Nonparametric Specification Testing of Conditional Asset Pricing Models

Francisco Peñaranda, Juan M. Rodríguez-Poo, and Stefan Sperlich

Nonparametric Specification Testing of Conditional Asset Pricing Models

ABSTRACT

This article presents an adaptive omnibus specification test of asset pricing models where the stochastic discount factor is conditionally affine in the pricing factors. These models provide constraints that conditional moments of returns and pricing factors must satisfy, but most of them do not provide information on the functional form of those moments. Our test is robust to functional form misspecification, and also detects any relationship between pricing errors and conditioning variables. We give special emphasis to the test implementation and calibration, and extensive simulation studies prove the functioning in practice. Our empirical applications show a conditional counterpart of a well-known problem of unconditional models. The lack of rejection of consumption based conditional models seems to be due to a poor conditional correlation between consumption and stock returns.

1. Introduction

Asset pricing theory must consider that asset returns and pricing factors are predictable, in the sense of a significant time-variation in their joint conditional distribution. For this purpose, conditional asset pricing models provide constraints that conditional moments of returns and pricing factors must satisfy. However, most of these models do not provide much information on the functional form of those conditional moments. This article presents an adaptive omnibus specification test for conditional asset pricing models based on kernel smoothing. In this way, our test is robust to functional form misspecification of the conditional moments and prices of risk, and has power against any relationship between pricing errors and conditioning variables.

There is a vast literature that uses parametric methods to evaluate conditional asset pricing models, and concludes that conditioning information substantially improves their empirical performance. For instance, Jagannathan and Wang (1996) found that the conditional capital asset pricing model (CAPM) can explain the cross section of stock returns even though the static CAPM cannot, while Lettau and Ludvigson (2001b) found that the value premium can be explained by a consumption CAPM (CCAPM) with a time-varying price of risk. Other authors suspected that this may be an illusion caused by low statistical power of standard asset pricing tests, see, for example, Lewellen and Nagel (2006).

In this line, there have been significant contributions that avoid specifying the conditional distribution of returns and pricing factors by using semi-parametric techniques. Nagel and Singleton (2011) estimated nonparametrically the first and second conditional moments, but work with asset pricing models where the prices of risk are parametric. Local GMM methods have been used in empirical asset pricing by Gagliardini, Gourieroux, and Renault (2011), Antoine, Proulx, and Renault (2018), and Gagliardini and Ronchetti (2020). Nonparametric prices of risk have been considered by Wang (2003) and Roussanov (2014) using kernel methods, and their work can be interpreted in terms of two stages. In the first stage, they estimate the first and second moments of returns and factors by kernel methods, and use those moments to estimate the prices of risk. However, in the second stage, they use parametric methods to test the asset pricing model. We start from their estimators of prices of risk, but we develop a formal nonparametric test of the model instead.

In particular, Wang (2003) used nonparametric methods to evaluate conditional variants of the CAPM and the Fama-French (FF) model. He uses the stochastic discount factor (SDF) approach to first estimate the prices of risk nonparametrically, and then test if pricing errors are linear in the conditioning variables. This is easy to implement but has zero power against pricing errors that exhibit nonlinear dependence on conditioning variables. We introduce a test that is general enough to detect nonparametric alternatives from independence. Roussanov (2014) estimated nonparametric prices of risk by minimizing a quadratic form of conditional pricing errors, which nests the prices of risk in Wang (2003), with a focus on linearized consumption based models. However, he does not develop a formal test of the asset pricing model. Instead, he tests a particular pricing error has a zero average, and displays figures of pricing errors with point-wise bootstrap bands. We derive a formal test and develop the asymptotic theory for it. Moreover, our test considers jointly all the pricing errors, and does not focus on unconditional means. Cai, Ren, and Sun (2015) developed the asymptotic
distribution of a locally linear estimator of the market price of risk in the CAPM, but they do not provide a formal test of this model.

In econometrics and statistics, there exists a plenitude of specification tests to test parametric models against a broad set of alternatives. Mainly two fundamental approaches have been considered: statistics that appear as a weighted average of the residuals, and statistics that compare parametric and nonparametric fits (see Hart 1997). We follow the second alternative, along the lines of Härdle and Mammen (1993). Their test is based on an integrated squared difference between the two estimated curves, where the parametric curve is also smoothed. They show that standard bootstrap does not work for this test, whereas wild bootstrap does.

Our test is not a trivial application of this test to asset pricing models. Our test allows richer data dynamics, but more importantly our null hypothesis is nonparametric because a complex structural semiparametric model explains the conditional risk premia of excess returns. So the null contains an asset pricing model in which both the risks, or factor sensitivities, and the prices of risk are nonparametric functions of some given observables requiring several non- and semiparametric estimation steps. The general problem of nonparametrically testing against structural non- or semiparametric alternatives was studied, among others, by Rodríguez-Poo, Sperlich, and Vieu (2015). While our null hypothesis and estimators are much more complex, we can partly follow their construction of a test statistic and its calibration to achieve reasonable power while holding the level, even for small samples.

We study two empirical applications. In the first one, we work with traded factors and monthly data. Both the conditional CAPM and the FF model are rejected with size and book-to-market sorted portfolios, even though the latter model yields lower pricing errors. In the second application, we work with a mixture of traded and non-traded factors, and quarterly data. We find strong empirical evidence against the conditional CAPM. However, we do not find evidence against zero pricing errors when we test a linearized CCAPM with consumption growth as the only pricing factor. In fact, in the case of a CCAPM with Epstein-Zin preferences, where both the market portfolio and consumption growth are considered, the market price of risk seems to be zero.

Importantly, low pricing errors with the consumption factor do not necessarily mean that this is an economically meaningful pricing factor. The lack of rejection seems to be due to a poor conditional correlation between consumption and the cross section of stock returns. We find a conditional counterpart of a well-known problem with unconditional asset pricing models. Kan and Zhang (1999) argued that some empirical models rely on useless factors with a zero unconditional covariance with returns. Recently, the empirical relevance of this problematic case has been emphasized by Peñaranda and Sentana (2015), Burnside (2016), Bryzgalova (2016), Gospodinov, Kan, and Robotti (2017), and Kleibergen and Zhan (2020).

We focus our analysis on asset pricing models whose SDF is conditionally affine in the pricing factors, but we are flexible in the conditional moments of returns and factors, and the prices of risk of such an SDF. There are earlier examples of different flexible SDF methods in empirical asset pricing. Bansal and Viswanathan (1993) did not start from a particular asset pricing model, instead they only impose that the SDF depends on some variables, and are flexible in that dependence. Other flexible approaches are Ait-Sahalia and Lo (1998) and Rosenberg and Engle (2002), who focused on option pricing and the corresponding state-price density or SDF projection, respectively, as a function of the underlying asset’s payoffs.

The rest of the article is organized as follows. Section 2 reviews the estimation of conditional asset pricing models. Section 3 develops our test, and the corresponding asymptotic behavior. Section 4 describes the implementation and calibration of the test, jointly with a Monte Carlo exercise, while Section 5 reports our applications. Finally, Section 6 concludes. Proofs and auxiliary results are deferred to appendices that are available online.

2. Conditional Asset Pricing Models

This section describes the nonparametric estimation of prices of risk along the lines of Wang (2003) and Roussanov (2014), who represent a conditional asset pricing model by means of its SDF. The simplicity of the SDF to present a general theory of asset pricing is well recognized, see Cochrane (2005).

2.1. Prices of Risk

The investment set is given by $N$ assets, with a vector of excess returns $r_{t+1}$ that is known at time $t+1$. The information set known at $t$ is given by a $d \times 1$ vector $z_t$ of conditioning variables that measure the state of the economy. For instance, the excess returns could be associated with stock portfolios, and a conditioning variable could be the default spread, which is high during recessions. Standard arguments such as the lack of arbitrage opportunities, or the first-order conditions of a representative investor, imply the pricing conditions

$$E(m_{t+1} | r_{t+1} = 0, r_{t+1} = (r_{1,t+1}, \ldots, r_{N,t+1})' \quad (1)$$

for some random variable $m_{t+1}$ called SDF, which discounts uncertain pay-offs in such a way that their expected discounted value equals their cost. Appendix E introduces a riskless asset in the investment set, and describes its SDF implications.

The usual approach in empirical finance is to model the SDF as an affine transformation of some $K < N$ observable risk factors $f_{t+1}$ such as the return on the market portfolio or aggregate consumption growth. There are two common variants to express that SDF, the uncentered SDF

$$m_{t+1} = 1 - f_{t+1}' \delta(z_t), \quad f_{t+1} = (f_{1,t+1}, \ldots, f_{K,t+1})' \quad (2)$$

and the centered SDF $m_{t+1} = 1 - (f_{t+1} - v(z_t))' \tau(z_t)$, where $v(z_t) = E(f_{t+1} | z_t)$. For instance, Wang (2003) uses the former variant, while Roussanov (2014) used the latter variant. In what follows we focus on the uncentered SDF as there are relevant situations for which this variant exists but the centered SDF does not, as we find in our empirical application. Appendix C develops the centered SDF and the relationship between the two variants.
The conditional extensions of the CAPM and the FF model are well-known examples of conditionally affine SDFs, see Wang (2003) and the references therein. In the former model, the excess return on the market portfolio is the only pricing factor, while the latter model has two additional factors associated with portfolios that capture size and value effects, denoted SMB (long/short in small/large capitalization stocks) and HML (long/short in high/low book-to-market stocks), respectively. The two commented models rely on traded factors only, and hence the SDF pricing conditions follow directly from the conditional mean–variance efficiency of a portfolio of the factors. But we can also use similar pricing conditions with models that include non-traded factors, such as linear approximations to the CCAPM and its extensions, which are often used in empirical studies, see Roussanov (2014) and the references therein. The linearized version of the canonical CCAPM with power utility has a single pricing factor, the aggregate consumption growth, and the price of risk is constant and equal to the relative risk aversion. Lettau and Ludvigson (2001b) motivated a time-varying price of risk by means of a linearized version of the habit model of Campbell and Cochrane (1999). The CCAPM with Epstein-Zin preferences identifies both the market return and consumption growth as pricing factors.

