On the Predictability of Global Stock Returns[1]

Erik Hjalmarsson
Department of Economics, Yale University
P.O. Box 208281, New Haven, CT 06520-8281

December 16, 2004

[1] This paper has greatly benefitted from comments by my advisors, Peter Phillips and Robert Shiller. I am also grateful for advice from Don Andrews, Ray Fair, Yuichi Kitamura, Taisuke Otsu, Randi Pintoff and participants in the econometrics seminar and workshop at Yale University.

Abstract

Stock return predictability is a central issue in empirical finance. Yet no comprehensive study of international data has been performed to test the predictive ability of lagged explanatory variables. In fact, most stylized facts are based on U.S. stock-market data. In this paper, I test for stock return predictability in the largest and most comprehensive data set analyzed so far, using four common forecasting variables: the dividend- and earnings-price ratios, the short interest rate, and the term spread. The data contain over 20,000 monthly observations from 40 international markets, including markets in 22 of the 24 OECD countries.

I also develop new asymptotic results for long-run regressions with overlapping observations. I show that rather than using auto-correlation robust standard errors, the standard t-statistic can simply be divided by the square root of the forecasting horizon to correct for the effects of the overlap in the data. Further, when the regressors are persistent and endogenous, the long-run OLS estimator suffers from the same problems as does the short-run OLS estimator, and similar corrections and test procedures as those proposed by Campbell and Yogo (2003) for the short-run case should also be used in the long-run; again, the resulting test statistics should be scaled due to the overlap.

The empirical analysis conducts time-series regressions for individual countries as well as pooled regressions. The results indicate that the short interest rate and the term spread are fairly robust predictors of stock returns in OECD countries. The predictive abilities of both the short rate and the term spread are short-run phenomena; in particular, there is only evidence of predictability at one and 12-month horizons. In contrast to the interest rate variables, no strong or consistent evidence of predictability is found when considering the earnings- and dividend-price ratios as predictors. Any evidence that is found is primarily seen at the long-run horizon of 60 months. Neither of these predictors yields any consistent predictive power for the OECD countries.

The interest rate variables also have out-of-sample predictive power that is economically significant; the welfare gains to a log-utility investor who uses the predictive ability of these variables to make portfolio decisions are substantial.

JEL classification: C22, C23, G12, G15

Keywords: Predictive regressions, long-horizon regressions, panel data, stock return predictability.

1 Introduction

Our empirical knowledge regarding the predictability of stock returns by valuation ratios or interest rate variables has been subject to constant updating over time. This updating has been mainly driven by the development of a number of new econometric methods that enable us to more accurately assess the evidence of stock return predictability.[1] Despite these methodological advances, little consensus regarding stock return predictability has emerged.

The typical forecasting regression for stock returns, or excess stock returns, is plagued by some difficult econometric problems due to the near persistence and endogeneity of the forecasting variables. This is especially true when the dividend- and earnings-price ratios are used as regressors. Early work such as Fama and French (1988a, 1989) and Campbell and Shiller (1988), which mostly ignored these issues, concluded that there is generally strong evidence of stock return predictability. Using the most efficient and robust methods to date, Campbell and Yogo (2003) and Lewellen (2003) still find evidence of predictability; however, their results are much less conclusive than the earlier studies. In particular, the predictive ability of the dividend- and earnings-price ratios appear sensitive to the sample period and the choice of frequency (annual to monthly).

Despite these substantial methodological advances, there have been surprisingly few attempts at furthering our understanding of stock return predictability using data other than that of the U.S. stock-market. Consequently, many take as given stylized facts regarding stock return predictability that are based on this limited source of data. Since the predictable component of stock returns must be small, if indeed one does exist, there seems to be little chance of reaching a decisive conclusion using U.S. data alone, which effectively provides only one time-series at the market level.

There has been some analysis of predictability in international stock returns, but many of the results are based on relatively small data sets and most studies rely on non-robust econometric methods. In addition, most international results are based on individual time-series regressions and very little analysis has been conducted using pooled panel data regressions; as is discussed below, pooling the data is useful both from a strictly econometric viewpoint and in terms of drawing overall conclusions. Harvey (1991,1995) and Ferson and Harvey (1993) consider various aspects of predictability in international stock returns. The data sets used in these studies cover a fair amount of different countries, but the overall time span is limited to the period 1969-1992 and most results are thus based on regressions using less than 20 years of data; no robust econometric methods are used and no pooling of the data is attempted. Ang and Bekaert (2003) analyze predictability in stock returns from four different countries, in addition to the U.S, and their international sample only dates back to 1975. In a survey article, Campbell (2003) briefly considers the evidence of stock return predictability using an international data set of 11 countries with observations going back to the 1970s. Polk et al. (2004) use an international sample consisting of 22 countries with observations dating back to 1975, but they only analyze the predictive ability of their cross-sectional beta-premiums and do not consider more

[1] Some early references are Mankiw and Shapiro (1986), Nelson and Kim (1993), and Goetzman and Jorion (1993). Recent work include Cavanagh et al. (1995), Stambaugh (1999), Lanne (2002), Lewellen (2003), Campbell and Yogo (2003), Janson and Moreira (2004), and Polk et al. (2004). Ferson et al. (2003) discuss spurious regressions and data mining.

traditional forecasting variables. However, I am not aware of any serious attempts to test the predictive ability of common forecasting variables, like the earnings-price ratio and short interest rate, using up-to-date econometric methods in a large data set of international stock returns.

The aim of this paper is twofold. By considering a large global data set, I provide the most comprehensive picture of stock return predictability to date. The data contain over 20,000 monthly observations from 40 countries, including markets in 22 of the 24 OECD countries.[2] The longest data series is for the U.K. stock-market and dates back to 1836 while data for eight other markets date back to before 1935. Secondly, I develop and apply a number of new results for short- and long-run forecasting regressions, with an extra emphasis on methods utilizing the panel structure of the data.

Since an international data set of stock returns and forecasting variables provides a panel, some alternative methods to those usually employed in the standard time-series case can be considered. As shown by Hjalmarsson (2004), pooling the data provides for a convenient solution to the traditional inferential problems encountered in regressions with highly endogenous regressors of unknown persistence. In particular, I apply panel data methods that do not require the use of inefficient bound procedures (Cavanagh et al., 1995). Pooling also enables one to draw more general conclusions regarding predictability in the face of contradicting evidence. For instance, when using individual time-series regressions, the null of no predictability can often be rejected in some countries but not in others. As argued by Hjalmarsson (2004), a rejection of the null hypothesis of a zero slope coefficient in a pooled predictive regression indicates that there is *on average* a predictive relationship in the data. In addition, tests based on pooled estimates can have more power than individual time-series tests (Hjalmarsson, 2004).

The main theoretical contribution of this paper is the development of new asymptotic results for long-run regressions with overlapping observations. Typically, auto-correlation robust estimation of the standard errors (e.g. Newey and West, 1987) is used to perform inference in long-run regressions. However, these robust estimators tend to perform poorly in finite samples since the serial correlation induced in the error terms by overlapping data is often very strong.[3] In a time-series setting, I show that rather than using robust standard errors, the standard $t-$statistic can simply be divided by the square root of the forecasting horizon to correct for the effects of the overlap in the data.[4] Further, when the regressors are persistent and endogenous, the long-run OLS estimator suffers from the same problems as does the short-run OLS estimator, and similar corrections and test procedures as those proposed by Campbell and Yogo (2003) for the short-run case should also be used in the long-run; again, the resulting test statistics should be scaled due to the overlap. Thus, these results lead to simple and more efficient inference in long-run regressions by obviating the need for robust standard error estimation methods and controlling for the endogeneity and persistence of the regressors. These

[2] Included in the sample are the stock-markets in Hong Kong and Taiwan. Since Hong Kong is part of China and Taiwan is not a formally recognized sovereign state, the use of the term country for these markets is not entirely correct, but is used for convenience throughout the paper.

[3] Ang and Bekaert (2003) advocate the use of Hodrick (1992) auto-correlation robust standard errors. However, these rely on the regressors being covariance stationary, which is usually a restrictive assumption for forecasting variables like the short interest rate or the dividend-price ratio that are typically modeled as being nearly persistent processes.

[4] This result is similar to one by Hansen and Tuypens (2004), who consider the covariance stationary case.

long-run results are also extended to the panel data case.

The asymptotic distributions of the long-run estimators are derived not only under the null-hypothesis of no predictability, but also under an alternative of predictability. This gives a more complete characterization of the asymptotic properties of the long-run estimators than is typically found in the literature, where results for long-run estimators are often derived only under the null-hypothesis of no predictability. It is shown that, under the standard econometric model of stock return predictability, the long-run estimators converge to well defined quantities, but their asymptotic distributions are non-standard and fundamentally different from the asymptotic distributions under the null hypothesis of no predictability. The rates of convergence of the long-run estimators are also slower under the alternative hypothesis of predictability than under the null hypothesis, and slower than that of the short-run estimator. These results suggest that under the standard econometric specifications that are typically postulated, short-run inference is preferable to long-run inference. As discussed briefly, there may be cases where long-run estimation has some advantages; however, these scenarios may not necessarily be easily captured by formal econometric models.[5]

In the empirical analysis, I conduct time-series regressions for individual countries as well as pooled regressions. In both types of analyses, I estimate short- and long-run regressions for four of the most commonly used forecasting variables: the dividend- and earnings-price ratios, the short interest rate, and the term spread. In the pooled regressions, countries are either all grouped together in a global panel or split up into groups of OECD and Non-OECD countries. The short-run time-series analysis uses methods similar to those of Campbell and Yogo (2003) while the long-run and pooled portions of the analysis use the methods described above and in Hjalmarsson (2004), respectively. All results are based on predictive regressions for excess stock returns, although for convenience I will typically just write stock returns.

The results indicate that the short interest rate and the term spread are both fairly robust predictors of stock returns in OECD countries. The null of no predictability is clearly rejected in the OECD pooled regressions as well as in a number of time-series regressions for OECD countries. The predictive abilities of both the short rate and the term spread are short-run phenomena; in particular, there is only evidence of predictability at one and 12-month horizons. These results are generally in line with those found by Campbell and Yogo (2003) with U.S. data and with the limited international results of Ang and Bekaert (2003). These results are also quite consistent over time, as evidenced by rolling regressions.

In contrast to the interest rate variables, no strong or consistent evidence of predictability is found when considering the earnings- and dividend-price ratios as predictors. Any evidence that is found is primarily seen at the long-run horizon of 60 months. Specifically, there is rather weak evidence that the earnings-price ratio predicts stock returns and the majority of evidence that does exist is for Non-OECD countries. The results for the dividend-price ratio are, in general, parallel to those of

[5] Mark and Sul (2004) analyze *local* alternatives to the null hypothesis of no predictability and find that there are cases in which a long-run specification has more power to detect deviations from the null hypothesis, than do the short-run specifications. No formal analysis of power properties is performed in this paper but simulation results indicate that, in standard models, the short-run tests dominate the long-run ones.

the earnings-price ratio, although they do contain a greater number of significant time series results. Neither predictor yields any consistent predictive power for the OECD countries; as seen in rolling regressions, this is particularly true for the dividend-price ratio.

In response to the Goyal and Welch (2003b, 2004) critique, on the poor out-of-sample performance of forecasting variables for stock returns, I also consider out-of-sample forecasts of stock returns using the short interest rate or term spread as predictors. For all countries where there is a significant in-sample predictive relationship, it is found that the forecasts based on either of these predictor variables beat the benchmark forecasts based on the average of past stock returns. The results are strongest for the short interest rate, which overall appears as the most robust predictor in international data. Moreover, the out-of-sample predictive power is economically significant, resulting in substantial welfare gains to a log-utility investor who uses the predictive ability of the short rate to make portfolio decisions; in most cases, the welfare gain for the investor is at least equivalent to that enjoyed from a one to two percentage point increase in the annual real risk-free interest rate.

The rest of the paper is organized as follows. Section 2 states the econometric model and main assumptions, Section 3 outlines the short-run inference methods, and Section 4 derives the new long-run estimation results. In Section 5, the practical implementation of some of the procedures is discussed. The data is described in Section 6, the main empirical results are provided in Section 7, out-of-sample performance and economic implications are discussed in Section 8, and Section 9 concludes. Technical proofs are found in the appendix.

2 Model and assumptions

Let the excess returns for stocks in country i, $i = 1, ..., n$, be denoted $r_{i,t}$, and the corresponding vector of regressors, $x_{i,t}$, where $x_{i,t}$ is an $m \times 1$ vector and $t = 1, ..., T$. The empirical section of this paper deals almost exclusively with the case of scalar regressors where $m = 1$; but since the econometric results developed in this paper are of general applicability, the model is formulated in more general terms. The behavior of $r_{i,t}$ and $x_{i,t}$ are assumed to satisfy,

$$r_{i,t} = \alpha_i + \beta_i x_{i,t-1} + u_{i,t}, \tag{1}$$

$$x_{i,t} = A_i x_{i,t-1} + v_{i,t}, \tag{2}$$

where $A_i = I + C_i/T$ is an $m \times m$ matrix, with diagonal elements $1 + c_{k,i}/T$, and off-diagonal elements $c_{kl,i}/T$, $k, l = 1, ..., m$, $k \neq l$. The error processes are assumed to satisfy the following conditions.

Assumption 1 *Let* $w_{i,t} = (u_{i,t}, \epsilon_{i,t})'$ *and* $\mathcal{F}_t = \{ w_{i,s} | \, s \leq t, i = 1, ..., n \}$ *be the filtration generated by* w_{it}, $i = 1, ..., n$. *Then, for all* $i = 1, ..., n$,

1. $v_{i,t} = D_i(L)\,\epsilon_{i,t} = \sum_{j=0}^{\infty} D_{i,j}\epsilon_{i,t-j}$, $\bar{D}_j \equiv \sup_i ||D_{i,j}|| < \infty$, *and* $\sum_{j=0}^{\infty} j^3 \left|\left|\bar{D}_j\right|\right| < \infty$.
2. $E[w_{it} | \, \mathcal{F}_{t-1}] = 0$, $\sup_t E\left[u_{i,t}^4\right] < \infty$ *and* $\sup_t E\left[||\epsilon_{i,t}||^4\right] < \infty$.
3. $E\left[w_{i,t}w_{i,t}' \middle| \, \mathcal{F}_{t-1}\right] = \Sigma_i = [(\sigma_{11i}, \sigma_{12i}), (\sigma_{21i}, I)]$.
4. $E\left[w_{i,t}w_{j,s}'\right] = 0$ *for all* t, s *and* $i \neq j$.

The model described by equations (1) and (2) and Assumption 1 captures the essential features of a predictive regression with nearly persistent regressors. It states the usual martingale difference (mds) assumption for the errors in the return processes but allows for a linear time-series structure in the errors of the predictor variables. The error terms $u_{i,t}$ and $v_{i,t}$ are also often highly correlated. The auto-regressive roots of the regressors are parametrized as being local-to-unity, which captures the near-unit-root behavior of many predictor variables, but is less restrictive than a pure unit-root assumption. In the cross-section, the innovation processes are assumed to be independent. This is clearly a restrictive assumption and methods for relaxing it is detailed in the section on pooled estimation below.

Similar models, for the time-series properties of the data, are used to analyze the predictability of stock returns by Cavanagh et al. (1995), Torous et al. (2005), Lanne (2002), Campbell and Yogo (2003), and Valkanov (2003).

Let $E_{i,t} = (u_{i,t}, v_{i,t})'$ be the joint innovations process. Under Assumption 1, by standard arguments (Phillips and Solo, 1992), for any i,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} E_{i,t} \Rightarrow B_i(r) = BM(\Omega_i)(r),$$

where $\Omega_i = [(\sigma_{11i}, \omega_{12i}), (\omega_{21i}, \Omega_{22i})]$, $\omega_{21i} = D_i(1)\sigma_{12i}$, $\omega_{12i} = \omega_{21i}'$, $\Omega_{22i} = D_i(1) D_i(1)'$, and $B_i(\cdot) = (B_{i,1}(\cdot), B_{i,2}(\cdot))'$ denotes an $1+m-$dimensional Brownian motion. Also, let $\Lambda_{22i} = \sum_{k=1}^{\infty} E\left(v_{i,k} v_{i,0}'\right)$ be the one-sided long-run variance of $v_{i,t}$. The following lemma sums up the key asymptotic results for the nearly integrated model in this paper (Phillips 1987,1988).

Lemma 1 *Under Assumption 1, as $T \to \infty$,*

(a) $T^{-1/2} x_{i,[Tr]} \Rightarrow J_{i,C_i}(r)$,

(b) $T^{-3/2} \sum_{t=1}^{T} x_{i,t} \Rightarrow \int_0^1 J_{i,C_i}(r)\, dr$,

(c) $T^{-2} \sum_{t=1}^{T} x_{i,t} x_{i,t}' \Rightarrow \int_0^1 J_{i,C_i}(r) J_{i,C_i}(r)'\, dr$,

(d) $T^{-1} \sum_{t=1}^{T} u_{i,t} x_{t-1}' \Rightarrow \int_0^1 dB_{i,1}(r) J_{i,C_i}(r)'$,

(e) $T^{-1} \sum_{t=1}^{T} v_{i,t} x_{t-1}' \Rightarrow \int_0^1 dB_{i,2}(r) J_{i,C_i}(r)' + \Lambda_{22i}$,

where $J_{i,C_i}(r) = \int_0^r e^{(r-s)C_i} dB_{i,2}(s)$.

Analogous results hold for the demeaned variables $\underline{x}_{i,t} = x_{i,t} - T^{-1} \sum_{t=1}^{n} x_{i,t}$, with the limiting process J_{i,C_i} replaced by $\underline{J}_{i,C_i} = J_{i,C_i} - \int_0^1 J_{i,C_i}$. These results are used repeatedly below.

The greatest problem in dealing with regressors that are near-unit-root processes is the nuisance parameter C_i; it is generally unknown and not consistently estimable. It is nevertheless useful to first derive inferential methods under the assumption that C_i is known, and then use the methods of Cavanagh et al. (1995) to construct feasible tests. The following two sections derive and outline the inferential methods used for estimating and performing tests on β_i in equation (1), treating C_i as known. I consider both time-series methods, where individual $\beta_i s$ for each country are estimated, and panel data methods where the data are pooled across countries and a common estimate, β, for all i

is obtained. Section 5 discusses how the methods of Cavanagh et al. (1995) can be used to construct feasible tests with C_i unknown.

In line with much of the previous literature on stock return predictability, I consider both short- and long-run regressions. The main econometric contributions of the paper are on long-run regressions in both the time-series and panel data case. The section on short-run time-series inference is mainly a review section, but sets the stage for the long-run results. The section on short-run panel inference summarizes some of the results in Hjalmarsson (2004), which have not been used before.

3 Short-run inference

3.1 The time-series case

Let $\hat{\beta}_i$ denote the standard OLS estimate of β_i in equation (1). By Lemma 1 and the continuous mapping theorem (CMT), it follows that

$$T\left(\hat{\beta}_i - \beta\right) \Rightarrow \left(\int_0^1 dB_{i,1}\underline{J}'_{i,C_i}\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}'_{i,C_i}\right)^{-1}, \tag{3}$$

as $T \to \infty$. Analogously to the case with pure unit-root regressors, the OLS estimator does not have an asymptotically mixed normal distribution due to the correlation between $B_{i,1}$ and $B_{i,2}$, which causes $B_{i,1}$ and J_{i,C_i} to be correlated. Therefore, standard test procedures cannot be used.

In the pure unit-root case, one popular inferential approach is to "fully modify" the OLS estimator as suggested by Phillips and Hansen (1990) and Phillips (1995). In the near-unit-root case, a similar method can be considered. Define the quasi-differencing operator

$$\Delta_{C_i} x_{i,t} = x_{i,t} - x_{i,t-1} - \frac{C_i}{T} x_{i,t-1} = v_{i,t}, \tag{4}$$

and let $r^+_{i,t} = r_{i,t} - \hat{\omega}_{12i}\hat{\Omega}^{-1}_{22i}\Delta_{C_i}x_{i,t}$ and $\hat{\Lambda}^+_{12i} = -\hat{\omega}_{12i}\hat{\Omega}^{-1}_{22i}\hat{\Lambda}_{22i}$, where $\hat{\omega}_{12i}$, $\hat{\Omega}^{-1}_{22i}$, and $\hat{\Lambda}_{22i}$ are consistent estimates of the respective parameters.[6] The fully modified OLS estimator is now given by

$$\hat{\beta}^+_i = \left(\sum_{t=1}^T \underline{r}^+_{i,t}\underline{x}'_{i,t-1} - T\hat{\Lambda}^+_{12i}\right)\left(\sum_{t=1}^T \underline{x}_{i,t-1}\underline{x}'_{i,t-1}\right)^{-1}, \tag{5}$$

where $\underline{r}^+_{i,t} = \underline{r}_{i,t} - \hat{\omega}_{12i}\hat{\Omega}^{-1}_{22i}\Delta_{C_i}x_{i,t}$ and $\underline{r}_{i,t} = r_{i,t} - T^{-1}\sum_{t=1}^{t} r_{i,t}$. The only difference in the definition of (5), to the FM-OLS estimator for the pure unit-root case, is the use of the quasi-differencing operator, as opposed to the standard differencing operator. Once the innovations $v_{i,t}$ are obtained from quasi-differencing, the modification proceeds in exactly the same manner as in the unit-root case.

Define $\sigma_{11\cdot 2,i} = \sigma_{11i} - \omega_{12i}\Omega^{-1}_{22i}\omega_{21i}$ and the Brownian motion $B_{i,1\cdot 2} = B_{i,1} - \omega_{12i}\Omega^{-1}_{22i}B_{i,2} = BM\left(\sigma_{11\cdot 2,i}\right)$. The process $B_{i,1\cdot 2}$ is now orthogonal to $B_{i,2}$ and J_{i,C_i}. Using the same arguments as

[6] The definition of $\hat{\Lambda}^+_{12i}$ is slightly different from the one found in Phillips (1995). This is due to the predictive nature of the regression equation (1), and the martingale difference sequence assumption on $u_{i,t}$.

Phillips (1995), it follows that

$$T\left(\hat{\beta}_i^+ - \beta_i\right) \Rightarrow \left(\int_0^1 dB_{i,1\cdot 2}\underline{J}_{i,C_i}{}'\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right)^{-1} \equiv MN\left(0, \sigma_{11\cdot 2,i}\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right)^{-1}\right). \quad (6)$$

The corresponding $t-$statistics and Wald statistics will now have standard distributions asymptotically. For instance, the $t-$test of the null hypothesis $\beta_{i,k} = \beta_{i,k}^0$ satisfies

$$t_i^+ = \frac{\hat{\beta}_{i,k}^+ - \beta_{i,k}^0}{\sqrt{\hat{\sigma}_{11\cdot 2,i} a' \left(\sum_{t=1}^T \underline{x}_{t-1}\underline{x}_{t-1}'\right)^{-1} a}} \Rightarrow N(0,1) \quad (7)$$

under the null, as $T \to \infty$. Here a is an $m \times 1$ vector with the k'th component equal to one and zero elsewhere.

The t_i^+-statistic is identical to the unfeasible $Q-$statistic of Campbell and Yogo (2003). Whereas Campbell and Yogo (2003) attack the problem from a test point-of-view, the derivation in this paper starts with the estimation problem and delivers the test-statistic as an immediate consequence. However, presenting the derivation in this manner makes clear that this approach is a generalization of fully modified estimation.

In the empirical analysis of the paper, the $\hat{\beta}_i^+$ estimator is implemented under the assumption that $b_i(L) v_{i,t} = u_{2,i,t}$, where $b_i(L) = \sum_{h=0}^p b_{i,h} L^h$, and $b_{i,0} = I_m$. That is, the innovations to the regressors are assumed to follow an $AR(p)$ process, rather than the general linear process specified in Assumption 1. Although this imposes slightly stronger conditions on the error terms, it allows for the parametric estimation of ω_{12i}, Ω_{22i}, and Λ_{22i}, and avoids non-parametric estimation, which might perform poorly in some of the shorter time-series in the sample. The lag length p is determined by the BIC model selection approach, and ω_{12i}, Ω_{22i}, and Λ_{22i} are estimated as in Campbell and Yogo (2003).

The inferential approach described above is based on asymptotic arguments for local-to-unity processes; effectively this leads to an extension of well established procedures for unit root data. Janson and Moreira (2004) and Polk et al. (2004) consider a fundamentally different approach to testing in models with endogenous and nearly persistent regressors, relying, in part, on finite sample arguments for Gaussian models. The discussion in Polk et al. (2004) reveals that in terms of power properties of the tests, neither approach dominates the other, however.

3.2 Pooled estimation

As an alternative to analyzing each time-series regression individually, data from several countries can be pooled together. In the pooled regressions considered in this paper, a common intercept β is estimated, but the individual intercepts α_i are allowed to vary across countries. Since a common β for all i is *estimated,* most panel data studies also assume that the *true* $\beta_i s$ are in fact identical. Rather than making this often unrealistic assumption, it is useful to start with an assumption of heterogenous

$\beta_i s$, and consider pooled estimation and testing under such conditions. That is, suppose $\beta_i = \beta + \theta_i$, where $\{\theta_i\}_{i=1}^n$ are *iid* random variables with mean zero and variance $\Omega_{\theta\theta}$, and $\{\theta_i\}_{i=1}^n$ is independent of $w_{i,t}$. The parameter β is now the average slope coefficient in the panel. As shown by Hjalmarsson (2004), in terms of practical inference, the assumption of heterogenous $\beta_i s$ does not change anything. The same pooled estimator and corresponding $t-$test can be used both in the homogenous case where $\beta_i = \beta$ for all i, and in the heterogenous case. The interpretations are different, however. For heterogenous $\beta_i s$, the pooled estimator converges to the average parameter β and the estimated value should thus be interpreted as an average relationship. When performing tests, this becomes even more important. When the $\beta_i s$ are heterogenous, the hypothesis of the typical pooled $t-$test is $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$. That is, the pooled $t-$test evaluates whether the average parameter β is different from zero; it has no power against the alternative that some β_i are different from zero, as long as the average value $\beta = 0$.

