Journal of Econometrics 164 (2011) 79–91

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

# A control function approach for testing the usefulness of trending variables in forecast models and linear regression

Graham Elliott *

*Department of Economics, University of California, San Diego, 9500 Gilman Drive, LA JOLLA, CA, 92093-0508, USA*

ARTICLE INFO

*Article history:*
Available online 8 March 2011

*JEL classification:*
C12
C22
C53

*Keywords:*
Near unit root
Cointegration
Prediction regressions

ABSTRACT

Many predictors employed in forecasting macroeconomic and finance variables display a great deal of persistence. Tests for determining the usefulness of these predictors are typically oversized, overstating their importance. Similarly, hypothesis tests on cointegrating vectors will typically be oversized if there is not an exact unit root. This paper uses a control variable approach where adding stationary covariates with certain properties to the model can result in asymptotic normal inference for prediction regressions and cointegration vector estimates in the presence of possibly non-unit root trending covariates. The properties required for this result are derived and discussed.

Published by Elsevier B.V.

## 1. Introduction

A common problem in constructing forecasting regressions is understanding the sampling uncertainty of the coefficients in the prediction regression when the predictors appear to be persistent. Forecasters care about this sampling uncertainty because it is often unclear whether or not the predictors are useful for forecasting. Given the near unit root behavior of the predictors, standard asymptotic normal distributions have been shown to be very poor approximations to the sampling distributions of the coefficients in the prediction regression. Such approximations lead to oversized tests and are hence potentially misleading as guides as to whether or not the variables actually have any predictive usefulness.

The classic example in the forecasting literature is the prediction of stock market returns with the dividend price ratio—the dividend price ratio appears to have a root on or near the unit circle. For this regression the problem extends to other popular predictors such as interest rate differentials and the earnings price ratio as well. Similar issues arise with forecasting exchange rates with the forward premium, forecasting changes in income or consumption with interest rate differentials and ratios of macroeconomic variables.

A number of approaches have been taken to provide tests for inclusion of the predictive variables that control size even when the predictive variable has a unit root or near unit root. The common tool has been to use local to unity asymptotics — using limit theory for sequences of models where the largest root for the predictive regression remains in the neighborhood of one — to approximate distributions. The tests then differ in the precise statistic to be computed and how they handle the unknown root. One approach has been to use Bonferroni or related methods (Cavanagh et al., 1995; Lewellen, 2004; Campbell and Yogo, 2006). Alternatively Moreira and Jansson (2006) condition on a sufficient statistic for the root to remove the dependence in their test. None of these methods dominates each other theoretically or in practice. The precise nature of the problem and these methods are reviewed in Section 2 of this paper.

The methods we have so far have some limitations. First, for each of these methods extending the methods beyond the bivariate regression is extremely challenging, and no results exist in the literature for more than a single predictor in the regression. This is in part because the methods themselves are somewhat cumbersome to apply.

This paper suggests a different approach to obtaining tests which control size. In the method presented here additional covariates are added to the regression. The problems of size distortion and inference are shown to depend on a convolution of the parameters of the trending process and nuisance parameters that describe the relationships between the shocks to the forecasting regression and the shocks to the variables used for forecasting. Judiciously chosen, the covariates have the potential to remove the dependence of the hypothesis tests on the nuisance parameters describing the trend, and so provide inference that is robust to lack of knowledge over the trending behavior of the data. The method is discussed in Section 2 in a simplified bivariate case and examined for very general models in Section 3.

* Tel.: +858 534 4481; fax: +858 534 7040.
*E-mail address:* grelliott@ucsd.edu.

0304-4076/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.jeconom.2011.02.014

The method is applicable for a very wide set of problems. First, we allow for the inclusion of a general number of regressors in the prediction regression, hence we are not (as in the methods above) restricted to a single regressor. Second, for the prediction regression the method involves only the running of OLS regressions, and hence even for the general case is straightforward to apply. Third, inference is standard asymptotic (mixed) normal and so hypothesis tests are also straightforward to apply.

We also show, in Section 4, that the method can also be used when examining 'cointegrating' regressions where there is uncertainty over a unit root. Such uncertainty is ubiquitous in the cointegration literature, hence the use of unit root and rank pretests. With use of additional stationary covariates, we show similar results as for the prediction equation problem—asymptotically standard (mixed) normal inference can still apply even when the roots are not exactly on the unit circle.

The suggested method, as in the case in instrumental variables, requires finding additional data with particular properties. Such data are referred to as orthogonalizing covariates. In Section 5 we go into more detail regarding these properties. We suggest how such variables can be found in practice. We also show, in Theorem 5, that the use of such covariates that satisfy these properties have at minimum the same power (and generally better power) as would be obtained if we knew the size of the largest root of the predictor variable. Since all of the methods discussed in Section 2 cannot have better power than the case of the root known, this leads directly to the implication that none of the current methods can have better power against any alternative than the method suggested here when we indeed have orthogonalizing covariates.

Monte Carlo results are presented in Section 6 to examine both size and power properties of the suggested method. We also examine the implications of Theorem 5 for the relationship between these methods and the popular Campbell and Yogo (2006) procedure in terms of power.

Proofs of the results are contained in an Appendix.