The vectors $\delta (z_t)$ and $\tau (z_t)$ represent the prices of risk, that is, the risk premium that is obtained per unit of risk, where risk is measured as sensitivity with respect to the factors. In the simple setting of a single-factor SDF in Equation (2), we can rewrite the pricing conditions (1) as all excess returns satisfying

$$
\frac{E(r_{t+1} | z_t)}{E(f_{t+1} | z_t)} = \delta (z_t), \quad i = 1, \ldots, N.
$$

If the factor is traded, then the factor $f_{t+1}$ itself should satisfy the pricing conditions when the model holds, and we can interpret the price of risk as

$$
\delta (z_t) = \frac{E(f_{t+1} | z_t)}{E(f_{t+1}^2 | z_t)},
$$

where the factor’s risk is its second moment. If the factor is non-traded, we still have a similar interpretation of risk premium per unit of risk when the asset pricing model holds, but in terms of the corresponding factor mimicking portfolio (the projection of the factor on the returns)

$$
\delta (z_t) = \frac{E(f_{t+1}^2 | z_t)}{E(f_{t+1}^2 | z_t)},
$$

$$
f_{t+1}^* = E(f_{t+1} | z_t) \left[ E(r_{t+1} f_{t+1} | z_t) \right]^{-1} r_{t+1}.
$$

In empirical work, the prices of risk are usually chosen to minimize some metric of pricing errors. Let us denote the vector of risk premia and, respectively, the risks or factor sensitivities

$$
\mu (z_t) = E(r_{t+1} | z_t) \quad \text{and} \quad \sigma (z_t) = E(r_{t+1} f_{t+1} | z_t).
$$

Given some prices of risk $\delta (z_t)$, the model pricing errors are

$$
e (z_t | \delta (z_t)) = E \left[ \left( 1 - f_{t+1}^* \delta (z_t) \right) r_{t+1} | z_t \right] = \mu (z_t) - \sigma (z_t) \delta (z_t)
$$

and the particular vector $\delta (z_t)$ can be chosen by minimizing the quadratic form

$$
q (z_t | \delta (z_t)) = e (z_t | \delta (z_t))' W (z_t) e (z_t | \delta (z_t))
$$

for some weighting matrix $W (z_t)$ that may depend on $z_t$, but not on $\delta (z_t)$. To keep our notation simple, we also denote $\delta (z_t)$ the particular prices of risk that minimize the criterion $q (z_t | \delta (z_t))$. The first-order conditions that pin down these prices of risk are

$$
D (z_t)' W (z_t) e (z_t | \delta (z_t)) = 0,
$$

which are equivalent to the exact pricing of the portfolios with excess returns

$$
D (z_t)' W (z_t) r_{t+1},
$$

and we obtain

$$
\delta (z_t) = \left[ D (z_t)' W (z_t) D (z_t) \right]^{-1} D (z_t)' W (z_t) \mu (z_t).
$$

The chosen metric of pricing errors is irrelevant under the null hypothesis, when there is a vector of risk prices that makes the pricing errors equal to zero. However, under the alternative hypothesis different weighting matrices pin down different prices of risk. Two widely used choices of $W (z_t)$ are the identity matrix, which makes the criterion (7) equal to the sum of squared pricing errors, and

$$
W (z_t) = \left[ E \left( r_{t+1} r_{t+1}' | z_t \right) \right]^{-1},
$$

which makes the criterion (7) equal to the conditional counterpart of the squared Hansen and Jagannathan (1997) distance (HJD). These authors show that the mean-square distance between the SDF implied by a model and the set of SDFs with zero pricing errors is equal to a quadratic form in the model pricing errors, with the commented weighting matrix. The HJD is widely used in empirical finance, see Gagliardini and Ronchetti (2020) and the references therein. On the other hand, in the simpler setting of a parametric SDF, we could think of optimal IV like Nagel and Singleton (2011), and weight the pricing conditions with the inverse of $E (m_{t+1}^2 r_{t+1} r_{t+1}' | z_t)$ evaluated at the true SDF $m_{t+1}$ to optimize the large-sample properties of the estimators.

The HJD was developed to compare competing models, but the choice of weighting matrix (10) is also useful to interpret the prices of risk. This choice makes the first-order conditions (8) equivalent to the exact pricing of the factor mimicking portfolios

$$
f_{t+1}^* = E \left( f_{t+1} r_{t+1}' | z_t \right) \left[ E \left( r_{t+1} r_{t+1}' | z_t \right) \right]^{-1} r_{t+1},
$$

and the corresponding prices of risk can be written as

$$
\delta (z_t) = \left[ E \left( f_{t+1} r_{t+1}' | z_t \right) \right]^{-1} E \left( r_{t+1}' | z_t \right).
$$

That is, we obtain prices of risk like in Equation (4), even if the asset pricing model does not hold. In fact, the choice of prices of risk in Wang (2003) with traded factors

$$
\delta (z_t) = \left[ E \left( f_{t+1} r_{t+1}' | z_t \right) \right]^{-1} E \left( f_{t+1}' | z_t \right)
$$

is equivalent to the choice of the weighting matrix (10), even though he follows mean–variance arguments without an explicit weighting matrix. In the case of traded factors, the pricing conditions (1) should also include $f_{t+1}$, pricing the vector $x_{t+1} = (r_{t+1}, f_{t+1})'$ instead of only $r_{t+1}$. We can show that, if we use

$$
W (z_t) = \left[ E \left( x_{t+1} x_{t+1}' | z_t \right) \right]^{-1}
$$

to weight the pricing errors of $x_{t+1}$, then the corresponding first-order conditions (8) yield Wang’s prices of risk.
2.2. SDF Estimation

A fully parametric prespecified SDF would inherit all problems of the various potential misspecifications. We relax such a framework, which actually requires a series of different nonparametric (pre-)estimates that entail the typical issues like bandwidth choice, boundary problems, or bias reduction. Practically, we have to estimate $\mu(z)$ in Equation (5), and $D(z)$ in Equation (5), and maybe also $W(z)$, to estimate the prices of risk $\delta(z)$ in Equation (9).

Our estimation strategy is the following. We estimate the standard regression problem $\mu(z)$ in Equation (5) by local linear kernel smoothing with bandwidth $h_{1}$. In this way, we nest linear regression models also for finite samples, see Fan and Gijbels (1995). We denote $\mu_{h_{1}}(z)$ the corresponding estimator. For the second-order moments $D(z)$ in Equation (5), and potentially $W(z)$ in Equation (10), we follow Yin et al. (2010) using a local constant kernel regression with bandwidth $h_{2}$. We denote $D_{h_{2}}(z)$ and $W_{h_{2}}(z)$ the corresponding estimators. In particular, for a kernel function $K(\cdot)$, all proposed estimators are the respective minimizer $\hat{b}_{0}$ of

$$\min_{b_{0},b_{1}} \sum_{t=1}^{T} \left\{ y_{t+1} - b_{0} - b_{1}(z_{t}-z_{i}) \right\}^{2} K \left( \frac{z-z_{i}}{h} \right).$$ (11)

For instance, for $y_{t+1} = r_{t+1}$ and $h_{1} = h_{2}$, $b_{0}$ gives the first element of the vector estimate $\mu_{h}(z)$. For making $b_{0}$ a Nadaraya-Watson estimate of matrix element $D_{i j}(z)$, set $b_{1} = 0$, $h_{1} = h_{2}$, and $y_{t+1} = r_{t+1}$. Analogously, we can obtain an estimate for $E \left(r_{t+1}f_{i t+1} \mid z_{i} \right)$. The estimators and bandwidths have to fulfill conditions (C.1), (C.2), (A.6), (A.7), and those mentioned in Theorem 1 of Section 3.2. To satisfy the bandwidth conditions, higher dimensional $z$ may require higher order kernels $K(\cdot)$. In our simulations and applications, $d$ is 1 so that we can take the Quartic kernel. Bandwidths $h_{1}$ and $h_{2}$ should be chosen to calibrate the test, and therefore unlike in Wang (2003) or Roussanov (2014). See Section 4 for the corresponding details of our data adaptive bandwidths.

Given the previous estimators, the estimator of the prices of risk (9) is

$$\hat{\delta}(z_{t}) = \left[ D_{h_{2}}(z_{t}^{'}) W_{h_{2}}(z_{t}) D_{h_{2}}(z_{t}) \right]^{-1} D_{h_{2}}(z_{t}^{'}) W_{h_{2}}(z_{t}) \mu_{h_{1}}(z_{t}) .$$ (12)

This estimator depends on bandwidths ($h_{1}, h_{2}$) through the estimator of the first moments $\mu_{h_{1}}(z_{i})$ and the second moments $D_{h_{2}}(z_{i})$ of the matrix $W_{h_{2}}(z_{t})$ may not require a bandwidth as it could be a known matrix, or even a constant matrix that does not depend on $z_{i}$.

These prices of risk can be interpreted as a two-pass procedure where first the conditional moments $\mu(z_{i})$ and $D(z_{i})$ are estimated from time series data, and second the prices of risk are obtained from cross-sectional data, with a weighted least-squares projection of $\mu_{h_{1}}(z_{i})$ on $D_{h_{2}}(z_{i})$. This interpretation is similar in spirit to Fama-MacBeth regressions, which are an alternative to SDF methods which may also allow for time-variation in risk and risk premia, see Cochrane (2005) for instance.

