# Proliferation of Anomalies and Zoo of Factors – What does the Hansen–Jagannathan Distance Tell Us ? *

Xiang Zhang†

January 2, 2018

**Abstract**

Recent research finds that prominent asset pricing models have mixed success in evaluating the cross-section of anomalies, which highlights proliferation of anomalies and zoo of factors. In this paper, I investigate that how is the relative pricing performance of these models to explain anomalies, when comparing their misspecification errors – the Hansen–Jagannathan (HJ) distance measure. I find that a single model dominates others in a specific anomaly by incorporating the multiple HJ distance comparing inference; the model relative pricing performance is sensitive to the test portfolios choice. However, different from the current research, I result that the HJ distance is a general statistic measure to compare models and it can be divided into the 'relative comparison' (sensitive to test portfolios) and the 'absolute comparison' (independent to test portfolios) parts. Second, there is a large variation in the shape and curvature of these confidence sets of anomalies, which makes any single SDF difficult to satisfy confidence sets of anomalies all. My results imply that further work is required not only in pruning the number of priced factors but also in building models that explain the data better.

*JEL Classification:* G1, G12

*Keywords:* asset pricing models, Hansen–Jagannathan distance, misspecification errors

---

*I would like to thank Raymond Kan, Cesare Robotti, Abhay Abhyankar, Michael Creel, Rajesh Tharyan, Motohiro Yogo and Guofu Zhou for their comments and suggestions. I am especially grateful to Raymond Kan and Cesare Robotti for their helpful comments. I also gratefully acknowledge financial support from the Spanish Ministry of Science and Innovation through grant ECO2008-04756 (Grupo Consolidado-C) and FEDER.

†School of Finance, Southwestern University of Finance and Economics, 611130, Chengdu, Sichuan, China. Email: xiangzhang@swufe.edu.cn

# 1 Introduction and Motivation

Over the last few decades empirical studies report that over three hundred factors that claim to explain anomalies and this number is growing. The leading papers by Harvey et al. (2016) and Hou et al. (2017) have studied reasons why different factor models have mixed success in explaining the cross-sectional anomalies. Harvey et al. (2016) reports that there are more than 300 factors that seem to price the cross-section of expected returns. They result that more stringent standards than conventional levels of significance need to be used in statistical tests in order to focus on a set of factors. The estimation of their model suggests that today a newly discovered factor needs to clear a much higher hurdle, with a t-ratio greater than 3.0. Hou et al. (2017) replicates 447 anomalies in the literature. They find that most these anomalies are insignificant with microaps alleviating via NYSE breakpoints and value-weighted returns. In this paper, I investigate the above research question in a different perspectives – why there is no dominated factor models in explaining the cross-sectional anomalies by evaluating their misspecification errors after Hansen and Jagannathan (1997) propose Hansen – Jagannathan (HJ) distance – a misspecification measure of models.

In early work, researchers use the HJ distance plus the pairwise HJ distance comparison inference to indicate that their new factor models perform better than the benchmarks to explain the cross-section of equity anomalies. Jagannathan and Wang (1996) result that their labor income conditional capital asset pricing model (CAPM) outperforms other linear factor models to explain returns on Fama and French 100 size-beta portfolios based on the HJ distance misspecification measure. Lettau and Ludvigson (2001) show that their consumption-wealth ratio conditional consumption-based CAPM (CCAPM) obtains the smallest HJ distance compared to the CCAPM and the Fama–French three-factor model to explain returns on the Fama–French 10 deciles portfolios and the 6 size-book/market ratio portfolios. Vassalou and Xing (2004) find that their Fama–French model augmented by the default risk may be less misspecified than the Fama–French model to explain the 27 size, book/market ratio and default likelihood indicator portfolios. Moreover, Parker and Julliard (2005) show that their ultimate consumption risk conditional CCAPM has the smaller HJ distance than the Lettau and Ludvigson (2001) conditional model and the contemporaneous CCAPM to explain the Fama–French 25 size and book/market ratio portfolios.

In this paper, I aim to contribute to the existing literature by providing reasons that why prominent theoretically and empirically motivated factor models explain cross-sectional anomalies differently when comparing their HJ distance measures. I use two recent developments, one is the formal tests of multiple model comparison and the other is the set inference methods. My work is closely related papers, such as Hodrick and Zhang (2001) evaluate the specification errors of several asset pricing models that are rivals to the standard CAPM. Other related work includes Wang (2005) and Zhang (2006) who evaluate the cross-sectional pricing performances of several asset pricing models using the HJ distance metric. Finally, Li et al. (2010) study a sequence of model selection procedures for non-nested and nested models based on the second HJ distance. However, these papers rely on comparisons of point estimates of the HJ distance for factor models, which does not account for sampling errors and model misspecification uncertainty.

I employ the methodologies of Chen and Ludvigson (2009) and Gospodinov et al. (2013), which allow for multiple model selection tests with improved finite-sample properties. 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 not only the pairwise but the multiple model comparison inference based on the sample HJ distance measures, and show new pivotal specification and model comparison tests that are asymptotically chi-squared distributed. To my limited knowledge, this paper is the first one that applies the Gospodinov, Kan and Robotti (2013) multiple model comparison and incorporates the appropriate null hypotheses to re-evaluate classical linear factor asset pricing models on a common set of test portfolios.

Since most earlier papers show no economic value of comparison asset pricing models, this paper uses set inference methods proposed by Chernozhukov, Kocatulum and Menzel (Chernozhukov et al. (2012)) that allow me to construct confidence regions for the HJ sets of admissible stochastic discount factors (SDFs) arising from my candidate models. These set inference methods allow me to further explore the economic reasons that underlie statistical rejections of candidate models. Different from Chernozhukov, Kocatulum and Menzel (2012) who investigate various set inference methods for achieving tight confidence regions and more powerful inference procedures, I apply their set inference methods to analyze interactions between confidence sets of the HJ set and asset pricing models. Specifically, I empirically check which families of SDFs price test portfolios correctly and which do not by seeing whether the mean and standard deviation of SDFs are admissible.

I find, in this paper, that different from the problem of a “zoo of factors”, a set of prominent asset pricing models that are both theoretically and empirically motivated perform differently in pricing basic asset classes like stock portfolios; the relative pricing performance of linear factor models is sensitive to the test portfolios choice, which "highlight the challenges of evaluating empirical factor models" (Kogan and Tian (2013)). However, comparing five empirical factor models and using the same set of assets, the Carhart four-factor model has a better performance than alternative models like the Fama–French factor models (Fama and French (1992), Fama and French (2015)) and the Hou et al. (2015) investment-based model. Among consumption-based models, Yogo (2006) non-durable and durable CCAPM does better than the Piazzesi et al. (2007) housing CCAPM and the classical CCAPM model when pricing the Fama–French size/book-to-market ratio and the industry sorted portfolios. In comparisons of conditional asset pricing models, I find that the Lettau and Ludvigson (2001) model obtain smaller HJ distances than the other two candidates when the test assets are the size/book-to-market and industry portfolios.

This paper has two contributions as follows:

Firstly, I show that the HJ distance is a general measure, which can be divided into the 'relative comparison' (sensitive to test portfolios) and the 'absolute comparison' (independent to test portfolios) parts. The 'relative comparison' is defined as the comparison among factor models to explain the cross-sectional anomalies; the 'absolute comparison' is defined as the comparison among factor models. The 'relative comparison' or the 'absolute comparison' is like to compare heights between two people for playing basketball or to compare two people's heights. I evidence that test assets influence the comparison results if linear asset pricing models are compared by explaining several cross-sectional anomalies. Specifically, I show that the modified HJ distance of Kan and Robotti (2008) is the general form of the Gibbons, Ross and Shanken test (GRS, Gibbons et al. (1989)), and the modified HJ distance can be divided into two parts: the one is related to the factor structure GRS statistics and the other is related to the test assets GRS statistics. Hence, I can interpret my empirical results into two perspectives: first, misspecification errors comparison results depend on the choice of test assets; second, the GRS test can tell which mimicking portfolios that spanning candidate models are the ex post minimum-variance frontiers, only if the model comparison scenario focuses on that which model is able to price the factors in the other model (Barillas and Shanken (2016a)).

The difference between the 'relative comparison' and the 'absolute comparison' comes from the difference between the GRS test and the HJ distance test. The GRS test basically evaluates the improvement chances of factors portfolios to the efficient mean-variance frontier. Hence, their null hypothesis states that the candidate factor-models perfectly match the mean-variance frontier of test assets. However, the Barillas and Shanken (2016a) shows that we do not need to match the mean-variance frontier of test assets for every factors portfolios, but we can evaluate the relative 'match' of each two pair of candidate factors portfolios. If we only compare how 'match' of each two pair of candidate portfolios, then test assets will not participate into the comparison; the choice of any test assets will not influence results of any two factor-models. The HJ distance measure tries to find the minimum second moment portfolios of each candidate factor models. Specifically, the HJ distance measure is equal to the second moment of the difference 'factor structure' GRS test statistics and the 'net effect' of the test assets on the GRS test statistics. Better than these two GRS separation statistics, the HJ distance measure can be jointly tested by the Wald statistics (Ren and Shimotsu (2009)). Recently, Barillas and Shanken (2016b) develop a Bayesian procedure that allows to compute model probabilities for the collection of all possible pricing models that can be formed from a given set of factors. They contribute the literature by analyzing the joint alpha restriction for a set of test assets in a Bayesian setting. Their paper's evidence casts strong doubt on the validity of their six-factor model.

Secondly, I show that there is a large variation in the shape and curvature of these HJ confidence sets, which makes it difficult for any single SDF to satisfy them. For example, industry portfolios have a higher HJ confidence set than that of the size/book-to-market portfolios. When using excess returns on test assets, the shape of derived confidence sets for the HJ set is flatter than that when using gross returns. This analysis helps us to understand why results that the different least misspecified model to explain cross-sectional returns on test assets may be different. Cochrane (2005) assert that there exist a proper set of SDFs that can price existing assets, because the entire universe of possible random payoffs may not be spanned by payoffs to existing assets, especially when markets are incomplete. About a decade ago it was lamented that "asset pricing (theory) has fallen on hard times" (Lettau and Ludvigson (2001)) - I find that this observation is still valid. Furthermore, using entropy to decompose the SDF, I explore the high variations of the filtered SDFs on our least misspecified consumption-based asset pricing models during financial bear and bull periods, but for those empirical factors models, entropy may not be the proper economic interpretation criteria.

