



HAL
open science

Tradable or nontradable factors: what does the Hansen–Jagannathan distance tell us?

Xiang Zhang, Yangyi Liu, Kun Wu, Bertrand Maillet

► To cite this version:

Xiang Zhang, Yangyi Liu, Kun Wu, Bertrand Maillet. Tradable or nontradable factors: what does the Hansen–Jagannathan distance tell us?. *International Review of Economics and Finance*, 2021, 71, 853-879 p. hal-03287946

HAL Id: hal-03287946

<https://hal.science/hal-03287946>

Submitted on 16 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tradable or Nontradable Factors—What Does the Hansen–Jagannathan Distance Tell Us? *

Xiang Zhang[†]
Yangyi Liu[‡]
Kun Wu[§]
Bertrand Maillet[¶]

September 14, 2020

Abstract

We investigate the difference in pricing performance between tradable and nontradable factors in terms of explaining cross-sectional portfolio returns by comparing the Hansen–Jagannathan (HJ) distance misspecification measures. By constructing nontradable factors mimicking portfolios and incorporating them into the least misspecified tradable stochastic discount factor (SDF), we provide cross-country empirical evidence that this single proxy SDF dominates others to price cross-sectional risky assets. Since nontradable factors mimicking portfolios (FMPs) are functions of current risky factors information about the economic state, therefore FMPs “hedge” the state variable risks and FMPs’ returns describe the risk premiums.

JEL Classification: G1, G12

Keywords: Tradable and Nontradable Factors, Hansen–Jagannathan Distance, Misspecification Errors, Mimicking Portfolios

*We would like to thank Raymond Kan, Cesare Robotti, Abhay Abhyankar, Michael Creel, Bing Han, Robert Anderson, Rajesh Tharyan, Motohiro Yogo and Guofu Zhou for their comments and suggestions. We are especially grateful to Raymond Kan and Cesare Robotti for their helpful comments.

[†]School of Finance/Institute of Big Data, Southwestern University of Finance and Economics, 611130, Chengdu, Sichuan, China. Email: xiangzhang@swufe.edu.cn

[‡]Ph.D. Student, School of Finance, Southwestern University of Finance and Economics, 611130, Chengdu, Sichuan, China.

[§]Chengdu Intelligent Car City Development Co. LTD, 610000, Chengdu, Sichuan, China.

[¶]Corresponding Author, QUANT Research Center, Emlyon Business School, Paris 75012, France; CEMOL, University of La Reunion, Saint Denis 97744, France; Variances, Paris 75013, France.

1 Introduction and Motivation

Empirical studies often use a linear combination of factors to explain asset returns. Those factors may be the excess returns on some traded securities (tradable factors), nontraded economy-wide sources of uncertainty related to macroeconomic variables, or a combination of the two. In view of the empirical success of such factors, researchers have entertained a progressively broader set of them, which has resulted in several claims (Harvey et al. (2016); Hou et al. (2017)). While Barillas and Shanken (2018) and Barillas et al. (2019) determine which of tradable factors are successful by comparing improvements in Sharpe ratios, some further questions remain: whether a determined linear stochastic discount factor (SDF), which consists of admissible tradable or nontradable factors, outperforms others to price cross-sectional risky assets when all SDFs are misspecified? Do any nontradable factors (macro factors) are successful by comparing improvements in misspecification errors to tradable factors?

Few studies find that nontradable factors empirically outperform tradable factors, which makes nontradable factors seem useless for explaining cross-sectional returns. Despite the theoretical importance of macroeconomic risk factors in explaining the cross-section of expected asset returns, macro factor-based asset pricing models fail to explain certain cross-sectional stock return anomalies, such as momentum (Griffin et al. (2003)), and the profitability premium (Wang and Yu (2013)). Most studies commonly attribute the empirical failure of the macro factor-based asset pricing model to the large measurement errors in macroeconomic factors, the differences between a theoretical definition and its empirical counterpart, or the low frequency in reporting macroeconomic variables. Hence, not only is it necessary to perform a rank test for avoiding ‘useless’ nontradable factors (Gospodinov et al. (2017)), but also it is more interesting to investigate nontradable factors’ pricing ability against tradable factors through another perspectives: misspecification errors – the Hansen-Jagannathan (HJ) distance.

Using the HJ distance to compare models’ pricing ability has a large literature, but there is lack of papers to discuss the pricing ability difference between nontradable and tradable factors. One of reasons is the misleading HJ distance inference used by early papers. Early papers use the point estimates of HJ distance and pairwise HJ distance comparison to show their models’ better pricing performances, including Jagannathan and Wang (1996), Lettau and Ludvigson (2001), Hodrick and Zhang (2001), Vassalou and Xing (2004), Parker and Julliard (2005), Wang (2005), Zhang (2006), Li et al. (2010). Chen and Ludvigson (2009) first show that the pairwise HJ distance comparison inference cannot jointly test of correct specification of two or more asset pricing models, “a general statistical procedure for model

comparison is still missing”. Gospodinov et al. (2013) further point out there exist the sampling and model misspecification uncertainty when we compare two or more models using sample HJ distances. Therefore, they improve the pairwise to the multiple model comparison inference based on the sample HJ distance measures.

Another reason is that since macro factors usually exhibit small correlations with asset returns, the misspecification error measure (the HJ distance) is distorted when it is used for linear asset-pricing models even in large samples (Kong (2019)). Theoretically, macroeconomics factors (e.g., Chen et al. (1986)) and other non-traded factors (e.g., Adrian et al. (2014)) capture the fundamental risks in the economy and thus should also explain the cross-sectional expected returns. However, observed changes in these factors contain measurement errors and provide only weak predictions of asset returns. To reduce factor noise, the previous literature recommends factor-mimicking portfolios (FMPs), which contain traded assets that are representations of underlying non-traded factors (e.g., Barillas et al. (2019); Pukthuanthong et al. (2019)). Balduzzi and Robotti (2008) conclude that using the time-series formulation of FMPs performs better in term of estimating risk premiums than using the original factors with the one-step cross-sectional approach. Kleibergen and Zhan (2018) propose a test of the risk premiums of FMPs constructed by a time-series approach that does not depend on the magnitude of betas. Barillas et al. (2019) also use the time-series formulation of FMPs in order to construct Sharpe ratio of candidate models for comparison. Intuitively, since the return covariances with the factors are functions of current information about the economic state, therefore FMPs “hedge” the state variable risks and FMPs’ returns describe the risk premiums.

This paper makes two contributions as follows:

First, this paper finds that there exists a determined least misspecified SDF which outperforms others to price several countries’ cross-sectional risky assets. The paper uses 230 portfolios and 21 factors over the period 1967-2015 in the U.S. market, 122 portfolios with 16 factors monthly from 1990 to 2017 in the U.K. market, and 22 factors to price 83 portfolios monthly from 2000 to 2018 in the Chinese market. By employing the methodologies of Chen and Ludvigson (2009) and Gospodinov et al. (2013), which allow multiple model selection tests with improved finite-sample properties, the paper finds that unlike the problem of a “zoo of factors,” the Fama-French five-factor plus momentum model has a better performance than alternatives such as the CAPM, the Fama-French three-factor model (Fama and French (1992)), the Carhart four-factor model (Carhart (1997)), the Fama-French three-factor plus liquidity model from Pástor and Stambaugh (2003), the Fama-French five-factor model (Fama and French (2015)), the four-factor

Q model (Hou et al. (2015)), the betting-against-beta model (Frazzini and Pedersen (2014)) and the mispricing factors model (Stambaugh and Yuan (2016)). None of the nested nontradable factors can outperform the nested tradable factor SDFs in terms of these countries' pricing results.

Second, this paper gives an international empirical evidence that nontradable factors are able to improve the pricing performance of SDFs while incorporating a mix of tradable and nontradable factors. With the help of factor-mimicking portfolios (FMPs), the paper contributes to those studies by providing further evidence for the incremental value of the nontradable factors: given common tradable factors, the nontradable factors can outperform tradable factors to decrease SDFs' misspecification errors. Given the Fama-French five tradable factors, the paper finds that in the U.S. market, Yogo (2006)'s durable consumption growth, Piazzesi et al. (2007)'s housing expenditure variable and the liquidity innovation variable actually outperform the momentum factor when augmented by the Fama-French five-factor model. For the cross-country analysis, we find that the Fama-French five plus the inflation FMPs in the U.K. and the the Fama-French plus yield level FMPs in China, each of them has the smallest HJ distance measures among alternatives when pricing 122 and 83 portfolios, respectively. In addition, nontradable factors are especially vulnerable to the "limited T versus large N " problem and become weak factors (Kleibergen and Zhan (2020)). For instance, when we include 83 portfolios as test assets for the Chinese market, we find that nontradable factors are all useless and provide little incremental value in improving misspecification errors.

Non-tradable factors helping to improve misspecification errors has economic foundations. Macroeconomic variable offer key insight into the general state of the economy, but they may not sufficiently capture the most accurate correlation structure of price movement across stocks. However, factor-mimicking portfolios (FMPs) are functions of current risky factors information about the economic state, therefore FMPs "hedge" the state variable risks and FMPs' returns describe the risk premiums. As a result, after controlling on tradable factors, nontradable factors (macro variables) are informative about both pricing errors and future expected returns (risk premiums); nontradable factors usually show a relatively higher predictive power of assets' cumulative returns in the long run than tradable factors do. Several papers apply macro mimicking factors to explain risk-return relationship in the beta-pricing models. Ferson et al. (2006) derives and characterizes mimicking portfolios in the presence of predetermined state variables. Their time-varying weight solutions are affected more by parameter estimation errors when the wrong Data Generator Process (DGP) is assumed. Our FMPs construction follows Pukthuanthong et al. (2019) and Jurczenko and Teiletche (2019). Pukthuanthong et al. (2019) find that equity returns

are priced by consumption growth, inflation, and the unemployment rate when they use the factor-mimicking portfolios to the beta-pricing model. Jurzenko and Teiletche (2019) incorporate empirical estimation improvements through machine learning methodologies, and they provide an application to the construction of tradable portfolios mimicking three global macro factors, namely growth, inflation surprises, and financial stress indicators. They show that these macro mimicking factors can be used to improve the risk-return profile of a typical endowment multi-asset portfolio. Different from the above, this paper aims to investigate SDFs misspecification improvements through macro mimicking factors.

Papers close to our study also include Barillas and Shanken (2017) (BS 2017) and Barillas and Shanken (2018) (BS 2018). BS (2017) shows that researchers do not need to match the mean-variance frontier of test assets for every factor portfolio; rather, they can evaluate the relative ‘match’ of each pair of candidate factor portfolios. Therefore, the choice of test assets will not influence the results of any pairwise factor models. However, the difference between the ‘relative comparison’ in BS (2017) and the ‘absolute comparison’ in our paper comes from the fundamental difference between the GRS test and the HJ distance test. The simple GRS test in BS (2017) basically evaluates the chances of improvement in excluded factor portfolios against the benchmark factor portfolios¹. However, the HJ distance test aims to find the minimum variance portfolios among candidate factor mimicking portfolios given the same risky assets to price. More importantly, the HJ distance test does not choose a zero-beta rate to minimize the difference in the squared Sharpe ratios of the two tangency portfolios; instead, the zero-beta rate is chosen to minimize the difference in the squared Sharpe ratios of the two tangency portfolios divided by the squared zero-beta rate. Hence, the zero-beta rates are estimated differently by the HJ distance test and by the GRS test. Moreover, Barillas and Shanken (2018) develop a procedure that allows for the analysis of the joint alpha restriction for a set of test assets in a Bayesian setting. Their evidence in this paper casts strong doubt on the validity of their six-factor model.

In the rest of the paper, we introduce in Section 2 the HJ distance, the multiple HJ distance comparison tests and the set inference-based confidence intervals. Section 3 describes the data and candidate models. Section 4 shows the empirical test results. Finally, we conclude the paper in Section 5.

¹Barillas et al. (2019) test the improvements in the squared Sharpe ratio by correctly specified and misspecified nested and nonnested models. They find that a variant of the Fama and French (2018) six-factor model emerges as the dominant model.

2 Methodology

We use three different Hansen–Jagannathan (HJ) distance measures to compare candidate linear asset-pricing models: the original HJ distance (HJ^O) introduced by Hansen and Jagannathan (1997), the modified HJ distance (HJ^M) from Kan and Robotti (2008), and the constrained HJ distance (HJ^C) from Gospodinov et al. (2012).

We begin with the basic asset-pricing model in the stochastic discount factor (SDF) representation

$$p_t = E_t[m_{t+1}x_{t+1}], \quad (2.1)$$

where p_t is the price of any stock, m_{t+1} is the true SDF, x_{t+1} is the future payoff of the stock, and E_t is the conditional expectation operator.

An asset-pricing model identifies a particular SDF that is a function of observable variables and the model parameters. Empirical estimation on this model can be done by using the two-stage GMM,

$$\min_b [g_T(b)' W g_T(b)], \quad (2.2)$$

where g_T represents the moment conditions and W is a weighting matrix. Most earlier papers use Hansen’s J_T test statistics to estimate and test each model on the same set of asset returns—testing correct specification against the alternative of incorrect specification. Using the J-statistics, we find that the overidentification restrictions are not rejected for one model but are rejected for another. However, as Hansen’s J_T test statistic depends on the model-specific S matrix (Ludvigson (2013)), a model can look better just because its SDF and the pricing errors are more volatile than those of its competitor.

2.1 The Hansen–Jagannathan Distance

Hansen and Jagannathan (1997) suggest a solution to this problem. They assume that the proposed SDF y_{t+1} can be approximated as a linear function of factors

$$y_{t+1} = \theta' f_{t+1}, \quad (2.3)$$

where f_t denotes the pricing factors. By using the pricing equation, we can derive the following equation:

$$\alpha_t(\theta) = R_t y_t(\theta) - I_N = R_t \theta f_t' - I_N, \quad (2.4)$$

where $R_t = [R_{1,t}, R_{2,t}, \dots, R_{N,t}]'$ are the gross returns on N assets and $\alpha_t(\theta)$ is the vector of pricing errors. Hence, the maximum pricing error per unit

norm of any portfolio of N assets (or HJ^O) is given by

$$\left[HJ^O \right]^2 = E \left[(\alpha_t(\theta))' \right] \left[E(R_t R_t') \right]^{-1} E[\alpha_t(\theta)]. \quad (2.5)$$

The HJ^O measure is equivalent to a GMM estimator with the moment condition $E[\alpha_t(\theta)] = 0$ and the weighting matrix $\left[E(R_t R_t') \right]^{-1}$, which is different from the optimal matrix (see Appendix A.1 for details). There are two advantages that lead us to choose to use the HJ distance. The first advantage is stated by Ludvigson (2013), namely, that the HJ distance does not reward SDF volatility. As a result, it is suitable for conducting model comparisons. Second, the HJ distance provides a measure of model misspecification. The HJ distance also gives the maximum pricing error of any portfolio formed from N assets.