3. Omnibus Specification Test

We introduce a test for conditional asset pricing models, addressing the two-dimensional nonparametric nature of such a test: the SDF that represents the model, and the pricing conditions that represent the null hypothesis.

3.1. Test Statistic

The main problem in nonparametric testing is to (i) estimate the distribution of the test statistic under the null, and to (ii) calibrate this distribution estimate for a given significance level to guarantee that the practitioner can control the error of the first type, that is, the rejection rate when the null hypothesis is correct. Finally, it has to be checked that (iii) this calibrated version exhibits some power against all interesting alternatives. Out of a huge set of theoretically valid test statistics we have to find one for which we can provide solutions to these points (i) to (iii).

From the pricing conditions (1) and an uncentered SDF (2), the null hypothesis of a correctly specified asset pricing model is

$$H_{0} : \mu(z) = D(z) \delta(z)$$ (13)

for all $z \in \mathcal{Z}$, or equivalently

$$H_{0} : e(z; \delta(z)) = 0$$ (14)

by Equation (6). Therefore, our statistic must check whether the pricing errors are significantly different from zero, or any other given constant. This can be done by a statistic that integrates the quadratic form Equation (7)

$$S = \int e(z; \delta(z))^{'} W(z) e(z; \delta(z)) \pi_{1}(z) dz ,$$ (15)

where $\pi_{1}(z)$ is a given weight function, typically used for trimming. Under the null hypothesis, $S$ is small, while it is large under the alternative. Note that, while it is theoretically possible to choose a weighting matrix $W(z)$ in the test statistic (15) that is different from that in the criterion function (7) that pins down the prices of risk, in practice we recommend the same choice. The reason is that $W(z)$ defines the metric of pricing errors that we are interested in, and this metric should drive both the choice of prices of risk and the test of the model. In other words, we should compare the null model to a nonparametric alternative under the same metric as the one defining the null model estimate. Note further that the standard arguments from parametric inference for efficient estimation, and those for standardization to obtain chi-squared tests, do not directly apply to the analysis of nonparametric functions. Therefore, we favor the choice of $W(z)$ in terms of the interpretation of the corresponding metric of pricing errors metric and the implied prices of risk. For instance, by choosing Equation (10), our statistic integrates the squared conditional HJD, and the prices of risk are given by the pricing of the factor mimicking portfolios.

To obtain a sample analogue of Equation (15), we start from the standard approach used in specification testing for regression functions. We look at the residuals under the null model

$$u_{t+1} = r_{t+1} - \mu(z_{t}) + e(z_{t}; \delta(z_{t})) = r_{t+1} - D(z_{t}) \delta(z_{t}).$$ (16)
Note that \( E (\mathbf{u}_{t+1} \mid \mathbf{z}_t) = \mathbf{e} (\mathbf{z}_t; \delta (\mathbf{z}_t)) = 0 \) under the null. Let us define the residuals

\[
\hat{\mathbf{u}}_{t+1} = \mathbf{r}_{t+1} - \mathbf{D}_{h_2} (\mathbf{z}_t) \hat{\delta} (\mathbf{z}_t) , \tag{17}
\]

where \( \hat{\delta} (\mathbf{z}_t) \) is the nonparametric estimator (12) of the prices of risk. These residuals depend on bandwidth \( h_2 \) through \( \mathbf{D}_{h_2} (\mathbf{z}_t) \), while \( \hat{\delta} (\mathbf{z}_t) \) adds the bandwidth \( h_1 \).

Then, with \( K_{h_2} (\mathbf{v}) = h^{-d} K (h^{-1} \mathbf{v}) \), \( K \cdot (\cdot) \) being a standard \( d \)-variate kernel function and \( h \) a bandwidth, we obtain a sample counterpart of (15)

\[
\hat{S}_h = \int \frac{1}{T^2} \left[ \sum_t K_h (\mathbf{z} - \mathbf{z}_t) \hat{\mathbf{u}}_{t+1} \right] \mathbf{W}_{h_2} (\mathbf{z}_t) \times \left[ \sum_t K_h (\mathbf{z} - \mathbf{z}_t) \hat{\mathbf{u}}_{t+1} \right] \pi_2 (\mathbf{z}) d\mathbf{z}, \tag{18}
\]

in which \( \pi_2 (\mathbf{z}) \) corresponds to \( \pi_2 (\mathbf{z}) p^2 (\mathbf{z}) \). Once again, the matrix \( \mathbf{W}_{h_2} (\mathbf{z}_t) \) may not require a bandwidth. This test statistic can be interpreted as a standard approach in nonparametric specification testing, based on the difference between a direct estimator of \( \mu (\mathbf{z}) \) and an estimator of the null \( \mathbf{D} (\mathbf{z}) \delta (\mathbf{z}) \) convoluted with the same smoother as the raw output \( \mathbf{r}_{t+1} \). Note that this statistic considers \( \hat{\mathbf{u}}_{t+1} \) only at the values of \( \mathbf{z}_t \) within the support of \( p (\cdot) \), and pays substantially less attention to areas where the data are sparse. This not only simplifies the theoretical derivations, but also makes the statistic usable in practice.

The statistic \( \hat{S}_h \) depends directly on \( h \), but also indirectly on \( (h_1, h_2) \) through \( \hat{\mathbf{u}}_{t+1} \). As explained in Spéralich (2014), \( (h_1, h_2) \) have to be chosen in a way that the test is calibrated, that is, does not over- (or too much under-) reject under the null. We maximize the power along \( h \) by looking at

\[
\widehat{S} = \max_h \frac{\hat{S}_h - E [\hat{S}_h \mid H_0]}{\sqrt{V [\hat{S}_h \mid H_0]}} \tag{19}
\]

which requires the estimation of \( E [\hat{S}_h \mid H_0] \) and \( V [\hat{S}_h \mid H_0] \) for any \( h \) and given \( (h_1, h_2) \). As bootstrap is used in Section 4 to approximate the distribution of our statistic under \( H_0 \), the resulting bootstrap samples can also be used for estimating these quantities. In sum we have a test statistic which is quite attractive with a natural interpretation, an automatic procedure for maximizing power, and two calibration parameters, \( (h_1, h_2) \). We review previous approaches to testing asset pricing models with nonparametric prices of risk in Appendix B.

### 3.2. Asymptotic Behavior

In this section, we show the asymptotic normal distribution of the (recentered and scaled) test statistic under the null hypothesis, as well as under the local alternatives to the null hypothesis for which the test is consistent.

The level and power of our test depend on the properties of \( \mu_{h_1} (\mathbf{z}) \) and \( \mathbf{D}_{h_2} (\mathbf{z}) \), which appear in the test through the prices of risk \( \hat{\delta} (\mathbf{z}) \) defined in Equation (12). Therefore, we first analyze the properties of this estimator, and afterwards the asymptotic properties of the test. To do this, we introduce two different sets of assumptions. The first one is introduced to obtain the asymptotic properties of \( \hat{\delta} (\mathbf{z}) \), whereas the second set of assumptions is introduced to derive the asymptotic properties of our test statistic. Specifically, we assume the following:

(C.1) Let \( \mathcal{Z} \) be a compact subset of \( \mathbb{R}^d \) and let \( \mu_{h_1} (\mathbf{z}) \) be an estimator of \( \mu (\mathbf{z}) \) at observation points \( \mathbf{z} \) such that \( p (\mathbf{z}) > 0 \), then

\[
\sup_{\mathbf{z} \in \mathcal{Z}} |\mu_{h_1} (\mathbf{z}) - \mu (\mathbf{z})| = O_p \left( \frac{\log T}{Th_{h_1}} \right) + O_p (h_{h_1}^{\alpha_1}) ,
\]

for some \( \alpha_1 > 0 \), as \( T \) tends to infinity.

(C.2) Let \( \mathcal{Z} \) be a compact subset of \( \mathbb{R}^d \) and let \( \mathbf{D}_{h_2} (\mathbf{z}) \) be an estimator of \( \mathbf{D} (\mathbf{z}) \) at observation points \( \mathbf{z} \) such that \( p (\mathbf{z}) > 0 \), then

\[
\sup_{\mathbf{z} \in \mathcal{Z}} |\mathbf{D}_{h_2} (\mathbf{z}) - \mathbf{D} (\mathbf{z})| = O_p \left( \frac{\log T}{Th_{h_2}} \right) + O_p (h_{h_2}^{\alpha_2}) ,
\]

for some \( \alpha_2 > 0 \), as \( T \) tends to infinity.

The assumptions above give uniform asymptotic bounds for \( \mu_{h_1} (\mathbf{z}) \) and \( \mathbf{D}_{h_2} (\mathbf{z}) \). They are fulfilled for many of the existing nonparametric estimators of \( \mu_{h_1} (\mathbf{z}) \) and \( \mathbf{D}_{h_2} (\mathbf{z}) \), including our proposals in Section 2.2. As it is well known, these bounds depend on the bandwidth (either \( h_1 \) or \( h_2 \)), the dimensionality \( d \), and a parameter \( (\alpha_1 \) or \( \alpha_2) \) that is related to the smoothness of \( \mu (\mathbf{z}) \) and \( \mathbf{D} (\mathbf{z}) \), respectively. Specifically, it is implicitly assumed that the \( \alpha_1 \)th derivatives of \( \mu \) are Lipschitz continuous, and analogously the \( \alpha_2 \)th derivatives for \( \mathbf{D} \). In many articles dealing with nonparametric estimators, conditions (C.1) and (C.2) are shown to hold for estimators that rely on rather primitive conditions. For example, for a local linear estimator of \( \mu (\mathbf{z}) \), Masry (1996) showed under some mild regularity conditions a similar bound as the one in (C.1). In the case of a local constant estimator for \( \mathbf{D} (\mathbf{z}) \), Yin et al. (2010) show a similar bound as the one in (C.2). In sum, we have introduced (C.1) and (C.2) for the sake of generality, that is, to manifest that our test is valid for a broad range of nonparametric estimators of \( \mu (\mathbf{z}) \) and \( \mathbf{D} (\mathbf{z}) \). The next conditions are needed for identification:

(C.3) The elements of \( \mathbf{W} (\mathbf{z}) \), \( \mathbf{D} (\mathbf{z}) \) and \( \mu (\mathbf{z}) \), at observation points \( \mathbf{z} \) such that \( p (\mathbf{z}) > 0 \), fulfill

\[
\sup_{\mathbf{z} \in \mathcal{Z}} |\mathbf{W} (\mathbf{z}) \mu (\mathbf{z})| < \infty, \quad \sup_{\mathbf{z} \in \mathcal{Z}} |\mathbf{D} (\mathbf{z})' \mathbf{W} (\mathbf{z})| < \infty,
\]

for fixed \( N, K \) and \( d \).

