and one can directly test with statistics having this asymptotic distribution the null hypothesis of a zero tradable risk premium. This is a sharp difference with, e.g., the approach in Gospodinov et al. (2014), in which useless factors have first to be consistently eliminated with a sequential testing procedure, before testing risk premium hypotheses involving factors that are not useless.

3.2 Tradable factor risk premium estimation with useless and weak factors

The stationarity assumption in Proposition 3.1 does not cover nearly-singular settings, in which matrix $\boldsymbol{V}_{RF}$ is approximated by a sequence of matrices $\boldsymbol{V}_{RF}^{(T)}$ depending on the sample size and having rank strictly greater than $\boldsymbol{V}_{RF}$. This is the context in which some factors may be weak, in the sense that while they exhibit a non degenerate covariance structure with returns in finite samples, they may give rise to a degenerate covariance structure as the sample size grows. Such a weak factor setting has been shown in Kleibergen (2009), among many others, to be empirically relevant for many macro-based asset pricing factors.

To study the properties of tradable risk premia estimators in weak factor settings, we make the following assumption, in order to allow for the presence of both useless and weak factors.

Assumption 2. The sample size-dependent joint covariance matrix of factors

[10] As we demonstrate in Section 4, this robustness of marginal asymptotic distributions with respect to the presence of additional factors in an asset pricing model is a key property, which is instrumental for building consistent exclusion procedures of useless or weak factors that perform reliably in finite samples.

and returns is given by:

$$\boldsymbol{V}^{(T)} = \begin{bmatrix} \boldsymbol{V}_R & \boldsymbol{V}_{RF}^{(T)} \\ \boldsymbol{V}_{FR}^{(T)} & \boldsymbol{V}_F \end{bmatrix} , \tag{26}$$

where submatrix $\boldsymbol{V}_{RF}^{(T)}$ is of the form:

$$\boldsymbol{V}_{RF}^{(T)} := \boldsymbol{V}_{RF} + \boldsymbol{\Gamma}/T^{\alpha} + \boldsymbol{\Delta}/\sqrt{T} ,$$

for some $\alpha > 1/2$ and $N \times K$ matrices $\boldsymbol{\Gamma}, \boldsymbol{\Delta}$ such that $\mathbf{0} = \boldsymbol{V}_{FR}\boldsymbol{\Gamma} = \boldsymbol{V}_{FR}\boldsymbol{\Delta}$. In Assumption 2, matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Delta}$ parametrize the way how covariance matrix $\boldsymbol{V}_{RF}^{(T)}$ converges to its (possibly reduced-rank) limit $\boldsymbol{V}_{RF}$. By construction,

$$\mathrm{Rank}\left(\boldsymbol{V}_{RF}^{(T)}\right) \geq \mathrm{Rank}\left(\boldsymbol{V}_{RF}\right) ,$$

with strict inequality holding when $\boldsymbol{\Gamma} \neq \mathbf{0}$ or $\boldsymbol{\Delta} \neq \mathbf{0}$. The orthogonality restriction imposed on $\boldsymbol{\Gamma}$ ($\boldsymbol{\Delta}$) embeds weak-factor settings in which covariance matrix $\boldsymbol{V}_{RF}$ contains columns of zeros but the corresponding columns of $\boldsymbol{V}_{RF}^{(T)}$ are nonzero and vanishing at a rate $1/T^{\alpha}$ with $\alpha > 1/2$ ($1/\sqrt{T}$). More generally, the orthogonality restriction imposed on $\boldsymbol{\Gamma}$ and $\boldsymbol{\Delta}$ also embeds nearly singular settings in which some columns of $\boldsymbol{V}_{RF}$ are collinear, but the corresponding columns of $\boldsymbol{V}_{RF}^{(T)}$ are only weakly collinear.

The next proposition establishes the asymptotic properties of tradable factor risk premium estimator (18) in the weak factor setting, under the general assumption of factor and excess return near-epoch dependence.[11]

Proposition 3.2. *Let Assumptions 1 and 2 hold, $\boldsymbol{Y}_t$ be near-epoch dependent with finite fourth moments, and regularity conditions detailed in Davidson (1992) hold. Let further matrix $\boldsymbol{\Sigma}$ defined in Proposition 3.1 be positive definite. It then follows:*

$$\sqrt{T}(\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i) \to_d \mathcal{N}\left(\boldsymbol{\lambda}(\boldsymbol{\Delta}), \boldsymbol{\Sigma} - \boldsymbol{\lambda}(\boldsymbol{\Delta})\boldsymbol{\lambda}(\boldsymbol{\Delta})'\right) , \tag{27}$$

where

$$\boldsymbol{\lambda}(\boldsymbol{\Delta}) := \boldsymbol{\Delta}'\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R . \tag{28}$$

[11] This assumption, together with the assumption of finite fourth moments of random vector $\boldsymbol{Y}_t$ and regularity conditions detailed in Davidson (1992), is needed to invoke a suitable central limit theorem for jointly near-epoch dependent random vector $((\boldsymbol{Y}_t - \mathbb{E}[\boldsymbol{Y}_t])', \mathrm{vec}(\boldsymbol{Y}_t\boldsymbol{Y}_t' - \boldsymbol{V}^{(T)})')'$.

Proposition 3.2 shows that weak factors vanishing at rate $1/T^{\alpha}$ with $\alpha > 1/2$ do not affect the asymptotic distribution of sample tradable factor risk premia.[12] This property is remarkable, because weak factors giving rise to rapidly vanishing covariance distortions do instead crucially affect the asymptotic distribution of risk premium estimators based on two-step cross-sectional regression approaches.

A distinct situation emerges with respect to weak factors associated with slowly vanishing covariance distortions $\boldsymbol{\Delta}/\sqrt{T}$ in Assumption 1. Indeed, a comparison of asymptotic distributions (25) and (27) shows that the latter distribution displays both a bias term $\boldsymbol{\lambda}(\boldsymbol{\Delta})$ and a smaller variance covariance matrix because of component $-\boldsymbol{\lambda}(\boldsymbol{\Delta})\boldsymbol{\lambda}(\boldsymbol{\Delta})'$. Therefore, weak factors associated with slowly vanishing covariance distortions generate a bias in the asymptotic distribution of sample tradable risk premia, but they do not affect the estimator's mean square error. Importantly, these biases are concentrated in the marginal asymptotic distribution of estimated risk premia of weak factors with slowly vanishing covariance distortions, because matrix $\boldsymbol{\Delta}$ in Assumption 1 is by construction nonzero only along these very same components.

In summary, the asymptotic distribution of Proposition 3.2 shows that sample tradable risk premia tend to estimate a spuriously non zero finite sample risk premium exclusively for the weak factors with a slowly vanishing rate. Therefore, in the next section we borrow from the properties of sample tradable risk premia to build an associated Oracle estimator that consistently eliminates these distortions.

3.3 Oracle tradable factor risk premium estimation with useless and weak factors

This section studies Oracle tradable risk premium estimators. Borrowing the terminology from high dimensional statistics (see, e.g., Fan and Li, 2001), we define an Oracle tradable risk premium estimator as an estimator satisfying two key properties. First, it consistently selects in finite samples factors that are not weak or useless. Second, it implies an efficient asymptotic distribution for the estimated risk premia of the selected factors.

Thanks to the convenient asymptotic properties of sample tradable risk premia, Oracle tradable risk premium estimators can be built with a simple

[12] This result directly follows from the fact that asymptotic distribution (27) is independent of matrix $\boldsymbol{\Gamma}$, which models in Assumption 1 the rapidly vanishing components of the covariance distortions generated by weak factors.

approach. Denote by

$$\widehat{\boldsymbol{\rho}} := [\widehat{\boldsymbol{\rho}}_1, \ldots, \widehat{\boldsymbol{\rho}}_K] := \widehat{\mathbb{C}\mathrm{or}}[\boldsymbol{R}_t, \boldsymbol{F}_t'] \ , \tag{29}$$

the $N \times K$ matrix of sample correlations between returns and factors. We propose an Oracle tradable risk premium estimator built by means of a convenient minimum distance correction of sample tradable risk premia, in which estimated risk premia of factors having small sample correlations with all asset returns are shrank using a suitable data-driven penalty.

Definition 2. Given a penalty parameter $\tau_T > 0$, consider the penalized estimator defined by:

$$\check{\boldsymbol{\lambda}}^i := (\check{\lambda}_1^i, \ldots, \check{\lambda}_K^i)' := \underset{\boldsymbol{\lambda} \in \mathbb{R}^K}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}\|_2^2 + \tau_T \sum_{k=1}^{K} \frac{|\lambda_k|}{\|\widehat{\boldsymbol{\rho}}_k\|_2^2} \right\} \ . \tag{30}$$

Estimator (30) is defined by a penalized minimum Euclidean distance correction of sample tradable risk premia and belongs to the class of proximal estimators studied in Quaini and Trojani (2022). The optimization problem in equation (30) is solvable in closed-form and gives rise to the soft-thresholding formula:

$$\check{\lambda}_k^i = \mathrm{sign}(\widehat{\lambda}_k^i) \max \left\{ |\hat{\lambda}_k^i| - \frac{\tau_T}{\|\hat{\boldsymbol{\rho}}_k\|_2^2}, 0 \right\} \ ; \ k = 1, \ldots K \ . \tag{31}$$

Therefore, estimator $\check{\boldsymbol{\lambda}}^i$ implies a zero estimated risk premium for all factors associated with a sample tradable risk premium $\hat{\lambda}_k^i$ that is smaller in absolute value than scaled penalty parameter $\tau_T / \|\widehat{\boldsymbol{\rho}}_k\|_2^2$.

We next show that an appropriate choice of tuning parameter τ_T in equation (30) implies the Oracle property. To see this, let

$$\mathcal{S} := \{k \in \{1, \ldots, K\} : \boldsymbol{V}_{RF_k} \neq \boldsymbol{0}\} \ , \tag{32}$$

be the active set indexing components of factor vector $\boldsymbol{F}_t = (F_{1t}, \ldots, F_{Kt})'$ that are neither useless nor weak in the sense of Assumption 2. Accordingly, for any vector $\boldsymbol{x} \in \mathbb{R}^K$ we denote by $\boldsymbol{x}_{\mathcal{S}}$ the subvector consisting only of components of $\boldsymbol{x}$ with index in $\mathcal{S}$. Finally, the estimated active set implied by estimator $\check{\boldsymbol{\lambda}}^i$ is denoted by

$$\check{\mathcal{S}} := \{k \in \{1, \ldots, K\} : \check{\lambda}_k^i \neq 0\} \ . \tag{33}$$

Using this notation, we characterize in the next proposition the asymptotic distribution of estimator (30) and its Oracle property.

Proposition 3.3. *Let the assumptions of Proposition 3.2 hold. Let further the sequence of tuning parameter τ_T be such that $\tau_T\sqrt{T} \to 0$ and $\tau_T T \to \infty$ as $T \to \infty$. It then follows, as $T \to \infty$:*

1. *Asymptotic distribution:*

$$\sqrt{T}\left(\check{\boldsymbol{\lambda}}^i_{\mathcal{S}} - \boldsymbol{\lambda}^i_{\mathcal{S}}\right) \to_d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathcal{S}}) \ , \tag{34}$$

with

$$\boldsymbol{\Sigma}_{\mathcal{S}} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{E}[\boldsymbol{h}_{\mathcal{S}t} \boldsymbol{h}'_{\mathcal{S}s}] \ , \tag{35}$$

where

$$\boldsymbol{h}_{\mathcal{S}t} := \boldsymbol{V}_{F_{\mathcal{S}}R} \boldsymbol{V}_R^{-1} \bar{\boldsymbol{R}}_t - \boldsymbol{V}_{F_{\mathcal{S}}R} \boldsymbol{V}_R^{-1} \bar{\boldsymbol{R}}_t \bar{\boldsymbol{R}}'_t \boldsymbol{V}_R^{-1} \boldsymbol{\mu}_R + \bar{\boldsymbol{F}}_{\mathcal{S}t} \bar{\boldsymbol{R}}'_t \boldsymbol{V}_R^{-1} \boldsymbol{\mu}_R \ , \tag{36}$$

and $\boldsymbol{V}_{F_{\mathcal{S}}R}$ is the submatrix containing all rows of matrix $\boldsymbol{V}_{FR}$ with index in $\mathcal{S}$.

2. *Consistent variable selection:*

$$\mathbb{P}(\check{\mathcal{S}} = \mathcal{S}) \to 1 \ .$$

A direct consequence of Proposition 3.3 is that proximal tradable risk premium estimator (30) is an Oracle estimator. Indeed, Property 2 states that the set of factors that are neither useless nor weak is correctly selected with a probability converging to one as the sample size grows. Property 1 further implies that the asymptotic distribution of the estimated risk premia of factors that are neither useless nor weak is identical to the asymptotic distribution of an Oracle sample tradable risk premium estimator which knows ex ante which factors are useless or weak.[13] In this way, our Oracle estimator directly produces a valid inference for the tradable risk premia of factors that are not weakly correlated with returns. Alternatively, it can be used to

[13]Such asymptotic distribution is reported in Proposition 3.1. The Oracle property of our proposed proximal tradable risk premium estimator is a natural consequence of the desirable properties of sample tradable risk premia. To ensure a consistent selection of factors that are not useless or weak, we exploit the fact that the sample correlations with returns of factors that are useless or weak are at most of order $1/\sqrt{T}$ in probability. Once a consistent factor selection is ensured, the efficiency of the asymptotic distribution for the estimated risk premia of the selected factors follows directly, because the marginal asymptotic distributions of sample tradable risk premia are independent of the properties of the other factors in an asset pricing model.

perform a consistent preliminary screening of factors weakly correlated with returns, in order to reliably apply in a second step standard cross-sectional inference methodologies to the risk premia of all other factors in an asset pricing model.

Remark 1. Notice that while according to Proposition 3.3 the shrinkage of sample tradable risk premia implied by soft-thresholding formula (31) is asymptotically negligible for factors that are neither useless nor weak, it may give rise to a finite-sample bias. Such bias can be removed by computing sample risk premia only for factors that have been selected in finite samples by our proximal estimator. By construction, under the assumptions of Proposition 3.3 the resulting "relaxed" proximal estimator is asymptotically equivalent to proximal estimator (30).

4 Monte Carlo simulation evidence

We study by Monte Carlo simulation the finite-sample distribution of sample tradable risk premium estimator (18), Oracle estimator (30) and its relaxed version.[14] In addition, we investigate the finite-sample factor selection properties induced by Oracle estimator (30).

4.1 Benchmark estimation and inference approaches

We benchmark our methodology to other important approaches in the literature. First, we consider the misspecification-robust estimation and inference approach in Kan et al. (2013). This approach relies on a factor risk premium defined with equation (13) for a weighting matrix $\boldsymbol{W} = \boldsymbol{V}_R^{-1}$. It gives rise to the factor risk premium estimator:

$$\hat{\boldsymbol{\lambda}} := \hat{\boldsymbol{\lambda}}(\hat{\boldsymbol{V}}_R^{-1}) := (\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{V}}_R^{-1}\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\beta}}\hat{\boldsymbol{V}}_R^{-1}\hat{\boldsymbol{\mu}}_R \ , \tag{37}$$

which is implied by the sample version of equation (13). Second, we document how our approach compares with respect to the penalized two-step Fama-MacBeth estimation and inference approach introduced in Bryzgalova (2015). This approach relies on a penalized factor risk premium estimator defined by:

$$\check{\boldsymbol{\lambda}} := \check{\boldsymbol{\lambda}}(\hat{\boldsymbol{V}}_R^{-1}) := \arg\min_{\boldsymbol{\lambda}\in\mathbb{R}^K} \left\{ (\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda})'\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{V}}_R^{-1}\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) + \tau_T \sum_{k=1}^{K} \frac{|\lambda_k|}{||\hat{\boldsymbol{\beta}}_k||_1^d} \right\}, \tag{38}$$

[14] Recall, that the relaxed Oracle estimator is obtained by computing sample tradable risk premia only for those factors that have been first selected by the Oracle estimator.

where $\tau_T > 0$, $d > 0$ are tuning parameters and $||\hat{\boldsymbol{\beta}}_k||_1$ denotes the l_1-norm of the $k-$column in the $N \times K$ matrix $\hat{\boldsymbol{\beta}}$ of sample betas.

By definition, this estimator is given by a penalized weighted minimum distance correction of misspecification-robust factor risk premium estimator $\hat{\boldsymbol{\lambda}}$. Since in unidentified models weighting matrix $\hat{\boldsymbol{W}} := \hat{\boldsymbol{\beta}}'\hat{\boldsymbol{V}}_R^{-1}\hat{\boldsymbol{\beta}}$ does not converge to a positive definite limit matrix, the penalty term in equation (38) is chosen to ensure that estimator $\check{\boldsymbol{\lambda}}$ gives rise asymptotically to a zero risk premium for all useless and weak factors. As shown in Proposition 2 of Bryzgalova (2015), estimator (38) is oracle in the sense that (i) it consistently selects in finite samples factors that are neither weak nor useless, and (ii) it induces for all other factor risk premia an asymptotic distribution as efficient as the one of an oracle misspecification-robust factor risk premium estimator which knows ex-ante the factors that are neither useless not weak. As a consequence, estimator (38) offers a second obvious benchmark for the finite-sample performance of our approach.[15]

4.2 Monte Carlo simulation setting

We simulate four linear asset pricing models with varying dimension and factor properties:[16]

- Model 1 is a 1-factor model with no useless or weak factor;
- Model 2 (Model 3) is a 3-factor model with two useless (weak) factors and one factor that is neither useless nor weak;
- Model 4 is a higher dimensional model with eleven useless factors, one weak factor and three factors that are neither useless nor weak.

The two weak factors in Model 3 have covariances with returns that are fastly and slowly vanishing at rate $1/T^{3/4}$ and $1/\sqrt{T}$, respectively, while the single weak factor in Model 4 has slowly vanishing covariances at rate $1/\sqrt{T}$.

[15] In our implementation of estimator (38) we use $d = 4$, consistently with the choice in the simulation section of Bryzgalova (2015), and we define the adaptive penalty weights based on sample betas. Results for different choices of tuning parameter d and penalty weights – e.g., based on correlations or partial correlations of factors and returns – are similar to the ones obtained below, and are available upon request.

[16] In Online Appendix 9.2.2, we further report results for the Monte Carlo simulation settings considered in Gospodinov et al. (2014) of a 2-factor model, where the useless factor has mean zero and unit variance, while the factor that is neither useless nor weak has moments calibrated to the market factor.

For each of the above models, we generate a correctly specified and a misspecified version of the model. Returns on the test assets and factors are drawn from a multivariate normal distribution

$$\begin{pmatrix} \boldsymbol{R}_t \\ \boldsymbol{F}_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_R \\ \boldsymbol{\mu}_F \end{pmatrix}, \begin{bmatrix} \boldsymbol{V}_R & \boldsymbol{V}_{RF} \\ \boldsymbol{V}_{FR} & \boldsymbol{V}_F \end{bmatrix} \right) .$$

Moments in this distribution are calibrated to match those of following set of factors and test asset excess returns:

- The five Fama-French factors, momentum, the q-factors in Hou et al. (2015), the intermediary factors of He et al. (2017), the betting-against-beta factors in Frazzini and Pedersen (2014), and the real per capita nondurable consumption growth;
- The 25 double-sorted portfolios constructed on size and book-to-market and 17 industry portfolios.

Data for the moment calibration are obtained from the Kenneth French data library and consist of monthly observations from January 1970 to October 2021 for the considered set of assets and factors, which gives rise to a panel including 624 observations across time. Accordingly, in our Monte Carlo simulation we produce $10'000$ random samples consisting each of 624 time series observations of factors and excess returns.

Regarding factors that are neither weak nor useless, in Model 1–3, the single such factor is calibrated to the market factor. In Model 4, they are calibrated to the three available factors implying the highest sum of squared sample correlations with test asset returns. These are the size factor (R_{ME}), the expected growth factor (R_{EG}) and the profitability factor (R_{ROE}).