Thus, rejecting the null hypothesis of $\beta = 0$ does not reveal whether the variable $x_{i,t-1}$ predicts $r_{i,t}$ for a specific i, but it does say that on average there is a predictive relationship in the panel. The interpretation of β as an average relationship in the panel has the advantage of resolving evidence from individual time-series regressions. It is often the case that stock returns are found predictable in some countries, but not in others. The interpretations of such results are not straightforward but the results from a pooled regression provide an answer; if the average slope coefficient β is significant, then *on average* there is a significant relationship in the panel. Of course, the most complete picture of the empirical evidence is given by careful consideration of both the panel data and the time-series evidence.

The actual estimators and tests are identical regardless of whether one assumes that the $\beta_i s$ are homogenous or not. The convergence rates of the pooled estimator do differ, however, and for brevity, the results for just the homogenous case are given here. As shown by Hjalmarsson (2004), in the case with no individual effects in equation (1), such that $\alpha_i = \alpha$ for all i, the pooled estimator is asymptotically normally distributed as $(n, T \to \infty)$. The problems arising from the endogeneity of the regressors in the time-series case are thus no longer an issue in a plain pooled estimation without fixed effects. Intuitively, when summing up over a large cross-section, the endogeneity effects are diluted by independent cross-sectional information, and disappear asymptotically as the cross section grows large. However, when allowing for individual intercepts in each equation in the panel, the resulting fixed effects pooled estimator does suffer from a second order bias, due to the endogeneity and persistence of the regressors. That is, the asymptotic distribution is not centered around zero and standard tests will be invalid. In fact, it follows directly from the results of Hjalmarsson (2004) that in the most common forecasting regressions, with dividend- or earnings-price ratios as predictors and the covariances ω_{12i} negative, the pooled estimator has an upward bias. In this case, tests of predictability using standard pooled estimation with individual effects will tend to over-reject the null of no predictability.

One way of dealing with the problem caused by the inclusion of individual effects would be to consider fully modified methods similar to those in the time-series case. However, in the panel case, it is possible to get around the endogeneity problems without resorting to procedures that make use of the persistence parameters C_i. When demeaning each time-series in the panel, which is effectively what is done when fitting individual intercepts, information after time t is used to form the demeaned regressor $\underline{x}_{i,t}$; this induces a correlation between $\underline{x}_{i,t-1}$ and $u_{i,t}$, which gives rise to the second order bias in the fixed effects estimator. This effect can be avoided by using recursively demeaned data. Let

$$\hat{\beta}_{n,T}^{rd} = \left(\sum_{i=1}^{n}\sum_{t=1}^{T} \underline{r}_{i,t}^{dd} \underline{x}_{i,t-1}^{d\prime}\right)\left(\sum_{i=1}^{n}\sum_{t=1}^{T} \underline{x}_{i,t-1}^{dd} \underline{x}_{i,t-1}^{d\prime}\right)^{-1}, \tag{8}$$

where, $\underline{x}_{i,t}^{d} = x_{i,t} - \frac{1}{t}\sum_{s=1}^{t} x_{i,s}$, and $\underline{x}_{i,t}^{dd} = x_{i,t} - \frac{1}{T-t}\sum_{s=t}^{T} x_{i,s}$, and $\underline{r}_{i,t}^{dd} = r_{i,t} - \frac{1}{T-t}\sum_{s=t}^{T} r_{i,s}$. It now follows, under the null, as $(n, T \to \infty)$,

$$\sqrt{n}T\left(\hat{\beta}_{n,T}^{rd} - \beta\right) \Rightarrow N\left(0, \left(\underline{\Omega}_{xx}^{rd\prime}\right)^{-1} \underline{\Phi}_{ux}^{rd} \left(\underline{\Omega}_{xx}^{rd}\right)^{-1}\right), \tag{9}$$

where the expressions for $\underline{\Phi}_{ux}^{r}$ and $\underline{\Omega}_{xx}^{r}$ are given in Hjalmarsson (2004). By using information only up till time t in the demeaning of $x_{i,t}$ and only information after time t in the demeaning of $r_{i,t}$, the distortive effects arising from standard demeaning are eliminated. Standard $t-$tests can now be performed, using consistent estimators of $\underline{\Phi}_{ux}^{rd}$ and $\underline{\Omega}_{xx}^{rd}$, also given in Hjalmarsson (2004). Thus, the panel-based inference does not require any knowledge of the parameters C_i, either for estimation or testing. The nuisance parameter problem arising from C_i in time-series inference is therefore no longer an issue.

The international panel of stock returns analyzed in this paper is unbalanced, with data for some countries dating further back than for others. The methods just described extend readily to unbalanced panels; the details are given in Hjalmarsson (2004).

Under Assumption 1, the innovations $u_{i,t}$ and $v_{i,t}$ are cross-sectionally independent. This is in general too restrictive when dealing with international financial data where, for instance, global shocks might be present. In order to account for the possibility of cross-sectional dependence, the model in equations (1) and (2) is extended as follows. Let,

$$r_{i,t} = \beta_i x_{i,t-1} + \gamma_i \Lambda_t + u_{i,t}, \tag{10}$$

and

$$x_{i,t} = z_{i,t} + \delta_i \Pi_t, \tag{11}$$

where

$$z_{i,t} = A_i z_{i,t} + v_{i,t}, \tag{12}$$

and

$$\Pi_t = G_\Pi \Pi_{t-1} + \eta_t. \tag{13}$$

The idiosyncratic error terms $u_{i,t}$ and $v_{i,t}$ still satisfy Assumption 1, but the common factors Λ_t and Π_t are now part of the return and regressor processes, respectively. The factor Λ_t is assumed to be stationary and satisfy the functional law, $T^{-1/2}\sum_{s=1}^{t}\Lambda_s \Rightarrow B_{\Lambda}(r)$, for $t=[Tr]$, as $T\rightarrow\infty$. The process Π_t is an auto-regressive process where the parameter G_{Π} is assumed to be local-to-unity, so that $T^{-1/2}\Pi_t \Rightarrow J_{\Pi}(r)=\int_0^r e^{(r-s)C_{\Pi}}dB_{\Pi}(r)$, as $T\rightarrow\infty$. The above specification allows for a general factor structure in both the regressand and the regressor. Hjalmarsson (2004) shows that the effects of these common factors can be controlled for by removing the common factor Λ_t from the error term of the returns process. This is done by performing a first-stage OLS time-series regression for each time-series i, and obtain estimates of the residuals $\gamma_i\Lambda_t+u_{i,t}$. From the estimated residuals, estimates of $\{\gamma_i\}_{i=1}^{n}$ and $\{\Lambda_t\}_{t=1}^{T}$ can be obtained and the 'de-factored' data $r_{i,t}^{df}=r_{i,t}-\hat{\gamma}_i\hat{\Lambda}_t$ is created. Using $r_{i,t}^{df}$ instead of $r_{i,t}$ in the pooled regression controls for the effects of the common factors in the returns. The common factors in the regressors will be implicitly accounted for when estimating the variance-covariance matrix and standard $t-$tests will be normally distributed. In the empirical analysis, estimation of the common factors is done through an extension of the principal component method to unbalanced panels, described in Stock and Watson (2000).

4 Long-run estimation

4.1 The time-series case

In long-run regressions, the focus of interest is fitted regressions of the type

$$r_{i,t+q}(q)=\alpha_i^U(q)+\beta_i^U(q)\,x_{i,t}+u_{i,t+q}(q), \tag{14}$$

and

$$r_{i,t+q}(q)=\alpha_i^B(q)+\beta_i^B(q)\,x_{i,t}(q)+u_{i,t+q}(q), \tag{15}$$

where $r_{i,t}(q)=\sum_{j=1}^{q}r_{i,t-q+j}$ and $x_{i,t}(q)=\sum_{j=1}^{q}x_{i,t-q+j}$. In equation (14), long-run future returns are regressed onto a one period predictor, whereas in equation (15), long-run future returns are regressed onto long-run past regressors. Equation (14) is the specification most often used for testing stock return predictability, although Fama and French (1988b) use (15) in a univariate framework where sums of future returns are regressed onto sums of past returns. The theoretical results of both Hansen and Tuypens (2004) and Valkanov (2003) suggest that equation (15) may have some desirable properties and I will consider both kinds of specifications here. The regressions in equation (14) and (15) will be referred to as the unbalanced and balanced regressions, respectively, since in the former case long-run returns are regressed onto short-run predictors and in the latter long-run returns are regressed onto long-run predictors. This choice of terminology, i.e unbalanced and balanced, is used purely as a mnemonic device; 'unbalanced' is not meant to convey anything negative about this specification.

Let the OLS estimators of $\beta_i^U(q)$ and $\beta_i^B(q)$ in equations (14) and (15), using overlapping observations, be denoted by $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$, respectively. A long-standing issue in the return-forecasting literature is the calculation of correct standard errors for $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$.[7] Since overlapping observations are used to form the estimates, the residuals $u_{i,t}(q)$ will exhibit serial correlation; standard errors failing to account for this fact will lead to biased inference. The common solution to this problem has been to calculate auto-correlation robust standard errors, using methods described by Hansen and Hodrick (1980) and Newey and West (1987). However, these robust estimators tend to have rather poor finite sample properties; this is especially so in cases when the serial correlation is strong, as it often is when overlapping observations are used. In this section, I derive the asymptotic properties of $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$ under the assumption that the forecasting horizon q grows with the sample size but at a slower pace. The results complement those of Valkanov (2003), who treats the case where the forecasting horizon grows at the same rate as the sample size, and those of Hansen and Tuypens (2004) who also consider the case where $q/T \to 0$ as $q, T \to \infty$, but in a covariance stationary setup. As is seen below, the endogeneity of the forecasting variables plays as important a role in long-run regressions as they do in short-run regressions. This point is somewhat obscured by the asymptotic results derived in Valkanov (2003), and effectively not treated in the covariance stationary model of Hansen and Tuypens (2004).

Given that equations (14) and (15) are estimated with overlapping observations, created from short-run data, they should be viewed as fitted regressions rather than actual data generating processes (dgp); the use of overlapping observations effectively necessitates the specification of a dgp for the observed short-run data. The results below are derived under the assumption that the true dgp satisfies equations (1) and (2), and that the long-run observations are formed by summing up data generated by that process. Under the null hypothesis of no predictability, the one period dgp is simply $r_{i,t} = u_{i,t}$, in which case the long-run coefficients $\beta_i^U(q)$ and $\beta_i^B(q)$ will also be equal to zero. It follows, that under the null, both equations (14) and (15) are correctly specified and the analysis of $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$ simplifies. It is therefore common in the literature to only derive asymptotic results for long-run estimators under the null of no predictability. By considering the properties of the estimators both under the null and the alternative, however, a more complete picture of the properties of the long-run estimators emerges. Of course, equation (1) is only one possible alternative to the null of no predictability, but it provides a benchmark case.

Theorem 1 *Suppose the data is generated by equations (1) and (2), and that Assumption 1 holds.*

1. *Under the null hypothesis that $\beta_i = 0$, as $q, T \to \infty$, such that $q/T \to 0$,*

(a)

$$\frac{T}{q}\left(\hat{\beta}_i^U(q) - 0\right) \Rightarrow \left(\int_0^1 dB_{i,1}\underline{J}_{i,C_i}'\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right)^{-1}, \tag{16}$$

(b)

$$T\left(\hat{\beta}_i^B(q) - 0\right) \Rightarrow \left(\int_0^1 dB_{i,1}\underline{J}_{i,C_i}'\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right)^{-1}. \tag{17}$$

[7] There is now a large literature on regressions with overlapping observations. Classiscal references include Hansen and Hodrick (1980), Richardson and Stock (1989), Richardson and Smith (1991), and Hodrick (1992). Some examples of recent research are Campbell (2001), Daniel (2001), Valkanov (2003), Britten-Jones and Neuberger (2004), Hansen and Tuypens (2004), Mark and Sul (2004), Moon et al. (2004), and Torous et al. (2005).

2. Under the alternative hypothesis that $\beta_i \neq 0$, as $q, T \rightarrow \infty$, such that $q/T \rightarrow 0$,

(a)

$$\frac{2T}{q^2}\left(\hat{\beta}_i^U(q) - \beta_i^U(q)\right) \Rightarrow \beta_i \left(\int_0^1 dB_{i,2}\underline{J}'_{i,C_i} + \Lambda_{22i}\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}'_{i,C_i}\right)^{-1}, \tag{18}$$

(b)

$$\frac{T}{q}\left(\hat{\beta}_i^B(q) - \beta_i^B(q)\right) \Rightarrow \beta_i \left(\int_0^1 dB_{i,2}\underline{J}'_{i,C_i} + \Omega_{22i}\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}'_{i,C_i}\right)^{-1}, \tag{19}$$

where $\beta_i^U(q) = \beta_i\left(I + A_i + ... + A_i^{q-1}\right)$ and $\beta_i^B(q) = \beta_i A_i^{q-1}$. Since $A_i = I + C_i/T$, it follows that $\beta_i^U(q)/q = \beta_i + O(q/T) \rightarrow \beta_i$ and $\beta_i^B(q) = \beta_i + O(q/T) \rightarrow \beta_i$, as $q, T \rightarrow \infty$, such that $q/T \rightarrow 0$.

Theorem 1 shows that under the null of no predictability, the limiting distributions of $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$ are identical to that of the plain short-run OLS estimator $\hat{\beta}_i$, although $\hat{\beta}_i^U(q)$ needs to be standardized by q^{-1}, since, as seen in part *2* of the theorem, the estimated parameter $\beta_i^U(q)$ is of an order q times larger than the original short-run parameter β_i. Under the alternative hypothesis of predictability, the limiting distributions of $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$ are quite different from the short-run result, and are in fact similar to the distribution of the OLS estimator of the first order auto-regressive root in $x_{i,t}$, although the rate of convergence is slower. The estimators still converge to well defined parameters under the alternative hypothesis, but their asymptotic distributions are driven by the auto-regressive nature of the regressors and the fact that the fitted regressions in (14) and (15) are effectively miss-specified, under the assumption that the true relationship takes the form of equation (1).

It is apparent that, under the null hypothesis, the long-run OLS estimators suffer from the same endogeneity problems as does the short-run estimator. Similar remedies to those discussed for the short-run case, such as the fully modified approach, can be considered. Estimates of Ω_{22i} and Λ_{22i} can be obtained directly from the short-run specification of $x_{i,t}$ in equation (2). By a first-stage long-run OLS regression, estimates of the residuals $u_{i,t}(q)$ can be obtained and the covariance ω_{12i} can be estimated. However, simulations not reported in the paper show that estimates of ω_{12i} based on the long-run residuals $u_{i,t}(q)$ are often very poor for typical values of q. Thus, unless one resorts to a *short-run* OLS first-stage regression, which seems unappealing in a long-run estimation exercise, the actual finite sample properties of the fully-modified long-run estimators appear unsatisfactory. However, by imposing somewhat more restrictive assumptions on the error process $v_{i,t}$ in the regressors, an alternative solution can be considered.

Assumption 2 *Let $w_{i,t} = (u_{i,t}, v_{i,t})'$ satisfy conditions 2-4 in Assumption 1, with $\epsilon_{i,t}$ replaced by $v_{i,t}$ in condition 2. That is, $v_{i,t}$, as well as $u_{i,t}$, are martingale difference sequences with finite fourth order moments.*

Under Assumption 2, the long-run covariance matrix Ω_i is now identical to the short-run covariance Σ_i, although I continue to use the long-run notation to be consistent with previous notation.

Consider the fitted augmented regression equations

$$r_{i,t+q}(q) = \alpha_i^U(q) + \beta_i^U(q)\, x_{i,t} + \gamma_i^U(q)\, v_{i,t+q}(q) + u_{i,t+q\cdot 2}(q), \tag{20}$$

and

$$r_{i,t+q}(q) = \alpha_i^B(q) + \beta_i^B(q)\, x_{i,t}(q) + \gamma_i^B(q)\, v_{i,t+q}(q) + u_{i,t+q\cdot 2}(q), \tag{21}$$

where $v_{i,t}(q) = \sum_{j=1}^{q} v_{i,t-q+j}$. Let $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$ be the OLS estimators of $\beta_i^U(q)$ and $\beta_i^B(q)$ in equations (20) and (21).

Theorem 2 *Suppose the data is generated by equations (1) and (2), and that Assumption 2 holds.*

1. Under the null hypothesis that $\beta_i = 0$, as $q, T \to \infty$, such that $q/\sqrt{T} \to 0$,

$$T\left(\hat{\beta}_i^{B+}(q) - 0\right), \frac{T}{q}\left(\hat{\beta}_i^{U+}(q) - 0\right) \Rightarrow MN\left(0, \sigma_{11\cdot 2,i}\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right)^{-1}\right), \tag{22}$$

where $\sigma_{11\cdot 2,i} = \sigma_{11i} - \omega_{12i}\Omega_{22i}^{-1}\omega_{21i}$.

2. Under the alternative hypothesis that $\beta_i \neq 0$, as $q, T \to \infty$, such that $q/\sqrt{T} \to 0$,

$$\frac{T}{q}\left(\hat{\beta}_i^{B+}(q) - \beta_i^B(q)\right), \frac{2T}{q^2}\left(\hat{\beta}_i^{U+}(q) - \beta_i^U(q)\right) \Rightarrow \beta_i\left(\int_0^1 dB_{i,2}\underline{J}_{i,C_i}'\right)\left(\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right)^{-1}. \tag{23}$$

Under the null hypothesis of no predictability, the estimators $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$ have asymptotically mixed normal distributions, although under the alternative hypothesis of predictability, the asymptotic distributions are still non-standard. The long-run estimators $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$, which correct for the endogeneity effects of nearly persistent regressors, are to the best of my knowledge the first to appear in the literature. Given the asymptotically mixed normal distributions of $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$ under the null hypothesis, standard test procedures can now be applied to test the null of no predictability. In fact, the following convenient result is easy to prove.

Corollary 1 *Let $t_i^{U+}(q)$ and $t_i^{B+}(q)$ denote the standard $t-$statistics corresponding to $\hat{\beta}^{U+}(q)$ and $\hat{\beta}^{B+}(q)$. That is,*

$$t_i^{U+}(q) = \frac{\hat{\beta}_{i,k}^{U+}(q) - \beta_{i,k}^{U,0}(q)}{\sqrt{\left(\frac{1}{T}\sum_{t=1}^{T}\hat{u}_i^{U+}(q)^2\right) a'\left(\sum_{t=1}^{T}\underline{z}_{i,t}\underline{z}_{i,t}'\right)^{-1} a}}, \tag{24}$$

and

$$t_i^{B+}(q) = \frac{\hat{\beta}_{i,k}^{B+}(q) - \beta_{i,k}^{B,0}(q)}{\sqrt{\left(\frac{1}{T}\sum_{t=1}^{T}\hat{u}_i^{B+}(q)^2\right) a'\left(\sum_{t=1}^{T}\underline{z}_{i,t}(q)\,\underline{z}_{i,t}(q)'\right)^{-1} a}}, \tag{25}$$

where $\hat{u}^{U+}(q)$ and $\hat{u}^{B+}(q)$ are the estimated residuals, $\underline{z}_{i,t} = \left(\underline{x}_{i,t}, v_{i,t+q}(q)\right)$, $\underline{z}_{i,t}(q) = \left(\underline{x}_{i,t}(q), v_{i,t+q}(q)\right)$ and a is an $2m \times 1$ vector with the k'th component equal to one and zero elsewhere. Then, under the *null-hypotheses of* $\beta_i = 0$,

$$\frac{t_i^{U+}(q)}{\sqrt{q}}, \frac{t_i^{B+}(q)}{\sqrt{q}} \Rightarrow N(0,1). \tag{26}$$

Thus, long-run inference can be performed by simply scaling the corresponding standard $t-$statistic by $q^{-1/2}$. In the case with covariance stationary regressors, where endogeneity effects play no role, Hansen and Tuypens (2004) derive a similar scaling result for the standard $t-$statistics corresponding to $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$.

The results in Theorems 1 and 2 bring some clarity to the properties of long-run regressions with nearly persistent regressors. Under the null of no predictability, the long-run estimators have identical asymptotic distributions to the short-run estimators. Under the alternative hypothesis of predictability, however, the asymptotic properties of the long-run estimators change substantially and the results are now driven by the *de facto* miss-specification of the long-run regressions, and the auto-regressive nature of the regressors; this is manifest in both the slower rate of convergence as well as the non-standard limiting distribution.

All of the above asymptotic results are derived under the assumption that the forecasting horizon grows with the sample size, but at a slower rate. Torous et al. (2005) and Valkanov (2003) also study long-run regressions with near-integrated regressors, but derive their asymptotic results under the assumption that $q/T \rightarrow \kappa \in (0,1)$ as $q, T \rightarrow \infty$. That is, they assume that the forecasting horizon grows at the same pace as the sample size. Under such conditions, the asymptotic properties of $\hat{\beta}_i^U(q)$ and $\hat{\beta}_i^B(q)$ are quite different from those derived in this paper. There is, of course, no right or wrong way to perform the asymptotic analysis; what matters in the end is how well the asymptotic distributions capture the properties of actual finite sample estimates. To this end, a brief Monte Carlo simulation is therefore conducted.

Equations (1) and (2) are simulated, with $u_{i,t}$ and $v_{i,t}$ drawn from an *iid* bivariate normal distribution with mean zero, unit variance and correlation $\delta = -0.9$. The large negative correlation is chosen to assess the effectiveness of the endogeneity corrections in $\hat{\beta}^{U+}(q)$ and $\hat{\beta}^{B+}(q)$, as well as to reflect the sometimes high endogeneity of regressors such as the dividend- or earnings-price ratio. The intercept α_i is set to one and the auto-regressive root A_i is also set to unity. Three different estimators, and their corresponding $t-$statistics, are considered: the long-run estimators, $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$, as well as the short-run OLS estimator in the augmented regression equation (20) (or equivalently (21)).[8] Since the aim of the simulation is to determine how well the asymptotic distributions derived above reflect actual finite sample distributions, all estimation and testing is done under the assumption that the root A_i is known. The sample sizes are chosen as $T = 100$ and $T = 500$.

The first part of the simulation study evaluates the finite sample properties of the three estimators under an alternative of predictability, where the true β_i is set equal to 0.05. The second part analyzes the size and power properties of the scaled $t-$tests, and the third part shows the properties of the long-run estimators as the forecasting horizon grows but the sample size is kept fixed. In all except the last exercise, the forecasting horizon is set to $q = 12$ and $q = 60$ for the $T = 100$ and $T = 500$

[8] As shown by Phillips (1991), in the case of normally distributed errors, the OLS estimator in the short-run ($q = 1$) augmented regression equation (20) will in fact be equal to the maximum likelihood estimator.

samples, respectively. These forecasting horizons are similar to those often used in practice for similar sample sizes. All results are based on $10,000$ repetitions.

The results are shown in Figure 1. In the top two graphs, A1 and A2, the kernel estimates of the densities of the estimated coefficients are shown. To enable a comparison, the $\hat{\beta}^{U+}(q)$ estimate is scaled by q^{-1}. The non-standard distributions of $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$ under the alternative are evident, especially so for $\hat{\beta}_i^{B+}(q)$. The fact that $\hat{\beta}_i^{U+}(q)$ converges faster than $\hat{\beta}_i^{B+}(q)$ under the alternative, after scaling $\hat{\beta}_i^{U+}(q)$ by q^{-1} is also clear. The short-run estimator outperforms both long-run estimators, however. In the middle graphs, B1 and B2, the rejection rates of the 5% two-sided $t-$tests, for tests of the null of no predictability, are given. For both $T = 100$ and $T = 500$, all three tests have a rejection rate very close to 5% under the null, so the scaling of the long-run $t-$statistics by $q^{-1/2}$ appears to work well in practice, as well as the endogeneity correction implicit in $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$. The test based on the $\hat{\beta}_i^{U+}(q)$ estimator has similar power properties to the short-run test, although the short-run test performs better in all instances. The test based on $\hat{\beta}_i^{B+}(q)$ performs rather poorly, especially in the larger sample with the longer forecasting horizon. In the bottom graphs, C1 and C2, I illustrate the effects, on the long-run point estimates, of an increase in the forecasting horizon as the sample size stays fixed. Three different cases are considered, $\beta_i = 0.05, 0.00, -0.05$. It is interesting to note that as q grows larger relative to the sample size, the long-run estimates tend to drift towards the opposite sign. Only under the null-hypothesis do they not tend to drift, although $\hat{\beta}_i^{B+}(q)$ does so a bit in the small sample. These last simulation results show the importance of not using too large a horizon relative to the sample size.

In summary, the simulation results for the size and power properties, in particular, show that the endogeneity correction performed in $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$ appears to work well and that the scaling of the $t-$statistic, as suggested by Corollary 1, achieves the correct size.

Both the asymptotically slower rate of convergence for $\hat{\beta}_i^{U+}(q)$ and $\hat{\beta}_i^{B+}(q)$ under the alternative of predictability and the finite sample results given in Figure 1 indicate that there is little reason to consider long-run tests if one believes that the alternative model of stock return predictability is given by equation (1). This is not to say that long-run procedures are not useful. Long-run regressions effectively perform a smoothing of the data, for both the dependent and independent variables in the case of $\hat{\beta}_i^{B+}(q)$ and for the dependent variable only in the case of $\hat{\beta}_i^{U+}(q)$. It is likely that there are situations where such smoothing is desirable and the relationship in the smoothed data reveals properties not easily detected in the short-run data. For instance, one could consider a setting where asset prices are prone to temporarily drift away from their 'fundamental', or rational, values. Suppose that the earnings-price ratio predicts future stock returns when returns and prices reflect fundamentals. Then short-run tests of stock return predictability may be less likely to capture this relationship than long-run tests, since, in the short run, the prices and returns might not be determined by fundamentals. But, as long as the forecasting horizon is large enough, the long-run regression will capture the relationship.

Likewise, the advantages of $\hat{\beta}_i^{B}(q)$ over $\hat{\beta}_i^{U}(q)$ indicated by the work of Valkanov (2003) and Hansen and Tuypens (2004) do not appear in the results above, but might also be realized under different alternative models of stock return predictability. Intuitively, the smoothing of the regressor in $\hat{\beta}_i^B(q)$, versus no smoothing in $\hat{\beta}_i^U(q)$, provides a trade-off between reducing noise and using the latest available information in forming the forecasts. The best approach will depend on the relative importance of these factors.

