Hitotsubashi University

Global COE Hi-Stat Discussion Paper Series 192

Research Unit for Statistical and Empirical Analysis in Social Sciences (Hi-Stat)

Estimation and Inference in Predictive Regressions

Eiji Kurozumi and Kohei Aono

May 2011

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

Hi-Stat Discussion Paper

Estimation and Inference in Predictive Regressions[1]

Eiji Kurozumi[2]

Department of Economics
Hitotsubashi University

Kohei Aono

College of Business Administration
Ritsumeikan University

May 18, 2011

Abstract

This paper proposes new point estimates for predictive regressions. Our estimates are easily obtained by the least squares and the instrumental variable methods. Our estimates, called the plug-in estimates, have nice asymptotic properties such as median unbiasedness and the approximated normality of the associated t-statistics. In addition, the plug-in estimates are shown to have good finite sample properties via Monte Carlo simulations. Using the new estimates, we investigate U.S. stock returns and find that some variables, which have not been statistically detected as useful predictors in the literature, are able to predict stock returns. Because of their nice properties, our methods complement the existing statistical tests for predictability to investigate the relations between stock returns and economic variables.

JEL classification: C13; C22; C58; G17

Key words: Unit root; near unit root; bias; median unbiased; stock return

[1]Correspondence: Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. E-mail: kurozumi@stat.hit-u.ac.jp

[2]This research was supported by the Global COE program, Research Unit for Statistical and Empirical Analysis in Social Sciences, Hitotsubashi University.

1. Introduction

Predictive regressions have long been studied in the financial and econometric literature. One of the difficulties in predicting stock returns by financial variables is that the ordinary least squares estimate (OLSE) is severely biased, as pointed out by Mankiw and Shapiro (1986) and Elliott and Stock (1994). Because of the upward bias, the usual t test suffers from over-size distortions, and hence, we tend to erroneously find evidence of strong predictability.

The bias of predictive regressions comes from the fact that typical predictive variables are strongly serially correlated and also contemporarily correlated with prediction errors. Because such persistent variables are known to be well characterized by nearly integrated models (Phillips, 1987), we consider the following predictive regression model in this paper:

$$r_t = \mu_r + \beta x_{t-1} + u_t, \tag{1}$$

$$x_t = \mu_x + \rho x_{t-1} + e_t, \tag{2}$$

where $\mu_x = m_x(1-\rho)$ with $m_x = E[x_t]$ and $\rho = 1-\theta/T$ with some fixed value of θ. Because the predictive variable depends on the localizing parameter θ, the OLSE of β also depends on it even asymptotically. As a result, the t-statistic of β depends on the nuisance parameter θ, and thus, we cannot control the size of the test.

The problem in the above model is that we cannot consistently estimate the localizing parameter, and consequently, many efforts have been made to overcome this problem. Elliott and Stock (1994) considered approximating the distribution of test statistics by the Bayesian mixture procedure, while Cavanagh, Elliott and Stock (1995) proposed to construct test statistics that are free from the localizing parameter by using three intervals—sup-bound, Bonferroni, and Scheffe-type. The latter methods have been applied to stock returns by Torous, Valkanov and Yan (2004), who found evidence of predictability only at short horizons. Stambaugh (1999) tackled this problem from the Bayesian point of view and based on his theory, Lewellen (2004) developed the bias adjustment of the OLSE of β. He found stock returns predictability using new tests by combining the t tests based on the OLSE of β and the bias-adjusted estimate. Campbell and Yogo (2006) discussed the optimality of predictive regressions. Theoretically, we can construct conditional optimal tests as developed by Jansson and Moreira (2006) but such tests require advanced computational methods. Instead, Campbell and Yogo (2006) proposed Bonferroni intervals based on conditionally optimal tests and found stock return predictability. On the other hand, Lanne (2002) tested predictability not by estimating (1) but by applying stationarity tests, and found it difficult to predict stock returns.

Most of these existing articles are mainly concerned with controlling the size of tests with strongly serially correlated predictors but do not provide the estimation methods. However, it is often the case that we need the point estimates of β once the evidence of predictability is observed. For example, we need the point estimates when actually forecasting future returns by predictive regressions. Another example is the case where a change in predictability is observed, as is often the case with the U.S. stock returns. In this case, we may want to examine the magnitude of change in predictability, which can be measured by using the point estimates of β in two sub-samples. Note that although the existing testing methods may give the confidence intervals of β, they do not necessarily give the point estimates of β. Hence, we need to develop estimation methods for further investigation of predictive regressions.

In this paper, we propose new estimates for predictive regressions. We show via Monte Carlo simulations that our estimates behave quite well in finite samples. The important finite sample properties of our estimates are (i) the bias is relatively small and the estimates are almost median unbiased, (ii) the empirical size of the one-sided t tests based on the new estimates is close to the nominal one, (iii) the coverage rate of the confidence intervals is close to the theoretical one, and (iv) the power of our t tests is comparable to that of the Bonferroni Q test by Campbell and Yogo (2006) if the contemporaneous correlation between two innovations is not so strong, although the latter test dominates the former in other cases. Considering these favorable finite sample properties, our new estimation methods will complement the existing testing methods to investigate the predictability of stock returns.

In the empirical analysis, we first apply our methods to the U.S. monthly and annual stock returns investigated by Campbell and Yogo (2006). We find the predictability of stock returns using the same predictor variables as observed by Campbell and Yogo (2006). In addition, we also find strong evidence of predictability by the dividend-price ratio, which is not considered as a statistically significant predictor. We then investigate the same stock returns by extending the sample period. Again, we clearly observe the predictability of stock returns by the dividend-price ratio.

The rest of the paper is organized as follows. Section 2 explains our new estimation methods. We investigate the finite sample properties of our new estimates in Section 3. The predictability of the U.S. stock returns is investigated in Section 4, and the concluding remarks are given in Section 5. The technical proofs and derivations of the statistical properties of the new estimates are relegated to the Appendix.

2. Estimation Methods for Predictive Regressions

Let us first consider the following simple predictive regression without a constant:

$$r_t = \beta x_{t-1} + u_t, \tag{3}$$

$$x_t = \rho x_{t-1} + e_t, \tag{4}$$

for $t = 1, \cdots, T$, where $\rho = 1 - \theta/T$, $u_t \sim i.i.d.(0, \sigma_u^2)$, $e_t \sim i.i.d.(0, \sigma_e^2)$, and u_t and e_t are contemporaneously correlated with covariance σ_{ue}. We use this simple model in order to explain our main logic underlying the construction of new point estimates. The extensions to general autoregressive models of order p (AR(p)) with a constant will be discussed later in this section and in the Appendix. Since typical predictors are highly persistent but do not necessarily have a unit root by economic theory, we consider the local-to-unity specification for the AR(1) parameter as given by $\rho = 1 - \theta/T$.

For model (3), it is well known that the OLSE of β,

$$\hat{\beta} = \frac{\sum_{t=1}^{T} x_{t-1} r_t}{\sum_{t=1}^{T} x_{t-1}^2} = \beta + \frac{\sum_{t=1}^{T} x_{t-1} u_t}{\sum_{t=1}^{T} x_{t-1}^2}, \tag{5}$$

is shifted and skewed to the right when σ_{ue} is negative, and hence, $\hat{\beta}$ is positively biased. The main reason of the bias is the strong correlation between u_t and e_t. In other words, the OLSE is asymptotically unbiased if those two innovations are uncorrelated. In the case of infeasible regressions with a known ρ, model (3) can be transformed as

$$r_t = \beta x_{t-1} + \frac{\sigma_{ue}}{\sigma_e^2}(x_t - \rho x_{t-1}) + \left(u_t - \frac{\sigma_{ue}}{\sigma_e^2} e_t\right). \tag{6}$$

Then, by regressing r_t on x_{t-1} and $(x_t - \rho x_{t-1})$, the OLSE of β, denoted by $\hat{\beta}^*$, becomes

$$\hat{\beta}^* = \beta + \frac{\sum_{t=1}^{T} x_{t-1}(u_t - \frac{\sigma_{ue}}{\sigma_e^2} e_t)}{\sum_{t=1}^{T} x_{t-1}^2} + o_p\left(\frac{1}{T}\right). \tag{7}$$

Since $u_t - (\sigma_{ue}/\sigma_e^2)e_t$ is an i.i.d. sequence uncorrelated with a sequence of e_t, which drives x_t, we can show that $\hat{\beta}^*$ is asymptotically unbiased. Moreover, $\hat{\beta}^*$ is more efficient than $\hat{\beta}$ as can be seen from the fact that the variance of the regression error in (6) is $\sigma_u^2 - \sigma_{ue}^2/\sigma_e^2$, which is smaller than σ_u^2 in (3). We may also construct the efficient estimate by ignoring the $o_p(1/T)$ term in (7), which is given by

$$\begin{aligned} \tilde{\beta}^* &= \beta + \frac{\sum_{t=1}^{T} x_{t-1}(u_t - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} e_t)}{\sum_{t=1}^{T} x_{t-1}^2} && (8) \\ &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} \frac{\sum_{t=1}^{T} x_{t-1}(x_t - \rho x_{t-1})}{\sum_{t=1}^{T} x_{t-1}^2}, && (9) \end{aligned}$$

where $\hat{\sigma}_{ue}$ and $\hat{\sigma}_e^2$ are consistent estimates of σ_{ue} and σ_e^2, which can be constructed from the regression residuals in (3) and (4). See also Campbell and Yogo (2006).

In practice, neither $\hat{\beta}^*$ nor $\tilde{\beta}^*$ is feasible because we do not know the true value of ρ. A natural alternative estimate may be obtained by replacing ρ in (9) with some estimate. That is, by letting $\tilde{\rho}$ be an estimate of ρ, we construct a new estimate $\tilde{\beta}$ as

$$\begin{aligned} \tilde{\beta} &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} \frac{\sum_{t=1}^{T} x_{t-1}(x_t - \tilde{\rho} x_{t-1})}{\sum_{t=1}^{T} x_{t-1}^2} \\ &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} \frac{\sum_{t=1}^{T} x_{t-1}\{e_t - (\tilde{\rho} - \rho) x_{t-1}\}}{\sum_{t=1}^{T} x_{t-1}^2} \\ &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} \left\{ \frac{\sum_{t=1}^{T} x_{t-1} e_t}{\sum_{t=1}^{T} x_{t-1}^2} - (\tilde{\rho} - \rho) \right\}. && (10) \end{aligned}$$

Since $\hat{\rho} - \rho = \sum_{t=1}^{T} x_{t-1} e_t / \sum_{t=1}^{T} x_{t-1}^2$, where $\hat{\rho}$ is the OLSE of ρ in (4), $\tilde{\beta}$ becomes

$$\begin{aligned} \tilde{\beta} &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} \{(\hat{\rho} - \rho) - (\tilde{\rho} - \rho)\} \\ &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} (\hat{\rho} - \tilde{\rho}) . \end{aligned} \tag{11}$$

Note that there is no meaning in replacing ρ with the OLSE $\hat{\rho}$ because in this case, $\tilde{\beta}$ becomes equal to $\hat{\beta}$, from expression (11).