If excess returns are used to measure model misspecification, one cannot specify a proposed SDF in such a way that it can be zero for some values of θ . Kan and Robotti (2008) suggest defining the SDF as a linear function of the demeaned factors to avoid the affine transformation problem. Hence, the modified HJ distance (HJ^M) measure is defined as

$$\left[HJ^M \right]^2 = \min_{\theta} E \left[\alpha_T(\theta)' \right] V_{22,T}^{-1} E[\alpha_T(\theta)], \quad (2.6)$$

where $V_{22,T}^{-1}$ is the covariance matrix of the test portfolios.

Another problem we need to consider is that all of the above SDFs can be either positive or negative. For instance, if markets are incomplete, as suggested by Ross (1973), candidate SDFs such as the CAPM and linear factor models do not need to be strictly positive (Cochrane and Hansen (1993)). It is, however, possible for an SDF to price all the test assets correctly and yet take on negative values with positive probability. This happens when arbitrage opportunities exist among test portfolios (e.g., derivatives on test assets), and it could be problematic to set the SDF to price payoffs. Therefore, it is necessary to constrict the admissible SDFs to be nonnegative.

Gospodinov et al. (2012) solve for the constrained HJ distance as

$$\left[HJ^C \right]^2 = \min_{m_t, t=1, \dots, T} \frac{1}{T} \sum_{t=1}^T (y_t - m_t)^2, \quad (2.7)$$

subject to

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T m_t R_t &= \bar{q}, \\ m_t &\geq 0, t = 1, \dots, T, \end{aligned}$$

where y_t denotes the candidate SDF, m_t stands for an admissible SDF in the

set \mathbb{N}_+ , q_{t-1} is the vector of corresponding costs of N assets and $E[q_{t-1}] \neq 0$.

We examine whether the SDF is positive, which implies the absence of arbitrage. If the SDF is negative, then markets are incomplete and there are arbitrage opportunities; otherwise, there is something wrong with SDF theory itself.

2.2 Multiple Comparisons Tests for the HJ Distance

The traditional HJ distance test compares HJ distance measures statistically by making pairwise model comparison the null hypothesis: i.e., HJ_2^O may be less than HJ_1^O . Specifically, are the models significantly different from one another once we account for sampling error?

Gospodinov et al. (2013) propose a new Lagrange multiplier test for joint testing of misspecification of more than two asset-pricing models. They develop chi-squared versions of model comparison tests for strictly nonnested, nested and overlapping models. They also provide a multiple model comparison test that allows us to compare a benchmark model with a set of alternative models in terms of their HJ distance metrics. They suggest that we should separate models into three categories—nested, strictly nonnested and overlapping—and introduce test methods for different types of models.

In addition to the new tests developed in Gospodinov et al. (2013) for multiple model comparison, we also apply Chen and Ludvigson (2009)'s test method. The details of these models are given in Appendix A.4.

3 Data and Candidate Models

We investigate international empirical evidences based on data of two representative developed countries (US and UK) and of one representative developing country (China).

Although the asset-pricing literature has proposed an extremely large number of trading factors (McLean and Pontiff (2016); Harvey et al. (2016); Barillas et al. (2019)), we focus on a few representative ones. Here is a list of the models and corresponding risk factors considered for the US market:

1. The capital asset-pricing model (*CAPM*) is the value-weighted market return, constructed from the Center for Research in Security Prices (CRSP) for all stocks listed on the NYSE, AMEX, or NASDAQ.
2. The Fama-French three-factor model (*FF3*) includes, in addition to the market return, SMB (size) and HML (value).
3. The Carhart four-factor model (*Carhart*) adds a momentum factor (MOM) to *FF3*.
4. The Fama-French five-factor model (*FF5*), from Fama and French

(2015), adds to $FF3$ the RMW (operating profitability) and CMA (investment) factors.

5. The four factors from the Q-factor model (HXZ) of Hou et al. (2015) include the market return, ME (size), IA (investment), and ROE (profitability).

6. The betting-against-beta factor (BAB) from Frazzini and Pedersen (2014) is included.

7. Four factors from Stambaugh and Yuan (2016) (SY), who extends the $CAPM$ by adding a size factor (SMBSY) and two mispricing factors, ‘management’ (MGMT) and ‘performance’ (PERF).

8. The Fama-French three-factor model plus the liquidity factor from Pástor and Stambaugh (2003) ($FF3 + LIQ$) is included.

9. The six-factor model adds the up-minus-down (UMD) momentum factor motivated by the work of Jegadeesh and Titman (1993) to the $FF5CP$ model ($FF5CP + UMD$) following Ball et al. (2016) and Fama and French (2018).

10. The HXZ model substitutes RMWCP for ROE ($HXZCP$), given that the choice of the profitability factor is key to the performance of the five-factor model of Fama and French.

11. The final model ($FF5CPM + UMD$) includes the more timely value factor HML^m from Asness and Frazzini (2013) instead of the standard HML .

The empirical analysis uses both monthly and quarterly return data over the period 1967-2015 for the US market. 230 portfolios are included as test assets: 25 portfolios sorted by size and book-to-market ratio, 30 industry portfolios, 100 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, 25 portfolios sorted by size and momentum, and 25 portfolios sorted by size and beta².

We include some nontradable factors. The first one is Yogo (2006)’s durable consumption (CG). To update the data to 2015, we use personal consumption expenditures (PCE) on nondurables and services (obtained from the Bureau of Economic Analysis), including food, clothing and shoes, housing, utilities, transportation, and medical care. This series is then deflated by a weighted average of the price index for nondurables and services. Other macro factors include the housing collateral ratio (HCG) in Lustig and Van Nieuwerburgh (2005), the nonhousing consumption expenditure share EXR in Piazzesi et al. (2007) and the liquidity innovation for the non-tradable factor in Pástor and Stambaugh (2003) ($LIQINNOV$).

For each nontradable factor, we have a mimicking portfolio by regressing the nontradable factor on a constant and all the traded-factors. We include

²The industry portfolios are included to provide a greater challenge to the various asset-pricing models, as recommended by Lewellen et al. (2010)

all tradable factors in the models that we have compared. Additional basis assets could be considered but are not required. Although some of these returns are highly correlated, we are interested in the fitted value (the overall mimicking return), rather than the individual weights.

For studying the UK equity market, we follow Hanauer (2020) to replicate all following UK tradable factors:

1. The capital asset-pricing model (*CAPM*).
2. The Fama-French three-factor model (*FF3*) includes *CAPM*, in addition to the market return, *SMB* (size) and *HML* (value).
3. The Carhart four-factor model (*Carhart*) adds a momentum factor (*MOM*) to *FF3*.
4. The Fama-French five-factor model (*FF5*) adds to *FF3* the *RMW* (operating profitability) and *CMA* (investment) factors.
5. The four factors from the Q-factor model (*HXZ - FF*) of Hou et al. (2015) include the market return, *SMB* (size), *CMA* (investment), and *RMW_ROE* (profitability). Specifically, different from the original version of Hou et al. (2015), the three factors are constructed based on 2×3 independent sort following Fama and French (2015), instead of the triple sort of Hou et al. (2015).
6. Four factors from Stambaugh and Yuan (2016) (*SY - FF*), who extend the *CAPM* by adding a size factor, *MGMT* and *PERF*.
7. The model (*FF5CP*) following Ball et al. (2016) and Fama and French (2018).
8. The Fama-French five-factor plus the the up-minus-down (*UMD*) momentum factor (*FF5 + UMD*).
9. The six-factor model adds the up-minus-down (*UMD*) momentum factor to the *FF5CP* model (*FF5CP + UMD*).
10. The final model (*FF5CPM + UMD*) includes the more timely value factor HML^m instead of the standard *HML*.

The test assets for the UK market include 122 portfolios from January 1990 to December 2017, in which there are 25 portfolios sorted by size and book-to-market ratio, 25 portfolios sorted by size and momentum, 10 portfolios sorted by size, 10 portfolios sorted by book-to-market ratio, 10 portfolios sorted by momentum, 25 portfolios sorted by standard deviation, and 27 portfolios sorted by size, book-to-market ratio and momentum.

The nontradable factors considered for the UK market are all monthly. We firstly include the UK yield level (*Yield*) that is calculated as the average yield to maturity of the government bond with maturity of 1 year, 5 years and 10 years. The second macro factor is the maturity spread (*MS*). We calculated it as the difference of the yield to maturity between the government bonds with a maturity of 10 years and 5 years. The last three macro variables are *CPI*, *PPI* and the inflation (*INF*).

China has become the largest stock market outside the US, as well as the

largest emerging market. Hence, comparing the models' performance in the developing market like China should offer us some valuable information. We fetch the data for the Chinese stock market from 2000 to 2018 in the Wind database, which is one of the widely used financial databases in China. We exclude financial firms, stocks under special treatment, and firms going public within 12 months as well as stocks that stay at their daily price limits for a whole day and stocks for which trading is suspended at the end of each month. We also exclude the bottom 30% stocks to avoid so-called shell-value contamination following Liu et al. (2019).

We replicate all the models for the US market mentioned except model 8 ($FF3 + LIQ$), Appendix A.5 provides the details of the factor models in China. We consider 83 portfolios as the test assets: 25 portfolios sorted by size and book-to-market ratio, 10 portfolios sorted by size, 10 portfolios sorted by book-to-market ratio, 10 portfolios sorted by operating profitability, and 28 industry portfolios³. For the nontradable factors, we consider six macro factors, including the quarter-over-quarter consumption growth (CG), the year-over-year wage growth (WG), the yield level ($Yield$), the maturity spread (MS), the credit spread (CS), and the inflation (INF).

4 Main Results

We present main empirical results in Section 4.1 for the US market, and Section 4.2 and 4.3 for the UK and the China equity markets, respectively.

4.1 The US Market

4.1.1. Sequential Selection Procedure and Rank Tests. In Table 1, we report the results of the rank tests of our candidate factors and the model misspecification tests. Panel A presents the results of the rank restriction test of each individual factor and the corresponding p -value of the null hypothesis that the $N \times K$ matrix $B = E[x_t(1, f_{it})]$ is of column rank 1. Gospodinov et al. (2014) propose this sequential procedure that allows us to eliminate the useless factors from the model, since the presence of a useless factor would lead to a violation of the crucial condition for identification, which states that B is of full rank. The results show that we can reject the null of the factor being of column rank of one at the 5% level of significance for all of the risk factors that the paper considers, which suggests that all of the factors can be considered potentially useful.

³We proxy the industry portfolios by the SWS industry indexes compiled by the Shenwan Hongyuan Securities. The SWS industry indexes are widely accepted in China, and the Shenwan Hongyuan Securities is one of the largest securities companies in China.

Panel B reports the results of the rank test of the models' misspecification. The tests based on the HJ distance measures show that none of the models pass the tests. The null of the HJ distance equal to zero is rejected at the 1% level of significance for all the models. The implication is that all the models we consider are potentially misspecified based on the HJ distance measures. Since the HJ distance measures have been shown to substantially overreject the null, we also conduct an LM test, which has better size and power properties. The results are similar to those of the HJ tests in that all the models reject the null, suggesting that the HJ distance results are not driven by the finite sample properties of the HJ distance tests. This is consistent with the notion that all asset-pricing models are approximate representations of reality and are therefore potentially misspecified.

In the rank test of the models, we find that none of our candidate models suffer from identification problems at the 1% level. This is consistent with the fact that the factors we consider are correlated with returns on the test assets. Gospodinov et al. (2014) also note that the *MKT*, *SMBC* and *HML* factors are highly correlated with the test asset returns. Overall, these findings from Panel A and Panel B suggest that the empirical factor models do not suffer from identification issues, although all of the models are misspecified.

In Table 1, we also report the results of the sequential selection procedure under correct model specification (Panel C) and potential model misspecification (Panel D) assumptions. The sequential testing methodology uses the Bonferroni correction to allow for multiple testing. The *t*-statistics under correct specification in Panel C show that while the *MKT* factor in the CAPM is significant, in the FF3 model, the *MKT* and the *HML* are significant. In the Carhart model, *MKT*, *HML* and *MOM* are significant. In the FF5 model, *MKT* and *HML* are useful, while *SMB*, *RMWC* and *CMA* do not survive the sequential testing procedure. In the HXZ model, the *MKT* and the *IA* factors survive the sequential testing procedure at the 5% significance level. The liquidity factor *LIQTRADED* does not pass the sequential test. The factors that survive all the specifications are momentum *MOM* (*UMD*), the betting-against-beta factor *BAB*, operating profitability *RMWC*, and the two mispricing factors *MGMT* and *PERF*, followed by *MKT* and *HML*. The results in Panel D are identical to the results in Panel C, when potential model misspecification is taken into account.

4.1.2. HJ Distance Comparison Results. In Table 2, we report the results of the model comparison tests using both the gross returns and the excess returns on our 230 portfolios assets.

In Panel A, we report the HJ distance (HJ^O), the modified HJ distance (HJ^M) and the constrained HJ distance (HJ^C) measures for all the mod-

els that we consider. These results allow us to compare the misspecification measures across candidate SDFs with the point estimation and corresponding inference. Here, lower values of the measures are preferred, as they show lower levels of misspecification. The results show that across all candidate models that we consider, the HJ distance measures are greater than zero. These results suggest that all the candidate SDFs are potentially misspecified. The model with the lowest HJ^O , HJ^M and HJ^C measures is the Fama-French five-factor plus up-minus-down momentum model (FF5+UMD) when we consider both gross returns and excess returns on 230 portfolios.

