Predictive Regressions For Macroeconomic Data

Fukang Zhu*, Zongwu Cai† and Liang PENG‡

ABSTRACT

Researchers have constantly asked whether stock returns can be predicted by some macroeconomic data. On the other hand, it is known that macroeconomic data may exhibit non-stationarity and/or heavy tails, which complicates existing testing procedures for predictability. In this paper, we propose novel empirical likelihood methods based on some weighted score equation to test whether the monthly CRSP value-weighted index can be predicted by the log dividend-price ratio or the log earnings-price ratio. The new methods work well both theoretically and empirically regardless of the predicting variables being stationary or non-stationary or having an infinite variance.

Keywords: Autoregressive process, empirical likelihood, long memory process, nearly integrated, predictive regressions, unit root, weighted estimation.

1 Introduction

It is well documented in the literature that predictive regression models have been widely used in economics and finance for the evaluation of the mutual fund performance, the optimization of the asset allocations, the conditional capital asset pricing and others. In particular, it is used to check the predictability of asset returns by various lagged financial and economic variables, such as the log dividend-price ratio, the log earnings-price ratio, the log book-to-market ratio, the dividend yield, the term spread and default premium, the interest rates as well as other financial and state economic variables.

*School of Mathematics, Jilin University, Changchun, Jilin, China.
†Department of Economics, University of Kansas, Lawrence, KS 66045, USA and Wang Yanan Institute for Studies in Economics and Fujian Key Laboratory of Statistical Sciences, Xiamen University, China.
‡Corresponding author. School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA.

Our motivation for this research is trying to answer the question in the financial econometrics literature on whether the monthly CRSP (Center for Research in Security Prices) value-weighted index can be predicted by using the macroeconomic data such as the log dividend-price ratio or the log earnings-price ratio as well as other economic data like interest rates. To answer this question, we need a statistical model. By following the usual convention in the financial econometrics literature, we use the following simple predictive regression model which assumes that observations $\{(X_t, Y_t)\}_{t=1}^n$ follow the following structural model

$$\begin{cases} Y_t = \alpha + \beta X_{t-1} + U_t, \\ X_t = \theta + \phi X_{t-1} + V_t \end{cases} \tag{1}$$

with X_0 being a constant. Here, Y_t denotes a predictable variable, say the asset return like the CRSP value-weighted index, X_t denotes a predicting variable, such as financial instruments like the log dividend-price ratio or the log earnings-price ratio, and $(U_1, V_1), \cdots, (U_n, V_n)$ are independent and identically distributed (iid) innovations with zero means but U_t and V_t might be correlated. Our main purpose of this study is to examine the existence of the predictability of asset returns by some financial variables such as the log dividend-price ratio or the log earnings-price ratio. To achieve our goal, we need to construct a confidence interval for β in (1) or to test the null hypothesis of no predictability ($H_0 : \beta = 0$). The detailed report of analyzing the aforementioned real example is given in Section 4.

The empirical literature on the predictability of asset returns is rather large. In particular, estimating β and testing the null hypothesis of no predictability $H_0 : \beta = 0$ are receiving much attention in the recent literature of financial econometrics. For example, Stambaugh (1999) showed that the least squares estimator for β based on the first equation in (1) is biased in finite sample since the estimation procedure ignores the dependence between U_t and V_t. Since then, several bias-corrected estimation procedures and corresponding hypothesis

tests have been proposed in the literature when the sequence $\{X_t\}$ is stationary (i.e. $|\phi| < 1$) and/or integrated/nearly integrated (i.e., $\phi = 1 - \gamma_\phi/n$ for some $\gamma_\phi \geq 0$). Some references include but not limit to Amihud and Hurvich (2004), Campbell and Yogo (2006), Chen and Deo (2009), Jansson and Moreira (2006), Lewellen (2004), Amihud, Hurvich and Wang (2009), Cai and Wang (2013), and the references therein.

By assuming that the joint distribution of the two innovations (U_t, V_t) in (1) is a bivariate normal, Campbell and Yogo (2006) proposed a new Bonferroni Q-test, based on the infeasible uniform most powerful test and showed that this new test is more powerful than the Bonferroni t-test of Cavanagh, Elliott and Stock (1995) in the sense of Pitman efficiency. However, the normality assumption might not be satisfied for real applications and the implementation of the Bonferroni Q-test can be somewhat complicated, because it requires searching several tables as in Campbell and Yogo (2005), which depend heavily on both the Dickey-Fuller generalized least squares (DF-GLS) statistic and δ being the correlation coefficient between U_t and V_t. Moreover, the theoretical justification of the Bonferroni Q-test given in Campbell and Yogo (2006) heavily depends on the assumptions of known covariance of innovations, known shifts in the model and the predicting variable is non-stationary and has a finite variance. It remains unjustified when these unknown quantities are replaced by some estimators and/or the predicting variable is stationary or has an infinite variance.

Now, the question is how to construct a confidence interval for β or to test whether β equals a given value, say zero without knowing that the predicting variable is stationary or non-stationary or has an infinite variance. Obviously, none of those methods mentioned above works since the asymptotic limit of any one of them depends on whether the predicting variable is stationary or non-stationary or has an infinite variance. Moreover, it is impossible to distinguish these cases without imposing further model assumptions. To illustrate this

difficulty, let us look at the simple least squares estimator of β in (1), given by

$$\hat{\beta}_{LSE} = \frac{n\sum_{t=1}^{n} Y_t X_{t-1} - (\sum_{t=1}^{n} Y_t)(\sum_{t=1}^{n} X_{t-1})}{n\sum_{t=1}^{n} X_{t-1}^2 - (\sum_{t=1}^{n} X_{t-1})^2}.$$

Clearly, $\hat{\beta}_{LSE}$ can be re-expressed as follows

$$\hat{\beta}_{LSE} - \beta = \frac{n\sum_{t=1}^{n} U_t X_{t-1} - (\sum_{t=1}^{n} U_t)(\sum_{t=1}^{n} X_{t-1})}{n\sum_{t=1}^{n} X_{t-1}^2 - (\sum_{t=1}^{n} X_{t-1})^2}.$$

It is known that $n^{-1}\sum_{t=1}^{n} X_{t-1}$ and $n^{-1}\sum_{t=1}^{n} X_{t-1}^2$ do not converge in probability to some constants when the AR(1) process $\{X_t\}$ is integrated/nearly integrated. Therefore, the asymptotic limit of $\hat{\beta}_{LSE}$ is totally different for the stationary and non-stationary cases; see Campbell and Yogo (2006) and Cai and Wang (2013). On the other hand, when $\{X_t\}$ and $\{U_t\}$ are two independent random samples with heavy tails, Samorodnitsky et al. (2007) derived the asymptotic limit of $\hat{\beta}_{LSE}$, which is very complicated too. Therefore, if one wants to construct a confidence interval for β or to test $H_0 : \beta = \beta_0$ for a given value β_0 based on the asymptotic limit of $\hat{\beta}_{LSE}$, one has to distinguish the case between stationarity and non-stationarity, and between finite variance and infinite variance. This seems infeasible in the real implementation. Moreover, even if one can distinguish these cases, it is still a difficult task to obtain critical points by directly estimating or simulating the asymptotic limit when the sequence $\{X_t\}$ is integrated/nearly integrated and/or has an infinite variance. As an alternative way, a bootstrap method may be employed to obtain critical values. However, it is well known in the literature that the full sample bootstrap method is inconsistent for a nearly integrated or infinite variance AR process. Instead, one has to employ the subsample bootstrap method and so to face the difficulty of choosing the subsample size; see Datta (1996) and Hall and Jing (1998) for details.