Instead of using the OLSE of ρ, we propose to replace ρ with the Cauchy estimate, $\hat{\rho}_c$, by So and Shin (1999), which is obtained by the instrumental variable method for (4) with $\mathrm{sign}(x_{t-1})$ as an instrument, where $\mathrm{sign}(z) = 1$ if $z \geq 0$ and $\mathrm{sign}(z) = 0$ if $z < 0$[3]. The advantage of $\hat{\rho}_c$ over the OLSE $\hat{\rho}$ is that it is an asymptotically median unbiased estimate of ρ even for nearly integrated models. Then, $\tilde{\beta}$ becomes

$$\begin{aligned} \tilde{\beta}_c &= \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} (\hat{\rho} - \hat{\rho}_c) \\ &= \tilde{\beta}^* + \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2} (\hat{\rho}_c - \rho). \end{aligned} \tag{12}$$

Thus, the proposed plug-in estimate can be seen as the sum of the efficient estimate of β and the centered Cauchy estimate of ρ. Note that since $\tilde{\beta}^*$ and $\hat{\rho}_c$ are T consistent for β and ρ, the plug-in estimate is also T consistent.

The limiting properties of the plug-in estimate (mean-adjusted) are investigated in Appendix B. Although the expression of the limiting distribution of $\tilde{\beta}_c$ is complicated, it is almost median unbiased. That is, the limiting probability of taking positive values, $\lim_{T \to \infty} P(T(\tilde{\beta}_c - \beta) \geq 0)$, is almost 0.5. This implies that when the true value of β equals 0, the probability of $\tilde{\beta}_c \geq 0$ is almost the same as the probability of $\tilde{\beta}_c < 0$. This contrasts sharply with the asymptotic property of the OLSE $\hat{\beta}$. It is well known in the literature and

[3] It is also possible to replace ρ with other estimates. We considered the lag-augmented method by Choi (1993) and Toda and Yamamoto (1995) and its modified version by Kurozumi and Yamamoto (2000), combined with several detrending methods. Although the estimates of β based on these methods can be shown to be asymptotically normal, they do not perform better than the estimate based on the Cauchy estimation in finite samples in view of size and power. The weighted symmetric estimate by Park and Fuller (1995) may be a possible estimate, but it assumes normality in e_t. Because this assumption is too strong in empirical finance, we did not consider this estimate.

also confirmed in Appendix B that the OLSE of β is severely biased and that the limiting probability of $T(\hat{\beta}-\beta)\geq 0$ is close to one for the nearly integrated predictors, and as such, we tend to find evidence of predictability even if the true value of β equals zero. Note that the median unbiasedness has been considered as one of the nice properties of estimates in the econometric and statistical literature (in particular, in the unit root literature).

Another nice property of the plug-in estimate is that the limiting distribution of the t-statistic based on the plug-in estimate is well approximated by a standard normal distribution not only for the covariance stationary predictors but also for the nearly integrated variables. As a result, we do not have to tabulate critical values depending on θ; we only need to refer to a table of standard normal distribution and can also easily construct the confidence interval for any significance level. These two nice properties, the median unbiasedness of the plug-in estimate and the asymptotic normality of the t-statistic, are very useful in investigating the predictability of stock returns.

In practice, we usually have a constant term in the regressions, and thus, $\hat{\beta}$ and $\hat{\rho}$ should be obtained from regressions (1) and (2). On the other hand, we do not use the conventional demeaning method for the Cauchy estimation but implement the recursive OLS mean adjustment as suggested by So and Shin (1999). For AR(1) model (2), the recursive OLS mean adjusted process is given by

$$x_t-\bar{x}_{t-1}=\rho(x_{t-1}-\bar{x}_{t-1})+\epsilon_t,$$

where $\bar{x}_{t-1}=\sum_{s=1}^{t-1}x_s/(t-1)$ and $\epsilon_t=e_t+(m_x-\bar{x}_{t-1})(1-\rho)$. Then, the Cauchy estimate is obtained by the instrumental variable method with $\text{sign}(x_{t-1}-\bar{x}_{t-1})$ as an instrument:

$$\hat{\rho}_{c,rols}=\frac{\sum_{t=1}^{T}\text{sign}(x_{t-1}-\bar{x}_{t-1})(x_t-\bar{x}_{t-1})}{\sum_{t=1}^{T}|x_{t-1}-\bar{x}_{t-1}|}. \tag{13}$$

Using $\hat{\rho}_{c,rols}$, we have the following plug-in estimate:

$$\tilde{\beta}_{c,rols}=\hat{\beta}-\frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\left(\hat{\rho}-\hat{\rho}_{c,rols}\right),$$

where $\hat{\sigma}_{ue}$ and $\hat{\sigma}_e^2$ are the consistent estimates of σ_{ue} and σ_e^2 constructed from the regression residuals of (1) and (2). This plug-in estimate with the OLS mean adjustment is valid for both stationary and nearly integrated predictors.

The above recursive OLS mean adjustment is based on the OLS demeaning up to the time $t-1$. However, if we make use of the information that ρ is characterized as a local-to-unity system, we may be able to efficiently estimate ρ. For this purpose, we combine the recursive mean adjustment and the GLS detrending method by Elliott, Rothenberg and Stock (1996). Let x_t^{qd} and 1_t^{qd} be the quasi-differenced series of x_t and 1, that is, $x_t^{qd} = x_1$ and $1_t^{qd} = 1$ for $t = 1$ and $x_t^{qd} = x_t - (1 - 7/T)x_{t-1}$ and $1_t^{qd} = 7/T$ for $t \geq 2$. Then, by regressing x_s^{qd} on 1_s^{qd} for $s = 1, \cdots, t$, we obtain the recursive estimates of m_x at time t, denoted by $\hat{m}_{x,gls,t}$, for $t = 2, \cdots, T$. Using a sequence of these estimates, we construct the Cauchy estimate of ρ with the recursive GLS mean adjustment and then obtain

$$\tilde{\beta}_{c,rgls} = \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\left(\hat{\rho} - \hat{\rho}_{c,rgls}\right),$$

where $\hat{\rho}_{c,rgls}$ is defined in the same way as $\rho_{c,rols}$ in (13) with $\bar{x}_{t-1}$ being replaced by $\hat{m}_{x,rgls,t-1}$. Since ρ is estimated more efficiently, we can expect $\tilde{\beta}_{c,rgls}$ to be more efficient than $\tilde{\beta}_{c,rols}$, which, in fact, is confirmed in the next section and in Appendix B. Note that $\hat{\rho}_{c,rgls}$ is a valid estimate only for nearly integrated predictors because the GLS detrending is meaningful only for the local-to-unity system.

To summarize the asymptotic properties of the two plug-in estimates, both $\tilde{\beta}_{c,rols}$ and $\tilde{\beta}_{c,rgls}$ are almost median unbiased and the distributions of the t-statistics based on these estimates can be approximated by a standard normal distribution. While $\tilde{\beta}_{c,rgls}$ is more efficient than $\tilde{\beta}_{c,rols}$ for nearly integrated predictors, $\tilde{\beta}_{c,rols}$ is valid for both covariance stationary and nearly integrated variables. Because of this trade-off between robustness and efficiency, we recommend using $\tilde{\beta}_{c,rgls}$ if we are confident that ρ is close to unity. Otherwise, we may rely on $\tilde{\beta}_{c,rols}$ to investigate the predictability.

3. Finite Sample Properties

In this section, we investigate the finite sample properties of the estimates developed in the previous section by Monte Carlo simulations. The data generating process is the same as (3) and (4) with $u_t \sim i.i.d.N(0,1)$, $e_t \sim i.i.d.N(0,1)$, $Cov(u_t, e_t) = \sigma_{ue}$, and $x_0 = 0$. We set $\beta = 0$, $\rho = 0.7$, 0.8, 0.9, 0.95, and 0.99; $\sigma_{ue} = -0.95$ and -0.55; and $T = 50$, 100, 250, and 500. Although the true values of μ_r and μ_x equal zero, the models are estimated with a constant term. The number of replications is 10,000 and all computations are conducted by using the GAUSS matrix language.

Table 1 summarizes the simulation results when $\sigma_{ue} = -0.95$. In Panel (a), we report the bias of $\tilde{\beta}_{c,rols}$, $\tilde{\beta}_{c,rgls}$, and $\hat{\beta}$, and the mean squared error (MSE) multiplied by 100 of the corresponding estimates in parentheses. From the table, we can see that the bias is successfully reduced by the new methods. In particular, the plug-in estimate with the recursive OLS mean adjustment is less biased in a wide range of ρ and for $T \geq 100$, while the plug-in estimate with the recursive GLS mean adjustment has a small bias only when ρ is close to unity. This tendency is consistent with the asymptotic theory developed in the previous section. On the other hand, the MSE of the OLSE is smaller than our plug-in estimates when ρ moderately deviates from unity. However, as ρ approaches unity, our plug-in estimates have smaller MSE than the OLSE. Since $\tilde{\beta}_{c,rgls}$ has the smallest MSE when $\rho \geq 0.95$, we can see that $\tilde{\beta}_{c,rgls}$ is more efficient than the other two estimates.

Panel (b) shows the probabilities of taking positive values: $P(\tilde{\beta}_{c,rols} \geq 0)$, $P(\tilde{\beta}_{c,rgls} \geq 0)$, and $P(\hat{\beta} \geq 0)$. We can see that the OLSE of β takes positive values with probabilities greater than 0.8 when ρ is close to one, whereas both $P(\tilde{\beta}_{c,rols} \geq 0)$ and $P(\tilde{\beta}_{c,rgls} \geq 0)$ are close to 0.5, and hence, we confirm that our plug-in estimates are almost median unbiased even in finite samples.

Panel (c) reports the sizes of the one-sided t tests based on $\tilde{\beta}_{c,rols}$ and $\tilde{\beta}_{c,rgls}$ and the Bonferroni Q (BF-Q) test proposed by Campbell and Yogo (2006). As in the table, the size of $t_{c,rols}$ can be well controlled for any ρ and T, while the t test based on $\tilde{\beta}_{c,rgls}$ tends to be slightly oversized when ρ moderately deviates from unity. On the other hand, we can control the size of the BF-Q test only when ρ is close to unity; it is oversized for smaller sample sizes but undersized for larger sample sizes when ρ is not close to one. Again, these properties are consistent with the asymptotic theory; $\tilde{\beta}_{c,rols}$ is robust to the values of ρ while $\tilde{\beta}_{c,rgls}$ and the BF-Q test are valid only for the nearly integrated models.

