Improved Inference and Estimation in Regression With Overlapping Observations

Mark Britten-Jones
Head of European Equity Research,
Barclays Global Investors, Murray House,
1 Royal Mint Court,
London EC3N 4HH
United Kingdom.
Tel: (44) 20 7668 8965.
email: Mark.Britten-Jones@barclaysglobal.com.

Anthony Neuberger*
Professor of Finance
Warwick Business School
The University of Warwick
Coventry CV4 7AL
United Kingdom
Tel: (44) 24 7652 2955
email: anthony.neuberger@wbs.ac.uk

* Corresponding author.

ABSTRACT

We present an improved method for inference in linear regressions with overlapping observations. The method transforms the regression so that observations of the dependent variable are non-overlapping, allowing standard inference procedures to be used. These procedures are asymptotically valid when applied to the transformed regression, and Monte Carlo analysis shows they perform better in finite samples than the more sophisticated methods applied to the original regression that are in common usage. The transformation is also applicable to panel and Fama-MacBeth regressions.

Keywords: Long horizon, stock return predictability, induced autocorrelation

JEL Classification: C20, G12

1. INTRODUCTION

Researchers in empirical finance often regress long-horizon returns onto explanatory variables. Such regressions have been used to assess stock return predictability, to test the expectations theory of the term structure of interest rates, to test the cross-sectional pricing implications of the CAPM and consumption-CAPM, and to test the efficiency of foreign exchange markets. These regressions involve overlapping observations which raise econometric issues that are addressed in this paper.

Regressions with long horizon returns often show much higher R^2 than regressions with one-period returns. But work by Valkanov (2003), Boudoukh, Richardson and Whitelaw (forthcoming) and Hjmarlsson (2006) suggests that long-horizon return regressions have no greater statistical power to reject the null of no predictability than their short-horizon counterparts. For *testing* predictability the use of long-horizon returns (as opposed to one-period returns) would appear to be of little value.

Nonetheless, the analysis of long-horizon returns can contribute significantly to *understanding* predictability (or dependence) and its economic significance. For example, one concern with the interpretation of short-horizon regressions is measurement error. Cochrane and Piazzesi (2007, p. 139) forecast annual bond returns using monthly data and claim that "to see the core results you must look directly at the one-year horizon" and further find that estimating a typical one-month return model "completely misses the single factor representation" due to measurement error.

Another reason for analyzing longer horizons is lengthy and uncertain response times. A number of recent studies of the consumption-CAPM, including Daniel and Marshall (1997), Parker (2001), Julliard and Parker (2005), Jagannathan and Wang (2005, 2007) and Malloy, Moskowitz and Vissing-Jorgensen (2005) measure consumption risk using consumption growth and returns measured over several periods. In general this approach works better than the standard consumption-CAPM. If consumers face costs associated with changing consumption, or if information acquisition is constrained, then consumption may change more slowly than implied by the standard consumption-CAPM, and this justifies a focus on returns and consumption growth measured over longer horizons. Even though long-horizon regressions may be no better than single-period regressions for testing predictability, the above examples show that researchers are likely to continue using long-horizon regressions in order to better understand the nature of predictability.

Long-horizon return regressions potentially suffer from two econometric problems. The first is bias in the usual OLS slope estimate. The bias is not caused by the presence of overlapping observations but arises when the predictor variable is persistent and its innovations are strongly correlated with returns (see Mankiw and Shapiro (1986) and Stambaugh (1999)). These conditions may also arise in short-horizon regressions. A number of studies address the question of how to correct this bias (see Campbell and Yogo (2003) and Amihud and Hurvich (2004) and the references therein). Our paper does not address the problem of bias but focuses instead on a problem specific to overlapping observations - the strong autocorrelation induced by overlapping observations. It is now well known that commonly used methods to deal with the autocorrelation are inadequate and can be quite misleading in finite samples.