4.2 The panel case

As discussed in Section 3.2, the demeaning of the data in the standard short-run pooled estimator causes a second order bias when the regressors are endogenous. The same will be true in the long-run case, but the recursive demeaning solution used for the short-run estimator is less practical in the long-run; the overlapping nature of the data would necessitate a large loss of observations in the calculations of the recursively demeaned data. Instead, an approach similar to that in the time-series case will be used. Let $\hat{\beta}_{n,T}^{U+}(q)$ and $\hat{\beta}_{n,T}^{B+}(q)$ be the pooled estimators of $\beta^U(q)$ and $\beta^B(q)$ in the augmented regressions (20) and (21), respectively, allowing α_i and γ_i to vary across i. The exact expressions for $\hat{\beta}_{n,T}^{U+}(q)$ and $\hat{\beta}_{n,T}^{B+}(q)$ are given in the proof of the following theorem.

Theorem 3 *Suppose the data is generated by equations (1) and (2), and that Assumption 1 holds. Further, suppose the slope coefficients are homogenous so that $\beta_i = \beta$ for all i. Under the null hypothesis that $\beta = 0$, as $(T, n \to \infty)_{\text{seq}}$ and $q \to \infty$ such that $q/\sqrt{T} \to 0$,*

$$\sqrt{n}\frac{T}{q}\left(\hat{\beta}_{n,T}^{U+}(q) - 0\right), \sqrt{n}T\left(\hat{\beta}_{n,T}^{B+}(q) - 0\right) \Rightarrow N\left(0, \underline{\Omega}_{xx}^{-1}\underline{\Phi}_{u\cdot v,x}\underline{\Omega}_{xx}^{-1}\right),$$

where $\underline{\Omega}_{xx} = E\left[\int_0^1 \underline{J}_{i,C_i}\underline{J}_{i,C_i}'\right]$ *and* $\underline{\Phi}_{u\cdot v,x} = E\left[\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}_{i,C_i}'\right)\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}_{i,C_i}'\right)'\right]$.

This result most likely also holds in joint limits, under some extra rate restrictions on n and T, and could probably be proved using similar methods to those of Hjalmarsson (2004) and Phillips and Moon (1999).[9] However, the extra technical detail required for such a proof does not seem justified in the present context. For brevity, the results are given under the null-hypothesis of no predictability and under the assumption of homogenous slope coefficients.

Let $t_{n,T}^{U+}(q)$ and $t_{n,T}^{B+}(q)$ denote the pooled $t-$statistic from the augmented pooled estimation, defined in the proof of the following corollary.

Corollary 2 *Under the null hypothesis that $\beta = 0$, as $(T, n \to \infty)_{\text{seq}}$ and $q \to \infty$ such that $q/\sqrt{T} \to 0$,*

$$t_{n,T}^{U+}(q), t_{n,T}^{B+}(q) \Rightarrow N(0,1).$$

Thus, in the pooled case, no scaling of the $t-$statistics is required. This result follows from the fact that the natural way of estimating $\underline{\Phi}_{u\cdot v,x}$ for the pooled $t-$statistic leads to a heteroskedasticity

[9] The notation $(T, n \to \infty)_{\text{seq}}$ follows that of Phillips and Moon (1999) and indicates a sequential limit result derived by first keeping n fixed and letting T go to infinity and then letting n go to infinity.

and autocorrelation consistent (HAC) estimator, although the panel structure avoids the usual non-parametric shape of the estimator. One could consider a non-HAC estimator by deriving more explicit expressions for $\underline{\Phi}_{u\cdot v,x}$, but in this case the convenient result that the test works both in the homogenous and heterogenous slope coefficient cases no longer hold, as is evident from the analysis in Hjalmarsson (2004).

The effects of common factors can be dealt with in the same manner as for the short-run case; by de-factoring the returns data. As pointed out above, it is somewhat undesirable to rely on a first-stage short-run regression in a long-run analysis. In the case of obtaining estimates of the common factors in the returns residuals, however, first-stage short-run OLS time-series regressions seem more appropriate than long-run time-series regressions; the overlapping nature of the long-run residuals is likely to prove troublesome when attempting to extract common factors. The de-factored data, $r_{i,t}^{df}$ is thus created as in the short run, and the long-run pooled regressions are then estimated using long-run returns formed from $r_{i,t}^{df}$.

5 Feasible methods

To implement the methods described in the two previous sections, with the exception of the short-run pooled estimator in Section 3.2, knowledge of the parameters $\{C_i\}_{i=1}^{n}$ is required. Since C_i is typically unknown and not estimable in general, I rely on the bounds procedures of Cavanagh et al. (1995) and Campbell and Yogo (2003) to obtain feasible procedures. The following discussion assumes a scalar regressor, as do the above studies.

Although C_i is not estimable, a confidence interval for C_i can be obtained, as described by Stock (1991). By evaluating the estimator and test-statistic for each value of C_i in that confidence interval, a range of possible estimates and values of the test-statistic are obtained. A conservative test can then be formed by choosing the most conservative value of the test statistic, given the alternative hypothesis. If the confidence interval has a coverage rate of $100\,(1-\alpha_1)\,\%$ and the nominal size of the test is α_2, then by Bonferroni's inequality the final conservative test will have a size no greater than $\alpha = \alpha_1 + \alpha_2$. In general, the size of the test will be less than α, and a test with a pre-specified size can be achieved by fixing α_2 and adjusting α_1. In practice, the test-statistics used in this paper are monotone in C_i and only the end points of the confidence interval for C_i need be evaluated. A drawback of this method is that no clear-cut point estimate is produced, but rather a range of estimates. In the result section of the paper I therefore report the standard OLS point estimate, or in the case of the long-run pooled estimation, the standard fixed effects estimate.

For practical inference in the time-series regressions, I adopt a similar approach to Campbell and Yogo (2003), but rule out the possibility of explosive roots $C_i > 0$. A confidence interval for C_i is obtained by inverting the DF-GLS statistic. Table 2 of Campbell and Yogo (2003) is used to find the desired significance level of this confidence interval in order for the final test to have a one-sided 5% size. If the upper bound is greater than zero, it is simply replaced by zero. The test statistics, either the $t_i^{+}-$statistic in the short-run case or the scaled $t_i^{U+}(q)$ and $t_i^{B+}(q)$ statistics in the long-run, are evaluated for the upper and lower bounds of this confidence interval, and the more conservative of the two values, given the alternative one-sided hypothesis, is reported as a conservative one-sided test and normal significance levels apply. That is, if the reported statistic is greater than 1.65, the null can be rejected at the 5% level against an alternative positive hypothesis, and analogously for the negative alternative. In practice, ruling out the possibility of explosive roots have relatively little effect on the results and does not affect the overall conclusions.

In the case of the long-run pooled estimators, $\hat{\beta}^{U+}_{n,T}(q)$ and $\hat{\beta}^{B+}_{n,T}(q)$, and the corresponding $t-$statistics, the bounds procedures just described for the time-series case become cumbersome, especially if one wants to chose α_1 in a manner that produce tests with the correct size. Instead, in the panel case, I use an approach similar to that of Lewellen (2003). It is easy to show, for the pooled estimators, that the estimates will be decreasing in C_i if the average ω_{12i}, denoted ω_{12}, is negative, and increasing in C_i if ω_{12} is positive. Thus, if $C_i \in (-\infty, 0]$, and $\omega_{12} < 0$, a conservative test of $H_0 : \beta = 0$ versus $H_1 : \beta > 0$ is obtained by evaluating $t^{U+}_{n,T}(q)$ and $t^{B+}_{n,T}(q)$ at $C_i = 0$ for all i. Similarly, if one wants to construct a conservative test of $H_0 : \beta = 0$ versus $H_1 : \beta < 0$, for $\omega_{12} < 0$, the test-statistics should be evaluated at $C_i \to -\infty$ for all i. As $C_i \to -\infty$, the process $x_{i,t}$ become stationary and the fixed effects $t-$statistics obtained from pooling the standard long-run regressions in equations (14) and (15) can be used, rather than the augmented regressions in equations (20) and (21). Conservative tests when $\omega_{12} > 0$ are obtained by reversing the arguments for $\omega_{12} < 0$.

6 Data description

All of the data used in this paper come from the Global Financial Data database. Total returns, including direct returns from dividends, on market wide indices in 40 countries were obtained, as well as the corresponding dividend- and earnings-price ratios. In addition, for each country, measures of the short and long interest rates are obtained. All returns and interest rate data are on a monthly frequency. For a few of the older observations, the dividend- and earnings-price ratios are given on an annual basis; these are transformed to monthly data by filling in the monthly dividends or earnings over the year with the previous year's values.

With the exception of Spain, the dividend-price ratio data is available over the same sample period as the total stock returns. But, the other predictor variables are typically not available during the whole sample of total stock returns. Due to the two world wars, France, Germany, Japan, and the U.K. have some years during which no observations are available. Further, Spain's total returns data start in 1940, but no dividends data is available during 1968-1983. In the time-series analysis, separate regressions are fitted for each sample period for these countries, and in the pooled estimation separate intercepts are estimated. In Tables 2-4, which state the pooled results, the row listing the number of 'countries' in each panel can therefore include more than one count of some countries.

The variable definitions follow the usual conventions in the literature. The dividend-price ratio is defined as the sum of dividends during the past year, divided by the current price. The earnings-price ratio is defined in the same manner, i.e. the current price divided by the latest 12 months of earnings available at the time. Both the dividend- and earnings-price ratios are thus expressed in annual units. The measure of the short interest rate comes from the interest rate series constructed by Global Financial Data. These use rates on 3-month T-bills when available or, otherwise, private discount rates or interbank rates. The long rate is measured by the yield on long-term government bonds. When available, a 10 year bond is used; otherwise, I use that with the closest maturity to 10 years. To conform with the one month stock returns, the short and long interest rates are also expressed on a monthly basis. The term spread is defined as the log difference between the long and the short rate.

In defining excess stock returns, two approaches are available: *(i)* the return on stocks, in the local currency, over the local short rate, or *(ii)* the return on stocks in dollar terms, over the U.S. short interest rate. In the latter case, returns in local currencies are transformed to dollar denominated returns using exchange rate data, also obtained from Global Financial Data. During large parts of the sample period, official exchange rates poorly reflect the actual market rates; so, in cases when unofficial black market rates are available, they are used. The results presented in this paper rely mainly on the former approach since it provides the international analogue of the typical forecasting regressions estimated for U.S. data. In addition, using excess returns expressed in domestic currencies avoids the possibility that the results are driven by predictability in exchange rates rather than in stock returns. In practice, either approach delivers similar results in most cases; however, I point out in the text those instances where the results do differ substantially.

Four different forecasting variables are considered: the dividend- and earnings-price ratios, the short interest rate and the term spread.[10] All regressions are run using log-transformed variables with the log excess returns over the domestic short rate as the dependent variable.

As pointed out in the method sections above, there are advantages to pooling data from several countries; these advantages are greatest when the panel is reasonably homogenous. So in addition to pooling all countries into a global panel, I also consider pooling based on OECD membership. These different groups of countries will be referred to as the Global, OECD, and Non-OECD panels. Pooling according to geographical areas such as Europe and Asia was also considered, but the results were similar to those for the OECD pooling and yielded no extra insights. Table 1 provides summary statistics of the data and indicates what countries belong in each group. Table 8 in the Appendix lists the stock-index in each country to which the total returns and dividend- and earnings-price ratios correspond.

7 Empirical results

In the empirical analysis, I conduct pooled regressions as well as time-series regressions for individual countries. Once again, the dependent variable is always excess stock returns over the domestic short

[10] In addition to the standard earnings-price ratio, I also considered the 10 year smoothed earnings price ratio, as suggested by Campbell and Shiller (1988). This is defined as the average earnings over the past 10 years divided by the current price. The results were similar to those of the standard earnings-price ratio, however, and are not included in the paper.

rate; in the discussion, I just write stock returns for simplicity. In both types of analyses, I estimate short- and long-run regressions. While the short-run regressions use the original monthly data, the long-run regressions use overlapping observations at the 12 and 60-month horizons. In addition, both kinds of long-run specifications, given in equations (14) and (15), are evaluated. For the 1-month horizons, time-series with less than 60 observations are not considered, and for the 12 and 60-month horizons a minimum of 120 and 240 observations is required, respectively.

To control for cross-sectional dependence or, more precisely, common factors in the pooled data, I use the approach described previously, considering the cases with one, two, or three common factors in the residuals; the results are fairly similar for all three cases. In general, however, the evidence of stock return predictability tends to be somewhat *stronger* when the de-factored data are used. Given the uncertainty regarding the existence of common factors in the data, I present the somewhat more conservative pooled estimates and test-statistics obtained when the data are not de-factored.[11]

The results from the pooled regressions and summaries of the time-series results are presented in Tables 2-4. The time-series results for individual countries are given in Tables 5 and 6. Each table contains multiple panels, which correspond to different forecasting variables. To preserve space, not all of the results for the individual time-series regressions are presented. In particular, I present all the results for the 1-month horizon, but for the long-run regressions, I only show the 60-month horizon results for the dividend- and earnings-price ratios. The remaining long-run country level results are summarized in Tables 2-4, and described briefly in the text; a complete set of results is available on request. In each case, I present the long-run time-series results based on both the balanced and unbalanced specifications, as indicated in the tables. The unbalanced long-run specification is given in equation (14) and regresses long-run future returns onto the current short-run regressor while the balanced long-run specification is given in equation (15) and regresses long-run future returns onto the long-run past regressor.

For the pooled short-run regressions, two sets of pooled results are given; the fixed effects estimate using recursively demeaned data, $\hat{\beta}^{rd}_{n,T}$, and the corresponding $t-$statistic, $t^{rd}_{n,T}$, as well as the standard fixed effects estimate, $\hat{\beta}_{n,T}$, and $t-$statistic, $t_{n,T}$. The latter $t-$test is not robust to the endogeneity and persistence of the regressors, but in the case of the short interest rate and the term spread, it does in fact provide a conservative test against a negative and positive alternative, respectively. This follows immediately from the discussion in Section 5 and the average negative correlation between $u_{i,t}$ and $v_{i,t}$ for the short rate and the average positive correlation for the term spread. For the dividend- and earnings-price ratio, however, the standard $t-$test will tend to over-reject the null in favour of a positive alternative.

For the pooled long-run regressions, the standard long-run fixed effects estimates $\hat{\beta}^{U}_{n,T}(q)$ and $\hat{\beta}^{B}_{n,T}(q)$ are presented, along with the corresponding pooled $t-$statistics, $t^{U}_{n,T}(q)$ and $t^{B}_{n,T}(q)$. In addition, the pooled $t-$statistics, $t^{U+}_{n,T}(q)$ and $t^{B+}_{n,T}(q)$, from the augmented long-run regressions in equations (20) and (21), with $C_i = 0$ for all i, are given. In the case of the dividend- and earnings-

[11] Bai and Ng (2002) propose a model selection method for determining the number of common factors in a panel. For the data set used in this paper, these methods seemed to consistently over-parametrize the model, almost always delivering the maximum number of factors allowed in the procedure.

price ratios, these will provide conservative tests against a positive alternative. As in the short-run case, the standard test-statstics $t^{U}_{n,T}(q)$ and $t^{B}_{n,T}(q)$ can be used in evaluating the significance of the short rate and the term spread given a negative and a positive alternative, respectively. In the tables, the argument q is suppressed for both the estimators and test-statistics.

The results from the individual time-series regressions are presented in a similar manner to the long-run pooled regressions. The OLS point estimate, $\hat{\beta}_i$, or in the long-run, either $\hat{\beta}^{U}_i(q)$ or $\hat{\beta}^{B}_i(q)$, is presented along with corresponding standard R^2 and scaled $t-$statistic. Further, the robust and conservative test-statistic, t_i^{+}, or in the long-run, $t_i^{U+}(q)/\sqrt{q}$ or $t_i^{B+}(q)/\sqrt{q}$, is given. These are constructed as described in Section 5, and provide a one-sided conservative test against the particular alternative hypothesis that is considered; i.e. a positive alternative for the dividend- and earnings-price ratio as well as the term-spread and a negative alternative for the short interest rate. Finally, for the short-run regressions, an estimate of the correlation, $\hat{\delta}_i$, between $u_{i,t}$ and $v_{i,t}$ in each regression is given.

In all tables, the robust test-statistic from which proper inference can be drawn is in bold type. The number of individual time-series regressions that yield significant coefficients according to this test-statistic, in a one-sided $5\%-$level test, is indicated for each pooled regression in the column labeled $t^{+}_{0.05}$.

7.1 The earnings-price ratio

The results for the earnings-price ratio are presented in Panel A of Tables 2-4. At the 1-month horizon, there is minimal evidence of a positive predictive relationship. Specifically, pooling the data at either the Global or OECD levels does not yield a significant coefficient; however, there is evidence of an on average predictive relationship when pooling at the Non-OECD level. To ensure that the OECD results are not driven by the longer earnings-price ratio time-series available for the U.K. and the U.S., I also estimate these pooled regressions when restricting the sample to observations after 1950. The individual country time-series results confirm the lack of evidence of a predictive relationship in the pooled regressions. In particular, in the post-1950 sample, only two of the 38 time-series regressions (Argentina and the U.K.) yield any significant coefficients.

On the 12-month horizon, there is still no support for a predictive relationship in the pooled Global or OECD regressions. However, when pooling the Non-OECD countries, there is evidence of a significant relationship in both the balanced regression, where long-run returns are regressed onto long-run regressors, and the unbalanced regression, where long-run returns are regressed onto short-run regressors. In both of these cases, though there is only one country for which the time-series regressions yield a significant coefficient. For the OECD countries, the balanced and unbalanced time-series regressions only show a significant predictive relationship in the U.S.

However, there is somewhat stronger evidence of predictability at the 60-month horizon. Pooling at the Global level delivers a significant coefficient in the full sample balanced regression, but significance disappears when restricting the analysis to the post-1950 sample. The significant relationship seen at the Global level is primarily being driven by that seen in the Non-OECD countries, which is evident in both the balanced and unbalanced regressions. The time series regressions indicate a significant predictive relationship for Japan, the Philippines, and Thailand in the balanced specifications and Japan, South Africa, and the U.K. in the unbalanced specifications.

In summary, there is rather weak evidence that the earnings-price ratio predicts stock returns; the majority of evidence that does exist is for Non-OECD countries. It is noteworthy that the null of no predictability would have been rejected in all but three pooled regressions if one relied on non-robust methods that fail to control for the endogeneity and persistence of the regressors. In particular, at the short horizons, the discrepancy between the robust and non-robust test statistics is quite large.

7.2 The dividend-price ratio

Panel B in Tables 2-4 shows the results from pooled regressions with the dividend-price ratio as the regressor. On average, there is no evidence of predictability at a 1-month horizon; this can be seen in Table 2. Nor does the overall picture depicted by the individual time-series regressions indicate a pattern of predictability; a significant predictive relationship is only observed for post-WWII Japan, the U.K., and post-1950 U.S.

Panel B of Table 3 shows that, on average, there is limited evidence of a predictive relationship at the 12-month horizon in the balanced regression Specifically, the coefficient is significant when pooling both at the post-1950 Global level and the Non-OECD level. The individual time-series regressions provide some extra support for predictability, both in OECD and non-OECD countries. Evidence of predictability at the long-run horizon of 60 months is similar to that seen at the 12-month horizon. At the 60-month horizon, a significant predictive relationship is seen in either the balanced or unbalanced time-series regressions for the following countries: Canada, Chile, pre-WWII France and Germany, Greece, Hong Kong, post-WWII Japan, South Africa, Thailand, and the U.K..

Thus, there is no evidence of 'on average' predictability at short horizons and only limited evidence of such a relationship at longer horizons. The individual time-series regressions provide evidence of predictability at both short and long horizons, although this evidence is much stronger at the longer horizons. Though the results for the dividend-price ratio are, in general, parallel to those of the earnings-price ratio, they appear to be slightly stronger overall due to the greater number of significant time series results. In addition, as in the case of the earnings-price ratio, the null of no predictability would have been often rejected when using non-robust tests.

7.3 The short interest rate

In light of the empirical evidence seen in U.S. data, one would expect there to be a negative relationship between the current short rate and future stock returns. Hence, the robust $t-$tests are all obtained under the assumption of a negative alternative. The data used in all interest rate regressions are restricted to start in 1952 or after, following the convention used in studies with U.S. data.[12] The pooled

[12] In the U.S., the interest rate was pegged by the Federal Reserve before this date. Of course, in other countries, deregulation of the interest rate markets occurred at different times, most of which are later than 1952. As seen in the international finance literature (e.g. Kaminsky and Schmukler, 2002), however, it is often difficult to determine the

results for the short interest rate are presented in Panel C of Tables 2-4. Three pooled specifications are considered: Global, OECD, and Non-OECD.

At the 1-month horizon, the null of no predictability is strongly rejected in the pooled sample of OECD countries. In contrast to the strongly significant negative relationship seen in these countries, the average relationship in the Non-OECD countries is not significant. Given the rather capricious character of interest rates in many Non-OECD countries (e.g. Argentina), I focus strictly on the OECD results. As seen in Panel C of Table 5, this finding of predictability, at the 1-month horizon, is supported by the results of the individual time-series regressions for the OECD countries. In particular, a significant predictive relationship is found in eight out of 21 OECD countries, including: Canada, Germany, the Netherlands, New Zealand, Portugal, Spain, Switzerland, and the U.S.. In addition, a closer look at the individual country level results further strengthens this pattern; in particular, it reveals that the estimates for 15 of the OECD countries are more than one standard deviation away from zero while the robust $t-$statistic is negative for 18 of the 21 countries.

As seen in Panel C of Table 3, there is also, on average, predictability at the 12-month horizon for OECD countries. The individual time-series regressions provide some support of this finding, though the results are weaker than those for the 1-month horizon. At the 60-month horizon, all evidence of the negative predictive relationship seen in the previous pooled regressions disappears and there is a tendency for the coefficients to change sign. This is consistent with the simulation results, presented in Figure 1, for the behaviour of the long-run estimators as the forecasting horizon increases, when there is in fact a short-run relationship.

For the most part, the predictive ability of the short interest rate appears to be a short-run phenomenon. This conclusion is supported by Ang and Bekaert (2003) who find evidence of a significant relationship in the short run for the U.S. and Germany. However, my results indicate that their findings are not limited to these countries, but rather, they are representative of other OECD countries.

7.4 The term spread

Based on the U.S. experience, one would expect there to be a positive predictive relationship, if any, between the term spread and stock returns. As in the case of the short interest rate, I find a positive predictive relationship only in OECD countries and primarily in the short run. As shown in Panel D of Table 2, there is evidence of an average predictive relationship at the 1-month horizon when pooling the OECD countries. As this relationship is not evident for the Non-OECD countries, I once again focus on the results for the OECD countries. As seen in Panel D of Table 5, this finding of predictability, at the 1-month horizon, is supported by the results of the individual time-series regressions for the OECD countries. For nine of 21 individual time-series regressions, there is a positive and significant predictive relationship: Canada, France, Germany, Italy, the Netherlands, New Zealand, Spain, Switzerland, and the U.S.. Furthermore, 13 countries have a coefficient that is more than one standard deviation from

exact date of deregulation. And, if one follows classification schemes, such as those in Kaminsky and Schmukler (2002), then most markets are not considered to be fully deregulated until the 1980s, resulting in a very small sample period to study. Thus, the extent to which observed interest rates reflect actual market rates is hard to determine and one should keep this caveat in mind when interpreting the results.

zero.

The pooled evidence of predictability is quite similar at the 12-month horizon when using either the balanced or unbalanced regressions. However, the time-series evidence is somewhat weaker than that seen at the 1-month horizon. Specifically, in the unbalanced regressions, just five of the 21 OECD countries have significant coefficients.

In contrast, such evidence of predictability is not found in the pooled or time-series regressions at the long-run horizon of 60-months. But, similar to the 60-month horizon results for the short interest rate, there is some indication of a change in the direction of the predictive relationship. Thus, as with the short interest rate, the term spread has predictive powers in the short run and for OECD countries.

7.5 Stability over time

How robust are these patterns of predictability to different sample periods? To analyze this, I consider rolling pooled regressions for the OECD countries. Based on the evidence above, I use the 60-month horizon for the dividend- and earnings-price ratios. For these variables, a new country is added to the rolling regression when there are fifteen years of observations available. However, for the interest rate variables, just five years of data are required since the 1-month horizon is used. Conservative confidence intervals, with a 90% coverage rate, are calculated in a manner analogous to the conservative test statistics described above. For the dividend- and earnings-price ratios, the long-run pooled estimates $\hat{\beta}_{n,T}^{B+}(q)$ and $\hat{\beta}_{n,T}^{U+}(q)$, evaluated at $C_i = 0$ for all i, are presented together with a conservative confidence interval. Although the standard long-run pooled estimates, $\hat{\beta}_{n,T}^{B}(q)$ and $\hat{\beta}_{n,T}^{U}(q)$, are presented in Tables 2-4, I show $\hat{\beta}_{n,T}^{B+}(q)$ and $\hat{\beta}_{n,T}^{U+}(q)$ here since they tend to be less variable over time and better reflect the actual variation in the predictive relationship over time; the confidence intervals are, of course, not affected. Similarly, for the interest rate regressions, the standard short-run fixed effects estimate is used rather than the pooled estimator for recursively demeaned data since recursive demeaning of the often short time-series in the panels could lead to noisy estimates; again, conservative confidence intervals are presented. Since the upper and lower bounds of the conservative confidence intervals are based on two different test-statistics, that yield conservative tests in the respective direction, the confidence bounds are generally not symmetric.

The results are presented in Figure 2. The confidence interval from the unbalanced regressions for the earnings-price ratio lies above zero during most of the sample period after 1950; the results for the balanced specification are less conclusive. The confidence bounds for the dividend-price ratio are virtually centered around zero, both in the balanced and unbalanced regressions, for the whole sample period. Even though the time-series evidence for the earnings- and dividend-price ratios presented above indicates that the dividend-price ratio is a somewhat better predictor, this is not supported by the rolling pooled regressions for the OECD sample. In fact, the opposite appears to be true.

The term spread coefficient fluctuates around zero until the late 1970s, after which the lower bound of the confidence interval hovers above zero. It is only for the short rate that the coefficient is significantly different from zero during the whole sample period that is analyzed. In addition, I present rolling regression estimates for the short interest rate, with corresponding conservative confidence intervals, in Figure 3, for the twelve OECD countries with the longest sample periods; thus, these plots are based on the time-series rather than the pooled estimates. For all of the countries considered, except Japan, a very similar pattern is evident. After around 1980, the estimated coefficients and confidence intervals stabilize; in most cases, this occurs on or below zero, indicating a significant or near significant negative relationship. The main exception is Sweden, where the point estimate is actually above zero for most of the period after 1980.

In summary, this analysis solidifies the case for the predictive ability of the interest rate variables and, in particular, for the short rate. But, it also gives further evidence that the earnings-price ratio does have some predictive ability, even though it may have weakened over the past ten years.

7.6 The U.S. effect

As seen above, the interest rate variables appear to be the strongest predictors of stock returns. Given the global influence of the U.S. economy, it is natural to ask whether the U.S. interest rate has predictive power in other markets. To this end, I estimate individual country forecasting regressions and pooled regressions using either the U.S. short rate or the U.S. term spread as predictors. A summary of these results are presented in Panels E and F of Tables 2-4. Since the same regressor is used for all countries, the analytics for the pooled regressions are somewhat different than those described previously. It is easy to show that for exogenous regressors the same test-statistics as before can be applied, even though the actual asymptotic distributions differ; for the case of endogenous regressors the situation becomes somewhat more complicated. However, since both the U.S. short rate and the U.S. term spread are almost exogenous (the average correlations between the residuals are 0.031 and 0.033, respectively), there will be only minor effects from ignoring the endogeneity. Consequently, only the standard test-statistics are presented for the pooled estimates.

Though somewhat weaker, the results for the U.S. short rate and term spread are similar to those of their domestic counterparts. Using either the U.S. short rate or term spread, at the 1-month horizon, there is a significant predictive relationship in 7 of the 39 countries considered. These country level results are also supported by very strong significance in the pooled inference. As in the case of the domestic interest rates, the pooled evidence is strongest for the OECD countries. As in the case of the domestic variables, the predictive ability of the U.S. interest rate variables is primarily seen in the short run.

When using the U.S. short rate and term spread as predictors, the results vary greatly depending on whether the dependent variable is defined as domestic excess stock returns or dollar transformed excess returns over the U.S. short interest rate. In contrast to the results described in the preceding paragraph, evidence of the predictive ability of U.S. interest rate variables becomes much stronger when using dollar transformed excess returns. In fact, the U.S. short rate is a significant predictor in 24 out of the 39 countries considered. However, this result is somewhat misleading since a substantial share of that predictive ability can be attributed to predictions of the change in the exchange rate rather than predictions of the excess stock returns.[13]

[13] Complete results for these regressions using dollar transformed excess returns over the U.S. short rate are available

8 Out-of-sample evidence and economic implications

Goyal and Welch (2003a,b, 2004) present what appears to be a very strong critique against the presence of stock return predictability. They argue that the significant evidence of predictability found in U.S. data, based on traditional 'in-sample' statistical econometric procedures, does not hold up in out-of-sample exercises. In particular, they conclude that forecasting the next-period stock returns with the historical sample mean of past returns will almost always outperform a conditional forecast based on an estimated forecasting regression, using any of the standard forecasting variables proposed in the literature. Since significant in-sample results coupled with a lack of out-of-sample results are usually a sign of data-mining, the results already presented in this paper go a long way towards addressing the Goyal and Welch critique by analyzing large amounts of new data. Nevertheless, it is interesting to explicitly consider the out-of-sample implications of the above results and give an economic interpretation of the importance of the predictable component that stock returns do seem likely to possess.

Campbell and Thompson (2004) also show that the evidence against stock return predictability presented by Goyal and Welch might not be as relevant as it seems. In the out-of-sample tests performed by Goyal and Welch, it is assumed that the investor mechanically forecasts stock returns using either the historical average of past returns or the estimated regression equation. In both cases, the forecasts are based on estimates using data available up to the time of the forecast. The estimates are updated each period, incorporating the latest data. Campbell and Thompson (2004) argue that by imposing some weak, but common sense, restrictions on the forecasts, their performance can be vastly improved. Specifically, if an estimated coefficient does not have the expected sign, they set it equal to zero, and if the forecast of the equity premium is negative, the forecast is set equal to zero. Using U.S. aggregate data, they show that forecasts based on forecasting variables with significant in-sample coefficients typically outperform the forecasts based on the historical sample mean of past returns.

Here I use international data to perform an analysis similar to that of Campbell and Thompson (2004). Specifically, I present results concerning the out-of-sample performance of the domestic short rate and term spread, which yielded the best in-sample performance. To allow for a sufficient sample size in the out-of-sample analysis, which requires an initial 'training-sample' to obtain the estimates on which the first round of forecasts is based, I exclude all countries with less than 40 years of data; this allows for a 20-year training period and a minimum of a 20-year forecasting period.

The out-of-sample exercise is performed as follows. The first twenty years in each time-series is used to form the initial estimates of the regression coefficients. In each period following the first twenty years of observations, the coefficients are re-estimated, with the latest observations included. Next period's returns are forecasted based on the estimated regression equation; the restrictions suggested by Campbell and Thompson (2004) are imposed in every period. These 'conditional' forecasts, based on the regression model, are compared to the 'unconditional' forecasts, which in each period are identical to the sample mean of the then available past returns; the restriction of a positive forecast of the equity premium is imposed on the unconditional forecasts as well. To compare the statistical performances

upon request.

of the conditional and unconditional forecasts, an out-of-sample R^2 is calculated. This is defined as

$$R_{i,OS}^2 = 1 - \frac{\sum_{t=s}^{T} \left(y_{i,t} - \hat{y}_{i,t}\right)^2}{\sum_{t=s}^{T} \left(y_{i,t} - \bar{y}_{i,t}\right)^2}, \tag{27}$$

where $\hat{y}_{i,t}$ and $\bar{y}_{i,t}$ are the conditional and unconditional forecasts, respectively, and s is the length of the training sample. The $R_{i,OS}^2$ statistic will be positive when the conditional forecast outperforms the unconditional one. Thus, the out-of-sample R^2 is positive when the root mean squared error of the conditional forecast is less than that of the unconditional forecast. Given that the out-of-sample R^2 and a comparison of the root mean squared errors yield identical qualitative results, I focus on the out-of-sample R^2 since it is measured in comparable units to the in-sample R^2. In addition, I consider out-of-sample forecasts that are based on both the time-series and pooled estimates of the slope coefficients in the forecasting regressions. Although in the general case of non-identical slope coefficients the latter estimate only identifies the average coefficient, it is possible that it yields better out-of-sample performance if the time-series estimate is imprecise due to fewer available observations.

The results are presented in Table 7. As a comparison to the out-of-sample measures just described, the standard in-sample R^2 from the time-series regressions is provided. Again, one should note that, due to insufficient sample sizes, not all of those countries for which significant in-sample evidence was seen in Table 5, are considered in the out-of-sample analysis. As seen in Panel A of Table 7, there are five countires for which the domestic short rate has a significant negative predictive realtionship: Canada, Germany, the Netherlands, Spain, and the U.S.. With the exception of Japan and Sweden, the coefficients for the other countries have the correct sign. For the five countries in which there is a significant in-sample relationship, the out-of-sample R^2 are all positive, although typically smaller than the in-sample R^2; this holds true when using conditional forecasts that are based either on the time-series or pooled estimates. In all of the remaining countries but Australia, Japan, and Sweden, the out-of-sample R^2 is also positive for both the time-series and pooled cases. The results for Japan and Sweden are not surprising given that the corresponding full-sample estimates have the wrong sign. Thus, for 11 of the 12 countries for which the full-sample estimates have the correct sign, the conditional forecasts that are based on the time series estimates beat the unconditional forecasts. This pattern is even stronger when considering the pooled conditional forecasts; in this case, the conditional forecasts beat the unconditional in all 12 countries.

Panel B of Table 7 presents the results for the term spread. In this case, a significant in-sample predictive relationship is seen in seven countries. Again, only Japan and Sweden exhibit the wrong sign. The pattern of out-of-sample results is similar to that of the short rate, although the results for the forecasts based on the time-series estimates are somewhat weaker; in only five of the seven countries, for which there is a significant in-sample relationship, does the conditional forecast beat the unconditional forecast. The pooled results, however, are stronger in that the conditional forecast beats the unconditional one in all seven of these countries. In addition, the out-of-sample R^2 are positive for all countries but Japan, Sweden, and the U.K. when using the pooled estimates in the forecast.

Overall, the results in Table 7 provide evidence that a significant in-sample relationship is associated with out-of-sample predictive power. In general, it also seems that forecasts based on the pooled estimates out-perform forecasts based on the time-series estimates. Thus, it is possible that pooled methods could be useful in reducing risk when using conditional forecasts in portfolio choice. Lastly, this out-of-sample evidence strengthens the previous conclusions, i.e. that the interest rate variables, and in particular the short rate, are robust predictors of excess stock returns. Although not presented here, these strong results are not present when the forecast restrictions of Campbell and Thompson are not imposed.

Campbell and Thompson (2004) show how the out-of-sample R^2 can be directly related to the economic gains for a myopic log-utility investor. In particular, the absolute difference in expected returns for a myopic investor who uses the conditional forecasts and for one who relies on the unconditional forecasts is approximately equal to the out-of-sample R^2; the *proportional* increase in expected returns is approximately equal to $R^2_{i,OS}/S^2_i$ where S_i is the Sharpe ratio for market i. Since the increase in expected returns is partly due to the investor taking on more risk, it does not represent a pure welfare gain for a risk-averse investor. However, the welfare gain for the log-utility investor, from using the conditional forecast, turns out to be equivalent to the welfare gain from increasing the real risk-free interest rate by one half of the increase in absolute expected returns.

These results are shown in the last two time-series and pooled columns of Table 7. The second to last column shows the proportional increase in expected returns, calculated as $R^2_{i,OS}/S^2_i$, and is expressed as a percentage increase. The last column shows the increase in the real annual risk free rate, calculated as $12 \times R^2_{i,OS}/2$ and expressed in percentage points, that is equivalent to the welfare gain from using the conditional forecasts. Thus, a myopic log-utility investor investing in, say, the U.S. stock-market and using conditional forecasts based on the pooled estimates for the short rate would have increased her average returns by approximately 39%, or, equivalently, enjoyed the welfare gain of an increase by about three annual percentage points in the real risk-free rate. Overall, the gains from using the interest rate variables as predictors are economically significant and typically lead to substantial increases in expected returns; in most cases, the average returns increase by at least 20%, and sometimes considerably more. More importantly, however, are the risk-adjusted welfare gains, expressed as equivalent increases in the annual real risk-free interest rate. In most countries, the welfare gains for an investor investing in that country, are similar to at least a one to two percentage point increase in the annual real risk-free rate. Historically, the U.S. real risk-free interest rate, for instance, has been around 1% annually (Campbell and Thompson, 2004), so the one to two percentage point equivalent increase is rather substantial.

In summary, the results found using traditional in-sample methods appear to hold up well in an out-of-sample analysis, which also illustrates the great economic importance of these results.

9 Conclusions

The predictability of stock returns by lagged regressors is probably one of the most researched empirical topics in financial economics. Despite this interest, relatively little is known about the evidence in international data; almost all stylized 'facts' are based more or less on the U.S. experience. Ang and Bekaert (2003) are an important recent exception, and they consider a data set with four additional countries, apart from the U.S. In this paper, I consider a large global data set with data from 40 different markets and I develop new econometric methods to analyze the data. In particular, I derive several new results for long-run regressions, that use overlapping observations, when the regressors are endogenous and nearly persistent. I show how to properly correct for the overlap in the data in a simple manner that obviates the need for auto-correlation robust standard error methods in these regressions. Further, when the regressors are persistent and endogenous, I show how to correct the long-run OLS estimators and test procedures in a manner similar to that proposed by Campbell and Yogo (2003) for the short-run case.

The empirical analysis delivers three main findings: *(i)* Traditional valuation measures such as the dividend- and earnings-price ratios have very limited predictive ability in international data. It is evident, however, that using methods that do not account for the persistence and endogeneity of these variables would lead one to vastly misjudge their predictive powers. *(ii)* Interest rate variables are more robust predictors of stock returns, although their predictive power is mostly evident in OECD countries. The interest rate variables are primarily short-run predictors, with the strongest evidence found at a 1-month horizon, and the short rate tends to be a better predictor than the term spread. In addition, the U.S. interest rate variables have some predictive power, but it is not as strong as that of the domestic interest rate variables. *(iii)* The domestic short rate and term spread also have out-of-sample predictive power that is economically significant; the welfare gains to a log-utility investor who uses their predictive ability to make portfolio decisions are substantial.

The international results for the interest rate variables are similar to those of the U.S. while the overall findings for the earnings- and dividend-price ratios are substantially weaker. As in the U.S., there is some evidence that the earnings- and dividend-price ratios significantly predict stock returns in Japan and the U.K.; thus, restricting the analysis to such major economies may not give a representative picture of global stock return predictability. In summary, the results presented in this paper provide strong evidence that there is a predictable component in stock returns, which is captured at least partially by interest rate variables, and that it is large enough to be practically useful and economically meaningful for investors.

A Derivation of the long horizon results

Proof of Theorem 1. For ease of notation, the i subscript is suppressed, and the case with no intercept is treated. The results generalize immediately to regressions with fitted intercepts by replacing all variables by their demeaned versions. All limits as $q, T \to \infty$ are under the condition that $q/T \to 0$.

1. (a) Under the null hypothesis,

$$\hat{\beta}^U(q) = \left(\sum_{t=1}^{T} u_{t+q}(q)\, x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} = \left(\sum_{t=1}^{T}\sum_{j=1}^{q} u_{t+j} x_t'\right)\left(\sum_{t=1}^{T} x_t x_t'\right)^{-1}.$$

Thus,

$$\frac{T}{q}\left(\hat{\beta}^U(q) - 0\right) = \left(\frac{1}{qT}\sum_{t=1}^{T}\sum_{j=1}^{q} u_{t+j} x_t'\right)\left(\frac{1}{T^2}\sum_{t=1}^{T} x_t x_t'\right)^{-1}.$$

By standard arguments,

$$\frac{1}{qT}\sum_{t=1}^{T}\sum_{j=1}^{q} u_{t+j} x_t' = \frac{1}{qT}\sum_{t=1}^{T}\left(u_{t+1}x_t' + ... + u_{t+q}x_t'\right) \Rightarrow \int_0^1 dB_1 J_C',$$

as $q, T \to \infty$, such that $q/T \to 0$, since for any $h > 0$,

$$\frac{1}{T}\sum_{t=1}^{T} u_{t+h} x_t' \Rightarrow \int_0^1 dB_1 J_C'.$$

Therefore,

$$\frac{T}{q}\left(\hat{\beta}^U(q) - 0\right) \Rightarrow \left(\int_0^1 dB_1 J_C'\right)\left(\int_0^1 J_C J_C'\right)^{-1}.$$

(b) Under the null hypothesis,

$$\begin{aligned}
\hat{\beta}^B(q) &= \left(\sum_{t=1}^{T} u_{t+q}(q)\, x_t(q)'\right)\left(\sum_{t=1}^{T} x_t(q)\, x_t(q)'\right)^{-1} \\
&= \left(\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} u_{t+j} x_{t-q+k}'\right)\left(\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} x_{t-q+j} x_{t-q+k}'\right)^{-1}.
\end{aligned}$$

Thus

$$T\left(\hat{\beta}^B(q) - 0\right) = \left(\frac{1}{q^2 T}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} u_{t+j} x_{t-q+k}'\right)\left(\frac{1}{q^2 T^2}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} x_{t-q+j} x_{t-q+k}'\right)^{-1}.$$

By some algebraic manipulations,

$$
\begin{aligned}
& \frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} x_{t-q+j}x'_{t-q+k} \\
= \; & \frac{1}{T^2}\sum_{t=1}^{T}\left(x_{t-q+1}x'_{t-q+1} + x_{t-q+2}x'_{t-q+1} + x_{t-q+3}x'_{t-q+1} + ... + x_{t-1}x'_{t-q+1} + x_t x'_{t-q+1}\right) \\
& \vdots \\
& + \frac{1}{T^2}\sum_{t=1}^{T}\left(x_{t-q+1}x'_t + x_{t-q+2}x'_t + x_{t-q+3}x'_t + ... + x_{t-1}x'_t + x_t x'_t\right) \\
= \; & \hat{\gamma}^1_{xx}(0) + \hat{\gamma}^2_{xx}(0) + ... + \hat{\gamma}^{q-1}_{xx}(0) + \hat{\gamma}^q_{xx}(0) \\
& + \hat{\gamma}^2_{xx}(-1) + \hat{\gamma}^3_{xx}(-1) + ... + \hat{\gamma}^q_{xx}(-1) \\
& \vdots \\
& + \hat{\gamma}^q_{xx}(-(q-1)) + \hat{\gamma}^1_{xx}(q-1) \\
& \vdots \\
& + \hat{\gamma}^{q-1}_{xx}(1) + \hat{\gamma}^{q-2}_{xx}(1) + ... + \hat{\gamma}^1_{xx}(1),
\end{aligned}
$$

where

$$\hat{\gamma}^k_{xx}(h) = \frac{1}{T^2}\sum_{t=q}^{T-q+1} x_{t-q+k}x'_{t-q+k+h}.$$

Further, define

$$\gamma_{xx}(h) = \int_0^1 J_C J'_C,$$

and

$$\hat{\gamma}_{xx}(h) = \hat{\gamma}^q_{xx}(h) = \frac{1}{T^2}\sum_{t=q}^{T-q+1} x_t x'_{t+h}.$$

By standard arguments (Phillips and Solo, 1992), for any fixed k and h,

$$\hat{\gamma}^k_{xx}(h) \Rightarrow \gamma_{xx}(h) = \int_0^1 J_C J'_C,$$

since $q/T = o(1)$. Now, using Skorohod's representation theorem, there exists a probability space with random variables $\left\{\hat{\gamma}^{k*}_{xx}(h), \gamma^*_{xx}(h)\right\}$, for which

$$\hat{\gamma}^{k*}_{xx}(h) \overset{a.s.}{\to} \gamma^*_{xx}(h), \quad \hat{\gamma}^{k*}_{xx}(h) \equiv \hat{\gamma}^k_{xx}(h),$$

and

$$\int_0^1 J^*_C J^{*\prime}_C \equiv \int_0^1 J_C J'_C,$$

where '≡' denotes distributional equivalence. Since the asymptotic limit of $\hat{\gamma}_{xx}^{k*}(h)$ is identical for all k,

$$\hat{\gamma}_{xx}^{k*}(h) = \hat{\gamma}_{xx}^{*}(h) + o_{a.s.}(1),$$

and

$$\begin{aligned}
& \frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} x_{t-q+j}x'_{t-q+k} \\
\equiv\; & [q\hat{\gamma}_{xx}^{*}(0) + qo_{a.s.}(1)] + [(q-1)\hat{\gamma}_{xx}^{*}(1) + (q-1)o_{a.s.}(1)] + ... + [\hat{\gamma}_{xx}^{*}(-(q-1))] \\
& + [\hat{\gamma}_{xx}^{*}(q-1)] + ... + [(q-1)\hat{\gamma}_{xx}^{*}(1) + (q-1)o_{a.s.}(1)] \\
=\; & \sum_{h=-q+1}^{q-1}(q-|h|)\hat{\gamma}_{xx}^{*}(h) + q^2 o_{a.s.}(1).
\end{aligned}$$

Since $\hat{\gamma}_{xx}^{*}(h) \overset{a.s.}{\to} \int_0^1 J_C^* J_C^{*\prime}$ for all h, and $\frac{1}{q}\sum_{h=-q+1}^{q-1}\left(1-\frac{|h|}{q}\right) = 1$, it follows that as $q, T \to \infty$,

$$\frac{1}{q^2T^2}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} x_{t-q+j}x'_{t-q+k} \equiv \frac{1}{q}\sum_{h=-q+1}^{q-1}\left(1-\frac{|h|}{q}\right)\hat{\gamma}_{xx}^{*}(h) + o_{a.s.}(1) \overset{a.s.}{\to} \int_0^1 J_C^* J_C^{*\prime},$$

by Toeplitz's lemma and the assumption that $q/T = o(1)$. On the original probability space, therefore,

$$\frac{1}{q^2T^2}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} x_{t-q+j}x'_{t-q+k} \Rightarrow \int_0^1 J_C J_C',$$

as $q, T \to \infty$. Similarly, let

$$\gamma_{ux}(h) = \int_0^1 dB_1 J_C',$$

$$\hat{\gamma}_{ux}^{k}(h) = \frac{1}{T}\sum_{t=q}^{T-q+1} u_{t-q+k}x'_{t-q+k-h},$$

and

$$\hat{\gamma}_{ux}^{q}(h) = \hat{\gamma}_{ux}(h) = \frac{1}{T}\sum_{t=q}^{T-q+1} u_t x'_{t-h}.$$

For any fixed k and h,

$$\hat{\gamma}_{ux}^{k}(h) \Rightarrow \gamma_{ux}(h) = \int_0^1 dB_1 J_C'.$$

Using the same methods as for the denominator, as $q, T \to \infty$,

$$\frac{1}{q^2T}\sum_{t=1}^{T}\sum_{j=1}^{q}\sum_{k=1}^{q} u_{t+j}x'_{t-q+k} \Rightarrow \int_0^1 dB_1 J_C'.$$

Combining these results,

$$T\left(\hat{\beta}^{B}(q)-0\right)\Rightarrow\int_{0}^{1}dB_{1}J_{C}'\left(\int_{0}^{1}J_{C}J_{C}'\right)^{-1}.$$

2. (a) Next, consider the properties of $\hat{\beta}_{i}^{U}(q)$ under the alternative hypothesis. By summing up on both sides in equation (1),

$$\begin{aligned}
&r_{t+q}(q)=\beta\left(x_{t}+x_{t+1}+...+x_{t+q-1}\right)+u_{t+q}(q)\\
=\;&\beta\left(\left(x_{t}+Ax_{t}+...+A^{q-1}x_{t}\right)+v_{t+1}+\left(Av_{t+1}+v_{t+2}\right)+...+\sum_{p=2}^{q}A^{q-p}v_{t+p-1}\right)+u_{t+q}(q)\\
=\;&\beta\left(I+A+...+A^{q-1}\right)x_{t}+\beta\left(v_{t+1}+\left(Av_{t+1}+v_{t+2}\right)+...+\sum_{p=2}^{q}A^{q-p}v_{t+p-1}\right)+u_{t+q}(q)\\
=\;&\beta^{U}(q)\,x_{t}+\beta\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q}A^{q-p}v_{t+p-q+j}+u_{t+q}(q),
\end{aligned}$$

where

$$\beta^{U}(q)=\beta\left(I+A+...+A^{q-1}\right).$$

Thus,

$$\begin{aligned}
\hat{\beta}^{U}(q)&=\left(\sum_{t=1}^{T}r_{t+q}(q)\,x_{t}'\right)\left(\sum_{t=1}^{T}x_{t}x_{t}'\right)^{-1}\\
&=\beta^{U}(q)+\left(\beta\sum_{t=1}^{T}\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q}A^{q-p}v_{t+p-q+j}x_{t}'+\sum_{t=1}^{T}u_{t+q}(q)\,x_{t}'\right)\left(\sum_{t=1}^{T}x_{t}x_{t}'\right)^{-1}.
\end{aligned}$$

Re-arrange and standardize by T/q^{2},

$$\frac{T}{q^{2}}\left(\hat{\beta}^{U}(q)-\beta^{U}(q)\right)=\left(\beta\frac{1}{q^{2}T}\sum_{t=1}^{T}\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q}A^{q-p}v_{t+p-q+j}x_{t}'+\frac{1}{q^{2}T}\sum_{t=1}^{T}u_{t+q}(q)\,x_{t}'\right)\left(\frac{1}{T^{2}}\sum_{t=1}^{T}x_{t}x_{t}'\right)^{-1}.$$

By previous results, as $q,T\rightarrow\infty$,

$$\frac{1}{qT}\sum_{t=1}^{T}u_{t+q}(q)\,x_{t}'\Rightarrow\int_{0}^{1}dB_{1}J_{C}',$$

so that

$$\frac{1}{q^{2}T}\sum_{t=1}^{T}u_{t+q}(q)\,x_{t}'=O_{p}\left(q^{-1}\right).$$

Further,

$$\begin{aligned}
&\frac{1}{q^2T}\sum_{t=1}^{T}\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q}A^{q-p}v_{t+p-q+j}x_t' \\
= \ & \frac{1}{q^2T}\sum_{t=1}^{T}\left(v_{t+1}+\left(Av_{t+1}+v_{t+2}\right)+...+\sum_{p=3}^{q}A^{q-p}v_{t+p-2}+\sum_{p=2}^{q}A^{q-p}v_{t+p-1}\right)x_t' \\
= \ & \frac{1}{q^2T}\sum_{t=1}^{T}\left(\left(\sum_{p=2}^{q}A^{q-p}\right)v_{t+1}+...+\left(\sum_{p=q}^{q}A^{q-p}\right)v_{t+q-1}\right)x_t' \\
= \ & \frac{1}{q}\left[\left(\frac{1}{q}\sum_{p=2}^{q}A^{q-p}\right)\hat{\gamma}_{vx}\left(1\right)+...+\left(\frac{1}{q}\sum_{p=q}^{q}A^{q-p}\right)\hat{\gamma}_{vx}\left(q-1\right)\right],
\end{aligned}$$

where

$$\hat{\gamma}_{vx}\left(h\right)=\frac{1}{T}\sum_{t=1}^{T}v_tx_{t-h}'.$$

By the local-to-unity property of A,

$$\frac{1}{q}\sum_{p=h}^{q}A^{q-p}=\frac{1}{q}\sum_{p=h}^{q}\left(I+\frac{C}{T}\right)^{q-p}=\frac{1}{q}\sum_{p=h}^{q}\left(I+O\left(\frac{q}{T}\right)\right)=\frac{q-h+1}{q}I+O\left(qT^{-1}\right).$$