In Panel (d), we report the coverage rates of the 90% confidence intervals based on $\tilde{\beta}_{c,rols}$, $\tilde{\beta}_{c,rgls}$, and the BF-Q test. We can see that the coverage rates based on our estimates are close to 0.9 for any ρ and T, while those based on the BF-Q test tend to be greater than 0.9 for large sample sizes, even if ρ is close to one.

The above estimates are less biased in the case of $\sigma_{ue} = -0.55$ as compared to in the case of $\sigma_{ue} = -0.95$. However, the sizes and the coverage rates of the tests in the case of $\sigma_{ue} = -0.55$ are similar to those in the case of $\sigma_{ue} = -0.95$. Since the relative performances of the estimates and the tests are preserved in this case, we do not report the details to save space.

Finally, we compare the powers of $t_{c,rols}$, $t_{c,rgls}$, and the BF-Q test. From Figure 1, we can see that the BF-Q test is more powerful than our tests when $\rho = -0.95$, whereas the difference between the powers of the BF-Q test and $t_{c,rgls}$ is only slight when the innovations are moderately negatively correlated. This is because the BF-Q test takes into account the possible range of θ (or ρ) based on the ADF-GLS test by Elliott, Rothenberg and Stock (1996), whereas our plug-in estimate with the recursive GLS mean adjustment uses only the fact that ρ is characterized as the local-to-unity system and does not restrict the possible range of θ. It might be possible to construct a new estimate using the information about θ based on the ADF-GLS test as suggested by Cavanaugh, Elliott and Stock (1995), but such an estimate may have a complicated distribution and may not be median unbiased. Because our main purpose is not to develop a powerful test but to construct reliable estimates, we do not pursue such an estimate in this paper.

To summarize the simulation results, our plug-in estimates perform quite well in finite samples and the t tests based on them are comparable with the existing test in some cases. Because of the good finite sample properties, our plug-in estimates complement the existing testing procedures in empirical analysis.

4. Predictability of U.S. Stock Returns

In this section, we implement our estimation method on U.S. equity data, taking the previous findings of Campbell and Yogo (2006) as the benchmark of the comparison.

4.1. Description of data

We use five different series of the U.S. stock returns. The first three series are the returns on the annual S&P 500 index, the monthly and annual Center for Research in Security Prices (CRSP) data. These return data, along with the financial ratio variables—the dividend-price ratio and the earnings-price ratio—are used as the predictor variables. The series and the same sample period are the same as in Campbell and Yogo (2006)[4]. These data are taken from Motohiro Yogo's website[5]. The estimation results using these data series are presented in Panels A, B, and C of Table 2, respectively. In addition, we also use the updated annual return on the S&P 500 index and the monthly data. These data and the financial ratio variables are taken from Robert Shiller's website[6]. The estimation results using these data series are presented in Panels D and E of Table 2, respectively.

To compute the excess returns of stocks over risk-free return, we use the one-month T-bill rate for monthly series and roll over the one-month T-bill rate for the annual series. In our analysis, we use two additional predictor variables, the three month T-bill rate and the long-short yield spread, following Campbell and Yogo (2006). As in Fama and French (1989) and other previous researches, the long yield used to compute the yield spread is Moody's seasoned Aaa corporate bond yield. The short rate is the one-month T-bill rate. The T-bill rates and the Moody's seasoned Aaa corporate bond yield are from Yogo's and FRB's websites[7]. Following the usual convention, the excess returns and the predictor variables are in logs.

4.2. Persistence of the predictor variables

[4]We use the 1880-2002 and 1880-1994 samples for the annual S&P 500 index; the 1926M12-2002M12, 1926M12-1994M12, and 1952M1-2002M12 samples for the monthly CRSP data; and the 1926-2002, 1926-1994, and 1952-2002 samples for the annual CRSP data.

[5]http://www.nber.org/~myogo/

[6]http://www.econ.yale.edu/~shiller/data.htm

[7]http://www.federalreserve.gov/releases/h15/data.htm

In the fifth and sixth columns of Table 2, we report the estimated coefficients and standard errors for the autoregressive root ρ for the log dividend-price ratio (d-p), the log earnings-price ratio (e-p), the three-month T-bill rate (3my), and the long-short yield spread (ys) using two different methods. As discussed in Section 2, we call them the plug-in estimate with the recursive OLS mean adjustment ($\hat{\rho}_{c,rols}$) and the plug-in estimate with the recursive GLS mean adjustment ($\hat{\rho}_{c,rgls}$). The difference in these two methods is described in Section 2 and the Appendix. The autoregressive lag length $p \in [1, \bar{p}]$ for the predictor variable is determined by the Bayes information criterion (BIC). We set the maximum lag length $\bar{p}$ as four for annual data and eight for monthly data. The estimated lag lengths are reported in the fourth column of Table 2.

As discussed in the literature, the high persistence of these typical predictor variables suggests that the first-order asymptotics, that is, the t-statistic based on the OLSE being approximately normal in large samples, can possibly lead to misleading results. It should also be noted that whether or not the conventional inference based on the t-test is reliable depends on the correlation (δ) between the innovations to the excess returns and to the predictor variable, in addition to the true value of ρ. We report the point estimates of δ in the seventh column of Table 2. The correlations of the innovations to stock returns with the financial ratios are negative and large, but those with the interest rate variables (3my and ys) are much smaller. The former result indicates that the movements in stock returns and in financial ratios mostly come from the movements in the stock price. The latter finding suggests that for the interest rate variables, the conventional t-test perhaps provides approximately valid inference. These findings are essentially the same as those obtained by Campbell and Yogo (2006).

4.3. Estimating and testing the predictability of the U.S. stock returns

Following the methodology discussed in Section 2, we calculate our plug-in estimates ($\tilde{\beta}_{c,rols}$, $\tilde{\beta}_{c,rgls}$) to test the predictability on the U.S. stock returns. In the eighth and ninth columns of Table 2, we report our plug-in estimates and the corresponding p-values. As mentioned in

Section 3, $\tilde{\beta}_{c,rgls}$ is more efficient for persistent series, and hence, we discuss our estimation results based on $\tilde{\beta}_{c,rgls}$. We see from Panel A that the log dividend-price ratio (d-p) has return predictability at the 5% significant level and the log earnings-price ratio (e-p) at the 1% significant level for the U.S. annual S&P 500 index data with the Campbell and Yogo (2006) sample. Drawing a comparison between our findings and those of Campbell and Yogo (2006), we also report the confidence intervals of each of our plug-in estimates in the tenth and eleventh columns of Table 2. In the last column of Table 2, we report the result of the BF-Q test proposed by Campbell and Yogo (2006)[8]. Compared to these confidence intervals, our test based on $\tilde{\beta}_{c,rgls}$ rejects the null of no predictability for the log dividend-price ratio (d-p) and the log earnings-price ratio (e-p), whereas the BF-Q test does not reject the null hypothesis for the log dividend-price ratio (d-p) for the annual S&P 500 index data series in both samples.

For the monthly CRSP series with the Campbell and Yogo (2006) sample, reported in Panel B, the p-values of our test for the log earnings-price ratio (e-p) are almost 5% except for in the 1952M1-2002M12 sample, whereas both tests reject the null for the log dividend-price ratio (d-p). These findings are almost the same as those observed by Campbell and Yogo (2006). For the three-month T-bill rate (3my) and the long-short yield spread (ys), we reject the null hypothesis. Given that the correlation estimate $\hat{\delta}$ reported in the seventh column is negative but very small, the conventional inference based on the t-test leads to an approximately valid inference, as discussed in the previous subsection[9].

In panel C, we report the result for the annual CRSP series with the Campbell and Yogo (2006) sample. We see that the log dividend-price ratio (d-p) has return predictability at the 5% significant level and the log earnings-price ratio (e-p) at the 1% significant level for the 1926-2002 and the 1926-1994 samples, but both predictors are insignificant for the 1952-2002 sample. Compared to these confidence intervals, our test based on $\tilde{\beta}_{c,rgls}$ and the BF-Q test reject the null of no predictability for the log dividend-price ratio (d-p) and the log earnings-price ratio (e-p) for the 1926-2002 and the 1926-1994 samples, but not for the

[8] The figure is somewhat different in Campbell and Yogo (2006), but the conclusion is the same.
[9] This point is also discussed in Campbell and Yogo (2006).

1952-2002 sample.

For the longer and more recent monthly sample, reported in Panel D, we reject the null hypothesis at the 1% level for the log dividend-price ratio (d-p), the log earnings-price ratio (e-p), and the long-short yield spread (ys) based on the $\tilde{\beta}_{c,rgls}$ point estimate. According to the 5% confidence intervals by the plug-in estimates and the BF-Q test, we also reject the null for these three series. Note that there is no difference between the confidence intervals based on $\tilde{\beta}_{c,rgls}$ and the BF-Q test. Judging from these findings, we observe that the log dividend-price ratio (d-p) and the log earnings-price ratio (e-p) are more likely to has predictability on this monthly sample. However, as to the long-short yield spread (ys), there is room for reconsidering the interpretation, as stated above.

For the updated longer annual sample reported in Panel E, we reject the null at the 5% level for the log dividend-price ratio (d-p), but not for the log earnings-price ratio (e-p) based on the $\tilde{\beta}_{c,rgls}$ point estimate. According to the confidence interval based on $\tilde{\beta}_{c,rgls}$, the log dividend-price ratio has predictability, whereas the BF-Q test fails to reject the null.

In summary, with regard to the empirical results using the U.S. equity data series, we get that our plug-in estimates are as well as or better than the BF-Q test by Campbell and Yogo (2006) in finding predictability. Since our estimation method gives asymptotically median unbiased point estimates, our estimates and tests are useful to investigate the predictability of stock returns and complement the existing statistical methods such as the BF-Q test.

5. Conclusion

In this paper, we proposed new point estimates for predictive regressions. They are easily obtained by the least squares and the instrumental variable methods. Our estimates have nice asymptotic properties such as the (almost) median unbiasedness and the approximated normality of the associated t-statistics. In addition, the proposed estimates are shown to have good finite sample properties via Monte Carlo simulations. Using the new estimates, we investigated the U.S. stock returns and found that some variables, which have not been statistically detected as useful predictors in the literature, are able to predict stock returns.

Because of their nice properties, our methods complement the existing statistical tests for predictability to investigate the relations between stock returns and economic variables.

Appendix

A. Construction of the estimates and their variances

A.1. AR(1) case

We first explain the construction of the plug-in estimates for an AR(1) model and then consider a more general AR(p) model. Note that if we consider only the nearly integrated predictors, we do not have to construct the plug-in estimates as given in A.2 but it is sufficient to follow A.1 even when x_t is an AR(p) process. See also Appendix B.2. A.2. is required for $\hat{\beta}_{c,rols}$ to be valid even when x_t is covariance stationary.

1. Estimate (1) and (2) by OLS and obtain $\hat{\beta}$ and $\hat{\rho}$. Based on the regression residuals $\hat{u}_t$ and $\hat{e}_t$, construct $\hat{\sigma}_{ue} = \sum_{t=1}^{T} \hat{u}_t \hat{e}_t / T$, $\hat{\sigma}_e^2 = \sum_{t=1}^{T} \hat{e}_t^2 / T$, and $\hat{\sigma}_u^2 = \sum_{t=1}^{T} \hat{u}_t^2 / T$.

2. Estimate (2) by the instrumental variable method. The estimate with the recursive OLS mean adjustment is defined as

$$\hat{\rho}_{c,rols} = \frac{\sum_{t=2}^{T} \operatorname{sign}(x_{t-1} - \bar{x}_{t-1})(x_t - \bar{x}_{t-1})}{\sum_{t=2}^{T} |x_{t-1} - \bar{x}_{t-1}|},$$

where $\bar{x}_{t-1} = \sum_{s=1}^{t-1} x_s / (t-1)$. For the estimate with the recursive GLS mean adjustment, construct x_t^{qd} and 1_t^{qd} as follows:

$$x_t^{qd} = \begin{cases} x_1 & : \quad t = 1 \\ x_t - \frac{7}{T} x_{t-1} & : \quad t \geq 2, \end{cases} \qquad 1_t^{qd} = \begin{cases} 1 & : \quad t = 1 \\ \frac{7}{T} & : \quad t \geq 2. \end{cases}$$

Obtain $\hat{m}_{x,gls,t}$ by regressing x_s^{qd} on 1_s^{qd} for $s = 1, \cdots, t$:

$$\hat{m}_{x,gls,t} = \frac{\sum_{s=1}^{t} 1_s^{qd} x_s^{qd}}{\sum_{s=1}^{t} (1_s^{qd})^2}.$$

The Cauchy estimate of ρ with the recursive GLS mean adjustment is defined as

$$\hat{\rho}_{c,rgls} = \frac{\sum_{t=2}^{T} \operatorname{sign}(x_{t-1} - \hat{m}_{x,gls,t-1})(x_t - \hat{m}_{x,gls,t-1})}{\sum_{t=2}^{T} |x_{t-1} - \hat{m}_{x,gls,t-1}|}.$$

3. The plug-in estimates of β are obtained as

$$\tilde{\beta}_{c,rols} = \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\left(\hat{\rho} - \tilde{\rho}_{c,rols}\right) \quad \text{and} \quad \tilde{\beta}_{c,rgls} = \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\left(\hat{\rho} - \tilde{\rho}_{c,rols}\right).$$

4. The variance estimates of $\tilde{\beta}_{c,rols}$ and $\tilde{\beta}_{c,rgls}$ are defined as the sum of the variances of the efficient estimates $\hat{\beta}^*$ and the Cauchy estimates of ρ. That is,

$$Var(\tilde{\beta}_{c,rols}) = Var(\hat{\beta}^*) + \left(\frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\right)^2 Var(\hat{\rho}_{c,r,ols}), \tag{14}$$

$$\text{where} \quad Var(\hat{\beta}^*) = \frac{\hat{\sigma}_e^2\hat{\sigma}_u^2 - \hat{\sigma}_{ue}^2}{\hat{\sigma}_e^2\sum_{t-1}^T \tilde{x}_{t-1}^2} \quad \text{and} \quad Var(\hat{\rho}_{c,rols}) = \frac{\sum_{t=2}^T \hat{e}_{c,rols,t}^2}{\left(\sum_{t=2}^T |x_{t-1} - \bar{x}_{t-1}|\right)^2},$$

where $\tilde{x}_t = x_t - \bar{x}$, and $\hat{e}_{c,rols,t} = (x_t - \bar{x}_{t-1}) - \hat{\rho}_{c,rols}(x_{t-1} - \bar{x}_{t-1})$ is the residual from the instrumental variable estimation. Similarly, the variance of the plug-in estimate with the recursive GLS mean adjustment is obtained by replacing $Var(\hat{\rho}_{c,r,ols})$ with

$$Var(\hat{\rho}_{c,rgls}) = \frac{\sum_{t=2}^T \hat{e}_{c,rgls,t}^2}{\left(\sum_{t=2}^T |x_{t-1} - \hat{m}_{x,gls,t-1}|\right)^2},$$

where $\hat{e}_{c,rgls,t} = (x_t - \hat{m}_{x,gls,t-1}) - \hat{\rho}_{c,rgls}(x_{t-1} - \hat{m}_{x,gls,t-1})$ is the residual from the Cauchy estimation. We can construct $t_{c,rols}$ and $t_{c,rgls}$ by using $\tilde{\beta}_{c,rols}$, $Var(\tilde{\beta}_{c,rols})$, $\tilde{\beta}_{c,rgls}$, and $Var(\tilde{\beta}_{c,rgls})$.

A.2. AR(p) case

We now consider the case where x_t is an AR(p) process given by

$$x_t = \mu_x + \rho x_{t-1} + \psi_1 \Delta x_{t-1} + \cdots + \psi_{p-1}\Delta x_{t-p+1} + e_t, \tag{15}$$

where $\mu_x = m_x(1-\rho)$, Δ is a differencing operator, and $\rho = 1 - \theta/T$.

1. As in the AR(1) case, estimate (1) and (15) by OLS and obtain $\hat{\beta}$, $\hat{\sigma}_{ue}$, $\hat{\sigma}_e^2$, and $\hat{\sigma}_u^2$.

2. In order to estimate (15) by the instrumental variable method, express (15) in the recursive OLS mean adjustment form as follows:

$$x_t - \bar{x}_{t-1} = \rho(x_{t-1} - \bar{x}_{t-1}) + +\psi_1 \Delta x_{t-1} + \cdots + \psi_{p-1}\Delta x_{t-p+1} + \epsilon_t. \tag{16}$$

Obtain the instrumental variable estimates $\hat{\rho}_{c,rols}, \hat{\psi}_{1,c,rols}, \cdots, \hat{\psi}_{p-1,c,rols}$ with sign$(x_t - \bar{x}_{t-1})$ and $\Delta x_{t-1}, \cdots, \Delta x_{t-p+1}$ as instruments. Denote the estimate of the $p \times p$ variance covariance matrix by the instrumental variable estimation as $\hat{\Sigma}_{c,rols}$. Similarly, $\hat{\rho}_{c,rgls}$ and $\hat{\Sigma}_{c,rgls}$ are obtained by replacing $\bar{x}_{t-1}$ in (16) with $\hat{m}_{x,gls,t-1}$, as in the AR(1) case.

3. Let $\tilde{e}_{c,rols,t} = \tilde{x}_t - \hat{\rho}_{c,rols}\tilde{x}_{t-1} - \hat{\psi}_{1,c,rols}\Delta x_{t-1} - \cdots - \hat{\psi}_{p-1,c,rols}\Delta x_{t-p+1}$, where $\tilde{x}_t$ and $\tilde{x}_{t-1}$ are obtained by regressions x_t and x_{t-1} on a constant. Similarly, define $\tilde{e}_{c,rgls,t}$ by replacing $\hat{\rho}_{c,rols}, \hat{\psi}_{1,c,rols}, \cdots, \hat{\psi}_{p-1,c,rols}$ in $\tilde{e}_{c,rols,t}$ with $\hat{\rho}_{c,rgls}, \hat{\psi}_{1,c,rgls}, \cdots, \hat{\psi}_{p-1,c,rgls}$. Then, the plug-in estimates are defined as

$$\tilde{\beta}_{c,rols} = \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\frac{\sum_{t=2}^T \tilde{x}_{t-1}\tilde{e}_{c,rols,t}}{\sum_{t=2}^T \tilde{x}_{t-1}^2} \quad \text{and} \quad \tilde{\beta}_{c,rgls} = \hat{\beta} - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\frac{\sum_{t=2}^T \tilde{x}_{t-1}\tilde{e}_{c,rgls,t}}{\sum_{t=2}^T \tilde{x}_{t-1}^2}.$$

4. The variance estimate of $\tilde{\beta}_{c,rols}$ is given by (14) with the same $Var(\hat{\beta}^*)$ but with $Var(\hat{\rho}_{c,rols})$ being replaced by

$$Var(\hat{\rho}_{c,rols}) = J'\hat{\Sigma}_{c,rols}J \quad \text{where} \quad J = \left[1, \frac{\sum_{t=1}^T \tilde{x}_{t-1}\Delta x_{t-1}}{\sum_{t=1}^T \tilde{x}_{t-1}^2}, \cdots, \frac{\sum_{t=1}^T \tilde{x}_{t-1}\Delta x_{t-p+1}}{\sum_{t=1}^T \tilde{x}_{t-1}^2}\right].$$

Similarly, $Var(\hat{\rho}_{c,rgls})$ is obtained by replacing $\hat{\Sigma}_{c,rols}$ with $\hat{\Sigma}_{c,rgls}$.

B. Asymptotic distributions of the plug-in estimates

B.1. Case where $\rho < 1$ is fixed

For the AR(1) case, as given in (12), $\tilde{\beta}_{c,rols}$ is expressed as

$$(\tilde{\beta}_{c,rols} - \beta) = (\tilde{\beta}^* - \beta) + \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}(\hat{\rho}_{c,rols} - \rho). \tag{17}$$

From (8), the weak law of large numbers (WLLN), and the central limit theorem (CLT), we can see that

$$\sqrt{T}(\tilde{\beta}^* - \beta) = \frac{\frac{1}{\sqrt{T}}\sum_{t=1}^T (x_{t-1} - m_x)(u_t - \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}e_t)}{\frac{1}{T}\sum_{t=1}^T (x_{t-1} - m_x)^2} + o_p(1) \xrightarrow{d} N\left(0, \frac{\sigma_{u\cdot e}^2}{\gamma_x}\right), \tag{18}$$

where $\sigma_{u\cdot e}^2 = \sigma_u^2 - \sigma_{ue}^2/\sigma_e^2$ and $\gamma_x = Var(x_t)$, while

$$\sqrt{T}(\hat{\rho}_{c,rols} - \rho) = \frac{\frac{1}{\sqrt{T}}\sum_{t=2}^{T} \mathrm{sign}(x_{t-1} - m_x)e_t}{\frac{1}{T}\sum_{t=2}^{T} |x_{t-1} - m_x|} + o_p(1) \xrightarrow{d} N\left(0, \frac{\sigma_e^2}{(E|x_t - m_x|)^2}\right). \quad (19)$$

Moreover, since a sequence of $[(x_{t-1} - m_x)\{u_t - (\sigma_{ue}/\sigma_e)e_t\},\ \mathrm{sign}(x_{t-1} - m_x)e_t]$ for $t = 2, \cdots, T$ forms a martingale difference sequence with zero covariance, we can see that the limiting distributions (18) and (19) are independent. As a result, we have

$$\sqrt{T}(\tilde{\beta}_{c,rols} - \beta) \xrightarrow{d} N\left(0, \frac{\sigma_{u\cdot e}^2}{\gamma_x} + \frac{\sigma_{ue}^2}{\sigma_e^2}\frac{1}{(E|x_t - m_x|)^2}\right),$$

and then, we can see that the associated t-statistic is asymptotically distributed as $N(0, 1)$.

For the AR(p) case, $\tilde{\beta}_{c,rols}$ is expressed as

$$(\tilde{\beta}_{c,rols} - \beta) = (\tilde{\beta}^* - \beta) + \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\frac{\sum_{t=2}^{T}\tilde{x}_{t-1}\tilde{\eta}_{t-1}'(\hat{\phi}_{c,rols} - \phi)}{\sum_{t=2}^{T}\tilde{x}_{t-1}^2} + o_p\left(\frac{1}{\sqrt{T}}\right), \quad (20)$$

where $\tilde{\eta}_{t-1} = [\tilde{x}_{t-1}, \Delta x_{t-1}, \cdots, \Delta x_{t-p+1}]$, $\phi = [\rho, \psi_1, \cdots, \psi_{p-1}]$, and $\hat{\phi}_{c,rols} = [\hat{\rho}_{c,rols}, \hat{\psi}_{1,c,rols}, \cdots, \hat{\psi}_{p-1,c,rols}]$. Then, we can see that (18) still holds whereas the second term on the right-hand side multiplied by $\sqrt{T}$ becomes

$$\begin{aligned} \sqrt{T}\frac{\sum_{t=2}^{T}\tilde{x}_{t-1}\tilde{\eta}_{t-1}'(\hat{\phi}_{c,rols} - \phi)}{\sum_{t=2}^{T}\tilde{x}_{t-1}^2} &= J'\sqrt{T}(\hat{\phi}_{c,rols} - \phi) \\ &\xrightarrow{d} N\left(0, E[J']\Sigma_{c,rols}E[J]\right), \end{aligned} \quad (21)$$

because $J \xrightarrow{p} E[J]$ by the WLLN, while letting $\eta_{t-1} = [x_{t-1} - \bar{x}_{t-1}, \Delta x_{t-1}, \cdots, \Delta x_{t-p+1}]$ and $\eta_{c,t-1} = [\mathrm{sign}(x_{t-1} - \bar{x}_{t-1}), \Delta x_{t-1}, \cdots, \Delta x_{t-p+1}]$, we have

$$\sqrt{T}(\hat{\phi}_{c,rols} - \phi) = \left(\frac{1}{T}\sum_{t=2}^{T}\eta_{c,t-1}\eta_{t-1}'\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\eta_{c,t-1}e_t\right) + o_p(1) \xrightarrow{d} N\left(0, \Sigma_{c,rols}\right), \quad (22)$$

by the WLLN and the CLT, where $\Sigma_{c,rols} = \sigma_e^2(E[\eta_{c,t}\eta_t'])^{-1}E[\eta_{c,t}\eta_{c,t}'](E[\eta_t\eta_{c,t}'])^{-1}$. Then, from (18) and (22), we conclude that

$$\sqrt{T}(\tilde{\beta}_{c,rols} - \beta) \xrightarrow{d} N\left(0, \frac{\sigma_{u\cdot e}^2}{\gamma_x} + \frac{\sigma_{ue}^2}{\sigma_e^4}E[J']\Sigma_{c,rols}E[J]\right),$$

and that the associated t-statistic weakly converges to a standard normal distribution.

B.2. Case where $\rho = 1 - \theta/T$

For the AR(1) case, we have, by the functional central limit theorem,

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]}(u_t - \frac{\sigma_{ue}}{\sigma_e^2}e_t) \xrightarrow{d} W_{u\cdot e}(s), \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{[Ts]} e_t \xrightarrow{d} W_e(s), \quad \text{and} \quad \frac{1}{\sqrt{T}}(x_{[Ts]} - m_x) \xrightarrow{d} W_\theta(s) \tag{23}$$

jointly for $0 \leq s \leq 1$, where $W_{u\cdot e}(s)$ and $W_e(s)$ are Brownian motions with variances $\sigma_{u\cdot e}^2$ and σ_e^2, respectively, $W_\theta(s)$ is the Ornstein-Uhlenbeck process defined by $W_\theta(s) = \int_0^s e^{\theta(s-t)}dW_e(t)$ and $W_{u\cdot e}(s)$ is independent of $W_e(s)$ and $W_\theta(s)$. Using (23), we can see that

$$T(\tilde{\beta}^* - \beta) \xrightarrow{d} \frac{\int_0^1 \tilde{W}_\theta(s)dW_{u\cdot e}(s)}{\int_0^1 \tilde{W}_\theta^2(s)ds}, \tag{24}$$

where $\tilde{W}_\theta(s) = W_\theta(s) - \int_0^1 W_\theta(t)dt$. On the other hand, as in So and Shin (1999), we have

$$T(\hat{\rho}_{c,rols} - \rho) = \frac{\frac{1}{\sqrt{T}}\sum_{t=2}^T \mathrm{sign}(x_{t-1} - \bar{x}_{t-1})e_t}{\frac{1}{T\sqrt{T}}\sum_{t=2}^T |x_{t-1} - \bar{x}_{t-1}|} + o_p(1) \xrightarrow{d} \frac{\int_0^1 \mathrm{sign}(\tilde{W}_{\theta,c}(s))dW_e(s)}{\int_0^1 |\tilde{W}_{\theta,c}(s)|ds} \tag{25}$$

from (23) and the continuous mapping theorem, where $\tilde{W}_{\theta,c}(s) = W_\theta(s) - (1/s)\int_0^s W_\theta(t)dt$. From (24) and (25), the plug-in estimate with the recursive OLS mean adjustment weakly converges to

$$T(\tilde{\beta}_{c,rols} - \beta) \xrightarrow{d} \frac{\int_0^1 \tilde{W}_\theta(s)dW_{u\cdot e}(s)}{\int_0^1 \tilde{W}_\theta^2(s)ds} + \frac{\sigma_{ue}}{\sigma_e^2}\frac{\int_0^1 \mathrm{sign}(\tilde{W}_{\theta,c}(s))dW_e(s)}{\int_0^1 |\tilde{W}_{\theta,c}(s)|ds}. \tag{26}$$

For the AR(p) case, we can see that as in (20),

$$T(\tilde{\beta}_{c,rols} - \beta) = T(\tilde{\beta}^* - \beta) + \frac{\hat{\sigma}_{ue}}{\hat{\sigma}_e^2}\frac{\frac{1}{T}\sum_{t=2}^T \tilde{x}_{t-1}\tilde{\eta}_{t-1}'(\hat{\phi}_{c,rols} - \phi)}{\frac{1}{T^2}\sum_{t=2}^T \tilde{x}_{t-1}^2} + o_p(1). \tag{27}$$

In this case, since the weak convergences in (23) hold with $W_\theta(s)$ being redefined by $W_\theta(s) = \int_0^s e^{\theta(s-t)}dW_x(t)$, where $W_x(t)$ is a Brownian motion with the long-run variance induced from (15), we can see that (24) still holds. On the other hand, as in So and Shin (1999), we can see that

$$T(\hat{\rho}_{c,rols} - \rho) = \frac{\frac{1}{\sqrt{T}}\sum_{t=2}^T \mathrm{sign}(x_{t-1} - \bar{x}_{t-1})e_t}{\frac{1}{T\sqrt{T}}\sum_{t=2}^T |x_{t-1} - \bar{x}_{t-1}|} + o_p(1) \xrightarrow{d} \frac{\int_0^1 \mathrm{sign}(\tilde{W}_{\theta,c(s)})dW_e(s)}{\int_0^1 |\tilde{W}_{\theta,c}(s)|ds}, \tag{28}$$

whereas $\hat{\psi}_{j,c,rols} - \psi_j = O_p(1/\sqrt{T})$ for $j = 1, \cdots, p-1$. Then, since $\sum_{t=2}^{T} \tilde{x}_{t-1}^2 = O_p(T^2)$ while $\sum_{t=2}^{T} \tilde{x}_{t-1} \Delta x_{t-j} = O_p(T)$ for $j = 1, \cdots, p-1$, we can show that the second term on the right-hand side of (27) becomes

$$\frac{\frac{1}{T}\sum_{t=2}^{T} \tilde{x}_{t-1} \tilde{\eta}_{t-1}'(\hat{\phi}_{c,rols} - \phi)}{\frac{1}{T^2}\sum_{t=2}^{T} \tilde{x}_{t-1}^2} = T(\hat{\rho}_{c,rols} - \rho) + o_p(1). \tag{29}$$

Then, from (27), (28), and (29), we have the same convergence as given by (26) with $W_\theta(s) = \int_0^s e^{\theta(s-t)} dW_x(t)$.

For the plug-in estimate with the recursive GLS mean adjustment, we can show that

$$\frac{1}{\sqrt{T}}(x_{[Ts]} - \hat{m}_{x,gls,t-1}) \xrightarrow{d} W_\theta(s) \tag{30}$$

for both the AR(1) and AR(p) cases. As a result, we can see that

$$T(\tilde{\beta}_{c,rgls} - \beta) \xrightarrow{d} \frac{\int_0^1 \tilde{W}_\theta(s) dW_{u\cdot e}(s)}{\int_0^1 \tilde{W}_\theta^2(s) ds} + \frac{\sigma_{ue}^2}{\sigma_e^2} \frac{\int_0^1 \mathrm{sign}(W_\theta(s)) dW_e(s)}{\int_0^1 |W_\theta(s)| ds}. \tag{31}$$

In order to see the asymptotic properties of the plug-in estimates, we draw the probability density functions of the limiting distributions given by (26) and (31) for various values of θ with $\sigma_u^2 = \sigma_e^2 = 1$ and $\sigma_{ue} = -0.95$ and -0.55 in Figure 2[10]. As is well known in the literature, the OLS estimate is severely biased when $d = -0.95$ and θ is small. However, our plug-in estimates are located to the left, and hence, the biases of the plug-in estimates are smaller than that of the OLS estimate. Moreover, the plug-in estimates are almost median unbiased as we have $\lim P(T(\tilde{\beta}_{c,rols} - \beta) \geq 0) \simeq 0.5$ and $\lim P(T(\tilde{\beta}_{c,rgls} - \beta) \geq 0) \simeq 0.5$ in Table 3. Note that the median unbiasedness has been considered as one of the desirable properties of estimates in the econometric and statistical literature; in fact, the median unbiased estimates of the AR(1) coefficient have been developed by, for example, Andrews (1993) and So and Shin (1999). We can also see from Figure 2 that $\tilde{\beta}_{c,rgls}$ is more efficient than $\tilde{\beta}_{c,rols}$ because the probability density function of a normalized $\tilde{\beta}_{c,rgls}$ is more concentrated around the 0 axis than $\tilde{\beta}_{c,rols}$.

[10] Brownian motions are approximated by the scaled partial sums from 2,000 i.i.d. standard normal random variables. The densities are drawn for the range 1-99% points by the kernel method with a Gaussian kernel. The smoothing parameter, h, is decided by equation (3.31) in Silverman (1986): $h = 0.9AT^{-1/5}$, where $A = \min(\text{standard deviation}, \text{interquartile range}/1.34)$.

We now investigate the deviation of the t-statistics based on the plug-in estimates from the standard normal distribution. Figure 3 draws the q-q plots of the limiting distributions of the t-statistics against a standard normal distribution. From the figure, we can see that the q-q plots of the t-statistics are almost linear, which implies that the limiting distributions of the t-statistics are well approximated by $N(0, 1)$. In particular, the q-q plots based on the plug-in estimate with the recursive GLS mean adjustment are almost on the diagonal line. On the other hand, the limiting distribution of the t-statistic based on the plug-in estimate with the recursive OLS mean adjustment are located slightly to the left as compared to the standard normal distribution when θ is large. This implies that the one-sided test based on this t-statistic would be slightly conservative if the critical values of $N(0, 1)$ are used.

References

[1] Andrews, D. W. K. (1993). Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models. *Econometrica* 61, 139-165.

[2] Campbell, J. Y. and M. Yogo (2006). Efficient Tests if Stick Return Predictability. *Journal of Financial Economics* 81, 27-60.

[3] Cavanagh, C. L., G. Elliott and J. H. Stock (1994). Inference in Models with Nearly Integrated Regressors. *Econometric Theory* 11, 1131-1147.

[4] Choi, I. (1993). Asymptotic Normality of the Lest-Squares Estimates for Higher Order Autoregressive Integrated Processes with Some Applications. *Econometric Theory* 9, 263-282.

[5] Elliott, G., T. J. Rothenberg and J. Stock (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64, 813-836.

[6] Elliott, G. and J. H. Stock (1994). Inference in Time Series Regression When the Order of Integration of a Regressor is Unknown. *Econometric Theory* 10, 672-700.

[7] Fama, E. F. and K. R. French (1989). Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics* 25, 23-49.

[8] Jansson, M. and M. J. Moreira (2006). Optimal Inference in Regression Models with Nearly Integrated Regressors. *Econometrica* 74, 681-714.

[9] Kurozumi, E. and T. Yamamoto (2000). Modified Lag Augmented Vector Autoregressions. *Econometric Reviews* 19, 207-231.

[10] Lanne, M. (2002). Testing the Predictability of Stock Returns. *Review of Economics and Statistics* 84, 407-415.

[11] Lewellen, J. (2004). Predicting Returns with Financial Ratios. *Journal of Financial Economics* 74, 209-235.

[12] Mankiw, N. G. and M. D. Shapiro (1986). Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models. *Economics Letters* 20, 139-145.

[13] Park, H. J. and W. A. Fuller (1995). Alternative Estimators and Unit Root Tests for the Autoregressive Process. *Journal of Time Series Analysis* 16, 415-429.

[14] Phillips, P. C. B. (1987). Towards a Unified Asymptotic Theory for Autoregression. *Biometrika* 74, 535-547.

[15] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

[16] So, B. S. and D. W. Shin (1999). Cauchy Estimators for Autoregressive Processes with Applications to Unit Root Tests and Confidence Intervals. *Econometric Theory* 15, 165-176.

[17] Stambaugh, R. F. (1999). Predictive Regressions. *Journal of Financial Economics* 54, 375-421.

[18] Toda, H. Y. and T. Yamamoto (1995). Statistical Inference in Vector Autoregressions with Possibly Integrated Processes. *Journal of Econometrics* 66, 225-250.

[19] Torous, W., R. Valkanov and S. Yan (2004). On Predicting Stock Returns with Nearly Integrated Explanatory Variables. *Journal of Business* 77, 937-966.

Table 1: Finite Sample Properties of the Tests

(a) Bias and MSE of the estimates

	$T = 50$			$T = 100$			$T = 250$			$T = 500$		
ρ	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$
0.70	0.011	0.049	0.061	0.001	0.025	0.029	−0.002	0.011	0.012	−0.002	0.006	0.006
	(2.149)	(2.390)	(1.702)	(0.905)	(0.972)	(0.665)	(0.339)	(0.348)	(0.232)	(0.161)	(0.164)	(0.111)
0.80	0.015	0.041	0.069	0.003	0.021	0.033	−0.002	0.008	0.013	−0.002	0.004	0.007
	(1.755)	(1.843)	(1.618)	(0.700)	(0.722)	(0.565)	(0.248)	(0.249)	(0.178)	(0.116)	(0.116)	(0.082)
0.90	0.022	0.035	0.081	0.007	0.017	0.038	0.000	0.007	0.015	−0.001	0.004	0.007
	(1.346)	(1.273)	(1.578)	(0.460)	(0.442)	(0.465)	(0.143)	(0.141)	(0.118)	(0.065)	(0.063)	(0.049)
0.95	0.025	0.031	0.091	0.011	0.015	0.043	0.002	0.006	0.016	0.000	0.003	0.008
	(1.067)	(0.956)	(1.607)	(0.335)	(0.290)	(0.432)	(0.090)	(0.082)	(0.089)	(0.037)	(0.034)	(0.032)
0.99	0.024	0.025	0.097	0.013	0.013	0.049	0.005	0.005	0.020	0.002	0.003	0.010
	(0.834)	(0.702)	(1.614)	(0.227)	(0.174)	(0.430)	(0.046)	(0.034)	(0.077)	(0.014)	(0.011)	(0.020)

(b) Probabilities of taking positive values

	$T = 50$			$T = 100$			$T = 250$			$T = 500$		
ρ	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$
0.70	0.491	0.594	0.673	0.473	0.576	0.621	0.469	0.556	0.575	0.475	0.549	0.558
0.80	0.492	0.575	0.720	0.473	0.563	0.653	0.462	0.540	0.599	0.465	0.534	0.575
0.90	0.508	0.565	0.804	0.482	0.548	0.726	0.454	0.532	0.647	0.464	0.529	0.614
0.95	0.523	0.553	0.884	0.502	0.544	0.808	0.468	0.530	0.714	0.465	0.525	0.663
0.99	0.513	0.524	0.940	0.525	0.529	0.931	0.534	0.526	0.891	0.509	0.521	0.828

Table 1: (Continued)

(c) Size of the one-sided tests

	$T=50$			$T=100$			$T=250$			$T=500$		
ρ	$t_{c,rols}$	$t_{c,rgls}$	BF-Q	$t_{c,rols}$	$t_{c,rgls}$	BF-Q	$t_{c,rols}$	$t_{c,rgls}$	BF-Q	$t_{c,rols}$	$t_{c,rgls}$	BF-Q
0.70	0.050	0.077	0.163	0.046	0.072	0.099	0.044	0.064	0.000	0.042	0.060	0.000
0.80	0.055	0.072	0.125	0.047	0.068	0.103	0.043	0.061	0.000	0.040	0.058	0.000
0.90	0.061	0.065	0.106	0.052	0.061	0.076	0.043	0.059	0.059	0.040	0.057	0.000
0.95	0.059	0.064	0.102	0.053	0.057	0.069	0.048	0.056	0.054	0.043	0.053	0.043
0.99	0.054	0.057	0.096	0.050	0.056	0.062	0.053	0.051	0.052	0.046	0.056	0.051

(d) Coverage rates of 90% confidence intervals

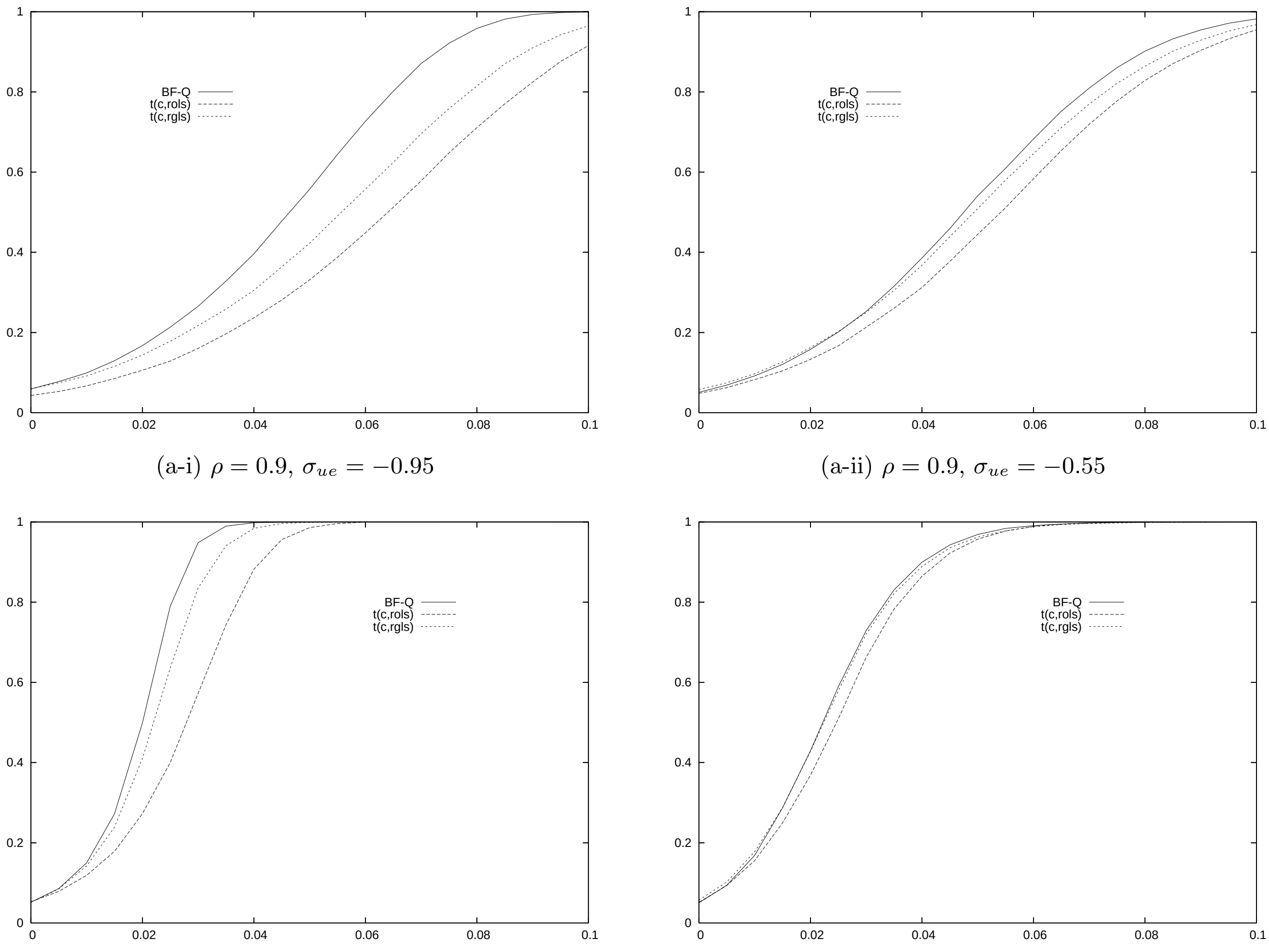
	$T=50$			$T=100$			$T=250$			$T=500$		
ρ	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	BF-Q	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	BF-Q	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	BF-Q	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	BF-Q
0.70	0.893	0.897	0.829	0.895	0.899	0.894	0.891	0.897	0.101	0.898	0.900	0.000
0.80	0.888	0.902	0.866	0.895	0.901	0.887	0.889	0.899	0.990	0.897	0.900	0.000
0.90	0.888	0.910	0.886	0.889	0.905	0.913	0.891	0.899	0.928	0.894	0.901	0.984
0.95	0.905	0.912	0.894	0.897	0.910	0.922	0.887	0.904	0.932	0.894	0.906	0.939
0.99	0.910	0.903	0.903	0.916	0.903	0.934	0.906	0.910	0.937	0.906	0.904	0.937

Table 2: Empirical Results for Stock Returns

r_t	x_t	sample	p	$\hat{\rho}_{c,rols}$ (std)	$\hat{\rho}_{c,rgls}$ (std)	$\hat{\delta}$	$\tilde{\beta}_{c,rols}$ (p-value)	$\tilde{\beta}_{c,rgls}$ (p-value)	CI by $\tilde{\beta}_{c,rols}$	CI by $\tilde{\beta}_{c,rgls}$	CI by BF-Q
Panel A	d-p	1880-2002	3	0.932 (0.084)	0.859 (0.086)	−0.851	0.092 (0.071)	0.142 (0.014)	(−0.011, 0.196)	(0.035, 0.248)	(−0.020, 0.142)
CY(SP & 500,Y)		1880-1994	3	0.905 (0.103)	0.770 (0.104)	−0.842	0.090 (0.117)	0.174 (0.012)	(−0.034, 0.214)	(0.047, 0.301)	(−0.017, 0.236)
	e-p	1880-2002	1	0.846 (0.064)	0.832 (0.065)	−0.962	0.146 (0.010)	0.160 (0.006)	(0.043, 0.249)	(0.055, 0.266)	(0.043, 0.225)
		1880-1994	1	0.805 (0.076)	0.784 (0.078)	−0.958	0.185 (0.007)	0.205 (0.004)	(0.063, 0.308)	(0.079, 0.332)	(0.092, 0.323)
Panel B	d-p	1926.12-2002.12	2	1.000 (0.007)	0.994 (0.005)	−0.951	−0.003 (0.679)	0.003 (0.283)	(−0.013, 0.008)	(−0.006, 0.012)	(−0.005, 0.011)
CY(CRSP,M)		1926.12-1994.12	2	0.999 (0.009)	0.990 (0.007)	−0.949	−0.008 (0.811)	0.001 (0.467)	(−0.021, 0.007)	(−0.011, 0.012)	(−0.007, 0.017)
		1952.01-2002.12	1	0.998 (0.007)	0.999 (0.004)	−0.967	0.006 (0.190)	0.005 (0.125)	(−0.005, 0.017)	(−0.002, 0.011)	(−0.004, 0.010)
	e-p	1926.12-2002.12	1	0.999 (0.007)	0.993 (0.006)	−0.987	0.003 (0.310)	0.009 (0.058)	(−0.008, 0.014)	(−0.000, 0.019)	(0.001, 0.018)
		1926.12-1994.12	2	0.997 (0.009)	0.988 (0.008)	−0.984	0.005 (0.278)	0.013 (0.050)	(−0.009, 0.019)	(0.000, 0.026)	(0.006, 0.029)
		1952.01-2002.12	1	0.998 (0.006)	0.998 (0.005)	−0.982	0.003 (0.326)	0.003 (0.301)	(−0.008, 0.013)	(−0.006, 0.011)	(−0.004, 0.012)
	3my	1952.01-2002.12	2	0.998 (0.010)	0.999 (0.005)	−0.071	−0.018 (0.994)	−0.018 (0.995)	(−0.030, −0.006)	(−0.030, −0.006)	(−0.029, −0.006)
	ys	1952.01-2002.12	1	0.947 (0.020)	0.964 (0.019)	−0.066	0.045 (0.002)	0.044 (0.003)	(0.019, 0.071)	(0.018, 0.070)	(0.020, 0.072)
Panel C	d-p	1926-2002	1	0.972 (0.065)	0.935 (0.058)	−0.721	0.096 (0.048)	0.123 (0.011)	(0.001, 0.191)	(0.035, 0.211)	(0.013, 0.185)
CY(CRSP,Y)		1926-1994	1	0.930 (0.088)	0.895 (0.079)	−0.693	0.144 (0.034)	0.169 (0.012)	(0.014, 0.274)	(0.047, 0.291)	(0.059, 0.333)
		1952-2002	1	1.013 (1.038)	1.038 (0.054)	−0.749	0.069 (0.141)	0.051 (0.171)	(−0.037, 0.139)	(−0.037, 0.139)	(−0.006, 0.179)
	e-p	1926-2002	1	0.896 (0.078)	0.817 (0.078)	−0.957	0.130 (0.050)	0.206 (0.004)	(0.080, 0.332)	(0.080, 0.332)	(0.041, 0.272)
		1926-1994	1	0.852 (0.102)	0.736 (0.102)	−0.959	0.192 (0.031)	0.303 (0.001)	(0.138, 0.468)	(0.138, 0.468)	(0.131, 0.455)
		1952-2002	1	0.946 (0.086)	0.905 (0.079)	−0.955	0.059 (0.242)	0.098 (0.102)	(−0.029, 0.226)	(−0.029, 0.226)	(−0.032, 0.225)
	3my	1952-2002	1	0.921 (0.123)	0.948 (0.091)	0.006	−0.094 (0.877)	−0.094 (0.877)	(−0.228, 0.040)	(−0.228, 0.040)	(−0.228, 0.039)
	ys	1952-2002	1	0.490 (0.163)	0.508 (0.157)	−0.243	0.161 (0.093)	0.156 (0.098)	(−0.043, 0.356)	(−0.043, 0.356)	(−0.073, 0.354)
Panel D	d-p	1926.01-2009.12	2	0.996 (0.004)	0.996 (0.004)	−0.560	0.032 (0.000)	0.032 (0.000)	(0.026, 0.037)	(0.026, 0.037)	(0.026, 0.038)
US(M)	e-p	1926.01-2009.12	3	0.998 (0.004)	0.995 (0.003)	−0.556	0.026 (0.000)	0.028 (0.000)	(0.020, 0.032)	(0.022, 0.033)	(0.021, 0.034)
	3my	1926.12-2009.12	2	1.000 (0.006)	0.997 (0.005)	0.011	−0.003 (0.795)	−0.003 (0.797)	(−0.010, 0.003)	(−0.010, 0.003)	(−0.010, 0.003)
	ys	1926.01-2009.12	8	0.990 (0.007)	0.990 (0.006)	0.018	0.105 (0.000)	0.105 (0.000)	(0.100, 0.110)	(0.100, 0.110)	(0.100, 0.110)
Panel E	d-p	1881-2009	3	0.914 (0.066)	0.818 (0.068)	−0.835	0.037 (0.247)	0.108 (0.026)	(−0.052, 0.126)	(0.016, 0.199)	(−0.055, 0.115)
US(Y)	e-p	1881-2009	1	0.899 (0.053)	0.905 (0.051)	−0.932	0.066 (0.099)	0.060 (0.116)	(−0.019, 0.151)	(−0.023, 0.143)	(−0.007, 0.156)
	3my	1926-2009	3	0.952 (0.071)	0.923 (0.068)	0.171	−0.021 (0.671)	−0.025 (0.704)	(−0.098, 0.057)	(−0.103, 0.052)	(−0.099, 0.045)
	ys	1919-2009	2	0.580 (0.113)	0.836 (0.079)	−0.268	0.006 (0.466)	−0.049 (0.751)	(−0.117, 0.129)	(−0.169, 0.070)	(−0.164, 0.089)

Table 3: Limiting Probabilities of Taking Positive Values

	$\sigma_{ue} = -0.95$			$\sigma_{ue} = -0.55$		
c	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$	$\tilde{\beta}_{c,rols}$	$\tilde{\beta}_{c,rgls}$	$\hat{\beta}$
0	0.495	0.515	0.947	0.503	0.530	0.809
5	0.496	0.503	0.815	0.501	0.515	0.691
10	0.468	0.502	0.736	0.483	0.510	0.641
20	0.451	0.501	0.672	0.470	0.505	0.601
30	0.447	0.501	0.641	0.464	0.505	0.582
50	0.445	0.502	0.609	0.462	0.502	0.561