To overcome the aforementioned difficulties and problems, in this paper, by applying the empirical likelihood method to some weighted score equations, we propose new methods to

construct a confidence interval for β or to test $H_0 : \beta = \beta_0$ without distinguishing whether the sequence $\{X_t\}$ is stationary or non-stationary (integrated or nearly integrated) or has an infinite variance. As a powerful nonparametric likelihood approach, empirical likelihood method has been extended and applied to many different settings including time series models since Owen (1988, 1990) introduced the method. See Owen (2001) for an overview.

The rest of this paper is organized as follows. Section 2 is devoted to presenting the methodologies and some asymptotic results. A simulation study is reported in Section 3, which shows the good finite sample performance of the new methods. The detailed analysis of the monthly CRSP value-weighted index is reported in Section 4 to highlight the practical usefulness of the proposed methods. Section 5 concludes the paper. All theoretical proofs are relegated to the Appendix.

2 Methodology and Asymptotic Properties

First, we consider that observations $\{(X_t, Y_t)\}$ follow the model

$$\begin{cases} Y_t = \beta X_{t-1} + U_t \\ X_t = \theta + \phi X_{t-1} + e_t \\ B(L)e_t = V_t, \end{cases} \tag{2}$$

where $L^i e_t = e_{t-i}$, $B(L) = 1 - (\sum_{i=1}^{p} b_i L^i)$, $B(1) \neq 0$, all the roots of $B(L)$ are fixed and less than one in absolute value, and $(U_1, V_1), \cdots, (U_n, V_n)$ are iid random vectors with zero means.

As shown in Chuang and Chan (2003), the empirical likelihood method fails for non-stationary AR processes in the sense that Wilks theorem does not hold. It is also known that the asymptotic limit of the least squares estimator for ϕ in the second equation of (2) is a stable law rather than a normal distribution when e_t has an infinite variance. Hence, it is expected that Wilks theorem fails for a direct application of the empirical likelihood

method to the score equation via the first equation in (2) when the sequence $\{X_t\}$ is either non-stationary or has an infinite variance.

Recently, Ling (2005) proposed minimizing the weighted least squares $\sum_{t=1}^{n}\{X_t - \theta - \phi X_{t-1}\}^2 w(X_{t-1})$ for some weight function $w(\cdot)$ so as to ensure a normal limit whenever $e_t = V_t$ has a finite or infinite variance. Chan, Li and Peng (2012) combined the weighted idea with the empirical likelihood method to construct a confidence interval for ϕ whenever the sequence $\{X_t\}$ is stationary or nearly integrated, but has a finite variance. Here, we propose using the weighted idea together with the empirical likelihood method to construct a confidence interval for β rather than ϕ regardless of the sequence $\{X_t\}$ being stationary or nearly integrated or having an infinite variance. More specifically, we define the empirical likelihood function for β as

$$L_n(\beta) = \sup\left\{\prod_{t=1}^{n}(np_t) : p_1 \geq 0, \cdots, p_n \geq 0, \sum_{t=1}^{n} p_t = 1, \sum_{t=1}^{n} p_t Z_t(\beta) = 0\right\}, \tag{3}$$

where $Z_t(\beta) = (Y_t - \beta X_{t-1})X_{t-1}/\sqrt{1 + X_{t-1}^2}$. It follows from the Lagrange multiplier technique that

$$l_n(\beta) = -2\log L_n(\beta) = 2\sum_{t=1}^{n}\log\{1 + \lambda Z_t(\beta)\},$$

where $\lambda = \lambda(\beta)$ satisfies

$$\sum_{t=1}^{n}\frac{Z_t(\beta)}{1 + \lambda Z_t(\beta)} = 0.$$

The following theorem shows that Wilks theorem holds for the above proposed empirical likelihood method.

Theorem 1. *Suppose model (2) holds with either $|\phi| < 1$ or $\phi = 1 - \gamma_\phi/n$ for some $\gamma_\phi \geq 0$. Furthermore, we assume that $EU_1 = 0, E|U_1|^{2+q} < \infty$ for some $q > 0$, and the distribution of V_t is in the domain of attraction of a stable law with index $\alpha^* \in (0, 2]$. Then, $l_n(\beta_0)$ converges in distribution to a chi-square limit with one degree of freedom as $n \to \infty$, where β_0 denotes the true value of β.*

Remark 1. *If $EV_t^2 < \infty$, then the distribution of V_t is in the domain of attraction of a stable law with index $\alpha^* = 2$. When the distribution of V_t is in the domain of attraction of a stable law with index $\alpha^* = 2$, EV_t^2 may be infinite, but $E|V_t|^{\gamma^*} < \infty$ for any $\gamma^* \in (0, 2)$. When the distribution of V_t is in the domain of attraction of a stable law with index $\alpha^* \in (0, 2)$, we have $E|V_t|^{\gamma^*} < \infty$ for $\gamma^* < \alpha^*$ and $E|V_t|^{\gamma^*} = \infty$ for $\gamma^* > \alpha^*$. The reader is referred to Feller (1971) for details on stable laws.*

Next, we consider a more general model than (2) by including an intercept for Y_t:

$$\begin{cases} Y_t = \alpha + \beta X_{t-1} + U_t \\ X_t = \theta + \phi X_{t-1} + e_t \\ B(L)e_t = V_t, \end{cases} \tag{4}$$

where $L^i e_t = e_{t-i}$, $B(L) = 1 - (\sum_{i=1}^p b_i L^i)$, $B(1) \neq 0$, all the roots of $B(L)$ are fixed and less than one in absolute value, and $(U_1, V_1), \cdots, (U_n, V_n)$ are iid random vectors. Once again, our observations are $\{(X_t, Y_t)\}_{t=1}^n$.

As before, one may apply the empirical likelihood method to the following estimating equations:

$$\sum_{t=1}^n (Y_t - \alpha - \beta X_{t-1}) = 0 \quad \text{and} \quad \sum_{t=1}^n (Y_t - \alpha - \beta X_{t-1}) X_{t-1} / \sqrt{1 + X_{t-1}^2} = 0.$$

It is clear that when $\{X_t\}$ is integrated/nearly integrated, $n^{-1} \sum_{t=1}^n U_t X_{t-1} / \sqrt{1 + X_{t-1}^2}$ does not converge in probability to a constant. Instead, it converges in distribution. Therefore, the joint limit of $\frac{1}{\sqrt{n}} \sum_{t=1}^n (Y_t - \alpha_0 - \beta_0 X_{t-1})$ and $\frac{1}{\sqrt{n}} \sum_{t=1}^n (Y_t - \alpha_0 - \beta_0 X_{t-1}) X_{t-1} / \sqrt{1 + X_{t-1}^2}$ is no longer a bivariate normal distribution. Hence, Wilks theorem for the above empirical likelihood method fails when $\{X_t\}$ is non-stationary, which is due to the intercept α.

To overcome the above difficulty, one may employ the difference method to get rid of α by using $Y_{t+1} - Y_t$. In such a case, the sequence $\{X_{t+1} - X_t\}$ becomes stationary when $\phi = 1$. Therefore, inferences for β based on the differences become much less efficient with rate $\sqrt{n}$ instead of n when the sequence $\{X_t\}$ is non-stationary. Another issue on applying

the empirical likelihood method based on the difference $Y_{t+1} - Y_t$ is that the new errors $\{U_{t+1} - U_t\}s$ are not independent any more. Here, we propose to split the sample into two parts and then to use the differences with a very large lag to get rid of the intercept before applying the empirical likelihood method. More specifically, put $m = [n/2]$, $\tilde{Y}_t = Y_t - Y_{t+m}$, $\tilde{X}_t = X_t - X_{t+m}$, and $\tilde{U}_t = U_t - U_{t+m}$ for $t = 1, \cdots, m$. Then, we have

$$\tilde{Y}_t = \beta \tilde{X}_{t-1} + \tilde{U}_t \quad \text{for} \quad t = 1, \cdots, m.$$

Based on the above equation, we define the empirical likelihood function for β as

$$\tilde{L}_n(\beta) = \sup \left\{ \prod_{t=1}^{m} (mp_t) : p_1 \geq 0, \cdots, p_m \geq 0, \sum_{t=1}^{m} p_t = 1, \sum_{t=1}^{m} p_t \tilde{Z}_t(\beta) = 0 \right\}, \tag{5}$$

where $\tilde{Z}_t(\beta) = (\tilde{Y}_t - \beta \tilde{X}_{t-1})\tilde{X}_{t-1}/\sqrt{1 + \tilde{X}_{t-1}^2}$. By the Lagrange multiplier technique, we have

$$\tilde{l}_n(\beta) = -2 \log \tilde{L}_n(\beta) = 2 \sum_{t=1}^{m} \log\{1 + \tilde{\lambda} \tilde{Z}_t(\beta)\},$$

where $\tilde{\lambda} = \tilde{\lambda}(\beta)$ satisfies

$$\sum_{t=1}^{m} \frac{\tilde{Z}_t(\beta)}{1 + \tilde{\lambda} \tilde{Z}_t(\beta)} = 0.$$

The following theorem shows that Wilks theorem holds for the above proposed empirical likelihood method.

Theorem 2. *Under conditions of Theorem 1, $\tilde{l}_n(\beta_0)$ converges in distribution to a chi-square distribution with one degree of freedom as $n \to \infty$, where β_0 denotes the true value of β.*

Based on the above theorems, an empirical likelihood confidence interval for β_0 with level b can be obtained as

$$I_b = \{\beta : l_n(\beta) \leq \chi^2_{1,b}\} \quad \text{and} \quad \tilde{I}_b = \{\beta : \tilde{l}_n(\beta) \leq \chi^2_{1,b}\}$$

for models (2) and (4), respectively, where $\chi^2_{1,b}$ denotes the b-th quantile of a chi-square distribution with one degree of freedom. Therefore, the implementation for constructing the confidence interval is straightforward without estimating any additional quantities. Indeed, the function *"emplik"* in the R package (see Zhou (2012)) can be employed to compute $l_n(\beta)$ and $\tilde{l}_n(\beta)$ easily as we do in the simulation study below.

3 A Monte Carlo Simulation Study

In this section, we investigate the finite sample behavior of the proposed empirical likelihood methods for testing $H_0 : \beta = 0$ against $H_a : \beta \neq 0$. We compare our new methods with the bootstrap method and the Bonferroni Q-test proposed in Campbell and Yogo (2006) in terms of both size and power.

First, we calculate the rejection region based on the least squares estimator $\hat{\beta}_{LSE}$ by using bootstrap method to obtain critical points. More specifically, we first estimate $\alpha, \beta, \theta, \phi, b_j's$ in (4) by least squares estimators, which results in an estimator for (U_t, V_t), say $(\hat{U}_t, \hat{V}_t)$. Next, we draw 1000 random samples with size $n-1$ from $(\hat{U}_t, \hat{V}_t)$, say $(U_t^{*(j)}, V_t^{*(j)})$ for $t = 1, \cdots, n-1$ and $j = 1, \cdots, 1000$. Using model (4) with estimated $\alpha, \beta, \theta, \phi, b_j's$, we obtain the bootstrap samples $\{(X_t^{*(j)}, Y_t^{*(j)})\}_{t=1}^{n-1}$. For each j, we use the bootstrap sample $X_1^{*(j)}, \cdots, X_{n-2}^{*(j)}, Y_2^{*(j)}, \cdots, Y_{n-1}^{*(j)}$ to estimate β by the least squares approach again. Therefore, the rejection region can be obtained based on these 1000 bootstrapped least squares estimators for β. Note that such a bootstrap method is theoretically inconsistent when the sequence $\{X_t\}$ is either nearly integrated or has an infinite variance.

Next, we implement the Bonferroni Q-test given in Campbell and Yogo (2006). Note that the theoretical derivation of the tests in Campbell and Yogo (2006) assumes that α, θ and the covariance of (U_t, V_t) are known and ϕ is near one although the implementation of the Bonferroni Q-test given in Campbell and Yogo (2005) has no such requirements.

Theoretically, one may suspect that the Bonferroni Q-test is inconsistent when α and θ are replaced by their corresponding estimators and ϕ is not close to one. In order to validate this conjecture, we compute the Bonferroni Q-test by using both the true values and the estimated values of α and θ. Since the implementation of the Bonferroni Q-test requires to search several tables in Campbell and Yogo (2005), which depend on both the DF-GLS statistic and δ being the correlation coefficient between U_t and V_t, and are only designed for constructing a 90% two-sided confidence interval or 95% one-sided confidence interval, we fix $\delta = -0.75$ in the model setup. That is, we consider model (4) with $U_t \sim N(0,1)$, $\epsilon_t \sim t(\nu)$, $\delta = -0.75$, $V_t = \delta U_t + \frac{\sqrt{1-\delta^2}}{\sqrt{\nu/(\nu-2)}}\epsilon_t$ if $\nu > 2$ and $V_t = \delta U_t + \epsilon_t$ if $\nu \leq 2$, where $U_1, \cdots, U_n$ and $\epsilon_1, \cdots, \epsilon_n$ are two independent random samples. We also choose $\alpha = 0$, $\beta = a/\sqrt{n}$, $\theta = 0$, $\phi = 0.9,\ 0.99,\ 1$, $p = 1$, $b_1 = 0,\ -0.5$, $\nu = 4,\ 1.5,\ 0.5$ and repeat $10,000$ times with sizes $n = 100$ and 300 from the above setting. Hence, results for $a = 0$ correspond to the size.

We also calculate the empirical likelihood functions in both (3) and (5) by using the R package *"emplik"* in Zhou (2012), i.e., we consider both known and unknown α. In Tables 1 and 2, we report the sizes for these tests. From these two tables, we observe that the proposed empirical likelihood methods have a size close to the nominal level 0.1 whenever the sequence $\{X_t\}$ is stationary or near-integrated or has an infinite variance. The normal approximation method via the bootstrap method only works for the case of $(\phi, \nu) = (0.9, 4)$, i.e., it fails when the sequence $\{X_t\}$ is either nearly integrated or has an infinite variance. This is not surprising because this empirical evidence is in line with the theory provided by Datta (1996) and Hall and Jing (1998). Furthermore, it is interesting to see that the Bonferroni Q-test seems only working for the case of $(\phi, \nu, b_1) = (1, 4, 0)$ with known α and θ. Therefore it remains cautious to employ the Bonferroni Q-test in Campbell and Yogo (2006) due to the complicated implementation and lack of theoretical justification.

In Tables 3 and 4, we report the powers for these tests. We choose $a = -0.3,\ -0.1,$

-0.002 for $\nu = 4$, 1.5, 0.5, respectively. From these two tables, we observe that the proposed empirical likelihood method with known α is much more powerful than the one with unknown α especially for the case of $\nu = 4$. When the normal approximation method produces a consistent size (i.e., the case of $(\phi, \nu) = (0.9, 4)$, it is more powerful than the proposed empirical likelihood methods in both (3) and (5). When the Bonferroni Q-test with known α and θ has a consistent size (i.e., the case of $(\phi, \nu, b_1) = (1, 4, 0)$), it is more powerful than the proposed empirical likelihood method in (5), but less powerful than the empirical likelihood method in (3).

It is easy to verify that Theorems 1 and 2 still hold when $Z_t(\beta)$ in (3) and $\tilde{Z}_t(\beta)$ in (5) are replaced by $Z_t(\beta) = (Y_t - \beta X_{t-1})X_{t-1}/w(X_{t-1})$ and $\tilde{Z}_t(\beta) = (\tilde{Y}_t - \beta \tilde{X}_{t-1})\tilde{X}_{t-1}/w(\tilde{X}_{t-1})$, respectively for some weight function $w(t)$ satisfying that $w(t)/t$ converges to a positive constant as $t \to \infty$. A theoretical optimal weight function will be chosen to minimize the coverage probability error. Without doubt, it is impossible to obtain such an optimal one. Here we consider the class $w(t) = (1 + |t|^h)^{1/h}$ for some $h > 0$. Under the same setup as above, we compute the size and power for the proposed empirical likelihood methods for $h = 1$, 2, 4. From Table 5, we observe that the methods are not quite sensitive to the choice of h especially when X_t has an infinite variance.

To summarize the simulation results, we find the reliable evidence that the proposed empirical likelihood method in (5) can deliver an accurate size and a non-trivial power regardless of the predicting variable being stationary or near-integrated, or having an infinite variance.

Table 1: Empirical sizes are reported for testing $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ with level 10% for the proposed empirical likelihood test in (3) with known α (EL1), the proposed empirical likelihood test in (5) with unknown α (EL2), the normal approximation based on bootstrap method (NA), the Bonferroni Q-test in Campbell and Yogo (2006) with known α and θ (BQ1), and the Bonferroni Q-test with unknown α and θ (BQ2). Sample size $n = 100$.

(a, ϕ, ν, b_1)	EL1	EL2	NA	BQ1	BQ2
$(0, 0.9, 4, 0)$	0.1019	0.1084	0.0971	0.0460	0.0297
$(0, 0.99, 4, 0)$	0.1022	0.0976	0.0682	0.0955	0.0341
$(0, 1, 4, 0)$	0.1048	0.1155	0.0666	0.1085	0.0363
$(0, 0.9, 1.5, 0)$	0.1005	0.1039	0.0458	0.0613	0.0418
$(0, 0.99, 1.5, 0)$	0.1036	0.1053	0.0358	0.1576	0.0445
$(0, 1, 1.5, 0)$	0.1038	0.1077	0.0417	0.1881	0.0449
$(0, 0.9, 0.5, 0)$	0.1007	0.1028	0.0125	0.0864	0.0634
$(0, 0.99, 0.5, 0)$	0.1025	0.1035	0.0168	0.2124	0.0600
$(0, 1, 0.5, 0)$	0.1035	0.1031	0.0283	0.2406	0.0630
$(0, 0.9, 4, -0.5)$	0.1052	0.1070	0.1088	0.0688	0.0156
$(0, 0.99, 4, -0.5)$	0.1042	0.0963	0.0817	0.2326	0.0137
$(0, 1, 4, -0.5)$	0.1055	0.1172	0.0750	0.2505	0.0126
$(0, 0.9, 1.5, -0.5)$	0.1008	0.1000	0.0500	0.0593	0.0354
$(0, 0.99, 1.5, -0.5)$	0.1036	0.1051	0.0377	0.1809	0.0427
$(0, 1, 1.5, -0.5)$	0.1053	0.1065	0.0414	0.2078	0.0440
$(0, 0.9, 0.5, -0.5)$	0.0978	0.1009	0.0130	0.0902	0.0639
$(0, 0.99, 0.5, -0.5)$	0.1041	0.1030	0.0159	0.2200	0.0568
$(0, 1, 0.5, -0.5)$	0.1018	0.1023	0.0248	0.2532	0.0700

Table 2: Empirical sizes are reported for testing $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ with level 10% for the proposed empirical likelihood test in (3) with known α (EL1), the proposed empirical likelihood test in (5) with unknown α (EL2), the normal approximation based on bootstrap method (NA), the Bonferroni Q-test in Campbell and Yogo (2006) with known α and θ (BQ1), and the Bonferroni Q-test with unknown α and θ (BQ2). Sample size $n = 300$.

(a, ϕ, ν, b_1)	EL1	EL2	NA	BQ1	BQ2
$(0, 0.9, 4, 0)$	0.1036	0.1048	0.1061	0.0251	0.0213
$(0, 0.99, 4, 0)$	0.1035	0.0869	0.0752	0.0636	0.0335
$(0, 1, 4, 0)$	0.1063	0.1051	0.0627	0.0936	0.0311
$(0, 0.9, 1.5, 0)$	0.1058	0.1087	0.0580	0.0424	0.0383
$(0, 0.99, 1.5, 0)$	0.1004	0.1055	0.0362	0.1105	0.0438
$(0, 1, 1.5, 0)$	0.0980	0.1072	0.0391	0.1914	0.0424
$(0, 0.9, 0.5, 0)$	0.1005	0.1031	0.0081	0.0617	0.0551
$(0, 0.99, 0.5, 0)$	0.0966	0.1012	0.0079	0.1454	0.0565
$(0, 1, 0.5, 0)$	0.0970	0.0989	0.0194	0.2334	0.0552
$(0, 0.9, 4, -0.5)$	0.1052	0.1070	0.1088	0.0688	0.0156
$(0, 0.99, 4, -0.5)$	0.1043	0.0885	0.0790	0.2358	0.0193
$(0, 1, 4, -0.5)$	0.1084	0.1071	0.0656	0.3239	0.0208
$(0, 0.9, 1.5, -0.5)$	0.1032	0.1076	0.0573	0.0443	0.0374
$(0, 0.99, 1.5, -0.5)$	0.1010	0.1034	0.0392	0.1288	0.0442
$(0, 1, 1.5, -0.5)$	0.0975	0.1049	0.0378	0.2104	0.0430
$(0, 0.9, 0.5, -0.5)$	0.1044	0.1031	0.0082	0.0610	0.0534
$(0, 0.99, 0.5, -0.5)$	0.0958	0.1020	0.0082	0.1454	0.0566
$(0, 1, 0.5, -0.5)$	0.0963	0.0994	0.0192	0.2363	0.0558

Table 3: Empirical powers are reported for testing $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ with level 10% for the proposed empirical likelihood test in (3) with known α (EL1), the proposed empirical likelihood test in (5) with unknown α (EL2), the normal approximation based on bootstrap method (NA), the Bonferroni Q-test in Campbell and Yogo (2006) with known α and θ (BQ1), and the Bonferroni Q-test with unknown α and θ (BQ2). Sample size $n = 100$.

(a, ϕ, ν, b_1)	EL1	EL2	NA	BQ1	BQ2
$(-0.3, 0.9, 4, 0)$	0.1831	0.1582	0.2097	0.0556	0.0428
$(-0.3, 0.99, 4, 0)$	0.3872	0.1578	0.2167	0.1750	0.0908
$(-0.3, 1, 4, 0)$	0.4613	0.1721	0.2312	0.2566	0.1412
$(-0.1, 0.9, 1.5, 0)$	0.2366	0.1957	0.1529	0.1957	0.1670
$(-0.1, 0.99, 1.5, 0)$	0.5168	0.3175	0.2770	0.4347	0.2865
$(-0.1, 1, 1.5, 0)$	0.5991	0.3478	0.3495	0.5159	0.3448
$(-0.002, 0.9, 0.5, 0)$	0.6002	0.5814	0.4149	0.6869	0.6673
$(-0.002, 0.99, 0.5, 0)$	0.7925	0.7176	0.6168	0.8113	0.7461
$(-0.002, 1, 0.5, 0)$	0.8215	0.7235	0.6740	0.8370	0.7679
$(-0.3, 0.9, 4, -0.5)$	0.1495	0.1348	0.1694	0.1348	0.0498
$(-0.3, 0.99, 4, -0.5)$	0.2834	0.1170	0.1512	0.3804	0.0569
$(-0.3, 1, 4, -0.5)$	0.3518	0.1214	0.1541	0.4551	0.0769
$(-0.1, 0.9, 1.5, -0.5)$	0.1798	0.1589	0.1027	0.1679	0.1281
$(-0.1, 0.99, 1.5, -0.5)$	0.3907	0.2330	0.1707	0.3990	0.1986
$(-0.1, 1, 1.5, -0.5)$	0.4683	0.2538	0.2180	0.4733	0.2366
$(-0.002, 0.9, 0.5, -0.5)$	0.5337	0.5149	0.3513	0.6321	0.6108
$(-0.002, 0.99, 0.5, -0.5)$	0.7350	0.6536	0.5297	0.7627	0.6801
$(-0.002, 1, 0.5, -0.5)$	0.7648	0.6634	0.5903	0.7921	0.7072

Table 4: Empirical powers are reported for testing $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ with level 10% for the proposed empirical likelihood test in (3) with known α (EL1), the proposed empirical likelihood test in (5) with unknown α (EL2), the normal approximation based on bootstrap method (NA), the Bonferroni Q-test in Campbell and Yogo (2006) with known α and θ (BQ1), and the Bonferroni Q-test with unknown α and θ (BQ2). Sample size $n = 300$.

(a, ϕ, ν, b_1)	EL1	EL2	NA	BQ1	BQ2
$(-0.3, 0.9, 4, 0)$	0.1787	0.1461	0.2182	0.0587	0.0541
$(-0.3, 0.99, 4, 0)$	0.4674	0.2742	0.3933	0.2788	0.1953
$(-0.3, 1, 4, 0)$	0.6547	0.3258	0.4457	0.5156	0.3933
$(-0.1, 0.9, 1.5, 0)$	0.2318	0.1955	0.2367	0.2457	0.2337
$(-0.1, 0.99, 1.5, 0)$	0.6937	0.5122	0.5717	0.6361	0.5411
$(-0.1, 1, 1.5, 0)$	0.8484	0.6328	0.7032	0.7972	0.6867
$(-0.002, 0.9, 0.5, 0)$	0.8285	0.8153	0.8367	0.9240	0.9204
$(-0.002, 0.99, 0.5, 0)$	0.9730	0.9568	0.9495	0.9794	0.9717
$(-0.002, 1, 0.5, 0)$	0.9870	0.9698	0.9738	0.9898	0.9829
$(-0.3, 0.9, 4, -0.5)$	0.1437	0.1262	0.1728	0.1651	0.1026
$(-0.3, 0.99, 4, -0.5)$	0.3314	0.1895	0.2636	0.4296	0.1054
$(-0.3, 1, 4, -0.5)$	0.5156	0.1992	0.2779	0.6398	0.2266
$(-0.1, 0.9, 1.5, -0.5)$	0.1695	0.1555	0.1535	0.2030	0.1836
$(-0.1, 0.99, 1.5, -0.5)$	0.5276	0.3659	0.3593	0.5354	0.3705
$(-0.1, 1, 1.5, -0.5)$	0.7258	0.4767	0.5041	0.7153	0.5116
$(-0.002, 0.9, 0.5, -0.5)$	0.7595	0.7492	0.7562	0.8866	0.8831
$(-0.002, 0.99, 0.5, -0.5)$	0.9510	0.9278	0.9080	0.9611	0.9488
$(-0.002, 1, 0.5, -0.5)$	0.9764	0.9472	0.9487	0.9783	0.9649

Table 5: Empirical sizes and powers are reported for testing $H_0 : \beta = 0$ against $H_a : \beta \neq 0$ with level 10% for the proposed empirical likelihood test in (3) with known α (EL1) and the proposed empirical likelihood test in (5) with unknown α (EL2), where the general weight function $w(t) = (1 + |t|^h)^{1/h}$ is employed. Sample size $n = 300$.

(a, ϕ, ν, b_1)	EL1 $h = 1$	EL1 $h = 2$	EL1 $h = 4$	EL2 $h = 1$	EL2 $h = 2$	EL2 $h = 4$
$(0, 0.9, 4, 0)$	0.0970	0.0963	0.0956	0.1005	0.1006	0.0989
$(0, 0.99, 4, 0)$	0.1025	0.1008	0.1006	0.0868	0.0853	0.0833
$(0, 1, 4, 0)$	0.1097	0.1067	0.1053	0.1099	0.1043	0.1022
$(0, 0.9, 1.5, 0)$	0.0951	0.0968	0.0967	0.0964	0.0969	0.0960
$(0, 0.99, 1.5, 0)$	0.1001	0.0973	0.0964	0.0938	0.0925	0.0928
$(0, 1, 1.5, 0)$	0.0995	0.0993	0.0994	0.0983	0.00991	0.0979
$(0, 0.9, 0.5, 0)$	0.1008	0.1004	0.0999	0.0978	0.0982	0.0977
$(0, 0.99, 0.5, 0)$	0.0991	0.0980	0.0983	0.1033	0.1033	0.1032
$(0, 1, 0.5, 0)$	0.0969	0.0963	0.0968	0.1033	0.1032	0.1032
$(-0.3, 0.9, 4, 0)$	0.1687	0.1675	0.1657	0.1435	0.1436	0.1443
$(-0.3, 0.99, 4, 0)$	0.4747	0.4658	0.4617	0.2669	0.2664	0.2658
$(-0.3, 1, 4, 0)$	0.6717	0.6669	0.6623	0.3187	0.3173	0.3163
$(-0.1, 0.9, 1.5, 0)$	0.2505	0.2343	0.2290	0.1994	0.1922	0.1915
$(-0.1, 0.99, 1.5, 0)$	0.7035	0.6879	0.6850	0.5203	0.5124	0.5112
$(-0.1, 1, 1.5, 0)$	0.8475	0.8401	0.8378	0.6384	0.6324	0.6308
$(-0.002, 0.9, 0.5, 0)$	0.8341	0.8306	0.8304	0.8088	0.8079	0.8072
$(-0.002, 0.99, 0.5, 0)$	0.9728	0.9723	0.9722	0.9547	0.9545	0.9545
$(-0.002, 1, 0.5, 0)$	0.9885	0.9888	0.9888	0.9683	0.9682	0.9682

4 Predictability of monthly CRSP value-weighted index

An often asked question in financial econometrics is whether asset returns can be predicted by some macroeconomic data such as the dividend-price ratio and the earnings-price ratio as well as other state variables like interest rates. In this section, we apply the empirical likelihood method in (5) to re-visit the data set analyzed by Campbell and Yogo (2006). More specifically, the predictable variable Y_t is the monthly CRSP value-weighted index data (1926:12–2002:12) from the Center for Research in Security Prices, and the predicting variable X_t is either the log dividend-price ratio (ldp) or the log earnings-price ratio (lep). The dividend-price ratio is computed as dividends over the past year divided by the current price, and the earnings-price ratio is computed as a moving average of earnings over the past ten years divided by the current price. There are 913 observations in total. The detailed description of this data set can be found in the paper by Campbell and Yogo (2006). Similar to Campbell and Yogo (2006), we consider three time periods as 1926:12–2002:12, 1926:12–1994:12, and 1952:12–2002:12. The main purpose of re-visiting this particular data set is to argue that the proposed methodology in this paper can provide more accurate statistical inference than that in Campbell and Yogo (2006).

Based on the above data set and model (4) with $p = 0$, Campbell and Yogo (2006) calculated the Bonferroni Q-test for $\tilde{\beta} = \beta\sigma_V/\sigma_U$ rather than β by simply scaling the test by $\hat{\sigma}_V/\hat{\sigma}_U$, where $\sigma_V, \hat{\sigma}_V$ and $\sigma_U, \hat{\sigma}_U$ denote the standard deviation and estimated standard deviation of V_t and U_t in (1), respectively. Hence, the results in Table 5 of Campbell and Yogo (2006) ignored the effect of the plug-in estimators $\hat{\sigma}_U$ and $\hat{\sigma}_V$. It is natural to conjecture that such an effect should result in wider intervals for β than those reported in Table 5 of Campbell and Yogo (2006). Moreover, due to the complicated implementation and too simplified

theoretical derivations in Cambell and Yogo (2006), one may question the reliability of the empirical findings in Campbell and Yogo (2006). Here, we employ the proposed empirical likelihood method in (5) to compute intervals for β rather than $\tilde{\beta}$. Since the new method works for all cases with sound theory and is easy to implement, we believe that the analysis under the new method is more robust and reliable.

Table 6 reports confidence intervals with levels 0.90 in the fifth column and 0.95 in the last column for the monthly CRSP value-weighted index with periods 1926–2002, 1926–1994, and 1952–2002 as in Table 5 of Campbell and Yogo (2006). It is not surprising to observe from Table 6 that the new intervals are indeed wider than those reported in Table 5 of Campbell and Yogo (2006) because as demonstrated earlier, the Bonferroni Q-test gives a tighter conference interval. Similar to Campbell and Yogo (2006), the null hypothesis of no predictability ($H_0 : \beta = 0$) is not rejected by the new method for the log dividend-price ratio for all three time periods and for the log earnings-price ratio in the sub-sample 1952–2002. Also, the null hypothesis of no predictability is rejected by the new method for the log earnings-price ratio for the full sample 1926–2002 at both levels 90% and 95%. However, interestingly, the null hypothesis of no predictability is not rejected by the proposed new method for the log earnings-price ratio in the sub-sample 1926–1994, while it is rejected by Campbell and Yogo (2006). Indeed, our finding for this sub-sample is similar to the conclusion in Cai and Wang (2013) for the period 1930:12 – 1990:12. That is, the asset return is not predictable in the sub-sample through the early of 1990s. The source of this difference between our finding and the result in Campbell and Yogo (2006) can be explained by the following arguments. For this sub-sample, the confidence interval for ϕ (see Table 4 in Campbell and Yogo (2006)) is $[0.970, 0.997]$ and it does not cover $\phi = 1$ so that X_t might be stationary and is a less persistent series. As indicated earlier, the Bonferroni Q-test may not perform well when X_t is stationary or nearly integrated.

Table 6: Confidence intervals for the monthly CRSP value-weighted index are computed for the proposed empirical likelihood method.

CRSP Series	Variable	$\hat{\beta}_{LSE}$	$\hat{\sigma}_V/\hat{\sigma}_U$	$I_{0.9}$	$I_{0.95}$
1926–2002	d-p	0.0083	1.0367	[-0.0042, 0.0231]	[-0.0068, 0.0259]
1926–2002	e-p	0.0129	1.0428	[0.0034, 0.0317]	[0.0008, 0.0346]
1926–1994	d-p	0.0123	1.0342	[-0.0134, 0.0297]	[-0.0175, 0.0342]
1926–1994	e-p	0.0211	1.0373	[-0.0059, 0.0401]	[-0.0102, 0.0449]
1952–2002	d-p	0.0116	1.0324	[-0.0105, 0.0181]	[-0.0133, 0.0208]
1952–2002	e-p	0.0088	1.0117	[-0.0134, 0.0118]	[-0.0159, 0.0142]

5 Conclusion

Researchers have constantly asked whether stock returns can be predicted by macroeconomic data. On the other hand, macroeconomic data may exhibit non-stationarity and heavy tails. Therefore, it is important to have a unified method to test predictability in regressions without distinguishing whether the predicting variable is stationary or non-stationary or has an infinite variance.

In this paper, we study a predictive regression model which has an ability to include the regressors to be stationary or non-stationary (integrated/nearly integrated) process and/or has an infinite variance and allows the so-called two innovations to be correlated. We propose novel empirical likelihood methods based on some weighted score equation to construct a confidence for the coefficient or to test the predictability. We show that Wilks theorem holds for the proposed empirical likelihood methods regardless of the predicting variable being stationary, or non-stationary or having an infinite variance. The proposed new methods are easy to implement without any ad hoc method such as bootstrap method for obtaining critical values. Therefore, the proposed new methods provide more robust findings than other existing methods in the literature of predictive regressions and have wide applications

in financial econometrics.

6 Appendix: Proofs

We only prove Theorem 2 since the proof of Theorem 1 is easier.

Proof of Theorem 2. Put $\tilde{V}_j = V_j - V_{j+m}$ and let $\mathcal{F}_t$ denote the σ-field generated by $\{(\tilde{U}_s, \tilde{V}_s) : 1 \le s \le t\} \cup \{V_s : s \le 0\}$. Write $B(L) = \prod_{j=1}^{p}(1 - \tilde{b}_j L)$. Then we have $B^{-1}(L) = \prod_{j=1}^{p}(1 - \tilde{b}_j L)^{-1} = \sum_{k=0}^{\infty} a_k L^k$ and

$$e_t = \sum_{k=0}^{\infty} a_k V_{t-k} = \sum_{k=0}^{t-1} a_k V_{t-k} + \sum_{k=t}^{\infty} a_k V_{t-k}.$$

Note that

$$|a_k| \le k^p (\max_{1\le i\le p} |\tilde{b}_i|)^k \quad \text{and} \quad \max_{1\le i\le p} |\tilde{b}_i| < 1. \tag{6}$$

Put $e_{t,1} = \sum_{k=0}^{t-1} a_k \tilde{V}_{t-k} + \sum_{k=t}^{\infty} a_k V_{t-k} - \sum_{k=t+m}^{\infty} a_k V_{t+m-k}$ and $e_{t,2} = -\sum_{k=t}^{t+m-1} a_k V_{t+m-k}$. Then we have

$$e_t - e_{t+m} = e_{t,1} + e_{t,2} \quad \text{for} \quad t = 1, \cdots, m. \tag{7}$$

Write

$$X_t = \frac{1-\phi^t}{1-\phi}\theta + \sum_{j=1}^{t} \phi^{t-j} e_j + \phi^t X_0 \tag{8}$$

and

$$\begin{aligned} X_{t+m} &= \tfrac{1-\phi^{t+m}}{1-\phi}\theta + \textstyle\sum_{j=1}^{t+m} \phi^{t+m-j} e_j + \phi^{t+m} X_0 \\ &= \tfrac{1-\phi^{t}}{1-\phi}\theta + \textstyle\sum_{j=1}^{t} \phi^{t-j} e_{j+m} + \phi^{t} X_0 + \tfrac{\phi^t - \phi^{t+m}}{1-\phi}\theta + \textstyle\sum_{j=1}^{m} \phi^{t+m-j} e_j + (\phi^{t+m} - \phi^t) X_0. \end{aligned} \tag{9}$$

Put $W_{t,1} = \sum_{j=1}^{t} \phi^{t-j} e_{j,1}$ and

$$W_{t,2} = \sum_{j=1}^{t} \phi^{t-j} e_{j,2} - \frac{\phi^t - \phi^{t+m}}{1-\phi}\theta - \sum_{j=1}^{m} \phi^{t+m-j} e_j - (\phi^{t+m} - \phi^t) X_0.$$

Then, it follows from (7)–(9) that

$$\tilde{X}_t = W_{t,1} + W_{t,2} \quad \text{for} \quad t = 1, \cdots, m. \tag{10}$$

When $|\phi| < 1$, it follows from (6) that as $n \to \infty$

$$\begin{aligned}
&\frac{1}{m}\sum_{t=1}^{m}\frac{W_{t-1,1}^2}{1+W_{t-1,1}^2}\\
=&\frac{1}{m}\sum_{t=1}^{m}\left(\sum_{j=1}^{t}\phi^{t-j}e_{j,1}\right)^2 \Big/ \left\{1+\left(\sum_{j=1}^{t}\phi^{t-j}e_{j,1}\right)^2\right\}\\
=&\lim_{t\to\infty} E\frac{(\sum_{j=1}^{t}\phi^{t-j}e_{j,1})^2}{1+(\sum_{j=1}^{t}\phi^{t-j}e_{j,1})^2}+o_p(1)\\
:=&\sigma_0^2+o_p(1).
\end{aligned} \tag{11}$$

When $\phi = 1 - \gamma_\phi/n$ for some constant $\gamma_\phi \geq 0$, we have

$$W_{t,1} \xrightarrow{p} \infty, \quad W_{t,1} = O_p(t^{1/\alpha^*}) \quad \text{and} \quad \frac{|W_{t,1}|}{t^{1/\alpha^*-\delta_0}} \xrightarrow{p} \infty \tag{12}$$

for any $\delta_0 > 0$ as $t \to \infty$ by using (6) and the fact that the distribution of V_t lies in the domain of attraction of a stable law with index α^*. Hence

$$\frac{W_{t-1,1}^2}{1+W_{t-1,1}^2} \xrightarrow{p} 1 \quad \text{as} \quad t \to \infty,$$

i.e.,

$$\frac{1}{m}\sum_{t=1}^{m}\frac{W_{t-1,1}^2}{1+W_{t-1,1}^2} \xrightarrow{p} 1 \quad \text{as} \quad n \to \infty. \tag{13}$$

By (11) and (13), we have as $n \to \infty$

$$\begin{aligned}
&\frac{1}{m}\sum_{t=1}^{m}E\left(\frac{\tilde{U}_t^2 W_{t-1,1}^2}{1+W_{t-1,1}^2}\middle|\mathcal{F}_{t-1}\right) = 2EU_1^2\frac{1}{m}\sum_{t=1}^{m}\frac{W_{t-1,1}^2}{1+W_{t-1,1}^2}\\
\xrightarrow{p}&\begin{cases} 2EU_1^2\sigma_0^2, & \text{if} \quad |\phi|<1,\\ 2EU_1^2, & \text{if} \quad \phi = 1-\gamma_\phi/n.\end{cases}
\end{aligned}$$

Similarly, for any $c > 0$,

$$\begin{aligned}
&\frac{1}{m}\sum_{t=1}^{m} E\left(\frac{\tilde{U}_t^2 W_{t-1,1}^2}{1+W_{t-1,1}^2} I\left(\frac{\tilde{U}_t^2 W_{t-1,1}^2}{1+W_{t-1,1}^2} > c^2 m\right)\middle|\mathcal{F}_{t-1}\right)\\
\leq&\frac{1}{(c\sqrt{m})^q}\frac{1}{m}\sum_{t=1}^{m} E\left(\left|\frac{\tilde{U}_t W_{t-1,1}}{\sqrt{1+W_{t-1,1}^2}}\right|^{2+q}\middle|\mathcal{F}_{t-1}\right)\\
=&\frac{E|\tilde{U}_1|^{2+q}}{(c\sqrt{m})^q}\frac{1}{m}\sum_{t=1}^{m}\frac{|W_{t-1,1}|^{2+q}}{(1+W_{t-1,1}^2)^{(2+q)/2}}\\
\xrightarrow{d}&0 \quad \text{as} \quad n\to\infty.
\end{aligned}$$

By Corollary 3.1 of Hall and Heyde (1980), we have as $n \to \infty$

$$\frac{1}{\sqrt{m}}\sum_{t=1}^{m}\frac{\tilde{U}_t W_{t-1,1}}{\sqrt{1+W_{t-1,1}^2}} \xrightarrow{d} \begin{cases} N\left(0, 2E(U_1^2)\sigma_0^2\right), & \text{if} \quad |\phi|<1,\\ N(0, 2EU_1^2), & \text{if} \quad \phi = 1-\gamma_\phi/n. \end{cases} \tag{14}$$