Regarding factors that are weak or useless, their means and variances are calibrated to the SMB and HML factors in Model 2–3, and to the twelve factors implying the lowest sum of square sample correlations with test asset returns in Model 4. For these factors, means and marginal variances are calibrated from the data, while the covariances between factors are set to zero. The factor covariances with asset returns are set to zero when simulating useless factors and appropriately scaled by $1/T^{3/4}$ or by $1/\sqrt{T}$ when simulating weak factors with fastly and slowly vanishing covariances with returns.

Our simulations cover both correctly specified and misspecified models. Following, e.g., Gospodinov et al. (2014), the misspecified version of each model is simulated by setting the expected asset return vector $\boldsymbol{\mu}_R$ equal to the sample asset returns means in the data. Instead, correctly specified models are simulated by setting

$$\boldsymbol{\mu}_R = \boldsymbol{\beta}_{\mathcal{S}} \boldsymbol{\lambda}_{\mathcal{S}}^{KRS} = \boldsymbol{\beta}^i \boldsymbol{\lambda}^i \ .$$

Here, $\boldsymbol{\beta}_{\mathcal{S}}$ ($\boldsymbol{\lambda}_{\mathcal{S}}^{KRS}$) is the matrix (vector) of factor betas (misspecification-robust factor risk premia) for the subset of factors that are not useless or weak.[17] All these parameters act as population parameters in our simulations, and are appropriately calibrated to the data. In Models 2–4, which include useless or weak factors, we set $\boldsymbol{\lambda}_{\mathcal{S}^c}^{KRS} = \mathbf{0}$ for the factor risk premia of factors that are weak or useless, as Kan et al. (2013) risk premia are not identified in these models.

Finally, to select the tuning parameter τ_T of our proximal tradable risk premium estimator in a purely data-driven way, we make use of a Generalized Cross Validation (GCV; Wahba, 1990) scheme based on following criterion:

$$\text{GCV} = \frac{\left\| \hat{\boldsymbol{\mu}}_R - \hat{\boldsymbol{\beta}}^i_{\check{\mathcal{S}}} \check{\boldsymbol{\lambda}}^i_{\check{\mathcal{S}}} \right\|_2^2}{(1 - |\check{\mathcal{S}}|/T)^2} \ ,$$

where $\check{\mathcal{S}}$ is the active set estimated by $\check{\boldsymbol{\lambda}}^i$, $\hat{\boldsymbol{\beta}}^i_{\check{\mathcal{S}}} := \hat{\boldsymbol{V}}_{RF_{\check{\mathcal{S}}}} (\hat{\boldsymbol{V}}_{F_{\check{\mathcal{S}}}R} \hat{\boldsymbol{V}}_R^{-1} \hat{\boldsymbol{V}}_{RF_{\check{\mathcal{S}}}})^{-1}$, and $|\check{\mathcal{S}}|$ is the cardinality of set $\check{\mathcal{S}}$, which is a proxy for a model's number of degrees of freedom.[18]

4.3 Finite sample bias-variance tradeoff and factor selection

We start by studying the finite sample bias-variance tradeoffs emerging for the KRS factor risk premium estimator and for our sample, Oracle and Relaxed estimators of tradable factor risk premia. For each simulated model, Figures 1–4 report the decomposition of the estimators' MSE into its squared bias and variance components.[19] Overall we observe that, with the exception of Model 1 where the estimators' MSE are almost identical, the KRS

[17] Recall from Proposition 2.1 that when the model is correctly specified, expected returns are equivalently given by $\boldsymbol{\beta}^i \boldsymbol{\lambda}^i$, where $\boldsymbol{\beta}^i$ ($\boldsymbol{\lambda}^i$) is the matrix (vector) of factor mimicking portfolios return betas (tradable factor risk premia).

[18] When comparing our Oracle estimator to the Pen-FM estimator of Bryzgalova (2015), we further consider 5-fold Cross Validation (CV). Following, e.g., James et al. (2021), we select the tuning parameters of the penalized estimators in our analysis according to the "one-standard-error" rule: we select the most parsimonious model among the models with a CV/GCV score that is not more than one standard error above the CV/GCV score of the best model.

[19] The bias of each estimator is computed with respect to the associated population factor risk premium definition.

estimator's MSE is considerably higher than the one of sample, Oracle and relaxed estimators of tradable factor risk premia. This feature is a direct consequence of the fact that the KRS risk premia of useless and weak factors are not estimated consistently, which gives rise to a substantially larger estimator variance in the box plots of Figures 5–8.[20] When further comparing our sample and Oracle estimators of tradable factor risk premia, the latter estimator gives rise to some degree of bias for the risk premia of factors that are not weak nor useless. This feature is a natural consequence of the shrinkage properties of this estimator. This shrinkage effect also implies that the tradable risk premia of useless and weak factors are estimated with a clearly lower variance, which results in a lower overall MSE with respect to the MSE of sample tradable factor risk premia. Finally, the relaxed estimator successfully removes the finite-sample bias implied by the Oracle estimator, while giving rise to a higher (lower) overall mean square error in misspecified (correctly specified) models.

Given the problematic finite-sample properties of misspecification-robust factor risk premium estimator (37) in weak or not identified models, we next ask how much its Oracle version (38) improves on them. Figure 9 reports the corresponding estimators' MSE decompositions for Models 2 and 4. Overall, we observe that Oracle estimator (38) nicely improves on estimator (37) in the low-dimensional Model 2, by providing comparable MSEs as our Oracle estimator. In contrast, for the higher dimensional Model 4, it still implies large MSEs, which exhibit a large bias component under both cross-validation schemes used. On the other hand, our Oracle estimator gives rise to clearly lower MSEs and virtually no bias when using 5-fold cross-validation. This evidence indicates that estimating precisely misspecification-robust factor risk premia in weakly or not identified asset pricing models may be a challenge also for Oracle-type estimators of these risk premia.

Figure 10 summarizes the finite sample selection properties of our Oracle estimator of tradable factor risk premia, based on a Generalized cross-validation scheme. To this end, we report the Monte Carlo simulation frequency with which the active set estimated by our Oracle estimator equals or contains, respectively, the unknown active subset of factors that are nei-

[20] For brevity, when reporting these results for Models 2–4 in the main text, we focus on correctly specified versions of the models. Similarly, for Model 4 we do not report the evidence for estimated tradable weak factor risk premia. The results for misspecified versions of the models and for the estimated weak factor risk premia in Model 4 are analogous to those in the main text for correctly specified models and for useless factors, respectively. Therefore, they are collected in Online Appendix 9.2.

ther useless nor weak. Overall, we find that our methodology produces a quite reasonable factor selection. Indeed, we find that the finite sample factor selection is virtually exact in Model 1. In Model 2–3 it is correct in more than 98% (around 75-80%) of the cases in correctly specified (misspecified) models, and it virtually always contains the unknown active subset of factors that are neither useless nor weak. Even in the large dimensional Model 4, a correct factor selection is obtained in 95% (74%) of the cases for correctly specified (misspecified) models, while the unknown active subset of factors that are neither useless nor weak is contained in the estimated active set in 95% of the cases

Figure 11 sets the above evidence in relation to the finite sample factor selection properties of Oracle estimator (38) of misspecification-robust factor risk premia. Coherently with the above MSE results, we find that for both cross-validation schemes used this estimator produces a satisfactory factor selection in low-dimensional Model 2. In contrast, it gives rise often to incorrect factor selections in higher dimensional Model 4. On the other hand, we find that 5-fold cross-validation further improves the factor selection properties of our Oracle estimator under Generalized cross-validation, by providing a virtually perfect selection for all models considered.

In summary, we conclude that a reliable estimation and factor selection based on misspecification-robust factor risk premia may be hardly achievable in weak or not identified models, even when using Oracle-type estimators of these risk premia. Nonetheless, a trustworthy Oracle estimation and factor selection can still be produced based on tradable factor risk premia.

4.4 Finite-sample inference

We next document the finite-sample inference properties implied by KRS factor risk premium estimators and by the sample, Oracle and Relaxed estimators of tradable factor risk premia. For the estimators of tradable risk premia, we obtain a feasible finite sample inference based on the asymptotic standard error formulas implied by Propositions 3.1 and 3.3. For the KRS estimator, we rely on the misspecification-robust standard error formula implied by Proposition 2 of the internex appendix of Kan et al. (2013).

Figure 12 reports coverage probabilities for the event that a candidate risk premium parameter λ^* in a neighborhoud of true population risk premium λ_0 is contained in a finite sample 95% confidence interval implied by the different methods. Therefore, when $\lambda^* = \lambda_0$, one minus this probability is the finite sample size of a test for testing null hypothesis $\lambda = \lambda_0$. Conversely, when $\lambda^* \neq \lambda_0$, one minus this probability gives the finite sample power of a test for testing null hypothesis $\lambda = \lambda^*$ against alternative hypothesis $\lambda = \lambda_0$. Since we obtain a similar and coherent evidence for all simulated models, we focus for brevity in the sequel on the evidence produced for the correctly specified versions of Model 4.

A first general pattern is that, relative to misspecification-robust risk premia, sample tradable risk premium estimators imply an improved ability to reject the hypothesis of a non-zero risk premium for weak and useless factors. This feature is reflected in the coverage probabilities in the lower panels of Figure 12. Indeed, while the size behavior of both tests under the null of a zero risk premium for the useless factor gives rise to a conservative empirical size lower than the nominal one, the null hypothesis of a non zero tradable risk premium is correctly rejected with a rapidly increasing probability using sample tradable risk premia, as the distance between the null of a nonzero risk premium and the alternative of a zero risk premium increases. In contrast, the same null hypothesis for KRS misspecification-robust risk premia is rarely rejected, even for quite large absolute risk premia under the null. The well-known conservative behavior of the inference implied by KRS misspecification-robust risk premia for testing the null of a zero risk premium for the useless factors, is the foundation of the sequential testing procedure proposed in Gospodinov et al. (2014) for consistently eliminating these factors from an asset pricing model. In our methodology, the consistent elimination of all useless and weak factors is performed jointly by our Oracle estimator of tradable factor risk premia. This feature is directly visible in the lower panels of Figure 12, by the fact that such estimator produces in virtually almost all simulations a zero estimated tradable risk premium.

The evidence in the upper panels of Figure 12 suggests that the inference regarding the tradable risk premia of factors that are not weak nor useless is unaffected by the presence of weak or useless factors in an asset pricing model. This feature holds exactly asymptotically for our estimators, as shown in Propositions 3.1 and 3.3. In contrast, Gospodinov et al. (2014) show that the same feature does not apply to estimators of KRS misspecification-robust risk premia.[21] Indeed, as displayed in the upper panels of Figure 12, the finite sample inference implied by sample tradeable risk premia is well-behaved and sharper than the one implied by misspecification-robust risk premia, in the sense that their coverage probabilities are decaying faster as we move away from the corresponding population risk premium. The coverage probabilities of our Oracle estimator present a similar sharp

[21] This is the reason why they propose in the first place a sequential testing procedure for consistently eliminating the useless factors from an asset pricing model.

behavior, but they are shifted with respect to the population tradable risk premium due to the finite sample bias induced by the shrinkage effect. As shown in the figure, this finite sample bias can be reduced by means of the Relaxed estimator.

5 Empirical analysis

We exploit our factor selection and inference methodology based on the Oracle proximal estimator of Definition 2, in order to explore and analyze a wide range of empirical asset pricing models generated from the factor zoo. With this analysis, we ideally aim to first isolate a relevant subset of low-dimensional, well-identified asset pricing models. Second, we aim to understand whether a robust subset of factors exists, which consistently appears in such subset of well-identified asset pricing models. Finally, if such a subset of factors exists, we can try to study which of these factors are priced.

5.1 Empirical setting

Our analysis is based on a universe of potential factors given by the 51 tradable and nontradable factors compiled in Bryzgalova et al. (2023).[22] As test assets, we consider the 25 portfolios sorted on size and book-to-market and the 17 industry portfolios. Section 9.3 of the Online Appendix reports results for test assets including the 25 portfolios sorted on size and book-to-market and the 25 portfolios sorted on operating profitability and investment.

All data on test assets is sourced from the Kenneth French data library. Our dataset consists of monthly observations of test asset excess returns and risk factors collected from October 1973 to December 2016.[23] Using this dataset, we build randomized families of factor models having factor dimension growing from 1 to 10. To these families of models, we apply our factor selection methodology based on tradable factor risk premia. Intuitively, models of higher dimension are more likely to be weakly or not identified. In this respect, a main goal of our analysis is to understand

[22] We are grateful to the authors for making their data publicly accessible. Detailed descriptions of this factor universe can be found in their Table B.1 and Table IA.XIII from their Online Appendix.

[23] Since tradable risk premia satisfy the identity $\boldsymbol{\lambda}^i = \boldsymbol{V}_F^{1/2}\, \mathbb{C}\mathrm{or}[\boldsymbol{F}, \boldsymbol{R}] \boldsymbol{V}_R^{-1/2} \boldsymbol{\mu}_R$, we scale all factors so that they have a unit sample variance, in order not to spuriously select factors with a large variance.

whether our methodology is effective in mitigating these identification problems, exclusively by consistently pruning the useless and the weak factors, in a way that may ideally give rise to a selected subset of economically relevant and well-identified lower dimensional models.

In order to verify the identification properties of the various candidate models in our empirical study, we make use of the rank test proposed in Chen and Fang (2019).[24] We perform Chen and Fang (2019) test both before and after our consistent selection procedure for factors that are neither weak nor useless is applied. Precisely, let $\boldsymbol{\beta} = \boldsymbol{V}_{RF}\boldsymbol{V}_F^{-1}$ be the $N \times K$ matrix of factor betas of a set of candidate factors, before our selection procedure has been applied. With Chen and Fang (2019) test we first test the null hypothesis:

$$\mathcal{H}_0: \ \text{Rank}(\boldsymbol{\beta}) \leq K - 1 \ , \tag{39}$$

against the alternative hypothesis $\text{Rank}(\boldsymbol{\beta}) = K$, using the corresponding matrix $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{V}}_{RF}\hat{\boldsymbol{V}}_F^{-1}$ of estimated factor betas. Let further $\boldsymbol{\beta}_{\mathcal{S}} = \boldsymbol{V}_{RF_{\mathcal{S}}}\boldsymbol{V}_{F_{\mathcal{S}}}^{-1}$ be the matrix of factor betas of factors that are neither useless nor weak. We then test the null hypothesis

$$\mathcal{H}_{0\mathcal{S}}: \ \text{Rank}(\boldsymbol{\beta}_{\mathcal{S}}) \leq |\mathcal{S}| - 1 \ , \tag{40}$$

against the alternative hypothesis $\text{Rank}(\boldsymbol{\beta}_{\mathcal{S}}) = |\mathcal{S}|$, using the estimated cardinality $|\hat{\mathcal{S}}|$ of active set $\mathcal{S}$ and the corresponding matrix $\hat{\boldsymbol{\beta}}_{\hat{\mathcal{S}}} = \hat{\boldsymbol{V}}_{RF_{\hat{\mathcal{S}}}}\hat{\boldsymbol{V}}_{F_{\hat{\mathcal{S}}}}^{-1}$ of estimated factor betas after the factor selection has been performed.

By comparing the rank test results before and after performing the factor selection, we can finally analyze whether our methodology is effective in mitigating potential identification problems, exclusively by pruning the useless and the weak factors. Moreover, by inspecting the distribution of selected factors across randomized models, we can study whether a robust subset of economically relevant, well-identified low dimensional asset pricing models appears. Finally, we can study the pricing properties of asset pricing factors appearing in such well-identified low dimensional models.

5.2 Identification evidence pre- and post-factor screening

We first quantify the degree to which our consistent factor selection methodology can help improve the identification of a candidate asset pricing model.

[24] This test first employs the iterative Kleibergen and Paap (2006) rank test procedure to compute a first-step rank estimate. In a second step, it implements a bootstrap procedure that allows to directly test the null hypothesis of a reduced column rank in the matrix of factor betas of a candidate asset pricing model.

Intuitively, if a model's weak or no identification is exclusively due to the existence of some useless or weak factors, then our methodology is able to consistently select an associated set of well-identified submodels, each including exclusively factors that are neither useless nor weak. In such a situation, the rejection frequency of the null of no identification by Chen and Fang (2019) test should be substantially larger after applying our factor selection procedure over randomized factor models from the factor zoo. Moreover, after performing the factor selection, a potential identification problem should depend to a less extent on the initial dimension of an asset pricing model.

Figure 13 reports rejection frequencies of the null of no identification by Chen and Fang (2019) test, before and after applying our factor selection methodology, for different randomization procedures over factor models of dimension between 5 and 10 from the factor zoo.[25] Each null rejection is based on asymptotic critical values for a nominal size $\alpha = 0.05$. Note that this analysis is not aimed to directly identify a subset of well-identified models from an initial larger set of models in a multiple hypothesis testing problem, but rather to quantify how often a test of identification applied to a randomly selected model would indicate an identification problem. Therefore, we do not introduce a correction for multiple testing at this stage, even though such a correction could be in principle developed using existing methodologies in the literature.

The upper panels of Figure 13 show that the pre-screening identification frequency decreases rapidly in the number of candidate factors of an asset pricing model. Moreover, this frequency is lower for the testing setting based on the lower dimensional set of test assets. For instance, identification frequencies decrease from about 0.95 (0.99) for single factor models, to about 0.05 for eight, nine and ten (ten) factor models, when test assets comprise the 25 size/book-to-maket (25 size/book-to-maket and the 17 industry) portfolios. This evidence is partly expected, as the probability of observing a reduced column rank in the matrix of factor betas of randomly selected models may be naturally higher for larger dimensional models and for testing frameworks based on less test assets. However, it crucially indicates that the identification problem may appear quite frequently already across randomly selected low dimensional models including no more than, e.g., five factors.

A clearly distinct picture emerges for the post-screening identification frequencies, which are uniformly higher than the pre-screening identifica-

[25] Results for lower dimensions are reported in the supplemental appendix.

tion frequency, in many cases by a lare extent. For instance, the post-screening identification frequency decrease from about 0.95 (0.99) for single factor models to about 0.65 (0.8) for ten factor models, when test assets comprise the 25 size/book-to-maket (25 size/book-to-maket and the 17 industry) portfolios. This evidence suggests that our factor screening methodology is indeed effective in selecting factors that are neither useless nor weak. Moreover, it indicates that after the consistent selection of these factors the weak or no identification problem emerges much less frequently across randomly selected models.

The lower panels of Figure 13 further consider model randomizations that always include the market factor in the initial model. Also in this case, the pre-screening identification frequencies feature a dramatic degradation as the asset pricing model size increases. In contrast, all post-screening identification frequencies are always very large and above 0.9 (0.95) even for ten factor models, when test assets comprise the 25 size/book-to-maket (25 size/book-to-maket and the 17 industry) portfolios. This additional evidence further confirms that our screening methodology is quite successful in selecting factors that are neither useless nor weak and thereby largely improving the identification properties of empirical asset pricing models.

5.3 Pre- and post-screening factor space dimension

While the post-screening dimension of an asset pricing model is by construction not larger than the pre-screening dimension, the different pre- and post-screening identification properties documented above have to be related to strictly lower model dimensions after screening. This feature may imply that the typical dimension of an identified asset pricing models from the factor zoo is relatively low. To understand this issue properly, we study in Figure 14 the distribution of post-screening factor space dimensions emerging across randomized asset pricing models.

The upper panels of Figure 14 show that, as expected, the post-screening factor space dimension increases with the initial dimension of an asset pricing model. More importantly, regardless of the pre-screening dimension, lower post-selection factor space dimensions occur with higher frequencies than higher dimensions. For instance, for settings with five initial factors, we find that models with at most one, at most two, and more than four selected factor(s) occur around 30% (60%), 50% (73%) and 30% (16%) of the cases, respectively, when the test assets comprise the the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios. The post-screening factor dimension further shrinks uniformly for model randomizations that always include the market factor. For instance, for settings with five initial factors, we find that models with at most one, at most two, and more than four selected factor(s) occur 33% (30%), 77% (73%) and 6% (8%) of the cases, respectively, when the test assets comprise the 25 size/book-to-market portfolios (the 25 size/book-to-market and the 17 industry portfolios).[26]

Overall, we conclude that the most likely factor space dimensions of identified asset pricing models from the factor zoo are relatively low, and even more so for models including the market factor. As a consequence, standard tests of asset pricing models based on two-step cross-sectional regression methodologies are likely to be harmed by a weak identification problem already when testing low-dimensional models with, e.g., between five and seven factors.