## 2. The testing problem and conditioning on stationary covariates

This paper examines two models which have been treated rather differently in the literature but for which common methods apply. The first regression is where a relatively non-trending variable is regressed on lagged trending variables for purposes of prediction. The second is a cointegrating regression where we are uncertain that there are unit roots in the data. In this case the 'cointegrating' vector is a trending variable being predicted by contemporaneous trending variables. These problems are similar in the sense that it is not the behavior of the left-hand side variable (which differs here in properties) but the terms on the right-hand side that dominate the theoretical properties of the estimators and tests. In this sense we are able to generate a solution to both problems from a similar idea, as detailed in following sections. The differences lie mostly in the auxiliary assumptions of the model, and the perceived reasonable coefficients on the trending covariates. The remaining difference is that in terms of constructing methods that deal with uncertainty over the trending process, many papers have appeared that examine the prediction regression whereas only a few have examined the cointegrating regression. We focus on the predictive regression in this review section.

The simplest predictive regression test would be to examine if a variable measured today predicts the outcome we are interested in seeing in the next period. For example there is a large literature on forecasting stock returns using the log of the dividend price ratio in the previous period as a predictor. For inference on predictive ability we require assumptions on both the properties of the predictor variable and on the errors to the regression to justify any inferential procedure on the predictiveness or lack thereof of the regressor. Consider the model

$$(1-\rho L)(y_{1t}-\varphi_1)=\varepsilon_{1t} \quad t=2,\ldots,T$$
$$y_{2t}=\varphi_2+\beta y_{1t-1}+\varepsilon_{2t} \quad t=2,\ldots,T \tag{1}$$

where we observe $\{y_t\}_{t=1}^T=\{[y_{1t},y_{2t}]'\}_{t=1}^T$ where both $y_{1t}$ and $y_{2t}$ are univariate, $\varphi_1,\varphi_2,\beta$ and $\rho$ are unknown parameters, and $\varepsilon_{1t}$ and $\varepsilon_{2t}$ are unknown residual terms. Define the parameter space for $\rho$ as $P$. Define $E[\varepsilon_{it}\varepsilon_{jt}']=\Sigma_{ij}$ for $i,j=1,2$. The correlation between $\varepsilon_{1t}$ and $\varepsilon_{2t}$ is $\delta=\Sigma_{12}/(\Sigma_{11}\Sigma_{22})^{1/2}$. In addition, we must initialize the processes, set the initial value of $\varepsilon_{1t}=\xi$ such that $T^{-1/2}\xi\to^p 0$. Some of these restrictions are made for expository purposes in this review—in the next section $y_{2t}$ can be multivariate, the specification of the deterministic terms is more general, and we also allow serial correlation in $(1-\rho L)y_{1t}$.

This regression is typically run to test the notion of unpredictability of $y_{2t}$ – the typical null hypothesis derived from economic theory – so we assume that $\varepsilon_{2t}$ is uncorrelated. Regressions of this form appear as noted above in the prediction of stock prices by such covariates as the dividend price ratio, earnings price ratio and interest rates. Fama and French (1988) for example reject the hypothesis that the dividend price ratio cannot predict returns. Similar issues arise in testing term structure models of the interest rate, where the lagged interest rate spread is used to predict interest rates. Fama (1975) uses interest rates to predict inflation. Hall (1978) examines the predictability of changes in consumption with lagged income and stock prices. Large literatures reexamining these findings has followed these seminal papers.

In each of these cases, the regressor is typically a persistent or trending variable in the sense that there is clearly a low frequency component to the data. Unit root or near unit root behavior affects the sampling distribution of least squares estimators for $\beta$ and also hypothesis tests based on this estimate. We can estimate $\beta$ by OLS, yielding

$$\hat\beta-\beta=\frac{\sum_{t=2}^T y_{1t-1}^m\varepsilon_{2t}}{\sum_{t=2}^T (y_{1t-1}^m)^2}$$

where $y_t^m=y_t-(T-1)^{-1}\sum_{t=2}^T y_t$.

Projecting $\varepsilon_{2t}$ onto $\varepsilon_{1t}$ yields $\varepsilon_{2t}=(\Sigma_{12}/\Sigma_{11})\varepsilon_{1t}+\varepsilon_{2.1t}$ where $E[\varepsilon_{1t}\varepsilon_{2.1t}]=0$ by construction and the variance of $\varepsilon_{2.1t}$ is $\Sigma_{2.1}=\Sigma_{22}-\Sigma_{12}^2/\Sigma_{11}$. Hence the above expression can be rewritten as

$$\hat\beta-\beta=\frac{\sum_{t=2}^T y_{1t-1}^m\varepsilon_{1t}(\Sigma_{12}/\Sigma_{11})}{\sum_{t=2}^T (y_{1t-1}^m)^2}+\frac{\sum_{t=2}^T y_{1t-1}^m\varepsilon_{2.1t}}{\sum_{t=2}^T (y_{1t-1}^m)^2}.$$

$$=(\hat\rho-\rho)(\Sigma_{12}/\Sigma_{11})+\frac{\sum_{t=2}^T y_{1t-1}^m\varepsilon_{2.1t}}{\sum_{t=2}^T (y_{1t-1}^m)^2} \tag{2}$$

where $\hat\rho$ is the least squares estimator for $\rho$ using data on $y_{1t}$.

A number of authors have pointed out issues with inference over $\beta$ in regressions that take this form. Examining the predictability of changes in consumption using lagged income as a predictor, Mankiw and Shapiro (1986) showed in Monte Carlo experiments that hypothesis tests on $\beta$ did not control size when standard critical values were employed. Stambaugh (1999) showed a similar result for prediction of returns with the dividend

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات