



DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY



Time-varying parameters: New test tailored to applications in finance and macroeconomics

Russell Davidson and Niels S. Grønborg

CREATES Research Paper 2018-22

Time-varying parameters: New test tailored to applications in finance and macroeconomics*

Russell Davidson^{a,b}

Niels S. Grønborg^{c,d}

^a *Department of Economics and CIREQ, McGill University, Montréal, Québec H3A 2T7, Canada.*

^b *Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, AMSE, France.*

^c *Department of Economics and Business Economics and CREATES, Aarhus University, Denmark.*

^d *Danish Finance Institute.*

Abstract

Many economic theories imply a linear relationship with constant parameters between financial or macroeconomic variables. While the linear model with constant parameters is often disputed in the literature, this model specification is rarely tested. This paper proposes a new and intuitively appealing test for model specification tailored for applications in finance and macroeconomics. Importantly, the test allows for autocorrelation, which is often present in these applications. We demonstrate impressive properties of the test in a realistic simulation study and obtain important insights from empirical applications.

JEL classification: C12, C14, C32, E61, G12.

Keywords: Nonparametric estimator, Time-varying parameters, Bootstrap, Finance, Macroeconomics.

This version: August 13, 2018.

*We are grateful to Mikkel Bennedsen, Jonas N. Eriksen, Stig V. Møller, participants at the SoFiE conference (Lugano, 2018), and seminar participants at Aarhus University and CREATES for valuable comments and suggestions. Any mistakes are our own. Corresponding author: N. S. Grønborg, ngroenborg@econ.au.dk. Aarhus BSS, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. N. S. Grønborg acknowledges support by Center for Research in Econometric Analysis of Time Series (CREATES), funded by the Danish National Research Foundation, Grant No. DNFR 78.

1 Introduction

The financial and macroeconomic literatures contain many examples of theories that imply linear relationships with constant parameters between financial or macroeconomic variables. A few prominent examples include: tests of theories such as the CAPM of Sharpe (1964) and Lintner (1965); the arbitrage pricing theory of Ross (1976) or factor models in the spirit of Fama and French (1993), all of which are typically centered around linear regression models with constant parameters; much of the literature concerned with predicting the equity premium, as summarized in Welch and Goyal (2008), which is based on simple regression models; factor models in the spirit of Stock and Watson (2002), which rely on a linear model with constant factor loadings; several papers in the macro-finance and macroeconomics literatures, where a simple linear policy rule with constant coefficients in the spirit of Taylor (1993) is assumed, see for example Ang and Piazzesi (2003) and Clarida et al. (2000).

The linear model with constant parameters is often disputed in the literature, and several interesting papers consider variations of the linear model with time-varying parameters. OLS estimates based on rolling windows are often used as evidence against constant parameters, but formal testing procedures are rarely carried out. As an example, Figure 1 presents an update of Figure 1 in the influential paper by Ang and Chen (2007).¹ The top panel of Figure 1 presents a time-series plot of rolling 60 month window OLS estimates of the market beta of a portfolio of value stocks (red line) and a line representing the OLS estimates based on the full sample of observations (blue line). The bottom panel of Figure 1 presents the corresponding plot for a portfolio of growth stocks. The rolling window estimates of market beta of the value portfolio varies considerably over time, while the market beta of the growth portfolio is closer to the full sample OLS estimate. The graphical inspection of rolling window estimates indicates time-variation, but ultimately leaves the analyst to make a judgement call. A formal test would be more convincing.²

This paper introduces a new way to test for time variation in model parameters. Typically, we have a simple null hypothesis in mind, namely that the model is linear and parameters are constant. We compare the estimates obtained under the null hypothesis to those obtained by a very flexible nonparametric regression model. Our test allows the user to specify a model under the null and compare the resulting estimates to those of a nonparametric alternative. If the resulting standardized test statistic is too large, it means that the model assumed under the null

¹The figure in Ang and Chen (2007) presents the time series of rolling window market betas for both value and growth in the same plot and does not include the full sample OLS estimates. The sample has been updated and ends in July 2016.

²We formally test for parameter stability in the CAPM over the long run in in section 4.1.

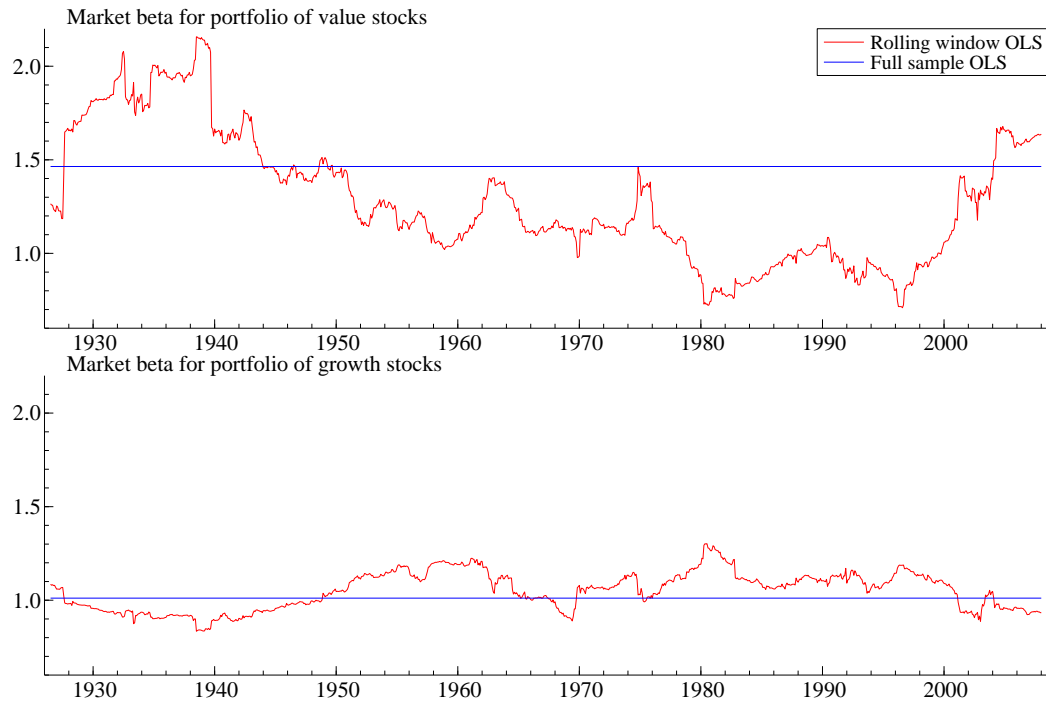


Figure 1: Time-varying market betas. The figure reproduces and updates Figure 1 of Ang and Chen (2007) and illustrates the potential time variation in the beta from the regression model, $R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}$, where $R_{i,t}$ is the excess return on a portfolio, $i = value, growth$, $R_{m,t}$ is the excess return on the market portfolio and $\varepsilon_{i,t}$ is the residual. The top panel presents the market beta of a portfolio of value stocks. The red line is the 60 month rolling window estimate of beta and the blue line is the full sample OLS estimate. The bottom panel presents the same estimates for a portfolio of growth stocks. The estimates are based on a sample of monthly returns from July 1926 to July 2016.

hypothesis is too restrictive. The test is easily implemented with bootstrap methods, where the distribution of the test statistic under the null hypothesis can be simulated.

In an extensive simulation study, we show that the empirical rejection rates are very close to nominal levels when the null hypothesis is true and that the test has good power against various realistic alternatives.

Testing for time variation in the parameters of regression models is by no means a new idea. In the following we highlight some important contributions to the field, but we do not present an exhaustive list.³

Early work in the field include the CUSUM tests of Brown et al. (1975), the supremum tests of Andrews (1993), and the double maximum test of Bai and Perron (1998), all of which are used frequently in the literature. Lin and Teräsvirta (1994) propose an LM-type test, where parameter constancy is tested against a parametric alternative. This test is similar to the one we propose, but instead of a parametric alternative we chose a nonparametric one. Elliott and Müller (2006) present an LR-type test in which parameters are allowed to vary over time in a very flexible way. They prove that asymptotically no power is lost by leaving the particular form of time variation unspecified. Their test is very easy to implement, and the authors document impressive small-sample properties. Cai (2007) and Chen and Hong (2012) present tests very similar to the one we propose. As with our approach, their tests are based on the difference between OLS and a nonparametric regression model, but, where we test on the estimated parameters directly, Chen and Hong (2012) and Cai (2007) test on the fitted values of the regression model and the residuals, respectively. Du et al. (2013) develop a very general framework, where both equality and inequality constraints can be imposed in nonparametric regression models, and propose using bootstrap methods to test the validity of the constraints.