This paper presents a simple procedure that can markedly improve inference in regressions with overlapping observations. We consider an overlapping regression in which a multi-period return is regressed onto a set of regressors, and for which observations are available each period. This regression is transformed into a non-overlapping regression in which one-period returns are regressed onto a set of transformed regressors. The OLS coefficient estimates from the original and transformed regressions are numerically identical, but inference based on the transformed regression is simplified because the autocorrelation induced by overlapping observations is no longer present. The procedure is applicable to time-series regressions and to panel regressions. It can be applied to both predictive (forecasting) and contemporaneous (explanatory) regressions. We also show how the method can be adapted to improve the efficiency of regression coefficient estimates.

We show that standard inference techniques, such as OLS and the White (1980) adjustment for heteroscedasticity, are asymptotically valid when applied to the transformed regression. To assess the finite-sample performance of our procedure we run Monte Carlo simulations. The simulations show that standard inference procedures (OLS and White

1980) perform substantially better when based on the transformed regression rather than on the original specification, and indeed these simple procedures when applied to the transformed regression perform better than more sophisticated techniques such as Hansen-Hodrick (1980) and Newey-West (1987) applied to the original regression. The superior performance of our procedure is most marked when the return horizon in the original specification is long in comparison to the sample length, and Hansen-Hodrick and Newey-West standard errors tend to be severely biased down.

To better understand the bias in the Hansen-Hodrick and Newey-West standard errors, we show that it derives in part from a bias in the estimate of the auto-covariance matrix and we provide a simple analytical expression for it. In contrast, the standard errors obtained from our transformed regression have much less bias and also exhibit lower sampling variability. The result is that confidence intervals using our method have coverage probabilities much closer to their nominal levels than confidence intervals constructed using standard techniques.

Other papers have documented problems with conventional inference applied to long-horizon regressions (for example Ang and Bekaert 2007; Nelson and Kim 1993; and Hodrick 1992) and utilize or advocate simulation techniques for inference. Another strand of the literature develops covariance estimators for specific cases, estimators that impose more structure on the serial correlation of moment conditions. These structured estimators generally have excellent small-sample properties, but their applicability is limited. For example, the estimator of Richardson and Smith (1991) provides valid inference only under the null hypothesis that returns are serially uncorrelated, and only when the explanatory variables are past returns. Even then, valid inference requires the unpalatable assumption (for asset returns) of conditional homoscedasticity.

The methodology that is mathematically most similar to ours is that presented in Hodrick (1992). He presents a structured covariance estimator that generalizes Richardson and Smith (1991) in that regressors need not be past returns and returns need not be conditionally homoscedastic. We show that this estimator is in fact a special case of our methodology and produces standard error estimates that are identical to White (1980) heteroscedasticity-consistent standard errors from our transformed regression. A drawback of Hodrick's approach is that it is fairly complex and is valid only under the null hypothesis of no-predictability of returns. It has not gained widespread acceptance. For example the well-known text by Campbell, Lo, and Mackinlay (1997) discusses statistical inference in long-horizon regressions at length but does not present the estimator of Hodrick (1992). Ang and Bekaert (2007) is an exception. They use Hodrick (1992) standard errors and argue that much of the empirical evidence for the time-series predictability of stock returns has been overstated in the literature due, in part, to the use of OLS or Hansen-Hodrick (1980) standard errors which they find 'lead to severe over-rejections of the null hypothesis'.

An approach similar in appearance to ours is advocated by Jegadeesh (1991) and Cochrane (1991). They bypass the problem of overlapping observations by regressing one-period returns onto the sum of lags of the explanatory variable. However this is strictly a procedure for testing the null of no-predictability. It does not provide a coefficient estimate for a long-horizon regression, and it is restricted to regressions with a single explanatory variable, so it is of little use for understanding the sources of long-horizon predictability.

Our method is easily extended to the analysis of panel data. We show how Fama-MacBeth style regressions involving multi-period returns can be transformed into regressions involving one-period returns. The transformation has a natural interpretation in terms of 'rolling portfolios' as used by Jegadeesh and Titman (1993). Our approach is likely to be particularly useful in the panel setting, where the complexity and size of the data precludes some of the more sophisticated methods for dealing with overlapping observations such as bootstrapping and the Hodrick (1992) procedure. We also show how to adapt the procedure to estimate the regression coefficients in an overlapping regression more accurately by making better use of the data available.