5.4 Oracle factor selection and factor types

The previous evidence of a low post-screening factor space dimension suggests that the majority of identified asset pricing models from the factor zoo is relatively low-dimensional. Therefore, we next ask whether we can detect asset pricing factors that most consistently appear in the subset of identified low-dimensional models.

Figure 15 reports factor marginal selection frequencies, with which every individual risk factor is selected from a randomized initial model after screening, as a function of the model's dimension. The upper panels of Figure 15 show that factors such as the market, the SMB, the ICR in He et al. (2017), and the market factor with a hedged unpriced component in Daniel et al. (2020) (MKT_star) are those with the highest and very consistent marginal selection frequencies across randomized models and test asset choices. For instance, when using the 25 size/book-to-market and 17 industry portfolios, the marginal selection frequencies of these factors when randomizing 10-factor models are all 100%, except those for factor SMB which has a selection frequency of 93.4%. Analogously, when using the 25 size/book-to-market portfolios, these selection frequencies are 99.4% and 92.9% for SMB and MKT_star, respectively, and 100% for the other factors. Overall, we conclude that this set of factors is unlikely to generate an identification problem when included in an asset pricing model from the factor zoo.

[26] Analogous evidence is obtained in Section 9.3 of the Online Appendix for test assets including the 25 size/book-to-market portfolios and 25 portfolios sorted on operating profitability and investment.

A second tier of factors display high, but less consistent, marginal selection frequency across randomized models and choices of test assets. These factors may be crowd out in some identified models by other factors more strongly co-moving with test asset returns. When using the 25 size/book-to-market and 17 industry portfolios as test assets, there are just a few of these factors, namely the long-term reversal factor in Jegadeesh and Titman (2001) (LTRev), the liquidity factor in Pástor and Stambaugh (2003) (LIQ_NT), nondurable consumption (NONDUR) and the systematic skeweness factor in Langlois (2020) (SKEW). When using just the 25 size/book-to-market portfolios as test assets, additional such factors are the SMB factor with a hedged unpriced component in Daniel et al. (2020) (SMB_star), the long-term behavioral factor in Daniel et al. (2020) (BEH_FIN), the profitability factor measured by gross profits-to-assets in Novy-Marx (2013) (GR_PROF), the mispricing factor in Stambaugh and Yuan (2017) (MGMT), and the distress risk factor in Campbell et al. (2008) (DISSTR).

An even sharper evidence arises in the lower panels of Figure 15, for model randomizations that always include the market factor. Here, factors including SMB, ICR, MKT_star and MGMT are those having by far the highest and most consistent marginal selection frequencies, with, e.g., selection frequencies above 97% when randomizing ten-factor asset pricing models. Second tier factors with high, but less consistent, marginal selection frequencies include BEH_FIN, the composite equity issuance in Daniel and Titman (2006) (COMP_ISSUE) and the conservative minus aggressive in Fama and French (2015) (CMA). For randomizations of ten-factor models, these factors have marginal selection frequencies of 89.1% (85.4%), 69.3% (89.2%) and 58% (73.4%), respectively, when using the 25 size/book-to-market (25 size/book-to-market and 17 industry) portfolios as test assets. A similar, and to some extent stronger, evidence is obtained in Section 9 of the Online Appendix for test assets including the 25 size/book-to-market portfolios and 25 portfolios sorted on operating profitability and investment. Here, all these factors have marginal selection frequencies above 95% when randomizing ten-factor asset pricing models.

Overall, we conclude that models including robust asset pricing factors MKT, SMB, ICR, MKT_star, and MGMT, together with some of the factors BEH_FIN, COMP_ISSUE and CMA, are natural well-identified benchmarks for explaining the cross-section of test assets' expected returns.

5.5 Tradable factor risk premia

Given the above evidence regarding the set of well-identified candidate models and corresponding factors for explaining the cross-section of asset returns, we next expore the associated tradable factor risk premia. Exploiting the properties of our Oracle estimator in Definition 2, we can test whether a factor gives rise to a tradeable risk premium component. Indeed, recall that conditionally on a factor being identified as neither useless nor weak by our Oracle factor selection, we can still make use of the marginal asymptotic distributions from Proposition 3.3 to test the null hypothesis of a zero individual tradeable risk premium. Since these asymptotic distributions are independent of the dimension of the factor space in an asset pricing model, this inference is also invariant across randomized models.

Figure 16 reports the resulting inference, with estimated tradable factor risk premia and their confidence intervals ranked according to a factor's selection frequency. In the set of robust asset pricing factors MKT, SMB, ICR, MKT_star, and MGM with highest selection frequencies, we find that they all exhibit a statistically significant tradable risk premium at the 10% significance level, for both sets of test assets. The tradeable risk premium of factor CMA is as well significant for both sets of test assets, while the risk premium of factor COMP_ISSUE is significant for the 25 size/book-to-market and 17 industry portfolios. In contrast, factor BEH_FIN, is (marginally) insignificant at the 10% significance level for both sets of test assets. An analogous, and to some extent stronger, evidence is obtained in Section 9 of the Online Appendix for test assets including the 25 size/book-to-market portfolios and 25 portfolios sorted on operating profitability and investment. Here, all tradeable risk premia of the above factors are significant at the 10% level and each of these factors is selected in more than 97% of randomized ten-factor models.

Overall, this evidence points to a robust subset of economically relevant and well-identified low dimensional models from the factor zoo, which can be built out of selected factors giving rise to a nonzero tradable risk premium.

5.6 Misspecification-robust factor risk premia

Given the above evidence regarding the set of low-dimensional well-identified candidate models, we next explore the properties of misspecification-robust factor risk premia estimated by two-step cross-sectional regression methods. In doing so, we build on the fact that our Oracle factor selection ensures a good identification of asset pricing models after screening.

Recall that a factor has different misspecification-robust and tradable risk premia only when it is partly unspanned by test asset returns. Furthermore, the latter risk premium depends in general on the factor composition of a model. Therefore, misspecification-robust risk premia in general imply different factor risk premium estimators across asset pricing models. Figures 17–18 summarize the composite post screening effects of unspanned factor risks for such factor risk premium estimates, across well-identified and potentially misspecified models from the factor zoo.

We find that while factors MKT, SMB, MGMT, CMA and COMP_ISSUE imply a relatively robust statistical significance of estimated risk premia, factors like ICR, MKT_star, and to some extent BEH_FIN, give rise to a uniformly weak significance. In addition, factors such as ICR, MKT_star or CMA give rise to a large variability of estimated risk premia across models, which is hardly interpretable economically. This evidence emerges despite the above findings that all these factors consistently appear in the subset of well-identified asset pricing models with a statistically significant tradable risk premium.

Overall, we conclude that the misspecification of identifiable models from the factor zoo generates a challenge for interpreting notions of risk premia allocating a non zero price to unspanned factor risks. In contrast, tradable factor risk premia are economically-founded in general and can be studied empirically based on our Oracle inference framework.

6 Conclusions

This paper addresses the challenge of testing misspecified and weakly (or not) identified asset pricing models. Existing research has emphasized the limitations of two-step testing procedures for factor risk premia defined by the negative factor covariance with a Stochastic Discount Factor (SDF) projection on the factor space. In such settings, misspecification introduces non-zero pricing errors that affect the asymptotic distribution of estimated factor risk premia, while weak (or no) identification hinders the unique identification of the candidate SDF projection, and, consequently, the associated factor risk premia. In contrast, our inference approach relies on the well-established concept of a tradable factor risk premium, which is defined by the negative factor covariance with the SDF projection on the asset return space. We show that tradable factor risk premia give rise to a robust and well-defined framework for Oracle inference in potentially misspecified or weakly identified asset pricing models.

Our methodology consistently identifies the set of factors that are weakly correlated with returns and simultaneously enables a reliable and efficient inference on the risk premia of all other factors in an asset pricing model. We achieve this by leveraging the properties of a simple sample version of tradeable factor risk premia, which exhibits an asymptotically Gaussian distributions even in models including factors weakly correlated with returns. Since this distribution may imply some asymptotic biases only for the estimated risk premia of weakly correlated factors, we develop an Oracle estimator of tradeable factor risk premia that overcomes these distortions.

Our Oracle estimator is derived through a closed-form penalized minimum distance correction of sample tradable risk premia. It improves upon existing estimators by producing a consistent identification of the set of weak factors and simultaneously generating an efficient asymptotic distribution for the estimated risk premia of other factors. By design, our Oracle estimator directly provides a valid inference for the tradable risk premia of factors that are not weakly correlated with returns. Alternatively, it can be used for a consistent preliminary screening of potentially weakly correlated factors, which enhances the identification of a model and hence facilitates the reliable application of efficient two-step cross-sectional tests of asset pricing models in a subsequent step of the analysis.

We make use of our Oracle estimation and inference methodology for tradable factor risk premia, in order to build a coherent and easily applicable framework for studying the asset pricing properties of a broad class of factor models from the factor zoo. For this purpose, we first form randomized families of candidate models of largest factor space dimension, in order to single out a set of well-identified submodels consisting exclusively of factors that are neither useless nor weak. Based on this set, we pin down the properties of the resulting distribution of factors and factor risk premia across well-identified models, for various benchmark choices of test assets in the literature.

Our empirical analysis reveals that the proposed factor selection procedure greatly enhances model identification. This allows us to detect a robust subset of economically relevant and well-identified models that are all built out of factors with a nonzero tradable risk premium. Such factors include market, size, intermediaries capital ratio, market with a hedged unpriced component, a mispricing factor, a long-term behavioral factor, a liquidity factor, a conservative minus aggressive factor and a composite equity issuance factor. At the same time, we find that the relatively low factor space dimension of well-identified models is associated with some degree of misspecification. This feature harms the interpretation of other established notions of a factor risk premium in the literature, which may allocate a non zero price to factor risks that are not tradable.

References

Attouch, H. (1984). *Variational convergence for functions and operators*, Volume 1. Pitman Advanced Publishing Program.

Bryzgalova, S. (2015). Spurious factors in linear asset pricing models. *LSE manuscript 1*(3).

Bryzgalova, S., J. Huang, and C. Julliard (2023). Bayesian solutions for the factor zoo: We just ran two quadrillion models. *The Journal of Finance 78*(1), 487–557.

Campbell, J. Y., J. Hilscher, and J. Szilagyi (2008). In search of distress risk. *The Journal of finance 63*(6), 2899–2939.

Chen, Q. and Z. Fang (2019). Improved inference on the rank of a matrix. *Quantitative Economics 10*(4), 1787–1824.

Daniel, K., D. Hirshleifer, and L. Sun (2020). Short-and long-horizon behavioral factors. *The review of financial studies 33*(4), 1673–1736.

Daniel, K., L. Mota, S. Rottke, and T. Santos (2020). The cross-section of risk and returns. *The Review of Financial Studies 33*(5), 1927–1979.

Daniel, K. and S. Titman (2006). Market reactions to tangible and intangible information. *The Journal of Finance 61*(4), 1605–1643.

Davidson, J. (1992). A central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes. *Econometric theory 8*(3), 313–329.

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

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

Fan, J. and R. Li (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association 96*(456), 1348–1360.

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

Geyer, C. J. (1994). On the asymptotics of constrained m-estimation. *The Annals of statistics*, 1993–2010.

Gospodinov, N., R. Kan, and C. Robotti (2014). Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors. *The Review of Financial Studies 27*(7), 2139–2170.

Grinblatt, M. and S. Titman (1987). The relation between mean-variance efficiency and arbitrage pricing. *Journal of Business*, 97–112.

He, Z., B. Kelly, and A. Manela (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics 126*(1), 1–35.

Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies 28*(3), 650–705.

Huberman, G., S. Kandel, and R. F. Stambaugh (1987). Mimicking portfolios and exact arbitrage pricing. *The Journal of Finance 42*(1), 1–9.

James, G., D. Witten, T. Hastie, R. Tibshirani, G. James, D. Witten, T. Hastie, and R. Tibshirani (2021). Statistical learning. *An introduction to statistical learning: with applications in R*, 15–57.

Jegadeesh, N. and S. Titman (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of finance 56*(2), 699–720.

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

Kan, R. and C. Zhang (1999). Two-pass tests of asset pricing models with useless factors. *the Journal of Finance 54*(1), 203–235.

Kandel, S. and R. F. Stambaugh (1995). Portfolio inefficiency and the cross-section of expected returns. *The Journal of Finance 50*(1), 157–184.

Kleibergen, F. (2009). Tests of risk premia in linear factor models. *Journal of econometrics 149*(2), 149–173.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics 133*(1), 97–126.

Kleibergen, F. and Z. Zhan (2015). Unexplained factors and their effects on second pass r-squared's. *Journal of Econometrics 189*(1), 101–116.

Langlois, H. (2020). Measuring skewness premia. *Journal of Financial Economics 135*(2), 399–424.

Lehmann, B. N. and D. M. Modest (1988). The empirical foundations of the arbitrage pricing theory. *Journal of financial economics 21*(2), 213–254.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of financial economics 108*(1), 1–28.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political economy 111*(3), 642–685.

Quaini, A. and F. Trojani (2022). Proximal estimation and inference. *arXiv preprint arXiv:2205.13469*.

Shanken, J. (1992). On the estimation of beta-pricing models. *The review of financial studies 5*(1), 1–33.

Shanken, J. and G. Zhou (2007). Estimating and testing beta pricing models: Alternative methods and their performance in simulations. *Journal of Financial Economics 84*(1), 40–86.

Stambaugh, R. F. and Y. Yuan (2017). Mispricing factors. *The review of financial studies 30*(4), 1270–1315.

Wahba, G. (1990). *Spline models for observational data*. SIAM.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association 101*(476), 1418–1429.

7 Simulations: figures

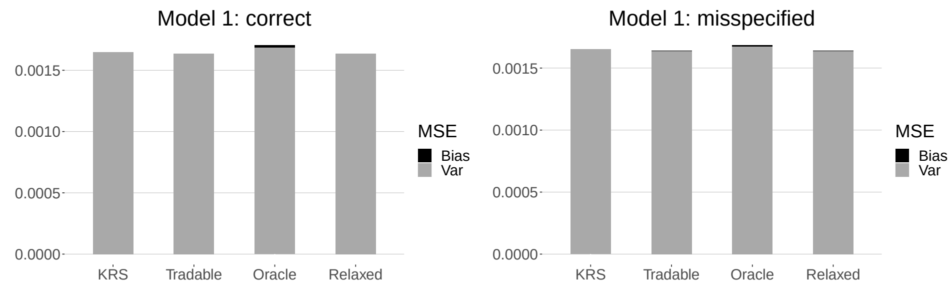

Figure 1: **Estimators' MSE decomposition, Model 1**: the figure presents the MSE bias-variance decomposition of KRS, tradable, Oracle and Relaxed factor risk premia estimators in simulation Model 1, for both correctly specified (left panel) and misspecified (right panel) version.

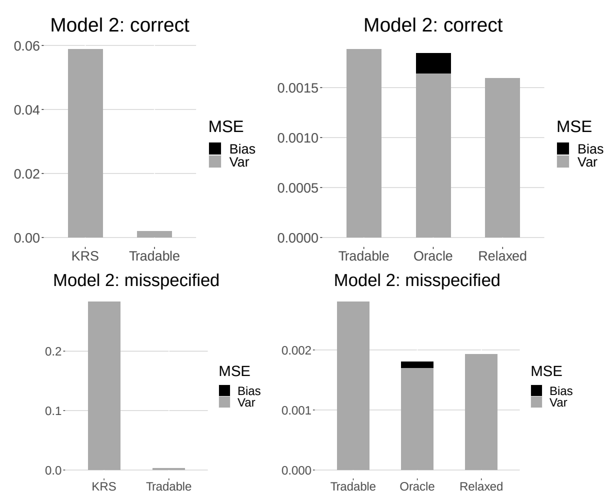

Figure 2: **Estimators' MSE decomposition, Model 2**: the figure presents the MSE bias-variance decomposition of KRS (left panels), tradable (in both left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators in simulation Model 2, for both correctly specified (upper panels) and misspecified (lower panel) version.

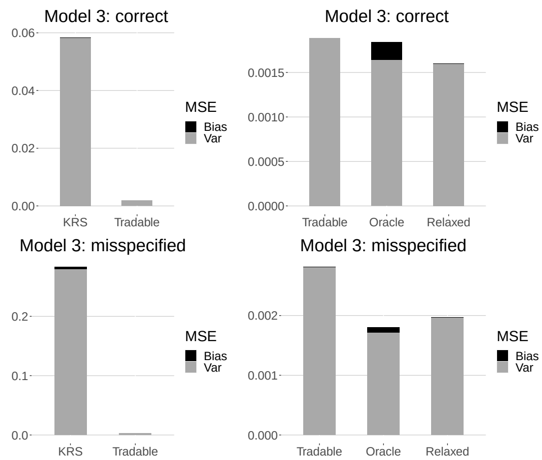

Figure 3: **Estimators' MSE decomposition, Model 3**: the figure presents the MSE bias-variance decomposition of KRS (left panels), tradable (in both left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators in simulation Model 3, for both correctly specified (upper panels) and misspecified (lower panels) version.

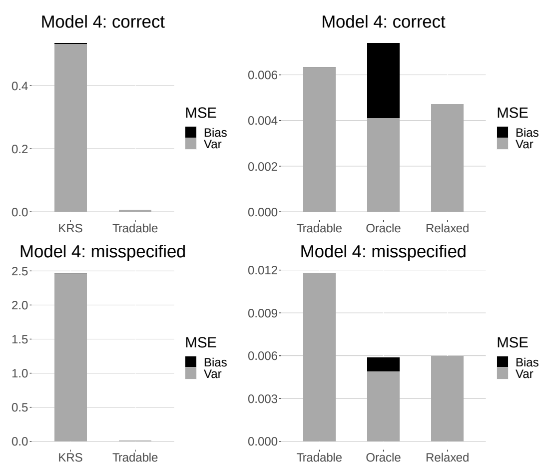

Figure 4: **Estimators' MSE decomposition, Model 4**: the figure presents the MSE bias-variance decomposition of KRS (left panels), tradable (in both left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators in simulation Model 4, for both correctly specified (upper panels) and misspecified (lower panels) version.