(C.4) At observation points \( \mathbf{z} \) with \( p (\mathbf{z}) > 0 \), we define and impose

\[
\kappa_T = \inf_{\mathbf{z} \in \mathcal{Z}} \mathbf{D} (\mathbf{z})' \mathbf{W} (\mathbf{z}) \mathbf{D} (\mathbf{z}) > 0.
\]

(C.5) At observation points \( \mathbf{z} \) such that \( p (\mathbf{z}) > 0 \) the weighting matrix \( \mathbf{W} (\mathbf{z}) \) is bounded from above and below.

Conditions (C.3) and (C.4) are standard bounds in first and second order moment conditions. Note also that the matrix \( \mathbf{W} (\mathbf{z}) \) needs to be bounded at values of \( \mathbf{z} \) such that \( p (\mathbf{z}) > 0 \). These conditions are similar to the assumptions imposed in Hansen (2008) to show weak uniform convergence for general nonparametric estimators. Now we can state that the estimator of the prices of risk satisfies the following:
Theorem 1. Let $\mathcal{F}$ be a compact subset of $\mathbb{R}^d$ and assume conditions (C.1) to (C.5) hold. Then, as $\kappa_T^{-1}h_1$ and $\kappa_T^{-1}h_2$ tend to zero in such a way that $\kappa_T^2 T h_1^d \to \infty$ and $\kappa_T^2 T h_2^d \to \infty$ we have that, under the null hypothesis,

$$
\sup_{z \in \mathcal{F}} \left| \hat{\delta}(z) - \delta(z) \right| = O_p \left( \kappa_T^{-1} \sqrt{\log T} \right) + O_p \left( \kappa_T^{-1} T h_1^2 \right),
$$

as $T$ tends to infinity.

Next, we introduce the assumptions for deriving the asymptotic properties of test statistic $S_n$.

(A.1) The process $(z_i, r_i)$ is absolutely regular, that is,

$$
\beta(s) \equiv \sup_{t \geq 1} \left\{ \sup_{A \in \mathcal{B}(\mathbb{R}^d)} \left| P \left( A \mid \mathcal{F}_t \right) - P(A) \right| \right\} \to 0, \quad \text{as} \quad s \to \infty,
$$

where $\mathcal{F}_t$ is the $\sigma$-field generated by $\{z_i, r_i : k = t, \ldots, s\}$. Furthermore, it is assumed that $\beta(s)$ exhibits a geometric rate of decay.

(A.2) $z_i$ has a bounded density function $p(z_i)$. Furthermore, the joint density of distinct elements of $(z_i, r_i, z_j, z_i, r_j, r_j)$, $(t > s > 1)$, is continuous and bounded by a constant independent of $s$ and $t$. Furthermore, $p(z_i)$ is twice continuously differentiable in all its arguments.

(A.3) Let $\epsilon_{i,t+1} = r_{i,t+1} - \mu(z_i)$. With probability one, $E \left( \epsilon_{i,t+1} \delta_{z_i} - 1 \right) = 0$, for all $t$ and $i = 1, \ldots, N$. Let $\sigma^2(z_i) = \text{Var}(r_{i,t+1} | z_i = z_i) = E \left( \epsilon_{i,t+1}^2 | z_i = z_i \right)$. Then $\sigma^2(z_i)$ satisfies some Lipschitz conditions: $|\sigma^2_{z_i}(z_i + \mathbf{u}) - \sigma^2_{z_i}(z_i)| \leq M(z_i) \| \mathbf{u} \|$, with $E \left( |M(z_i)|^{2+\eta} \right) < \infty$, for all $t$ and $i = 1, \ldots, N$. Furthermore, for all $i \neq j$, and $t \neq s$, $E \left( \epsilon_{i,t+1,1} \epsilon_{j,s+1,1} | z_i, z_j \right) = 0$. Finally, $E \left( |\mu(z_i)|^{16} + r_{i,t+1}^{16} \right) < \infty$.

(A.4) $K(\cdot)$ is a product kernel, that is, $K(z) = \prod_k K(z_k)$, and $k(\cdot)$ is a symmetric density function with bounded support in $\mathbb{R}$ and $[k(z_1) - k(z_2)] \leq c |z_1 - z_2|$ for all $(z_1, z_2)$ in its support.

(A.5) $h \sim T^{-2(1+\eta)}$ for some $1 > \eta > 0$.

(A.6) $h_{d/2}^{1/2} \kappa_T^{-2} \to 0$ and $h_{d/2}^{1/2} \kappa_T^{-2} \to 0$ as $T$ tends to infinity.

(A.7) $T h_{d+2\alpha_1}^d \to 0$ and $T h_{d+2\alpha_2}^d \to 0$ as $T$ tends to infinity.

There are many examples of absolutely regular processes in discrete and continuous time; see Bradley (1986) and Doukhan (1994). Further examples are renewal processes and solutions of stochastic differential equations; see Heinrich (1992) and Veretennikov (1987), respectively. Assumption (A.1) is needed to obtain the asymptotic distribution of the test statistic under the null. More precisely this assumption is necessary to apply a central limit theorem for $\hat{U}$-statistics under dependence. Several CLT’s under dependence are available in the literature, see among others Hjellvik, Yao, and Tjøstheim (1996), Fan and Li (1999) and Gao and Hong (2008). Assumption (A.2) specifies that the density $p(z_i)$ is bounded, while (A.3) controls the tail behavior of the conditional expectation of the errors $\epsilon_i$. We further need to make assumptions about the structure of the matrix $E \left( \epsilon_{i+1} \epsilon_{j+1} \mid z_i, z_j \right)$ for $t \neq s$. Assumption (A.4) on the kernel is also standard in nonparametrics. Finally, (A.5)–(A.7) relate all bandwidths included in the procedure. In particular, (A.5) is concerned with the rate of the bandwidth of the test. In fact, for bandwidths $(h_1, h_2)$ we assume rates that are arbitrarily close to the optimal rates for nonparametric estimation (minimizing the average mean squared error), and for bandwidth $h$ somewhat faster in accordance with nonparametric testing theory. While it generally cannot be directly compared to optimal rates for nonparametric testing when the null is fully parametric, it is interesting to note that it is close to the lower bound of optimal bandwidths found by Horowitz and Spokoiny (2001). More specifically, for $\kappa_T$ going arbitrarily slow to infinity, (A.5) to (A.7) imply that $(h_1, h_2)$ have rates slower than or equal to $T^{-1/2}$ but faster than $T^{-1/(\alpha_1+\eta)}$, $j = 1, 2$, while $h$ simply follows (A.5). Roughly, $\alpha_1$ and $\alpha_2$ should not be smaller than dimension $d$, which reflects the well-known curse of dimensionality in the nonparametric analysis.

Theorem 2. Under conditions (C.1) to (C.5), (A.1) to (A.7), and if the null hypothesis $H_0$ is true, then

$$
Th_{d/2}^{1/2} \kappa_T - Bh^{-d/2} \to d \mathcal{N}(0, V),
$$

where

$$
B = \int_{\mathbb{R}^d} K^2(u) \, du
\times \sum_i \sum_j \int_{\mathbb{R}^d} w_{ij}(z) E \left( \epsilon^2_{i,t+1,1} \mid z_i, z_j \right) p(\xi) \pi_2(z) \, dz
\times V = 2 \int_{\mathbb{R}^d} K(u) K(v) K(u-x) K(v-x) \, du \, dv \, dx
\times \int_i \sum_j \int_{\mathbb{R}^d} w_{ij}(z) E \left( \epsilon^2_{j,t+1,1} \mid z_i, z_j \right) \, dx
\times p^2(z) \pi_2(z) \, dz
$$

as $T$ tends to infinity.

For nonparametric methods, it is well known that, while the first-order approximations can be quite useful to understand the theoretical properties, they are less helpful for the empirical analysis. Typically, large samples are required before higher order terms become negligible; the same holds for the distributional approximation by normality. Furthermore, the asymptotic bias and variance have to be estimated nonparametrically which is not trivial but has an important impact on the test performance. All this makes the use of a bootstrap procedure recommendable in order to approximate the distribution of the test under the null.

Before we turn to the bootstrap and practical implementation issues, we study the power of our test against local alternatives to the null hypothesis. Let us define the sequence of local alternatives to the null hypothesis

$$
H_1: \mu(z) = D(z) \delta(z) + \gamma_T \Delta(z)
$$

where $\gamma_T$ is a sequence that tends to zero such that

$$
(\mathcal{H}1) \sqrt{T h_{d/2}^{1/2} \gamma_T} \to \infty, \text{ and } \gamma_T = o \left( \frac{1}{\sqrt{T h}} \right), \text{ as } T \text{ tends to infinity,}
$$
and \( \Delta(z) \) is a bounded function (uniformly in \( z \)) that fulfills

\[
\sup_{x \in \mathbb{S}} |W(z) \Delta(z)| < \infty.
\]

Assumption (H.1) is related to the bias problem that is mentioned above. Combined with (A.5) we obtain \( T^{1-(1-\epsilon)/3} \gamma_T \to \infty \) with \( \gamma_T = o(T^{1-(1-\epsilon)/3-1/2}) \). The power of our test is described in the following result.