In the financial literature notable contributions have been made by Ang and Kristensen (2012) and Guo et al. (2017). Ang and Kristensen (2012) propose testing whether nonparametric estimates of the parameters in a conditional version of the CAPM are constant, and develop a framework for doing so. Guo et al. (2017) rely on a different approach where the parameters of a conditional CAPM are estimated using splines. The shapes of the splines are determined by a relatively low number of parameters, and shape constraints on the splines can be tested based on tests of these parameters.

Chen and Hong (2012) carry out a large simulation study, and the authors demonstrate that their proposed test statistic performs satisfactorily when compared to the tests of Brown et al. (1975), Andrews (1993), Lin and Teräsvirta (1994), Bai and Perron (1998) and Elliott and Müller

³The interested reader is referred to Perron (2006) for an excellent review of the literature on structural changes.

(2006). Thus, Chen and Hong (2012) serve as our benchmark and we carry out simulations in the same framework and present comparable results.

The test of Chen and Hong (2012) does not allow for autocorrelation, but in our proposed framework it is possible to take this into account. We demonstrate that our test is superior to the test of Chen and Hong (2012) in the presence of autocorrelation, while the two tests have similar performance when errors are IID or heteroskedastic.

Autocorrelation is often present in the disturbances of economic models, and it should be handled in a reasonable way. A case can be made that, because autocorrelation is the result of model misspecification, the econometrician should try to adjust her model in order to alleviate this problem, for example by including lags of the dependent and explanatory variables, instead of accounting for the autocorrelation in the disturbances. Many existing tests assume a well-specified model, see Chen and Hong (2012) for example, and do not consider autocorrelation. The problem with this approach is that a particular economic model does not necessarily leave the econometrician with the freedom to make such adjustments. If the structure imposed by an economic model does not allow the econometrician to account for autocorrelation when specifying the model, then allowing for autocorrelation in the disturbances becomes essential. Our approach allows for the presence of autocorrelation, and simulation results show that the empirical level of the test is close to the nominal level in this case.

We uncover several interesting empirical findings. First we consider the long-run CAPM regression model in Figure 1, where our test rejects the linear model with constant parameters for both the portfolio of value stocks and for the portfolio of growth stocks. Ang and Chen (2007) find the book-to-market effect to be strong in the post-1963 sample. When testing on this period we reject the linear model for constant parameters for the portfolio of value stocks and for a book-to-market strategy. Thus, our results largely support the work of Ang and Chen (2007). Second, we test the specification of simple policy rules as presented in Taylor (1993) and present evidence against constant parameters. Third, we test for constant parameters in a predictive regression of equity premium on dividend yield, as considered in Welch and Goyal (2008), and reject the model with constant parameters. This result supports the methodology of Dangl and Halling (2012) and Johannes et al. (2014) and the findings of, among others, Pesaran and Timmermann (2002) and Paye and Timmermann (2006). Fourth, we consider the five-factor model of Fama and French (2015) and demonstrate that while inference based on full sample OLS estimates might indicate a well-specified asset pricing model in terms of zero alpha, it does not guarantee that the model does not suffer from other misspecifications. Finally, we test the relationship known as Okun's law.

The rest of this paper is organized as follows. Section 2 introduces the test statistic. Section 3 investigates the properties of our test in a large simulation study with realistic data generating processes. Section 4 considers five different applications of the proposed test. Finally, Section 5 concludes.

2 Testing Model Specification

Models with time-varying parameters have become very popular, see for example Ang and Chen (2007) or Dangl and Halling (2012), but this choice of model is rarely made based on a test for correct model specification. Notable exceptions include Ang and Kristensen (2012) and Guo et al. (2017), both of which test for time variation in the parameters of the CAPM. The proposed test in this paper is inspired by Cai (2007) and Chen and Hong (2012). The idea behind the test is to estimate the model by the nonparametric Local Linear Regression Model and compare the estimates to the ones obtained by a parametric estimation method.

The test we introduce is based on the time-varying parameter model, considered by, among others, Chen and Hong (2012). The time-varying parameter model is specified as

$$Y_t = \mathbf{X}_t \boldsymbol{\theta}_t + \varepsilon_t, \quad \text{for } t = \frac{1}{T}, \dots, 1, \quad (1)$$

where T denotes the sample size and \mathbf{X}_t contains d explanatory variables. We will estimate $\boldsymbol{\theta}_t$ using a nonparametric estimator and compare it to the OLS estimate, which is obtained by assuming that $\boldsymbol{\theta}_t$ is constant. The next section presents the nonparametric estimator of $\boldsymbol{\theta}_t$.

2.1 Nonparametric Regression

In (1) the time-varying parameters take the form $\boldsymbol{\theta}_t = (\theta_{t,1}, \dots, \theta_{t,d})'$, where we assume that $\theta_{j,t} = \theta_j(t)$ are unknown functions for $j = 1, \dots, d$. The model and its estimation was first considered by Robinson (1989, 1991). Nonparametric estimation of the model was further considered in Orbe et al. (2000, 2005) and Cai (2007).⁴

In this paper we choose the local linear estimator as our nonparametric estimator, where at any point in time t the unknown function $\boldsymbol{\theta}(s)$, where $s = \frac{1}{T}, \dots, 1$, is approximated around t by a linear function

$$\boldsymbol{\theta}(s) \approx \boldsymbol{\theta}(t) + \boldsymbol{\theta}'(t)(s - t). \quad (2)$$

⁴In Chapter 9 of Li and Racine (2007) this type of model is called a semiparametric smooth coefficient model to separate it from the full-fledged nonparametric cases where $Y_t = g(\mathbf{X}_t) + \varepsilon_t$, where $g(\cdot)$ is an unknown function. While this classification is more accurate we stick with the terminology normally used in the literature and refer to $Y_t = \mathbf{X}_t g(t) + \varepsilon_t$ as a nonparametric model to avoid confusion.

We adopt the following notation and define $\boldsymbol{\theta}(t) = \boldsymbol{\theta}_t$ and $\boldsymbol{\theta}'(t) = \boldsymbol{\theta}'_t$. At any point in time $t \in [0, 1]$, $\hat{\boldsymbol{\theta}}_t$ and $\hat{\boldsymbol{\theta}}'_t$ minimize the objective function:

$$\sum_{s=1/T}^1 [Y_s - \mathbf{X}_s \boldsymbol{\theta}_t + \mathbf{X}_s \boldsymbol{\theta}'_t(s-t)]^2 k_{st},$$

where $k_{st} = k\left(\frac{s-t}{h}\right)/h$. Here, k is a kernel function and h is the bandwidth. We focus exclusively on $\hat{\boldsymbol{\theta}}_t$. Further details are in Appendix A.

2.1.1 Linear approximation

The assumption in (2) is that the unknown function can be approximated well at a given point in time by a linear function. Ang and Kristensen (2012) relies on the local constant model, meaning that they rely on a simpler specification in (2). For specific applications the local constant specification might be sufficient, while other applications might require more flexible specifications. Note that rejection of the null hypothesis, as with any test, does not imply that the nonparametric regression model represents the true data generating process (DGP). The important thing is to select an estimation technique which, in case regression parameters are not constant, leads to consistent estimates significantly different from those obtained by OLS.

2.2 Hypothesis testing

Consider the following null-hypothesis, as presented in Cai (2007)

$$H_0 : \quad \theta_{t,j} - f_j(t, \boldsymbol{\phi}) = 0 \quad \text{for } j = 1, \dots, d,$$

where $f_j(t, \boldsymbol{\phi})$ is a family of candidate functions indexed by the unknown parameter vector $\boldsymbol{\phi}$. In the case of the linear model with constant parameters the candidate function is of a particularly simple form, namely $f_j(t, \boldsymbol{\phi}) = \phi_j$ for all j and $\hat{\phi}_j$ is the OLS estimate of the j -th parameter. We can then present the hypothesis of interest as

$$H_0 : \quad \boldsymbol{\theta}_t - \boldsymbol{\phi} = 0 \quad \text{for } t = \frac{1}{T}, \dots, 1,$$

$$H_1 : \quad \boldsymbol{\theta}_t - \boldsymbol{\phi} \neq 0 \quad \text{for at least one } t.$$

If we denote the vector of OLS estimates for the j parameters by $\hat{\phi}$, an asymptotically pivotal statistic for tests of this type of hypotheses can be formulated as

$$W = \begin{bmatrix} (\hat{\theta}_{1/T} - \hat{\phi})' & \dots & (\hat{\theta}_1 - \hat{\phi})' \end{bmatrix} \hat{\Omega}^{-1} \begin{bmatrix} \hat{\theta}_{1/T} - \hat{\phi} \\ \vdots \\ \hat{\theta}_1 - \hat{\phi} \end{bmatrix}, \quad (3)$$

where $\hat{\Omega}$ is an estimate of the $Td \times Td$ -dimensional covariance matrix of the difference between the two estimators. The elements of this matrix are fairly simple to calculate but the size of the matrix makes inversion both time consuming and potentially unreliable. This problem is similar to one faced in Patton and Timmermann (2010), where the authors develop a test for a monotonic relationship between returns on securities. The authors face the problem that testing on all potential pairs of security returns leads to a very high number of parameter restrictions. They solve their problem by focusing their test on the smallest difference between the returns of two adjacent securities, arguing that if the smallest difference is positive, then all differences are positive. Inspired by this paper we propose the following solution.