In the rest of the paper, I 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 describes the empirical analysis. Section 5 concludes the paper.

# 2 Methodology

I use three different Hansen–Jagannathan (HJ) distance measures, the original HJ distance $\left(HJ^O\right)$ introduced by Hansen and Jagannathan (1997), the modified HJ distance $\left(HJ^M\right)$ in Kan and Robotti (2008), and the constrained HJ distance $\left(HJ^C\right)$ in Gospodinov et al. (2012) to compare candidate linear asset pricing models.

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

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

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 to estimate,

$$min_b \left[g_T(b)^{'} W g_T(b)\right], \tag{2.2}$$

where $g_T$ are the moment conditions and $W$ is a weighting matrix. Most earlier papers use this 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. I assume that there are two models. One is the CCAPM with the SDF $y_{t+1}^{(1)} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}$, and the other is the CAPM with the SDF $y_{t+1}^{(2)} = a + \theta R_{m,t+1}$, where $R_{m,t+1}$ is the market return. Use the J-statistics, I can find that the over-identification restrictions are not rejected for $y_{t+1}^{(1)}$ but for $y_{t+1}^{(2)}$. However, does the above result mean that the CCAPM $y_{t+1}^{(1)}$ is superior? No. Ludvigson (2011) stated that Hansen's $J_T$ test statistic depends on the model specific $S$ matrix. Hence, the CCAPM can look better since its SDF and the pricing errors are more volatile than those of CAPM.

## 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

$$y_{t+1} = \theta^{'} f_{t+1}, \tag{2.3}$$

where $f_t$ denotes the pricing factors. By using the pricing equation, I can derive the below equation

$$\alpha_t(\theta) = R_t y_t(\theta) - I_N = R_t \theta f_t^{'} - I_N, \tag{2.4}$$

where $R_t = \left[R_{1,t}, R_{2,t}, ..., R_{N,t}\right]^{'}$ 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

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

The $HJ^O$ measure is equivalent to a GMM estimator with the moment condition $E\left[\alpha_t(\theta)\right] = 0$ and the weighting matrix $\left[E(R_t R_t^{'})\right]^{-1}$, which is different from the optimal matrix (see Appendix for details). There are two advantages that I choose to use the HJ distance. The first advantage is stated by Ludvigson (2011) that the HJ distance does not reward SDF volatility. As a result, it is suitable for model comparison. Second, the HJ distance provides a measure of model misspecification. The HJ distance also gives the maximum pricing error of any portfolio formed from the $N$ assets. From the economic point of view, the HJ distance can be explained by the minimum distance between a candidate discount factor $y$ and the space of true discount factors, which can be understood the same as the minimum value of the J-statistic criterion but with $W = E\left(xx^{'}\right)^{-1}$ as weighting matrix[1].

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

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

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

Another problem I need to consider about is that all of the above SDFs can be either positive or negative. For instance, if the markets are incomplete, candidate SDFs such as the CAPM and linear factor models as suggested by Ross (1973) 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 to take on negative values with positive probability. This happens when these exist arbitrage opportunities 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 non-negative.

[1] Ren and Shimotsu (2009) improve the finite sample properties of the HJ distance test by using the weighting matrix $W = E\left(xx^{'}\right)^{-1}$.

Gospodinov, Kan and Robotti (2010) (Gospodinov et al. (2010)) solve for the constrained HJ distance as

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

subject to

$$\frac{1}{T}\sum_{t=1}^{T} m_t R_t = \bar{q},$$

$$m_t \geq 0, t = 1, ..., T,$$

where $y_t$ denotes the candidate SDF, $m_t$ stands for admissible SDF in the set $\aleph_+$, $q_{t-1}$ is the vector of corresponding costs of $N$ assets and $E[q_{t-1}] \neq 0$.

## 2.2 Multiple Comparisons Tests for the HJ Distance

The traditional HJ distance test compares HJ distance measures statistically by making pairwise model comparison null hypothesis, i.e., $HJ_2^O$ may be less than $HJ_1^O$; are they statistically 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 non-nested, 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 non-nested and overlapping. For non-nested and overlapping models they introduce a multivariate inequality test based on Wolak (1987) and Wolak (1989).

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

$$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{A} N(0_p, \Omega_{\hat{\rho}}). \tag{2.8}$$

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

$$min_{\rho}(\hat{\rho} - \rho)^{'} \hat{\Omega}_{\hat{\rho}}^{-1}(\hat{\rho} - \rho), \tag{2.9}$$

$$s.t. \rho \leq 0_r, \tag{2.10}$$

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

$$LR = T(\hat{\rho} - \widetilde{\rho})^{'} \hat{\Omega}_{\hat{\rho}}^{-1}(\hat{\rho} - \widetilde{\rho}). \tag{2.11}$$

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 $\rho$, which is $\rho = 0_P$. Under this value, $LR$ follows a 'chi-bar-squared distribution',

$$LR \xrightarrow{A} \sum_{i=0}^{p} w_i(\Omega_{\hat{\rho}}^{-1})X_i, \tag{2.12}$$

where the $X_i$ are independent $\chi^2$ random variables with $i$ degrees of freedom and $\chi_0^2$ 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 $y_t^A(\lambda_1^*) = y_t^i(\lambda_i^*)$ 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 $H_0 : \varphi_2 = 0_{k_2-k_1}, ..., \varphi_{p+1}(\lambda_{p+1}^*) = 0_{k_{p+1}-k_1}$. 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

$$\delta_{1,T}^2 = min(d_{j,T}^2)_{j=1}^K. \tag{2.13}$$

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 $T^W = max_{2,...,5}\sqrt{T}(d_{1,T}^2 - d_{j,T}^2)$, 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, continues, 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

$$p_W = \frac{1}{B}\sum_{b=1}^{B} I_{(T^{W,b} > T^W)}, \tag{2.14}$$

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.

## 2.3 The Inference on Hansen-Jagannathan Mean-Variance Sets

Chernozhukov, Kocatulum and Menzel (Chernozhukov et al. (2012)) develop the inference methods on HJ sets defined by nonlinear inequality restrictions to check which 'families' of SDFs price the assets correctly and which do not.

They define $K$ be a compact convex body in $\mathbb{R}^2$, and the set of admissible means and standard deviations in $K$

$$\Theta_0 := \left\{admissible(\mu, \sigma^2) \in \mathbb{R}^2 \cap K\right\}, \tag{2.15}$$

is introduced by Hansen and Jagannathan (1991) and known as the Hansen-Jagannathan (HJ) set. The boundary of this set is known as the HJ bound.

They also define $\Gamma = (S_{vv}, S_{v1}, S_{11})'$, where $S_{vv} = v'\Sigma^{-1}v$, $S_{v1} = v'\Sigma^{-1}1_N$, $S_{11} = 1_N'\Sigma^{-1}1_N$, $v$ and $\Sigma$ denote the vector of mean returns and covariance matrix to assets $1, 2, ..., N$, which are assumed not to vary with information sets at each period $t$. Then the minimum variance $\sigma^2_{HJ}(\mu)$ achievable by a SDF given mean $\mu$ is equal to

$$\sigma^2_{HJ}(\mu) = S_{vv}\mu^2 - 2S_{v1}\mu + S_{11}. \tag{2.16}$$

Hence, the HJ set is equal to

$$\Theta_{HJ} = \left\{(\mu, \sigma) \in (\mathbb{R} \times \mathbb{R}_+) \cap K : \sigma_{HJ}(\mu) - \sigma \leq 0\right\}. \tag{2.17}$$

Chernozhukov, Kocatulum and Menzel construct a confidence region $R$ such that

$$lim_{n\to\infty} P\{\Theta_{HJ} \subseteq R\} = 1 - \alpha, \tag{2.18}$$

$$d_H(\Theta_{HJ}, R) = O_p\left(1/\sqrt{n}\right), \tag{2.19}$$

where $d_H$ is the distance measure. Following Chernozhukov, Kocatulum and Menzel (2012), I construct confidence regions for the set $\Theta_{HJ}$ using LR and Wald-type Statistics, and $\Theta_{HJ}$ can stand for the HJ bounds that determines which combinations of $\Theta_{HJ}$ the first two moments of the SDF generated by a given family of asset pricing models fall in the HJ set. Specifically, $\Theta_{HJ} = \left\{\mu(\varsigma), \sigma(\varsigma), \varsigma \in [0, \infty)\right\}$, where $\varsigma$ is the elasticity of power utility function, and we can check which values of $\gamma$ give us overlap of $\Theta_{HJ}$ with the confidence set $R$ for the HJ set.

Different from Chernozhukov, Kocatulum and Menzel (2012), however, my paper empirically applies different set estimate and inference to check which 'families' of SDFs price our test assets correctly. I firstly estimate confidence sets for the HJ sets of mean-variances of SDFs using unweighted Wald-type statistics that complements other approaches based on the direct Hausdorff distance suggested in Beresteanu and Molinari (2008), Hausdorff distance in Molchanov (1998) and a structured projection approach. The reason is that these three approaches show different test powers on set inference, where the structured projection approach seems much more natural and "economically appealing". However, the structured projection approach remains very conservative and is much less powerful than the approach based on the optimally weighted LR-type (the Hausdorff distance) and W-statistics (Chernozhukov, Kocatulum and Menzel (2012)).

# 3 Data and Candidate Models

## 3.1 Data

The main empirical analysis uses both monthly and quarterly return data over the period 1967 - 2012 for the U.S. market. I use two test assets, the Fama-French 25 size- and book-to-market sorted portfolios (FF25) and 30 industry portfolios. The industry portfolios are included to provide a greater challenge to the various asset-pricing models, as recommended by Lewellen et al. (2010). In the multiple comparison tests of asset pricing models, I use both gross and excess returns on test assets. I also use the one-month and three-month T-bill rate from FRED at the Federal Reserve Bank of St. Louis over the period January 1967 to December 2012 as the riskless rate of interest. Moreover, the Fama–French factors and the momentum factor are from the Kenneth French's website and the Hou, Xue and Zhang (2015) four factors obtain from the original paper of Hou et al. (2015).