It follows that

$$\begin{aligned}
&\frac{1}{q^2T}\sum_{t=1}^{T}\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q}A^{q-p}v_{t+p-q+j}x_t' \\
= \ & \frac{1}{q}\left[\left(\frac{q-1}{q}+O\left(qT^{-1}\right)\right)\hat{\gamma}_{vx}\left(1\right)+...+\left(\frac{1}{q}+O\left(qT^{-1}\right)\right)\hat{\gamma}_{vx}\left(q-1\right)\right] \\
= \ & \frac{1}{q}\sum_{h=1}^{q-1}\left(1-\frac{h}{q}\right)\hat{\gamma}_{vx}\left(h\right)+O_p\left(qT^{-1}\right).
\end{aligned}$$

For any h,

$$\hat{\gamma}_{vx}\left(h\right)=\frac{1}{T}\sum_{t=1}^{T}v_tx_{t-h}'\Rightarrow\int_0^1 dB_2J_C'+\Lambda_{22}\left(h\right)=\gamma_{vx}\left(h\right)+\Lambda_{22}\left(h\right),$$

where

$$\Lambda_{22}\left(h\right)=\sum_{k=h}^{\infty}E\left(v_{i,k}v_{i,0}'\right).$$

Since $\frac{1}{q}\sum_{h=1}^{q-1}\left(1-\frac{h}{q}\right)=\frac{1}{2}$, as $q\to\infty$, and $\Lambda_{22}\left(h\right)\to 0$ as $h\to\infty$, it follows that

$$\frac{2}{q}\sum_{h=1}^{q-1}\left(1-\frac{h}{q}\right)\Lambda_{22}\left(h\right)\to\Lambda_{22}\left(1\right)=\Lambda_{22},$$

as $q \to \infty$. Thus as $q, T \to \infty$,

$$\frac{2}{q} \sum_{h=1}^{q-1} \left(1 - \frac{h}{q}\right) \hat{\gamma}_{vx}(h) \Rightarrow \int_0^1 dB_2 J_C' + \Lambda_{22}.$$

Summing up,

$$\frac{2T}{q^2} \left(\hat{\beta}^U(q) - \beta^U(q)\right) \Rightarrow \beta \left(\int_0^1 dB_2 J_C' + \Lambda_{22}\right) \left(\int_0^1 J_C J_C'\right).$$

(b) As in part *2.(a)*,

$$r_{t+q}(q) = \beta(x_t + x_{t+1} + ... + x_{t+q-1}) + u_{t+q}(q) = \beta x_{t+q-1}(q) + u_{t+q}(q).$$

Thus,

$$\hat{\beta}^B(q) = \left(\beta \sum_{t=1}^T x_{t+q-1}(q) x_t(q)' + \sum_{t=1}^T u_{t+q}(q) x_t(q)'\right) \left(\sum_{t=1}^T x_t(q) x_t(q)'\right)^{-1}.$$

Observe that

$$x_{t+q-1} = A^{q-1} x_t + \sum_{p=2}^q A^{q-p} v_{t+p-1}, ..., x_t = A^{q-1} x_{t-q+1} + \sum_{p=2}^q A^{q-p} v_{t+p-q},$$

and one can therefore write

$$\begin{aligned}
& x_{t+q-1}(q) = x_{t+q-1} + ... + x_{t+1} + x_t \\
= \; & A^{q-1} x_t + \sum_{p=2}^q A^{q-p} v_{t+p-1} + ... + A^{q-1} x_{t-q+2} + \sum_{p=2}^q A^{q-p} v_{t+p-q+1} + A^{q-1} x_{t-q+1} + \sum_{p=2}^q A^{q-p} v_{t+p-q} \\
= \; & A^{q-1}(x_t + ... + x_{t-q+2} + x_{t-q+1}) + \sum_{p=2}^q A^{q-p}(v_{t+p-1} + ... + v_{t+p-q+1} + v_{t+p-q}) \\
= \; & A^{q-1} x_t(q) + \sum_{p=2}^q A^{q-p} \sum_{h=1}^q v_{t+p-h} \\
= \; & A^{q-1} x_t(q) + \sum_{p=2}^q A^{q-p} v_{t+p-1}(q).
\end{aligned}$$

Thus,

$$\sum_{t=1}^T x_{t+q-1}(q) x_t(q)' = A^{q-1} \sum_{t=1}^T x_t(q) x_t(q)' + \sum_{t=1}^T \sum_{p=2}^q A^{q-p} v_{t+p-1}(q) x_t(q)',$$

and

$$\hat{\beta}^B(q) = \beta A^{q-1} + \left(\beta \sum_{t=1}^T \sum_{p=2}^q A^{q-p} v_{t+p-1}(q) x_t(q)' + \sum_{t=1}^T u_{t+q}(q) x_t(q)'\right) \left(\sum_{t=1}^T x_t(q) x_t(q)'\right)^{-1}.$$

Let $\beta^B(q) = \beta A^{q-1}$ and consider

$$
\begin{aligned}
& \frac{T}{q}\left(\hat{\beta}^B(q) - \beta^B(q)\right) \\
= \quad & \left(\beta \frac{1}{q^3 T}\sum_{t=1}^{T}\sum_{p=2}^{q} A^{q-p} v_{t+p-1}(q)\, x_t(q)' + \frac{1}{q^3 T}\sum_{t=1}^{T} u_{t+q}(q)\, x_t(q)'\right)\left(\frac{1}{q^2 T^2}\sum_{t=1}^{T} x_t(q)\, x_t(q)'\right)^{-1}.
\end{aligned}
$$

Since, as $q, T \to \infty$,

$$
\frac{1}{q^2 T}\sum_{t=1}^{T} u_{t+q}(q)\, x_t(q)' \Rightarrow \int_0^1 dB_1 J_C',
$$

it follows that

$$
\frac{1}{q^3 T}\sum_{t=1}^{T} u_{t+q}(q)\, x_t(q)' = o_p\left(q^{-1}\right).
$$

Next,

$$
\frac{1}{q^3 T}\sum_{t=1}^{T}\sum_{p=2}^{q} A^{q-p} v_{t+p-1}(q)\, x_t(q)' = \frac{1}{q}\sum_{p=2}^{q} A^{q-p}\frac{1}{q^2 T}\sum_{t=1}^{T} v_{t+p-1}(q)\, x_t(q)'.
$$

Again, let $\hat{\gamma}_{vx}(h) = T^{-1}\sum_{t=1}^{T} v_t x_{t-h}'$ and use the Skorohod construction, such that

$$
\hat{\gamma}_{vx}^*(h) \overset{a.s.}{\to} \gamma_{vx}^*(h) + \Lambda_{22}(h) = \int_0^1 dB_2^* J_C^{*\prime} + \Lambda_{22}(h)
$$

where

$$
\hat{\gamma}_{vx}^*(h) \equiv \hat{\gamma}_{vx}(h) \text{ and } \int_0^1 dB_2^* J_C^{*\prime} \equiv \int_0^1 dB_2 J_C'.
$$

By similar arguments as in *1.(b)*,

$$
\frac{1}{q}\sum_{p=2}^{q} A^{q-p}\frac{1}{q^2 T}\sum_{t=1}^{T} v_{t+p-1}(q)\, x_t(q)' \equiv \frac{1}{q}\sum_{p=2}^{q} A^{q-p}\frac{1}{q}\sum_{h=-q+1}^{q-1}\left(1 - \frac{|h|}{q}\right)\hat{\gamma}_{vx}^*(p-1+h) + o_{a.s.}(1).
$$

Since $\Lambda_{22}(h) \to 0$ as $h \to \infty$ and $\Lambda_{22}(h) \to \Omega_{22} = \sum_{k=-\infty}^{\infty} E\left(v_{i,k} v_{i,0}'\right)$ as $h \to -\infty$, it follows that

$$
\frac{1}{q}\sum_{h=-q+1}^{q-1}\left(1 - \frac{|h|}{q}\right)\Lambda_{22}(p-1+h) \to \Omega_{22}
$$

as $q \to \infty$, for any p, and thus

$$
\frac{1}{q}\sum_{p=2}^{q} A^{q-p}\frac{1}{q}\sum_{h=-q+1}^{q-1}\left(1 - \frac{|h|}{q}\right)\hat{\gamma}_{vx}^*(p-1+h) + o_{a.s.}(1) \overset{a.s.}{\to} \int_0^1 dB_2^* J_C^{*\prime} + \Omega_{22}.
$$

On the original probability space

$$\frac{1}{q}\sum_{p=2}^{q}A^{q-p}\frac{1}{q^2T}\sum_{t=1}^{T}v_{t+p-1}\left(q\right)x_t\left(q\right)'\Rightarrow\int_0^1 dB_2J_C'+\Omega_{22},$$

and

$$\frac{T}{q}\left(\hat{\beta}^B\left(q\right)-\beta^B\left(q\right)\right)\Rightarrow\beta\left(\int_0^1 dB_2J_C'+\Omega_{22}\right)\left(\int_0^1 J_CJ_C'\right)^{-1}.$$

■

Proof of Theorem 2. *1. (a)* Let $\mathbf{r}_{+q}^q=\left(r_{1+q}\left(q\right),...,r_T\left(q\right)\right)'$ be the $T\times 1$ vector of observations, and define $\mathbf{x}^q$ and $\mathbf{v}_{+q}^q$ analogously. The OLS estimator of $\beta^B\left(q\right)$ in (21) is now given by

$$\hat{\beta}^{B+}\left(q\right)=\left(\mathbf{r}_{+q}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(\mathbf{x}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1}.$$

where $Q_{\mathbf{v}^q}=I-\mathbf{v}_{+q}^q\left(\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q\right)^{-1}\mathbf{v}_{+q}^{q\prime}$. Under the null-hypothesis,

$$Q_{\mathbf{v}^q}\mathbf{r}_{+q}^q=Q_{\mathbf{v}^q}\left(\mathbf{v}_{+q}^q\gamma'+\left(\mathbf{u}_{+q}^q-\mathbf{v}_{+q}^q\gamma'\right)\right)=Q_{\mathbf{v}^q}\mathbf{u}_{+q}^q$$

and

$$T\left(\hat{\beta}^{B+}\left(q\right)-0\right)=\left(q^{-2}T^{-1}\mathbf{u}_{+q}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(q^{-2}T^{-2}\mathbf{x}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1}.$$

As $q,T\to\infty$,

$$(qT)^{-2}\mathbf{x}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q=(qT)^{-2}\mathbf{x}^{q\prime}\mathbf{x}^q-qT^{-1}\left(q^{-2}T^{-1}\mathbf{x}^{q\prime}\mathbf{v}_{+q}^q\right)\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q\right)^{-1}\left(q^{-2}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{x}^q\right)\Rightarrow\int_0^1 J_CJ_C',$$

since $q/T\to 0$. Next,

$$\begin{aligned}Q_{\mathbf{v}_i^q}\mathbf{u}_{+q}^q &= \mathbf{u}_{+q}^q-\mathbf{v}_{+q}^q\left(\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q\right)^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{u}_{+q}^q\\ &= \mathbf{u}_{+q}^q-\mathbf{v}_{+q}^q\left(\frac{1}{qT}\sum_{t=1}^{T}v_{t+q}\left(q\right)v_{t+q}\left(q\right)'\right)^{-1}\left(\frac{1}{qT}\sum_{t=1}^{T}u_{t+q}\left(q\right)v_{t+q}\left(q\right)'\right).\end{aligned}$$

Let

$$\hat{\gamma}_{vv}\left(h\right)=\frac{1}{T}\sum_{t=q}^{T-q+1}v_tv_{t-h}',\text{ and }\hat{\gamma}_{uv}\left(h\right)=\frac{1}{T}\sum_{t=q}^{T-q+1}u_tv_{t-h}',$$

and define $\hat{\gamma}_{vv}^{k}(h)$ and $\hat{\gamma}_{uv}^{k}(h)$ in an analogous manner to the previous proof. As above,

$$\begin{aligned}
& \frac{1}{T}\sum_{t=1}^{T} v_{t+q}(q)\, v_{t+q}(q)' \\
= \; & \hat{\gamma}_{vv}^{1}(0) + \hat{\gamma}_{vv}^{2}(0) + \ldots + \hat{\gamma}_{vv}^{q-1}(0) + \hat{\gamma}_{vv}^{q}(0) \\
& + \hat{\gamma}_{vv}^{2}(-1) + \hat{\gamma}_{vv}^{3}(-1) + \ldots + \hat{\gamma}_{vv}^{q}(-1) \\
& \vdots \\
& + \hat{\gamma}_{vv}^{q}(-(q-1)) + \hat{\gamma}_{vv}^{1}(q-1) \\
& \vdots \\
& + \hat{\gamma}_{vv}^{q-1}(1) + \hat{\gamma}_{vv}^{q-2}(1) + \ldots + \hat{\gamma}_{vv}^{1}(1),
\end{aligned}$$

and it follows that as $q, T \to \infty$,

$$\frac{1}{qT}\sum_{t=1}^{T} v_{t+q}(q)\, v_{t+q}(q)' = \sum_{h=-q+1}^{q-1}\left(1 - \frac{|h|}{q}\right)\hat{\gamma}_{vv}(h) + O_p\left(qT^{-1/2}\right) \to_p \Omega_{22},$$

by standard results (e.g. Andrews, 1991), since $qT^{-1/2} = o(1)$. By identical arguments, as $q, T \to \infty$,

$$\frac{1}{qT}\sum_{t=1}^{T} u_{t+q}(q)\, v_{t+q}(q)' = \sum_{h=-q+1}^{q-1}\left(1 - \frac{|h|}{q}\right)\hat{\gamma}_{uv}(h) + O_p\left(qT^{-1/2}\right) \to_p \omega_{12}.$$

Again, using the same methods as in the proof of part *1.(b)* in Theorem 1, it follows that

$$\frac{1}{q^2 T}\mathbf{v}_{+q}^{q\prime}\mathbf{x}^{q} = \frac{1}{q^2 T}\sum_{t=1}^{T} v_{t+q}(q)\, x_t(q)' \Rightarrow \int_0^1 dB_2 J_C'.$$

Thus,

$$\begin{aligned}
q^{-2}T^{-1}\mathbf{u}_{+q}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q &= q^{-2}T^{-1}\mathbf{u}_{+q}^{q\prime}\mathbf{x}^q - \left(q^{-1}T^{-1}\mathbf{u}_{+q}^{q\prime}\mathbf{v}_{+q}^{q}\right)\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^{q}\right)^{-1}\left(q^{-2}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{x}^q\right) \\
&\Rightarrow \int_0^1 dB_1 J_C' - \omega_{12}\Omega_{22}^{-1}\int_0^1 dB_2 J_C' \\
&= \int_0^1 dB_{1\cdot 2} J_C',
\end{aligned}$$

and

$$T\left(\hat{\beta}^{B+}(q) - 0\right) = \left(q^{-2}T^{-1}\mathbf{u}_{+q}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(q^{-2}T^{-2}\mathbf{x}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1} \Rightarrow \int_0^1 dB_{1\cdot 2}J_C'\left(\int_0^1 J_C J_C'\right)^{-1}.$$

1. (b) Denote $\mathbf{x}^1 = \mathbf{x}$. The OLS estimator of $\beta_i^U(q)$ in (20) is now given by

$$\hat{\beta}^{U+}(q) = \left(\mathbf{r}_{+q}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}\right)\left(\mathbf{x}' Q_{\mathbf{v}^q}\mathbf{x}\right)^{-1}.$$

Under the null hypothesis,

$$Q_{\mathbf{v}^q}\mathbf{r}_{+q}^q = Q_{\mathbf{v}^q}\left(\mathbf{v}_{+q}^q\gamma' + \left(\mathbf{u}_{+q}^q - \mathbf{v}_{+q}^q\gamma'\right)\right) = Q_{\mathbf{v}^q}\mathbf{u}_{+q}^q,$$

and thus

$$\frac{T}{q}\left(\hat{\beta}^{U+}(q) - 0\right) = \left(q^{-1}T^{-1}\mathbf{u}_{+q}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}\right)\left(T^{-2}\mathbf{x}' Q_{\mathbf{v}^q}\mathbf{x}\right)^{-1}.$$

As $T \to \infty$,

$$T^{-2}\mathbf{x}' Q_{\mathbf{v}^q}\mathbf{x} = T^{-2}\mathbf{x}'\mathbf{x} - qT^{-1}\left(q^{-1}T^{-1}\mathbf{x}'\mathbf{v}_{+q}^q\right)\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q\right)^{-1}\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{x}\right) \Rightarrow \int_0^1 J_C J_C',$$

since $q/T \to 0$. The desired result now follows by the same arguments as in part *1.(a)*.

2.(a) Observe that under the alternative hypothesis,

$$Q_{\mathbf{v}^q}\mathbf{r}_{+q}^q = Q_{\mathbf{v}^q}\left(\mathbf{x}_{+q-1}^q\beta' + \mathbf{v}_{+q}^q\gamma' + \left(\mathbf{u}_{+q}^q - \mathbf{v}_{+q}^q\gamma'\right)\right) = Q_{\mathbf{v}^q}\mathbf{x}_{+q-1}^q\beta' + Q_{\mathbf{v}^q}\mathbf{u}_{+q}^q$$

and

$$\hat{\beta}^{B+}(q) = \beta\left(\mathbf{x}_{+q-1}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(\mathbf{x}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1} + \left(\mathbf{u}_{+q}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(\mathbf{x}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1}$$

Recall that

$$x_{t+q-1} = A^{q-1}x_t(q) + \sum_{p=2}^{q} A^{q-p}v_{t+p-1}(q).$$

Thus,

$$\mathbf{x}_{+q-1}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q = \left(A^{q-1}\mathbf{x}^{q\prime} + \sum_{p=2}^{q} A^{q-p}\mathbf{v}_{+p-1}^{q\prime}\right) Q_{\mathbf{v}^q}\mathbf{x}^q = A^{q-1}\mathbf{x}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q + \sum_{p=2}^{q} A^{q-p}\mathbf{v}_{+p-1}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q,$$

and

$$\hat{\beta}^{B+}(q) = \beta A^{q-1} + \beta\sum_{p=2}^{q} A^{q-p}\left(\mathbf{v}_{+p-1}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(\mathbf{x}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1} + \left(\mathbf{u}_{+q}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)\left(\mathbf{x}^{q\prime} Q_{\mathbf{v}^q}\mathbf{x}^q\right)^{-1}.$$

Define $\beta^{B}(q) = \beta A^{q-1}$ and write

$$\begin{aligned}
\frac{T}{q}\left(\hat{\beta}^{B+}(q) - \beta^{B}(q)\right) &= \beta \frac{1}{q}\sum_{p=2}^{q} A^{q-p}\left(q^{-2}T^{-1}\mathbf{v}_{+p-1}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q}\right)\left(q^{-2}T^{-2}\mathbf{x}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q}\right)^{-1} \\
&\quad + q^{-1}\left(q^{-2}T^{-1}\mathbf{u}_{+q}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q}\right)\left(q^{-2}T^{-2}\mathbf{x}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q}\right)^{-1} \\
&= \beta \frac{1}{q}\sum_{p=2}^{q} A^{q-p}\left(q^{-2}T^{-1}\mathbf{v}_{+p-1}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q}\right)\left(q^{-2}T^{-2}\mathbf{x}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q}\right)^{-1} + o_{p}\left(q^{-1}\right).
\end{aligned}$$

Consider

$$q^{-2}T^{-1}\mathbf{v}_{+p-1}^{q\prime}Q_{\mathbf{v}^{q}}\mathbf{x}^{q} = q^{-1}T^{-1}\left[\left(q^{-1}\mathbf{v}_{+p-1}^{q\prime} - \left(q^{-1}T^{-1}\mathbf{v}_{+p-1}^{q\prime}\mathbf{v}_{+q}^{q}\right)\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^{q}\right)^{-1}q^{-1}\mathbf{v}_{+q}^{q\prime}\right)\right]\mathbf{x}^{q}.$$

By standard arguments, as $q, T \to \infty$,

$$q^{-1}T^{-1}\mathbf{v}_{+p-1}^{q\prime}\mathbf{v}_{+q}^{q} \to_{p} \Omega_{22} \text{ and } q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^{q} \to_{p} \Omega_{22}.$$

Observe that,

$$v_{t+p-1}(q) - v_{t+q}(q) = \sum_{h=0}^{q-p} v_{t-h} - \sum_{h=0}^{q-p} v_{t+q-h},$$

and

$$\begin{aligned}
&\frac{1}{q^{3}T}\sum_{t=1}^{T}\sum_{p=2}^{q} A^{q-p}\left(\sum_{h=0}^{q-p} v_{t-h} - \sum_{h=0}^{q-p} v_{t+q-h}\right)x_{t}(q)' + o_{p}(1) \\
=\;& \frac{1}{q^{3}T}\sum_{t=1}^{T}\sum_{p=2}^{q}\left(\sum_{h=0}^{q-p} v_{t-h} - \sum_{h=0}^{q-p} v_{t+q-h}\right)x_{t}(q)' + O_{p}\left(T^{-1}\right) \\
=\;& \frac{1}{q^{3}T}\sum_{t=1}^{T}\left(\sum_{p=2}^{q}\sum_{h=0}^{q-p} v_{t-h} - \sum_{p=2}^{q}\sum_{h=0}^{q-p} v_{t+q-h}\right)x_{t}(q)' + O_{p}\left(T^{-1}\right) \\
=\;& \frac{1}{q^{2}T}\sum_{t=1}^{T}\left(\sum_{h=0}^{q-2}\left(1 - \frac{h}{q}\right)v_{t-h}x_{t}(q)' - \sum_{h=0}^{q-2}\left(1 - \frac{h}{q}\right)v_{t+q-h}x_{t}(q)'\right) + O_{p}\left(T^{-1}\right) \\
\Rightarrow\;& \frac{1}{2}\int_{0}^{1} dB_{2}J_{C}' + \frac{1}{2}\int_{0}^{1} dB_{2}J_{C}' \\
=\;& \int_{0}^{1} dB_{2}J_{C}',
\end{aligned}$$

as $q, T \to \infty$. Thus,

$$\frac{T}{q}\left(\hat{\beta}^{B+}(q) - \beta^{B}(q)\right) \Rightarrow \left(\int_{0}^{1} dB_{2}J_{C}\right)\left(\int_{0}^{1} J_{C}J_{C}'\right)^{-1}.$$

2. (b) Denote $\mathbf{x}^1 = \mathbf{x}$. The OLS estimator of $\beta_i^U(q)$ in (20) is then given by

$$\hat{\beta}^{U+}(q) = \left(\mathbf{r}_{+q}^{q\prime} Q_{\mathbf{v}^q} \mathbf{x}\right)\left(\mathbf{x}' Q_{\mathbf{v}^q} \mathbf{x}\right)^{-1}.$$

Observe that, by the arguments above,

$$\mathbf{r}_{+q}^{q\prime} = \beta^U(q)\,\mathbf{x}' + \beta \sum_{j=1}^{q-1} \sum_{p=q-j+1}^{q} A^{q-p} \mathbf{v}'_{+p-q+j} + \mathbf{u}_{+q}^{q\prime}.$$

Thus,

$$\begin{aligned}
\frac{T}{q^2}\left(\hat{\beta}^{U+}(q) - \beta^U(q)\right) &= \beta\left[\frac{1}{q}\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q} A^{q-p}\left(q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}Q_{\mathbf{v}^q}\mathbf{x}\right)\right]\left(T^{-2}\mathbf{x}'Q_{\mathbf{v}^q}\mathbf{x}\right)^{-1} \\
&\quad + q^{-1}\left(q^{-1}T^{-1}\mathbf{u}_{+q}^{q\prime}Q_{\mathbf{v}^q}\mathbf{x}\right)\left(T^{-2}\mathbf{x}'Q_{\mathbf{v}^q}\mathbf{x}\right)^{-1} \\
&= \beta\left[\frac{1}{q}\sum_{j=1}^{q-1}\sum_{p=q-j+1}^{q} A^{q-p}\left(q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}Q_{\mathbf{v}^q}\mathbf{x}\right)\right]\left(T^{-2}\mathbf{x}'Q_{\mathbf{v}^q}\mathbf{x}\right)^{-1} + o_p\left(q^{-1}\right).
\end{aligned}$$

Consider

$$q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}Q_{\mathbf{v}^q}\mathbf{x} = q^{-1}T^{-1}\left(\mathbf{v}'_{+p-q+j} - \left(q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}\mathbf{v}_{+q}^q\right)\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q\right)^{-1}\mathbf{v}_{+q}^{q\prime}\right)\mathbf{x}.$$

Observe that under the martingale difference assumption,

$$q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}\mathbf{v}_{+q}^q = \frac{1}{qT}\sum_{t=1}^{T} v_{t+q}(q)\,v'_{t+p-q+j} = \frac{1}{qT}\sum_{t=1}^{T}\left(v_{t+1} + v_{t+2} + ... + v_{t+q}\right)v'_{t+p-q+j} = O_p\left(q^{-1}\right).$$

Further,

$$q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q = \frac{1}{qT}\sum_{t=1}^{T} v_{t+q}(q)\,v_{t+q}(q)' \to_p \Omega_{22},$$

and

$$q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{x} \Rightarrow \int_0^1 dB_2 J_C',$$

as $q, T \to \infty$. Thus,

$$\begin{aligned}
q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}Q_{\mathbf{v}^q}\mathbf{x} &= q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}\mathbf{x} - \left(q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}\mathbf{v}_{+q}^q\right)\left(q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{v}_{+q}^q\right)^{-1}q^{-1}T^{-1}\mathbf{v}_{+q}^{q\prime}\mathbf{x} \\
&= q^{-1}T^{-1}\mathbf{v}'_{+p-q+j}\mathbf{x} + O_p\left(q^{-1}\right).
\end{aligned}$$

The result now follows in the same manner as above, noting that $\Lambda_{22} = 0$ under the martingale assumption. ■