To avoid dealing with Ω , we instead consider the following hypothesis

$$\begin{aligned} H_0 : \quad & \theta_t - \phi = 0 \quad \text{for } t = \frac{1}{T}, \dots, 1, \\ H_1 : \quad & \max_{t=1/T, \dots, 1} \theta_t - \phi \neq 0. \end{aligned} \quad (4)$$

This hypothesis can be tested by defining

$$w_t = (\hat{\theta}_t - \hat{\phi})' \hat{\Omega}_t^{-1} (\hat{\theta}_t - \hat{\phi}) \quad (5)$$

and then considering the following test statistic

$$W_{\max} = \max_{t=1/T, \dots, 1} w_t.$$

Note, that the calculation of W_{\max} only requires us to estimate the T different $d \times d$ blocks on the diagonal of Ω , and not all $Td \times Td$ elements. A test statistic based on the supremum is used by Andrews (1993) to detect unknown breakpoints. For any t , the covariance matrix for the difference of the estimators is given by

$$\Omega_t = V(\hat{\theta}_t - \hat{\phi}) = V(\hat{\theta}_t) + V(\hat{\phi}) - \text{cov}(\hat{\theta}_t, \hat{\phi}) - \text{cov}(\hat{\theta}_t, \hat{\phi})'. \quad (6)$$

A suitable estimator $\hat{\Omega}_t$ could be, for example, a White or Newey-West type estimator. This approach requires inversion of T individual $d \times d$ matrices in (5) and not a $Td \times Td$ matrix as in (3). Further details are available in Appendix B.

2.2.1 Testing jointly

Ang and Kristensen (2012) carry out their test on alpha and beta individually, while our procedure tests the coefficients of the model jointly. Testing on the coefficients individually can lead to rejection levels which are different from nominal levels of the test and the severity of this problem increases for regression models with more variables, as for example the five factor model of Fama and French (2015).

2.2.2 Other candidate functions

The linear model with constant parameters is not the only interesting candidate function. Some applications are simpler and might require a test based on a candidate function where parameters are constant and known to the analyst. We consider one such case in Section 4.2, where we apply our test to a policy rule in the spirit of Taylor (1993). Other applications might be more complicated. It is also possible to test whether the linear approximation in (2) is sufficient, and thus test the Local Linear model against other members of the class of Local Polynomial Regression models.

2.2.3 Asymptotic distribution

Delgado and Manteiga (2001) show that test statistics such as the one we propose, under certain conditions, converge in distribution to the supremum of the sum of a zero mean Gaussian process and a bias term. The authors note that such asymptotic tests are difficult to implement in practice and, in line with Patton and Timmermann (2010), advocate using bootstrap procedures for inference.

2.3 Bootstrap inference

Similarly to Cai (2007) and Chen and Hong (2012), finite sample inference will be based on the Wild Bootstrap. Under the null hypothesis the model from (1) is

$$Y_t = \mathbf{X}_t \boldsymbol{\phi} + \varepsilon_t,$$

and can be estimated by OLS. Let $\hat{\varepsilon}_{\phi,t}$ denote the estimated residuals under the null hypothesis. Bootstrap samples can be generated by

$$Y_t^* = \mathbf{X}_t \hat{\boldsymbol{\phi}} + s_t^* \hat{\varepsilon}_{\phi,t},$$

where s_t^* is a draw from the Rademacher distribution. W_{\max}^* can be estimated based on Y_t^* and by repeating the process the distribution of the test statistic under the null hypothesis can be constructed.

A notable difference to Cai (2007) and Chen and Hong (2012) is that both papers base their bootstrap on the non-parametric residuals, whereas we rely on the residuals obtained under the null.

If ε_t contains autocorrelation, another bootstrap approach is needed. We rely on the stationary bootstrap of Politis and Romano (1994), where the block length is chosen according to Politis and White (2004) and Patton et al. (2009).⁵

3 Simulation

This section presents the results of a series of experiments to assess the performance of our proposed test statistic in finite samples. We consider a bivariate specification of \mathbf{X}_t in (1), where the first column of \mathbf{X}_t is a vector of ones. We refer to the first element of $\boldsymbol{\theta}_t$ as α_t and to the second element as β_t . All simulations are based on the same common structure from Chen and Hong (2012) where

$$Y_t = \alpha_t + \beta_t X_t + \varepsilon_t \tag{7}$$

$$X_t = 0.5X_{t-1} + \eta_t \tag{8}$$

$$\varepsilon_t = \sqrt{h_t} u_t \tag{9}$$

We carry out simulations for six different specifications of α_t and β_t in (7). These specifications are illustrated in Figure 2 and presented in detail below. The first panel of Figure 2 represents a DGP where the null hypothesis is true, and the model can reasonably be estimated by OLS. The top right panel and the two middle panels, respectively, represent single, multiple, and temporal structural breaks, and are very important in a financial context. The bottom left panel represents a DGP with smoothly changing elements of $\boldsymbol{\theta}_t$. Finally, the bottom right panel represents a realization of a DGP where the elements of $\boldsymbol{\theta}_t$ follow independent unit root processes.

⁵Matlab codes for the stationary bootstrap and block length selection are kindly made available by Andrew Patton on his website, <http://public.econ.duke.edu/~ap172/>.