**Theorem 3.** Under conditions (C.1) to (C.5) and (A.1) to (A.7), if \( H_1 \) verifies (H.1) and (H.2), then for all sequences of random variables \( \{c_T : T \geq 1 \} \) with \( c_T = O_p(1) \) we have

\[
P \left( T^{d/2} S_h > c_T \right) \to 1
\]
as \( T \) tends to infinity.

**Theorem 3** indicates that our test has nontrivial power only against sequences of local alternatives for which \( \gamma_T \) tends to zero at a rate that is smaller than \( \sqrt{T} \). As seen in Andrews (1997), tests based on weighted parametric residuals have nontrivial power against local alternatives for which the rate is exactly \( \sqrt{T} \). Thus, at least in terms of the asymptotic local power these tests appear to dominate tests that require slower rates. However, as shown in Horowitz and Spokoiny (2001), at an exact rate of \( \sqrt{T} \), no test can have nontrivial power uniformly over reasonable classes of functions \( \Delta(\cdot) \) in (22).

### 4. Implementation and Calibration

The test distribution is estimated using wild bootstrap. We study different bootstrap procedures of our statistic, and show the consistency for the implemented procedure. Calibration and power are obtained by a proper combination of the bandwidth choices. Importantly, we develop a feasible testing procedure even though three bandwidths are required: Two for the estimation of the first and second conditional moments, and a third one for the test statistic.

#### 4.1. Wild Bootstrap-Type Resampling

Härdle and Mammen (1993) studied three different bootstrap procedures and conclude that wild bootstrap is the most pertinent method for testing the regression structure. Therefore, we adopt a wild bootstrap scheme to estimate the distribution of our statistic. Note that \( \hat{S}_h \) can be written as

\[
\hat{S}_h = \frac{1}{T^2} \int \left[ \sum_t K_h(z - z_t) \hat{u}_{t+1} \right] W_{h_2}(z) \quad (23)
\]

\[
\left[ \sum_t K_h(z - z_t) \hat{u}_{t+1}(z_t) \right] \pi_2(z) \, dz,
\]

with \( \hat{u}_{t+1} = r_{t+1} - D_{h_2}(D_{h_2} W_{h_2} D_{h_2})^{-1} D_{h_2} W_{h_2} H_{h_1} r_{t+1} := (I - \hat{P}) r_{t+1} \) where \( I \) is the identity matrix, and \( H_{h_1} \) is the so-called hat- or smoothing matrix of the kernel estimator used when estimating \( \mu(\cdot) \). Following Härdle and Mammen (1993), we generate \( B \) bootstrap analogues of \( \hat{S}_h \) defined exactly as above but replacing \( \hat{u}_{t+1} \) by \( u^*_{t+1} = (I - \hat{P}) r^*_{t+1} + \epsilon^*_{t+1} \) with \( r^*_{t+1} \) being generated under the null hypothesis. We may then set

\[
r^*_{t+1} = D_{h_2}(z) \hat{\delta}_h(z_t) + \epsilon^*_{t+1}
\]

when simulating the distribution of the statistic under the null hypothesis, and

\[
\hat{S}_h = \frac{1}{T^2} \int \left[ \sum_t K_h(z - z_t) \epsilon_{t+1} \right] W_{h_2}(z) \quad (25)
\]

\[
\left[ \sum_t K_h(z - z_t) \epsilon_{t+1} \right] \pi_2(z) \, dz,
\]

a quadratic form of them. It is not hard to show that under the null hypothesis, \( T^{d/2} \hat{S}_h \) is asymptotically normal, and more importantly, its asymptotic distribution is the same as that of \( T^{d/2} \hat{S}_0 \). This indicates that we could mimic the distribution of \( \hat{S}_h \) by bootstrapping \( \hat{S}_h \). The latter does not depend on the validity of the null hypothesis, although \( \hat{S}_h \) does, and hence it follows the guidelines set by Hall and Wilson (1991).

Kreiss, Neumann, and Yao (2008) discussed various advantages of their bootstrap testing approach compared to that of Härdle and Mammen (1993). In their simulations, however, they only consider cases where the null hypothesis is a parametric
model. That is, while their findings are useful for large samples, they miss the calibration problem one faces under non- and semiparametric null hypotheses highlighted in Sperlich (2014). For our case, this means the impact of $P$ on the test distribution matters in finite samples. This is confirmed by simulations that are not shown.

The obvious balance is to generate bootstrap replicates of the test by substituting $u_{t+1}^*$ for $\hat{u}_{t+1}$ in Equation (23). This can either be seen as a refinement of Kreiss, Neumann, and Yao (2008) by including $\tilde{P}$, or as a simplification of Härde’s and Mammen (1993) by reducing $t_{\nu_{t+1}}^*$ to $e_{t+1}^*$. In all versions the bootstrap innovations $e_{t+1}^*, \ldots, e_T^*$ have to be conditionally independent given the observed data $\{(r_{t+1}, z_t) : 1 \leq t \leq T\}$. Furthermore,

$$E^* [e_{t+1}^*] = 0,$$
$$E^* [e_{t+1}^* e_{t+1}^*'] = \{r_{t+1} - \mu_{h_{1}}(z_t) \} \{r_{t+1} - \mu_{h_{1}}(z_t) \}',$$

where $E^*$ denotes the expectation under the bootstrap distribution. In practice, we can define $e_{t+1}^* = \{r_{t+1} - \mu_{h_{1}}(z_t) \} \eta_{t+1}$ using bandwidth $h_{t}^{*}$ (see below for more details), where $\eta_{t+1}$ is a sequence of iid random variables following a standard Gaussian distribution.

More specifically, we work with the bootstrap statistic

$$\tilde{S}_h^* = \frac{1}{T^2} \int \left[ \sum_t K_h(z - z_t) u_{t+1}^* \right] W_{h_2}(z),$$

for which we can show the following bootstrap distribution.

**Theorem 4.** Assume that the conditions of Theorem 2 hold. For the bootstrap statistic $\tilde{S}_h^*$ defined in Equation (26), we have as $T$ tends to infinity,

$$T h^{d/2} \left[ \tilde{S}_h^* - E^* \left( \tilde{S}_h^* \right) \right] \rightarrow_d N(0, V)$$

conditionally on $\{(r_{t+1}, z_t) : 1 \leq t \leq T\}$, where $V$ is the same as given in Theorem 2, and $T h^{d/2} \left[ E(\tilde{S}_h^*) - E^*(\tilde{S}_h^*) \right] \rightarrow_p 0$.

We reject $H_0$ if $\tilde{S}_h^* > t_{\alpha}^*$, where $t_{\alpha}^*$ is the upper $\alpha$-quantile of the conditional distribution of $\tilde{S}_h^*$ obtained via bootstrap, that is, the $p$-value of the test is the relative frequency of the event $\tilde{S}_h^* > \tilde{S}_h^*$ in the bootstrap replications. Next we prove that the bootstrap is asymptotically correct in the sense that its significant level converges to $\alpha$ as $T$ tends to infinity.

**Theorem 5.** Assume that the conditions of Theorem 2 hold. Let $t_{\alpha}^*$ be the upper $\alpha$-quantile of the conditional distribution of $\tilde{S}_h^*$ given $\{(r_{t+1}, z_t) : 1 \leq t \leq T\}$ and $\alpha \in (0, 1)$. Then, as $T \rightarrow \infty$, under the null hypothesis $P \{ \tilde{S}_h^* > t_{\alpha}^* \} \rightarrow \alpha$.

In practice, the test performance depends certainly on bandwidths $(h_1, h_2, h)$ and $h_1^*$. To understand the calibration problem in nonparametric testing, the following facts must be recognized. While bootstrap methods in nonparametric inference work quite well for the variance estimation, they do less so for the approximation of the smoothing bias. Moreover, even if the variance is well approximated and maybe also the bias problem reduced, the bootstrap distribution may differ from the true one in some higher moments so that the critical values may not be estimated very well. There exist bootstrap methods that asymptotically adapt the higher moments, but there exists also an important literature showing that these methods only improve performance for either huge samples or very particular cases.

To solve these problems at once, we recommend to interpret the quantiles of the distribution of $\tilde{S}_h^* | H_0$, say $t_{\alpha}$, as a function of the smoothing parameters, see Sperlich (2014). Furthermore, it is recommended to maximize the power of the test along $h$ by using the statistic $\tilde{S}$ in Equation (19). Estimates for $E[\tilde{S}_h^* | H_0]$ and $V[\tilde{S}_h^* | H_0]$ are obtained from the bootstrap replicates. Finally, we are looking for $(h_1, h_2, h_1^*)$ such that for the corresponding quantile of $\tilde{S}_h^*$, $h$, $t_{\alpha}$, our test holds the nominal level $\alpha$, as implemented in the next section. Here, $\tilde{S}_h^*$ is defined as $\tilde{S}$ but replacing $\tilde{S}_h^*$ by $\tilde{S}_h^*$. We do not provide an extension of Theorem 2 to $\tilde{S}$ as such proof is purely technical and would follow similar steps as the extension in Rodriguez-Poo, Sperlich, and Vieu (2015).