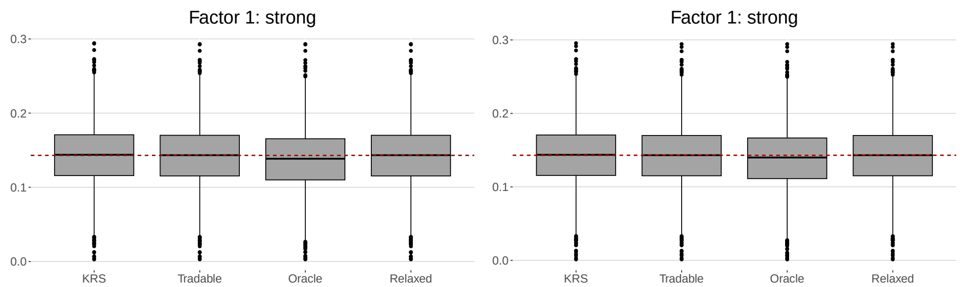

Figure 5: **Estimators' simulated distribution, Model 1**: the figure presents the simulated distribution of KRS, tradable, Oracle and Relaxed factor risk premia estimators of the strong factor of simulation Model 1, for both correctly specified (left panel) and misspecified (right panel) version. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

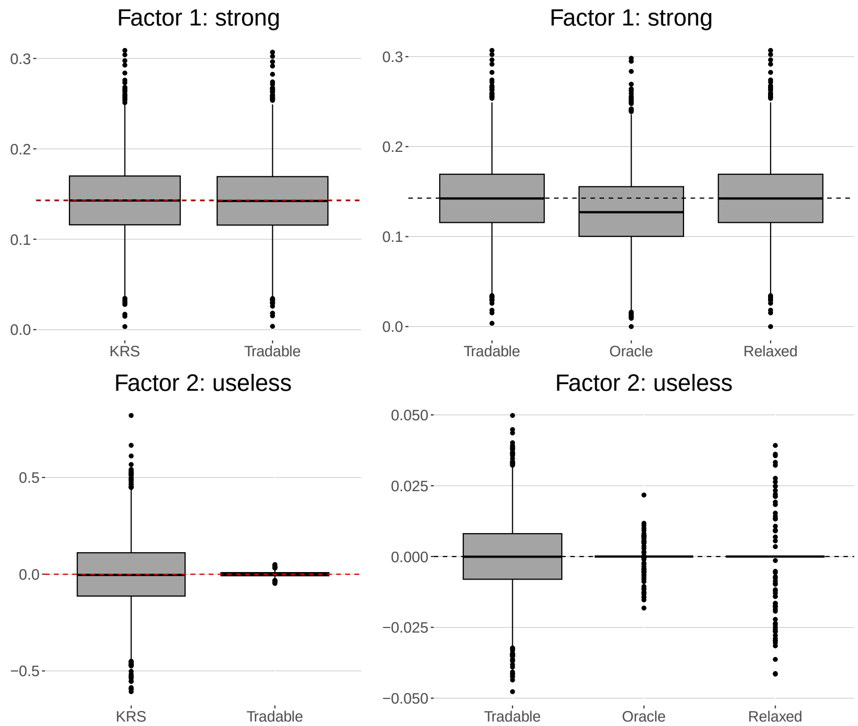

Figure 6: **Estimators' simulated distribution, Model 2 (correctly specified)**: the figure presents the marginal simulated distributions of KRS (left panels), tradable (left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators of the strong (upper panel) and one of the two useless (lower panel) factors of the correctly specified version of simulation Model 2. We omit the figure of the second useless factor as it is qualitatively identical to the one of the first useless factor. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

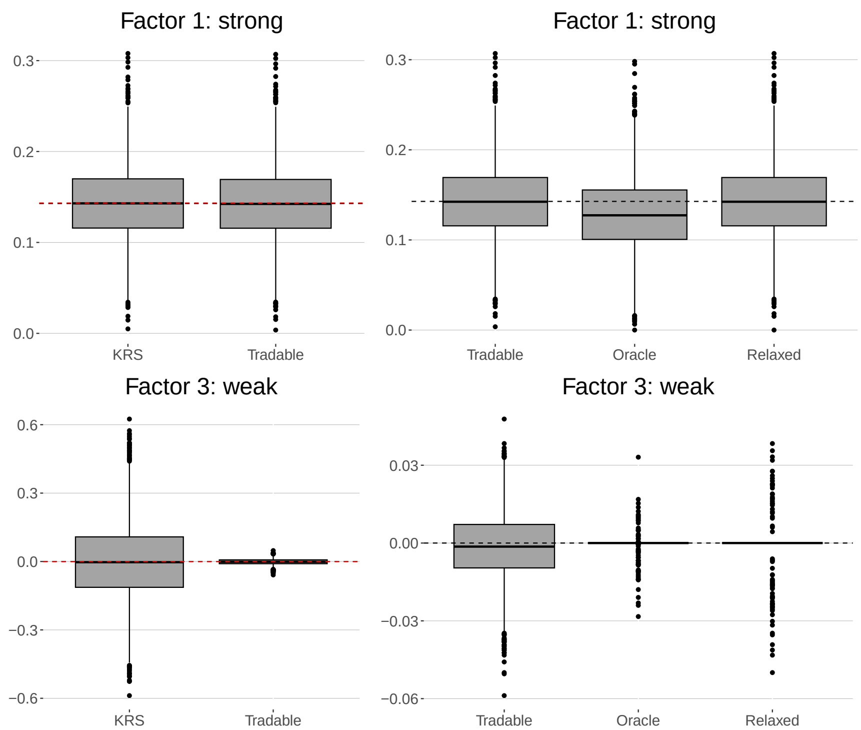

Figure 7: **Estimators' simulated distribution, Model 3 (correctly specified)**: the figure presents the marginal simulated distributions of KRS (left panels), tradable (left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators of the strong (upper panels)and the $1/\sqrt{T}-$weak (lower panels) factors of the correctly specified version of simulation Model 3. We omit the figure of the $1/T^{3/4}-$weak factor as it is qualitatively identical to the one of the $1/\sqrt{T}-$weak factor. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

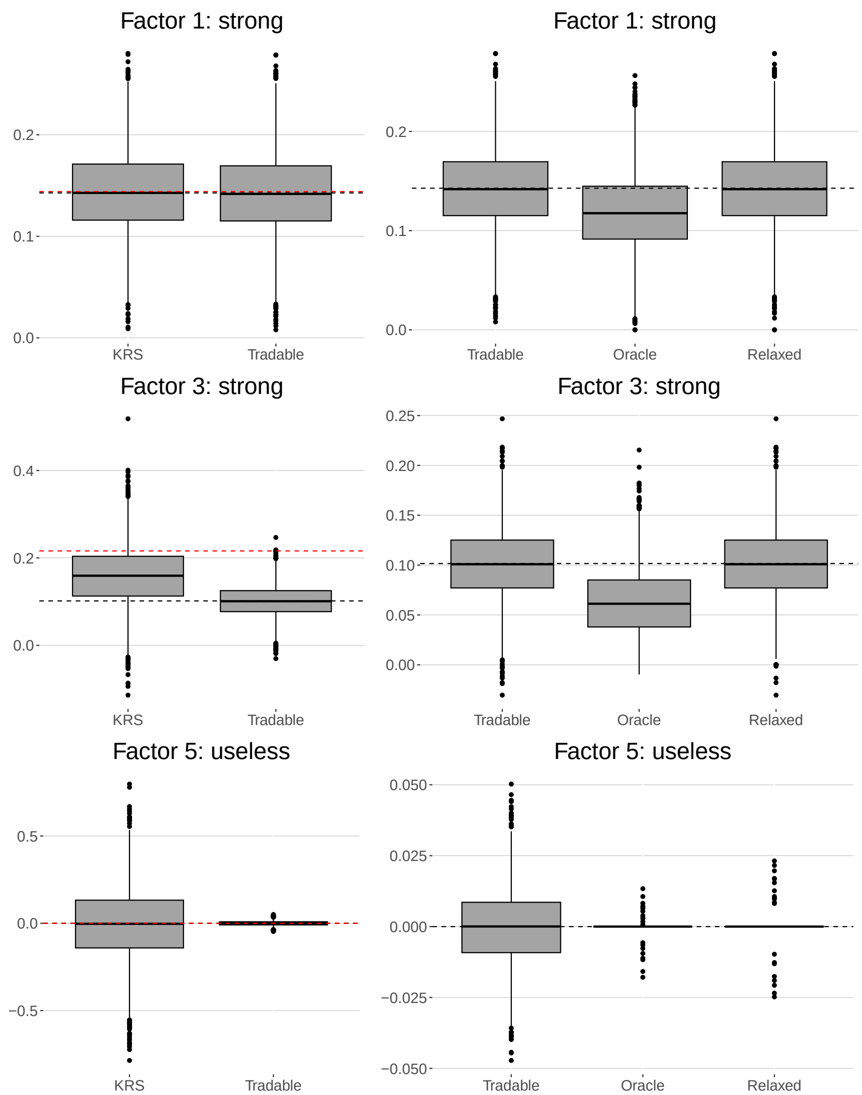

Figure 8: **Estimators' simulated distribution, Model 4 (correctly specified)**: the figure presents the marginal simulated distributions of KRS (left panels), tradable (left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators of the first strong (upper panels), the third strong (central panels) and the first useless (lower panels) factors of the correctly specified version of simulation Model 4. We omit the figure of the other useless factors and of the weak factor as they are qualitatively identical to the one of the first useless factor. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

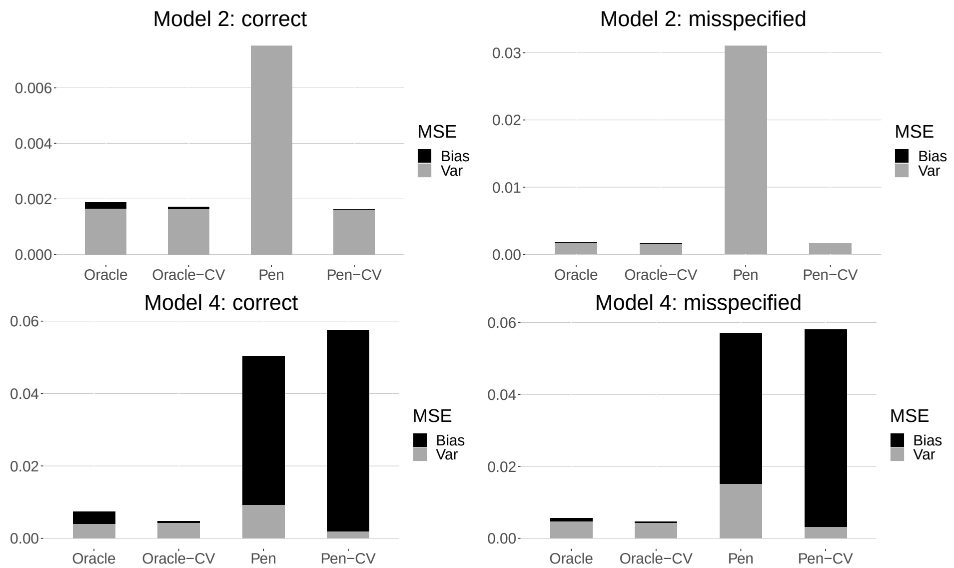

Figure 9: **Estimators' MSE decomposition, Models 2 and 4**: MSE bias-variance decomposition of the Pen-FM (Bryzgalova (2015)) and our Oracle factor risk premia estimators, for both correctly specified (left panel) and misspecified (right panel) versions of simulation Models 2 (upper panels) and 4 (lower panels). Pen and Oracle (Pen-CV and Oracle-CV) are tuned using Generalized Cross Validation (5-fold Cross Validation).

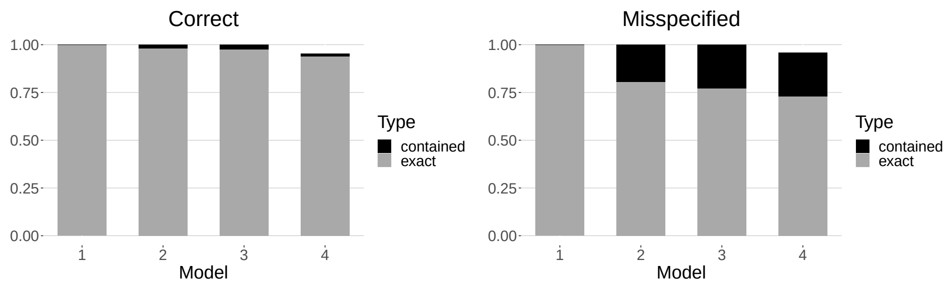

Figure 10: **Oracle factor risk premia selection properties, Models 1–4**: the figure presents the "exact" and "contained" variable selection properties of our Oracle factor risk premia estimator for both correctly specified (left panel) and misspecified (right panel) versions of simulation Models 1–4. The "exact" ("contained") variable selection property is computed as the frequency over the simulation runs that the active set of the population tradable factor risk premia coefficient equals (is contained in) the estimated active set of our Oracle risk premia estimator.

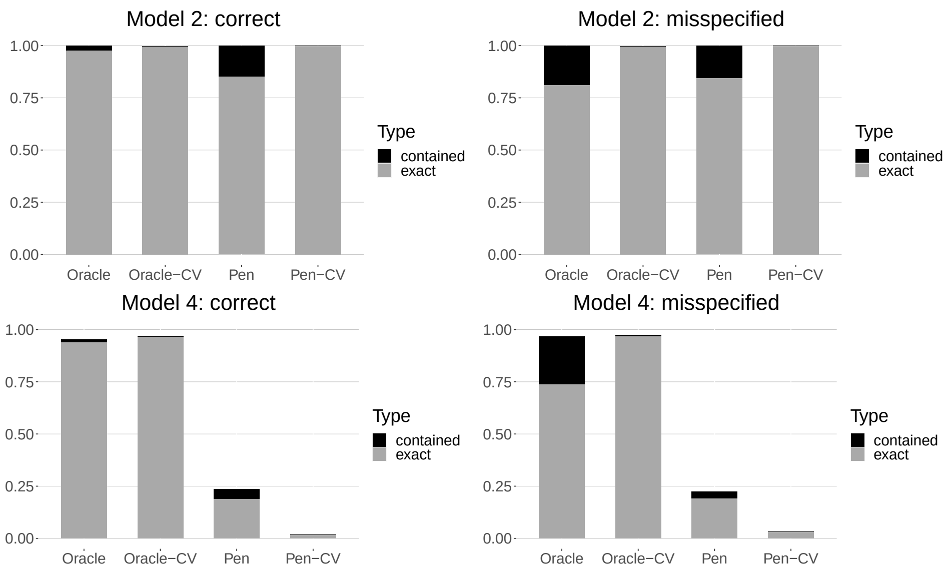

Figure 11: **Factor risk premia selection comparison, Models 2 and 4**: the figure presents the "exact" and "contained" variable selection properties of the Pen-FM (Bryzgalova (2015)) and our Oracle factor risk premia estimators, for both correctly specified (left panel) and misspecified (right panel) versions of simulation Models 2 (upper panels) and 4 (lower panels). The "exact" ("contained") variable selection property is computed as the fraction over the simulation runs that the active set of the KRS (tradable) population factor risk premia coefficient equals (is contained in) the estimated active set of the Pen-FM (our Oracle) estimator. Pen and Oracle (Pen-CV and Oracle-CV) are tuned using Generalized Cross Validation (5-fold Cross Validation).

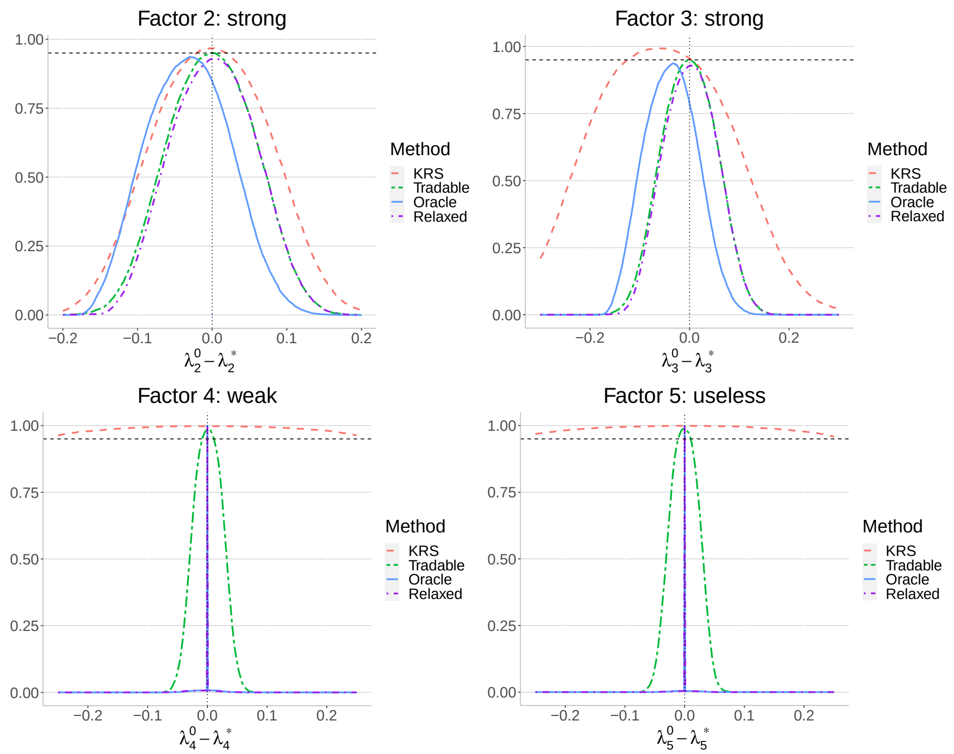

Figure 12: **Inclusion frequencies of estimators' 95% confidence intervals, Model 4 (correctly specified)**: the figure presents the frequencies with which various values of factor's j risk premium λ_j^* around the population coefficient λ_j^0 are contained in the corresponding KRS (red dashed line), tradable (green dashed-dotted line), Oracle (solid blue line) and Relaxed (dashed-dotted purple line) estimators' 95% confidence interval across the simulation runs, for the first strong factor (upper left panel, $j = 1$), the second strong factor (upper right panel, $j = 2$), the $1/\sqrt{T}-$weak factor (lower left panel, $j = 4$) and the first useless factor (lower right panel, $j = 5$) of the correctly specified version of simulation Model 4. To facilitate the comparison, we display these frequences as a function of the deviation from the estimators' corresponding population risk premia coefficient λ_j^0, which is the population KRS risk premia coefficient for the KRS estimator, and the population tradable risk premia coefficient for the tradable, Oracle and Relaxed estimators. The horizontal dashed black line represents the nominal level of the confidence intervals, namely 95%.

8 Empirics: figures

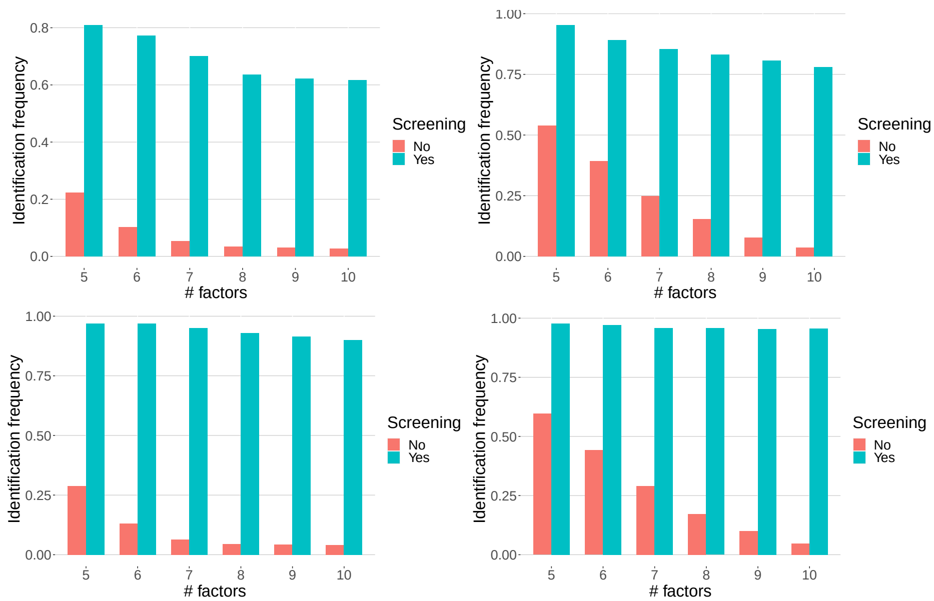

Figure 13: **Model identification frequencies**: Frequency of null hypothesis (39) rejections under a significance level $\alpha = 0.05$ by the Chen and Fang (2019) test, across randomized factor models including 5–10 initial factors. The red (blue) bars indicate model identification frequencies before (after) having applied our Oracle factor selection. The upper (lower) panels report the model identification frequencies for selections with no (the market as) ubiquitous factor across models. The left (right) panels report the results for test assets comprising the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios. For models having 1 (2) randomized factors, we consider all 52 (1326) possible model combinations. For models including more than 2 randomized factors, we consider 10'000 random factor combinations.

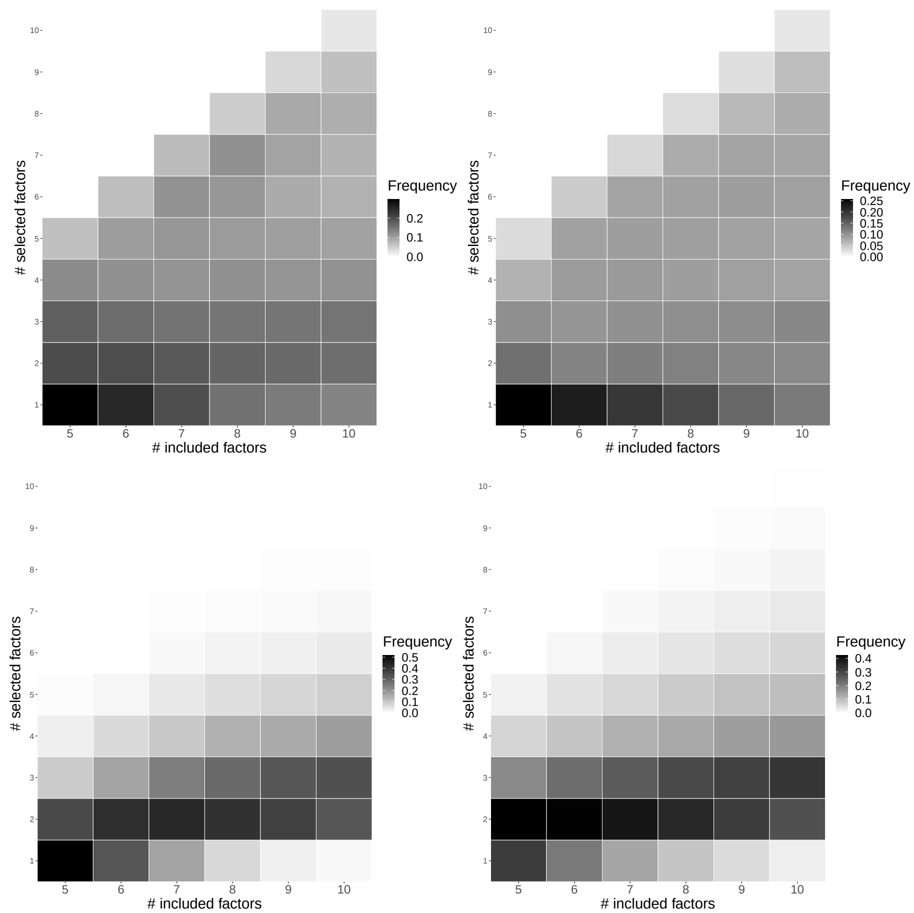

Figure 14: **Post-selection model size**: Frequency of post-selection model dimensions, i.e., number of factors selected by our Oracle factor selection, across randomized factor models including 5–10 initial factors. The red (blue) bars indicate model identification frequencies before (after) having applied our Oracle proximal factor selection. The upper (lower) panels report the model identification frequencies for selections with no (the market as) ubiquitous factor across models. The left (right) panels report the results for test assets comprising the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios.

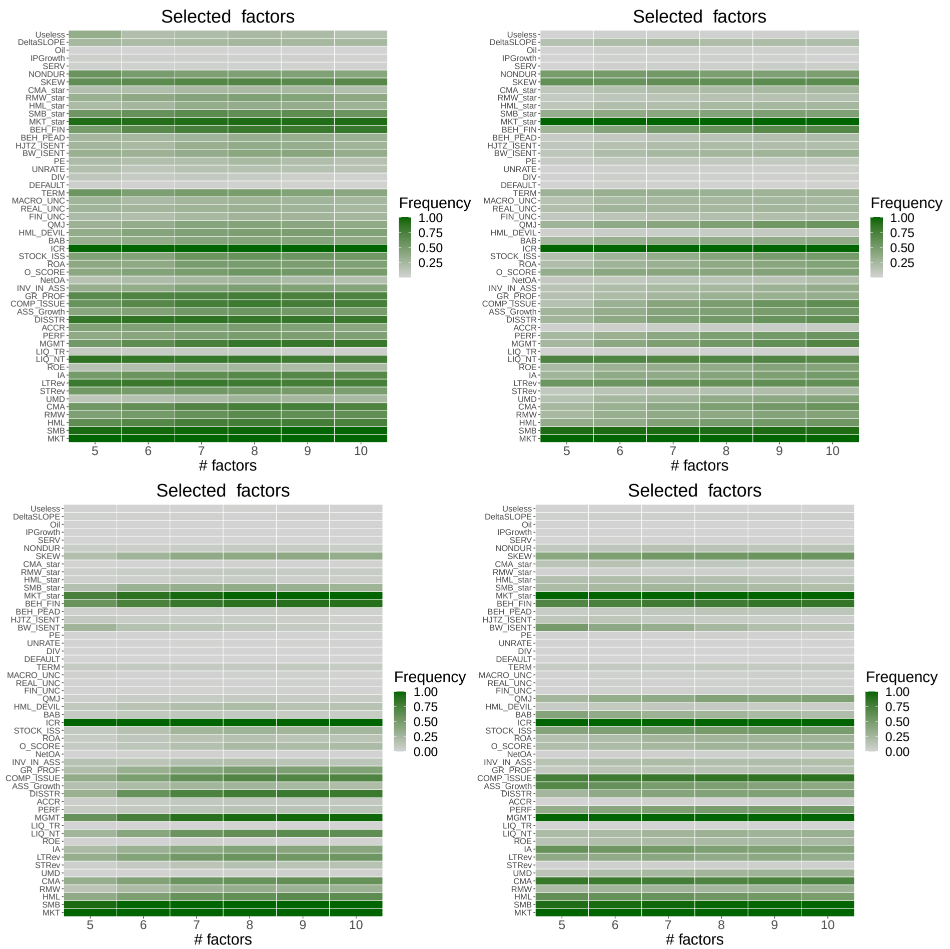

Figure 15: **Factor selection frequencies**: Selection frequencies of individual factors for our Oracle factor selection across randomized factor models including 5–10 initial number of factors. The upper (lower) panels report the frequencies of the post-selection model dimensions for selections with no (the market as) ubiquitous factor across models. The left (right) panels report the results for test assets comprising the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios.

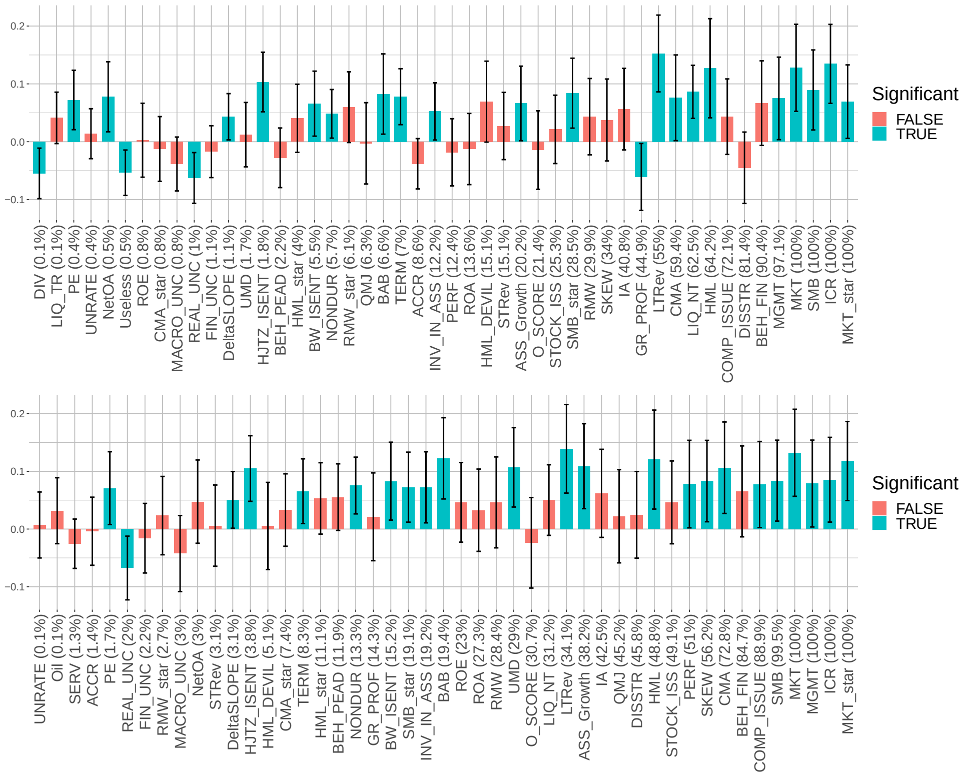

Figure 16: **Tradable factor risk premia**: Point estimates and confidence intervals at confidence level 90% of sample tradable factor risk premia. Factors are ordered by their selection frequency in randomized 10-factor models always including the market, reported in parenthesis, and only factors with positive selection frequency are retained. The upper (lower) panel reports the results obtained with test assets comprising the 25 size/book-to-market (25 size/book-to-market and the 17 industry portfolios).

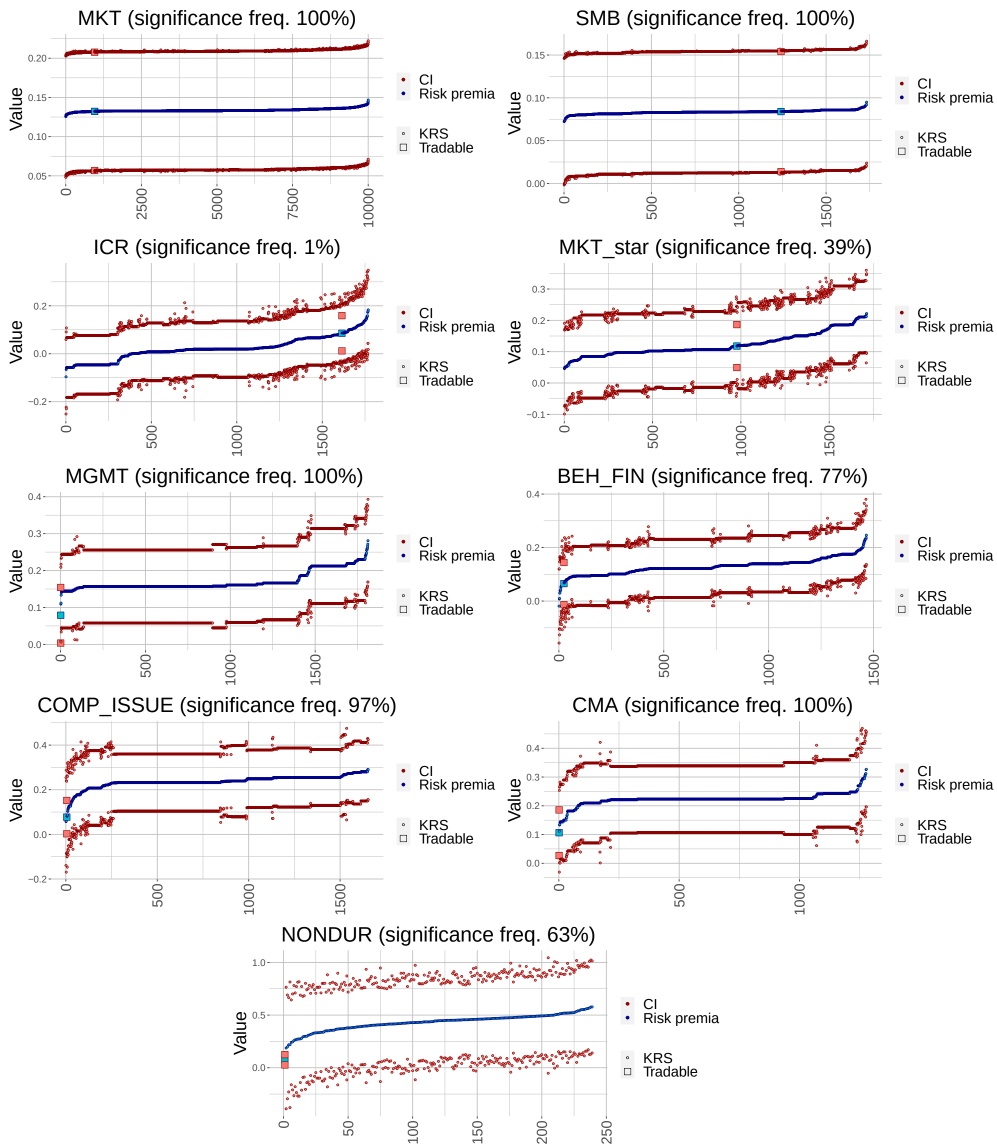

Figure 17: **Factor risk premia and** 90% **confidence intervals**: Misspecification-robust (represented by dots) and tradable (represented by squares) factor risk premia and corresponding 90% confidence intervals of a selection of factors over various randomized 10-factor models always including the market. The significance frequencies are reported in parenthesis, next to the factor's label. The results are obtained with test assets comprising the 25 size/book-to-market and the 17 industry portfolios.

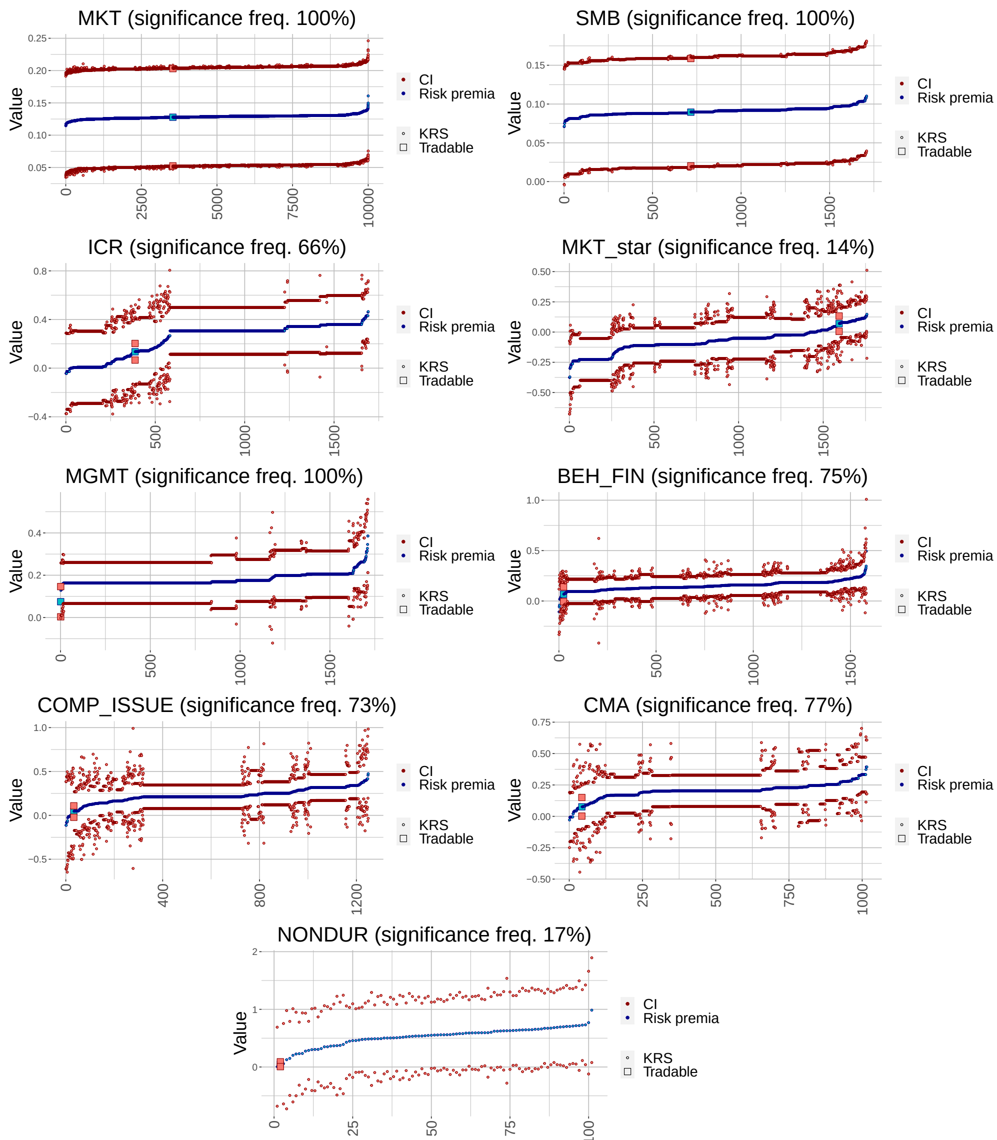

Figure 18: **Factor risk premia and** 90% **confidence intervals**: Misspecification-robust (represented by dots) and tradable (represented by squares) factor risk premia and corresponding 90% confidence intervals of a selection of factors over various randomized 10-factor models always including the market. The significance frequencies are reported in parenthesis, next to the factor's label. The results are obtained with test assets comprising the 25 size/book-to-market portfolios.

9 Online Appendix

9.1 Proofs of the mathematical results in the main text

Proof of Proposition 2.1. M_F is a valid SDF if and only if $\boldsymbol{\mu}_R = \boldsymbol{V}_{RF}\boldsymbol{\gamma}$, which implies:

$$\boldsymbol{\lambda}^i = \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF}\boldsymbol{\gamma} \ . \tag{41}$$

Furthermore, using the definition of generalized inverse, it follows:

$$\boldsymbol{V}_{FR} = \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF}\boldsymbol{\beta}^{i\prime} \ . \tag{42}$$

Together, this gives:

$$\boldsymbol{\mu}_R = \boldsymbol{V}_{RF}\boldsymbol{\gamma} = \boldsymbol{\beta}^i\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF}\boldsymbol{\gamma} = \boldsymbol{\beta}^i\boldsymbol{\lambda}^i \ . \tag{43}$$

Conversely, if $\boldsymbol{\mu}_R = \boldsymbol{\beta}^i\boldsymbol{\lambda}^i$, it follows:

$$\boldsymbol{\mu}_R = \boldsymbol{\beta}^i\boldsymbol{\lambda}^i = \boldsymbol{V}_{RF}(\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^+\boldsymbol{\lambda}^i =: \boldsymbol{V}_{RF}\boldsymbol{\gamma} \ , \tag{44}$$

with $\boldsymbol{\gamma} := (\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^+\boldsymbol{\lambda}^i$. This concludes the proof. □

Proof of Proposition 2.2. Under Assumption 1 we obtain, if matrix $\boldsymbol{V}_{RF}$ has full column rank:

$$\boldsymbol{\lambda}(\boldsymbol{V}_R^{-1}) = \boldsymbol{V}_F(\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^{-1}\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R = \boldsymbol{V}_F(\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^{-1}\boldsymbol{\lambda}^i \ .$$

To prove the second statement of the proposition, note first that if a factor F_j is tradable, then it is given without loss of generality by the payoff $\tilde{F}_j$ of its factor mimicking portfolio:

$$\tilde{F}_{jt} = \boldsymbol{V}_{F_jR}\boldsymbol{V}_R^{-1}\boldsymbol{R}_t \ .$$

Therefore, $\boldsymbol{V}_F\boldsymbol{e}_j = \boldsymbol{V}_{\tilde{F}}\boldsymbol{e}_j$, where $\boldsymbol{e}_j$ denotes the $j-$unit vector, $\tilde{\boldsymbol{F}}$ the vector of all factor-mimicking portfolio payoffs, and

$$\boldsymbol{V}_{\tilde{F}} = \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF} \ ,$$

its variance-covariance matrix. Using statement (14) of the proposition, this finally gives:

$$\boldsymbol{e}_j'\boldsymbol{\lambda}(\boldsymbol{V}_R^{-1}) = \boldsymbol{e}_j'\boldsymbol{V}_F(\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{V}_{RF})^{-1}\boldsymbol{\lambda}_i = \boldsymbol{e}_j'\boldsymbol{\lambda}_i \ .$$

If additionally the correct specification property from Proposition 2.1 holds, then factor risk premium $\boldsymbol{\lambda}(\boldsymbol{W})$ is independent of weighting matrix $\boldsymbol{W}$ and we can set without loss of generality $\boldsymbol{\lambda}(\boldsymbol{W}) = \boldsymbol{\lambda}(\boldsymbol{I})$. This gives:

$$\begin{aligned}\boldsymbol{\lambda}(\boldsymbol{W}) =& \boldsymbol{V}_F(\boldsymbol{V}_{FR}\boldsymbol{V}_{RF})^{-1}\boldsymbol{V}_{FR}\boldsymbol{\mu}_R \\ =&(\boldsymbol{\beta}'\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\boldsymbol{\mu}_R \\ =&(\boldsymbol{\beta}'\boldsymbol{\beta})^{-1}\boldsymbol{\beta}'\boldsymbol{\beta}^i\boldsymbol{\lambda}^i \ .\end{aligned}$$

Using the correct specification property from Proposition 2.1 and the fact that matrix $\boldsymbol{\beta}^i$ has full column rank under the given assumptions, we further obtain:

$$\boldsymbol{\lambda}^i = (\boldsymbol{\beta}^{i\prime}\boldsymbol{\beta}^i)^{-1}\boldsymbol{\beta}^{i\prime}\boldsymbol{\beta}^i\boldsymbol{\lambda}^i = (\boldsymbol{\beta}^{i\prime}\boldsymbol{\beta}^i)^{-1}\boldsymbol{\beta}^{i\prime}\boldsymbol{\mu}_R \ . \tag{45}$$

This concludes the proof. $\square$

Proof of Proposition 3.1. Given the assumptions on stochastic process $\boldsymbol{Y}_t = [\boldsymbol{F}_t', \boldsymbol{R}_t{}']'$, terms $\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1}$, $\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR}$ and $\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R$ are all $O_{\mathrm{Pr}}(1/\sqrt{T})$. Therefore, we can write:[27]

$$\begin{aligned}\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i =& \hat{\boldsymbol{V}}_{FR}\hat{\boldsymbol{V}}_R^{-1}\hat{\boldsymbol{\mu}}_R - \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \\ =&\boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R) + \boldsymbol{V}_{FR}(\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1})\boldsymbol{\mu}_R \\ &+ (\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR})\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R + O_{\mathrm{Pr}}(1/T) \ .\end{aligned}$$

Moreover:[28]

$$\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1} = \boldsymbol{V}_R^{-1}(\boldsymbol{V}_R - \hat{\boldsymbol{V}}_R)\boldsymbol{V}_R^{-1} + O_{\mathrm{Pr}}(1/T) \ .$$

This further gives, up to terms of order $O_{\mathrm{Pr}}(1/T)$:

$$\begin{aligned}\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i &= \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R) - \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}(\hat{\boldsymbol{V}}_R - \boldsymbol{V}_R)\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \\ &\quad + (\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR})\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \\ &= \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R) - \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\hat{\boldsymbol{V}}_R\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R + \hat{\boldsymbol{V}}_{FR}\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \ .\end{aligned}$$

Finally, note that:

$$\hat{\boldsymbol{V}}_R = \frac{1}{T}\sum_{t=1}^{T}\bar{\boldsymbol{R}}_t\bar{\boldsymbol{R}}_t' - (\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R)(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R)' = \frac{1}{T}\sum_{t=1}^{T}\bar{\boldsymbol{R}}_t\bar{\boldsymbol{R}}_t' + O_{\mathrm{Pr}}(1/T) \ .$$

[27] Here, the term $O_{\mathrm{Pr}}(1/T)$ collects $(\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR})(\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1})(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R)$, $(\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR})(\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1})\boldsymbol{\mu}_R$, $(\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR})\boldsymbol{V}_R^{-1}(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R)$ and $\boldsymbol{V}_{FR}(\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1})(\hat{\boldsymbol{\mu}}_R - \boldsymbol{\mu}_R)$.