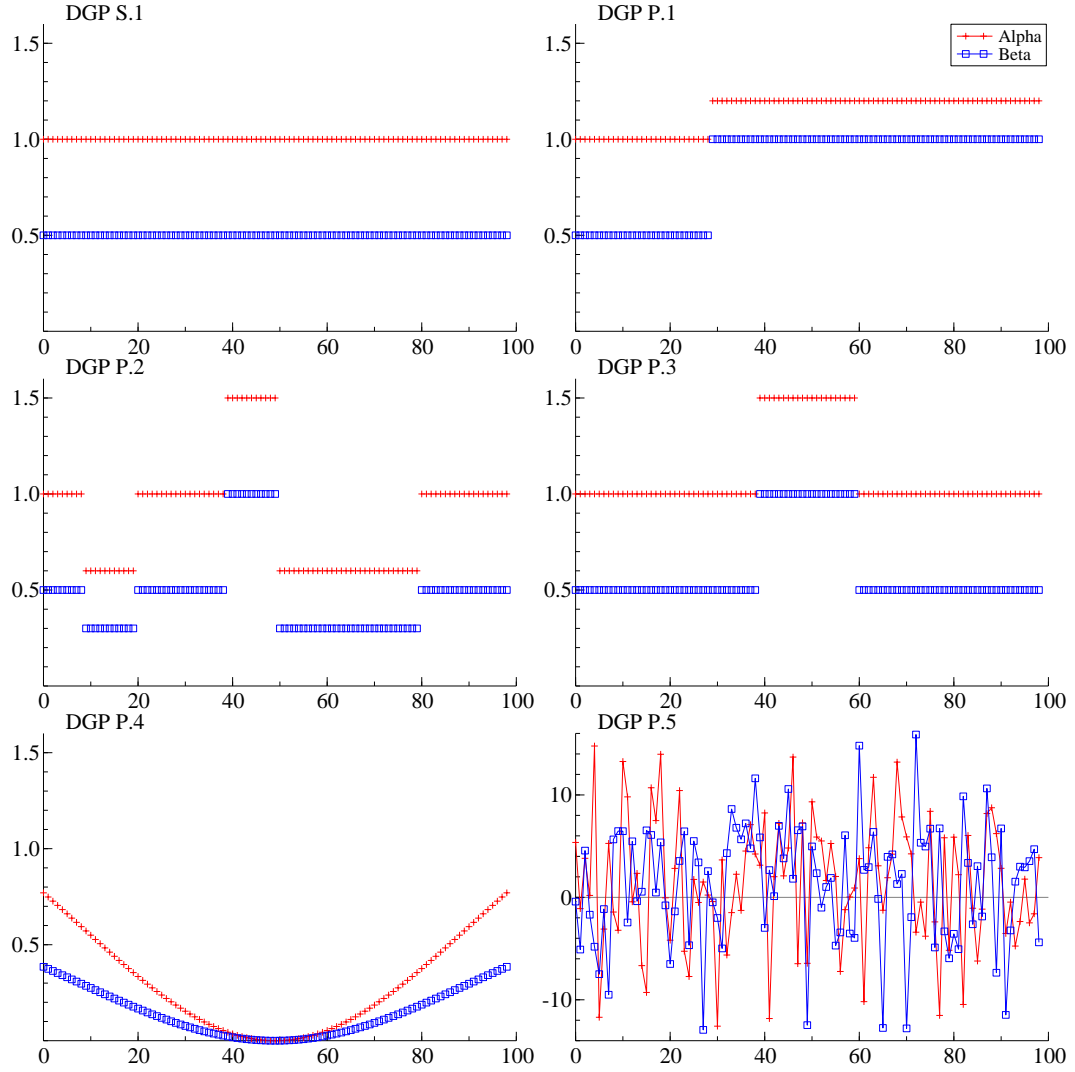


Figure 2: Parameters of the DGP. The figure presents the different specifications of α_t and β_t considered in the simulation study. The plots are constructed based on a sample size of 100 observations. The bottom right panel presents one potential realization of α_t and β_t with the unit root specification of DGP.P5.

The six different specifications of α_t and β_t are identical to those of Chen and Hong (2012) and formally defined as follows.

Constant parameters

$$Y_t = 1 + 0.5X_t + \varepsilon_t \quad (\text{DGP S.1})$$

Single structural break

$$Y_t = \begin{cases} 1 + 0.5X_t + \varepsilon_t, & t \leq 0.3T \\ 1.2 + X_t + \varepsilon_t, & \text{Otherwise} \end{cases} \quad (\text{DGP P.1})$$

Multiple structural breaks

$$Y_t = \begin{cases} 0.6 + 0.3X_t + \varepsilon_t, & 0.1T \leq t \leq 0.2T \text{ and } 0.7T \leq t \leq 0.8T \\ 1.5 + X_t + \varepsilon_t, & 0.4T \leq t \leq 0.5T \\ 1.0 + 0.5X_t + \varepsilon_t, & \text{Otherwise} \end{cases} \quad (\text{DGP P.2})$$

Non-persistent temporal structural breaks

$$Y_t = \begin{cases} 1 + 0.5X_t + \varepsilon_t, & t \leq 0.4T \text{ and } t \geq 0.6T, \\ 1.5 + X_t + \varepsilon_t, & \text{Otherwise} \end{cases} \quad (\text{DGP P.3})$$

Smooth structural changes

$$Y_t = F(\tau)(1 + 0.5X_t) + \varepsilon_t \quad (\text{DGP P.4})$$

$$F(\tau) = 1.5 - 1.5 \exp(-3(\tau - 0.5)^2), \quad \tau = \frac{t}{T}.$$

Unit root in parameters

$$Y_t = \rho_{1,t} + \rho_{2,t}X_t + \varepsilon_t \quad (\text{DGP P.5})$$

$$\rho_{i,t} = \rho_{i,t-1} + u_{i,t}, \quad u_{i,t} \sim \text{NID}(0, 1/15), \quad i = 1, 2.$$

We assume that $\eta_t \sim \text{NID}(0, 1)$ in (8), and in (9) we consider four different specifications of h_t and u_t .

Case 1: $h_t = 1$ and $u_t \sim \text{NID}(0, 1)$, so that ε_t is normal IID.

Case 2: $h_t = 0.2 + 0.8\varepsilon_{t-1}^2$ and $u_t \sim \text{NID}(0, 1)$, so that ε_t contains ARCH(1) effects.

Case 3: $h_t = 0.2 + 0.5X_t^2$ and $u_t \sim \text{NID}(0, 1)$, so that $\varepsilon_t | \mathbf{X}_t \sim \text{N}(0, f(\mathbf{X}_t))$, with $f(\mathbf{X}_t) = 0.2 + 0.5X_t^2$.

Case 4: $h_t = 0.2 + 0.8\varepsilon_{t-1}^2$ and $u_t = 0.9u_{t-1} + \xi_t$, where $\xi_t \sim \text{NID}(0, 0.2^2)$, so that ε_t is heteroskedastic and autocorrelated.

Cases 1-3 are similar to the specifications in Chen and Hong (2012), while case 4 is similar to one considered in Cai (2007).⁶ Note, that Chen and Hong (2012) consider only the IID specification, Case 1, in their simulation study of (DGP P.1)-(DGP P.5), whereas we consider all four cases for all six specifications of α_t and β_t .

In addition to the results for W_{\max} , we also present results for the heteroskedasticity consistent version of the test of Chen and Hong (2012), \hat{H}_{het} . The test is defined as

$$\hat{H}_{\text{het}} = \frac{\sqrt{h}T\hat{Q} - \hat{A}_H}{\sqrt{\hat{B}_H}},$$

where $\hat{Q} = T^{-1} \sum_{t=1}^T (\mathbf{X}'_t \boldsymbol{\theta}_t - \mathbf{X}'_t \boldsymbol{\phi})^2$, and \hat{A}_H and \hat{B}_H are defined in Chen and Hong (2012). In Chen and Hong (2012) the authors rely on a technique called reflection around the boundaries to make the behavior of the estimator at the boundary points similar to the behavior at interior points. We do not use this technique in the present simulation study. Inference for \hat{H}_{het} is based on bootstrap distributions obtained from the nonparametric residuals, as suggested by the authors.

3.1 Results

In order to ensure comparability, we use the uniform kernel and the rule-of-thumb bandwidth used in Chen and Hong (2012), meaning that the bandwidth is chosen as $h = (1/\sqrt{12})T^{-1/5}$. For Cases 2 and 3 the calculation of the test statistic is based on White standard errors, while no correction is used in Case 1. The calculation of the test statistic in Case 4 is based on Newey-West standard errors. Empirical rejection rates for (DGP S.1) at the ten percent level based on 5000 simulations and 499 bootstraps are presented in Table 1. The empirical rejection rates for (DGP S.1) from Cases 1-3 are very similar to the results presented in Chen and Hong (2012). W_{\max} and \hat{H}_{het} both perform satisfactorily in the first three cases. For the smaller sample sizes of 100 and 250 observations, W_{\max} performs somewhat better than \hat{H}_{het} . For 500 observations, both produce rejection rates relatively close to the nominal level of 10%. Results for Case 4 are very different. Rejection rates are further from nominal levels than for Cases 1-3. The test of Chen and Hong (2012) is not designed for autocorrelation, and, unsurprisingly, \hat{H}_{het} is severely size distorted in this case. The empirical rejection rates of W_{\max} for Case 4 are satisfactorily close to the nominal level, but more observations are required to achieve results comparable to

⁶We consider a greater amount of heteroskedasticity in Case 2. The parameter is 0.8 in this analysis and 0.5 in Chen and Hong (2012)

Cases 1-3. Based on the results in Table 1 we limit our analysis of (DGP P.1)-(DGP P.5) in the following to the W_{\max} statistic.

Table 1: Rejection rates of W_{\max} at the 10% level.

| | Case 1 | | | Case 2 | | | Case 3 | | | Case 4 | | |
|------------------------|---------|------|------|--------|------|------|--------|------|------|--------|------|------|
| | 100 | 250 | 500 | 100 | 250 | 500 | 100 | 250 | 500 | 100 | 250 | 500 |
| | DGP S.1 | | | | | | | | | | | |
| W_{\max} | 9.9 | 9.6 | 10.1 | 8.8 | 9.8 | 9.8 | 9.3 | 9.9 | 9.7 | 20.5 | 16.5 | 12.9 |
| \hat{H}_{het} | 17.7 | 13.6 | 12.3 | 17.0 | 12.9 | 11.8 | 18.1 | 14.8 | 12.7 | 96.4 | 89.5 | 74.4 |

Empirical rejection rates at the ten percent level (multiplied by 100) of W_{\max} and \hat{H}_{het} from 5000 simulations from (DGP S.1). The test statistic is based on White standard errors in Cases 2 and 3 and Newey-West standard errors in Case 4. Inference is based on 499 bootstrap replications.

Rejection rates for (DGP P.1)-(DGP P.5) are presented in Table 2 and based on 499 bootstrap replications and 1000 simulations. The empirical rejection rates for Case 1 are very similar to Chen and Hong (2012) for (DGP P.1)-(DGP P.5).⁷ The test suffers from low power in the case of multiple structural breaks, (DGP P.2), when the residuals contain autocorrelation, but apart from this case, power is, in general, good for samples of 500 observations.

Table 2: Rejection rates of W_{\max} at the 10% level.

| | Case 1 | | | Case 2 | | | Case 3 | | | Case 4 | | |
|---------|--------|------|------|--------|------|------|--------|------|------|--------|------|------|
| | 100 | 250 | 500 | 100 | 250 | 500 | 100 | 250 | 500 | 100 | 250 | 500 |
| DGP P.1 | 43.5 | 79.8 | 97.4 | 43.4 | 90.9 | 99.4 | 23.6 | 54.0 | 85.6 | 51.6 | 89.8 | 97.0 |
| DGP P.2 | 44.2 | 86.7 | 99.3 | 38.0 | 86.6 | 99.5 | 28.4 | 72.2 | 98.0 | 18.0 | 27.0 | 47.9 |
| DGP P.3 | 50.6 | 90.6 | 99.9 | 48.3 | 90.5 | 98.9 | 32.0 | 73.4 | 97.9 | 60.3 | 88.1 | 96.0 |
| DGP P.4 | 42.3 | 79.3 | 97.3 | 44.2 | 90.1 | 99.4 | 33.8 | 75.4 | 98.3 | 42.0 | 61.2 | 77.3 |
| DGP P.5 | 99.3 | 100 | 100 | 96.3 | 100 | 100 | 95.2 | 99.9 | 100 | 61.9 | 92.4 | 98.6 |

Empirical rejection rates at the ten percent level (multiplied by 100) of W_{\max} from 1000 simulations from (DGP P.1)-(DGP P.5). The test statistic is based on White standard errors in Cases 2 and 3 and Newey-West standard errors in Case 4. Inference is based on 499 bootstrap replications.

3.1.1 Bandwidth Selection

In the simulation exercise and in the empirical application we rely on a rule-of-thumb bandwidth for the nonparametric estimator. Very strong cases for data-driven bandwidth selection are made by Loader (1999) and Li and Racine (2004). We choose the rule-of-thumb specification in spite of its undesirable properties to ensure comparability with the simulation study in Chen and Hong (2012). Chen and Hong (2012) show that their results are not affected by changing the bandwidth.

⁷Chen and Hong (2012) consider rejections based on empirical critical values (ECV) in their power study, (DGP P.1)-(DGP P.5). In this case the 5%-level ECV is the 95% quantile of the distribution of the test statistic obtained from the 5000 simulations of (DGP S.1). We base our inference on bootstrap methods as discussed above.

4 Empirical applications

This section presents five empirical applications based on important financial and macroeconomic theories. The first application is the long-run CAPM of Ang and Chen (2007), the second is a test of the policy rule proposed in Taylor (1993), the third is a return predictability model from Welch and Goyal (2008), the fourth is the five factor model from Fama and French (2015), and finally the fifth application is a test of Okun’s law.

For empirical applications, we rely on the Breusch-Godfrey test to detect autocorrelation and the Breusch-Pagan test for heteroskedasticity. W_{\max} is based on White standard errors when the null hypothesis of homoskedasticity is rejected and on Newey-West standard errors when the null hypothesis of no autocorrelation is rejected.

4.1 CAPM in the long run

Ang and Chen (2007) introduce a conditional CAPM with time-varying betas that can account for the long-run spread in the average returns of book-to-market sorted portfolios. The model is motivated in part by the graphical inspection of 60 months rolling window OLS estimates of the market beta of two different portfolios as presented in Figure 1. The analysis is based on the book-to-market sorted decile portfolios of Davis (1994) and Davis et al. (2000), which together with market return and the T-bill rate are available from the data library of Kenneth French.⁸ The rolling window estimates of Ang and Chen (2007) are based on the following model

$$R_{i,t} = \alpha_i + \beta_i R_{M,t} + \varepsilon_{i,t}, \quad i = \{value, growth\}, \quad (10)$$

where $R_{i,t}$ is the excess return of portfolio i at time t , $R_{M,t}$ is the excess market return and $\varepsilon_{i,t}$ is an error term.

Figure 3 presents four panels. The top panels contain the full sample OLS estimates in blue, the rolling window OLS estimates in red and the nonparametric estimates of α_i in (10) in green, for the value portfolio on the left hand side and for growth on the right hand side, respectively. The bottom panels present the corresponding estimates of β in (10) from the same three estimation techniques. The estimates of α and β vary substantially for the portfolio of value stocks, while the variation is smaller for the portfolio of growth stocks.

It is interesting that the rolling window estimates are more volatile than the nonparametric ones. Even though the two methods are similar, they rely on a different kernel specification. The rolling window estimation can be seen as a one-sided uniform kernel based on 60 observations

⁸http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

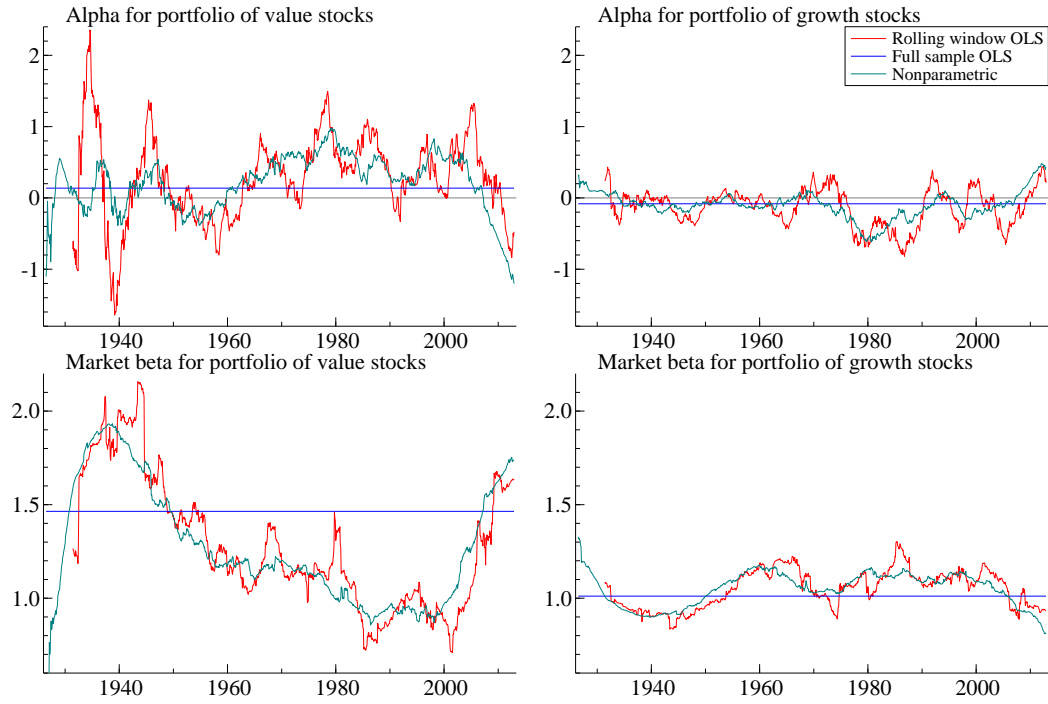


Figure 3: Time-varying parameters. The figure presents the estimates from the regression model, $R_{i,t} = \alpha_i + \beta_i R_{M,t} + \varepsilon_{i,t}$, where $R_{i,t}$ is the excess return on a portfolio, $i = value, growth$, $R_{M,t}$ is the excess return on the market portfolio and $\varepsilon_{i,t}$ is the residual. The top panels present the estimates of α of a portfolio of value stocks on the left hand side and a portfolio of growth stocks on the right hand side. The bottom panels present the estimates of β . The nonparametric estimates are presented with green lines, the full sample estimates are presented in blue, and 60 month rolling window estimates are presented in red. The estimates are based on a sample of monthly returns from July 1926 to July 2016.

and a constant approximation of the unknown parameters, while the nonparametric estimator is based on a linear approximation using a two-sided uniform kernel with a bandwidth of $(1/12)^{1/2} * (T)^{-1/5}$, meaning that the estimator uses 154 observations (two times bandwidth) in this particular application.

The graphical analysis in Figure 3 is complemented by a formal test, and the P values for W_{\max} are presented in Table 3. The table contains results for the portfolio of value stocks, the portfolio of growth stocks and the difference between the two. The difference between the two is called the book-to-market strategy in Ang and Chen (2007). The table also contains the P values of the Breusch-Godfrey test, which indicates that the presence of autocorrelation cannot be rejected in most cases. At a five percent level we reject the linear model with constant parameters for the portfolios of value and for the book-to-market strategy, while the null is rejected for the portfolio of growth stocks at a ten percent level.