Note that the proposed test statistic is the best among different alternative (consistent) tests that we study. In fact, this statistic exhibits numerically the best performance. Moreover, some alternatives turned out to be very hard to calibrate. Recall that, even though one has some freedom to choose $(h_1, h_2, h_1^*)$ in practice, it is not guaranteed that for a given sample there exist a bandwidth vector such that the test works. As outlined in Henderson and Sheehan (2018), Rodriguez-Poo, Sperlich, and Vieu (2015), and Sperlich (2014), this is a common problem entailed by facing non- or semiparametric null hypotheses.

### 4.2. Simulation Results

Appendix D describes the DGP of the Monte Carlo studies in detail. We perform two studies along our empirical applications. In the first application, we work with traded factors and monthly data; while in the second one, we study a mixture of traded and non-traded factors with quarterly data. Essentially, we design the experiments with our real datasets, which are described in the next section. That is, the data-generating functions in the model are estimated from the real datasets. Therefore, the simulation study below reflects a realistic situation, and we can use its outcome for the calibration of the test when it is applied to the real data.

There are several parameters and estimators to be chosen. The conditional expectations of factors and excess returns are calculated with local linear estimators in order to eliminate biases in the linear direction, and keep boundary effects small. The second moment matrices $\mathbf{D}$ and $\mathbf{W}$ (when we use the weighting matrix (10)) are estimated by a Nadaraya-Watson estimator. We use the Quartic kernel throughout and ran our simulations with bandwidths $h_j := c_j \cdot h_1^{VP}$ for $j = 1, 2$, that is, cross-validation bandwidths multiplied by constants $c_j > 0$. Most convenient is to simply set $c_1 = c_2 = 1$ and calibrate the
test only on $h^*_1$. In particularly small and/or noisy samples one might also use these constants for further calibration.

For calculating our test, the integral in Equation (18) is calculated using $\pi_2$ only for trimming but else setting $\pi_2 = 1$. Trimming is often done to avoid that boundary effects of the nonparametric estimators distort the test. However, often the interesting things happen right at the boundaries of the support. Therefore, in our simulations we compare the case of no trimming to a trimming by cutting off 5% (2.5% each side) of the most extreme values of $z_t$. In our first experiment, we generate time series with $T = 251$ as in the real monthly dataset. In particular, we use two DGPs, both with a predictor that represents the log of the default spread, and six excess returns where the first three represent the FF factors: (a) the CAPM, where the SDF depends only on the first FF factor (the market portfolio return); (b) the FF model, where the SDF depends on the three commented factors. The simulation was repeated 500 times under the null $H_0$ and 100 times under each $H_1$ scenario we considered. In these simulations, we used 250 bootstrap samples for estimating the $p$-values. It turns out that the sample is large enough for using $h_j := h^{CV}_{j, i}, j = 1, 2$, calibrating the test over $h^*_1 = c^* \cdot h_1$, $c^* > 0$ (typically $\geq 1$) if needed.

Appendix D.1 shows the pricing errors under the alternative hypothesis for the two models, which correspond to $\gamma_f \Delta (z)$ in the local alternatives (22) of Theorem 3. These two components are artificial constructs in order to define the local alternatives to which our test is consistent, while it is clearly consistent to global alternatives. In the simulations, we do not study the rate of convergence, nor in detail how the power increases with sample size for a fixed alternative. Therefore, to get an idea of $\Delta (z)$ in our simulations, we can simply think of the corresponding pricing errors, and $\gamma_f \equiv 1$ as sample size and bandwidths are fixed (i.e., no convergence to infinity or zero, respectively).

In Table 1 the simulation results are summarized for the test statistic $\tilde{S}$ with $W$ following (10) in the one- and three-factor models. Of course, this is just a small extract of all simulations we performed. For example, we also repeated the simulations with $W$ being the identity matrix, other bootstrap procedures, using local linear estimators for the second moments, etc. Table 1 shows that our tests meets quite well the error $\alpha$ of the first kind, and rejects in $100\%$ of all cases under the simulated alternatives. We also see how the calibration of the test works well only by a proper choice of $c^*$ for the bootstrap bandwidth, while taking the cross-validation bandwidths for $(h_1, h_2)$, and $h$ being selected along Equation (19). These findings hold for both models and independently from the trimming function $\pi_2$. What can happen if the sample size is much smaller is illustrated below.

In our second simulation study, we not only generate time series with $T = 251$, as in the real quarterly dataset, but also repeat our simulations with $T = 500$ to see how performance improves when moving from small to moderate samples. The predictor represents the log of the default spread again, and there are three excess returns that represent the three FF factors. We simulate data under two asset pricing models: (a) the CAPM, as explained above; (b) the linearized Epstein-Zin CCAPM, a two-factor model with both the market return and aggregate consumption growth. The simulation was repeated 500 times under the null $H_0$ and 100 times under $H_1$. In these simulations we used 250 bootstrap samples for estimating the $p$-values. As can be seen from Tables 2 and 3, at least when $T$ is small, we need $(c_1, c_2)$ for calibration. As cross-validation bandwidths are known for their tendency to undersmooth, it is not surprising that we achieve calibration for $c_1, c_2 > 1$. We show only results for $c_1 = c_2 = 5$ with $c^* = 1$, that is, $h^*_1 = h_1$. Not surprisingly, for the testing problem (b) containing more nonparametric function estimates, the calibration turns out to be harder. In Tables 2 and 3, the simulation results for test $\tilde{S}$ refers to $W$ following Equation (10). For the sake of brevity, we show results with $T = 500$ only for (b) but skip the values for the trimmed statistics. It can be seen clearly that in the two-factor model with only $T = 251$, the calibration does not work that well. As expected, this problem disappears for increasing samples. In any case, the low power in Table 3 is due to the calibration of $H_1$ to the data, where the pricing errors are low for this model due to reasons explained in the empirical application.

Like in the first experiment, this is just a small extract of all simulations we performed. We obtain similar results for

| $\text{trimming}$ | $c^*$ | av.pv | $H_0$ | $H_1$ |
|-------------------|------|-------|-------|-------|
|                   |      | 0.10  | 0.05  | 0.01  |
| Panel A. One-factor model |      |       |       |       |
| without           | 1    | 0.499 | 0.096 | 0.064 |
|                   | 2    | 0.503 | 0.096 | 0.060 |
| 5% of data        | 1    | 0.500 | 0.100 | 0.060 |
|                   | 2    | 0.503 | 0.098 | 0.058 |
| Panel B. Three-factor model |      |       |       |       |
| without           | 1    | 0.515 | 0.102 | 0.054 |
|                   | 2    | 0.520 | 0.100 | 0.056 |
| 5% of data        | 1    | 0.514 | 0.108 | 0.050 |
|                   | 2    | 0.517 | 0.102 | 0.050 |

| $\text{trimming}$ | $c_2$ | $c_1$ | av.pv | $H_0$ | $H_1$ |
|-------------------|-------|-------|-------|-------|-------|
|                   | 0.75  | 0.75  | 0.493 | 0.100 | 0.046 |
|                   | 1.0   | 1.0   | 0.491 | 0.102 | 0.044 |
| 5% of data        | 1.5   | 1.5   | 0.485 | 0.106 | 0.050 |
|                   | 1.0   | 1.0   | 0.493 | 0.098 | 0.044 |
|                   | 1.5   | 1.5   | 0.487 | 0.096 | 0.048 |
simulations with \( W \) being the identity matrix and/or using other conditioning variables. We also simulated a linearized CCAPM, where there is a single nontraded pricing factor, and obtain similar results. Generally, when using the same bandwidths as for the CAPM, the test becomes a bit more conservative, that is, the same bandwidths give slightly larger \( p \)-values. Therefore in the application we use slightly larger \((c_1, c_2)\) for the CCAPM. Putting all together, our test works well, even for relatively small samples. For appropriately chosen parameters \((c_1, c_2)\), the test is calibrated to hold the rejection level under the null and exhibit reasonable power.

5. Empirical Applications

We study two empirical applications. In the first one, we have asset pricing models with traded factors and monthly data. In the second application, we study a mixture of traded and nontraded factors with quarterly data. Note that our simulations can directly be used for the calibration of the tests, that is, the bandwidth choice that guarantees no over-rejection, because the DGP were fully based on our real data. For the first application, cf. Table 1, these are exactly the cross-validation bandwidths, that is, \( c_1 = c_2 = c^* = 1 \). In the second application, we take \( c_1 = 1.0 \) for the CAPM and 1.5 for the CCAPM, but 3.0 for the two-factor model. For robustness check, however, we also ran the tests with slightly larger and smaller \((c_1, c_2)\) and larger \( c^* \). The findings remained constantly the same.

5.1. Traded Factors: CAPM and FF Model

The vector of excess returns \( r_{t+1} \) is given by the six FF size and book-to-market sorted portfolios. They are constructed as the intersection of two portfolios formed on size (market equity) and three portfolios formed on the ratio of book equity to market equity. These portfolios are widely used in empirical finance, and the corresponding data are available from Kenneth French’s Data Library. We subtract the risk-free return from the portfolio returns to obtain the corresponding excess returns. For \( z_t \) we use alternate two prominent predictors: the logarithm of the default spread \( (lds) \) constructed from FRED data, using yields on AAA and BAA-rated bonds, and the logarithm of the price-earnings ratio \( (lpe) \) taken from Robert Shiller’s web page.

We apply our test to two popular asset pricing models, or equivalently two choices of \( f_{t+1} \). The first model is the conditional CAPM, where the excess return on the market portfolio \( (Mk) \) is the only pricing factor. The second model is the conditional FF model, where the SDF depends on the market, size \( (SMB) \), and value \( (HML) \) factors. We obtain the pricing factors from Ken French’s Data Library. See his web page, as well as Fama and French (1993) for details.

We use monthly nominal data from 1964 to 2016 \( (T = 648) \). We show the results when \( z_t \) is \( lds \), but the performance of the models is similar if we use \( lpe \) instead, and the corresponding figures are available upon request. We find that the estimated conditional means are positive, and clearly nonlinear in \( lds \). These functions increase for most values of \( lds \), which is expected as we can associate high values of this conditioning variable to “bad states,” and low values to “good states.” We also find the size and value effects: the conditional mean of small firms is higher than the mean for big firms, and the conditional mean of high book-to-market firms is higher than the mean for low book-to-market firms.

Figure 1 shows the prices of risk that are obtained from the estimated conditional moments with \( W \) following \((10)\). This figure and the following ones display smooth 90% point-wise confidence intervals obtained from wild bootstrap. For better appreciating the local behavior of our estimators and confidence bands, we indicate the kernel density of the conditioning variable in all figures (adapted to visibility convenience, not normalized). The price of risk of the CAPM is positive and increasing for most values of \( lds \). On the other hand, the FF model has three prices of risk, all of them positive. The market price of risk is relatively flat, but the size price of risk has a clear \( U \)-shape. The value price of risk is decreasing and becomes negative for high \( lds \), but that region does not seem statistically significant.

The first row of Figure 2 shows the square root of the criterion function that the prices of risk minimize, and the second row the conditional mean of the associated SDFs. As discussed, the square root of the criterion function \((7)\) can be interpreted as an HJD when we use the weighting matrix \((10)\). The HJD has an \( U \)-shape for both models, but is clearly lower for the FF model. The SDF mean is positive across the default spread values, even though an affine SDF does not necessarily yield a positive SDF. The importance of studying the SDF mean is explained in Appendix E, where it is related to the correlation between the SDF and returns.

Finally, using 1000 bootstrap replications, the \( p \)-values of our tests are always zero for both models, no matter if we conditioned on \( lds \) or \( lpe \), and independently of whether then we applied trimming or not. This does not change when trying some smaller and/or larger bandwidths. Therefore, we find clear empirical evidence against the conditional CAPM and FF model.
5.2. Traded and Nontraded Factors: CCAPM and Epstein-Zin Model

In our second application, we apply our test to models that have nontraded factors. In particular, we test three popular asset pricing models, or equivalently three choices of \( f_{t+1} \): the conditional CAPM, the linearized CCAPM where the per capita consumption growth (cgp) is the only pricing factor, and the linearized CCAPM with Epstein-Zin preferences (EZ), which nests the previous two models. Consequently, this application considers a model with a traded factor, a model with a nontraded factor, and a model with the two types of factors.

As the models in this section have at most two factors, we can directly test them with the vector of returns \( r_{t+1} \) given by three FF factors (the market, SMB, and HML), which capture the relevant properties of the FF size and book-to-market sorted portfolios. The conclusions of our analysis are similar if we use the corresponding six portfolios. We work with two prominent predictors as components of \( z_t \): the default spread, which is also used in the previous application, and the cointegrating residual of consumption and wealth (\( cay \)). The latter predictor was developed by Lettau and Ludvigson (2001a), and we take its data from Martin Lettau’s web page.
We work with quarterly real data from 1951 to 2014 ($T = 251$). Like in the previous application, we show the results with conditioning variable $lds$, but also computed results conditioning on $cay$. The estimated first and second moments are available from the authors. They show that the market risk premium increases with $lds$. This is expected as we can associate high values of $lds$ to “bad states,” and low values to “good states.” In fact, the market variance also increases with $lds$. The size premium increases for most of the predictor values, but decreases for extreme values. The value premium increases for most of the predictor values, but it is much flatter than the other two risk premia.

Like in the previous application, we show the results when looking at the HJD, that is, choosing $W$ as in (10), and the figures display smooth 90% point-wise confidence intervals obtained from wild bootstrap. Figure 3 shows the prices of risk that are obtained from the estimated conditional moments. Both the market and consumption factors have a positive and nonmonotonic price of risk in the CAPM and CCAPM, respectively. Nevertheless, when we join the two factors in the EZ model, the market factor does not seem to be a relevant pricing factor in the SDF. We provide a reason for such a crowding out effect below.

The first row of Figure 4 shows the HJD and the second row shows the conditional mean of the associated SDFs. The minimized criterion is clearly lower for the CCAPM and EZ models than for the CAPM, being close to zero for the CCAPM and EZ models at many values of $lds$. Also the SDF mean is closer to zero for those two models in those regions, and hence those low criterion functions must be interpreted with care, see Appendix E. There we describe the problematic case of an SDF that is uncorrelated with the cross section of returns, which translates into both zero pricing errors and a zero SDF mean. In fact, we find that confidence intervals show that the conditional covariances of returns and consumption are not significantly different from zero. This may also explain why consumption crowds out the market in the EZ model, see the bottom row of Figure 3. Importantly, Appendix E also clarifies that introducing a riskless asset in this investment set does not fully solve the problem.

Table 4 reports the $p$-values of our test for the three models with and without trimming based on 1000 bootstrap replications. The first row of Table 4 reports the results for $z := lds$, while the second row focuses on $cay$. We also computed the $p$-values for other admissible bandwidths, along the lines of Tables 2 and 3, obtaining essentially the same results: $p$-values 0 for the CAPM, about 0.5 for CCAPM and EZ when conditioning on $lds$ but around 0.1 if conditioning on $cay$. The trimming has no visible effect. As expected from the findings in the simulation study, larger $(c_1, c_2)$ tend to give smaller $p$-values.

In sum, we find that only the CAPM is clearly rejected. Roussanov (2014) found evidence against the correct pricing of some portfolios by the CCAPM and EZ models with similar data, but using the centered SDF approach. However, finding empirical evidence against a joint hypothesis (i.e., in a multiple testing problem) is certainly harder than it is for individual testing problems. Moreover, Peñaranda and Sentana (2015) found discrepancies between the centered and uncentered SDF approaches when testing unconditional asset pricing models. In their data, this seems to be due to a poor unconditional correlation between consumption growth and the cross-section of excess returns to be priced. Interestingly, we find a conditional counterpart of this problem.

Following Appendix E, nontraded factors that are conditionally uncorrelated with excess returns will automatically price those returns with an SDF whose conditional mean is 0. The second row of Figure 4 shows SDF means for the CCAPM and EZ models that are close to zero in the regions where the pricing errors are low. Note that the figures display point-wise
confidence bands, and uniform bands would be wider. We do not provide a formal test of a zero SDF mean, but we expect that such a test would not reject for the CCAPM and EZ model. In that case, the centered SDF approach is not well defined so that the corresponding asymptotic theory of the estimators and test would break down. When we implemented the centered approach, the procedure suffered from severe numerical problems simply because for some values of \( z \) (\( \text{lds} \) or \( \text{cay} \)) the nonparametric conditional covariances were almost zero. Further inference based on its inverse is then meaningless and actually corresponds to identification problems. Consequently, we rather recommend to use the uncentered SDF approach, which is well defined even in those problematic cases, and to carefully check the behavior of the SDF mean to detect an uncorrelated SDF.

6. Conclusions

In this article, we present an adaptive omnibus specification test of conditional asset pricing models. These models provide constraints that conditional moments of returns and pricing factors must satisfy, but frequently do not provide information on the functional form of those conditional moments. The main interest of our test is that it is not only robust to functional form misspecification of conditional moments, but it also detects any relationship between pricing errors and conditioning variables.

This last issue is of crucial interest for power in testing conditional models.

Our test statistic belongs to the class of consistent specification tests that compare parametric and nonparametric fits. However, our model under the null is also nonparametric, which is another crucial difference with respect to the existing tests. We develop the asymptotic theory of the test, and give special emphasis on practical issues. The test distribution is estimated using wild bootstrap, and calibration and power are obtained by proper bandwidth choices. Our test works well even for small samples.

We find interesting results in our empirical studies. In the first one, we study asset pricing models with traded factors and monthly data. Both the conditional CAPM and the FF model are rejected with size and book-to-market sorted portfolios, even though the latter model yields lower pricing errors. In the second application, we study a mixture of traded and non-traded factors with quarterly data. While we find strong empirical evidence against the conditional CAPM, we do not find such evidence against zero pricing errors with models that consider consumption growth. Yet, the low pricing errors do not necessarily mean that consumption is an economically meaningful pricing factor. The nonrejection seems to be due to the poor conditional correlation between consumption and the cross section of stock returns.

There are several interesting avenues for future research. We have focused our analysis on linear SDFs, which are prevalent in empirical finance, but we could extend our results to cover nonlinear models. This extension could be interesting given the problems that we have found with linearized variants of the CCAPM.

We have related our test statistic to the HJD for a particular choice of the metric of pricing errors. The HJD is often used to compare competing models that are likely misspecified, and
such comparisons represent another interesting extension of our theoretical setting. The choice of predictors is even more relevant in that setting, and hence it would be interesting to explore a larger set of conditioning variables. However, in the case of several covariates, bias (or dimension) reducing methods are needed like higher order kernels or separability structures.

Following Fama and French (2015) among others, the empirical asset pricing literature has enriched the usual cross section of size and book-to-market sorted portfolios with profitability and investment sorted portfolios. It would be interesting to apply our test to wider cross sections. In fact, recently there has been interest in testing asset pricing models with individual stocks instead of portfolios. We plan to extend our econometric methodology to the case of a number of assets that can grow with sample size.

Finally, there are recent applications of machine learning to asset pricing such as Kozak, Nagel, and Santos (2020) that consider both many conditioning variables and stocks. This new literature does not start from some given traded and non-traded factors like our testing methodology does, but instead extracts traded factors from the return data. In particular, firm-specific conditioning variables (e.g., size and book-to-market ratio) are used to obtain new pricing factors in the spirit of the FF model. We plan to extend our econometric methodology to accommodate both macroeconomic and firm-specific conditioning variables, and provide a formal asset pricing test to this new literature.