Using (6) and the fact that the distribution of V_t lies in the domain of attraction of a stable law with index α^*, it is easy to check that

$$|\phi^{-t}W_{t,2}| = \begin{cases} O_p(1), & \text{if} \quad |\phi|<1,\\ O_p(m^{1/\alpha^*}), & \text{if} \quad \phi = 1-\gamma_\phi/n \end{cases} \tag{15}$$

and

$$\frac{|W_{t,2}|}{t^{1/\alpha^*-\delta_0}} \xrightarrow{p} \infty \tag{16}$$

for any $\delta_0 > 0$ as $t \leq m$ goes to infinity. Hence, by (12), (15) and (16), we have

$$\begin{aligned}
&\frac{1}{\sqrt{m}}\sum_{t=1}^{m}\left(\frac{\tilde{U}_t\tilde{X}_{t-1}}{\sqrt{1+\tilde{X}_{t-1}^2}} - \frac{\tilde{U}_t W_{t-1,1}}{\sqrt{1+W_{t-1,1}^2}}\right)\\
=&-\frac{1}{\sqrt{m}}\sum_{t=1}^{m}\tilde{U}_t\frac{1}{\{1+(a_{t-1}\tilde{X}_{t-1}+(1-a_{t-1})W_{t-1,1})^2\}^{3/2}}W_{t-1,2}\\
=&-\frac{1}{\sqrt{m}}\sum_{t=1}^{m}\tilde{U}_t\frac{1}{\{1+(W_{t-1,1}-a_{t-1}W_{t-1,2})^2\}^{3/2}}W_{t-1,2}\\
=&\begin{cases} O_p(\frac{1}{\sqrt{m}}\sum_{t=1}^{m}\phi^t|\tilde{U}_t|), & \text{if} \quad |\phi|<1,\\ O_p(\frac{1}{\sqrt{m}}\sum_{t=1}^{m}|\tilde{U}_t|\frac{m^{1/\alpha^*+\delta_0}}{t^{3(1/\alpha^*-\delta_0)}}), & \text{if} \quad \phi = 1-\gamma_\phi/n \end{cases}\\
=&o_p(1) \quad \text{as} \quad n\to\infty,
\end{aligned} \tag{17}$$

where $a_{t-1} \in [0,1]$ may depend on $\tilde{X}_{t-1}$ and $W_{t-1,1}$, and $\delta_0 > 0$ is small enough. It follows from (14) and (17) that

$$\frac{1}{\sqrt{m}} \sum_{t=1}^{m} \tilde{Z}_t(\beta_0) \xrightarrow{d} \begin{cases} N\left(0, 2EU_1^2\sigma_0^2\right), & \text{if} \quad |\phi| < 1, \\ N(0, 2EU_1^2), & \text{if} \quad \phi = 1 - \gamma_\phi/n \end{cases}$$

as $n \to \infty$. Similarly we can show that

$$\frac{1}{m} \sum_{t=1}^{m} \tilde{Z}_t^2(\beta_0) \xrightarrow{p} \begin{cases} 2EU_1^2\sigma_0^2, & \text{if} \quad |\phi| < 1, \\ 2EU_1^2, & \text{if} \quad \phi = 1 - \gamma_\phi/n \end{cases}$$

as $n \to \infty$. The rest follows from the standard arguments in the proof of the empirical likelihood method (see Chapter 11 of Owen (2001)).

□

Acknowledgments

We thank the editor, Professor Susan Paddock, an associate editor and two reviewers for helpful comments. Zhu's research was supported by National Natural Science Foundation of China grants #11371168, #11001105, #11271155, Specialized Research Fund for the Doctoral Program of Higher Education #20110061110003, Science and Technology Developing Plan of Jilin Province #20130522102JH and 985 Project of Jilin University. Cai's research was supported, in part, by the National Natural Science Foundation of China grants #71131008 (Key Project) and #70971113. Peng's research was supported by NSF grant DMS 1005336.

References

Amihud, Y. and Hurvich, C.M. (2004). Predictive regressions: a reduced-bias estimation method. *Journal of Financial and Quantitative Analysis* **39**, 813–841.

Amihud, Y., Hurvich, C. and Wang, Y. (2009). Multiple-predictor regressions: Hypothesis testing. *Review of Financial Studies* **22**, 413-434.

Cai, Z. and Wang, Y. (2013). Testing predictive regression models with nonstationary regressors. *Journal of Econometrics*, to appear. It is now available at the web site http://www.sciencedirect.com/science/article/pii/S0304407613001528.

Campbell, J.Y. and Yogo, M. (2005). Implementing the econometric methods in "Efficient tests of stock return predictability". Unpublished working paper. University of Pennsylvania.

Campbell, J.Y. and Yogo, M. (2006). Efficient tests of stock return predictability. *Journal of Financial Economics* **81**, 27–60.

Cavanagh, C.L., Elliott, G. and Stock, J.H. (1995). Inference in models with nearly integrated regressors. *Econometric Theory* **11**, 1131-1147.

Chan, N.H., Li, D. and Peng, L. (2012). Toward a unified interval estimation of autoregressions. *Econometric Theory* **28**, 705–717.

Chuang, C. and Chan, N.H. (2002). Empirical likelihood for autoregressive models, with applications to unstable time series. *Statistica Sinica* **12**, 387–407.

Chen, W.W. and Deo, R.S. (2009). Bias reduction and likelihood-based almost exactly sized hypothesis testing in predictive regressions using the restricted likelihood. *Econometric Theory* **25**, 1143–1179.

Datta, S. (1996). On asymptotic properties of bootstrap for AR(1) processes. *Journal of Statistical Planning and Inference* **53**, 361–374.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications (Vol. 2, 2nd ed.)*. New York: John Wiley.

Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application. Academic Press.*

Hall, P. and Jing, B. (1998). Comparison of bootstrap and asymptotic approximations to the distribution of a heavy-tailed mean. *Statistica Sinica* **8**, 887–906.

Jansson, M. and Moreira, M.J. (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica* **74**, 681–714.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics* **74**, 209–235.

Ling, S. (2005). Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *Journal of the Royal Statistical Society Series B* **67**, 381–393.

Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237–249.

Owen, A. (1990). Empirical likelihood confidence regions. *Annals of Statistics* **18**, 90–120.

Owen, A. (2001). *Empirical Likelihood*. Chapman and Hall, New York.

Samorodnitsky, G., Rachev, S.T., Kurz-Kim, J. and Stoyanov, S.V. (2007). Asymptotic distribution of unbiased linear estimators in the presence of heavy-tailed stochastic regressors and residuals. *Probability and Mathematical Statistics* **27**, 275–302.

Stambaugh, R.F. (1999). Predictive regressions. *Journal of Financial Economics* **54**, 375–421.

Zhou, M. (2012). emplik: Empirical likelihood ratio for censored/truncated data. *R package version 0.9-8-2. http://CRAN.R-project.org/package=emplik*