Proof of Corollary 1. Observe that under the null hypothesis,

$$\begin{aligned}
\hat{\sigma}^U_{11\cdot 2} &= \frac{1}{qT}\sum_{t=1}^{T}\hat{u}^{U+}_t(q)^2 = \frac{1}{qT}\sum_{t=1}^{T}\left(u_{t+q\cdot 2}(q) + O_p\left(\frac{q}{T}\right)\right)^2 \\
&= \frac{1}{qT}\sum_{t=1}^{T}u_{t+q\cdot 2}(q)^2 + \frac{1}{qT}\sum_{t=1}^{T}O_p\left(\frac{q}{T}\right)^2 + 2O_p\left(\frac{q}{T}\right)\frac{1}{qT}\sum_{t=1}^{T}u_{t+q\cdot 2}(q) \\
&= \frac{1}{qT}\sum_{t=1}^{T}\left(u_{t+q}(q) - \omega_{12}\Omega_{22}^{-1}v_{t+q}(q)\right)^2 + O_p\left(\frac{q}{T}\right) \\
&\to {}_p\sigma_{11\cdot 2}.
\end{aligned}$$

Identical arguments hold for $\hat{\sigma}^B_{11\cdot 2}$. The following $t-$statistics can now be formed

$$\frac{\frac{T}{q}\left(\hat{\beta}^{U+}(q) - \beta^{U,0}_{i,k}(q)\right)}{\sqrt{\hat{\sigma}^U_{11\cdot 2}T^{-2}a'\left(\sum_{t=1}^{T}\underline{z}_t\underline{z}_t'\right)^{-1}a}} = \frac{\hat{\beta}^{U+}(q) - \beta^{U,0}_{i,k}(q)}{\sqrt{q^2\hat{\sigma}^U_{11\cdot 2}a'\left(\sum_{t=1}^{T}\underline{z}_t\underline{z}_t'\right)^{-1}a}} = t^{U+}(q)/\sqrt{q},$$

and

$$\frac{T\left(\hat{\beta}^{B+}(q) - \beta^{B,0}_{i,k}(q)\right)}{\sqrt{\hat{\sigma}^B_{11\cdot 2}(qT)^{-1}a'\left(\sum_{t=1}^{T}\underline{z}_t(q)\underline{z}_t(q)'\right)^{-1}a}} = \frac{\hat{\beta}^{B+}(q) - \beta^{B,0}_{i,k}(q)}{\sqrt{q^2\hat{\sigma}^B_{11\cdot 2}a'\left(\sum_{t=1}^{T}\underline{z}_t(q)\underline{z}_t(q)'\right)^{-1}a}} = t^{B+}(q)/\sqrt{q},$$

where $t^{U+}(q)$ and $t^{B+}(q)$ are the standard $t-$statistics corresponding to $\hat{\beta}^{U+}(q)$ and $\hat{\beta}^{B+}(q)$. The results now follow directly from the asymptotically mixed normal distributions of the estimators. ■

Proof of Theorem 3. Let $\underline{\mathbf{r}}^{q\prime}_{i,+q}$, $\underline{\mathbf{x}}^q_i$, and $\underline{\mathbf{x}}_i$ denote the time-series demeaned data. Define,

$$\hat{\beta}^{B+}_{n,T}(q) = \left(\sum_{i=1}^{n}\underline{\mathbf{r}}^{q\prime}_{i,+q}Q_{\mathbf{v}^q_i}\underline{\mathbf{x}}^q_i\right)\left(\sum_{i=1}^{n}\underline{\mathbf{x}}^{q\prime}_iQ_{\mathbf{v}^q_i}\underline{\mathbf{x}}^q_i\right)^{-1}, \tag{28}$$

and

$$\hat{\beta}^{U+}_{n,T}(q) = \left(\sum_{i=1}^{n}\underline{\mathbf{r}}^{q\prime}_{i,+q}Q_{\mathbf{v}^q_i}\underline{\mathbf{x}}_i\right)\left(\sum_{i=1}^{n}\underline{\mathbf{x}}_i'Q_{\mathbf{v}^q_i}\underline{\mathbf{x}}_i\right)^{-1}. \tag{29}$$

Start with $\hat{\beta}^{B+}_{n,T}(q)$. By the derivations in the proof of Theorem 2,

$$\begin{aligned}
&\sqrt{n}T\left(\hat{\beta}^{B+}_{n,T}(q)-0\right)\\
=&\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}q^{-2}T^{-1}\mathbf{u}^{q\prime}_{i,+q}Q_{\mathbf{v}^q_i}\underline{\mathbf{x}}^q_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}q^{-2}T^{-2}\underline{\mathbf{x}}^{q\prime}_iQ_{\mathbf{v}^q_i}\underline{\mathbf{x}}^q_i\right)^{-1}\\
\Rightarrow&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}'_{i,C_i}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\int_0^1\underline{J}_{i,C_i}\underline{J}'_{i,C_i}\right)^{-1}\\
\Rightarrow&N\left(0,\underline{\Omega}^{-1}_{xx}\underline{\Phi}_{u\cdot v,x}\underline{\Omega}^{-1}_{xx}\right),
\end{aligned}$$

in sequential limits as $(T,n\to\infty)_{\text{seq}}$, where

$$\underline{\Omega}_{xx}=E\left[\int_0^1\underline{J}_{i,C_i}\underline{J}'_{i,C_i}\right]$$

and

$$\underline{\Phi}_{u\cdot v,x}=E\left[\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}'_{i,C_i}\right)\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}'_{i,C_i}\right)'\right].$$

Similarly, as $(T,n\to\infty)_{\text{seq}}$,

$$\begin{aligned}
&\sqrt{n}\frac{T}{q}\left(\hat{\beta}^{U+}_{n,T}(q)-0\right)\\
=&\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}q^{-1}T^{-1}\mathbf{u}^{q\prime}_{i,+q}Q_{\mathbf{v}^q_i}\underline{\mathbf{x}}_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}T^{-2}\underline{\mathbf{x}}'_iQ_{\mathbf{v}^q_i}\underline{\mathbf{x}}_i\right)^{-1}\\
\Rightarrow&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}'_{i,C_i}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\int_0^1\underline{J}_{i,C_i}\underline{J}'_{i,C_i}\right)^{-1}\\
\Rightarrow&N\left(0,\underline{\Omega}^{-1}_{xx}\underline{\Phi}_{u\cdot v,x}\underline{\Omega}^{-1}_{xx}\right).
\end{aligned}$$

■

Proof of Corollary 2. Let $\hat{u}^U_{i,t}(q)$ and $\hat{u}^B_{i,T}(q)$ be the estimated residuals. Form the $t-$statistic,

$$t^{B+}_{n,T}(q)=\frac{\sqrt{n}T\left(\hat{\beta}^{B+}_{n,T}(q)-0\right)}{\sqrt{a'\left(\hat{\underline{\Omega}}^B_{xx}\right)^{-1}\hat{\underline{\Phi}}^B_{u\cdot v,x}\left(\hat{\underline{\Omega}}^B_{xx}\right)^{-1}a}}$$

$$\begin{aligned}
\hat{\underline{\Phi}}^B_{u\cdot v,x}&=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{q^4T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\hat{u}^B_{i,t+q}(q)\,\underline{z}_{i,t}(q)'\right)'\left(\hat{u}^B_{i,s+q}(q)\,\underline{z}_{i,s}(q)'\right)\\
&\to_p E\left[\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}'_{i,C_i}\right)\left(\int_0^1 dB_{i,1\cdot 2}\underline{J}'_{i,C_i}\right)'\right]
\end{aligned}$$

and

$$\underline{\hat{\Omega}}_{xx}^{B}=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{q^{2}T^{2}}\sum_{t=1}^{T}\underline{z}_{i,t}\left(q\right)\underline{z}_{i,t}\left(q\right)'\rightarrow_{p}E\left[\int_{0}^{1}\underline{J}_{i,C_{i}}\underline{J}_{i,C_{i}}'\right].$$

It follows that,

$$t_{n,T}^{B+}\left(q\right)=\frac{\sqrt{n}T\left(\hat{\beta}_{n,T}^{B+}\left(q\right)-0\right)}{\sqrt{a'\left(\underline{\hat{\Omega}}_{xx}^{B}\right)^{-1}\underline{\hat{\Phi}}_{u\cdot v,x}^{B}\left(\underline{\hat{\Omega}}_{xx}^{B}\right)^{-1}a}}=\frac{\sqrt{n}T\left(\hat{\beta}_{n,T}^{B+}\left(q\right)-0\right)}{\sqrt{a'\underline{\hat{\Omega}}_{xx}^{-1}\underline{\hat{\Phi}}_{u\cdot v,x}\underline{\hat{\Omega}}_{xx}^{-1}a}},$$

where $\underline{\hat{\Omega}}_{xx}^{-1}$ and $\underline{\hat{\Phi}}_{u\cdot v,x}$ are the standard estimates of $\underline{\Omega}_{xx}^{-1}$ and $\underline{\Phi}_{u\cdot v,x}$, respectively, formed as above but without any standardization by q. The $t_{n,T}^{U+}\left(q\right)$ is formed analogously and the result follows in an identical manner. Observe that these pooled $t-$statistics are standard heteroskedasticity and autocorrelation robust $t-$tests as described in Baltagi (1995). ■

References

Ang, A., and G. Bekaert, 2003. Stock Return Predictability: Is it There? Working paper, Columbia University.

Bai, J., and S. Ng, 2002. Determining the Number of Factors in Approximate Factor Models, Econometrica 70, 191-221.

Baltagi, B.H., 1995. Econometric Analysis of Panel Data, Wiley, Chichester, England.

Britten-Jones, M., and A. Neuberger, 2004. Improved inference and estimation in regression with overlapping observations, Working paper, London Business School.

Campbell, J.Y., 2001. Why long horizons? A study of power against persistent alternatives, Journal of Empirical Finance 8, 459-491.

Campbell, J.Y., 2003. Consumption-Based Asset Pricing, in: Constantinides, G.M., Harris M., and Stulz R. eds., *Handbook of the Economics of Finance*, Vol. 1B (North-Holland, Amsterdam) 803-888.

Campbell, J.Y., and R. Shiller, 1988. Stock prices, earnings, and expected dividends, Journal of Finance 43, 661-676.

Campbell., J.Y., and S.B. Thompson, 2004. Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?, Working Paper, Harvard University.

Campbell, J.Y., and M. Yogo, 2003. Efficient Tests of Stock Return Predictability, Working Paper, Harvard University.

Cavanagh, C., G. Ellliot, and J. Stock, 1995. Inference in models with nearly integrated regressors, Econometric Theory 11, 1131-1147.

Daniel, K., 2001. The power and size of mean reversion tests, Journal of Empirical Finance 8, 493-535.

Fama, E.F., and K.R. French, 1988a. Dividend yields and expected stock returns, Journal of Financial Economics 22, 3-25.

Fama, E.F., and K.R. French, 1988b. Permanent and Temporary Components of Stock Prices, Journal of Political Economy, 96, 256-73.

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

Ferson, W.E., and C.R. Harvey, 1993. The Risk and Predictability of International Equity Returns, Review of Financial Studies 6, 527-566.

Ferson, W.E., S. Sarkissian, and T.T. Simin, 2003. Spurious Regressions in Financial Economics? Journal of Finance 58, 1393-1413.

Goetzman W.N., and P. Jorion, 1993. Testing the Predictive Power of Dividend Yields, Journal of Finance 48, 663-679.

Goyal A., and I. Welch, 2003a. A Note on "Predicting Returns with Financial Ratios", Yale ICF Working Paper No. 04-02.

Goyal A., and I. Welch, 2003b. Predicting the Equity Premium with Dividend Ratios, Management Science 49, 639-654.

Goyal A., and I. Welch, 2004. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, Yale ICF Working Paper No. 04-11.

Hansen, L.P., and R.J. Hodrick, 1980. Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis, Journal of Political Economy 88, 829-853.

Hansen, C.S., and B.E. Tuypens, 2004. Predicting the S&P 500 - Simple Efficient Methods, Working Paper, Zicklin School of Business, Baruch College.

Harvey C.R., 1991. The World Price of Covariance Risk, Journal of Finance 46, 111-157.

Harvey C.R., 1995. Predictable Risk and Returns in Emerging Markets, Review of Financial Studies 8, 773-816.

Hodrick, R.J., 1992. Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement, Review of Financial Studies 5, 357-386.

Hjalmarsson, E., 2004. Predictive regressions with panel data, mimeo, Yale University.

Jansson, M., and M.J. Moreira, 2004. Optimal Inference in Regression Models with Nearly Integrated Regressors, Working Paper, Harvard University.

Kaminsky G.L., and S.L. Schmukler, 2002. Short-Run Pain, Long-Run Gain: The Effects of Financial Liberalization, Working Paper, George Washington University.

Lanne, M., 2002. Testing the Predictability of Stock Returns, Review of Economics and Statistics 84, 407-415.

Lewellen, J., 2003. Predicting returns with financial ratios, Journal of Financial Economics, forthcoming.

Mankiw, N.G., and M.D. Shapiro, 1986. Do we reject too often? Small sample properties of tests of rational expectations models, Economic Letters 20, 139-145.

Mark, N.C., and D. Sul, 2004. The Use of Predictive Regressions at Alternative Horizons in Finance and Economics, NBER Technical Working Paper 298.

Moon, R., A. Rubia, and R. Valkanov, 2004. Long-Horizon Regressions when the Predictor is Slowly Varying, Working Paper, UCLA, Anderson School of Management.

Nelson, C.R., and M.J. Kim, 1993. Predictable Stock Returns: The Role of Small Sample Bias, Journal of Finance 48, 641-661.

Newey, W., and K. West, 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55, 703-708.

Phillips, P.C.B, 1987. Towards a Unified Asymptotic Theory of Autoregression, Biometrika 74, 535-547.

Phillips, P.C.B, 1988. Regression Theory for Near-Integrated Time Series, Econometrica 56, 1021-1043.

Phillips, P.C.B, 1991. Optimal Inference in Cointegrated Systems, Econometrica 59, 283-306.

Phillips, P.C.B, 1995. Fully Modified Least Squares and Vector Autoregression, Econometrica 63, 1023-1078.

Phillips, P.C.B, and B. Hansen, 1990. Statistical Inference in Instrumental Variables Regression with I(1) Processes, Review of Economic Studies 57, 99-125.

Phillips, P.C.B., and H.R. Moon, 1999. Linear Regression Limit Theory for Nonstationary Panel Data, Econometrica 67, 1057-1111.

Phillips, P.C.B., and V. Solo, 1992. Asymptotics for Linear Processes, Annals of Statistics, 20, 971-1001.

Polk, C., S. Thompson, and T. Vuolteenaho, 2004. Cross-sectional Forecasts of the Equity Premium, Working Paper, Harvard University.

Richardson, M., and T. Smith, 1991. Tests of Financial Models in the Presence of Overlapping Observations, Review of Financial Studies 4, 227-254.

Richardson, M., and J.H. Stock, 1989. Drawing Inferences from Statistics Based on Multiyear Asset Returns, Journal of Financial Economics 25, 323-348.

Stambaugh, R., 1999. Predictive regressions, Journal of Financial Economics 54, 375-421.

Stock, J.H., 1991. Confidence intervals for the largest autoregressive root in U.S. economic time-series. Journal of Monetary Economics 28, 435-460.

Stock. J.H., and M.W. Watson, 2000. Macroeconomic Forecasting Using Diffusion Indexes, Working Paper, Princeton University.

Torous, W., R. Valkanov, and S. Yan, 2005. On Predicting Stock Returns with Nearly Integrated Explanatory Variables, Journal of Business 78, forthcoming.

Valkanov, R., 2003. Long-horizon regressions: theoretical results and applications, Journal of Financial Economics 68, 201-232.

Table 1: Data summary. The first two columns give the country and the longest available sample for that country. The O column indicates whether a country is included in the OECD panels. The remaining columns give summary statistics of the data. The first two of these show the mean and standard deviation of the excess log-returns over the domestic short interest rate (r^e). The last four give the sample means of the logged dividend price ratio ($d-p$), earnings price ratio ($e-p$), short interest rate (r_s), and term spread ($y-r_s$). The summary statistics are calculated using the longest available samples for all variables except the short interest rate and the term spread, for which only data after 1952 is used, since that is the sample period considered in the forecasting regressions using these variables.

Country	Sample	O	$\mu(r^e)$	$\sigma(r^e)$	$\mu(d-p)$	$\mu(e-p)$	$\mu(r_s)$	$\mu(y-r_s)$
Argentina	1988.3 – 2004.6		40.250	71.605	-3.913	-2.807	30.333	7.519
Australia	1882.12 – 2004.6	O	7.292	13.292	-2.814	-2.723	6.574	1.033
Austria	1970.2 – 2004.6	O	1.422	18.531	-3.751	-3.428	6.466	0.682
Belgium	1952.1 – 2004.6	O	2.559	14.813	-3.304	-2.604	6.572	0.452
Brazil	1988.3 – 2004.5		15.681	56.361	-3.654	-2.706	111.820	-9.436
Canada	1934.3 – 2004.6	O	4.737	15.207	-3.319	-2.852	6.169	1.033
Chile	1983.3 – 2004.6		11.835	23.829	-3.124	-2.616	16.083	-
Denmark	1970.2 – 2004.6	O	2.079	17.631	-3.639	-2.721	10.235	-0.074
Finland	1962.3 – 2004.6	O	6.344	21.705	-3.233	-3.257	8.410	0.209
France	1898.1 – 1914.7	O	2.342	8.081	-3.246	-	-	-
	1919.2 – 1940.3	O	4.880	19.935	-3.302	-	-	-
	1941.5 – 2004.6	O	6.412	20.314	-3.469	-2.822	6.638	0.757
Germany	1872.9 – 1942.3	O	-5.310	38.852	-3.033	-	-	-
	1953.2 – 2004.6	O	4.975	17.329	-3.425	-2.733	4.544	1.963
Greece	1977.3 – 2004.6	O	4.945	30.394	-2.919	-2.630	12.725	0.538
Hong Kong	1973.1 – 2004.6		6.263	36.287	-3.338	-2.661	5.336	1.541
Hungary	1993.12 – 2004.6		5.506	35.817	-4.202	-3.220	16.939	-1.824
India	1988.3 – 2003.12		7.733	31.677	-4.179	-2.994	10.092	0.883
Ireland	1990.7 – 2004.6	O	5.024	18.960	-3.809	-2.677	6.559	0.216
Israel	1994.1 – 2004.6		-2.135	25.173	-3.726	-2.626	10.179	-
Italy	1925.3 – 2004.6	O	4.045	25.862	-3.461	-3.384	8.098	0.784
Japan	1922.1 – 1942.1	O	4.521	14.168	-2.739	-	-	-
	1949.7 – 2004.6	O	6.592	19.947	-3.890	-3.289	4.742	1.396
Jordan	1988.2 – 2003.2		-0.148	14.743	-3.472	-2.671	6.117	-
Luxembourg	1985.2 – 1994.12	O	7.453	14.771	-3.273	-	-	-
Malaysia	1973.1 – 2004.6		3.001	31.271	-3.820	-3.060	4.597	2.293
Mauritius	1997.1 – 2002.12		-0.921	11.895	-2.903	-2.159	9.421	-
Mexico	1988.3 – 2004.6		4.886	28.924	-4.113	-2.565	23.849	0.340
Netherlands	1969.9 – 2004.6	O	5.252	17.705	-3.134	-2.336	4.606	1.624
New Zealand	1987.1 – 2004.6	O	-5.441	20.811	-3.064	-2.730	9.059	-0.630
Norway	1970.2 – 2001.9	O	1.805	24.799	-3.620	-2.661	8.334	0.130
Philippines	1982.3 – 2004.5		1.287	31.559	-3.992	-2.684	14.451	0.029
Poland	1993.12 – 2004.5		-7.625	41.668	-4.479	-2.511	21.289	-2.793
Portugal	1988.4 – 2004.6	O	-2.330	19.345	-3.689	-2.861	7.884	0.700
Singapore	1973.1 – 2004.6		1.168	28.376	-3.953	-3.045	3.807	1.625
South Africa	1960.4 – 2004.6		6.452	22.308	-3.337	-2.368	8.812	1.750
Spain	1940.6 – 1968.12	O	7.762	13.893	-3.221	-	-	-
	1981.3 – 2004.6	O	7.512	22.136	-2.931	-2.680	7.830	0.736
Sweden	1919.2 – 2004.6	O	4.548	17.713	-3.295	-2.591	6.607	0.766
Switzerland	1966.4 – 2004.6	O	4.215	16.885	-3.798	-2.665	3.568	0.883
Taiwan	1988.3 – 2004.1		1.256	39.475	-4.709	-3.271	5.136	0.721
Thailand	1976.1 – 2004.6		4.011	32.504	-3.313	-2.467	8.030	1.346
Turkey	1986.4 – 2004.6		7.302	59.953	-3.482	-2.705	50.398	-35.619
UK	1836.1 – 1916.12	O	0.219	5.585	-3.337	-	-	-
	1924.2 – 2004.6	O	4.874	16.789	-3.089	-2.419	7.213	0.548
USA	1871.3 – 2004.6	O	4.464	16.720	-3.183	-2.663	5.106	1.165

Table 2: 1-month horizon pooled results. The first column indicates which panel is being used, with (1950.1-) denoting that all observations before 1950 are dropped from that panel. The next two columns give the number of individual time series and total number of observations in the panel. The following four columns report the pooled estimate based on recursive demeaning of the data, the standard fixed effects estimate, and the two corresponding $t-$statistics. The last column gives the number of significant coefficients in the individual time-series regressions performed on each of the time-series in the panel. The robust $t-$statistics from which proper inference can be drawn are printed in bold type.

Panel	n	# obs	$\hat{\beta}^{rd}_{n,T}$	$\hat{\beta}_{n,T}$	$t^{rd}_{n,T}$	$t_{n,T}$	$t^{+}_{0.05}$
Panel A. The earnings-price ratio							
Global	38	13,610	0.004	0.010	**1.394**	4.079	2
Global (1950.1-)	38	12,400	0.003	0.010	**0.998**	3.905	2
OECD	21	9,395	0.000	0.003	**0.127**	2.531	1
OECD (1950.1-)	21	8,185	-0.002	0.003	**-0.729**	2.400	1
Non-OECD	17	4,274	0.013	0.026	**2.464**	6.832	1
Panel B. The dividend-price ratio							
Global	46	20,594	-0.005	0.006	**-1.155**	3.352	2
Global (1950.1-)	41	14,947	-0.007	0.009	**-1.311**	4.742	3
OECD	28	16,245	-0.006	0.003	**-1.346**	1.739	2
OECD (1950.1-)	23	10,598	-0.012	0.005	**-1.601**	5.231	3
Non-OECD	18	4,349	-0.003	0.015	**-0.399**	3.074	0
Panel C. The short rate							
Global	39	15,260	-0.140	0.100	**-1.604**	0.494	11
OECD	21	10,642	-2.049	-0.946	**-5.268**	-3.956	8
Non-OECD	18	4,618	-0.082	0.135	**-0.717**	0.597	3
Panel D. The term spread							
Global	35	13,048	14.174	0.902	**0.106**	1.247	11
OECD	21	10,470	1.752	1.932	**3.026**	3.799	9
Non-OECD	14	2,578	-0.212	0.338	**-0.087**	0.374	2
Panel E. The U.S. short rate							
Global	39	15,260	-	-1.207	-	**-4.414**	7
OECD	21	10,642	-	-1.199	-	**-8.172**	4
Non-OECD	18	4,618	-	-1.230	-	**-1.211**	3
Panel F. The U.S. term spread							
Global	35	13,060	-	2.636	-	**5.365**	7
OECD	21	10,471	-	2.001	-	**4.183**	6
Non-OECD	14	2,589	-	5.047	-	**3.446**	1

Table 3: 12-month horizon results for the pooled regressions. The first column indicates what sample is being used, and the next two columns give the number of separate time-series in the panel and the total number of observations. The following four columns state the results from the balanced long-run regressions: the pooled estimate of the slope coefficient, the corresponding standard $t-$statistic, the robust, and conservative, $t_{n,T}^{B+}-$statistic, and the number of significant slope coefficients found in the individual time-series regressions, based on the conservative $t-$statistics. The final four columns give the corresponding results for the unbalanced long-run regression. The robust $t-$statistics from which proper inference can be drawn are printed in bold type.

			Balanced regression, $q=12$				Unbalanced regression, $q=12$			
Panel	n	# obs	$\hat{\beta}_{n,T}^{B}$	$t_{n,T}$	$t_{n,T}^{B+}$	$t_{0.05}^{B+}$	$\hat{\beta}_{n,T}^{U}$	$t_{n,T}$	$t_{n,T}^{U+}$	$t_{0.05}^{U+}$
Panel A. The earnings-price ratio										
Global	37	13,528	0.010	4.736	**0.850**	2	0.126	4.426	**0.427**	2
Global (1950.1-)	37	12,318	0.010	4.525	**0.683**	1	0.126	4.275	**0.293**	1
OECD	21	9,395	0.005	3.273	**-0.962**	1	0.057	3.864	**-1.488**	1
OECD (1950.1-)	21	8,185	0.005	3.032	**-1.311**	0	0.053	3.684	**-1.843**	0
Non-OECD	16	4,133	0.025	5.306	**2.940**	1	0.291	4.479	**2.164**	1
Panel B. The dividend-price ratio										
Global	45	20,522	0.010	6.528	**1.397**	7	0.111	4.684	**-0.273**	7
Global (1950.1-)	39	14,756	0.012	8.478	**2.995**	5	0.141	6.295	**0.772**	7
OECD	28	16,245	0.008	5.377	**-0.039**	4	0.069	3.608	**-1.139**	3
OECD (1950.1-)	22	10,479	0.010	9.486	**0.704**	2	0.103	9.536	**-0.389**	3
Non-OECD	17	4,277	0.017	5.073	**2.763**	3	0.206	3.696	**1.396**	4
Panel C. The short rate										
Global	39	15,260	0.175	**1.067**	1.289	2	1.887	**1.049**	1.261	6
OECD	21	10,642	-0.205	**-1.173**	-2.317	1	-6.972	**-2.770**	-3.728	3
Non-OECD	18	4,618	0.192	**1.058**	1.270	1	2.169	**1.084**	1.321	3
Panel D. The term spread										
Global	27	12,222	1.138	**2.657**	3.460	3	16.446	**3.206**	3.793	6
OECD	21	10,470	1.475	**4.122**	5.229	3	19.580	**4.297**	5.106	5
Non-OECD	6	1,752	-0.167	**-0.136**	0.586	0	5.042	**0.363**	0.916	1
Panel E. The U.S. short rate										
Global	39	15,260	-0.552	**-2.734**	-	2	-8.864	**-3.105**	-	3
OECD	21	10,642	-0.636	**-3.777**	-	1	-10.259	**-6.365**	-	2
Non-OECD	18	4,618	-0.251	**-0.349**	-	1	-4.222	**-0.371**	-	1
Panel F. The U.S. term spread										
Global	28	12,346	4.044	**9.121**	-	3	35.470	**9.571**	-	4
OECD	21	10,471	3.950	**9.658**	-	2	35.765	**14.827**	-	3
Non-OECD	7	1,875	4.535	**2.632**	-	1	33.947	**1.779**	-	1

Table 4: 60-month horizon results for the pooled regressions. The first column indicates what sample is being used, and the next two columns give the number of separate time-series in the panel and the total number of observations. The following four columns state the results from the balanced long-run regressions: the pooled estimate of the slope coefficient, the corresponding standard $t-$statistic, the robust, and conservative, $t^{B+}_{n,T}-$statistic, and the number of significant slope coefficients found in the individual time-series regressions, based on the conservative $t-$statistics. The final four columns give the corresponding results for the unbalanced long-run regression. The robust $t-$statistics from which proper inference can be drawn are printed in bold type.