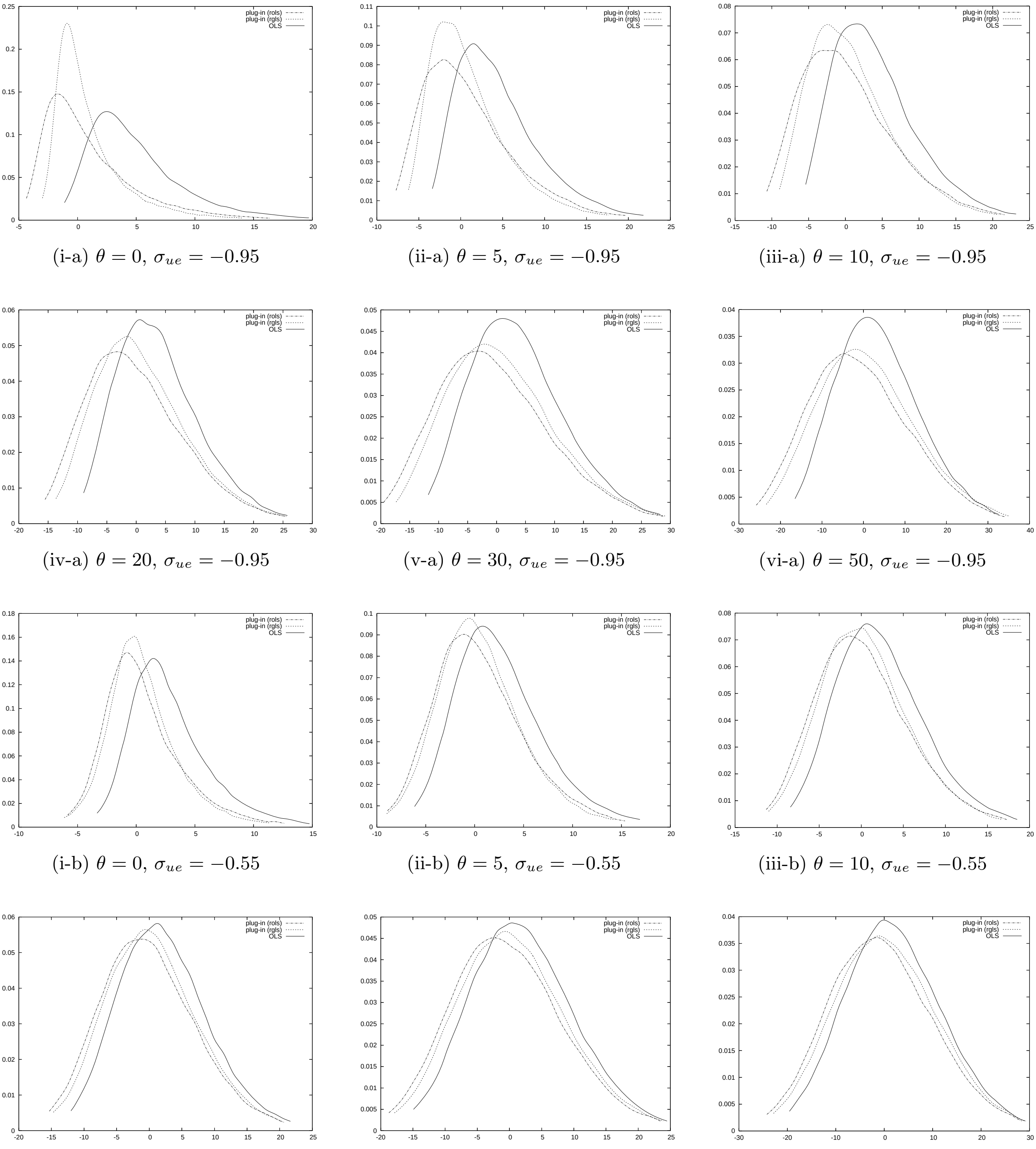
(a-i) $\rho = 0.9$, $\sigma_{ue} = -0.95$

(a-ii) $\rho = 0.9$, $\sigma_{ue} = -0.55$

(b-i) $\rho = 0.99$, $\sigma_{ue} = -0.95$

(b-ii) $\rho = 0.99$, $\sigma_{ue} = -0.55$

Figure 1: Finite Sample Power of the Tests ($T = 250$)



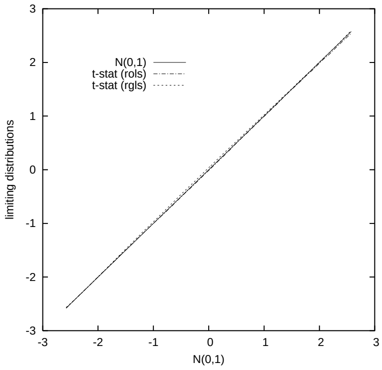
(i-a) $\theta = 0$, $\sigma_{ue} = -0.95$ (ii-a) $\theta = 5$, $\sigma_{ue} = -0.95$ (iii-a) $\theta = 10$, $\sigma_{ue} = -0.95$

(iv-a) $\theta = 20$, $\sigma_{ue} = -0.95$ (v-a) $\theta = 30$, $\sigma_{ue} = -0.95$ (vi-a) $\theta = 50$, $\sigma_{ue} = -0.95$

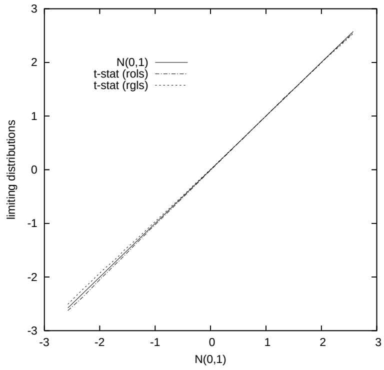
(i-b) $\theta = 0$, $\sigma_{ue} = -0.55$ (ii-b) $\theta = 5$, $\sigma_{ue} = -0.55$ (iii-b) $\theta = 10$, $\sigma_{ue} = -0.55$

(iv-b) $\theta = 20$, $\sigma_{ue} = -0.55$ (v-b) $\theta = 30$, $\sigma_{ue} = -0.55$ (vi-b) $\theta = 50$, $\sigma_{ue} = -0.55$

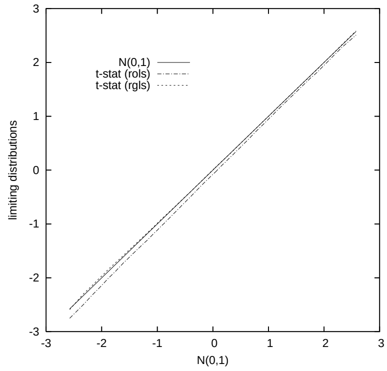
Figure 2: The limiting pdfs of the estimates



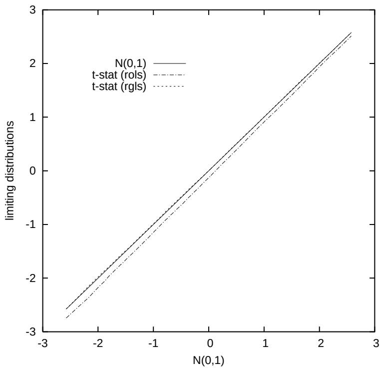
(i-a) $\theta = 0$, $\sigma_{ue} = -0.95$



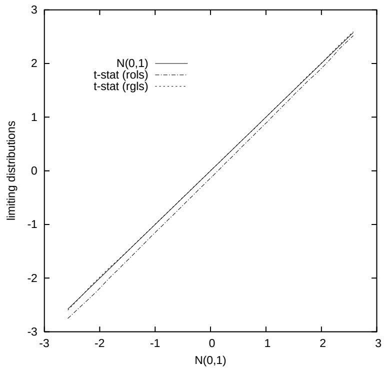
(ii-a) $\theta = 5$, $\sigma_{ue} = -0.95$



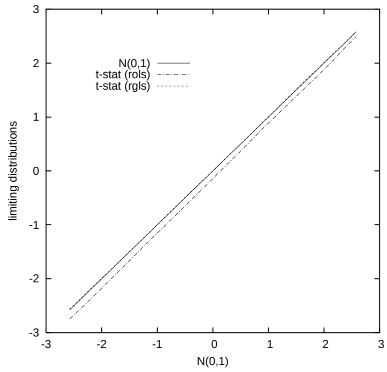
(iii-a) $\theta = 10$, $\sigma_{ue} = -0.95$



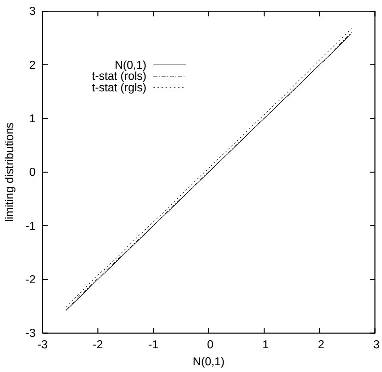
(iv-a) $\theta = 20$, $\sigma_{ue} = -0.95$



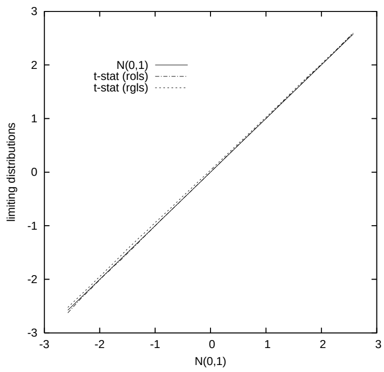
(v-a) $\theta = 30$, $\sigma_{ue} = -0.95$



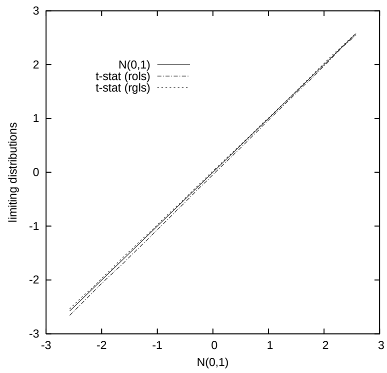
(vi-a) $\theta = 50$, $\sigma_{ue} = -0.95$



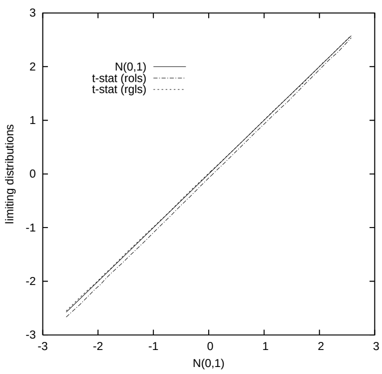
(i-b) $\theta = 0$, $\sigma_{ue} = -0.55$



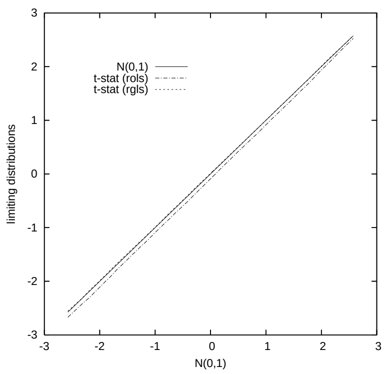
(ii-b) $\theta = 5$, $\sigma_{ue} = -0.55$



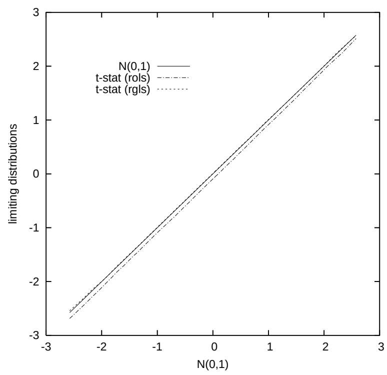
(iii-b) $\theta = 10$, $\sigma_{ue} = -0.55$



(iv-b) $\theta = 20$, $\sigma_{ue} = -0.55$



(v-b) $\theta = 30$, $\sigma_{ue} = -0.55$



(vi-b) $\theta = 50$, $\sigma_{ue} = -0.55$

Figure 3: The q-q plots