[28] Here, the term $O_{\mathrm{Pr}}(1/T)$ consists in $(\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1})(\boldsymbol{V}_R - \hat{\boldsymbol{V}}_R)\boldsymbol{V}_R^{-1}$.

Similarly:

$$\hat{\boldsymbol{V}}_{FR} = \frac{1}{T}\sum_{t=1}^{T} \bar{\boldsymbol{F}}_t \bar{\boldsymbol{R}}_t' + O_{\mathrm{Pr}}(1/T) \ .$$

Overall, we thus obtain:

$$\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{h}_t + O_{\mathrm{Pr}}(1/T) \ , \tag{46}$$

where

$$\boldsymbol{h}_t := \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\bar{\boldsymbol{R}}_t - \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\bar{\boldsymbol{R}}_t\bar{\boldsymbol{R}}_t'\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R + \bar{\boldsymbol{F}}_t\bar{\boldsymbol{R}}_t'\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \ , \tag{47}$$

defines a stationary and ergodic process with zero expectation and finite second moments. Using the Central Limit Theorem for stationary and ergodic processes, we conclude:

$$\sqrt{T}(\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i) \to_d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \ , \tag{48}$$

where

$$\boldsymbol{\Sigma} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{\infty}\sum_{m=1}^{\infty} \mathbb{E}[\boldsymbol{h}_t\boldsymbol{h}_m'] \ . \tag{49}$$

This concludes the proof. □

Proof of Proposition 3.2. To prove the statement of the proposition, we follow the same steps as in the proof of Proposition 3.1, noting that under the given assumptions on stochastic process $\boldsymbol{Y}_t = [\boldsymbol{F}_t', \boldsymbol{R}_t{}']'$, quantities $\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_R$, $\hat{\boldsymbol{V}}_R^{-1} - \boldsymbol{V}_R^{-1}$ and $\hat{\boldsymbol{V}}_{FR} - \boldsymbol{V}_{FR}$ are still all $O_{\mathrm{Pr}}(1/\sqrt{T})$. Therefore,

$$\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{h}_t + O_{\mathrm{Pr}}(1/T) \ , \tag{50}$$

where

$$\boldsymbol{h}_t = \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\bar{\boldsymbol{R}}_t - \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\bar{\boldsymbol{R}}_t\bar{\boldsymbol{R}}_t'\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R + \bar{\boldsymbol{F}}_t\bar{\boldsymbol{R}}_t'\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \ . \tag{51}$$

However the distribution of random vector $\boldsymbol{h}_t$ is different from the one obtained in Proposition 3.1. Indeed,

$$\boldsymbol{h}_t = \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}\bar{\boldsymbol{R}}_t - \boldsymbol{V}_{FR}\boldsymbol{V}_R^{-1}(\bar{\boldsymbol{R}}_t\bar{\boldsymbol{R}}_t' - \boldsymbol{V}_R)\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \ + (\bar{\boldsymbol{F}}_t\bar{\boldsymbol{R}}_t' - \boldsymbol{V}_{FR})\boldsymbol{V}_R^{-1}\boldsymbol{\mu}_R \ ,$$

where, for any $\alpha > 1/2$:

$$\begin{aligned} \bar{\boldsymbol{F}}_t \bar{\boldsymbol{R}}_t' - \boldsymbol{V}_{FR} &= \bar{\boldsymbol{F}}_t \bar{\boldsymbol{R}}_t' - \boldsymbol{V}_{FR}^{(T)} + \boldsymbol{\Gamma}'/T^{\alpha} + \boldsymbol{\Delta}'/\sqrt{T} \\ &= \bar{\boldsymbol{F}}_t \bar{\boldsymbol{R}}_t' - \boldsymbol{V}_{FR}^{(T)} + \boldsymbol{\Delta}'/\sqrt{T} + o_{\mathrm{Pr}}(1/\sqrt{T}) \ . \end{aligned}$$

Therefore, using the Central Limit Theorem for near-epoch dependent processes, we obtain:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{h}_t \to_d \mathcal{N}\left(\boldsymbol{\lambda}(\boldsymbol{\Delta}), \boldsymbol{\Sigma} - \boldsymbol{\lambda}(\boldsymbol{\Delta})\boldsymbol{\lambda}(\boldsymbol{\Delta})'\right) \ , \tag{52}$$

where $\boldsymbol{\lambda}(\boldsymbol{\Delta}) := \boldsymbol{\Delta}' \boldsymbol{V}_R^{-1} \boldsymbol{\mu}_R$. This concludes the proof. □

Proof of Proposition 3.3. Under the assumptions of Proposition 3.2, it follows:

$$\sqrt{T}(\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i) \to_d \boldsymbol{\xi} \sim \mathcal{N}\left(\boldsymbol{\lambda}(\boldsymbol{\Delta}), \boldsymbol{\Sigma} - \boldsymbol{\lambda}(\boldsymbol{\Delta})\boldsymbol{\lambda}(\boldsymbol{\Delta})'\right) \ . \tag{53}$$

Moreover,

$$\sqrt{T}(\check{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i) = \arg \min_{\boldsymbol{u} \in \mathbb{R}^K} \{g_T(\boldsymbol{u}) + q_T(\boldsymbol{u})\} \ , \tag{54}$$

where

$$g_T(\boldsymbol{u}) := \frac{1}{2} \left|\left| \sqrt{T}\left(\hat{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i\right) - \boldsymbol{u} \right|\right|_2^2 \ , \tag{55}$$

and

$$q_T(\boldsymbol{u}) := T\tau_T \left(f_T\left(\boldsymbol{\lambda}^i + \frac{\boldsymbol{u}}{\sqrt{T}}\right) - f_T(\boldsymbol{\lambda}^i) \right) \ , \tag{56}$$

where $f_T(\boldsymbol{\lambda}^i) = \sum_{k=1}^{K} |\lambda_k^i| / \|\hat{\boldsymbol{\rho}}_k\|_2^2$. Therefore,

$$g_T(\boldsymbol{u}) \to_d \frac{1}{2} ||\boldsymbol{\xi} - \boldsymbol{u}||_2^2 =: g_0(\boldsymbol{u}) \ , \tag{57}$$

uniformly on compact sets. Moreover, using the functional form of the Adaptive Lasso penalty f_T and the fact that, under the assumptions of Proposition 3.2, $\hat{\boldsymbol{\rho}}_{\mathcal{S}} \to_{\mathrm{Pr}} \boldsymbol{\rho}_{\mathcal{S}} \neq \boldsymbol{0}$, $\hat{\boldsymbol{\rho}}_{\mathcal{S}^c} \to_{\mathrm{Pr}} \boldsymbol{0}$ and that $\hat{\boldsymbol{\rho}}_{\mathcal{S}^c} = O_{\mathrm{Pr}}(1/\sqrt{T})$:

$$q_T(\boldsymbol{u}) \to_{\mathrm{Pr}} q_0(\boldsymbol{u}) := \begin{cases} +\infty & \boldsymbol{u}_{\mathcal{S}^c} \neq \boldsymbol{0} \\ 0 & \mathit{else} \end{cases} \ , \tag{58}$$

in epigraph. Using (Attouch, 1984, Thm. 2.15) and Geyer (1994), we thus obtain:

$$g_T(\boldsymbol{u}) + q_T(\boldsymbol{u}) \to_d g_0(\boldsymbol{u}) + q_0(\boldsymbol{u}) \ , \tag{59}$$

in epigraph, and:[29]

$$\sqrt{T}(\check{\boldsymbol{\lambda}}^i - \boldsymbol{\lambda}^i) \to_d \underset{\{\boldsymbol{u}:\boldsymbol{u}_{\mathcal{S}^c}=\boldsymbol{0}\}}{\operatorname{argmin}} \{g_0(\boldsymbol{u})\} = \begin{pmatrix} \boldsymbol{\xi}_{\mathcal{S}} \\ \boldsymbol{0} \end{pmatrix} \ . \tag{60}$$

Since the components of factor vector $\boldsymbol{F}_t$ with index in $\mathcal{S}$ are strong factors, the distribution of $\boldsymbol{\xi}_{\mathcal{S}}$ is unaffected by the presence of useless or weak factors, i.e.:

$$\boldsymbol{\xi}_{\mathcal{S}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\mathcal{S}}) \ , \tag{61}$$

where, from the definition of variance covariance matrix $\boldsymbol{\Sigma}$ in Proposition 3.1:

$$\boldsymbol{\Sigma}_{\mathcal{S}} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{m=1}^{T} \mathbb{E}[\boldsymbol{h}_{t\mathcal{S}} \boldsymbol{h}'_{m\mathcal{S}}] \ . \tag{62}$$

This proves the first statement in Proposition 3.2. The second statement follows from the first statement using standard proofs of model selection consistency in Zou (2006) for the Adaptive Lasso penalty. This concludes the proof. □

[29] The equality in equation (58) holds after a reordering of the vector components.

9.2 Additional Monte Carlo evidence

9.2.1 Additional evidence on Monte Carlo settings for models in the main text

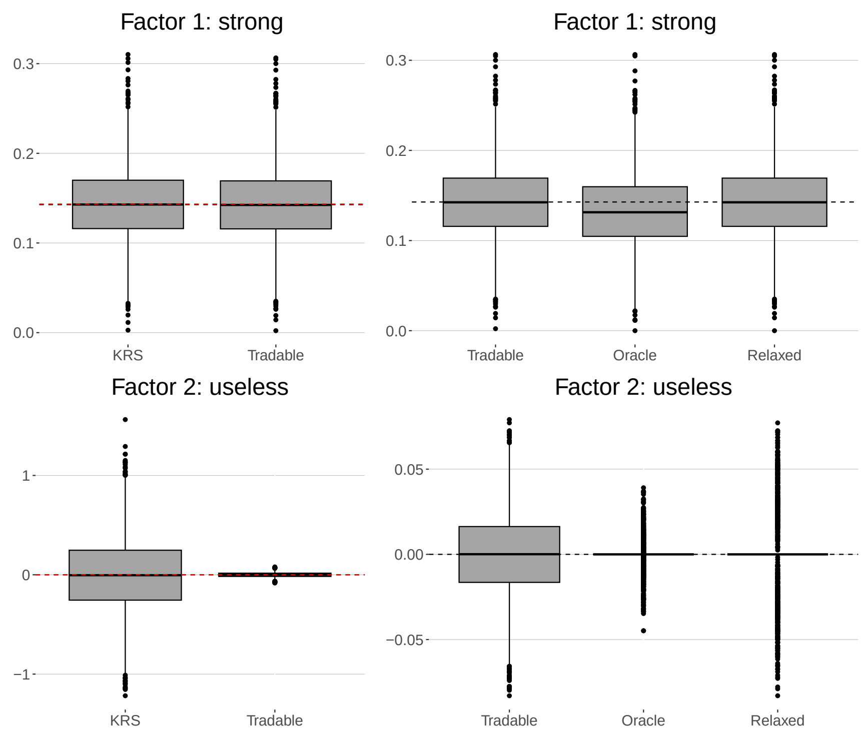

Figure 19: **Estimators' simulated distribution, Model 2 (misspecified)**: the figure presents the marginal simulated distributions of KRS (left panels), tradable (left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators of the strong (upper panel) and one of the two useless (lower panel) factors of the misspecified version of simulation Model 2. We omit the figure of the second useless factor as it is qualitatively identical to the one of the first useless factor. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

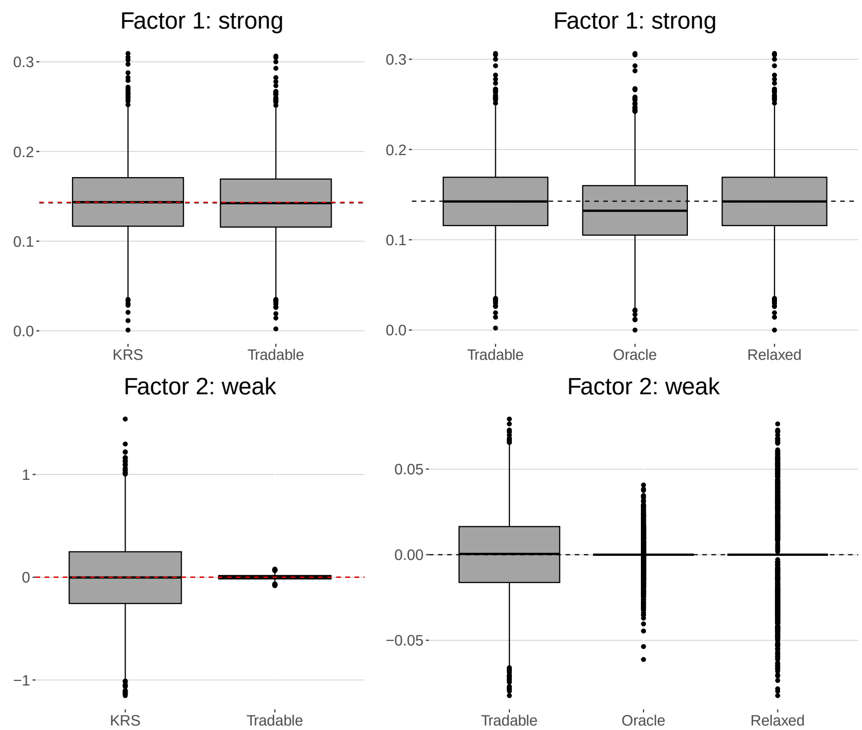

Figure 20: **Estimators' simulated distribution, Model 3 (misspecified)**: the figure presents the marginal simulated distributions of KRS (left panels), tradable (left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators of the strong (upper panels), the $1/T^{3/4}-$weak (central panels) and the $1/\sqrt{T}-$weak (lower panels) factors of the misspecified version of simulation Model 3. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

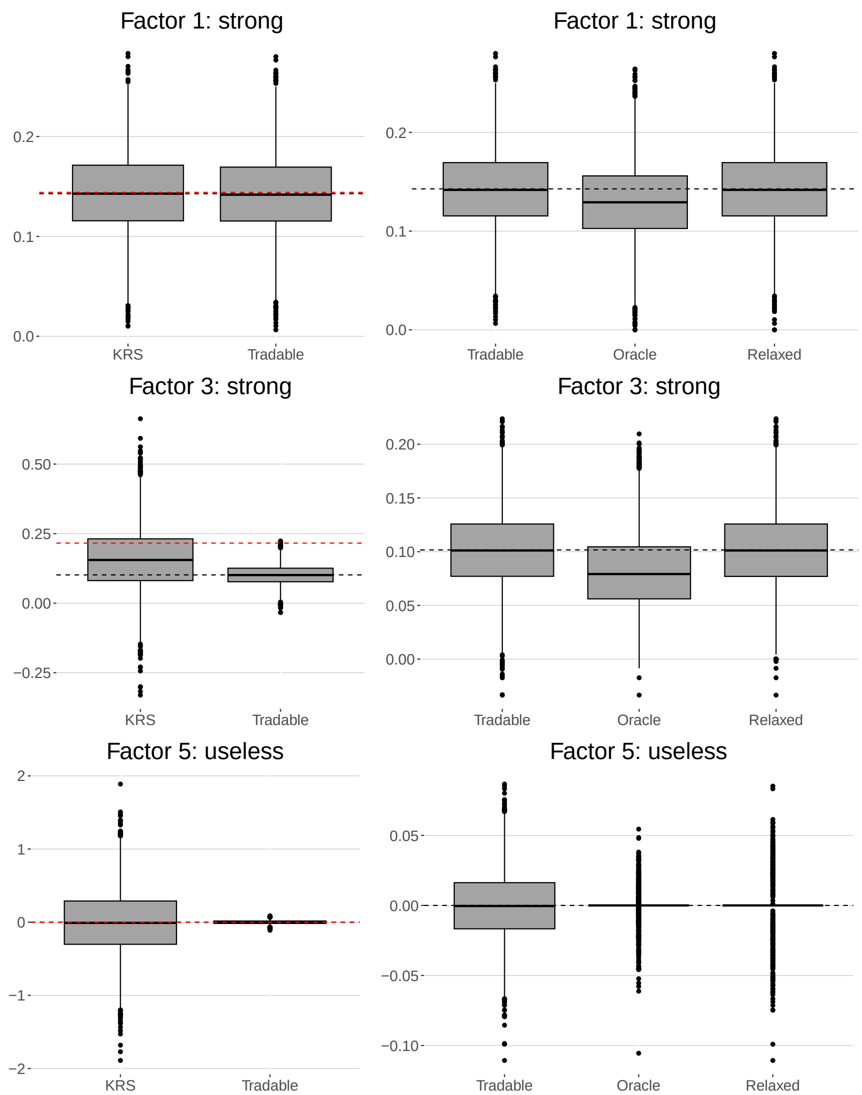

Figure 21: **Estimators' simulated distribution, Model 4 (correctly specified)**: the figure presents the marginal simulated distributions of KRS (left panels), tradable (left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators of the first strong (upper panels), the third strong (central panels) and the first useless (lower panels) factors of the correctly specified version of simulation Model 4. We omit the figure of the other useless factors and of the weak factor as they are qualitatively identical to the one of the first useless factor. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

9.2.2 Additional Monte Carlo setting

We report the simulation evidence for the two-factor model in Gospodinov et al. (2014), in which returns on the test assets and factors are simulated in a similar manner as in Section 4 of the main text, but with moments of the factor that is neither weak nor useless calibrated to the market factor and moments of the useless factor given by a zero mean and a unit variance. As shown in Figures 22–24 below, the analysis and conclusions regarding the MSEs and finite sample distributions of our sample, Oracle and Relaxed estimators of tradable factor risk premia, in relation to the misspecification-robust estimators of Kan et al. (2013) of misspecification-robust factor risk premia, remain unchanged. Similarly, Figure 25 displays the high degree of factor selection accuracy achieved by our Oracle estimator in this simulation setting.

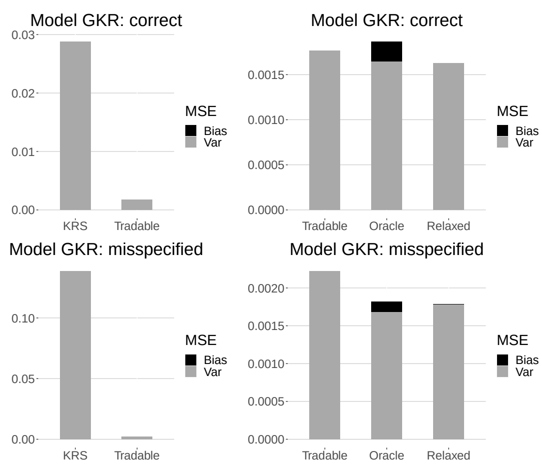

Figure 22: **Estimators' MSE decomposition, Model GKR**: the figure presents the MSE bias-variance decomposition of KRS (left panels), tradable (both left and right panels), Oracle and Relaxed (right panels) factor risk premia estimators in simulation Model GKR, for both correctly specified (upper panels) and misspecified (lower panels) version.

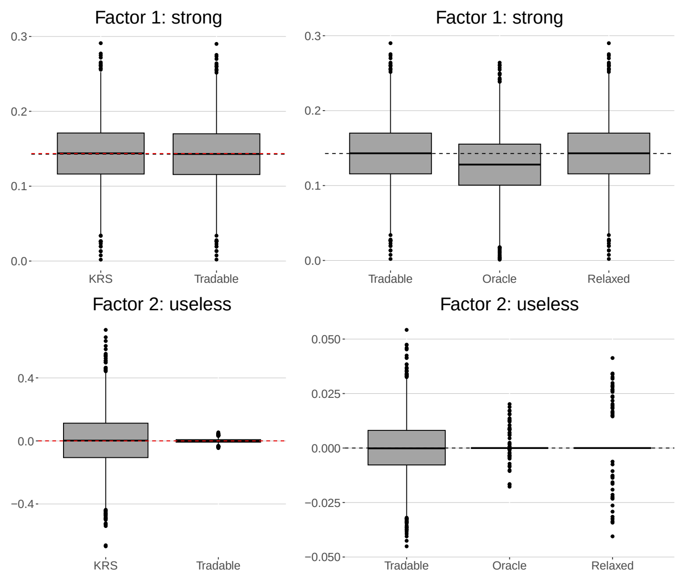

Figure 23: **Estimators' empirical distribution, Model GKR (correctly specified)**: the figure presents the empirical distribution of KRS (left panels), tradable (left panels), Oracle and Relaxed (right panels) factor risk premia estimators of the strong (upper panel) and the two useless (lower two panels) factors of the correctly specified version of simulation Model GKR. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

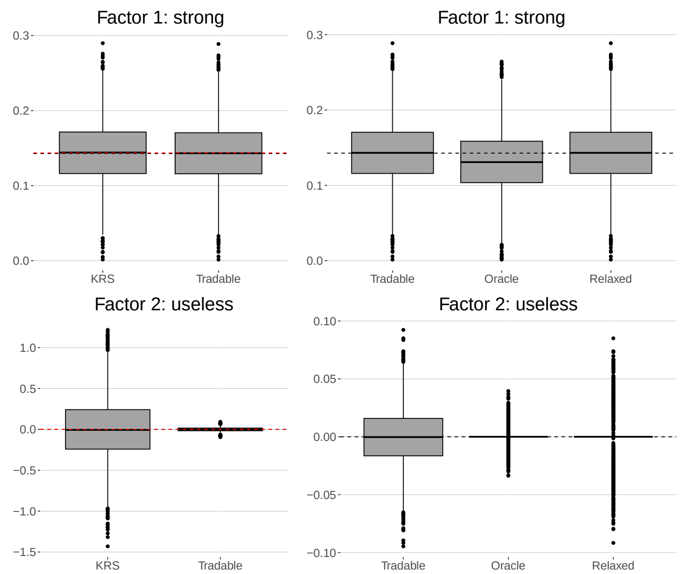

Figure 24: **Estimators' empirical distribution, Model GKR (misspecified)**: the figure presents the empirical distribution of KRS (left panels), tradable (left panels), Oracle and Relaxed (right panels) factor risk premia estimators of the strong (upper panel) and the two useless (lower two panels) factors of the misspecified version of simulation Model GKR. The population tradable and KRS factor risk premia are represented by the horizontal dashed black and red lines, respectively.

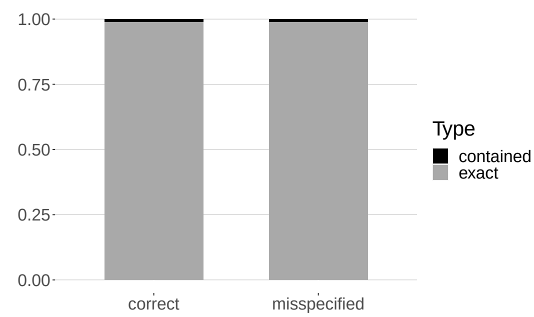

Figure 25: **Proximal factor risk premia selection properties, Model GKR**: the figure presents the "exact" and "contained" variable selection properties of our Oracle factor risk premia estimator for both correctly specified and misspecified simulation Model GKR. The "exact" ("contained") variable selection property is computed as the fraction over the simulation runs that the active set of the tradable population factor risk premia coefficient equals (is contained in) the estimated active set of our proximal risk premia estimator.

9.3 Additional empirical evidence

9.3.1 Identification and factor selection

This section complements the empirical analysis in Section 5 of the main text, by reporting model identification frequencies, post-screening factor dimensions and factor selection frequencies across randomized models with initial factor dimension between 1 and 4.

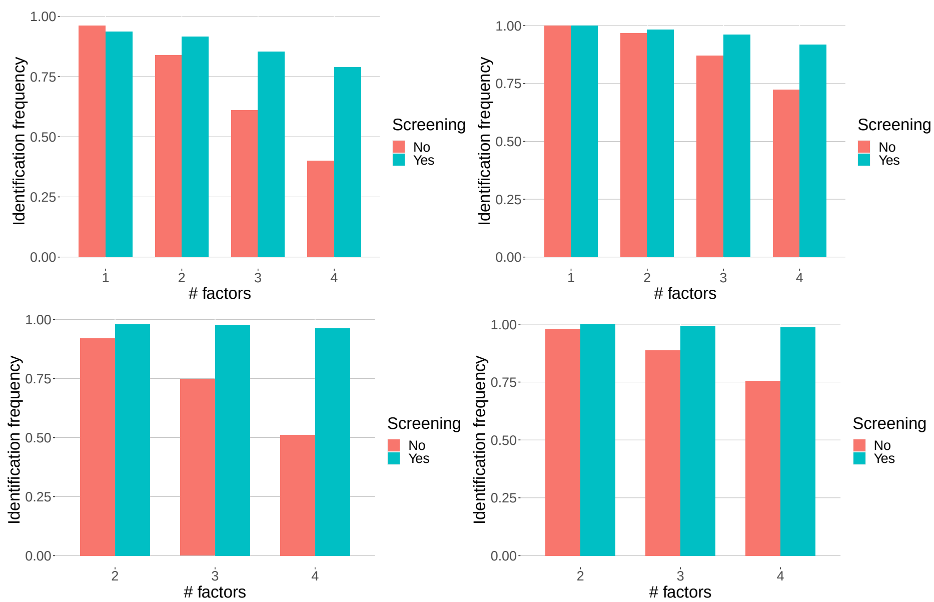

Figure 26: **Model identification frequencies**: Frequency of null hypothesis (39) rejections under a significance level $\alpha = 0.05$ by the Chen and Fang (2019) test, across randomized factor models including 1–4 initial factors. The red (blue) bars indicate model identification frequencies before (after) having applied our Oracle factor selection. The upper (lower) panels report the model identification frequencies for selections with no (the market as) ubiquitous factor across models. The left (right) panels report the results for test assets comprising the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios. For models having 1 (2) randomized factors, we consider all 52 (1326) possible model combinations. For models including more than 2 randomized factors, we consider 10'000 random factor combinations.

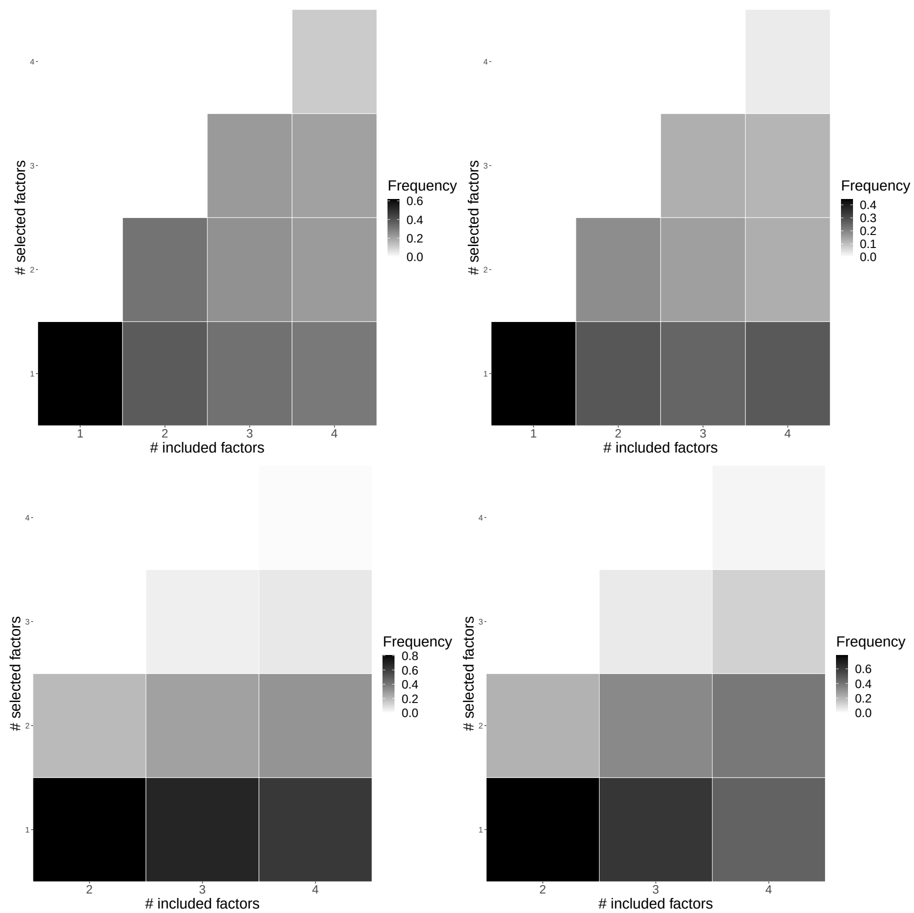

Figure 27: **Post-selection model size**: Frequency of post-selection model dimensions, i.e., number of factors selected by our Oracle factor selection, across randomized factor models including 1–4 initial factors. The red (blue) bars indicate model identification frequencies before (after) having applied our Oracle proximal factor selection. The upper (lower) panels report the model identification frequencies for selections with no (the market as) ubiquitous factor across models. The left (right) panels report the results for test assets comprising the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios.

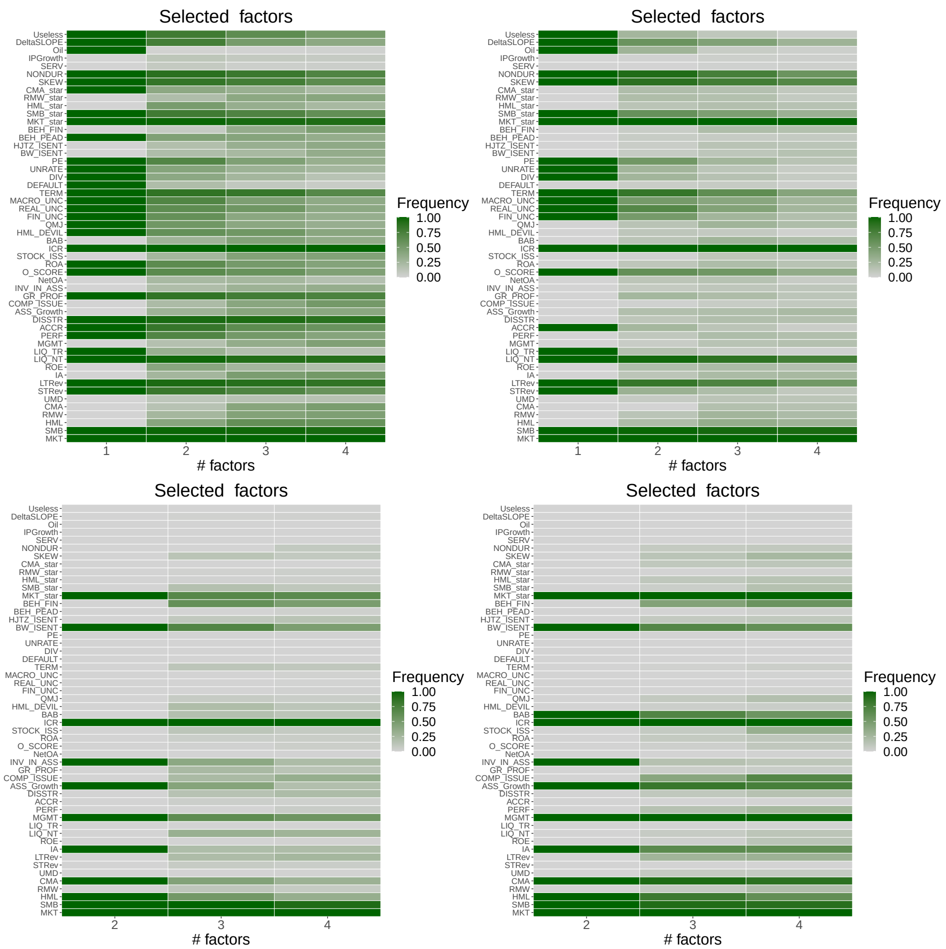

Figure 28: **Factor selection frequencies**: Selection frequencies of individual factors for our Oracle factor selection across randomized factor models including 1–4 initial number of factors. The upper (lower) panels report the frequencies of the post-selection model dimensions for selections with no (the market as) ubiquitous factor across models. The left (right) panels report the results for test assets comprising the 25 size/book-to-market (the 25 size/book-to-market and the 17 industry) portfolios.

9.3.2 Misspecification-robust factor risk premia

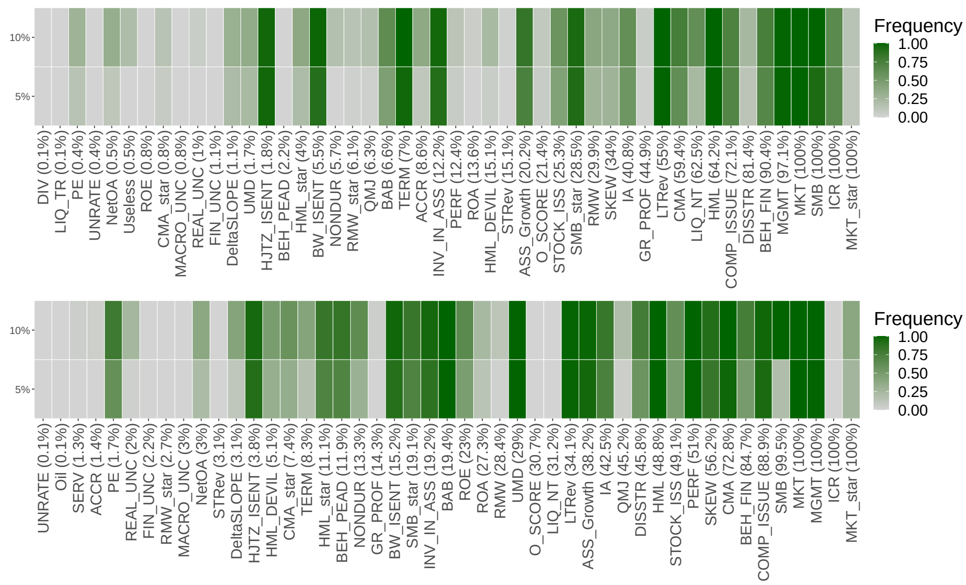

Figure 29: **Significance of misspecification-robust factor risk premia**: Significance frequencies of individual misspecification-robust factor risk premia in randomized 10-factor models always including the market. The 5% and 10% rows report the significance frequency of estimated misspecification-robust factor risk premia post-screening, based on marginal t-tests with significance level 5% and 10%, respectively. Factors are ordered by their selection frequency in randomized 10-factor models always including the market, reported in parenthesis, and only factors with positive selection frequency are retained. The upper (lower) panel reports the results obtained with test assets comprising the 25 size/book-to-market (25 size/book-to-market and the 17 industry) portfolios.

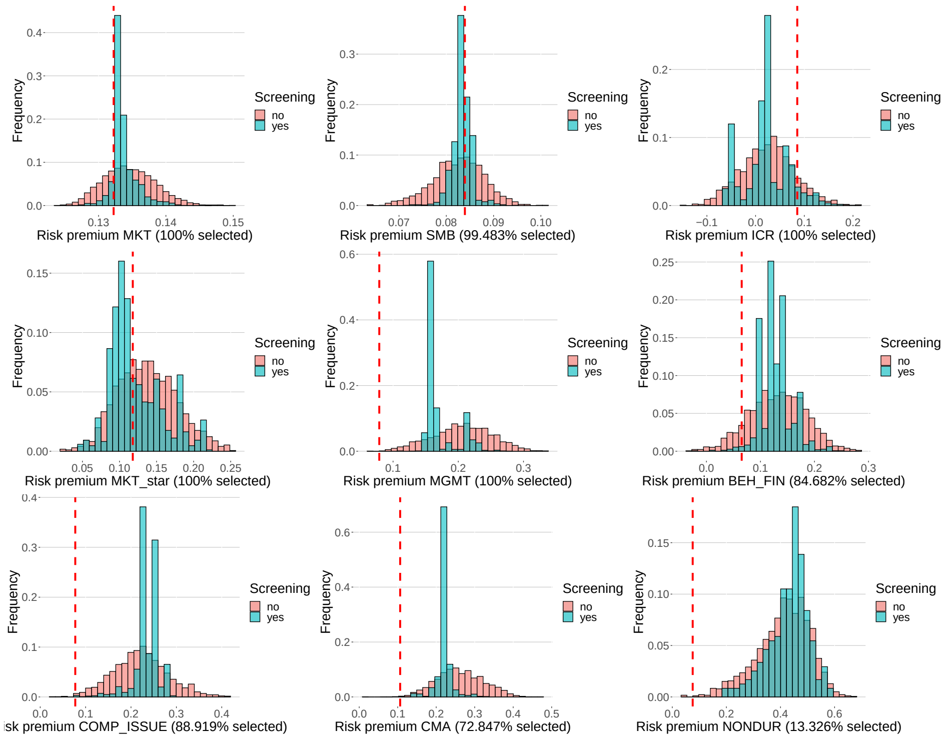

Figure 30: **Misspecification-robust factor risk premia**: Empirical distribution across randomized 10-factor models always including the market of estimated misspecification-robust factor risk premia for a selection of factors. The dashed red vertical line displays the estimated tradable factor risk premium. Factor selection frequencies are reported in parenthesis, next to the factor's label. The results are obtained with test assets comprising the 25 size/book-to-market and the 17 industry portfolios.

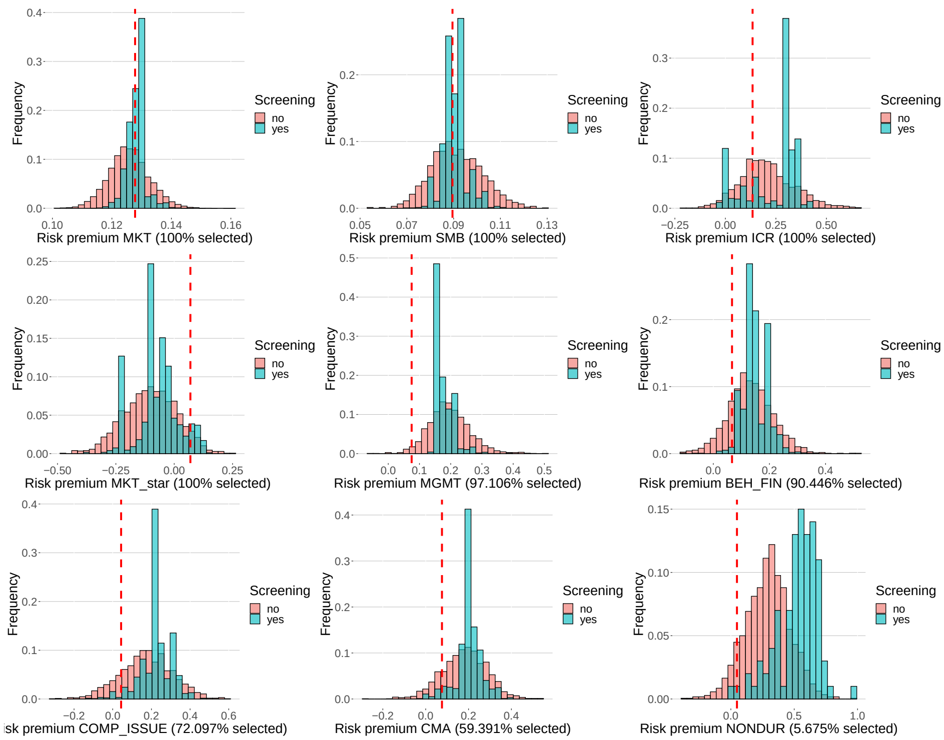

Figure 31: **Misspecification-robust factor risk premia**: Empirical distribution across randomized 10-factor models always including the market of estimated misspecification-robust factor risk premia for a selection of factors. The dashed red vertical line displays the estimated tradable factor risk premium. Factor selection frequencies are reported in parenthesis, next to the factor's label. The results are obtained with test assets comprising the 25 size/book-to-market portfolios.

9.3.3 Robustness to the choice of test assets

In this section we perform the same empirical analysis found in Section 5 of the main text, for a different set of test asset returns: the 25 size/book-to-market and the 25 operating profitability/investment portfolios. Overall, the analysis performed with this set of test assets confirms the conclusions obtained when test asset comprise only the 25 size/book-to-market or the 25 size/book-to-market and the 17 industry portfolios. The only major difference consists in the fact that, when considering models that always include the market factor, more factors seem to display a higher selection frequency compared to the selection frequencies obtained without necessarily always including the market factor in the model; see Figure 34.

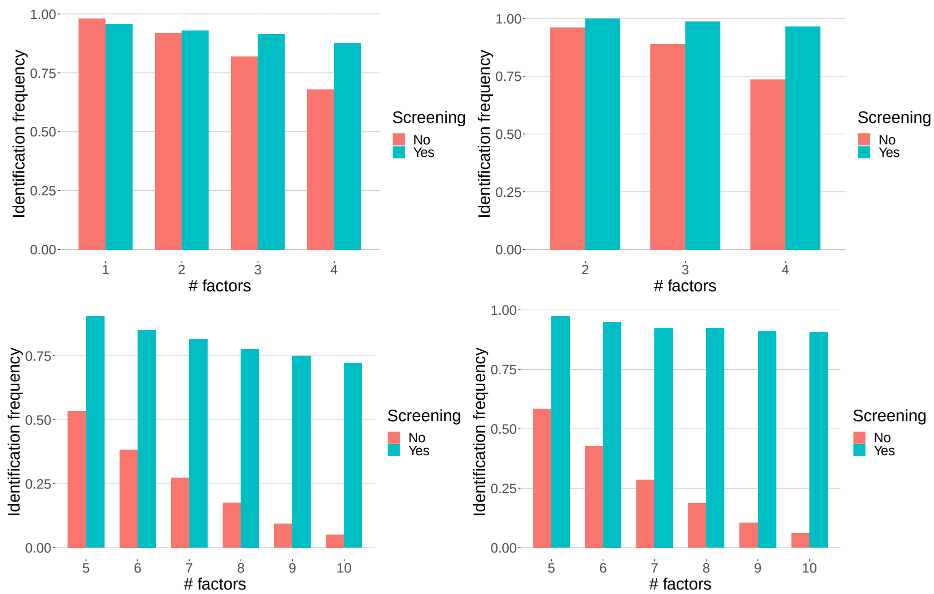

Figure 32: **Model identification frequencies**: Frequencies over various combinations of factor models having 1–10 initial number of factors that null hypothesis (39) is rejected, with a size $\alpha = 0.05$, by the Chen and Fang (2019) test. The red (blue) bars indicate the model identification frequencies before (after) having applied our proximal factor selection. The left (right) panels report the model identification frequencies for a model with no (the market as) common factor in the various models. The results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

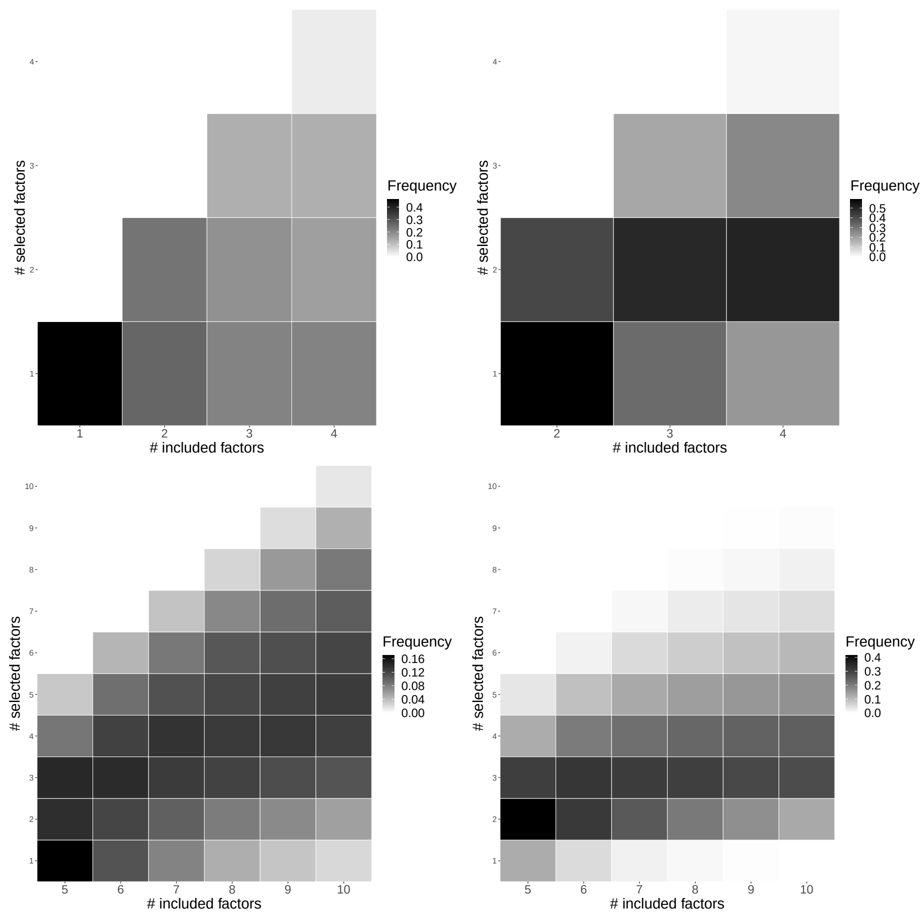

Figure 33: **Post-selection model size**: Frequencies of the post-selection model size, i.e., number of factors selected by our proximal factor selection, over the various combinations of factor models having 1–10 initial number of factors. The left (right) panels report the frequencies of the post-selection model size for a model with no (the market as) common factor in the various models. The results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

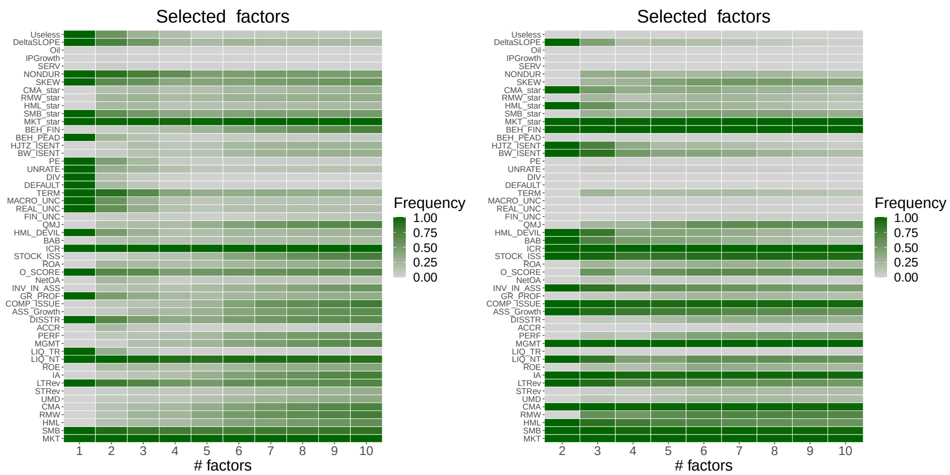

Figure 34: **Factor selection frequencies**: Selection frequencies of individual factors under our proximal factor selection procedure over the various combinations of factor models having 1–10 initial number of factors. The left (right) panel reports the factor selection frequencies for a model with no (the market as) common factor in the various models. These results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

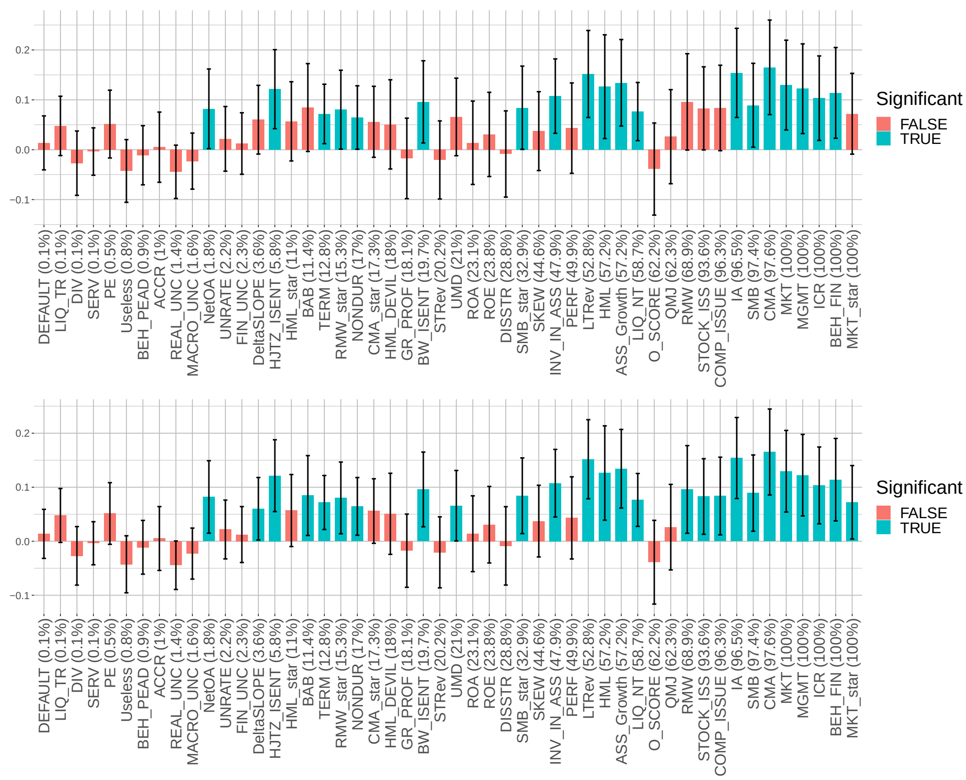

Figure 35: **Tradable factor risk premia**: Tradable factor risk premia and corresponding confidence intervals at approximate level 95% (90%) in the upper (lower) panel. Factors are ordered by their selection frequency in randomized 10-factor models always including the market, reported in parenthesis, and only factors with positive selection frequency are retained. The results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

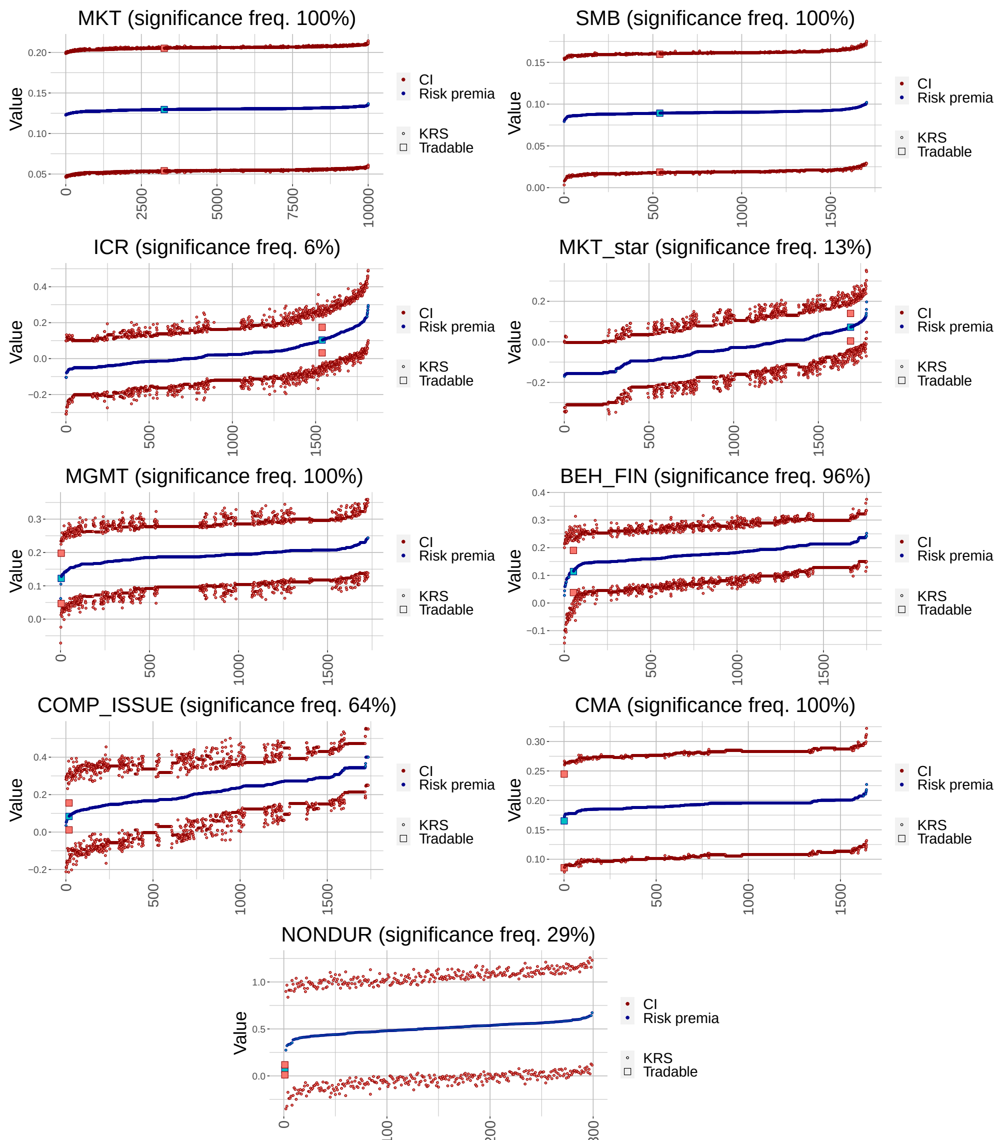

Figure 36: **Factor risk premia and** 95% **confidence intervals**: Misspecification-robust (represented by dots) and tradable (represented by squares) factor risk premia and corresponding 95% confidence intervals of a selection of factors over various random- ized 10-factor models always including the market. The significance frequencies are reported in parenthesis, next to the factor's label. The results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

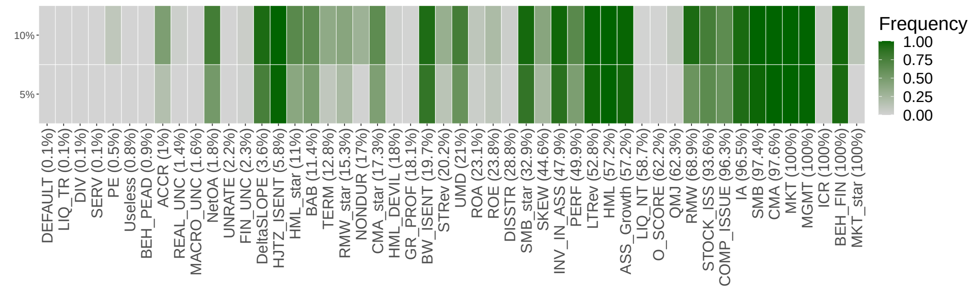

Figure 37: **Significance of misspecification-robust factor risk premia**: Significance frequencies of individual misspecification-robust factor risk premia in randomized 10-factor models always including the market. The 5% and 10% rows report the significance frequency of estimated misspecification-robust factor risk premia post-screening, based on marginal t-tests with significance level 5% and 10%, respectively. Factors are ordered by their selection frequency in randomized 10-factor models always including the market, reported in parenthesis, and only factors with positive selection frequency are retained. The results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

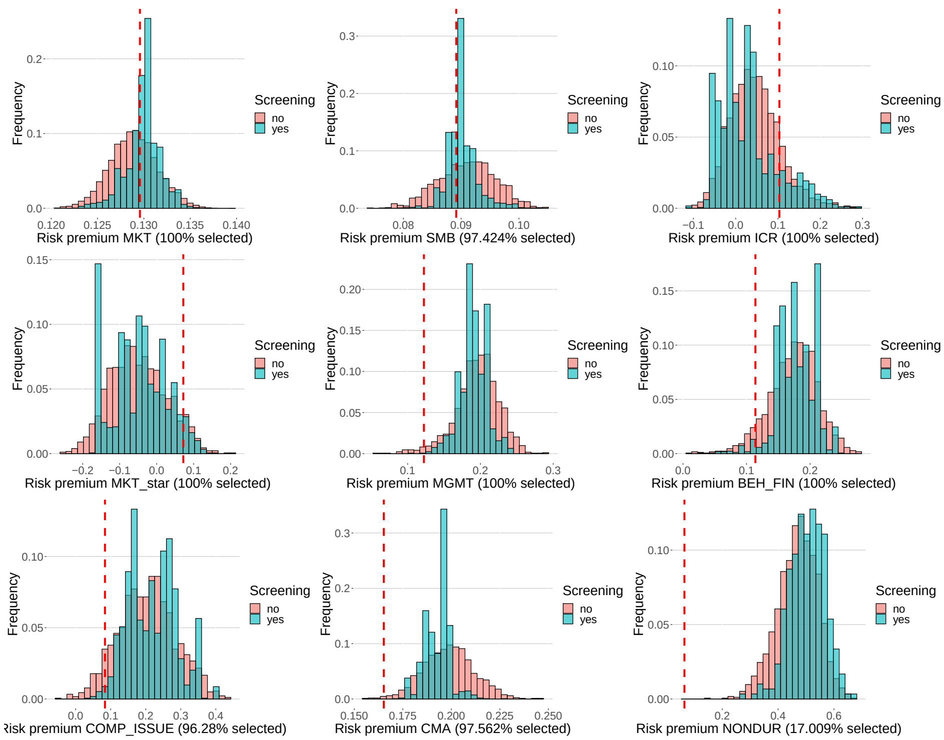

Figure 38: **Misspecification-robust factor risk premia**: Histograms of a selection of misspecification-robust factor risk premia of a selection of factors over various randomized 10-factor models always including the market. The dashed red vertical line displays the corresponding tradable factor risk premia. The selection frequencies are reported in parenthesis, next to the factor's label. The results are obtained with test assets comprising the 25 size/book-to-market and the 25 operating profitability/investment portfolios.

Swiss Finance Institute

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

swiss:finance:institute

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