Table 3: Empirical results for the long run CAPM.

| | 1926:07-2016:07 | | 1926:07-1963:06 | | 1963:07-2016:07 | |
|--------------|-----------------|--------|-----------------|--------|-----------------|--------|
| | W_{\max} | BG | W_{\max} | BG | W_{\max} | BG |
| Value | 0.0040 | 0.0000 | 0.0180 | 0.0171 | 0.0100 | 0.0000 |
| Growth | 0.1022 | 0.0065 | 0.0000 | 0.0048 | 0.0721 | 0.1187 |
| BM | 0.0020 | 0.0021 | 0.0000 | 0.1616 | 0.0140 | 0.0001 |
| Observations | 1083 | | 444 | | 639 | |

The table presents P values from the W_{\max} test in the model: $R_{i,t} = \alpha_i + \beta_i R_{M,t} + \varepsilon_{i,t}$, where $R_{i,t}$ is the excess returns of the value portfolio, the growth portfolio and the difference between the two, respectively and $r_{m,t}$ is the excess return on the market portfolio. P values for W_{\max} are based on 499 bootstrap replications. P values for the Breusch-Godfrey test for autocorrelation are presented in the BG columns. W_{\max} is based on White standard errors when autocorrelation is rejected and on Newey-West standard errors when it cannot be rejected. Results are presented for the full sample from July 1926 through July 2016 and two subsamples.

At a five percent level we reject the linear model with constant parameters for the portfolio of value stocks and the book-to-market strategy for the post-1963 subsample, while we reject the null hypothesis for the portfolio of growth stocks at a ten percent level.

We conclude that in particular for the post-1963 period, these results support the results in Ang and Chen (2007), and justify the use of a model with time-varying parameters to explain the book-to-market effect, which is found to be strongest in this period. Our findings support the results of Ang and Kristensen (2012) and Guo et al. (2017).

4.2 Taylor rule

Taylor (1993) suggests that a good policy rule would cause a central bank to change the federal funds rate when the price level changes or when real income changes. The following simple policy

rule is suggested

$$\begin{aligned} r_t &= p_t + 0.5y_t + 0.5(p_t - 2) + 2, \\ &= 4 + 0.5y_t + 1.5(p_t - 2). \end{aligned} \tag{11}$$

The rule relates the federal funds rate, r_t , to the rate of inflation p_t , and real GDP output gap, y_t . Data for this application are measured at a quarterly frequency and are available from 1954:Q3 to 2016:Q4, resulting in 250 observations. The data are available from the website of the Federal Reserve Bank of St. Louis.⁹ Four data series have been downloaded. r is the effective federal funds rate (FEDFUNDS), using the end of period option for quarterly aggregation. The inflation rate is calculated from the GDP deflator (GDPDEF) as the rate of inflation over the previous year. Finally, y_t is calculated as the percentage deviation of real GDP (GDPC1) to real potential GDP (GDPPOT).

The policy rule represents a linear model, where the parameters are known to the analyst *a priori*. Our proposed test is well suited for such a situation. To test whether the parameters in (11) are constant, we specify the test statistic slightly differently and let the vector $(4, 0.5, 1.5)'$ replace the OLS estimator, $\hat{\phi}$, in (5) and we remove any estimation uncertainty related to $\hat{\phi}$ in the estimate of Ω_t in (6).

Results of the analysis are presented in Table 4. The Breusch-Godfrey test strongly rejects the null hypothesis of no autocorrelation with a P value of 0.0000. Based on 499 bootstrap replications the null hypothesis of constant parameters is strongly rejected with a P value of 0.0000. The period after 1982 is of particular interest and is often awarded extra attention in the literature, see for example Clarida et al. (2000). Results for this period are also presented in Table 4 and confirm the results for the full sample period.

Based on this analysis we conclude that the parameters of the simple version of the Taylor rule in (11) are not constant.

Several papers consider the simple Taylor rule in a different form from that presented in (11), where parameters are not known, but have to be estimated. See for example Ang and Piazzesi (2003), who consider

$$r_t = \beta_0 + \beta_y y_t + \beta_p (p_t - 2) + \varepsilon_t, \tag{12}$$

to establish a baseline before introducing more advanced policy rules. Time variation of policy rules has also been considered by Ang and Bekaert (2002), who estimate a regime-switching model for the short rate. Our test, if the null hypothesis is rejected, can provide justification for

⁹fred.stlouisfed.org

such approaches. Results for our test for the specification in (12) are presented in Table 4. The Breusch-Godfrey test rejects the null hypothesis of no autocorrelation in both the full sample as well as in the subsample. Based on 499 bootstrap samples the null hypothesis of constant parameters is rejected for both samples on a 10% level.

Table 4: Empirical results for the Taylor rule.

| | 1954:Q3-2016:Q4 | | 1983:Q1-2016:Q4 | |
|---|-----------------|--------|-----------------|--------|
| | W_{\max} | BG | W_{\max} | BG |
| Known parameters, $(4, 0.5, 1.5)'$ | 0.0000 | 0.0000 | 0.0040 | 0.0000 |
| Unknown parameters, $(\hat{\alpha}, \hat{\beta}_y, \hat{\beta}_p)'$ | 0.0401 | 0.0000 | 0.0641 | 0.0000 |
| Observations | 250 | | 136 | |

The table presents values from the W_{\max} test in the model: $r_t = \alpha + \beta_y y_t + \beta_p(p_t - 2) + \varepsilon_t$, where r_t is the effective federal funds rate, y_t is real GDP output gap, and p_t is the rate of inflation over the previous year. The first row contains results for the tests, where α , β_y , and β_p are assumed known and equal to 4, 0.5, and 1.5, respectively. The second row contains results for the case in which α , β_y , and β_p are unknown and must be estimated. P values for W_{\max} are based on 499 bootstrap replications. P values for the Breusch-Godfrey test for autocorrelation are presented in the BG columns. W_{\max} is based on White standard errors when autocorrelation is rejected, and on Newey-West standard errors when it cannot be rejected. Results are presented for the full sample from third quarter of 1954 through last quarter of 2016 and for a subsample.

The rejection of the Taylor rules in (11) and (12) are relevant for several papers in the macro-finance and macroeconomic literatures. Among many others Clarida et al. (2000), Ang and Piazzesi (2003), Christiano et al. (2005), and Smets and Wouters (2007) all rely on the assumption of a policy rule that is linear and has constant parameters. The policy rules in these papers are more complicated than the ones tested here. The results in Table 4 support the use of alternative models compared to the simple ones considered in this application, but do not guarantee that such alternative models are well specified. It would be interesting to test whether these models are well specified and to investigate the effects of potential misspecification, but this lies beyond the scope of this small application.

4.3 Return prediction

When it comes to predicting the equity premium, many variables have been suggested in the literature. Welch and Goyal (2008) carry out an extensive study, where they test, one by one, whether a large set of economic variables have any predictive power. Their study is centered around the regression model

$$r_{t+1} = \gamma_0 + \gamma_1 X_t + \varepsilon_{t+1}, \quad (13)$$

which relates the future value of the equity premium, r_{t+1} , to the current value of a particular variable, X_t . We limit our attention to a single predictor that has been the focus of much attention, namely the dividend yield.

The data for this application are measured at monthly frequency and are available from January 1927 to December 2012.¹⁰ Following Welch and Goyal (2008), the equity premium is defined as the difference in the log return on the S&P 500 and the log of the Treasury-bill rate (lagged one month). The Dividend Yield is defined as the difference between the log of dividends and the log of the S&P 500 index (lagged one month). The sample size is 1066 observations and, based on the Breusch-Godfrey test, the null hypothesis of no autocorrelation is strongly rejected with a P value of 0.0002. The test of constant parameters is therefore based on the Newey-West estimator. The P value, resulting from 499 bootstrap replications, leads to rejection of the linear model with constant parameters with a P value of 0.0120.

Our results support, among others, the results of Paye and Timmermann (2006), Lettau and Nieuwerburgh (2008), and Farmer et al. (2018), where stability of (13) is also rejected. The rejection of the model with constant parameters is of particular relevance to studies such as Dangl and Halling (2012), who allow the parameters of (13) to vary over time. The results from our analysis justify this approach.