I 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 indices for nondurables and services. My durable consumption renews Yogo's (2006) and I update it to 2012. My non-housing consumption is measured by the nondurables consumption but excludes services such as shoes, clothing and housing.

The factors used as conditional variables include: (i) the aggregate consumption to wealth ratio $cay_t$ in Lettau and Ludvigson's (LL, 2001) conditional CCAPM (available on Professor Ludvigson's website); (ii) the housing collateral ratio $mymo_t$ in Lustig and Van Nieuwerburgh's (LVN, 2004) conditional CCAPM; $mymo_t$ is computed by the ratio of collateralizable housing wealth to non-collateralizable human wealth, which are from the Historical Statistics for the US (Bureau of the Census) and the Flow of Funds data (Federal Board of Governors), and (iii) the non-housing consumption expenditure share $s_t$ in Piazzesi, Schneider and Tuzel (SPST, 2007) conditional CCAPM; the expenditure share relies on per-period dollar expenditures on the item in NIPA.

## 3.2 Description of Candidate Models

I focus mainly on empirically motivated linear asset pricing models given their popularity in applied work. These include the Fama and French three-factor model, the Fama–French and Carhart momentum four-factor model, the Fama and French five-factor model and the Hou, Xue and Zhang (HXZ) investment-based model. In addition to these empirical models, I also include the Capital Asset Pricing Model (CAPM).

In their stochastic discount factor (SDF) representations, these models can be written as follows. The CAPM in which the expected excess return on an asset equals the market risk $\theta_1$ of the asset times the expected excess return on market portfolio (excess return (in excess of the one-month T-bill rate) on the value-weighted stock market index (NYSE-AMEX-NASDAQ)),

$$y_{t+1}^{CAPM} = \theta_0 + \theta_1 vw_{t+1}, \tag{3.1}$$

where $vw_{t+1}$ denotes excess returns on the market portfolios.

Fama and French (1992) (FF3) document the role of size and book/market ratio and Fama and French (2015) (FF5) the role of profitability and investment factors in explaining the cross-section of expected stock returns.

$$y_{t+1}^{FF3} = \theta_0 + \theta_1 vw_{t+1} + \theta_2 smb_{t+1} + \theta_3 hml_{t+1}, \tag{3.2}$$

and

$$y_{t+1}^{FF5} = \theta_0 + \theta_1 vw_{t+1} + \theta_2 smb_{t+1} + \theta_3 hml_{t+1} + \theta_4 rmw_{t+1} + \theta_5 cma_{t+1}, \tag{3.3}$$

where $smb$ is the difference between the return on diversified portfolios of small and large stocks, $hml$ is the difference between the return on diversified portfolios of high and low book-to-market stocks, $rmw$ is the difference between the return on diversified portfolios of stocks with robust profitability and weak profitability, and $cma$ is the difference between the return on diversified portfolios of stocks with low investment and high investment stocks.

HXZ propose an empirical $q$-factor model with following pricing factors:

$$y_{t+1}^{HXZ} = \theta_0 + \theta_1 vw_{t+1} + \theta_2 me_{t+1} + \theta_3 ia_{t+1} + \theta_4 roe_{t+1}, \tag{3.4}$$

where $me$ is the difference between the return on diversified portfolios of small size stocks and the return on a portfolio of big size stocks, $ia$ is the difference between the return on diversified portfolios of stocks with low and high investment, and $roe$ is the difference between the return on on diversified portfolios of stocks with high profitability and low profitability, where profitability is measured by the return on equity ($roe$).

Finally, I use the Carhart four-factor model with the momentum effect (Carhart (1997)),

$$y_{t+1}^{Carhart} = \theta_0 + \theta_1 vw_{t+1} + \theta_2 smb_{t+1} + \theta_3 hml_{t+1} + \theta_4 mom_{t+1}, \tag{3.5}$$

where $smb$ is the difference between the return on diversified portfolios of small and large stocks, $hml$ is the difference between the return on diversified portfolios of high and low book-to-market stocks, and $mom$ is the average of the return on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios.

My consumption-based linear factor models include the consumption-based CAPM (CCAPM) that can be written as

$$y_{t+1}^{CCAPM} = \theta_0 + \theta_1 c_{t+1}^{ndur}, \tag{3.6}$$

where $c_{t+1}^{ndur}$ is the growth rate of non-durable consumption.

I also use the Yogo (2006) CCAPM where includes durable consumption risk

$$y_{t+1}^{DUR} = \theta_0 + \theta_1 vw_{t+1} + \theta_2 c_{t+1}^{ndur} + \theta_3 c_{t+1}^{dur}, \tag{3.7}$$

where $vw_{t+1}$ is the excess returns on market portfolios and $c_{t+1}^{dur}$ denotes the consumption growth rate of durable goods.

Further I use the Piazzesi, Schneider and Tuzel (2007) that introduces the consumption-housing CCAPM

$$y_{t+1}^{PST} = \theta_0 + \theta_1 c_{t+1}^{nh} + \theta_2 s_{t+1}, \tag{3.8}$$

where $c_{t+1}^{nh}$ is the growth rate of non-housing consumption and $s_{t+1}$ denotes the log non-housing consumption expenditure share.

Finally I use a set of conditional asset pricing models. They include the conditional CCAPM of Lettau and Ludvigson (2002), which shows the consumption-wealth ratio can capture the time-varying risk premiums,

$$y_{t+1}^{LL} = \theta_0 + \theta_1 c_{t+1}^{ndur} + \theta_2 cay_t + \theta_3 c_{t+1}^{ndur} cay_t, \tag{3.9}$$

where $c_{t+1}^{ndur}$ is the growth rate of non-durable consumption and $cay_{t-1}$ is the consumption-wealth ratio.

My second candidate is the scaled CCAPM with the non-consumption expenditure share of Piazzesi, Schneider and Tuzel (2007). They find that while the non-housing expenditure ratio changes, the composition risk that relates changes in asset prices also changes,

$$y_t^{SPST} = \theta_0 + \theta_1 c_{t+1}^{ndur} + \theta_2 s_t + \theta_3 c_{t+1}^{ndur} s_t, \tag{3.10}$$

where $s_t$ is the non-housing consumption expenditure share.

Finally, I also use the scaled CCAPM with the collateral-consumption ratio of Lustig and Van Nieuwerburgh (2005) where shows the ratio of housing wealth to human wealth changes the conditional distribution of consumption growth across households in a model with collateralized borrowing and lending,

$$y_{t+1}^{LVN} = \theta_0 + \theta_1 c_{t+1}^{ndur} + \theta_2 my_t + \theta_3 c_{t+1}^{ndur} my_t, \tag{3.11}$$

where $my_t$ is the housing collateral ratio.

# 4 Main Results

## 4.1 The HJ Distance Comparison Results

In Table 2, I report the results of the model comparison tests using both the gross returns and the excess returns on test assets. To enable comparison with the earlier papers, we discuss the results mainly with respect to the combination of FF25 and 30 Fama–French industry portfolios as test assets.

In Panel A, I report the HJ distance $(HJ^O)$, the modified HJ distance $(HJ^M)$ and the constrained HJ distance $(HJ^C)$ measures for all the models that we consider. These results allow us to compare the HJ distance measures across the models without formally testing for their differences. Here, lower values of the measures are preferred as they show lower levels of misspecification. The results show that across all the models and HJ distance measures that I consider, all the HJ distance measures are greater than zero. These results suggest that all the models are potentially misspecified. The model with the lowest $HJ^O$, $HJ^M$ and $HJ^C$ measures is the Carhart model when using the gross returns on the test assets, and when using excess returns the Carhart model has lowest $HJ^O$ and $HJ^M$ measures. Overall the results suggest that the Carhart model is the least misspecified of all the models I consider. The FF5 is the next best followed by the FF3 model. The CAPM is the most misspecified irrespective of whether we use gross or excess returns and across the three HJ distance measures.

In Panel B, I present the results of the two formal tests of model comparison. In both these tests, I take as our model of choice, the one that has the lowest $HJ^O$, $HJ^M$ and $HJ^C$ measures among all the models in Table 2 Panel A. The three HJ distance measures of alternative models are then formally tested again the measures of our chosen model, the Carhart 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 greater than the distance measures of our chosen model. The null hypothesis in this test is that our chosen model has a $HJ^O$, $HJ^M$ and $HJ^C$ measure that is less than that of any of the alternative models. A failure to reject the null means that our chosen model is the least misspecified model. Using gross returns, I find that the null is not rejected for any of the HJ distance measures. This suggests that the Carhart model is the least misspecified. When considering excess returns, I find that for the $HJ^M$ and the $HJ^C$, the null rejects at 5% and 10% respectively. Noting, how the measures are computed, the overall conclusion that we can draw is that the Carhart model is the least misspecified model, particularly when a more conservative size of the test is considered.

The second test is a pairwise comparison test of whether the $HJ^O$, the $HJ^M$ and the $HJ^C$ measures of each of the alternative models are significantly greater than the distance measures of our chosen model. The null hypothesis in this test is that the chosen model has a $HJ^O$, $HJ^M$ and $HJ^C$ measure that is less than each of the alternative models. As in Panel A, the tests are conducted using gross and excess returns on the test assets. The results show that across tests using gross and excess assets, the comparison tests with the FF3 and the FF5 reject the null at the 5% or 10% level of significance. At 1% level of significance the Carhart model retains its position as the least misspecified of all the models.

In Table 3, I run tests for the HJ measures and the multiple comparison tests only on the FF25 portfolios. In Table 3 Panel A, I find that all the models are misspecified with the Carhart model the least misspecified of all the models. In Panel B, the multiple comparison tests using a conservative size of 1%, the results support the conclusions from the main results in Table 2. However, some models do better than the Carhart model at less conservative test sizes and depending on whether I look at gross or excess returns on the test assets. This suggests that the any conclusions drawn are potentially sensitive to the choice of test assets.

Table 4 shows the pricing performance of our set of consumption-based asset pricing models to explain returns on Fama–French 25 size/book-to-market plus 30 industry portfolios. The results show that all the HJ distance measures are greater than zero significantly. These results suggest that all the models are potentially misspecified. The model with the lowest $HJ^O$, $HJ^M$ and $HJ^C$ measures is the Yogo durable consumption-based CAPM model when using the gross returns on the test assets, and when using excess returns the Yogo model has lowest $HJ^M$ and $HJ^C$ measures while the $HJ^O$ measure for all models is not significant different from zero. Overall the results suggest that the Yogo model is the least misspecified of all the consumption-based CAPM models I consider.