Panel	n	# obs	Balanced regression, $q=60$				Unbalanced regression, $q=60$			
			$\hat{\beta}^{B}_{n,T}$	$t_{n,T}$	$t^{B+}_{n,T}$	$t^{+}_{0.05}$	$\hat{\beta}^{U}_{n,T}$	$t_{n,T}$	$t^{U+}_{n,T}$	$t^{+}_{0.05}$
Panel A. The earnings-price ratio										
Global	23	10,921	0.005	2.317	**1.661**	3	0.230	2.514	**1.273**	3
Global (1950.1-)	23	9,711	0.005	2.098	**1.444**	3	0.222	2.342	**1.010**	2
OECD	17	8,639	0.003	1.667	**1.048**	1	0.132	1.514	**0.291**	2
OECD (1950.1-)	17	7,429	0.003	1.408	**0.796**	1	0.116	1.314	**-0.002**	1
Non-OECD	6	2,282	0.022	3.348	**2.125**	2	0.937	8.900	**4.472**	1
Panel B. The dividend-price ratio										
Global	30	17,884	0.006	3.635	**1.023**	8	0.377	4.473	**1.082**	6
Global (1950.1-)	25	12,208	0.007	3.950	**0.993**	4	0.493	7.470	**2.358**	4
OECD	23	15,354	0.004	2.898	**0.075**	5	0.262	3.408	**-0.137**	4
OECD (1950.1-)	18	9,678	0.005	3.109	**-0.248**	1	0.368	7.155	**0.478**	2
Non-OECD	7	2,530	0.016	4.823	**2.489**	3	0.882	8.716	**5.200**	2
Panel C. The short rate										
Global	25	12,650	0.430	**2.476**	1.981	0	18.116	**1.800**	0.769	0
OECD	18	10,040	0.567	**3.206**	1.600	0	12.142	**1.431**	-0.636	0
Non-OECD	7	2,610	-0.270	**-0.840**	-0.128	0	32.867	**1.349**	2.106	0
Panel D. The term spread										
Global	20	10,967	-0.586	**-0.712**	-0.910	1	15.342	**1.218**	1.382	1
OECD	17	9,727	-1.288	**-1.669**	-1.944	0	9.224	**0.703**	1.454	1
Non-OECD	3	1,240	5.615	**2.432**	2.545	1	52.133	**1.296**	0.168	0
Panel E. The U.S. short rate										
Global	25	12,650	0.769	**2.956**	-	0	11.380	**1.234**	-	0
OECD	18	10,040	0.503	**2.187**	-	0	0.392	**0.056**	-	0
Non-OECD	7	2,610	2.270	**2.405**	-	0	62.964	**1.615**	-	0
Panel F. The U.S. term spread										
Global	20	10,968	-0.765	**-0.773**	-	0	53.014	**5.372**	-	0
OECD	17	9,727	0.029	**0.033**	-	0	54.934	**4.888**	-	0
Non-OECD	3	1,241	-6.633	**-1.746**	-	0	40.166	**7.006**	-	0

Table 5: 1-month horizon country level results. The first and second columns indicate the country and sample period, respectively, on which the estimates are based. The next four columns show the OLS estimates, the OLS R^2 expressed in percent, the OLS $t-$statistics, and the robust test-statistic, t_i^+. In the last column, the correlation between the residuals in the regressand and the regressor is given. The robust $t-$statistics from which proper inference can be drawn are printed in bold type.

Country	Sample	$\hat{\beta}_i$	R_i^2	t_i	t_i^+	$\hat{\delta}_i$
Panel A. The earnings-price ratio						
Argentina	1988.1 – 2004.6	0.041	2.577	2.277	**1.992**	-0.085
Australia	1962.1 – 2004.6	0.010	0.423	1.468	**0.407**	-0.665
Austria	1981.11 – 2004.6	0.002	0.066	0.423	**0.013**	-0.258
Belgium	1969.9 – 2004.6	0.012	0.408	1.305	**0.121**	-0.563
Brazil	1988.3 – 2004.5	0.022	1.361	1.632	**1.485**	-0.239
Canada	1956.3 – 2004.6	0.000	0.001	0.064	**-0.748**	-0.464
Chile	1988.3 – 2004.6	0.024	2.709	2.324	**1.467**	-0.513
Denmark	1970.1 – 2004.6	0.001	0.007	0.165	**-0.197**	-0.214
Finland	1988.3 – 2004.6	-0.005	0.195	-0.615	**-1.408**	-0.408
France	1971.11 – 2004.6	-0.001	0.003	-0.112	**-0.997**	-0.359
Germany	1969.9 – 2004.6	0.008	0.442	1.359	**0.758**	-0.570
Greece	1977.3 – 2004.6	-0.003	0.028	-0.302	**-1.940**	-0.471
Hong Kong	1973.1 – 2004.6	0.060	3.349	3.610	**0.182**	-0.755
Hungary	1993.3 – 2004.6	0.001	0.006	0.088	**-0.083**	-0.074
India	1988.3 – 2003.12	0.025	1.411	1.640	**0.790**	-0.639
Ireland	1990.7 – 2004.6	0.034	1.773	1.731	**0.580**	-0.563
Italy	1981.3 – 2004.6	0.005	0.330	0.959	**0.433**	-0.229
Japan	1956.3 – 2004.6	0.006	0.598	1.864	**0.942**	-0.555
Jordan	1988.3 – 2003.2	0.025	1.763	1.787	**1.362**	-0.320
Malaysia	1973.1 – 2004.6	0.022	0.637	1.553	**-1.783**	-0.720
Mauritius	1996.3 – 2002.12	0.003	0.063	0.225	**0.342**	-0.699
Mexico	1988.1 – 2004.6	0.017	0.467	0.959	**-0.466**	-0.574
Netherlands	1969.9 – 2004.6	0.002	0.046	0.439	**-0.171**	-0.574
New Zealand	1988.3 – 2004.6	-0.007	0.491	-0.979	**-1.699**	-0.177
Norway	1970.1 – 2001.9	-0.004	0.154	-0.764	**-1.242**	-0.304
Philippines	1982.3 – 2004.5	0.011	0.264	0.837	**0.408**	-0.300
Poland	1992.3 – 2004.5	0.025	1.331	1.398	**0.140**	-0.573
Portugal	1988.3 – 2004.6	0.010	0.356	0.832	**-0.333**	-0.464
Singapore	1973.1 – 2004.6	0.041	2.073	2.821	**-0.045**	-0.639
South Africa	1960.3 – 2004.6	0.020	0.969	2.277	**1.275**	-0.734
Spain	1980.1 – 2004.6	0.018	0.671	1.405	**-0.357**	-0.645
Sweden	1969.9 – 2004.6	0.002	0.028	0.343	**-0.058**	-0.298
Switzerland	1969.9 – 2004.6	-0.006	0.185	-0.879	**-2.492**	-0.558
Taiwan	1988.3 – 2004.1	0.043	1.614	1.761	**0.532**	-0.655
Thailand	1975.6 – 2004.6	0.013	0.485	1.300	**0.089**	-0.366
Turkey	1986.3 – 2004.6	0.035	3.250	2.706	**1.438**	-0.473
UK	1928.1 – 2004.6	0.010	0.662	2.471	**1.347**	-0.831
	1950.1 – 2004.6	0.011	0.960	2.513	**1.779**	-0.847
USA	1871.3 – 2004.6	0.011	0.718	3.398	**3.121**	-0.564
	1950.1 – 2004.6	0.007	0.536	1.874	**-0.034**	-0.822

Table 5: 1-month horizon country level results (continued).

Country	Sample	$\hat{\beta}_i$	R_i^2	t_i	t_i^+	$\hat{\delta}_i$
Panel B. The dividend-price ratio						
Argentina	1988.3 – 2004.6	-0.003	0.026	-0.224	**-0.884**	-0.378
Australia	1882.12 – 2004.6	0.005	0.138	1.416	**0.193**	-0.567
	1950.1 – 2004.6	0.013	0.560	1.916	**0.868**	-0.721
Austria	1970.2 – 2004.6	0.001	0.009	0.196	**-1.444**	-0.733
Belgium	1952.1 – 2004.6	0.003	0.077	0.694	**-0.874**	-0.769
Brazil	1988.3 – 2004.5	0.000	0.000	-0.016	**-0.962**	-0.426
Canada	1934.3 – 2004.6	0.006	0.300	1.592	**0.899**	-0.747
	1950.1 – 2004.6	0.007	0.336	1.483	**0.217**	-0.885
Chile	1983.3 – 2004.6	0.019	1.926	2.234	**1.645**	-0.388
Denmark	1970.2 – 2004.6	0.000	0.001	0.071	**-1.400**	-0.705
Finland	1962.3 – 2004.6	-0.004	0.109	-0.742	**-1.749**	-0.413
France	1898.1 – 1914.7	0.048	1.413	1.681	**0.438**	-0.845
	1919.2 – 1940.3	0.018	0.262	0.813	**-2.050**	-0.900
	1941.5 – 2004.6	-0.001	0.016	-0.353	**-1.910**	-0.609
	1950.1 – 2004.6	0.004	0.107	0.836	**-0.919**	-0.742
Germany	1872.9 – 1942.3	0.035	0.538	2.123	**1.520**	-0.143
	1953.2 – 2004.6	0.003	0.033	0.451	**-2.505**	-0.762
Greece	1977.3 – 2004.6	0.003	0.078	0.505	**0.127**	-0.272
Hong Kong	1973.1 – 2004.6	0.070	4.295	4.108	**1.433**	-0.826
Hungary	1993.12 – 2004.6	0.017	0.607	0.874	**-0.437**	-0.529
India	1988.3 – 2003.12	0.036	2.538	2.213	**1.034**	-0.616
Ireland	1990.7 – 2004.6	0.047	4.434	2.775	**1.447**	-0.659
Israel	1994.1 – 2004.6	0.018	1.036	1.139	**0.901**	-0.151
Italy	1925.3 – 2004.6	-0.006	0.276	-1.623	**-2.038**	-0.145
	1950.1 – 2004.6	0.006	0.150	0.988	**-0.696**	-0.593
Japan	1922.1 – 1942.1	-0.008	0.185	-0.665	**-2.226**	-0.549
	1949.7 – 2004.6	0.006	0.921	2.474	**2.083**	-0.528
Jordan	1988.2 – 2003.2	0.013	1.783	1.802	**1.479**	-0.281
Luxembourg	1985.2 – 1994.12	-0.007	0.224	-0.512	**-0.592**	-0.067
Malaysia	1973.1 – 2004.6	0.042	2.009	2.776	**-0.233**	-0.717
Mauritius	1997.1 – 2002.12	0.004	0.377	0.515	**0.078**	-0.148
Mexico	1988.3 – 2004.6	0.037	2.594	2.273	**1.459**	-0.468
Netherlands	1969.9 – 2004.6	0.004	0.124	0.720	**-0.713**	-0.776
New Zealand	1987.1 – 2004.6	0.039	2.458	2.290	**0.314**	-0.691
Norway	1970.2 – 2001.9	0.008	0.232	0.937	**-0.704**	-0.783
Philippines	1982.3 – 2004.5	0.003	0.152	0.635	**-0.053**	-0.359
Poland	1993.12 – 2004.5	0.045	4.618	2.451	**1.377**	-0.478
Portugal	1988.4 – 2004.6	-0.004	0.070	-0.367	**-1.349**	-0.441
Singapore	1973.1 – 2004.6	0.034	1.921	2.714	**0.395**	-0.810
South Africa	1960.4 – 2004.6	0.018	0.798	2.063	**0.919**	-0.423
Spain	1940.6 – 1968.12	0.002	0.015	0.230	**-0.758**	-0.558
	1981.3 – 2004.6	0.003	0.146	0.638	**-0.156**	-0.565
Sweden	1919.2 – 2004.6	-0.002	0.028	-0.532	**-2.075**	-0.587
	1950.1 – 2004.6	0.006	0.199	1.140	**-0.946**	-0.755
Switzerland	1966.4 – 2004.6	0.001	0.002	0.091	**-1.509**	-0.708
Taiwan	1988.3 – 2004.1	0.024	2.469	2.188	**1.640**	-0.415
Thailand	1976.1 – 2004.6	0.004	0.113	0.620	**-0.710**	-0.576
Turkey	1986.4 – 2004.6	0.041	3.219	2.687	**1.284**	-0.673
UK	1836.1 – 1916.12	0.006	0.258	1.585	**0.822**	-0.497
	1924.2 – 2004.6	0.018	0.908	2.970	**2.372**	-0.638
	1950.1 – 2004.6	0.026	1.896	3.550	**2.389**	-0.756
USA	1871.3 – 2004.6	0.004	0.076	1.105	**-1.387**	-0.767
	1950.1 – 2004.6	0.009	0.801	2.294	**2.929**	-0.954

Table 5: 1-month horizon country level results (continued).

Country	Sample	$\hat{\beta}_i$	R_i^2	t_i	t_i^+	$\hat{\delta}_i$
Panel C. The short interest rate						
Argentina	1988.1 – 2004.5	2.034	12.794	5.349	**6.425**	0.333
Australia	1952.1 – 2004.3	-0.711	0.254	-1.260	**-1.511**	-0.162
Austria	1970.1 – 2004.5	-1.739	0.337	-1.179	**-1.197**	-0.062
Belgium	1952.1 – 2004.6	-0.514	0.074	-0.682	**-0.738**	-0.138
Brazil	1988.1 – 2004.5	-0.042	0.089	-0.417	**-0.425**	-0.025
Canada	1952.1 – 2004.5	-1.489	1.019	-2.540	**-2.868**	-0.176
Chile	1983.1 – 2004.5	0.770	0.782	1.418	**0.883**	-0.249
Denmark	1970.1 – 2004.5	-0.528	0.204	-0.917	**-0.955**	-0.148
Finland	1962.1 – 2004.5	-1.369	0.466	-1.541	**-1.559**	-0.055
France	1952.1 – 2004.5	-1.022	0.311	-1.399	**-1.560**	-0.159
Germany	1953.1 – 2004.5	-3.201	1.064	-2.572	**-2.611**	-0.051
Greece	1977.1 – 2004.5	1.689	0.630	1.440	**1.439**	-0.091
Hong Kong	1970.1 – 2004.5	-3.076	0.445	-1.355	**-1.408**	-0.244
Hungary	1991.3 – 2004.5	-0.922	0.309	-0.698	**-0.735**	-0.021
India	1988.1 – 2003.12	-1.368	0.331	-0.794	**-0.787**	-0.121
Ireland	1988.3 – 2004.5	-0.338	0.044	-0.292	**-0.393**	-0.087
Israel	1993.1 – 2004.5	2.489	1.088	1.219	**1.191**	-0.017
Italy	1952.1 – 2004.5	-0.886	0.363	-1.512	**-1.592**	-0.108
Japan	1952.1 – 2004.5	1.182	0.221	1.178	**1.100**	-0.109
Jordan	1988.1 – 2003.2	-6.669	6.012	-3.393	**-3.394**	-0.002
Malaysia	1972.12 – 2004.5	-4.135	0.316	-1.091	**-1.229**	-0.117
Mauritius	1989.9 – 2004.5	-7.398	7.019	-3.635	**-3.576**	0.100
Mexico	1988.1 – 2004.6	-0.011	0.000	-0.030	**-0.385**	-0.025
Netherlands	1952.1 – 2004.5	-1.932	0.775	-2.213	**-2.381**	-0.108
New Zealand	1986.8 – 2004.3	-4.346	6.802	-3.915	**-4.074**	-0.109
Norway	1970.1 – 2001.9	-1.589	0.393	-1.223	**-1.357**	-0.071
Philippines	1982.1 – 2004.5	-0.173	0.011	-0.175	**-0.188**	-0.081
Poland	1991.6 – 2004.5	0.039	0.001	0.033	**0.129**	0.034
Portugal	1988.3 – 2004.5	-2.221	2.003	-1.986	**-2.049**	-0.107
Singapore	1970.1 – 2004.5	-0.308	0.006	-0.153	**-0.165**	-0.115
South Africa	1960.3 – 2004.5	-0.991	0.345	-1.354	**-1.524**	-0.131
Spain	1952.1 – 2004.5	-1.609	1.457	-3.045	**-3.141**	-0.117
Sweden	1952.1 – 2004.5	0.475	0.068	0.651	**0.527**	-0.162
Switzerland	1966.3 – 2004.6	-2.492	0.875	-2.010	**-2.061**	-0.144
Taiwan	1988.1 – 2004.1	-4.339	0.332	-0.798	**-0.803**	0.045
Thailand	1975.6 – 2004.6	-5.321	3.485	-3.540	**-3.490**	0.036
Turkey	1986.3 – 2004.6	0.518	0.162	0.595	**0.548**	-0.159
UK	1952.1 – 2004.5	-0.475	0.063	-0.629	**-1.181**	-0.242
USA	1952.1 – 2004.5	-1.825	1.032	-2.558	**-2.641**	-0.056

Table 5: 1-month horizon country level results (continued).

Country	Sample	$\hat{\beta}_i$	R_i^2	t_i	t_i^+	$\hat{\delta}_i$
Panel D. The term spread						
Argentina	1997.1 – 2004.6	2.682	5.077	2.169	**2.249**	0.118
Australia	1952.1 – 2004.6	1.428	0.177	1.056	**1.096**	0.047
Austria	1970.1 – 2004.6	-0.976	0.029	-0.345	**-0.372**	-0.018
Belgium	1952.1 – 2004.6	0.822	0.021	0.366	**0.327**	-0.026
Brazil	1994.1 – 2004.5	0.006	0.000	0.005	**0.226**	0.046
Canada	1952.1 – 2004.6	3.387	0.960	2.467	**2.577**	0.058
Denmark	1970.1 – 2004.6	0.591	0.066	0.520	**0.511**	-0.010
Finland	1962.1 – 2004.6	2.559	0.355	1.345	**1.312**	0.073
France	1952.1 – 2004.6	2.902	0.430	1.646	**1.804**	0.067
Germany	1953.1 – 2004.6	3.402	0.512	1.780	**1.743**	-0.032
Greece	1993.3 – 2004.6	-4.889	1.084	-1.212	**-1.141**	0.074
Hong Kong	1994.11 – 2004.6	0.271	0.001	0.029	**0.477**	0.340
Hungary	1997.4 – 2004.6	3.692	0.111	0.308	**-0.473**	-0.252
India	1988.1 – 2003.12	1.813	0.459	0.936	**0.895**	0.099
Ireland	1988.3 – 2004.6	0.111	0.002	0.069	**0.095**	0.050
Italy	1952.1 – 2004.6	2.656	0.436	1.659	**1.656**	-0.002
Japan	1952.1 – 2004.6	-3.154	0.631	-1.997	**-1.925**	0.042
Malaysia	1972.12 – 2004.6	3.579	0.220	0.911	**0.974**	0.117
Mexico	1995.3 – 2004.6	-0.488	0.047	-0.227	**-0.307**	0.020
Netherlands	1952.1 – 2004.6	2.922	0.527	1.824	**1.858**	0.059
New Zealand	1986.8 – 2004.6	10.240	7.312	4.099	**4.103**	-0.002
Norway	1970.1 – 2001.9	3.022	0.588	1.497	**1.611**	0.069
Philippines	1994.11 – 2004.5	-0.966	0.097	-0.331	**-0.364**	-0.025
Poland	1994.4 – 2004.5	3.553	0.233	0.529	**0.550**	0.160
Portugal	1988.3 – 2004.6	5.605	0.667	1.142	**1.140**	-0.001
Singapore	1988.1 – 2004.6	11.227	1.011	1.415	**1.634**	0.229
South Africa	1960.3 – 2004.6	2.240	0.382	1.425	**1.425**	-0.003
Spain	1952.1 – 2004.6	3.678	1.973	3.555	**3.594**	0.085
Sweden	1952.1 – 2004.6	-0.106	0.001	-0.061	**-0.119**	0.092
Switzerland	1966.3 – 2004.6	3.234	0.679	1.769	**1.856**	0.087
Taiwan	1995.3 – 2004.1	-4.869	0.143	-0.388	**-0.317**	0.201
Thailand	1977.2 – 2004.6	6.253	1.558	2.275	**2.266**	-0.044
Turkey	1997.11 – 2004.6	-1.474	1.434	-1.065	**-1.015**	0.033
UK	1952.1 – 2004.6	0.941	0.081	0.711	**0.780**	0.031
USA	1952.1 – 2004.6	4.946	1.256	2.826	**2.791**	-0.041

Table 6: 60-month horizon country level results. The first and second columns indicate the country and sample period, respectively, on which the estimates are based. The next four columns show the long-run OLS estimates, the OLS R^2 expressed in percent, the scaled OLS $t-$statistics, and the scaled robust test-statistic. The results in each panel are either for the balanced or unbalanced regression, as indicated by the superscript B or U, respectively. The robust $t-$statistics from which proper inference can be drawn are printed in bold type.

Country	Sample	$\hat{\beta}_i^B$	R_i^2	$t_i^B / \sqrt{q}$	$t_i^{B+} / \sqrt{q}$
	Panel A. The earnings-price ratio (Balanced)				
Australia	1962.1 – 2004.6	-0.001	0.238	-0.125	**-0.287**
Austria	1981.11 – 2004.6	0.009	18.174	0.753	**0.637**
Belgium	1969.9 – 2004.6	0.029	31.373	1.509	**1.144**
Canada	1956.3 – 2004.6	-0.002	3.530	-0.530	**-0.498**
Denmark	1970.1 – 2004.6	-0.006	31.394	-1.500	**-1.692**
France	1971.11 – 2004.6	0.009	15.552	0.915	**0.840**
Germany	1969.9 – 2004.6	0.007	13.411	0.879	**0.666**
Greece	1977.3 – 2004.6	0.046	36.809	1.424	**1.632**
Hong Kong	1973.1 – 2004.6	0.037	20.626	1.059	**1.426**
Italy	1981.3 – 2004.6	-0.002	1.421	-0.197	**0.065**
Japan	1956.3 – 2004.6	0.008	34.984	2.033	**1.836**
Malaysia	1973.1 – 2004.6	0.005	1.422	0.250	**-0.393**
Netherlands	1969.9 – 2004.6	0.005	6.522	0.590	**0.320**
Norway	1970.1 – 2001.9	-0.005	12.948	-0.806	**-0.530**
Philippines	1982.3 – 2004.5	0.030	61.140	1.970	**1.754**
Singapore	1973.1 – 2004.6	-0.013	9.123	-0.658	**-1.137**
South Africa	1960.3 – 2004.6	0.012	21.643	1.379	**1.169**
Spain	1980.1 – 2004.6	-0.012	3.236	-0.312	**-0.687**
Sweden	1969.9 – 2004.6	0.004	4.167	0.465	**0.525**
Switzerland	1969.9 – 2004.6	-0.001	0.088	-0.066	**-0.642**
Thailand	1975.6 – 2004.6	0.038	54.020	2.122	**1.778**
UK	1928.1 – 2004.6	0.007	17.508	1.681	**0.774**
	1950.1 – 2004.6	0.007	19.443	1.467	**0.280**
USA	1871.3 – 2004.6	0.005	4.463	1.074	**1.116**
	1950.1 – 2004.6	0.003	3.236	0.546	**-0.238**

Table 6: 60-month horizon country level results (continued).

Country	Sample	$\hat{\beta}_i^U$	R_i^2	$t_i^U / \sqrt{q}$	$t_i^{U+} / \sqrt{q}$
Panel B. The earnings-price ratio (Unbalanced)					
Australia	1962.1 – 2004.6	0.310	11.185	0.972	**0.592**
Austria	1981.11 – 2004.6	-0.074	1.360	-0.221	**-0.458**
Belgium	1969.9 – 2004.6	0.833	25.192	1.418	**0.915**
Canada	1956.3 – 2004.6	-0.058	1.452	-0.357	**-0.265**
Denmark	1970.1 – 2004.6	-0.110	4.670	-0.538	**-0.727**
France	1971.11 – 2004.6	-0.096	1.725	-0.312	**-0.698**
Germany	1969.9 – 2004.6	0.055	0.381	0.151	**-0.097**
Greece	1977.3 – 2004.6	1.396	40.350	1.738	**0.400**
Hong Kong	1973.1 – 2004.6	1.064	35.472	1.707	**0.969**
Italy	1981.3 – 2004.6	-0.375	34.772	-1.398	**-0.513**
Japan	1956.3 – 2004.6	0.454	39.912	2.399	**1.816**
Malaysia	1973.1 – 2004.6	0.879	21.871	1.218	**0.275**
Netherlands	1969.9 – 2004.6	0.253	8.036	0.722	**0.506**
Norway	1970.1 – 2001.9	0.010	0.023	0.035	**0.151**
Philippines	1982.3 – 2004.5	1.370	44.936	1.678	**1.335**
Singapore	1973.1 – 2004.6	0.699	21.190	1.194	**0.360**
South Africa	1960.3 – 2004.6	0.935	58.435	3.326	**2.983**
Spain	1980.1 – 2004.6	0.986	23.961	1.109	**-0.181**
Sweden	1969.9 – 2004.6	0.062	0.506	0.174	**0.213**
Switzerland	1969.9 – 2004.6	-0.119	1.234	-0.273	**-1.376**
Thailand	1975.6 – 2004.6	0.709	12.911	0.845	**0.360**
UK	1928.1 – 2004.6	0.511	35.353	2.796	**1.797**
	1950.1 – 2004.6	0.458	35.917	2.356	**1.472**
USA	1871.3 – 2004.6	0.311	6.990	1.389	**1.378**
	1950.1 – 2004.6	0.298	10.193	1.060	**0.159**

Table 6: 60-month horizon country level results (continued).