In Panel B, we present the results of the two formal tests of model comparison. In both of these tests, we take the least misspecified model, i.e., the one that has the lowest HJ^O , HJ^M and HJ^C measures among all the models in Table 2 Panel A as the benchmark. The three HJ distance measures of alternative models are then multiple tested against the measures of our chosen model, i.e., the FF5+UMD model.

The first test is a multiple comparison test of whether the HJ^O , HJ^M and HJ^C measures of any of the alternative models are significantly higher than the distance measures of our chosen model. The null hypothesis in this test is that the chosen model has HJ^O , HJ^M and HJ^C measures that are less than those of any of the alternative models. A failure to reject the null means that our chosen model is the least misspecified model. Using both the gross returns and excess returns on our 230 pooled portfolios, we find that the null is not rejected for any of the HJ distance measures. This suggests that the FF5+UMD model is the least misspecified.

The second test is a simultaneous pairwise comparison test of whether the HJ^O , HJ^M and HJ^C measures of each of the alternative models are significantly greater than the distance measures of our chosen model. The null hypothesis is that the chosen model has HJ^O , HJ^M and HJ^C measures that are less than those of each of the alternative models. The tests are conducted using both the gross and excess returns on the 230 portfolios of assets. The results show that at the 1% level of significance, the FF5+UMD model retains its position as the least misspecified of all the models.

In Table 3, we run tests for the multiple comparison tests only on the Fama-French 25 size and book-to-market ratio sorted portfolios, as well as other cross-sectional portfolios including 30 industry portfolios, 100 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, and 25 portfolios sorted by size and momentum. Each panel shows the result based on the candidate SDFs to price the cross-sectional risky assets to search for the least misspecified SDF. In Panel A, Panel B and Panel E, the table shows that the FF5+UMD model obtains the least misspecified measure of 11 factor models. The multiple comparison tests using a conservative size of 1% support the above conclusion. However,

in Panel C, Panel D and the excess returns on 100 operating profitability sorted portfolios (Panel E), the Fama-French five-factor including the updated value factor HML^m from Asness and Frazzini (2013) instead of the standard HML plus momentum factor model (FF5CPM+UMD) outperforms other candidate SDFs. The standard method calculates book-to-price (B/P) at portfolio formation using lagged book data, aligns price data using the same lag (ignoring recent price movements), and holds these values constant until the next rebalance. Thus, by the time the data are updated, the price used to determine “value” is 18 months old. Asness and Frazzini (2013) shows that B/P ratios based on more timely prices better forecast the true (unobservable) B/P ratios at fiscal year-end. Value portfolios based on the timeliest measures earn statistically significant alphas ranging from 305 to 378 basis points per year. This suggests that any conclusions drawn are potentially sensitive to the choice of test assets, although the Fama-French five factor can be treated as the common factor.

4.1.3. Nontradable Factors. We turn to the problem that this paper mainly focus on, i.e., whether any of the nontradable factors can reduce the HJ distance misspecification measures.

We find that the liquidity (nondurable consumption and services) mimicking portfolio has an average risk premium of 0.0005 (0.0129) per month over our sample period, and the associated t -statistic is 0.39 (2.2678). However, the durable consumption (housing consumption) mimicking portfolio has an average risk premium of -0.0614 (-0.0107) per month, and the associated t -statistic is -2.5791 (-7.4246). Insofar as marginal utility is low when the market is highly liquid or the market is full of nondurable consumption, holding these risks requires risk premia. However, as marginal utility is high when the economy is in bad times (durable consumption or housing consumption is low), asset-pricing theory suggests negative premia for these risks.

In Table 4, we first show that all the mimicking portfolios for nontradable factors, such as nondurable consumption growth (CG), durable consumption growth (DCG), housing consumption (HCG), the expenditure ratio (EXR) and liquidity innovation ($LIQINNOV$), are useful.

Panel B shows the horse race between the tradable and nontradable factors in the candidate SDFs to price the 230 pooled portfolios. For the gross returns, the misspecified measures for the Fama-French five plus durable consumption growth or plus housing consumption growth are quite close and smaller than others, even the Fama-French five plus up-minus-down momentum factor, which outperforms other tradable factors in Table 3. In the excess returns, the Yogo durable consumption growth factor is the least misspecified SDF in pricing the 230 portfolios, as it is the model with the lowest

HJ^M and HJ^C measures. These results suggest that given the Fama-French five common factors, the nontradable factors can help the SDFs decrease their misspecification error to an even greater extent than the up-minus-down momentum factor (a tradable factor). Overall, the results suggest that nontradable factors (housing consumption and durable consumption) outperform tradable factors (up-minus-down momentum) in pricing the large cross-section of risky assets with the least misspecified errors.

Panel C in Table 4 gives the multiple comparison test of whether the HJ^O , HJ^M and HJ^C measures of any of the alternative models are significantly greater than the distance measures of the candidate models. The null hypothesis in this test is that our chosen model (the FF5+UMD model)—the ‘winner’ SDF among the tradable factors—has HJ^O , HJ^M and HJ^C measures that are less than those of any of the alternative models. A failure to reject the null means that the FF5+UMD model is the least misspecified model. Then, we show the Chen and Ludvigson (2009) multiple comparison first and find that at the 5% significance level, the null hypothesis cannot be rejected. The test results followed by the Gospodinov et al. (2013) pairwise comparison test show that the last three models augmenting the Fama-French five-factor model with nontradable factors obtain smaller HJ distance measures at the 5% level of significance; their HJ^O , HJ^M and HJ^C measures are statistically even smaller than the ‘winner’ Fama-French five plus momentum model.

To explain why nontradable factors can improve SDF pricing errors, we conduct further a univariate predictive regression to check whether the nontradable factors can predict the future returns of the test assets significantly. For a specific nontradable factor and a given prediction period of k months, we first run univariate predictive regression of the future k -month returns of each portfolio on the current return of the factor mimicking portfolio of the nontradable factor and calculate the adjusted R-squared, then average across portfolios to arrive at the final result.

Here we focus on the durable consumption growth (DCG) and the housing consumption growth (HCG) factors based on the quarterly data from 1972 to 2011, and we consider prediction periods ranging from 1 quarter to 8 quarters. Table 5 reports the results. The two nontradable factors perform differently, where DCG mainly predicts the 1-quarter ahead asset returns, while HCG has a lower predictive power for the short-term future returns, but its impact is relatively long-lasting. For example, the average adjusted R-squared for the 8-quarter cumulative asset returns is 1.32%, which is much higher than that of a shorter prediction period.

4.2 The UK Market

We report the empirical results for the UK market in this section. Following the procedures for the US market described above, we first present the results for the tradable factors and related models, then turn to nontradable factors. The results are similar with those for the US market.

4.2.1. Sequential Selection Procedure and Rank Tests. We conduct tests for the tradable factors and related models based on the 122 pooled portfolios for the UK market. Table 6 presents results of the rank tests and model misspecification tests, as well as the sequential testing procedure, when only tradable factors are considered. The rank test results in Panel A shows that most of the tradable factors pass the rank test, except the *MKT* and *CMA* factors. The rank restriction test results of models in Panel B show that unlike those of the US market, many of the models with only tradable factors in UK do suffer from the identification problems. Moreover, the results of HJ distance tests and LM tests show that all models are potentially misspecified, as they all have significant HJ distance.

Panel C and Panel D of the table report the results of the sequential selection procedure under correct model specification and potential model misspecification assumptions, respectively. However, different from those of the US market, few factors survive the sequential selection procedure at the significance level of 5%. These results show that there is much room for the nontradable factors in UK.

4.2.2. HJ Distance Comparison Results. Table 7 further reports the model comparison results for models with only tradable factors. Similar to Table 2, we consider both the gross returns and excess returns of the 122 pooled portfolios. The results are similar to those of the US. When gross returns are considered, the FF5+UMD model always performs best with distance measures significantly lower than other models. While when excess returns are used, CAPM and FF3 models have a bit lower original HJ distance and modified HJ distance measures and the FF5CPM+UMD model has a slightly smaller constrained HJ distance. However, under these circumstances, the FF5+UMD model is still very competitive. As an interesting sidenote, the Carhart model performs better than the FF5 model for most of the situations, mainly due to the highly significance of the *UMD* factor and insignificance of the *CMA* factor in UK.

4.2.3. Nontradable Factors. We analyze the value-added of the nontradable factors based on 95 portfolios from January 1998 to December 2017.

Due to the smaller sample of nontradable factors, we exclude the 27 portfolios sorted by size, BM, and momentum, to ensure that we can get accurate estimates. Table 8 presents the test results. The rank test results in Panel A show that all nontradable factors alone can not pass the rank test. However, the model comparison results in Panel B show that, $FF5 + INF$, i.e., the Fama-French five-factor plus inflation model (FF5+INF) performs best under most circumstances, except that the FF5CPM+UMD model still has a slightly smaller constrained HJ distance when considering the excess returns. This shows that similar to that of the US, nontradable factors also help to improve SDF's pricing performance in UK, although the best performing factor is the inflation factor.

Then we conduct a predictive regression analysis to check how the nontradable factors help to explain the cross-section of asset returns. The test is similar to that of the US market in Section 4.1.3 but also with two differences. First, we use monthly data and we focus on the yield level and inflation factors in this test. Second, for ease of reading, we consider 12-month ahead prediction instead of 24-month ahead (which is equivalent to 8 quarters).

Table 9 reports the results. From the table, we can immediately figure out that both the two nontradable factors help to predict the future asset returns. Moreover, the predictive power drops with the increasing of prediction period, as the average adjusted R^2 declines. Since the explanatory variable is the current 1-month return of the factor's mimicking portfolio, we should not be surprised at the results. Lastly, although the results of the model comparison tests based on HJ distance suggest that the inflation factor performs better than the yield level, it does not guarantee that the former would always have higher predictive power than the latter one. And we indeed observe that for prediction periods ranging from 7 to 12 months, the yield level factor has a higher predictive power than the inflation factor. Moreover, for those longer prediction periods, the adjusted R-squared declines to nearly zero, showing that the predictive power for the very long periods is much smaller. Nevertheless, the results show that both factors have significant predictive power for the future cumulative returns of the test assets, even for the relatively long prediction periods.

4.3 The China Market

In this section, we report the empirical results for the UK market in this section. Following the procedures for the US and UK markets, we first present the results for the tradable factors and related models in Section 4.3.1, then we reports the model comparison test results in Section 4.3.2. Lastly, we further explore the performance of the nontradable factors in Sec-

tion 4.3.3. The results are very similar with those for the developed markets.

4.3.1. Sequential Selection Procedure and Rank Tests. Table 10 reports the results of the tests for tradable factors and related models using 83 portfolios as test assets. Panel A shows that all factors except BAB pass the rank restriction test, suggesting that most of these factors can be potentially useful in China, which is similar to the findings for the US and UK markets. Panel B indicates that all the candidate models may be misspecified based on the HJ distance measures and LM test, which is also consistent with the results for the US. However, the last two columns of Panel B show that a large number of the factor models cannot pass the rank test and hence may suffer from identification problems, which is similar to the situation in UK.

Panel C and Panel D of Table 10 present the results of the sequential selection procedure under correct model specification and potential model misspecification, respectively. Panel C shows that while all factors are insignificant in the CAPM, FF3, and Carhart 4 models, the *SMB*, *HML*, and *RMW* factors are significant in the FF5 and Fama-French five-factor plus up-minus-down momentum (FF5+UMD) models. The story is similar for the HXZ and HXZCP models, where the size factor (*ME*), profitability factors (*ROE* and *RMWC*), and market factor (*MKT*) are significant. However, when replacing *RMW* by *RMWCP* in FF5 and FF5+UMD, the profitability factor (*RMWC*) becomes insignificant, while the investment factor (*CMA*) becomes significant. When further replacing *HML* with the more updated version *HML^m*, four factors, including *MKT*, *SMB*, *HML^m* and *UMD*, are significant. The *BAB* factor and the two mispricing factors, i.e., *MGMT* and *PERF*, are insignificant. Moreover, no factors survive in all model specifications, with the *HML*, *SMB* and profitability factors (*ROE* and *RMW* factors) performing relatively better. Panel D shows that the results under potential model misspecification are similar. Hence, while the overall performance of factor models for the Chinese stock market is similar to that for the US market, there are also some differences.

4.3.2. HJ Distance Comparison Results. Table 11 reports the results of model comparison tests in China using both the gross returns and excess returns. Panel A reports the HJ distance measures for each model. The results using gross returns are very similar to those for the US market, where FF5+UMD has the smallest HJ distance and still outperforms when we consider the modified and constrained HJ distance using excess returns. However, the Fama-French five-factor with cash-based profitability plus up-minus-down momentum (FF5CP+UMD) model has the smallest distance when the original HJ distance is considered using excess returns. Nevertheless, FF5+UMD stands out under most circumstances. Panel B fur-

ther presents the results of the multiple model comparison test. These high p-values show that none of the alternative models have a smaller HJ distance than the models chosen above. Hence, the FF5+UMD model performs best under most circumstances, which is consistent with the results for both the US and UK markets.

4.3.3. Nontradable Factors. We also test how nontradable factors perform in China. Due to the availability of data on macro factors, the sample for the nontradable factors is shorter. Specifically, there are only 64 observations for consumption growth and wage growth and 154 observations for the other four factors; hence, we would encounter the “limited T versus large N” problem (Kleibergen and Zhan (2020)) if we still use the 83 portfolios as test assets to test the performance of the nontradable factors. Therefore, we conduct empirical tests using fewer portfolios as test assets. Specifically, we drop the 30 univariate sort portfolios and use the remaining 53 portfolios, i.e., 25 portfolios sorted by size and book-to-market ratio, and 28 industry portfolios as test assets. Under all circumstances, we always keep the industry portfolios in test assets, following Lewellen et al. (2010).

Table 12 reports the results. Panel A gives the results of the rank test. While consumption growth, wage growth, maturity spread, and credit spread are useless, the yield level and inflation factors are significant and therefore potentially useful. Panel B further presents the HJ distance measures for different models, where two models containing nontradable factors are added. The Fama-French five-factor plus yield level factor model (FF5+Yield) has the smallest HJ distance when we consider gross returns. Although it has larger original HJ distance and modified HJ distance measures than the FF5+UMD model when we consider excess returns, the FF5+Yield model is still very competitive under those circumstances. Panel C further presents the results of the multiple model comparison test. The FF5+Yield model has the smallest squared distance under all circumstances, and the high p-values indicate that it does perform best from the perspective of HJ distance. Hence, nontradable factors can help decrease the misspecification errors on SDFs given the Fama-French five-factor models as common factors and perform better than the factor models consisting only of tradable factors. In summary, the nontradable factors can help improve the performance of linear factor models for the Chinese market, which is consistent with the finding for the US market, although the nontradable factors proven to be useful may be different in the two markets.

We then check how the nontradable factors help improve the performance of SDFs in China. We conduct a predictive regression test, which is exactly what we have done for the UK market above. Table 13 shows the results of the predictive regression test. Again, we see a pattern similar

to that of the UK market reported in Table 9. Specifically, for both the yield level and inflation factors, the short-term predictive power is extremely high, while the predictive power drops a lot but still remains significant for longer prediction periods. And it is worth noting that the predictive power of the two nontradable factors are much higher than those of the UK market, for all prediction periods.

In summary, we find that the risk premium of nontradable factors has predictive power for future asset returns, and they work in the same way as the tradable factors. Specifically, they have the strongest predictive power for the short-term future asset returns, and the predictive power drops with the increasing of the prediction period.

5 Conclusions

In this paper, we use the HJ distance to compare the degree of misspecification among prominent linear factor models. We use recently developed multiple comparison inference to determine how well tradable or nontradable factors do in pricing a common set of risky portfolios.

We conclude that there is a unique proxy linear factor SDF that can consistently price a cross-section of large pooled risky portfolios. We also find cross-countries empirical evidences that nontradable factor mimicking portfolios can decrease misspecification errors when the common factors such as the Fama-French five-factor model are given. The risk premium of nontradable factors has predictive power for future asset returns, even for a relatively long prediction periods. Specifically, we arrive at similar results for the US, the UK and the China equity markets: models with only tradable factors, we show that the Fama-French five-factor plus up-minus-down momentum model (FF5+UMD) stands out under most circumstances. When nontradable factors are further considered, there always exist some nontradable factors that can improve the performance of SDFs and drive out the momentum factor, given the Fama-French five-factor model as benchmark.

References

- Adrian, T., E. Etula, and T. Muir (2014). Financial intermediaries and the cross-section of asset returns. *Journal of Finance* 69(6), 2557–2596.
- Asness, C. and A. Frazzini (2013). The devil in hml’s details. *Journal of Portfolio Management* 39(4), 49–68.
- Balduzzi, P. and C. Robotti (2008). Mimicking portfolios, economic risk premia, and tests of multi-beta models. *Journal of Business & Economic Statistics* 26(3), 354–368.
- Ball, R., J. Gerakos, J. T. Linnainmaa, and V. Nikolaev (2016). Accruals, cash flows, and operating profitability in the cross section of stock returns. *Journal of Financial Economics* 121(1), 28–45.
- Barillas, F., R. Kan, C. Robotti, and J. Shanken (2019). Model comparison with sharpe ratios. *Journal of Financial and Quantitative Analysis*, 1–35.
- Barillas, F. and J. Shanken (2017). Which alpha? *Review of Financial Studies* 30(4), 1316–1338.
- Barillas, F. and J. Shanken (2018). Comparing asset pricing models. *Journal of Finance* 73(2), 715–754.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance* 52(1), 57–82.
- Chen, N., R. Roll, and S. Ross (1986). Economic forces and the stock market. *Journal of Business*, 383–403.
- Chen, X. and S. Ludvigson (2009). Land of addicts? an empirical investigation of habit-based asset pricing models. *Journal of Applied Econometrics* 24(7), 1057–1093.
- Cochrane, J. H. and L. P. Hansen (1993). Asset pricing explorations for macroeconomics. *NBER Working Paper* (w4088).
- Fama, E. and K. French (1992). The cross-section of expected stock returns. *Journal of Finance* 47(2), 427–465.
- 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 K. R. French (2018). Choosing factors. *Journal of Financial Economics* 128(2), 234–252.
- Ferson, W., A. F. Siegel, and P. Xu (2006). Mimicking portfolios with conditioning information. *Journal of Financial and Quantitative Analysis*, 607–635.
- Frazzini, A. and L. H. Pedersen (2014). Betting against beta. *Journal of Financial Economics* 111(1), 1–25.
- Gospodinov, N., R. Kan, and C. Robotti (2013). Chi-squared tests for evaluation and comparison of asset pricing models. *Journal of Econometrics* 173(1), 108–125.
- Gospodinov, N., R. Kan, and C. Robotti (2014). Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors. *Review of Financial Studies* 27(7), 2139–2170.

- Gospodinov, N., R. Kan, and C. Robotti (2017). Spurious inference in reduced-rank asset-pricing models. *Econometrica* 85(5), 1613–1628.
- Gospodinov, N., C. Robotti, and R. Kan (2012). On the Hansen-Jagannathan distance with a no-arbitrage constraint. *Available at SSRN 1534727*.
- Griffin, J. M., X. Ji, and J. S. Martin (2003). Momentum investing and business cycle risk: Evidence from pole to pole. *The Journal of Finance* 58(6), 2515–2547.
- Hanauer, M. X. (2020). A comparison of global factor models. *Available at SSRN 3546295*.
- Hansen, L. and R. Jagannathan (1997). Assessing specification errors in stochastic discount factor models. *Journal of Finance*, 557–590.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.
- Harvey, C. R., Y. Liu, and H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies* 29(1), 5–68.
- Hodrick, R. and X. Zhang (2001). Evaluating the specification errors of asset pricing models. *Journal of Financial Economics* 62(2), 327–376.
- Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *Review of Financial Studies* 28(3).
- Hou, K., C. Xue, and L. Zhang (2017). Replicating anomalies. *Review of Financial Studies*.
- Jagannathan, R. and Z. Wang (1996). The conditional capm and the cross-section of expected returns. *Journal of Finance*, 3–53.
- Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48(1), 65–91.
- Jurczenko, E. and J. Teiletche (2019). Macro factor mimicking portfolios. *Available at SSRN 3363598*.
- Kan, R. and C. Robotti (2008). Specification tests of asset pricing models using excess returns. *Journal of Empirical Finance* 15(5), 816–838.
- Kleibergen, F. and Z. Zhan (2018). Identification-robust inference on risk premia of mimicking portfolios of non-traded factors. *Journal of Financial Econometrics* 16(2), 155–190.
- Kleibergen, F. and Z. Zhan (2020). Robust inference for consumption-based asset pricing. *Journal of Finance*.
- Kong, L. (2019). Weak (proxy) factors robust Hansen-Jagannathan distance for linear asset pricing models. *Working Paper*.
- Kudo, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* 50(3-4), 403–418.
- Lettau, M. and S. Ludvigson (2001). Resurrecting the (c) capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109(6), 1238–1287.
- Lewellen, J., S. Nagel, and J. Shanken (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics* 96(2), 175–194.

- Li, H., Y. Xu, and X. Zhang (2010). Evaluating asset pricing models using the second Hansen-Jagannathan distance. *Journal of Financial Economics* 97(2), 279–301.
- Liu, J., R. F. Stambaugh, and Y. Yuan (2019). Size and value in china. *Journal of Financial Economics* 134(1), 48–69.
- Ludvigson, S. C. (2013). Advances in consumption-based asset pricing: Empirical tests. In *Handbook of the Economics of Finance*, Volume 2, pp. 799–906. Elsevier.
- Lustig, H. and S. Van Nieuwerburgh (2005). Housing collateral, consumption insurance, and risk premia: An empirical perspective. *Journal of Finance* 60(3), 1167–1219.
- McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *Journal of Finance* 71(1), 5–32.
- Parker, J. and C. Julliard (2005). Consumption risk and the cross section of expected returns. *Journal of Political Economy* 113(1), 185–222.
- Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy* 111(3), 642–685.
- Piazzesi, M., M. Schneider, and S. Tuzel (2007). Housing, consumption and asset pricing. *Journal of Financial Economics* 83(3), 531–569.
- Pukthuanthong, K., R. Roll, J. L. Wang, and T. Zhang (2019). A toolkit for factor-mimicking portfolios. Available at SSRN 3341604.
- Ross, S. (1973). *The arbitrage theory of capital asset pricing*. Rodney L. White Center for Financial Research, University of Pennsylvania, The Wharton School.
- Stambaugh, R. F. and Y. Yuan (2016). Mispricing factors. *Review of Financial Studies* 30(4), 1270–1315.
- Vassalou, M. and Y. Xing (2004). Default risk in equity returns. *Journal of Finance* 59(2), 831–868.
- Wang, H. and J. Yu (2013). Dissecting the profitability premium. In *AFA 2013 San Diego Meetings Paper*.
- Wang, Y. (2005). A comparison of factor models for explaining the cross section of stock returns. Available at SSRN 769084.
- White, H. (2003). A reality check for data snooping. *Econometrica* 68(5), 1097–1126.
- Wolak, F. A. (1987). An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association* 82(399), 782–793.
- Wolak, F. A. (1989). Testing inequality constraints in linear econometric models. *Journal of Econometrics* 41(2), 205–235.
- Yogo, M. (2006). A consumption-based explanation of expected stock returns. *Journal of Finance* 61(2), 539–580.
- Zhang, X. (2006). Specification tests of international asset pricing models. *Journal of International Money and Finance* 25(2), 275–307.

A Appendix

A.1 Sample Estimates on the Hansen–Jagannathan Distance

In sample estimation, if the test portfolios are in gross returns, we can define

$$D_T = \frac{1}{T} \sum_{t=1}^T \frac{\partial \alpha_t(\theta)}{\partial \theta} = \frac{1}{T} R' f, \quad (\text{A.1})$$

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^T \alpha_t(\theta) = D_T \theta - I_N, \quad (\text{A.2})$$

$$G_T = \frac{1}{T} \sum_{t=1}^T R_t R_t' = \frac{1}{T} R' R, \quad (\text{A.3})$$

where

$$R = [R_1, R_2, \dots, R_T]',$$

$$f = [f_1, f_2, \dots, f_T].$$

The sample analog of the HJ distance is thus

$$\delta_T = \sqrt{\min_{\theta} g_T(\theta)' G_T^{-1} g_T(\theta)}. \quad (\text{A.4})$$

Taking the derivative of the above equation

$$D_T' G_T^{-1} g_T(\theta) = 0, \quad (\text{A.5})$$

which gives an analytic expression for the sample minimizer

$$\hat{\theta} = (D_T' G_T^{-1} D_T)^{-1} D_T' G_T^{-1} I_N. \quad (\text{A.6})$$

From Hansen (1982) the asymptotic variance of $\hat{\theta}$ is given by

$$\text{var}(\hat{\theta}) = \frac{1}{T} (D_T' G_T^{-1} D_T)^{-1} D_T' G_T^{-1} \Omega_T G_T^{-1} D_T (D_T' G_T^{-1} D_T)^{-1}, \quad (\text{A.7})$$

where, if the data is serially uncorrelated, the estimate of the variance matrix of pricing errors is given by

$$\Omega_T = \frac{1}{T} \sum_{t=1}^T \alpha_t(\hat{\theta}) \alpha_t(\hat{\theta})'. \quad (\text{A.8})$$

That is the estimator $\hat{\theta}$ that is equivalent to a GMM estimator defined by Hansen(1982) with the moment condition $E[g(\theta)] = 0$ and the weighting matrix G^{-1} .

Following Kan and Robotti (2008), if the test portfolios are in excess returns, we can define

$$y_{t+1}(\theta) = 1 - \theta' f_{t+1}, \quad (\text{A.9})$$

$$E_t[y_{t+1}(\theta) R_{t+1}] = 0_N, \quad (\text{A.10})$$

the estimates of risk premiums will change into

$$\hat{\theta} = -(D'_t G_t^{-1} D_t)^{-1} D'_t G_t^{-1} \bar{R}_t, \quad (\text{A.11})$$

where \bar{R}_t is the average excess return across N .

A.2 Testing the Hansen–Jagannathan Distance

If the weighting matrix is optimal in the sense of Hansen (1982), then $T\delta_T^2$ is asymptotically a random variable of χ^2 distribution with $N - K$ freedom, where is the dimension of θ .

However, if G is generally not optimal, $T\delta_T^2$ is not asymptotically a random variable of χ^2 . Instead, under the hypothesis that the SDF prices the returns correctly, the sample HJ distance follows:

$$T[\hat{\delta}^2] \xrightarrow{d} \sum_{j=1}^{N-K} \alpha_j \chi^2(1), \quad (\text{A.12})$$

where $\chi^2(1)$ are independent chi-squared random variables with one degree of freedom, and α_j are $N - K$ nonzero eigenvalues of the matrix A given by

$$A = \Omega^{\frac{1}{2}} G^{\frac{-1}{2}} [I_N - (G^{\frac{-1}{2}})' D (D' G^{-1} D)^{-1} D' G^{\frac{-1}{2}}] (G^{\frac{1}{2}})' (\Omega^{\frac{1}{2}})'. \quad (\text{A.13})$$

Here $\Omega = E[\alpha_t \alpha_t']$ denotes the variance of pricing errors, and $D = E(R'_t f_t)$. The $\frac{1}{2}$ means the upper-triangle matrices from the Cholesky decomposition. As long as we have a consistent estimate Ω_T of the matrix Ω , we can estimate the matrix A by replacing Ω and G by Ω_T and G_T , respectively. Under the hypothesis that the SDF prices the returns correctly, The Ω can be estimated consistently by $\Omega_T = T^{-1} \sum_{t=1}^T [\alpha_t \alpha_t']$.

Following Jagannathan and Wang (1996), to adjust for the small sample bias, we use Monte Carlo method to calculate the empirical distribution of HJ distance (under the null hypothesis). First, draw $M \otimes (N - K)$ independent random variables from $\chi^2(1)$ distribution. Then, calculate $u_j = \sum_{i=1}^{N-K} \alpha_i \chi^2(1)$. Here M is the number of simulation. Then the empirical p-value of the HJ distance is

$$p_{\hat{H}J} = \frac{1}{M} \sum_{j=1}^M I_{(u_j \geq T[HJ_T(\theta_T)]^2)}, \quad (\text{A.14})$$

where $I(\cdot)$ is an indicator function which equals one if the expression in the brackets is true and zero otherwise.

A.3 Testing the Constrained Hansen–Jagannathan Distance

To test the constrained HJ distance, we follow Gospodinov, Kan and Robotti (Gospodinov et al. (2012)). They state an asset pricing model is correctly specified if there exists a $\theta \in \Gamma$ such that $y_t(\theta) \in \mathbb{N}^+$, which implies that $\iota = 0_N$ and $\delta_+ = 0$; the model is misspecified if $y_t(\theta) \notin \mathbb{N}^+$ for all $\theta \in \Gamma$, which implies that $\delta_+ > 0$.

They show that

(a) if $\delta_+ = 0$, the pricing model is correctly specified,

$$T\hat{\delta}_+^2 \xrightarrow{A} \sum_{i=1}^{N-K} \zeta_i v_i, \quad (\text{A.15})$$

where the v_i are independent chi-squared random variables with one degree of freedom and the ζ_i are the eigenvalues of

$$A = P' U^{-\frac{1}{2}} S U^{-\frac{1}{2}} P, \quad (\text{A.16})$$

with $S = \sum_{j=-\infty}^{\infty} E[(x_t y_t(\theta^*) - q_{t-1})(x_{t+j} y_{t+j}(\theta^*) - q_{t+j-1})']$, $D = E[x_t \frac{\partial y_t(\theta^*)}{\partial \theta'}]$, $U = E[x_t x_t']$, and P being an $N \times (N - K)$ orthonormal matrix whose columns are orthogonal to $U^{-\frac{1}{2}} D$. This is the same as traditional HJ distance test.

(b) if $\delta_+ > 0$, the pricing model is misspecified,

$$\sqrt{T}(\hat{\delta}_+^2 - \delta_+^2) \xrightarrow{A} N(0, v), \quad (\text{A.17})$$

where $v = \sum_{j=-\infty}^{\infty} E[(\varphi_t(\lambda^*) - \delta_+^2)(\varphi_{t+j}(\lambda^*) - \delta_+^2)']$ and $\delta = [\theta', t']$.

To conduct inference, the variance matrix should be replaced by consistent estimator. In sample, we can replace A with \hat{A} , and $\hat{U} = \frac{1}{T} \sum_{t=1}^T x_t x_t'$, we also can obtain \hat{S} using a nonparametric heteroskedasticity and autocorrelation consistent estimator.

A.4 Multiple Model Comparison Test

Gospodinov et al. (2013) provide a multiple model comparison test that allows us to compare a benchmark model with a set of alternative models in terms of their HJ distance metrics. They suggest that we should separate models into three categories: nested, strictly non-nested and overlapping. For non-nested and overlapping models they introduce a multivariate inequality test based on Wolak (1987, 1989).

Let $\rho = (\rho_2, \dots, \rho_{p+1})$, where $\rho_i = \delta_1^2 - \delta_i^2$. We set δ_1^2 as the winner, and test $H_0: \rho \leq 0_p$. We assume that

$$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{A} N(0_p, \Omega_{\hat{\rho}}). \quad (\text{A.18})$$

Let $\tilde{\rho}$ be the optimal solution in the following quadratic programming problem:

$$\min_{\rho} (\hat{\rho} - \rho)' \hat{\Omega}_{\hat{\rho}}^{-1} (\hat{\rho} - \rho), \quad (\text{A.19})$$

$$s.t. \rho \leq 0_r, \quad (\text{A.20})$$

where $\hat{\Omega}_{\hat{\rho}}^{-1}$ is a consistent estimator of $\Omega_{\hat{\rho}}^{-1}$. The likelihood ratio test of the null hypothesis is

$$LR = T(\hat{\rho} - \tilde{\rho})' \hat{\Omega}_{\hat{\rho}}^{-1} (\hat{\rho} - \tilde{\rho}). \quad (\text{A.21})$$

Since the null hypothesis is composite, to construct a test with the desired size,

they require the distribution of LR under the least favorable value of ρ , which is $\rho = 0_p$. Under this value, LR follows a ‘chi-bar-squared distribution’,

$$LR \xrightarrow{A} \sum_{i=0}^p w_i(\Omega_{\hat{\rho}}^{-1})X_i, \quad (\text{A.22})$$

where the X_i are independent χ^2 random variables with i degrees of freedom and χ_0^2 is simply defined as the constant zero. An explicit formula for the weights $w_i(\Omega_{\hat{\rho}}^{-1})$ is given in Kudo (1963).

For nested models, Gospodinov et al. (2013) suppose that $y_i^A(\lambda_1^*) = y_i^i(\lambda_i^*)$ can be written as a parametric restriction of the form $\varphi_i(\lambda_i^*) = 0_{k_i - k_1}$, where $\varphi(\cdot)$ is a twice continuously differentiable function in its argument. The null hypothesis for multiple model comparison can therefore be formulated as $H_0 : \varphi_2 = 0_{k_2 - k_1}, \dots, \varphi_{p+1}(\lambda_{p+1}^*) = 0_{k_{p+1} - k_1}$. The comparison test statistic follows Wald test with the degree of freedom $(\sum_{i=2}^{p+1} k_i - pk_1)$.

Besides the new tests developed in Gospodinov et al. (2013) on multiple model comparison test, we also apply the Chen and Ludvigson (2009)’s test method. We denote the squared HJ distance for model j as

$$\delta_{1,T}^2 = \min(d_{j,T}^2)_{j=1}^K. \quad (\text{A.23})$$

Hence, the null hypothesis is stated as follows:

$$H_0 : \delta_{1,T}^2 - \delta_{2,T}^2 \leq 0,$$

where $d_{2,T}^2$ is the competing model with the next smallest squared distance. Now we define the test statistic as $T^W = \max_{2, \dots, 5} \sqrt{T}(d_{1,T}^2 - d_{j,T}^2)$, based on White (2003). The distribution of T^W is computed via block bootstrap. We note that the justification for the bootstrap rests on the existence of a multivariate, joint, continuous, limiting distribution for the set $(d_{j,T}^2)_{j=1}^K$ under the null.

By repeated sampling, the bootstrap estimate of the p -value is

$$p_W = \frac{1}{B} \sum_{b=1}^B I_{(T^{W,b} > T^W)}, \quad (\text{A.24})$$

where B is the number of bootstrap samples and $T^{W,b}$ stands for White’s original bootstrap test statistic. If the null is true, the historical value of T^W should not be unusually large, given sampling error. Given the distribution of T^W , reject the null if its historical value, T^W , is greater than the 95th percentile of the distributions for T^W . At a 5% level of significance, we reject the null if p_W is less than 0.05, but do not reject otherwise.

A.5 Factor Models in China

We replicate all the 11 models based on Chinese data except the Fama-French three-factor plus the liquidity factor model (FF3+LIQ), with a few minor differences in details due to data consideration. In this appendix, we give the details of the differences with the original models when we replicate those models in China. Except

for what mentioned as follows, we follow the procedures described in the original papers. Other tests show that these differences have little impact on the results.

- For those factors rebalanced annually, e.g., SMB and HML factors of the Fama-French five-factor model, we rebalance portfolio at the end of April, rather than the end of June. As the companies listed in the China's A-Share stock market are required to disclose their annual report no later than the end of April following the corresponding fiscal year, we can make use of the information in a more timely manner and construct a more robust factor by rebalancing portfolios at the end of April.
- For the MOM and UMD factors, we simply construct monthly rebalancing portfolios instead of calculating the average returns of a series of portfolios with a longer holding period.
- For the HXZ and HXZCP models, we construct monthly rebalancing portfolios using the most recent data available, while Hou et al. (2015) group stocks based on the most recent quarterly ROE, as well as the yearly market capitalization and investment.
- For Stambaugh and Yuan (2016)'s mispricing factors, we construct the two mispricing factors based on eight, instead of eleven anomalies. We do not consider financial distress, net issuance, and composite equity issues, due to lack of data.

A.6 Macro Factors in China

We consider six macro factors in this paper. In this appendix, we describe how the consumption growth, wage growth, and credit spread factors are constructed. The other three factors are defined in the same way as those for the UK market.

- Consumption growth (CG). As there are no historical data of quarterly consumption in China, we measure consumption as the total consumption of urban populations scaled by GDP. Specifically, the calculation of the consumption growth is as follows:
 - First, we interpolate the annual data of the urban population to get the quarterly urban population.
 - Then we multiply the urban per capita consumption by the urban population and arrive at the total quarterly consumption of urban populations.
 - Next, we scale the total consumption of urban populations by the GDP of the corresponding quarter, arriving at the share of urban consumption.
 - Finally, we calculate the quarter-over-quarter growth of the share of urban consumption and get consumption growth.
- Wage growth (WG). We calculate wage growth as the year-over-year growth of the total wage of urban populations. Moreover, as quarterly total wage and

wage growth data is unavailable for 2014 to 2018, we interpolate year-over-year wage growth data for that period.

- Credit spread (CS). We calculate the credit spread as the difference between the YTM of the AAA-rating enterprise bond and government bond with the same maturity. We calculate the spread for 1-year, 5-year, and 10-year, and then average across these maturities.

Table 1: Rank Test and Misspecification Identification in US

Notes: the table presents the results of the ranks tests of the individual factors, the model misspecification tests, and rank tests of the models in US. The models are estimated using monthly returns from 1967 to 2015 on the 230 pooled portfolios. Panel A reports the rank restriction test (w) and its p-value of the null that $E[x_i(1, fit)]$ has a column rank of one. Panel B reports the sample HJ distance ($\hat{\delta}$), the Lagrange multiplier (LM) test, and the rank restriction test (W^*) with the corresponding p-values for each model. Panels C and D show t-tests of the model selection procedures based on the standard errors under correct model specification and model misspecification, respectively. The boldface denotes the 5% significance level.

Panel A: Rank test for individual factors																	
Test	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	LIQTRADED	BAB	MGMT	PERF	
w	206.6	214.8	200.0	207.5	199.6	186.2	215.7	192.5	160.5	115.6	115.5	189.5	117.9	153.2	199.1	190.3	
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.016	0.000	0.011	0.000	0.000	0.000	
Panel B: HJ-distance, Lagrange multiplier, and rank tests																	
Model	delta	p-val	LM	p-val	w	p-val											
CAPM	0.701	0.00	176.91	0.00	206.6	0.00											
FF3	0.680	0.00	172.00	0.00	198.2	0.00											
FF3+LIQ	0.673	0.00	168.79	0.00	115.7	0.02											
BAB	0.686	0.00	167.90	0.00	147.6	0.00											
Carhart	0.648	0.00	154.67	0.00	102.3	0.06											
HXZ	0.684	0.00	171.78	0.00	122.3	0.00											
SY	0.628	0.00	142.00	0.00	132.0	0.00											
FF5	0.625	0.00	145.89	0.00	140.7	0.00											
FF5+UMD	0.607	0.00	140.21	0.00	149.5	0.00											
HXZCP	0.666	0.00	168.31	0.00	150.7	0.00											
FF5CPM+UMD	0.615	0.00	136.53	0.00	135.5	0.00											

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	LIQTRADED	BAB	MGMT	PERF
CAPM	-2.51															
FF3	-3.03	-1.03	-3.64													
	-3.30		-3.51													
FF3+LIQ	-3.11	-0.99	-3.69										-2.03			
	-2.60												-2.23			
	-2.36	-0.83											-2.21			
	-3.38		-3.58										-2.05			
BAB	-2.49														-2.93	
FF5	-4.09	-2.77	-3.55		0.28	-5.68										
	-2.81	-1.03	-2.17		0.11											
	-4.46	-2.73	-5.72			-5.69										
Carhart	-3.47	-0.98	-4.12													
	-2.83															
HXZ	-2.99						-2.06	-2.35	-2.20							
	-2.18						-1.23									
	-2.81						-1.33	-2.26								
	-2.22						-2.20		-2.18							
	-3.29							-2.15	-1.55							

Table continued.

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	LIQTRADED	BAB	MGMT	PERF
HXZCP	-4.39					-4.15		-3.33	0.90							
	-3.36					-3.46										
	-4.35					-4.34		-3.17								
	-3.34					-3.21			0.53							
SY	-4.49		-2.96												-4.86	-3.65
	-2.08		-1.77													
	-2.75		-1.89												-3.08	
	-3.77		-2.67													-4.50
FF5+UMD	-3.86	-2.55	-4.01		0.86	-4.73					-2.38					
FF5CPM+UMD	-4.17	-2.13			0.18	-4.99					-3.84	-3.65				

Panel D: Model selection procedure using model misspecification-robust standard errors

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	LIQTRADED	BAB	MGMT	PERF
CAPM	-2.51															
FF3	-3.02	-1.02	-3.63													
	-3.29		-3.51													
FF3+LIQ	-3.09	-0.98	-3.68													-1.51
	-2.59															-1.61
	-2.35	-0.83														-1.59
	-3.36	-3.57														-1.54

BAB	-2.49									-2.73
FF5	-3.95	-2.80	-3.19	0.24	-5.24					
	-2.74	-1.04	-1.89	0.09						
	-4.36	-2.73	-5.59		-5.25					
Carhart	-3.42	-0.98	-4.01						-3.01	
	-2.83								-2.48	
HXZ	-2.94					-1.98	-2.14	-2.06		
	-2.18					-1.22				
	-2.41					-1.81	-1.89			
	-2.22					-2.09		-1.92		
	-3.22						-1.94	-1.48		
HXZCP	-4.16				-3.58		-2.89	0.84		
	-3.33				-3.12					
	-4.14				-3.80		-2.79			
	-3.31				-2.84			0.49		
SY	-4.31		-2.89						-4.46	-3.43
	-2.08		-1.73							
	-2.73		-1.86						-2.86	
	-3.63		-2.60							-4.07
FF5+UMD	-3.76	-2.55	-3.71	0.77					-4.24	-2.30
FF5CP+UMD	-4.05	-2.14		0.15	-4.48				-3.32	-3.21

Table 2: HJ Distance and Large Cross-sectional Portfolios in US

Notes: the table represents the results of tradable factors SDF comparison tests using both gross returns and excess returns on the 230 pooled portfolios: 25 portfolios sorted by size and book-to-market ratio, 30 industry portfolios, 100 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, 25 portfolios sorted by size and momentum, and 25 portfolios sorted by size and beta. Panel A reports the HJ distance (HJ^O), the modified HJ distance (HJ^M) and the constrained HJ distance (HJ^C) measures. Panel B presents the results of the two formal tests of model comparison. The null hypothesis in the first test is that the chosen model has a HJ^O , HJ^M and HJ^C measure that is less than that of any of the alternative models. The second test is a pairwise comparison test of whether HJ^O , HJ^M and HJ^C measures of each of the alternative models are significantly greater than the distance measures of the chosen model.