Table 5 reports pricing errors in the case of the set of conditional consumption-based asset pricing models. In explaining returns on Fama–French 25 size/book-to-market ratio and 30 industry portfolios, the Lettau and Ludvigson model

(2001) has the smallest normalized pricing error relative to the other models based on the HJ, the modified HJ and the constrained HJ distance measures.

Panel B in Table 5 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 candidate models. The null hypothesis in this test is that our chosen model has a $HJ^O$, $HJ^M$ and $HJ^C$ measure that is less than that of any of the alternative models. A failure to reject the null means that the Lettau and Ludvigson model is the least misspecified model. Using both gross and excess returns, I find that the null is not rejected for any of the HJ distance measures. This suggests that the Lettau and Ludvigson model is the least misspecified. Then, I show the pairwise comparison test of whether the $HJ^O$, the $HJ^M$ and the $HJ^C$ measures of each of the alternative models are significantly greater than the distance measures of the Lettau and Ludvigson model.

## 4.2 Sequential Selection Procedure and Rank Tests

Gospodinov et al. (2014) show that in the presence of misspecification and the lack of identification, the finite sample distributions of the statistics of interest can depart substantially from the standard asymptotic approximations developed under the assumption of correctly specified and fully identified models. Moreover, they propose an easy-to-implement sequential procedure that allows us to eliminate the useless factors from the model and show its asymptotic validity.

In this section, I report the results of the ranks tests of the individual factors, the model misspecification tests, rank tests of the models and the sequential testing procedure.

In Table 6 Panel A, I report the results of the rank restriction test of the 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. This test is important since the presence of a useless factor leads to a violation of a crucial condition for identification which is 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 we consider. This suggests that all of the factors can be considered as potentially useful.

In Table 6 Panel B, we report 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 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 over reject the null, we also conduct a 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 but approximate representations of reality and are therefore potentially misspecified.

In rank test of the models, I find that three of the five models I consider, i.e. the CAPM, the FF3 and the Carhart models do not suffer from identification problems at 1% level. And indeed all of the models are identified at 5% level with p-values of 0.015 and 0.026 for the FF5 and the HXZ models. This is consistent with the fact that the factors we consider are correlated with returns on the test assets. Gospodinov et al. (2014) note that the $vw$, $smb$ and $hml$ factors are highly correlated with the test asset returns.

However, all the macro variables are useless, such as the non-durable consumption ($c^{ndur}$), the durable consumption ($c^{dur}$), the non-housing consumption ($c^{nh}$), the expenditure ratio ($s$) in the Piazzesi, Schneider and Tuzel model, the consumption-wealth ratio ($cay$) and the collateral consumption ratio ($my$) in the Lustig and Van Nieuwerburgh model.

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. However, all the macro factors are useless; the identification condition fails for macro factors and this may also affect the validity of any statistical inference. Hence, I will use the HJ set inference and entropy to further explain the consumption-based and conditional consumption-based asset pricing models' results.

In Table 6 Panel C and D, I 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 $vw$ factor in the CAPM is significant, in the FF3 model, the $vw$ and the $hml$ are significant. In the Carhart model, the $vw$, the $hml$ and the $mom$ are significant. In the FF5 model, the $vw$ and the $hml$ are useful while the $smb$, the $rmw$ and the $cma$ do not survive the sequential testing procedure. In the HXZ model the $vw$ and the $ia$ factors survive the sequential testing procedure at the 5% significance level. The factors that survive all the specifications they are included in are the $mom$ and the $ia$ factors, followed by the $vw$ and the $hml$ which is significant in ten out of the thirteen and three out of the five specifications they are included in. The results in panel D are identical to the results in Panel C, except for the $me$ factor in the Hou et al. (2015) that appears to be useful assuming correct model specification, turns to be not useful when potential model misspecification is taken into account.

### 4.3 The HJ Set Inference

Now I have analyzed the economic reasons that why theoretical motivated models outperform others to explain specific test portfolios by using the HJ set inference.

There are two families of SDFs on the candidate asset pricing models: the consumption-based and the conditional consumption-based models. I treat both the consumption and the conditional (scaled) consumption based models as the augmented CCAPM models. For the conditional consumption-based models, any of these asset pricing model can be expressed as multiple factor models by multiplying out the conditioning variables and the fundamental factor (CCAPM) (Ludvigson (2011), and Cochrane (2001)).

Figure 1 shows HJ sets using weighted LR statistic and the structured projection approaches when the consumption-based models explain the gross returns on Fama–French 25 size and book-to-market ratio portfolios. The first row gives the CCAPM, the second row is Piazzesi, Schneider and Tuzel (2007) housing consumption model and the third row shows the Yogo durable consumption model. The blue dashed line in the left is the 95% confidence region based on the weighted LR statistic, and the blue dashed in the right denotes the 95% confidence region based on the structured projection. Both approaches' confidence regions cover most of the bootstrap draws below the HJ bounds. However, it should also be noted that the confidence bound based on the weighted LR statistic is fairly tight relative to the confidence bound based on the projection; the structured projection confidence set performs quite poorly relative to the LR-based confidence set: in particular the latter is much smaller and lies strictly inside the former. In fact, the precision of the confidence set based on structured projection is poor enough to overturn the major empirical conclusion that the consumption-based CAPM cannot be reconciled with small values of risk aversion. This should be expected since the projection confidence bounds are based on a confidence set for the point-identified parameters that does not account for the specific shape of the bounds as a function of $\Gamma$. Again, I result that the Yogo durable consumption-based CCAPM outperforms because of the recursive utility function that requires a relatively smaller risk aversion to solve equity premiums.

Figure 2 shows the confidence set for HJ sets using excess returns on test portfolios. Comparing with the confidence set for HJ sets using gross returns, the current HJ sets are a little bit flatter. Hence, the HJ bound using the structured projection approach is much easier to arrive at for the Piazzesi, Schneider and Tuzel (2007) housing consumption model, whose utility function is the CES-composite power utility with the non-separability between the non-durable consumption and housing services. According to the second row in Figure 2, the housing consumption model requires the relative risk aversion equal to 5 to explain the cross-sectional equity premiums.

The HJ sets for industry portfolios look sharper than the HJ sets for Fama–French 25 size and book-to-market portfolios. This is the same result as we find in Figure 2, in which the industry portfolios obtain the highest HJ bound to reach. However, the structure projection that are based on a confidence set for the point-identified parameters makes the 95% confidence set like a 'bowl' as shown in all figures on the right side of Figure 3.

When using excess returns on industry portfolios, the 'bowl' will be much wider, therefore all conditional models can reach the bound by choosing small relative risk aversion. Here, both Lustig-Van Nieuwerburgh and Piazzesi-Schneider-Tuzel models follow the same power utilities with the non-separability between non-durable consumption and housing services or collateral housing ratio, but their intratemporal elasticity of substitution index are different, where the Lustig-Van Nieuwerburgh's intratemporal elasticity of substitution is smaller than one and the Piazzesi-Schneider-Tuzel model chooses the index above one. I result that the Lettau and Ludvigson scaled consumption-based CAPM outperforms other conditional consumption-based model to explain gross and excess returns on momentum portfolios in Figure 4 and Figure 5. Overall, to explain specific test assets, the least misspecification asset pricing models depend on two elements: the theoretical motivated model structure and the HJ bounds of test assets.

## 4.4 The Decomposition of the Modified HJ Distance

In this section, I show that the HJ distance results will depend on the choice of test assets if the asset pricing models need to explain returns on test assets, or the researchers only focus on the scenario that which asset pricing model is able to price the factors in the other asset pricing model.

I show that another form of the squared modified HJ distance (MHJD) is

$$\left[HJ^M\right]^2 = \alpha_{M1}^{'}\Sigma^{-1}\alpha_{M1} = Sh(f_1,f_2,R)^2 - Sh(f_1)^2, \tag{4.1}$$

where $Sh(\cdot)$ denotes the maximum squared Sharpe ratio (the mean excess return over standard deviation) obtainable from portfolios of the given returns, $\alpha_{M1} = (0,\alpha_{21},\alpha_{R1})$ denotes the alpha vector for evaluating the pricing of all these investments under the more parsimonious model $M$, $f_1$ could be the CAPM, $f_2$ could be the CAPM nested model Fama–French three-factor model and $R$ denotes returns on test assets. The first element shows that the alphas of $f_1$ on $f_1$ are necessarily 0. The second element $\alpha_{21}$ refers to the alphas of $f_2$ on $f_1$, and the last element $\alpha_{R1}$ is the alphas of $R$ on $f_1$, and $V$ is the covariance matrix for $(f_1,f_2,R)$.

We can rewrite the equation (4.1) into

$$\left[HJ^M\right]^2 = Sh(f_1, f_2, R)^2 - Sh(f_1, f_2)^2 + Sh(f_1, f_2)^2 - Sh(f_1)^2 = \alpha_R^{'}\Sigma^{-1}\alpha_R + \alpha_{21}^{'}\Sigma_{21}^{-1}\alpha_{21}. \tag{4.2}$$

The equation (4.2) shows that the squared MHJD is the sum of the quadratic form for adding the test assets $R$ to $(f_1, f_2)$ plus the quadratic form for adding the factors to $f_2$ to $f_1$. To explain the equation (4.2), I can establish the nested models, $M_1$ and $M$, for instance, $M_1$ is the CAPM and $M$ is the Fama–French three-factor model. First, I would like to see how excluded factors (the $SMB$ and the $HML$ factors) could influence the squared MHJD for a given test assets. We can obtain the squared MHJD of the CAPM ($M_1$) as

$$\left[HJ^M_{m,M_1}\right]^2 = \alpha_{R1}^{1}\Sigma_{R1}^{-1}\alpha_{R1}, \tag{4.3}$$

so the influence that $f_2$ has put on the squared MHJD is

$$\left[HJ^M\right]^2 - \left[HJ^M_{m,M_1}\right]^2 = \alpha_R^{'}\Sigma^{-1}\alpha_R - \alpha_{R1}^{1}\Sigma_{R1}^{-1}\alpha_{R1} + \alpha_{21}^{'}\Sigma_{21}^{-1}\alpha_{21} \equiv \left[HJ^M_{m,f}\right]. \tag{4.4}$$

Here, $\left[HJ^M_{m,f}\right]^2$ is denoted as the difference between the squared MHJD of model $M$ and that of $M_1$ when factor $f_2$ is added into $M_1$. It reflects the added-factors influence on the squared MHJD of $M$, or part of the squared MHJD of derived from added factors, when test assets remain constant.

According to the equation (4.4), we define $\alpha_R^{'}\Sigma^{-1}\alpha_R - \alpha_{R1}^{1}\Sigma_{R1}^{-1}\alpha_{R1}$ as the difference between two nested models' GRS test statistics. The second part $\alpha_{21}^{'}\Sigma_{21}^{-1}\alpha_{21}$ is defined as the GRS test statistics of the $SMB$ and the $HML$ factors effects when comparing the Fama–French three-factor and the CAPM models.

I can result that the choice of cross-sectional test portfolios influences the MHJD between two linear factor models through the GRS test statistics channel but not the excluded factors channel. Hence, to tell which linear factor model is better than the other linear factor model to explain cross-sectional returns given the test portfolios, if we want to use the MHJD for evaluating their pricing performances, we need to tell how much effect comes from the excluded factors and how much effect comes from the GRS test statistics. However, the MHJD statistics in Kan and Robotti (2008) generalized these two parts simultaneously and statistically. More importantly, they provide the asymptotic distribution of the MHJD.

## 4.5 The Entropy Explanation

Supposing the Euler equation will be hold in equilibrium for the probability measure which is unobserved, I non-parametrically extract filtered SDFs via a relative entropy minimization (Kitamura and Stutzer (1997)) based on Kullback-Leibler information criterion (KLIC) under the asset pricing restrictions coming from the Euler equation[2]. Interestingly, those least misspecified filtered SDF components follow some clear financial market cycle patterns; especially they have significant correlation with financial market crashes unrelated to economy-wide contractions.

I present that the difference between the least and the most misspecified candidates in entropy. The recession period data comes from NBER and the financial market crashes and bull periods data come from Mishkin and White (2002). Here financial market crashes means a 20% drop in the market; speed is another feature. Therefore, I look at declines over windows of three months and one year.

At three months, the Dow additionally identifies, several more months in the 1930-1932 slide, November and December 1987, October, November and December 1937, and June 1962. Similarly, the S&P500 finds crashes in 1930-1933, 1937, 1962, 1974, and 1987, with June 1940 narrowly missing at 19.2 percent. For the NASDAQ, there are crashes in a three-month window in 1974, 1987, 1990, 2000, 2001.

For the Dow Jones, declines in the 12 month window in excess of 20%, pick out 1930-1933, 1937, 1938, 1970, 1974, and 1988. Using the S&P500 and a 12 month window there are crashes, ending in months in 1930-1933, 1937, 1938, 1941, 1947, 1970, 1974, and 1975. Looking at the NASDAQ for one year crashes, we pick out the same years as using a three month window, plus 1973, 1975, 1982, 1983, and 1984.

The Figure 6 shows that movements of filtered SDF for the consumption-based asset pricing models. The green dash line plots the benchmark consumption-based CAPM. The least misspecified Yogo (2006) is plotted in blue line. The gray shaded areas represent NBER-dated recessions, the red line stands for the financial market crashes periods and the aquamarine line is the financial market bull periods. While obtaining the same time-varying volatile filtered SDF, its correlation with the financial market crashes and bull periods are almost near 92.31%.

The Figure 7 gives us another graph showing the least misspecified one comparing to the benchmark for conditional consumption-based models. The Lettau and Ludvigson conditional consumption model (the blue line) varies more sharply than the consumption-CAPM. Moreover, the Lettau and Ludvigson model captures the financial crashes in 1962, in 1974, and in 1987, the correlation with the financial market crashed is 76.2%.

The drawback of entropy is that it can only extract the filtered SDF within nested models. The model like Santos and Veronesi (2006) can be

[2]The Appendix shows details on entropy and the filtered pricing kernel process.

filtered via entropy only by using mimicking portfolios, i.e., the excess returns on market portfolios, therefore making a nested model to CAPM.

# 5 Conclusions

In this paper, I use an invariant measure, the HJ distance to compare the degree of misspecification between a set of prominent linear factor models. I create a level playing field by having a common set of the most equity test assets; size, the Fama–French size and book-to-market and industry portfolios over a long time period. I use recently developed tests that allow me to compare the HJ-distance across different models based on a statistical criterion. In addition, I use set inference techniques that allow us to coma pore families of SDFs.

The main result is that my set of candidate asset pricing models are unable to consistently price the cross-section of even the most basic set of test assets and that no model paper to dominate. Specifically, the Carhart model should be a reasonably good model to use in a practical setting where risk adjustment is necessary. Among consumption-based models, the Yogo model does better than the Piazzesi, Schneider and Tuzel and the classical CCAPM model when pricing most test portfolios. In comparisons of conditional asset pricing models, I find that the Lettau-Ludvigson model obtain smaller HJ distances than the other candidates when the test assets are the size and book to market and industry portfolios. I note that my conclusions are drawn from empirical tests where I use the size and book-to-market and industry portfolios as test assets. This choice mitigates concerns about there being a factor structure in the size and book-to-market portfolios (see for example, Lewellen et al. (2010)) but importantly, this allows for the comparison of results from the previous work that considers variants of the conditional and unconditional consumption-based CAPM and the ICAPM such as Gospodinov et al. (2014) and others by using a similar set of test assets.

I acknowledge that our model comparison and ranking could be sensitive to the choice of test assets. I study the confidence regions for the HJ set of admissible stochastic discount factors as in (Chernozhukov et al. (2012)). I find that there is a large variation in the shape and curvature of these confidence sets that makes it difficult for any single SDF to satisfy them. Second, when using excess returns on test assets, the shape of derived confidence sets for the HJ set is flatter than that when using gross returns. My analysis helps in understanding why none of candidate models is able to dominate in pricing the set of test assets. However, the tests we employ require us to make a specific choice of test assets. If asset pricing models are to be compared such that the tests are invariant to the test assets, one approach could be to go down the route which considers the ability of factors in one model to price factors in another model form as a basis for model comparison. My general conclusion is that while we may never have a model that can price the cross-section of a broad range of test assets, further efforts are required in building models that can do better across a set of basic test assets.

# References


Barillas, F. and J. Shanken (2016a). Which alpha? *Review of Financial Studies*, hhw101.

Barillas, F. and J. A. Shanken (2016b). Comparing asset pricing models.

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

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.

Chernozhukov, V., E. Kocatulum, and K. Menzel (2012). Inference on sets in finance. Technical report, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Cochrane, J. (2001). Asset pricing.

Cochrane, J. H. (2005). *Asset pricing*, Volume 1. Princeton university press Princeton.

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.

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

Gospodinov, N., R. Kan, and C. Robotti (2010). On the hansen-jagannathan distance with a no-arbitrage constraint.

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., C. Robotti, and R. Kan (2012). On the hansen-jagannathan distance with a no-arbitrage constraint. *Available at SSRN 1534727*.

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.

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

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. Technical report, National Bureau of Economic Research.

Jagannathan, R. and Z. Wang (1996). The conditional capm and the cross-section of expected returns. *Journal of Finance*, 3–53.

Kan, R. and C. Robotti (2008). Specification tests of asset pricing models using excess returns. *Journal of Empirical Finance 15*(5), 816–838.

Kitamura, Y. and M. Stutzer (1997). An information-theoretic alternative to generalized method of moments estimation. *Econometrica: Journal of the Econometric Society*, 861–874.

Kogan, L. and M. H. Tian (2013). Firm characteristics and empirical factor models: a data-mining experiment. *FRB International Finance Discussion 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.

Ludvigson, S. (2011). Advances in consumption-based asset pricing: Empirical tests. Technical report, National Bureau of Economic Research.

Mishkin, F. and E. White (2002). Us stock market crashes and their aftermath: implications for monetary policy. Technical report, National Bureau of Economic Research.

Parker, J. and C. Julliard (2005). Consumption risk and the cross section of expected returns. *Journal of Political Economy 113*(1), 185–222.

Piazzesi, M., M. Schneider, and S. Tuzel (2007). Housing, consumption and asset pricing. *Journal of Financial Economics 83*(3), 531–569.

Ren, Y. and K. Shimotsu (2009). Improvement in finite sample properties of the hansen–jagannathan distance test. *Journal of Empirical Finance 16*(3), 483–506.

Ross, S. (1973). *The arbitrage theory of capital asset pricing*. Rodney L. White Center for Financial Research, University of Pennsylvania, The Wharton School.

Santos, T. and P. Veronesi (2006). Labor income and predictable stock returns. *Review of Financial Studies 19*(1), 1–44.

Vassalou, M. and Y. Xing (2004). Default risk in equity returns. *The Journal of Finance 59*(2), 831–868.

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. *The 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

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

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

$$G_T = \frac{1}{T}\sum_{t=1}^{T}R_tR_t^{'} = \frac{1}{T}R^{'}R, \tag{A.3}$$

where

$$R = [R_1, R_2, ..., R_T]^{'},$$

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

The sample analog of the HJ distance is thus

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

Taking the derivative of the above equation

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

which gives an analytic expression for the sample minimizer

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

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

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

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

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

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

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

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

the estimates of risk premiums will change into

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

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 $\chi^2$ distribution with $N-K$ freedom, where is the dimension of $\theta$.

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

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

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

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

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 $\Omega_T$ of the matrix $\Omega$, we can estimate the matrix $A$ by replacing $\Omega$ and $G$ by $\Omega_T$ and $G_T$, respectively. Under the hypothesis that the SDF prices the returns correctly, The $\Omega$ 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 $u_j = \sum_{i=1}^{N-K} a_i\chi^2(1)$. Here M is the number of simulation. Then the empirical p-value of the HJ distance is

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

where $I(.)$ 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 $y_t(\theta) \in \aleph^+$, which implies that $\iota = 0_N$ and $\delta_+ = 0$; the model is misspecified if $y_t(\theta) \notin \aleph^+$ for all $\theta \in \Gamma$, which implies that $\delta_+ > 0$.

They show that

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

$$T\hat{\delta_+^2} \xrightarrow{A} \sum_{t=1}^{N-K} \varsigma_i v_i, \tag{A.15}$$

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

$$A = P^{'} U^{-\frac{1}{2}} S U^{-\frac{1}{2}} P, \tag{A.16}$$

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,

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

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

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

Table 1: **Statistical Summary on Pricing Factors**

*Notes:* I use monthly return data over the period 1967 - 2012 for the U.S. market. The table shows the statistical summary for all the pricing factors, including $vw$ the market portfolios in the CAPM; the size effect $smb$, the book-to-market ratio effect $hml$, the difference between the returns on diversified portfolios of stocks with robust and weak profitability $rmw$ and the difference between the returns on diversified portfolios of low and high investment stocks $cma$ in the Fama–French three- and five-factor models; the difference between the return on a portfolio of small size stocks and the return on a portfolio of big size stocks $me$, the difference between the return on a portfolio of low investment stocks and the return on a portfolio of high stocks $ia$, and the difference between the return on a portfolio of high profitability return on equity stocks and the return on a portfolio of low profitability $roe$ in the Hou, Xue and Zhang (2015); the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios $mom$ in the Carhart (1997). $c^{ndur}$ is the growth rate of non-durable consumption for the consumption-based CAPM. $c^{dur}$ denotes the consumption growth rate of durable goods for the Yogo (2006) CCAPM. $c^{nh}$ is the growth rate of non-housing consumption and $s$ denotes the log non-housing consumption expenditure share of Piazzesi, Schneider and Tuzel (2007). Moreover, $cay$ is the consumption-wealth ratio in the conditional CCAPM of Lettau and Ludvigson (2002). $my$ and $mymor$ are the housing collateral ratio in the Lustig and Van Nieuwerburgh (2005).

| Variable | Number Obs | Mean | Std.Dev. | Minimum | Maximum |
|---:|---:|---:|---:|---:|---:|
| $vw$ | 552 | 0.0046 | 0.0463 | -0.2324 | 0.161 |
| $smb$ | 552 | 0.0011 | 0.0319 | -0.22 | 0.1101 |
| $hml$ | 552 | 0.0046 | 0.0296 | -0.0978 | 0.1384 |
| $rmw$ | 552 | 0.0028 | 0.0233 | -0.1911 | 0.1352 |
| $cma$ | 552 | 0.0036 | 0.0207 | -0.0688 | 0.0955 |
| $mom$ | 552 | 0.0068 | 0.0440 | -0.3458 | 0.1838 |
| $me$ | 552 | 0.0034 | 0.0314 | -0.1439 | 0.2213 |
| $ia$ | 552 | 0.0045 | 0.0191 | -0.0716 | 0.0925 |
| $roe$ | 552 | 0.0058 | 0.0257 | -0.1385 | 0.1038 |
| $c^{ndur}$ | 263 | 0.4752 | 0.9028 | -4.87 | 10.16 |
| $c^{dur}$ | 263 | 1.0192 | 2.6564 | -1.13 | 11.63 |
| $c^{nh}$ | 263 | 1.467 | 0.958 | -2.952 | 8.419 |
| $s$ | 263 | 0.7655 | 0.0421 | 0.576 | 0.791 |
| $cay$ | 263 | -0.019 | 1.6707 | -4.06 | 3.93 |
| $my$ | 263 | 0.9776 | 0.0056 | 0.9651 | 0.9891 |

Table 2: **Test Portfolios: FF25 plus 30 Industry Portfolios**

*Notes:* the table represents the results of the model comparison tests using both the gross returns and the excess returns on 25 size and book-to-market sorted plus 30 industry sorted portfolios. Panel A reports the HJ distance $\left(HJ^O\right)$, the modified HJ distance $\left(HJ^M\right)$ and the constrained HJ distance $\left(HJ^C\right)$ measures for all the models described in Section 2.2. Panel B presents the results of the two formal tests of model comparison. The null hypothesis in the first test is that our 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 our chosen model.

| | Gross Returns | | | | Excess Returns | | |
|---|---|---|---|---|---|---|---|
| | $HJ^O$ | $HJ^M$ | $HJ^C$ | | $HJ^O$ | $HJ^M$ | $HJ^C$ |
| **Panel A: HJ Measures** | | | | | | | |
| CAPM | 0.5313 | 0.5313 | 0.5343 | CAPM | 0.5234 | 0.6142 | 0.292 |
| FF3 | 0.5099 | 0.5099 | 0.5132 | FF3 | 0.5056 | 0.5852 | 0.281 |
| FF5 | 0.5086 | 0.5086 | 0.512 | FF5 | 0.5003 | 0.5761 | 0.2779 |
| Carhart | 0.4896 | 0.4896 | 0.4956 | Carhart | 0.4752 | 0.5384 | 0.2807 |
| HXZ | 0.5073 | 0.5073 | 0.512 | HXZ | 0.508 | 0.585 | 0.2794 |
| **Panel B: HJ Multiple Comparison** | | | | | | | |
| $Carhart < (HXZ/FF5/FF3/CAPM)$ | | | | $Carhart < (FF5/HXZ/FF3/CAPM)$ | | | |
| $H_0 : (p_{val})$ | 0.1922 | 0.1922 | 0.1644 | $H_0 : (p_{val})$ | 0.1148 | 0.0561 | 0.04982 |
| $Carhart < Candidates$ | | | | | | | |
| CAPM $(p_{val})$ | 0.0598 | 0.0536 | 0.054 | CAPM $(p_{val})$ | 0.0994 | 0.109 | 0.0942 |
| FF3 $(p_{val})$ | 0.0318 | 0.0312 | 0.0268 | FF3 $(p_{val})$ | 0.0378 | 0.0446 | 0.0382 |
| FF5 $(p_{val})$ | 0.0384 | 0.0306 | 0.0306 | FF5 $(p_{val})$ | 0.0322 | 0.037 | 0.0308 |
| HXZ $(p_{val})$ | 0.0594 | 0.0516 | 0.0532 | HXZ $(p_{val})$ | 0.0958 | 0.1026 | 0.1004 |

Table 3: **Test Portfolios: FF 25 Size and Book-to-Market Ratio**

*Notes:* the table represents the results of the model comparison tests using both the gross returns and the excess returns on 25 size and book-to-market sorted portfolios. Panel A reports the HJ distance $\left(HJ^O\right)$, the modified HJ distance $\left(HJ^M\right)$ and the constrained HJ distance $\left(HJ^C\right)$ measures for all the models described in Section 2.2. Panel B presents the results of the two formal tests of model comparison. The null hypothesis in the first test is that our 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 our chosen model.

| | Gross Returns | | | | Excess Returns | | |
|---|---|---|---|---|---|---|---|
| | $HJ^O$ | $HJ^M$ | $HJ^C$ | | $HJ^O$ | $HJ^M$ | $HJ^C$ |
| **Panel A: HJ Measures** | | | | | | | |
| CAPM | 0.3586 | 0.3586 | 0.3598 | CAPM | 0.417 | 0.4588 | 0.1706 |
| FF3 | 0.3209 | 0.3209 | 0.3226 | FF3 | 0.3833 | 0.4147 | 0.1479 |
| FF5 | 0.3195 | 0.3195 | 0.3209 | FF5 | 0.3708 | 0.3989 | 0.1465 |
| Carhart | 0.3068 | 0.3068 | 0.3119 | Carhart | 0.3406 | 0.362 | 0.1476 |
| HXZ | 0.3153 | 0.3153 | 0.3195 | HXZ | 0.3739 | 0.402 | 0.1483 |
| **Panel B: HJ Multiple Comparison** | | | | | | | |
| $Carhart < (HXZ/FF5/FF3/CAPM)$ | | | | $Carhart < (FF5/HXZ/FF3/CAPM)$ | | | |
| $H_0 : (p_{val})$ | 0.0214 | 0.0221 | 0.0191 | $H_0 : (p_{val})$ | 0.013 | 0.057 | 0.048 |
| $Carhart < Candidates$ | | | | | | | |
| CAPM ($p_{val}$) | 0.0454 | 0.0428 | 0.0374 | CAPM ($p_{val}$) | 0.1744 | 0.1694 | 0.184 |
| FF3 ($p_{val}$) | 0.0156 | 0.017 | 0.012 | FF3 ($p_{val}$) | 0.0174 | 0.0214 | 0.0202 |
| FF5 ($p_{val}$) | 0.0158 | 0.0164 | 0.0406 | FF5 ($p_{val}$) | 0.0132 | 0.0142 | 0.0166 |
| HXZ ($p_{val}$) | 0.04 | 0.0382 | 0.0318 | HXZ ($p_{val}$) | 0.1552 | 0.1528 | 0.1642 |

Table 4: **Test Portfolios: FF25 plus 30 Industry Portfolios**

*Notes:* the table represents the results of the consumption-based CAPM models comparison tests using both the gross returns and the excess returns on 25 size and book-to-market sorted plus 30 industry sorted portfolios. Panel A reports the HJ distance $\left(HJ^O\right)$, the modified HJ distance $\left(HJ^M\right)$ and the constrained HJ distance $\left(HJ^C\right)$ measures for all the models described in Section 2.2. Panel B presents the results of the two formal tests of model comparison. The null hypothesis in the first test is that our 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 our chosen model.

| | Gross Returns | | | | Excess Returns | | |
|---|---|---|---|---|---|---|---|
| | $HJ^O$ | $HJ^M$ | $HJ^C$ | | $HJ^O$ | $HJ^M$ | $HJ^C$ |
| **Panel A: HJ Measures** | | | | | | | |
| CCAPM | 0.7877 | 0.7421 | 0.8267 | CCAPM | *0.753* | 0.18 | 0.2235 |
| Yogo | 0.7806 | 0.7806 | 0.8224 | Yogo | *0.7421* | 0.1798 | 0.223 |
| Piazzesi | 0.7853 | 0.7853 | 0.826 | Piazzesi | *0.6925* | 0.1822 | 0.2241 |
| **Panel B: HJ Multiple Comparison** | | | | | | | |
| $Yogo < (CCAPM/Piazeesi)$ | | | | $Yogo < (CCAPM/Piazeesi)$ | | | |
| $H_0 : (p_{val})$ | 0.2776 | 0.1922 | 0.1644 | $H_0 : (p_{val})$ | 0.8895 | 0.9127 | 0.8901 |
| $Yogo < Candidates$ | | | | | | | |
| CCAPM $(p_{val})$ | 0.0332 | 0.0332 | 0.054 | CCAPM $(p_{val})$ | 0.0286 | 0.109 | 0.0942 |
| Piazzesi $(p_{val})$ | 0.0334 | 0.0334 | 0.0268 | Piazzesi $(p_{val})$ | 0.0398 | 0.0446 | 0.0382 |

Table 5: **Test Portfolios: FF25 plus 30 Industry Portfolios**

*Notes:* the table represents the results of the conditional consumption-based asset pricing models comparison tests using both the gross returns and the excess returns on 25 size and book-to-market sorted plus 30 industry sorted portfolios. Panel A reports the HJ distance $\left(HJ^O\right)$, the modified HJ distance $\left(HJ^M\right)$ and the constrained HJ distance $\left(HJ^C\right)$ measures for all the models described in Section 2.2. Panel B presents the results of the two formal tests of model comparison. The null hypothesis in the first test is that our 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 our chosen model.

| | Gross Returns | | | | Excess Returns | | |
|---|---|---|---|---|---|---|---|
| | $HJ^O$ | $HJ^M$ | $HJ^C$ | | $HJ^O$ | $HJ^M$ | $HJ^C$ |
| **Panel A: HJ Measures** | | | | | | | |
| cayC | 0.7743 | 0.7743 | 0.8148 | cayC | *0.744* | 0.1764 | 0.2211 |
| LVNC | 0.7862 | 0.7862 | 0.8258 | LVNC | *0.7509* | 0.1774 | 0.2215 |
| PISC | 0.7862 | 0.7862 | 0.8261 | PISC | *0.6912* | 0.1784 | 0.2233 |
| **Panel B: HJ Multiple Comparison** | | | | | | | |
| $cayC < (LVNC/PISC)$ | | | | $cayC < (LVNC/PISC)$ | | | |
| $H_0 : (p_{val})$ | 0.2574 | 0.21 | 0.6639 | $H_0 : (p_{val})$ | 0.223 | 0.23 | 0.8901 |
| $cayC < Candidates$ | | | | | | | |
| LVNC ($p_{val}$) | 0.0282 | 0.028 | 0.0342 | LVNC ($p_{val}$) | 0.02 | 0.0216 | 0.0168 |
| PISC ($p_{val}$) | 0.0338 | 0.0336 | 0.041 | PISC ($p_{val}$) | 0.024 | 0.0272 | 0.0254 |

Table 6: **Rank Test and Misspecification Identification**

*Notes:* the table presents the results of the ranks tests of the individual factors, the model misspecification tests, and rank tests of the models. The models are estimated using monthly and quarterly gross returns from 1967:12 until 2012:12 on the 25 size and book-to-market Fama–French portfolios, the 30 Fama–French industry portfolios, and the one-month T-bill. 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. In Panel B, I report 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 shows 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 | vw | smb | hml | rmw | cma | mom | me | ia | roe |
| $w$ | 190.8 | 187.6 | 178.1 | 143.1 | 160.6 | 81.8 | 179.2 | 152.0 | 126.8 |
| p-val | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.011 | 0.000 | 0.000 | 0.000 |
| Test | $c^{ndur}$ | | $c^{dur}$ | | $c^{nh}(s)$ | | $cay$ | | $my$ |
| $w$ | 65.7 | | 61.9 | | 67.2/43.0 | | 61.7 | | 56.8 |
| p-val | 0.153 | | 0.245 | | 0.124/0.88 | | 0.249 | | 0.407 |
| **Panel B: HJ-distance, Lagrange multiplier, and rank tests** | | | | | | | | | |
| Model | | $\hat{\delta}$ | p-val | | LM | p-val | | $W^*$ | p-val |
| CAPM | | 0.611 | 0.000 | | 152.707 | 0.000 | | 190.8 | 0.000 |
| FF3 | | 0.580 | 0.000 | | 137.606 | 0.000 | | 186.2 | 0.000 |
| FF5 | | 0.568 | 0.000 | | 133.141 | 0.000 | | 75.3 | 0.015 |
| Carhart | | 0.534 | 0.000 | | 106.155 | 0.000 | | 82.7 | 0.004 |
| HXZ | | 0.573 | 0.000 | | 135.194 | 0.000 | | 73.7 | 0.026 |
| **Panel C: Selection by standard errors under correct model specification** | | | | | | | | | |
| Model | vw | smb | hml | rmw | cma | mom | me | ia | roe |
| CAPM | -2.39 | | | | | | | | |
| FF3 | **-3.09** | -0.65 | **-4.18** | | | | | | |
| | **-3.37** | | **-4.20** | | | | | | |
| FF5 | **-3.65** | -1.02 | -0.75 | -1.67 | -1.41 | | | | |
| | **-3.45** | -1.05 | **-4.40** | -2.08 | | | | | |
| | **-3.50** | -0.69 | -0.34 | | -1.82 | | | | |
| Carhart | **-3.54** | -1.47 | **-4.37** | | | **-3.44** | | | |
| | **-2.91** | | | | | **-2.81** | | | |
| HXZ | **-3.48** | | | | | | **-2.55** | **-3.60** | -2.47 |
| | **-3.22** | | | | | | -1.61 | **-3.56** | |
| | **-3.85** | | | | | | | **-3.32** | -1.65 |
| **Panel D: Selection by model misspecification-robust standard errors** | | | | | | | | | |
| Model | vw | smb | hml | rmw | cma | mom | me | ia | roe |
| CAPM | -2.38 | | | | | | | | |
| FF3 | **-3.03** | -0.61 | **-4.00** | | | | | | |
| | **-3.35** | | **-4.03** | | | | | | |
| FF5 | **-3.49** | -0.80 | -0.47 | -1.42 | -0.91 | | | | |
| | **-3.38** | -0.85 | **-4.10** | -1.79 | | | | | |
| | **-3.25** | -0.63 | -0.23 | | -1.22 | | | | |
| Carhart | **-3.44** | -1.36 | **-4.15** | | | **-3.44** | | | |
| | **-2.91** | | | | | **-2.70** | | | |
| HXZ | **-3.44** | | | | | | -2.43 | **-3.39** | -2.23 |
| | **-3.16** | | | | | | -1.59 | **-3.27** | |
| | **-3.80** | | | | | | | **-3.09** | -1.50 |

Figure 1: **Confidence Set for the HJ Set for Gross Returns on Fama–French 25 Size and Value Portfolios**

*Notes:* Figure shows Hansen-Jagannathan (HJ) sets using weighted LR statistic and the structured projection approaches when the consumption-based models explain the gross returns on Fama-French 25 size and book-to-market ratio portfolios. The first row gives the CCAPM, the second row is the Piazzesi, Schneider and Tuzel (2007) housing consumption model and the third row shows the Yogo (2006) durable consumption model. The blue dashed line in the left is the 95% confidence region based on the weighted LR statistic, and the blue dashed in the right denotes the 95% confidence region based on the structured projection. Both approaches' confidence regions cover most of the bootstrap draws below the HJ bounds.

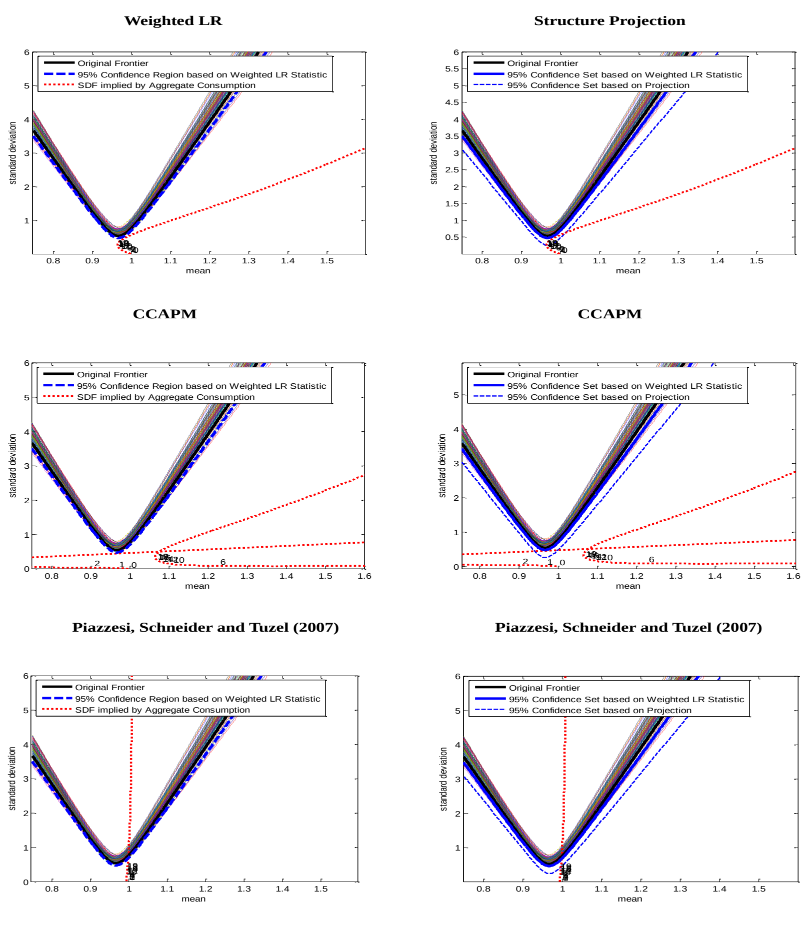


Figure 2: **Confidence Set for the HJ Set for Excess Returns on Fama-French 25 Size and Value Portfolios**

*Notes:* Figure shows Hansen-Jagannathan (HJ) sets using weighted LR statistic and the structured projection approaches when the consumption-based models explain the excess returns on Fama-French 25 size and book-to-market ratio portfolios. The first row gives the CCAPM, the second row is the Piazzesi, Schneider and Tuzel (2007) housing consumption model and the third row shows the Yogo (2006) durable consumption model. The blue dashed line in the left is the 95% confidence region based on the weighted LR statistic, and the blue dashed in the right denotes the 95% confidence region based on the structured projection. Both approaches' confidence regions cover most of the bootstrap draws below the HJ bounds.