Country	Sample	$\hat{\beta}_i^B$	R_i^2	$t_i^B / \sqrt{q}$	$t_i^{B+} / \sqrt{q}$
		Panel C. The dividend-price ratio (Balanced)			
Australia	1882.12 – 2004.6	0.007	7.766	1.371	**0.982**
	1950.1 – 2004.6	0.006	5.945	0.751	**0.407**
Austria	1970.2 – 2004.6	0.006	9.847	0.732	**0.119**
Belgium	1952.1 – 2004.6	0.009	13.476	1.152	**-0.112**
Canada	1934.3 – 2004.6	0.009	23.099	1.905	**1.723**
	1950.1 – 2004.6	-0.002	0.820	-0.272	**-0.676**
Chile	1983.3 – 2004.6	0.016	24.881	0.870	**0.575**
Denmark	1970.2 – 2004.6	-0.005	15.658	-0.954	**-2.318**
Finland	1962.3 – 2004.6	-0.001	0.214	-0.118	**-0.835**
France	1919.2 – 1940.3	0.099	58.484	1.780	**4.081**
	1941.5 – 2004.6	-0.001	0.309	-0.182	**-2.674**
	1950.1 – 2004.6	0.006	9.283	0.955	**-0.927**
Germany	1872.9 – 1942.3	0.032	6.356	0.900	**1.148**
	1953.2 – 2004.6	0.011	9.732	0.946	**0.385**
Greece	1977.3 – 2004.6	0.022	59.873	2.280	**2.448**
Hong Kong	1973.1 – 2004.6	0.034	25.397	1.212	**2.199**
Italy	1925.3 – 2004.6	-0.001	0.050	-0.083	**-0.662**
	1950.1 – 2004.6	0.003	1.086	0.313	**-0.072**
Japan	1922.1 – 1942.1	0.009	3.235	0.261	**-0.567**
	1949.7 – 2004.6	0.005	32.349	2.076	**1.229**
Malaysia	1973.1 – 2004.6	0.011	4.037	0.426	**0.381**
Netherlands	1969.9 – 2004.6	0.010	16.310	0.985	**-0.213**
Norway	1970.2 – 2001.9	-0.001	0.272	-0.109	**-0.634**
Philippines	1982.3 – 2004.5	0.012	64.051	2.096	**1.204**
Singapore	1973.1 – 2004.6	0.002	0.872	0.195	**-0.434**
South Africa	1960.4 – 2004.6	0.017	40.002	2.140	**1.706**
Spain	1940.6 – 1968.12	0.004	1.292	0.221	**-0.183**
	1981.3 – 2004.6	-0.010	27.329	-1.005	**-1.703**
Sweden	1919.2 – 2004.6	0.002	0.936	0.378	**0.564**
	1950.1 – 2004.6	0.006	6.757	0.804	**-0.011**
Switzerland	1966.4 – 2004.6	0.005	3.079	0.424	**-0.666**
Thailand	1976.1 – 2004.6	0.025	57.914	2.262	**2.519**
UK	1836.1 – 1916.12	0.007	24.299	2.136	**2.661**
	1924.2 – 2004.6	0.017	23.667	2.091	**2.091**
	1950.1 – 2004.6	0.015	19.824	1.485	**0.923**
USA	1871.3 – 2004.6	0.008	9.677	1.626	**1.549**
	1950.1 – 2004.6	0.005	5.622	0.729	**-0.061**

Table 6: 60-month horizon country level results (continued).

Country	Sample	$\hat{\beta}_i^U$	R_i^2	$t_i^U / \sqrt{q}$	$t_i^{U+} / \sqrt{q}$
Panel D. The dividend-price ratio (Unbalanced)					
Australia	1882.12 – 2004.6	0.440	13.799	1.932	**1.282**
	1950.1 – 2004.6	0.622	23.656	1.751	**1.175**
Austria	1970.2 – 2004.6	0.162	2.467	0.386	**-0.452**
Belgium	1952.1 – 2004.6	0.405	13.191	1.201	**-0.277**
Canada	1934.3 – 2004.6	0.457	23.551	2.006	**1.796**
	1950.1 – 2004.6	0.340	13.769	1.257	**0.615**
Chile	1983.3 – 2004.6	1.464	71.073	2.833	**3.169**
Denmark	1970.2 – 2004.6	-0.056	0.643	-0.195	**-1.042**
Finland	1962.3 – 2004.6	0.097	0.469	0.188	**-0.525**
France	1919.2 – 1940.3	1.229	14.541	0.742	**-0.425**
	1941.5 – 2004.6	0.025	0.094	0.105	**-1.475**
	1950.1 – 2004.6	0.455	16.810	1.414	**-0.394**
Germany	1872.9 – 1942.3	1.059	6.167	0.921	**1.987**
	1953.2 – 2004.6	0.656	15.458	1.303	**-0.091**
Greece	1977.3 – 2004.6	0.758	37.688	1.644	**1.641**
Hong Kong	1973.1 – 2004.6	1.413	52.768	2.433	**1.412**
Italy	1925.3 – 2004.6	-0.110	1.121	-0.411	**-1.509**
	1950.1 – 2004.6	0.390	6.830	0.852	**-0.062**
Japan	1922.1 – 1942.1	-0.242	2.316	-0.267	**-1.278**
	1949.7 – 2004.6	0.353	45.823	2.908	**2.177**
Malaysia	1973.1 – 2004.6	1.452	53.204	2.455	**0.987**
Netherlands	1969.9 – 2004.6	0.399	10.853	0.852	**-0.278**
Norway	1970.2 – 2001.9	0.334	10.901	0.808	**0.194**
Philippines	1982.3 – 2004.5	0.729	71.488	2.941	**1.295**
Singapore	1973.1 – 2004.6	0.765	34.956	1.688	**0.441**
South Africa	1960.4 – 2004.6	0.965	64.647	3.789	**2.679**
Spain	1940.6 – 1968.12	0.804	29.543	1.406	**0.607**
	1981.3 – 2004.6	0.250	9.656	0.626	**-0.011**
Sweden	1919.2 – 2004.6	0.041	0.144	0.153	**-0.855**
	1950.1 – 2004.6	0.277	6.215	0.810	**-0.525**
Switzerland	1966.4 – 2004.6	0.261	3.930	0.522	**-1.482**
Thailand	1976.1 – 2004.6	0.715	21.798	1.145	**1.057**
UK	1836.1 – 1916.12	0.282	12.554	1.477	**1.099**
	1924.2 – 2004.6	1.075	50.992	3.962	**3.616**
	1950.1 – 2004.6	1.014	53.273	3.360	**2.692**
USA	1871.3 – 2004.6	0.371	8.816	1.575	**0.733**
	1950.1 – 2004.6	0.445	20.194	1.583	**0.852**

Table 7: Out-of-sample results at the 1-month horizon. The first and second columns indicate the country and the sample period that are available, respectively. The next two columns give the in-sample robust $t-$statistic and standard in-sample R^2 expressed in percent that is obtained when using the entire sample. The following three columns show the out-of-sample results based on the time-series estimates. The first of these states the out-of-sample R^2 expressed in percent. The second gives the ratio of the out-of-sample R^2 to the squared Sharpe ratio, which is approximately equal to the *proportional* increase in expected returns, expressed in percent, for a log-utility investor utilizing the predictive power of the short rate or the term spread. The third column shows the increase in the real annual risk free rate, calculated as $12 \times R^2_{OS}/2$ and expressed in percentage points, that is equivalent to the welfare gain from using the conditional forecasts. The final three columns give the corresponding out-of-sample results based on the pooled estimates.

		In-sample		Time-series			Pooled		
Country	Sample	t_i^+	R_i^2	$R^2_{i,OS}$	$R^2_{i,OS}/S_i^2$	$6 \times R^2_{i,OS}$	$R^2_{i,OS}$	$R^2_{i,OS}/S_i^2$	$6 \times R^2_{i,OS}$
Panel A. The short interest rate									
Australia	1952.1 – 2004.3	-1.511	0.254	-0.149	-19.036	-0.892	0.150	19.195	0.900
Belgium	1952.1 – 2004.6	-0.738	0.074	0.008	3.343	0.050	0.117	47.053	0.702
Canada	1952.1 – 2004.5	-2.868	1.019	0.540	155.904	3.243	0.543	156.575	3.257
Finland	1962.1 – 2004.5	-1.559	0.466	0.149	21.635	0.894	0.157	22.773	0.941
France	1952.1 – 2004.5	-1.560	0.311	0.257	52.282	1.540	0.251	51.039	1.503
Germany	1953.1 – 2004.5	-2.611	1.064	0.481	70.773	2.885	0.692	101.913	4.154
Italy	1952.1 – 2004.5	-1.592	0.363	0.318	405.937	1.907	0.370	473.047	2.222
Japan	1952.1 – 2004.5	1.100	0.221	-0.022	-2.376	-0.134	-0.066	-6.987	-0.394
Netherlands	1952.1 – 2004.5	-2.381	0.775	0.097	7.488	0.582	0.260	20.067	1.559
South Africa	1960.3 – 2004.5	-1.524	0.345	0.096	14.749	0.575	0.164	25.332	0.987
Spain	1952.1 – 2004.5	-3.141	1.457	1.482	226.359	8.889	1.478	225.834	8.869
Sweden	1952.1 – 2004.5	0.527	0.068	0.040	4.181	0.241	-0.220	-22.880	-1.320
UK	1952.1 – 2004.5	-1.181	0.063	0.181	29.117	1.089	0.219	35.204	1.316
USA	1952.1 – 2004.5	-2.641	1.032	0.003	0.269	0.021	0.496	38.711	2.976
Panel B. The term spread									
Australia	1952.1 – 2004.6	1.096	0.177	0.058	7.262	0.345	0.032	4.068	0.193
Belgium	1952.1 – 2004.6	0.327	0.021	-0.085	-34.077	-0.508	0.055	21.959	0.328
Canada	1952.1 – 2004.6	2.577	0.960	0.516	146.416	3.097	0.484	137.433	2.907
Finland	1962.1 – 2004.6	1.312	0.355	0.300	42.289	1.800	0.572	80.625	3.431
France	1952.1 – 2004.6	1.804	0.430	0.204	40.667	1.224	0.272	51.317	1.544
Germany	1953.1 – 2004.6	1.743	0.512	0.395	56.860	2.368	0.428	61.697	2.570
Italy	1952.1 – 2004.6	1.656	0.436	0.276	331.597	1.657	0.281	337.650	1.687
Japan	1952.1 – 2004.6	-1.925	0.631	0.000	0.000	0.000	-0.112	-11.635	-0.675
Netherlands	1952.1 – 2004.6	1.858	0.527	-0.123	-9.386	-0.738	0.318	24.304	1.911
South Africa	1960.3 – 2004.6	1.425	0.382	-0.129	-20.554	-0.774	0.431	68.665	2.587
Spain	1952.1 – 2004.6	3.594	1.973	1.213	183.393	7.278	0.917	138.674	5.503
Sweden	1952.1 – 2004.6	-0.119	0.001	-0.421	-43.093	-2.528	-0.361	-36.882	-2.164
UK	1952.1 – 2004.6	0.780	0.081	-0.657	-104.714	-3.940	-0.054	-8.548	-0.322
USA	1952.1 – 2004.6	2.791	1.256	-1.316	-101.614	-7.894	0.538	41.514	3.225

Table 8: Stock indices. This table lists the stock-index in each country to which the returns and dividends- and earnings-price ratios correspond.

Country	Stock-index
Argentina	BUSE
Australia	ASX-All Ordinaries
Austria	Vienna SE Return Index
Belgium	CBB All-Share
Brazil	Sau Paulo IBX-50
Canada	Toronto SE-300
Chile	Santiago SE Return Index
Denmark	Copenhagen KAX
Finland	HEX All-Share
France	SBF-250
Germany	CDAX
Greece	Athens Main SE Return Index
Hong Kong	Hong Kong SE Return Index
Hungary	BUX
India	S&P CNX Nifty-50
Ireland	ISEQ
Israel	Tel Aviv SE Return Index
Italy	BCI
Japan	Nikko Securities Composite
Jordan	Jordan SE Return Index
Luxembourg	Luxembourg SE Return Index
Malaysia	Kuala Lumpur SE Return Index
Mauritius	Mauritius Semdex
Mexico	Mexico SE Return Index
Netherlands	Netherlands All-Share Index
New Zealand	NZSX 50
Norway	Oslo SE Return Index
Philippines	Philippines SE Return Index
Poland	WIG
Portugal	BVL
Singapore	SEI
South Africa	Johannesburg SE Return Index
Spain	Barcelona SE-30
Sweden	Affarsvärlden Return Index
Switzerland	Swiss Performance Index
Taiwan	FTSE/TSE-50
Thailand	Bangkok SE Return Index
Turkey	ISE-100
UK	FTA All-Share
USA	S&P 500

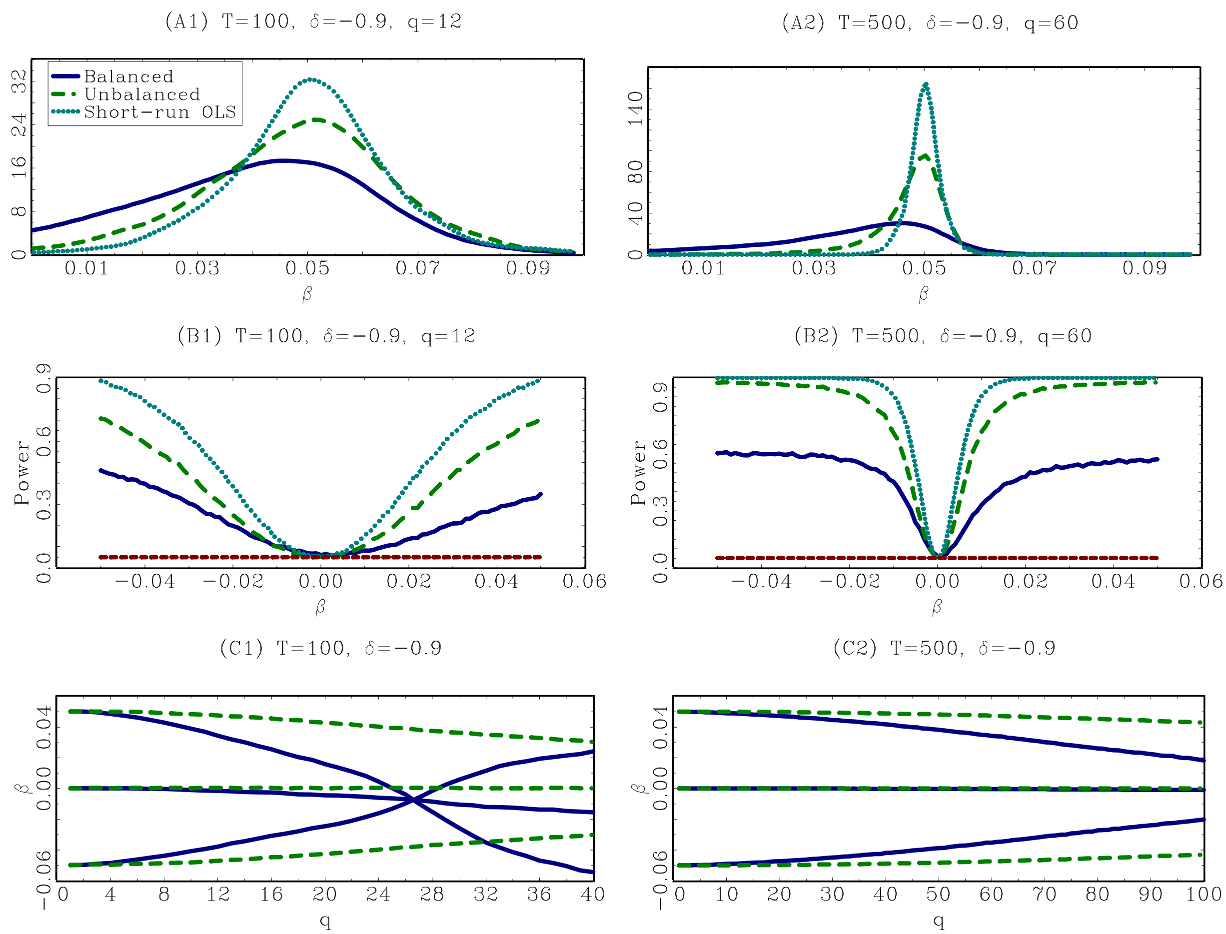

Figure 1: Results from the Monte Carlo simulation. The top two graphs, (A1) and (A2), show the kernel estimates of the densities of the estimated coefficients, using the estimators $\hat{\beta}^{B+}(q)$, $\hat{\beta}^{U+}(q)$, and $\hat{\beta}^{U+}(q=1)$, referred to as Balanced, Unbalanced and Short-run OLS, respectively, in the legend. The true value of β is equal to 0.05 and $\hat{\beta}^{U+}(q)$ is divided by q to enable comparison between the estimators. The middle graphs, (B1) and (B2), show the average rejection rates of the $t-$tests corresponding to the respective estimators; the flat dashed lines show the 5% level. The bottom two graphs plot the average estimated values of β as a function of the forecasting horizon q. The true values of β are $0.05, 0.00$, and -0.05, representing the upper, middle, and lower lines in the graphs, respectively.

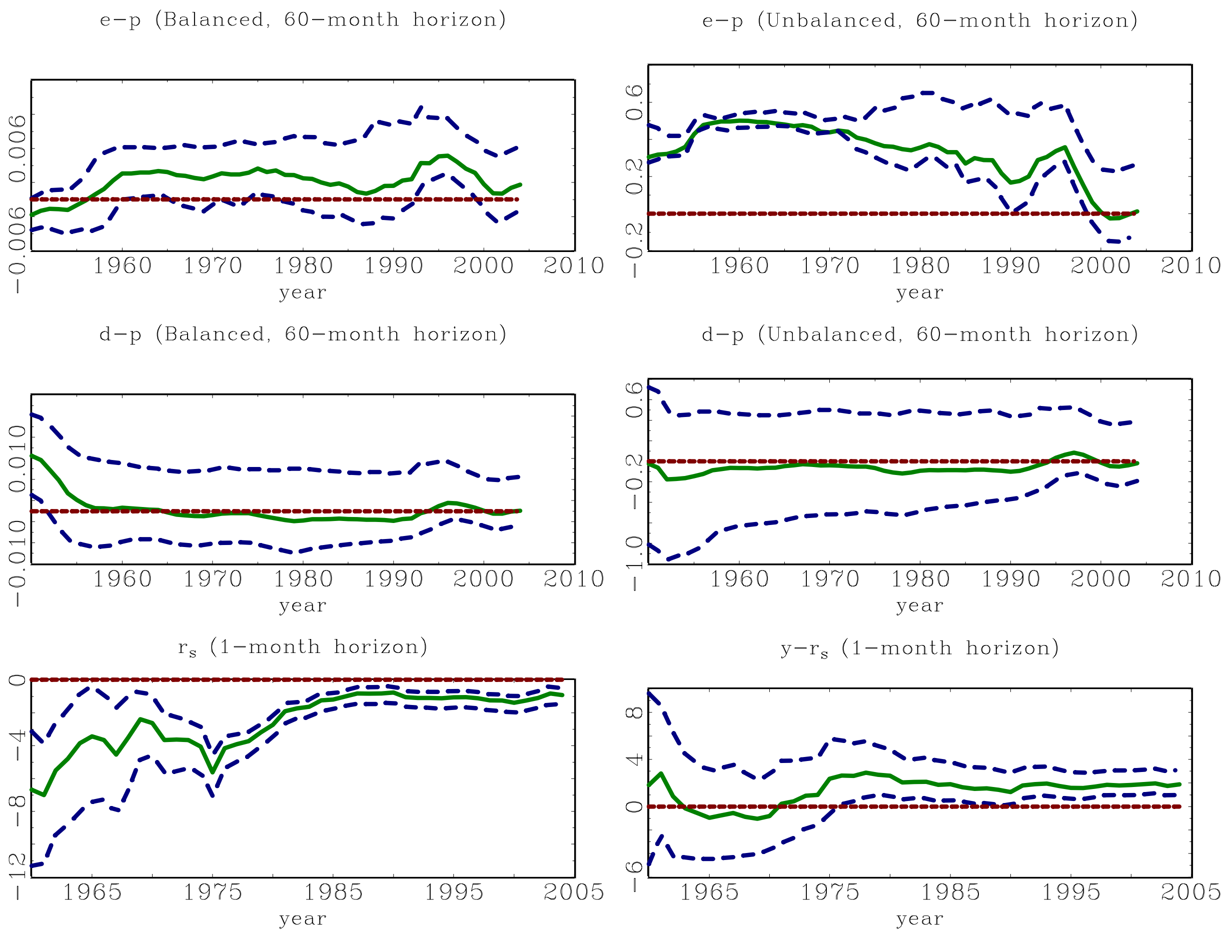

Figure 2: Rolling regression pooled estimates for the OECD panel. Each graph depicts the pooled estimate, and the corresponding conservative confidence bounds, with a coverage rate of 90%, that are obtained when pooling at the OECD level all observations available up until the year in the plot. For the earnings $(e - p)$ and dividend price ratios $(d - p)$, a time-series is added to the panel when 15 years of observations become available. For the short rate (r_s) and the term spread $(y - r_s)$, a time-series is added when there are five years of available observations. The flat dashed lines indicate a value of zero.

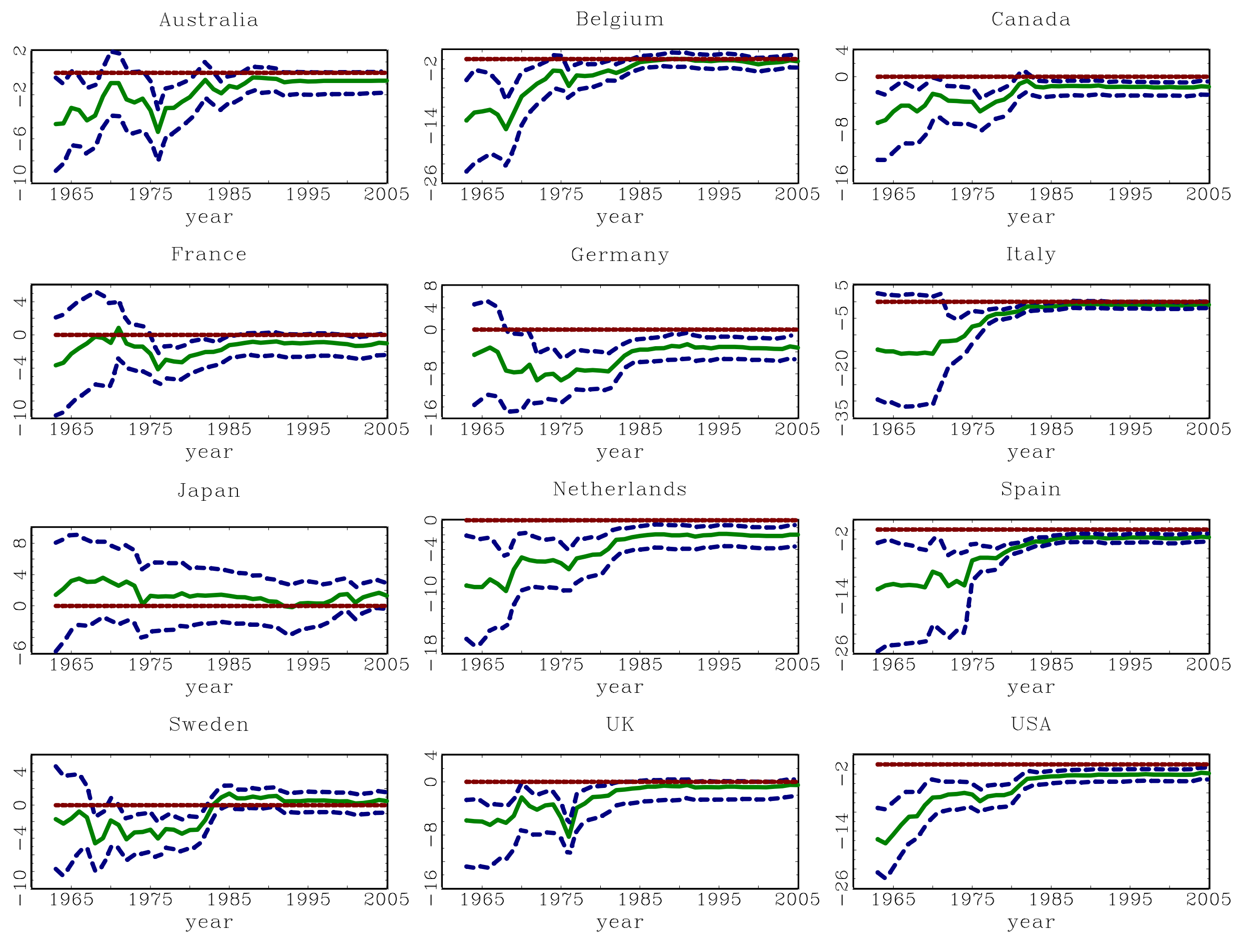

Figure 3: Rolling regression estimates for the short interest rate. Each graph depicts the point estimate and conservative confidence interval, with a coverage rate of 90%, that result from regressing excess returns, in the country indicated, onto the lagged value of the short interest rate. All regressions are at the 1-month horizon. The samples used in the estimation include data up till the year shown in the graph. The flat dashed lines indicate a value of zero.