4.4 A five-factor asset pricing model

Fama and French (2015) introduce a new asset pricing model relying on two additional factors compared to the celebrated model of Fama and French (1993). The five factor model is specified in the following way:

$$R_{i,t} = \alpha_i + \beta_{1,i}R_{M,t} + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}RMW_t + \beta_{5,i}CMA_t + \varepsilon_{i,t}, \quad (14)$$

where $R_{i,t}$ and $R_{M,t}$ are respectively the return on a portfolio, i , and the return on the market portfolio, M , in excess of the risk-free return. SMB_t is the return on a portfolio of small stocks minus the return on a portfolio of big stocks. HML_t is the orthogonal HML_t factor. HML_t is the return on a portfolio of high B/M stocks minus the return on a portfolio of low B/M stocks. The orthogonal HML_t factor is constructed as the sum of the intercept and the residuals of the regression of HML_t on the other four factors. RMW_t is the difference in returns on portfolios of stocks with robust and weak profitability, and CMA_t is the difference in return on stocks of conservative and aggressive firms. $\varepsilon_{i,t}$ is a disturbance with zero mean. The factors are available at the monthly frequency, and the sample used in Fama and French (2015) contains 606 observations, with the sample period starting in July of 1963 and ending in December 2013.¹¹

¹⁰The data are available from the website of Amit Goyal, <http://www.hec.unil.ch/agoyal/>.

¹¹The data are available from the website of Kenneth French, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Fama and French (2015) run the regression in (14) for a large number of different portfolios. In this example we limit ourselves to a single portfolio. 25 Size-B/M portfolios are considered in Table 7 of Fama and French (2015). Figure 4 plots the nonparametric estimates of the time-varying intercepts of (14) for a portfolio of big stocks with high B/M in green and compares the estimates to the full sample OLS estimates in blue. This particular portfolio is chosen because the authors cannot reject the null hypothesis of zero intercept in this case, indicating a well-specified asset-pricing model. Figure 4 illustrates that, although full sample OLS estimates might indicate a well-specified asset-pricing model, it might be on the basis of an unrealistic assumption of constant parameters.

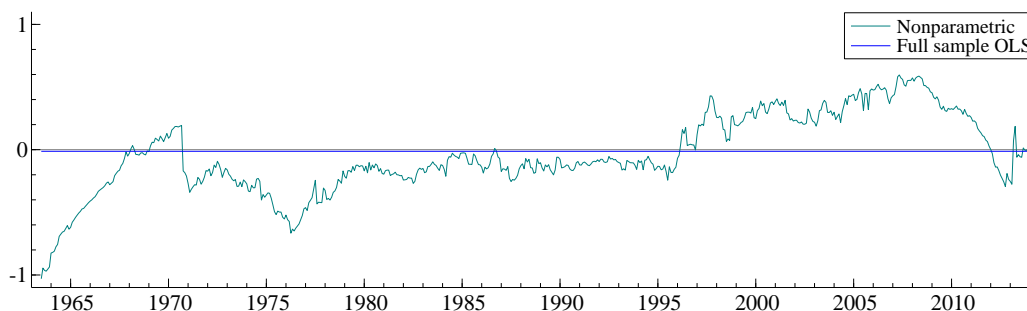


Figure 4: Time-varying alpha. The plot presents the estimates of α in the model: $R_{i,t} = \alpha_i + \beta_{i,1}R_{M,t} + \beta_{i,2}SMB_t + \beta_{i,3}HML_t + \beta_{i,4}RMW_t + \beta_{i,5}CMA_t + \varepsilon_{i,t}$ for the Big,High portfolio. Nonparametric estimates are presented in green and the full sample OLS estimates are presented in blue. Estimates are presented for the full sample of 606 monthly observations from the period 1963:7-2013:12.

The graphical analysis in Figure 4 motivates the formal testing of the linear model with constant parameters. With a P value of 0.526 from the Breusch-Godfrey test the null hypothesis of no autocorrelation cannot be rejected. Therefore the test for constant parameters is based on the White correction, and the null of constant parameters is strongly rejected with a P value of 0.0000 for 499 bootstrap replications.

4.5 Okun's law

This section investigates a relationship between changes in GDP and changes in the unemployment rate, known as Okun's law. If one assumes that the potential output and the potential unemployment rate are both constant, then Okun's law implies that

$$\frac{\Delta Y_t}{Y_t} = k - c\Delta u_t + \varepsilon_t,$$

where Y_t is real GDP, u_t is the unemployment rate and ε_t is an error term. Data for the U.S. real GDP and the unemployment rate are downloaded from the website of the Federal Reserve Bank of St. Louis.¹²

Quarterly percentage changes in real GDP are obtained for the period from 1948:Q2 to 2016:Q3. The Breusch-Godfrey test has a P value of 0.019, indicating that the data contain some autocorrelation and W_{\max} is therefore based on the Newey-West estimator. Based on 499 bootstrap replications the linear model with constant parameters is rejected at a 10% level with a P value of 0.084.

We conclude that our test provides some evidence against this version of Okun's Law.

5 Conclusion

This paper provides a simple test that allows researchers to judge the appropriateness of a linear model specification with constant parameters. The test is tailored for applications in finance and macroeconomics and allows for autocorrelation in the disturbances. In a large simulation study, the proposed test is shown to have empirical rejection rates very close to the nominal level when the data are generated by the model under the null hypothesis. Furthermore, the test has good power against various realistic alternatives. We consider five empirical applications and demonstrate that the test is applicable in many different areas, and that valuable insights can be obtained from the proposed test.

¹²fred.stlouisfed.org

References

- Andrews, D., 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.
- Ang, A., Bekaert, G., 2002. Regime switching in interest rates. *Journal of Business & Economics Statistics* 20 (2), 163–182.
- Ang, A., Chen, J., 2007. Capm over the long run: 1926-2001. *Journal of Empirical Finance* 14, 1–40.
- Ang, A., Kristensen, D., 2012. Testing conditional factor models. *Journal of Financial Economics* 106, 132–156.
- Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50, 745–787.
- Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.
- Brown, R., Durbin, J., Evans, J., 1975. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society. Series B (Methodological)* 37 (2), 149–192.
- Cai, Z., 2007. Trending time-varying coefficient time series models with serially correlated errors. *Journal of Econometrics* 136, 163–188.
- Chen, B., Hong, Y., 2012. Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica* 80, 1157–1183.
- Christiano, L. J., Eichenbaum, M., Evans, C. L., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113 (1), 1–45.
- Clarida, R., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics* 115 (1), 147–180.
- Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. *Journal of Financial Economics* 106, 157–181.
- Davis, J., 1994. The cross-section of realized stock returns: the pre-compustat evidence. *Journal of Finance* 49 (5), 1579–1593.

- Davis, J., Fama, E., French, K., 2000. Characteristics, covariances, and average returns: 1929-1997. *Journal of Finance* 55 (1), 389–406.
- Delgado, M. A., Manteiga, W. G., 2001. Significance testing in nonparametric regression based on the bootstrap. *The Annals of Statistics* 29 (5), 1469–1507.
- Du, P., Parmeter, C. F., Racine, J. S., 2013. Nonparametric kernel regression with multiple predictors and multiple shape constraints. *Statistica Sinica* 23, 1347–1371.
- Elliott, G., Müller, U., 2006. Efficient tests for general persistent time variation in regression coefficients. *Review of Economic Studies* 73, 907–940.
- Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.
- Fama, E., French, K., 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116, 1–22.
- Farmer, L., Schmidt, L., Timmermann, A., 2018. Pockets of predictability. Unpublished working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3152386.
- Guo, H., Wu, C., Yu, Y., 2017. Time-varying beta and the value premium. *Journal of Financial and Quantitative Analysis* 52 (4), 1551–1576.
- Johannes, M. S., Korteweg, A. G., Polson, N., 2014. Sequential learning, predictability, and optimal portfolio returns. *Journal of Finance* 69, 611–644.
- Lettau, M., Nieuwerburgh, S. V., 2008. Reconciling the return predictability evidence. *The Review of Financial Studies* 21 (4), 1607–1652.
- Li, Q., Racine, J., 2004. Cross-validated local linear nonparametric regression. *Statistica Sinica* 14, 485–512.
- Li, Q., Racine, J. S., 2007. *Nonparametric Econometrics: Theory and Practice*, 1st Edition. Princeton University Press, Princeton and Oxford.
- Lin, C., Teräsvirta, T., 1994. Testing the constancy of regression parameters against continuous structural change. *Journal of Econometrics* 62, 211–228.
- Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47, 13–37.