Acknowledgments

We thank participants at the 10th Conference on Computational and Financial Econometrics (Sevilla), the New Methods for the Empirical Analysis of Financial Markets (Comillas), the 26th Finance Forum (Santander), the 88th Meeting of the Southern Finance Association (Washington DC), the 68th Meeting of the Midwest Finance Association (Chicago), the CUNY Macroeconomics and Finance Colloquium (New York), the 12th Society for Financial Econometrics Conference (Shanghai), and specialist Yacine Aït-Sahalia, Stefanos Delikouras, and Nikolay Gospodinov, for helpful comments and suggestions. The comments from Christian Hansen (the editor), an associate editor, and two anonymous referees have also led to a substantially improved article.

Funding

The authors gratefully acknowledge financial support from the Research Projects APIE 1/2015-17: “New methods for the empirical analysis of financial markets” of the Santander Financial Institute (SANFI) of UCEIF Foundation resolved by the University of Cantabria and funded with sponsorship from Banco Santander; and PID2019-105986GB-C22: “Dependent risks: modeling and application” funded by the Spanish M. de Ciencia e Innovación.

Supplementary Material

The appendices are available online. Appendix A shows the proofs of Theorems 1–5. Appendix B reviews previous approaches to testing asset pricing models with nonparametric prices of risk. Appendix C develops the centered SDF variant and its relationship with the uncentered SDF variant. Appendix D describes the Monte Carlo design in detail. Finally, Appendix E studies the special case of an SDF that is conditionally uncorrelated with excess returns, and also the SDF implications of a riskless asset.

References

Aït-Sahalia, Y., and Lo, A. W. (1998), “Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices,” The Journal of Finance, 53, 499–547. [1456]
Andrews, D. W. K. (1997), “A Conditional Kolmogorov Test,” Econometrica, 48, 1097–1128. [1461]
Antoine, B., Proulx, K., and Renault, E. (2018), “Pseudo-True SDFs in Conditional Asset Pricing Models,” Journal of Financial Econometrics, 1, 1–59. [1455]
Bansal, R., and Viswanathan, S. (1993), “No Arbitrage and Arbitrage Pricing: A New Approach,” Journal of Finance, 48, 1231–1262. [1456]
Bradley, R. C. (1986), “Basic Properties of Strong Mixing Conditions,” in Dependence in Probability and Statistics (Obershof, 1985), Progr. Probab. Statist. (Vol. 11), eds. E. Eberlein and M.S. Taqqu, Boston, MA: Birkhäuser, pp. 165–192. [1460]
Bryzgalova, S. (2016), “Spurious Factors in Linear Asset Pricing Models,” Stanford University manuscript. [1456]
Burnside, C. (2016), “Identification and Inference in Linear Stochastic Discount Factor Models With Excess Returns,” Journal of Financial Econometrics, 14, 295–330. [1456]
Cai, Z., Ren, Y., and Sun, L. (2015), “Pricing Kernel Estimation: A Local Estimating Approach,” Econometric Theory, 31, 560–580. [1455]
Campbell, J. Y., and Cochrane, J. H. (1999), “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” Journal of Political Economy, 107, 205–251. [1457]
Cochrane, J. H. (2005), Asset Pricing, Princeton, NJ: Princeton University Press. [1456,1458]
Doukhan, P. (1994). Mixing: Properties and Examples, Lecture Notes in Statistics (Vol. 85), New York: Springer-Verlag. [1460]
Fama, E., and French, K. (1993), “Common Risk Factors in the Returns on Stocks and Bonds,” Journal of Financial Economics, 33, 3–56. [1464]
——— (2015), “A Five-Factor Asset Pricing Model,” Journal of Financial Economics, 116, 1–22. [1468]
Fan, J., and Gijbels, I. (1995), Local Polynomial Modelling and Its Applications, London: Chapman & Hall. [1458]
Fan, J., and Li, Q. (1999), “Central Limit Theorem for Degenerate U-Statistics of Absolutely Regular Processes With Applications to Model Specification Testing,” Journal of Nonparametric Statistics, 10, 245–271. [1460]
Gagliardini, P., Gourieroux, C., and Renault, E. (2011), “Efficient Derivative Pricing by the Extended Method of Moments,” Econometrica, 79, 1181–1232. [1455]
Gagliardini, P., and Ronchetti, D. (2020), “Comparing Asset Pricing Models by the Conditional Hansen-Jagannathan Distance,” Journal of Financial Econometrics, 18, 333–394. [1455,1457]
Gao, J., and Hong, Y. (2008), “Central Limit Theorems for Generalized U-Statistics With Applications in Nonparametric Specification,” Journal of Nonparametric Statistics, 20, 61–76. [1460]
Gospodinov, N., Kan, R., and Robotti, C. (2017), “Spurious Inference in Reduced-Rank Asset-Pricing Models,” Econometrica, 85, 1613–1628. [1456]
Hall, P., and Wilson, S. R. (1991), “Two Guidelines for Bootstrap Hypothesis Testing,” Biometrics, 47, 757–762. [1461]
Hansen, B. E. (2008), “Uniform Convergence Rates for Kernel Estimation with Dependent Data,” Econometric Theory, 24, 726–748. [1459]
Hansen, L. and Jagannathan, R. (1997), “Assessing Specification Errors in Stochastic Discount Factor Models,” Journal of Finance, 52, 557–590. [1457]
Härdle, W., and Mammen, E. (1993), “Comparing Nonparametric Versus Parametric Regression Fits,” Annals of Statistics, 21, 926–1947. [1456,1461,1462]
Hart, J. (1997), Nonparametric Smoothing Lack-of-Fit Tests, New York: Springer Verlag. [1456]
Heinrich, L. (1992), “Bounds for the Absolute Regularity Coefficient of a Stationary Renewal Process,” Yokohama Mathematical Journal, 40, 25–33. [1460]
Henderson, D., and Sheehan, A. (2018), “Kernel-Based Testing With Skewed and Heavy-Tailed Data: Evidence From a Nonparametric Test for Heteroskedasticity,” *Economics Letters*, 172, 8–11. [1462]

Hjellvik, V., Yao, Q., and Tjøstheim, D. (1996), “Linearity Testing Using Local Polynomial Approximation,” Sonderforschungsbereich 373, Discussion Paper 60, Humboldt Universität zu Berlin. [1460]

Horowitz, J., and Spokoiny, V. (2001), “An Adaptive, Rate-Optimal Test of a Parametric Mean-Regression Model Against a Nonparametric Alternative,” *Econometrica*, 69, 599–632. [1460,1461]

Jagannathan, R., and Wang, Z. (1996), “The Conditional CAPM and the Cross-Section of Expected Returns,” *Journal of Finance*, 51, 3–53. [1455]

Kan, R., and Zhang, C. (1999), “GMM Tests of Stochastic Discount Factor Models With Useless Factors,” *Journal of Financial Economics*, 54, 103–127. [1456]

Kleibergen, F., and Zhan, Z. (2020), “Robust Inference for Consumption-Based Asset Pricing,” *The Journal of Finance*, 75, 507–550. [1456]

Kozak, S., Nagel, S., and Santosh, S. (2020), “Shrinking the Cross-Section,” *Journal of Financial Economics*, 135, 271–292. [1468]

Kreiss, J.-P., Neumann, M. H., and Yao, Q. (2008), “Bootstrap Tests for Simple Structures in Nonparametric Time Series Regression,” *Statistics and Its Interface*, 1, 367–380. [1461,1462]

Lettau, M., and S. Ludvigson (2001a), “Consumption, Aggregate Wealth, and Expected Stock Returns,” *Journal of Finance*, 56, 815–849. [1465]

——— (2001b), “Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying,” *Journal of Political Economy*, 109, 1238–1287. [1455,1457]

Lewellen, J., and Nagel, S. (2006), “The Conditional CAPM Does Not Explain Asset-Pricing Anomalies,” *Journal of Financial Economics*, 82, 289–314. [1455]

Masry, E. (1996), “Multivariate Local Polynomial Regression for Time Series: Uniform Strong Consistency Rate,” *Journal of Time Series Analysis*, 17, 571–599. [1459]

Nagel, S., and Singleton, K. (2011), “Estimation and Evaluation of Conditional Asset Pricing Models,” *Journal of Finance*, 66, 873–909. [1455,1457]

Peñaranda, F., and Sentana, E. (2015), “A Unifying Approach to the Empirical Evaluation of Asset Pricing Models,” *Review of Economics and Statistics*, 97, 412–435. [1456,1466]

Rodriguez-Poo, J. M., Sperlich, S., and Vieu, P. (2015), “Specification Testing When the Null is Nonparametric or Semiparametric,” *Econometric Theory*, 31, 1281–1309. [1456,1462]

Rosenberg, J. V., and Engle, R. F. (2002), “Empirical Pricing Kernels,” *Journal of Financial Economics*, 64, 341–372. [1456]

Roussanov, N. (2014), “Composition of Wealth, Conditioning Information, and the Cross-Section of Stock Returns,” *Journal of Financial Economics*, 111, 352–380. [1455,1456,1457,1458,1466]

Sperlich, S. (2014), “On the Choice of Regularization Parameters in Specification Testing: A Critical Discussion,” *Empirical Economics*, 47, 427–450. [1459,1462]

Veretennikov, A. Y. (1987), “Estimates of the Mixing Rate for Stochastic Equations,” *Teor. Veroyatnost. i Primenen*, 32, 299–308. [1460]

Wang, K. (2003), “Asset Pricing With Conditioning Information: A New Test,” *Journal of Finance*, 58, 161–196. [1455,1456,1457,1458]

Yin, J., Geng, Z., Li, R., and Wang, H. (2010), “Nonparametric Covariance Model,” *Statistica Sinica*, 20, 469–479. [1458,1459]