Pooled Portfolios	Gross Returns			Excess Returns			
	HJ^O	HJ^M	HJ^C	HJ^O	HJ^M	HJ^C	
Panel A: HJ Measures							
CAPM	0.8841	0.8841	0.9631	CAPM	0.6849	0.9397	0.5562
FF3	0.8736	0.8736	0.9535	FF3	0.6817	0.929	0.5518
FF5	0.8388	0.8388	0.925	FF5	0.6761	0.8915	0.5446
Carhart	0.8541	0.8541	0.9361	Carhart	0.6734	0.9026	0.5505
HXZ	0.8672	0.8672	0.9482	HXZ	0.6939	0.9413	0.5487
HXZCP	0.8467	0.8467	0.9344	HXZCP	0.6841	0.9046	0.5472
FF3+LIQ	0.8732	0.8732	0.9534	FF3+LIQ	0.6847	0.9337	0.5518
BAB	0.8821	0.8821	0.9609	BAB	0.6875	0.9362	0.5522
SY	0.8417	0.8417	0.9328	SY	0.6718	0.8845	0.5478
FF5+UMD	0.8276	0.8276	0.914	FF5+UMD	0.6688	0.8759	0.5422
FFCP+UMD	0.828	0.828	0.9147	FFCP+UMD	0.6711	0.8763	0.5423
Panel B: HJ Multiple Comparison							
FF5+UMD<(FFCPUMD/FF5/SY/HXZCP				FF5+UMD<(FFCPUMD/SY/FF5/Carhart			
Carhart/HXZ/FF3+LIQ/FF3/BAB/CAPM)				HXZCP/FF3/FF3+LIQ/BAB/CAPM/HXZ)			
H_0 (p-value)	0.306	0.316	0.32	H_0 (p-value)	0.3134	0.19	0.302
FF5+UMD <				FF5+UMD <			
FF5 (p-value)	0.054	0.054	0.23	FF5CPM+UMD (p-value)	0.099	0.067	0.070
FF5CPM+UMD (p-value)	0.129	0.127	0.138	SY (p-value)	0.138	0.085	0.092
SY (p-value)	0.131	0.131	0.234	FF5 (p-value)	0.132	0.98	0.18
HXZCP (p-value)	0.153	0.153	0.182	HXZCP (p-value)	0.096	0.123	0.212
Carhart (p-value)	0.124	0.124	0.212	Carhart (p-value)	0.109	0.14	0.162
HXZ (p-value)	0.138	0.138	0.124	FF3+LIQ (p-value)	0.145	0.157	0.181
FF3+LIQ (p-value)	0.131	0.131	0.236	FF3 (p-value)	0.137	0.164	0.146
FF3 (p-value)	0.100	0.100	0.184	HXZ (p-value)	0.103	0.173	0.133
BAB (p-value)	0.131	0.131	0.222	BAB (p-value)	0.17	0.17	0.16
CAPM (p-value)	0.153	0.153	0.227	CAPM (p-value)	0.198	0.198	0.139

Table 3: HJ Distance and Specific Cross-sectional Portfolios in US

Notes: the table represents the results of the tradable factors SDFs comparison in US using both gross returns and excess returns on 25 portfolios sorted by size and book-to-market ratio, 30 industry portfolios, 100 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and operating profitability and 25 portfolios sorted by size and momentum. The table reports the HJ distance (HJ^O), the modified HJ distance (HJ^M) and the constrained HJ distance (HJ^C) measures. Then the table presents the results of the multiple comparison. The null hypothesis is that the chosen model has a HJ^O , HJ^M and HJ^C measure that is less than that of any of the alternative models.

	Gross Returns			Excess Returns			
	HJ^O	HJ^M	HJ^C	HJ^O	HJ^M	HJ^C	
Panel A: Size/BM/OP Sorted							
CAPM	0.573	0.573	0.578	CAPM	0.554	0.666	0.303
FF3	0.556	0.556	0.564	FF3	0.5418	0.6439	0.295
FF5	0.515	0.515	0.5265	FF5	0.5034	0.5785	0.285
Carhart	0.543	0.5432	0.5541	Carhart	0.5201	0.6062	0.2949
HXZ	0.5554	0.5554	0.5649	HXZ	0.5497	0.6521	0.2904
HXZCP	0.5308	0.5308	0.544	HXZCP	0.5194	0.6014	0.2891
FF3+LIQ	0.5555	0.5555	0.5634	FF3+LIQ	0.541	0.6415	0.2944
BAB	0.57	0.57	0.578	BAB	0.5515	0.6583	0.2963
SY	0.5275	0.5275	0.544	SY	0.4993	0.5725	0.2901
FF5+UMD	0.5067	0.5067	0.5201	FF5+UMD	0.4889	0.5576	0.2851
FFCPM+UMD	0.5161	0.5161	0.5304	FFCPM+UMD	0.4947	0.5656	0.2878
FF5+UMD<(FF5/FFCPM+UMD/SY/HXZCP				FF5+UMD<(FFCPM+UMD/SY/FF5/HXZCP			
Carhart/HXZ/FF3+LIQ/FF3/BAB/CAPM)				Carhart/FF3+LIQ/FF3/HXZ/BAB/CAPM)			
H_0 (p-value)	0.172	0.221	0.191	H_0 (p-value)	0.108	0.124	0.271
Panel B: Size/BM Sorted							
CAPM	0.3636	0.3636	0.3641	CAPM	0.4181	0.4602	0.1752
FF3	0.34	0.34	0.3407	FF3	0.3937	0.4281	0.1626
FF5	0.2911	0.2911	0.3068	FF5	0.3291	0.3494	0.1585
Carhart	0.3167	0.3167	0.3243	Carhart	0.3338	0.3542	0.1626
HXZ	0.3348	0.3348	0.3382	HXZ	0.3771	0.4061	0.1605
HXZCP	0.3135	0.3135	0.3236	HXZCP	0.3456	0.3686	0.1603

FF3+LIQ	0.3385	0.3385	0.3393	FF3+LIQ	0.3846	0.4161	0.1626
BAB	0.3624	0.3624	0.3629	BAB	0.394	0.4279	0.1726
SY	0.3129	0.3129	0.3216	SY	0.3221	0.3411	0.1588
FF5+UMD	0.2743	0.2743	0.2973	FF5+UMD	0.2947	0.311	0.1584
FFCPM+UMD	0.2883	0.2883	0.3085	FFCPM+UMD	0.3106	0.33	0.1606
FF5+UMD<(FFCPM+UMD/FF5/SY/HXZCP Carhart/HXZ/FF3+LIQ/FF3/BAB/CAPM)				FF5+UMD<(FFCPM+UMD/SY/FF5/HXZCP Carhart/FF3+LIQ/HXZ/FF3/BAB/CAPM)			
H_0 (p-value)	0.143	0.102	0.231	H_0 (p-value)	0.167	0.112	0.271
Panel C: Size/Mom Sorted							
CAPM	0.4187	0.4187	0.4271	CAPM	0.4276	0.473	0.2289
FF3	0.3937	0.3937	0.4079	FF3	0.4186	0.4607	0.2
FF5	0.3286	0.3286	0.3712	FF5	0.3344	0.3564	0.1993
Carhart	0.3733	0.3733	0.3824	Carhart	0.3671	0.3943	0.1998
HXZ	0.3441	0.3441	0.3534	HXZ	0.3324	0.3527	0.2069
HXZCP	0.3669	0.3669	0.3832	HXZCP	0.3559	0.3813	0.2092
FF3+LIQ	0.3878	0.3878	0.4018	FF3+LIQ	0.4183	0.4603	0.1988
BAB	0.388	0.388	0.3929	BAB	0.3642	0.3907	0.227
SY	0.3637	0.3637	0.3792	SY	0.3488	0.3722	0.1975
FF5+UMD	0.3285	0.3285	0.3606	FF5+UMD	0.3427	0.3664	0.1986
FFCPM+UMD	0.2913	0.2913	0.3411	FFCPM+UMD	0.3162	0.3355	0.1947
FFCPM+UMD<(FF5+UMD/FF5/HXZ/SY/HXZCP Carhart/FF3+LIQ/FF3/BAB/CAPM)				FFCPM+UMD<(FF5+UMD/FF5/HXZ/SY/HXZCP BAB/Carhart/FF3+LIQ/FF3/BAB/CAPM)			
H_0 (p-value)	0.119	0.129	0.133	H_0 (p-value)	0.108	0.321	0.22

Table continued.

Panel D: Industry Sorted

CAPM	0.2229	0.2229	0.223	CAPM	0.2359	0.2428	0.2134
FF3	0.2167	0.2167	0.2176	FF3	0.2207	0.2263	0.2005
FF5	0.1935	0.1935	0.1994	FF5	0.1974	0.2014	0.191
Carhart	0.2151	0.2151	0.2162	Carhart	0.2224	0.2281	0.1969
HXZ	0.2133	0.2133	0.2146	HXZ	0.2178	0.2232	0.1981
HXZCP	0.2041	0.2041	0.2052	HXZCP	0.2073	0.2118	0.1981
FF3+LIQ	0.2157	0.2157	0.2166	FF3+LIQ	0.2182	0.2236	0.1979
BAB	0.2226	0.2226	0.2226	BAB	0.234	0.2407	0.1974
SY	0.2115	0.2115	0.2121	SY	0.2133	0.2183	0.1968
FF5+UMD	0.1929	0.1929	0.1987	FF5+UMD	0.1981	0.2021	0.1894
FFCPM+UMD	0.1918	0.1918	0.1956	FFCPM+UMD	0.1967	0.2006	0.1907
FFCPM+UMD<(FF5+UMD/FF5/HXZCP/SY HXZ/Carhart/FF3+LIQ/FF3/BAB/CAPM)				FFCPM+UMD<(FF5+UMD/FF5/HXZCP/SY HXZ/FF3+LIQ/FF3/Carhart/BAB/CAPM)			
H_0 (p-value)	0.365	0.23	0.2666	H_0 (p-value)	0.163	0.212	0.240

Panel E: OP Sorted

CAPM	0.0935	0.0935	0.0935	CAPM	0.2359	0.2428	0.2134
FF3	0.0788	0.0788	0.08	FF3	0.2207	0.2263	0.2005
FF5	0.046	0.046	0.0794	FF5	0.1974	0.2014	0.191
Carhart	0.0768	0.0768	0.0768	Carhart	0.2224	0.2281	0.1969
HXZ	0.0801	0.0801	0.0805	HXZ	0.2178	0.2232	0.1981
HXZCP	0.0791	0.0791	0.0814	HXZCP	0.2073	0.2118	0.1981
FF3+LIQ	0.0768	0.0768	0.0792	FF3+LIQ	0.2182	0.2236	0.1979
BAB	0.083	0.083	0.0907	BAB	0.234	0.2407	0.1974
SY	0.071	0.071	0.0767	SY	0.2133	0.2183	0.1968
FF5+UMD	0.046	0.046	0.0679	FF5+UMD	0.1981	0.2021	0.1894
FFCPM+UMD	0.0465	0.0465	0.0681	FFCPM+UMD	0.1967	0.2006	0.1907
FF5+UMD<(FF5/FFCPM+UMD/SY/Carhart FF3+LIQ/FF3/HXZCP/HXZ/BAB/CAPM)				FFCPM+UMD<(FF5/FF5+UMD/HXZCP/SY HXZ/FF3+LIQ/FF3/Carhart/BAB/CAPM)			
H_0 (p-value)	0.396	0.365	0.233	H_0 (p-value)	0.212	0.267	0.313

Table 4: Mimicking Non-traded Factors and HJ Distance in US

Notes: the table represents the results of the non-tradable and tradable SDFs comparison in US using both gross returns and excess returns on the pooled 230 risk portfolios. Non-tradable mimicking portfolio are regressed by all the traded-factor returns by using adjusted R^2 and the F test. Panel A reports the rank restriction test (w) and its p-value of the null that $E[x_t(1, fit)]$ has a column rank of one. Panel B reports the HJ distance (HJ^O), the modified HJ distance (HJ^M) and the constrained HJ distance (HJ^C) measures among tradable and non-tradable factors SDFs. Panel C and D present results of the two formal tests of model comparison. The null hypothesis in the first test is that the chosen model has a HJ^O , HJ^M and HJ^C measure that is less than that of any of the alternative models. The second test is a pairwise comparison test of whether HJ^O , HJ^M and HJ^C measures of each of the alternative models are significantly greater than the distance measures of the chosen model.

Panel A: Rank test for individual factors						
Test	CG	DCG	HCG	EXR	LIQINNOV	
w	94.2	100.4	84.5	91.6	86.1	
p-val	0.001	0.000	0.006	0.001	0.005	
Panel B: HJ Measures						
	Gross Returns			Excess Returns		
	HJ^O	HJ^M	HJ^C	HJ^O	HJ^M	HJ^C
CAPM	0.8841	0.8841	0.9631	0.6947	0.9655	0.6234
FF3	0.8736	0.8736	0.9535	0.6919	0.9553	0.6188
FF5	0.8388	0.8388	0.925	0.6892	0.9117	0.6102
Carhart	0.8541	0.8541	0.9361	0.6903	0.9321	0.6172
HXZ	0.8672	0.8672	0.9482	0.7071	0.9756	0.6162
HXZCP	0.8467	0.8467	0.9344	0.6966	0.9326	0.6135
FF3+LIQ	0.8732	0.8732	0.9534	0.6935	0.9551	0.6182
BAB	0.8821	0.8821	0.9609	0.6994	0.9699	0.6202
SY	0.8417	0.8417	0.9328	0.6854	0.9154	0.6137
FF5+UMD	0.8276	0.8276	0.914	0.6796	0.8995	0.609
FFCPM+UMD	0.828	0.828	0.9147	0.682	0.9	0.6081
FF5CG+DCG	0.818	0.818	0.9045	0.6778	0.89	0.6079
FF5+HCG	0.8178	0.8178	0.9041	0.68	0.895	0.6098
FF5+LIQINO	0.8278	0.8278	0.9143	0.6759	0.899	0.619
Panel C: HJ Multiple Comparison						
FF5+UMD < (DCG/FFCPM + UMD)/HCG/ LIQINO/FF5/SY/HXZCP/Carhart HXZ/FF3 + LIQ/FF3/BAB/CAPM)	FF5+UMD < (FFCPM + UMD)/DCG/HCG/ LIQINO/FF5/SY/Carhart/HXZCP FF3/FF3 + LIQ/BAB/CAPM/HXZ)			H_0 (p-value)		
H_0 (p-value)	0.031	0.02	0.030	0.048	0.018	0.0404
FF5+UMD <	FF5+UMD <			H_0 (p-value)		
FF5 (p-value)	0.054	0.054	0.23	0.099	0.067	0.070
FF5CP+UMD (p-value)	0.127	0.127	0.138	0.138	0.085	0.092
SY (p-value)	0.131	0.131	0.234	0.132	0.98	0.18
HXZCP (p-value)	0.153	0.153	0.182	0.096	0.123	0.212
Carhart (p-value)	0.124	0.124	0.212	0.109	0.14	0.162
HXZ (p-value)	0.138	0.138	0.124	0.145	0.157	0.181
FF3+LIQ (p-value)	0.131	0.131	0.236	0.137	0.164	0.1464
FF3 (p-value)	0.100	0.104	0.184	0.103	0.173	0.133
BAB (p-value)	0.131	0.131	0.222	0.17	0.17	0.16
CAPM (p-value)	0.1532	0.1532	0.2272	0.198	0.198	0.139
FF5CG+DCG (p-value)	0.008	0.008	0.014	0.01	0.012	0.017
FF5+HCG (p-value)	0.014	0.013	0.011	0.014	0.017	0.015
FF5+LIQINO (p-value)	0.026	0.021	0.066	0.022	0.023	0.029

Table 5: Predictive regression of nontradable factors in US

Notes: the table presents the predictive regression estimation results of the nontradable factors in US. Specifically, for each factor, we run univariate predictive regression of n-month cumulative excess returns of test assets on the cumulative risk premium of the mimicking factor portfolio of the inflation factor over the same periods. The test assets include 230 pooled portfolios: 25 portfolios sorted by size and book-to-market ratio, 30 industry portfolios, 100 portfolios sorted by operating profitability and investment, 25 portfolios sorted by size and variance, 25 portfolios sorted by size and momentum, and 25 portfolios sorted by size and beta. We consider prediction periods ranging from 1 quarter to 8 quarters. For a given prediction period, we first run univariate predictive regression for each portfolio, and report the average adjusted R-squared (%) across different portfolios in the last column. The analysis is based on quarterly data from the first quarter of 1972 to the third quarter of 2011.

Quarter	DCG	HCG
1	0.57	0.40
2	-0.30	0.18
3	-0.52	0.13
4	-0.50	-0.01
5	-0.34	0.89
6	-0.40	0.14
7	-0.52	0.57
8	-0.52	1.32

Table 6: Rank Test and Misspecification Identification in UK

Notes: the table presents the results of the ranks tests of the individual factors, the model misspecification tests, and rank tests of the models for the UK stock market. The models are estimated using monthly returns from 1990 to 2017 on the 122 pooled portfolios. Panel A reports the rank restriction test (w) and its p-value of the null that $E[x_t(1, fit)]$ has a column rank of one. Panel B reports the sample HJ distance ($\hat{\delta}$), the Lagrange multiplier (LM) test, and the rank restriction test (W^*) with the corresponding p-values for each model.

Panel A: Rank test for individual factors

Test	MKT	SMB	HML	CMA	RMW	RMWC	RMW_ROE	UMD	HMLM	MGMT	PERF
w	141.0	169.6	173.8	138.1	159.1	161.8	165.1	155.4	172.6	153.3	158.3
p-val	0.103	0.002	0.001	0.137	0.012	0.008	0.005	0.019	0.001	0.025	0.013

Panel B: HJ-distance, Lagrange multiplier, and rank tests

Model	delta	p-val	LM	p-val	w	p-val
CAPM	0.829	0.000	135.640	0.156	141.0	0.203
FF3	0.826	0.000	135.020	0.135	142.3	0.072
FF5	0.826	0.000	135.160	0.108	117.2	0.476
Carhart	0.817	0.000	129.590	0.201	119.4	0.447
HXZ-FF	0.827	0.000	135.900	0.112	142.0	0.066
FF5CP	0.825	0.000	134.600	0.114	119.6	0.417
SY-FF	0.827	0.000	134.840	0.124	118.0	0.482
FF5+UMD	0.816	0.000	128.980	0.176	115.5	0.495
FF5CP+UMD	0.816	0.000	129.200	0.173	117.7	0.440
FF5CPM+UMD	0.818	0.000	130.090	0.159	117.6	0.440

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMB	HML	CMA	RMW	RMWC	RMW_ROE	MOM	HMLM	MGMT	PERF
CAPM	0.57										
FF3	0.92	-1.19	-1.00								
		-1.06	-0.73								
		-0.96									
FF5	0.68	-1.11	-0.62	0.07		-0.20					
	0.69	-1.19	-0.77			-0.19					
	0.92	-1.19	-1.00								
		-1.06	-0.73								
		-0.96									
Carhart	0.23	-1.48	-1.96					-1.93			
		-1.46	-1.90					-2.08			
			-1.51					-1.79			
								-1.29			
HXZ-FF	0.39	-1.02		-0.56			0.12				
	0.38	-1.07		-0.64							
		-1.05		-0.84							
		-0.96									
SY-FF	0.40	-1.09								-0.53	-0.16
	0.56	-1.08								-0.51	
	0.66	-1.02									
		-0.96									

Table continued.

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMB	HML	CMA	RMW	RMWC	RMW_ROE	MOM	HMLM	MGMT	PERF
FF5+UMD	0.47	-1.29	-1.31	0.26	0.55			-1.97			
	0.39	-1.28	-1.51		0.59			-1.94			
		-1.28	-1.50		0.50			-2.02			
		-1.46	-1.90					-2.08			
			-1.51					-1.79			
FF5CP+UMD	0.34	-1.48	-1.67	0.33		-0.33		-1.86			
	0.36	-1.49	-1.69	0.33				-1.96			
	0.23	-1.48	-1.96					-1.93			
		-1.46	-1.90					-2.08			
			-1.51					-1.79			
FF5CPM+UMD	0.28	-1.41		0.16		-0.29		-1.92	-1.38		
	0.23	-1.41				-0.29		-1.98	-1.54		
		-1.38				-0.30		-2.07	-1.50		
		-1.39						-2.06	-1.61		
								-1.76	-1.29		
							-1.29				

Panel D: Model selection procedure using standard errors under model misspecification

Model	MKT	SMB	HML	CMA	RMW	RMWC	RMW_ROE	MOM	HMLM	MGMT	PERF
CAPM	0.47										
FF3	0.74	-1.15	-0.92								
		-1.04	-0.69								
		-0.94									
FF5	0.51	-1.07	-0.53	0.06			-0.16				
	0.51	-1.15	-0.66				-0.15				
	0.74	-1.15	-0.92								
		-1.04	-0.69								
		-0.94									
Carhart	0.19	-1.42	-1.67					-1.67			
		-1.40	-1.65					-1.78			
			-1.34					-1.57			
								-1.19			
HXZ-FF	0.29	-1.00		-0.44			0.10				
	0.30	-1.05		-0.52							
		-1.03		-0.72							
		-0.94									
SY-FF	0.30	-1.07								-0.47	-0.13
	0.46	-1.06								-0.45	
	0.55	-1.00									
		-0.94									

Table continued.

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMB	HML	CMA	RMW	RMWC	RMW_ROE	MOM	HMLM	MGMT	PERF
FF5+UMD	0.35	-1.25	-1.11	0.21	0.47			-1.69			
	0.31	-1.24	-1.32		0.50			-1.67			
		-1.23	-1.31		0.43			-1.73			
		-1.40	-1.65					-1.78			
			-1.34					-1.57			
FF5CP+UMD	0.25	-1.41	-1.38	0.26		-0.26		-1.61			
	0.27	-1.43	-1.40	0.26				-1.70			
	0.19	-1.42	-1.67					-1.67			
		-1.40	-1.65					-1.78			
			-1.34					-1.57			
FF5CPM+UMD	0.21	-1.34		0.13		-0.24		-1.66	-1.16		
	0.19	-1.34				-0.23		-1.71	-1.29		
		-1.32				-0.24		-1.77	-1.26		
		-1.34						-1.76	-1.43		
								-1.54	-1.16		
							-1.19				

Table 7: HJ Distance and Large Cross-sectional Portfolios in UK

Notes: the table represents the results of tradable factors SDF comparison tests using both gross returns and excess returns on the 122 pooled portfolios in UK, including 25 portfolios sorted by size and book-to-market ratio, 25 portfolios sorted by size and momentum, 10 portfolios sorted by size, 10 portfolios sorted by book-to-market ratio, 10 portfolios sorted by momentum, 25 portfolios sorted by standard deviation, and 27 portfolios sorted by size, book-to-market ratio, and momentum. Panel A reports the HJ distance (HJ^O), the modified HJ distance (HJ^M) and the constrained HJ distance (HJ^C) measures. Panel B presents the results of the two formal tests of multiple model comparison. The null hypothesis in the first test is that the chosen model has a HJ^O , HJ^M and HJ^C measure that is less than that of any of the alternative models. The dataset ranges from January 1990 to December 2017.

Pooled Portfolios	Gross Returns				Excess Returns		
	HJ^O	HJ^M	HJ^C		HJ^O	HJ^M	HJ^C
Panel A: HJ Measures							
CAPM	0.8293	0.8293	0.8949	CAPM	0.6846	0.9388	0.6572
FF3	0.8257	0.8257	0.8913	FF3	0.6846	0.9388	0.6521
FF5	0.8256	0.8256	0.8912	FF5	0.6891	0.9443	0.6497
Carhart	0.8169	0.8169	0.8837	Carhart	0.6872	0.9347	0.6507
HXZ-FF	0.8266	0.8266	0.8926	HXZ-FF	0.6896	0.9432	0.6512
FF5CP	0.8246	0.8246	0.8906	FF5CP	0.6929	0.9512	0.6487
SY-FF	0.8267	0.8267	0.8911	SY-FF	0.6913	0.9476	0.6540
FF5+UMD	0.8161	0.8161	0.8825	FF5+UMD	0.6870	0.9316	0.6469
FF5CP+UMD	0.8165	0.8165	0.8830	FF5CP+UMD	0.6897	0.9361	0.6466
FF5CPM+UMD	0.8182	0.8182	0.8837	FF5CPM+UMD	0.6906	0.9382	0.6450
Panel B: HJ Multiple Comparison							
FF5+UMD < (FF5CP+UMD/FF5CPM+UMD)/FF5/SY-FF			FF5+UMD < (FF5CP+UMD/FF5CPM+UMD)/FF5/SY-FF				
Carhart/HXZ - FF/FF3/CAPM			Carhart/HXZ - FF/FF3/CAPM				
H_0 (p-value)	0.257	0.278		H_0 (p-value)	0.267	0.208	

Table 8: Mimicking Non-traded Factors and HJ Distance in UK

Notes: the table presents the rank test results of the non-tradable factors on the 95 pooled portfolios in UK, including 25 portfolios sorted by size and book-to-market ratio, 25 portfolios sorted by size and momentum, 10 portfolios sorted by size, 10 portfolios sorted by book-to-market ratio, 10 portfolios sorted by momentum, and 25 portfolios sorted by standard deviation. Non-tradable mimicking portfolio are regressed by all the traded-factor returns by using adjusted R^2 and the F test. Panel A reports the rank restriction test (w) and its p-value of the null that $E[x_t(1, fit)]$ has a column rank of one. We consider 6 nontradable factors in total. Yield is the yield level factor; MS is the maturity spread; CPI and PPI are year-over-year growth of the consumer price index and producer price index, respectively, and INF is inflation. For the macro factors, as monthly data are available, we estimate the factor risk premiums and conduct the rank test using monthly data from January 1998 to December 2017.

Panel A: Rank test for individual factors						
Test	Yield	MS	CPI	PPI	INF	
w	99.4	103.1	102.7	95.7	98.2	
p-val	0.332	0.244	0.253	0.433	0.362	

Panel B: HJ Measures							
	Gross Returns				Excess Returns		
	HJ^O	HJ^M	HJ^C		HJ^O	HJ^M	HJ^C
CAPM	0.8889	0.8889	0.9749	CAPM	0.7160	1.0256	0.7681
FF3	0.8815	0.8815	0.9708	FF3	0.7167	1.0274	0.7621
FF5	0.8773	0.8773	0.9674	FF5	0.7308	1.0621	0.7600
Carhart	0.8809	0.8809	0.9703	Carhart	0.7244	1.0439	0.7615
HXZ-FF	0.8841	0.8841	0.9716	HXZ-FF	0.7318	1.0611	0.7633
FF5CP	0.8773	0.8773	0.9683	FF5CP	0.7308	1.0633	0.0.7595
SY-FF	0.8844	0.8844	0.9725	SY-FF	0.7251	1.0458	0.7617
FF5+UMD	0.8754	0.8754	0.9667	FF5+UMD	0.7313	1.0597	0.7597
FF5CP+UMD	0.8756	0.8756	0.9673	FF5CP+UMD	0.7317	1.0612	0.7592
FF5CPM+UMD	0.8737	0.8737	0.9653	FF5CPM+UMD	0.7323	1.0622	0.7557
FF5+Yield	0.8772	0.8772	0.9669	FF5+Yield	0.7315	1.0640	0.7594
FF5+INF	0.8728	0.8728	0.9611	FF5+INF	0.7312	1.0519	0.7592

Panel C: HJ Multiple Comparison						
FF5+INF<(FF5+Yield/FF5+UMD/FF5/ FF5CPM+UMD/FF5CP+UMD/FF3/ Carhart/HXZ-FF/SY-FF/CAPM)				FF5+INF<(FF5+Yield/FF5+UMD/FF5/ FF5CPM+UMD/FF5CP+UMD/FF3/ Carhart/HXZ-FF/SY-FF/CAPM)		
H_0 (p-value)	0.2880	0.2810		H_0 (p-value)	0.2170	0.1620

Table 9: Predictive regression of inflation factor in UK

Notes: the table presents the predictive regression estimation results of the yield level (Yield) and inflation (INF) factors in UK. Specifically, for each factor, we run univariate predictive regression of n-month cumulative excess returns of test assets on the cumulative risk premium of the mimicking factor portfolio of the inflation factor over the same periods. The test assets include the 95 pooled portfolios in UK, including 25 portfolios sorted by size and book-to-market ratio, 25 portfolios sorted by size and momentum, 10 portfolios sorted by size, 10 portfolios sorted by book-to-market ratio, 10 portfolios sorted by momentum, and 25 portfolios sorted by standard deviation. We consider prediction periods ranging from 1 month to 12 months. For a given prediction period, we first run univariate predictive regression for each portfolio, and report the average adjusted R-squared (%) across different portfolios in the last column. The analysis is based on monthly data from January 1998 to December 2017.

Month	Yield	INF
1	7.50	12.31
2	4.77	6.82
3	4.22	5.20
4	4.88	6.02
5	4.44	5.07
6	4.24	4.42
7	5.22	3.26
8	3.29	1.86
9	3.16	1.53
10	3.04	1.38
11	2.51	0.94
12	1.72	0.72

Table 10: Rank Test and Misspecification Identification in China

Notes: the table presents the results of the ranks tests of the individual factors, the model misspecification tests, and rank tests of the models for the China's stock market. The models are estimated using monthly returns from 2000 to 2018 on the 83 pooled portfolios. Panel A reports the rank restriction test (w) and its p-value of the null that $E[x_t(1, fit)]$ has a column rank of one. Panel B reports the sample HJ distance ($\hat{\delta}$), the Lagrange multiplier (LM) test, and the rank restriction test (W^*) with the corresponding p-values for each model. Panels C and D show t-tests of the model selection procedures based on the standard errors under correct model specification and model misspecification, respectively. The boldface denotes the 5% significance level.

Panel A: Rank test for individual factors																
Test	MKT	SMB	HML	SMBSY	CMA	RMW	RMWC	ME	IA	ROE	MOM	UMD	HMLM	BAB	MGMT	PERF
w	111.9	116.3	111.2	116.6	123.1	116.9	104.9	121.3	111.8	109.3	113.2	115.4	118.8	101.3	106.6	124.2
p-val	0.016	0.008	0.018	0.007	0.002	0.007	0.045	0.003	0.016	0.024	0.013	0.009	0.005	0.073	0.036	0.002

Panel B: HJ-distance, Lagrange multiplier, and rank tests						
Model	delta	p-val	LM	p-val	w	p-val
CAPM	1.038	0.000	119.700	0.003	111.9	0.016
FF3	1.029	0.000	118.750	0.003	101.8	0.051
BAB	1.032	0.000	123.290	0.001	95.3	0.132
Carhart	1.014	0.000	119.090	0.002	86.1	0.274
HXZ	1.013	0.000	129.420	0.000	97.5	0.077
HXZCP	1.011	0.000	125.200	0.001	107.0	0.020
SY	1.024	0.000	127.760	0.000	89.4	0.198
FF5	0.952	0.000	109.950	0.008	82.3	0.347
FF5+UMD	0.954	0.000	107.170	0.013	92.7	0.122
FF5CP+UMD	0.951	0.000	107.020	0.011	102.3	0.029

FF5CPM+UMD 0.957 0.000 113.770 0.003 92.9 0.105

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	BAB	MGMT	PERF
CAPM	-1.88														
FF3	-1.09	-1.40	-1.82												
		-1.74	-2.13												
			-1.73												
BAB	-1.91												-1.78		
	-1.88														
FF5	-0.90	-4.50	-4.51		1.21	-2.75									
		-4.48	-4.99		1.64	-2.40									
		-4.84	-4.73			-4.69									
Carhart	-1.18	-2.04	-2.25												
		-2.35	-2.51												
			-1.82												
			-1.73												
HXZ	-2.49						-2.28	0.25	-2.69						
	-2.51						-2.53		-2.99						
HXZCP	-2.66					-2.60	-2.45	0.46							
	-2.74					-3.32	-2.65								
SY	-2.41			-1.55										-0.77	-2.53
	-2.24			-1.35											-2.37
	-2.37														-2.10

Table continued.

Panel C: Model selection procedure using standard errors under correct model specification

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	BAB	MGMT	PERF
	-1.88														
FF5+UMD	-0.77	-4.19	-4.48		1.41	-1.91									-1.05
		-4.15	-4.91		1.80	-1.60									-1.14
		-4.48	-4.99		1.64	-2.40									
		-4.84	-4.73			-4.69									
FF5CP+UMD	-0.76	-3.98	-4.22		3.28	-1.76									-1.07
		-3.98	-4.64		3.64	-1.56									-1.19
		-4.25	-4.54		3.77	-2.41									
		-3.99	-4.30		4.55										
FF5CPM+UMD	-1.21	-4.21			3.08	-2.35									-2.45
		-4.18			3.47	-2.02									-2.77
		-3.96			3.77										-3.68
															-4.64

Panel D: Model selection procedure using model misspecification-robust standard errors

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	BAB	MGMT	PERF
CAPM	-1.61														
FF3	-0.95	-1.36	-1.84												
		-1.74	-2.13												
			-1.73												

BAB	-1.62							-1.36
	-1.61							
FF5	-0.80	-4.39	-4.23	0.97	-2.44			
		-4.36	-4.73	1.34	-2.18			
		-4.74	-4.67		-4.57			
Carhart	-1.04	-1.89	-2.15					-1.95
		-2.18	-2.35					-1.83
			-1.76					-1.09
			-1.73					
HXZ	-1.94				-2.02	0.21	-2.19	
	-1.97				-2.41		-2.59	
HXZCP	-2.17				-2.12	-2.17	0.41	
	-2.24				-2.82	-2.45		
SY	-1.96		-1.48					-0.61
	-1.84		-1.34					-2.23
	-1.99							-1.94
	-1.61							
FF5+UMD	-0.67	-4.10	-4.12	1.13	-1.69			-0.88
		-4.06	-4.61	1.49	-1.47			-0.97
		-4.36	-4.73	1.34	-2.18			
		-4.74	-4.67		-4.57			
FF5CP+UMD	-0.65	-3.66	-3.95	2.85	-1.50			-0.92
		-3.67	-4.35	3.14	-1.37			-1.03
		-3.91	-4.24	3.21	-2.08			

Table continued.

Panel D: Model selection procedure using model misspecification-robust standard errors

Model	MKT	SMBC	HML	SMBSY	CMA	RMWC	ME	IA	ROE	MOM	UMD	HMLM	BAB	MGMT	PERF
		-3.86	-3.96		3.87										
FF5CPM+UMD	-1.01	-3.76			2.70	-1.95					-2.07	-3.97			
		-3.76			2.99	-1.73					-2.35	-4.30			
		-3.80			3.24						-3.08	-4.21			

Table 11: HJ Distance and Large Cross-sectional Portfolios in China

Notes: the table represents the results of tradable factors SDF comparison tests using both gross returns and excess returns on the 83 pooled portfolios in China, including 25 portfolios sorted by size and book-to-market ratio, 10 portfolios sorted by size, 10 portfolios sorted by book-to-market ratio, 10 portfolios sorted by operating profitability, and 28 industry portfolios. Panel A reports the HJ distance (HJ^O), the modified HJ distance (HJ^M) and the constrained HJ distance (HJ^C) measures. Panel B presents the results of the two formal tests of multiple model comparison. The null hypothesis in the first test is that the chosen model has a HJ^O , HJ^M and HJ^C measure that is less than that of any of the alternative models.

Pooled Portfolios	Gross Returns			Excess Returns			
	HJ^O	HJ^M	HJ^C	HJ^O	HJ^M	HJ^C	
Panel A: HJ Measures							
CAPM	1.0082	1.0082	1.1822	CAPM	0.7111	1.0115	0.8412
FF3	0.9970	0.9970	1.1723	FF3	0.7098	1.0067	0.8340
FF5	0.9305	0.9305	1.1218	FF5	0.6911	0.9444	0.8149
Carhart	0.9857	0.9857	1.1640	Carhart	0.7072	0.9978	0.8328
HXZ	0.9905	0.9905	1.1669	HXZ	0.7333	1.0475	0.8389
HXZCP	0.9865	0.9865	1.1700	HXZCP	0.7227	1.0264	0.8375
BAB	0.9994	0.9994	1.1795	BAB	0.7136	1.0169	0.8411
SY	0.9958	0.9958	1.1703	SY	0.7171	1.0235	0.8379
FF5+UMD	0.9263	0.9263	1.1203	FF5+UMD	0.6901	0.9422	0.8148
FF5CP+UMD	0.9279	0.9279	1.1318	FF5CP+UMD	0.6897	0.9425	0.8176
FF5CPM+UMD	0.9366	0.9366	1.1396	FF5CPM+UMD	0.7001	0.9575	0.8253
Panel B: HJ Multiple Comparison							
FF5+UMD<(FF5CP+UMD/FF5CPM+UMD/FF5/SY				FF5+UMD<(FF5CP+UMD/FF5CPM+UMD/FF5/SY			
Carhart/HXZ/HXZCP/FF3/BAB/CAPM)				Carhart/HXZ/HXZCP/FF3/BAB/CAPM)			
H_0 (p-value)	0.257	0.240		H_0 (p-value)	0.321	0.282	

Table 12: Mimicking Non-traded Factors and HJ Distance in China

Notes: the table presents the rank test results of the non-tradable factors on the pooled 53 risk portfolios in China, including 25 portfolios sorted by size and book-to-market ratio, and 28 industry portfolios. Non-tradable mimicking portfolio are regressed by all the traded-factor returns by using adjusted R^2 and the F test. Panel A reports the rank restriction test (w) and its p-value of the null that $E[x_t(1, fit)]$ has a column rank of one. We consider 6 nontradable factors in total. CG represents the quarter-over-quarter consumption growth, and WG is the year-over-year growth of the quarterly total wage of urban populations, scaled by the quarterly GDP (gross domestic production) of the same quarter. Yield is the yield level factor; MS and CS are the maturity spread and credit spread, respectively, while INF is inflation. The details of the computation of these macro factors are given in Appendix A.6. For CG and WG, as there are only quarterly data available, we estimate the factor risk premiums and conduct the rank test based on quarterly data from 2003 to 2018. For the other four macro factors, as monthly data are available, we estimate the factor risk premiums and conduct the rank test using monthly data from March 2006 to December 2018.

Panel A: Rank test for individual factors						
Test	CG	WG	Yield	MS	CS	INF
w	51.8	49.5	71.8	68.6	68.6	71.5
p-val	0.483	0.572	0.036	0.062	0.059	0.038

Panel B: HJ Measures							
	Gross Returns			Excess Returns			
	HJ^O	HJ^M	HJ^C	HJ^O	HJ^M	HJ^C	
CAPM	0.8788	0.8788	0.9946	CAPM	0.6646	0.8894	0.4601
FF3	0.8693	0.8693	0.9882	FF3	0.6634	0.8861	0.4562
FF5	0.7919	0.7919	0.9331	FF5	0.6271	0.8010	0.4465
Carhart	0.8500	0.8500	0.9683	Carhart	0.6537	0.8634	0.4500
HXZ	0.8518	0.8518	0.9793	HXZ	0.6704	0.8822	0.4579
HXZCP	0.8629	0.8629	0.9892	HXZCP	0.6708	0.8956	0.4582
BAB	0.8719	0.8719	0.9883	BAB	0.6671	0.8922	0.4597
SY	0.8553	0.8553	0.9811	SY	0.6580	0.8711	0.4594
FF5+UMD	0.7870	0.7870	0.9280	FF5+UMD	0.6253	0.7976	0.4451
FF5CP+UMD	0.8352	0.8352	0.9579	FF5CP+UMD	0.6543	0.8612	0.4479
FF5CPM+UMD	0.8243	0.8243	0.9538	FF5CPM+UMD	0.6569	0.8576	0.4468
FF5+Yield	0.7758	0.7758	0.9259	FF5+Yield	0.6700	0.8966	0.4459
FF5+INF	0.7878	0.7878	0.9315	FF5+INF	0.7047	0.9252	0.4455

Panel C: HJ Multiple Comparison			
	H_0 (p-value)		H_0 (p-value)
FF5+Yield<(FF5+INF/FF+5UMD/FF5CP+UMD/FF5CPM+UMD)/ FF5/SY/Carhart/HXZ/HXZCP/FF3/BAB/CAPM)	0.3060	0.2920	0.2750
FF5+Yield<(FF5+INF/FF5+UMD/FF5CP+UMD/FF5CPM+UMD)/ FF5/SY/Carhart/HXZ/HXZCP/FF3/BAB/CAPM)			0.3430

Table 13: Predictive regression of nontradable factors in China

Notes: the table presents the predictive regression estimation results of the yield level (Yield) and inflation (INF) factors in China. Specifically, for each factor, we run univariate predictive regression of n-month cumulative excess returns of test assets on the cumulative risk premium of the mimicking factor portfolio of the inflation factor over the same periods. The test assets include the 53 risk portfolios in China, i.e., 25 portfolios sorted by size and book-to-market ratio, and 28 industry portfolios. We consider prediction periods ranging from 1 month to 12 months. For a given prediction period, we first run univariate predictive regression for each portfolio, and report the average adjusted R-squared (%) across different portfolios in the last column. The analysis is based on monthly data from March 2006 to December 2018.

Month	Yield	INF
1	29.17	35.05
2	16.01	29.48
3	16.45	23.73
4	11.42	18.80
5	12.94	15.75
6	9.74	15.04
7	9.87	11.26
8	11.47	11.68
9	9.73	10.00
10	7.92	8.33
11	7.78	7.43
12	9.67	7.01