- Loader, C., 1999. Bandwidth selection: Classical or plug-in. *The Annals of Statistics* 27 (2), 415–438.
- Orbe, S., Ferreira, E., Rodriguez-Póo, J., 2000. A nonparametric method to estimate time varying coefficients under seasonal constraints. *Journal of Nonparametric Statistics* 12, 779–806.
- Orbe, S., Ferreira, E., Rodriguez-Póo, J., 2005. A nonparametric method to estimate time varying parameters under shape restrictions. *Journal of Econometrics* 126, 53–77.
- Patton, A., Politis, D., White, H., 2009. Correction to "automatic block-length selection for the dependent bootstrap" by d. politis and h. white. *Econometric Reviews* 28 (4), 372–375.
- Patton, A., Timmermann, A., 2010. Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts. *Journal of Financial Economics* 98, 605–625.
- Paye, B. S., Timmermann, A., 2006. Instability of return prediction models. *Journal of Empirical Finance* 13, 274–315.
- Perron, P., 2006. Dealing with structural breaks. volume 1: Econometric theory. In: Hassani, H., Mills, T. C., Patterson, K. (Eds.), *Palgrave Handbook of Econometrics*. Palgrave Macmillian, Ch. 8, pp. 278–358.
- Pesaran, M. H., Timmermann, A., 2002. Market timing and return prediction under model instability. *Journal of Empirical Finance* 9, 495–510.
- Politis, D., Romano, J., 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.
- Politis, D., White, H., 2004. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23 (1), 53–70.
- Robinson, P., 1989. Nonparametric estimation of time-varying parameters. In: Hackl, P. (Ed.), *Statistical Analysis and Forecasting of Economic Structural Change*. Springer, pp. 253–264.
- Robinson, P., 1991. Time-varying nonlinear regression. In: Hackl, P. (Ed.), *Statistical Analysis and Forecasting of Economic Structural Change*. Berlin: Springer, pp. 179–190.
- Ross, S. A., 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13 (3), 341–360.
- Sharpe, W., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19, 425–442.

- Smets, F., Wouters, R., 2007. Shocks and frictions in us business cycles: A bayesian dsge approach. *The American Economic Review* 97 (3), 586–606.
- Stock, J., Watson, M., 2002. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97 (460), 1167–1179.
- Taylor, J. B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.
- Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. *The review of financial studies* 21 (4), 1455–1508.

A Nonparametric Regression

This appendix presents a reformulation of the least squares problem and the corresponding Local Linear estimator, which is very practical for implementing the test.

The time-varying parameter model is specified as in (1). As mentioned in the main text, at any point in time $t \in [0, 1]$, the Local Linear model can be viewed as the least squares estimator of the linear regression model

$$k_{st}^{1/2} Y_s = k_{st}^{1/2} \mathbf{X}_s \gamma_{0,t} + k_{st}^{1/2} (s - t) \mathbf{X}_s \gamma_{1,t} + k_{st}^{1/2} \varepsilon_s, \quad \text{for } s = \frac{1}{T}, \dots, 1,$$

where $k_{st} = k\left(\frac{s-t}{h}\right)/h$. k is a kernel function and h is the bandwidth. It is convenient to reformulate the model. Let $u_{st} = k_{st}^{1/2} Y_s$, $\mathbf{W}_{st} = k_{st}^{1/2} \mathbf{X}_s$, $\mathbf{Z}_{st} = k_{st}^{1/2} (s - t) \mathbf{X}_s$ and $\eta_{st} = k_{st}^{1/2} \varepsilon_s$, which gives

$$u_{st} = \mathbf{W}_{st} \gamma_{0,t} + \mathbf{Z}_{st} \gamma_{1,t} + \eta_{st} \quad \text{for } s = \frac{1}{T}, \dots, 1.$$

or equivalently, in vector form

$$\mathbf{u}_t = \mathbf{W}_t \gamma_{0,t} + \mathbf{Z}_t \gamma_{1,t} + \boldsymbol{\eta}_t. \quad (15)$$

The estimates of $\gamma_{0,t}$ are estimates of $\boldsymbol{\theta}(t)$, whereas the estimates of $\gamma_{1,t}$ are estimates of $\boldsymbol{\theta}'(t)$. We are interested in the estimates of $\gamma_{0,t}$ only. Therefore, we define $\mathbf{M}_Z = \mathbf{I} - \mathbf{Z}_t (\mathbf{Z}_t' \mathbf{Z}_t)^{-1} \mathbf{Z}_t'$ and premultiply by this matrix in (15) to obtain

$$\mathbf{M}_Z \mathbf{u}_t = \mathbf{M}_Z \mathbf{W}_t \gamma_{0,t} + \mathbf{M}_Z \boldsymbol{\eta}_t. \quad (16)$$

The least squares estimator of $\gamma_{0,t}$ is

$$\hat{\gamma}_{0,t} = (\mathbf{W}_t' \mathbf{M}_Z \mathbf{W}_t)^{-1} \mathbf{W}_t' \mathbf{M}_Z \mathbf{u}_t,$$

B Estimation of Ω

Following the calculations in Appendix A we can find the variance of the difference between the nonparametric estimator and the OLS estimator as

$$\begin{aligned}
\Omega_t &= V(\hat{\theta}_t - \hat{\phi}) = V(\hat{\theta}_t) + V(\hat{\phi}) - cov(\hat{\theta}_t, \hat{\phi}) - cov(\hat{\theta}_t, \hat{\phi})' \\
&= (\mathbf{W}_t' \mathbf{M}_Z \mathbf{W}_t)^{-1} \mathbf{W}_t' \mathbf{M}_Z \mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\eta}_t') \mathbf{M}_Z \mathbf{W}_t (\mathbf{W}_t' \mathbf{M}_Z \mathbf{W}_t)^{-1} \\
&\quad + (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \\
&\quad - (\mathbf{W}_t' \mathbf{M}_Z \mathbf{W}_t)^{-1} \mathbf{W}_t' \mathbf{M}_Z \mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\varepsilon}_t') \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \\
&\quad - (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\eta}_t') \mathbf{M}_Z \mathbf{W}_t (\mathbf{W}_t' \mathbf{M}_Z \mathbf{W}_t)^{-1}.
\end{aligned} \tag{17}$$

$\hat{\Omega}_t$ can be found by replacing $\mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\eta}_t')$, $\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$, $\mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\varepsilon}_t')$ and $\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\eta}_t')$ with suitable (autocorrelation and) heteroskedasticity-robust estimators.

Research Papers

2018



- 2018-05: Torben G. Andersen, Martin Thyrsgaard and Viktor Todorov: Time-Varying Periodicity in Intraday Volatility
- 2018-06: Niels Haldrup and Carsten P. T. Rosenskjold: A Parametric Factor Model of the Term Structure of Mortality
- 2018-07: Torben G. Andersen, Nicola Fusari and Viktor Todorov: The Risk Premia Embedded in Index Options
- 2018-08: Torben G. Andersen, Nicola Fusari and Viktor Todorov: Short-Term Market Risks Implied by Weekly Options
- 2018-09: Torben G. Andersen and Rasmus T. Varneskov: Consistent Inference for Predictive Regressions in Persistent VAR Economies
- 2018-10: Isabel Casas, Xiuping Mao and Helena Veiga: Reexamining financial and economic predictability with new estimators of realized variance and variance risk premium
- 2018-11: Yunus Emre Ergemen and Carlos Velasco: Persistence Heterogeneity Testing in Panels with Interactive Fixed Effects
- 2018-12: Hossein Asgharian, Charlotte Christiansen and Ai Jun Hou: Economic Policy Uncertainty and Long-Run Stock Market Volatility and Correlation
- 2018-13: Emilio Zanetti Chini: Forecasting dynamically asymmetric fluctuations of the U.S. business cycle
- 2018-14: Cristina Amado, Annastiina Silvennoinen and Timo Teräsvirta: Models with Multiplicative Decomposition of Conditional Variances and Correlations
- 2018-15: Changli He, Jian Kang, Timo Teräsvirta and Shuhua Zhang: The Shifting Seasonal Mean Autoregressive Model and Seasonality in the Central England Monthly Temperature Series, 1772-2016
- 2018-16: Ulrich Hounyo and Rasmus T. Varneskov: Inference for Local Distributions at High Sampling Frequencies: A Bootstrap Approach
- 2018-17: Søren Johansen and Morten Ørregaard Nielsen: Nonstationary cointegration in the fractionally cointegrated VAR model
- 2018-18: Giorgio Mirone: Cross-sectional noise reduction and more efficient estimation of Integrated Variance
- 2018-19: Kim Christensen, Martin Thyrsgaard and Bezirgen Veliyev: The realized empirical distribution function of stochastic variance with application to goodness-of-fit testing
- 2018-20: Ruijun Bu, Kaddour Hadri and Dennis Kristensen: Diffusion Copulas: Identification and Estimation
- 2018-21: Kim Christensen, Roel Oomen and Roberto Renò: The drift burst hypothesis
- 2018-22: Russell Davidson and Niels S. Grønborg: Time-varying parameters: New test tailored to applications in finance and macroeconomics