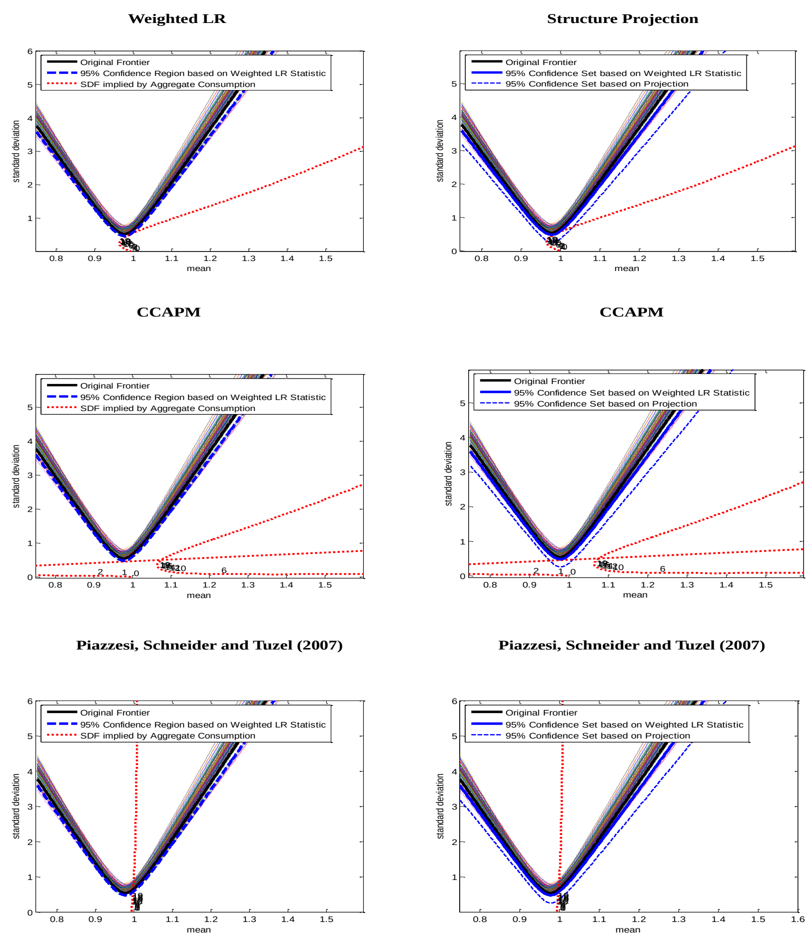


Figure 3: **Confidence Set for the HJ Set for Gross Returns on Industry Portfolios**

*Notes:* Figure shows Hansen-Jagannathan (HJ) sets using weighted LR statistic and the structured projection approaches when the conditional consumption-based models explain the gross returns on 30 industry portfolios. The first row gives the Lettau and Ludvigson (2001) scaled CCAPM, the second row is the Piazzesi, Schneider and Tuzel (2007) housing consumption conditional CCAPM model and the third row shows the Lustig and Van Nieuwerburgh (2005) housing collateral conditional CCAPM. The blue dashed line in the left is the 95% confidence region based on the weighted LR statistic, and the blue dashed in the right denotes the 95% confidence region based on the structured projection. Both approaches' confidence regions cover most of the bootstrap draws below the HJ bounds.

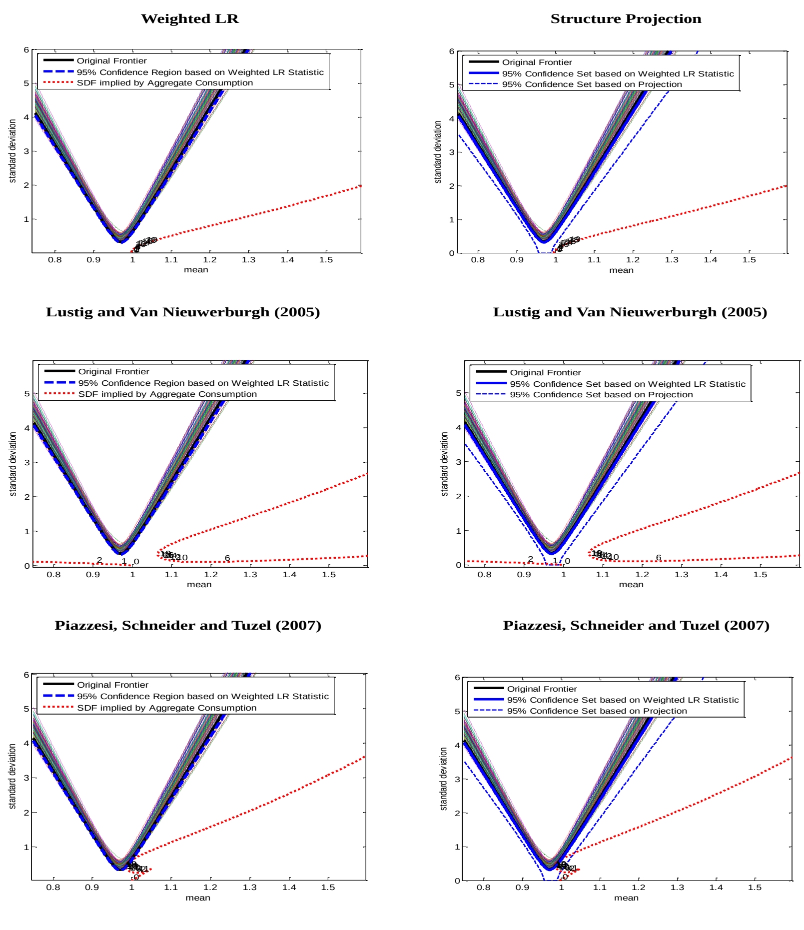
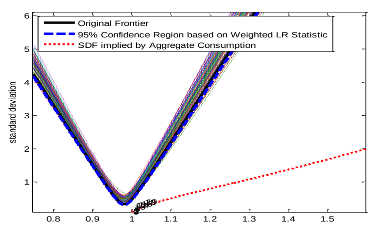
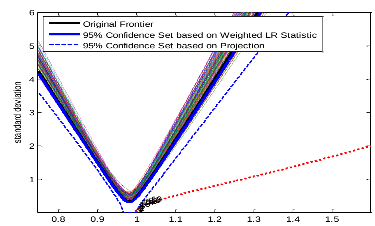


Figure 4: **Confidence Set for the HJ Set for Excess Returns on Industry Portfolios**

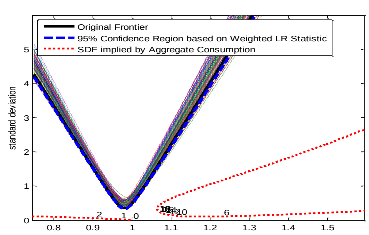
*Notes:* Figure shows Hansen-Jagannathan (HJ) sets using weighted LR statistic and the structured projection approaches when the conditional consumption-based models explain the excess returns on 30 industry portfolios. The first row gives the Lettau and Ludvigson (2001) scaled CCAPM, the second row is the Piazzesi, Schneider and Tuzel (2007) housing consumption conditional CCAPM model and the third row shows the Lustig and Van Nieuwerburgh (2005) housing collateral conditional CCAPM. The blue dashed line in the left is the 95% confidence region based on the weighted LR statistic, and the blue dashed in the right denotes the 95% confidence region based on the structured projection. Both approaches' confidence regions cover most of the bootstrap draws below the HJ bounds.

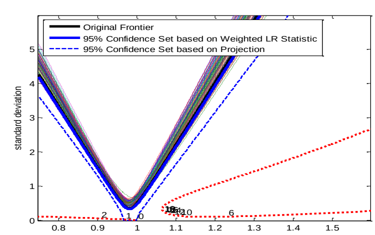
**Weighted LR** **Structure Projection**



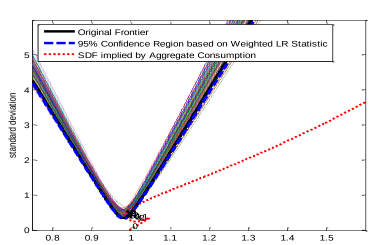



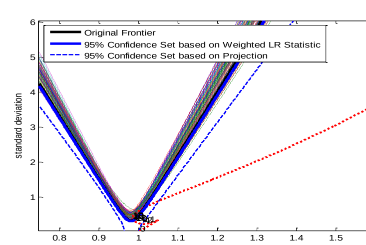

**Lustig and Van Nieuwerburgh (2005)** **Lustig and Van Nieuwerburgh (2005)**




**Piazzesi, Schneider and Tuzel (2007)** **Piazzesi, Schneider and Tuzel (2007)**




**Lettau and Ludvigson (2001)** **Lettau and Ludvigson (2001)**

Figure 5: **Yogo and CCAPM in the Entropy Scenario**

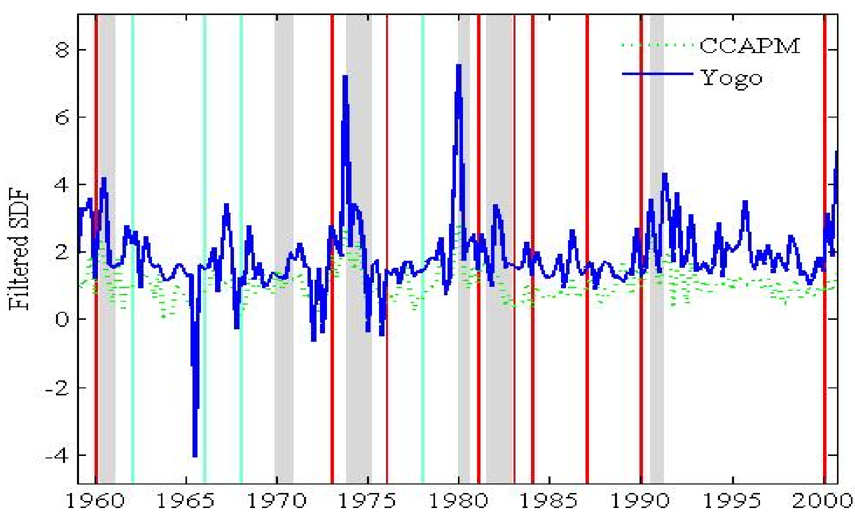


Figure 6: **Lettau and Ludvigson and CCAPM in the Entropy Scenario**