Section 2 develops the basic idea in the context of inference for a linear regression with overlapping observations. Section 3 presents results from Monte Carlo studies demonstrating the advantages of our approach. Section 4 shows how the transformation method can be used for Fama-MacBeth and panel regressions and Section 5 develops an improved estimator of the regression coefficients. Section 6 illustrates our approach with two empirical examples. The first example analyses the predictability of long-horizon US stock market returns and the second example analyses reversal in relative country stock index returns. Section 7 concludes.

2. LINEAR REGRESSION WITH OVERLAPPING OBSERVATIONS

Consider a regression in which the future k-period log return $y_{t,k}$ is regressed onto a l-dimensional row vector x_t of time t explanatory variables (this generally includes a constant term):

$$y_{t,k} = x_t \beta + u_{t,k} \,. \tag{1}$$

The k-period log return $y_{t,k}$ is the sum of k one-period log returns $y_{t,k} = r_{t+1} + r_{t+2} + \ldots + r_{t+k}$ (where r_t denotes the one-period log return). We require that these one-period returns be available to the researcher. This is generally the case, as multi-period returns are normally constructed from one-period returns, though our approach is not applicable in cases such as Hansen and Hodrick's (1980) study of the foreign exchange market.

The OLS parameter estimate in (1) can be expressed in terms of one-period returns. If $\{y_{t,k},\ x_t\}_{t\ =\ 1,\ldots,T\text{-}k}$ is a sample of T–k observations of k-period returns and explanatory variables, the OLS estimate of β in (1) is

$$\hat{\beta} = \left(\sum_{t=1}^{T-k} x_t \,' x_t \right)^{-1} \sum_{t=1}^{T-k} x_t \,' y_{t,k} = \left(\sum_{t=1}^{T-k} x_t \,' x_t \right)^{-1} \sum_{t=1}^{T-k} x_t \,' \left(r_{t+1} + \ldots + r_{t+k} \right). \tag{2}$$

Now define $\hat{x}_t$ as the sum of x_t and its first k – 1 lags: $\hat{x}_t = \sum_{i=0}^{k-1} x_{t-i}$, where we define $x_t = 0$, for $t \leq 0$ and for $t > T$–k. Then the OLS estimate of β can be written in terms of one-period returns:

$$\hat{\beta} = \left(\sum_{t=1}^{T-k} x_t \,' x_t \right)^{-1} \sum_{t=1}^{T-1} \hat{x}_t \,' r_{t+1} \,, \tag{3}$$

showing that the OLS estimate of β in (1) is a linear function of one-period returns.

In fact, the OLS estimate of β can be obtained from a regression in which the dependent variable is one-period returns. This can be seen most clearly using matrix notation. Define the matrix of original regressors **X** as the $(T\text{–}k) \times l$ matrix with the t^{th} row given by x_t, and define $\widehat{\mathbf{X}}$ as the $(T\text{–}1) \times l$ matrix with the t^{th} row given by $\widehat{x}_t$. We now construct a matrix of transformed regressors **Z** as

$$\mathbf{Z} = \widehat{\mathbf{X}}\left(\widehat{\mathbf{X}}'\widehat{\mathbf{X}}\right)^{-1}\mathbf{X}'\mathbf{X}, \tag{4}$$

and define **r** as a $(T\text{–}1)$-dimensional column vector with t^{th} element given by r_{t+1}.

The transformed regression of one-period returns onto transformed regessors:

$$\mathbf{r} = \mathbf{Z}\beta + \mathbf{u} \tag{5}$$

has an OLS parameter estimate given by $\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{r}$ and this can be expressed as

$$\begin{aligned}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{r} &= \left(\mathbf{X}'\mathbf{X}\left(\widehat{\mathbf{X}}'\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}'\widehat{\mathbf{X}}\left(\widehat{\mathbf{X}}'\widehat{\mathbf{X}}\right)^{-1}\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\left(\widehat{\mathbf{X}}'\widehat{\mathbf{X}}\right)^{-1}\widehat{\mathbf{X}}'\mathbf{r} \\ &= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\widehat{\mathbf{X}}'\mathbf{r}\end{aligned} \tag{6}$$

which is equivalent to the expression in (3), showing that the OLS estimate of β in the transformed non-overlapping regression (5) is numerically identical to that obtained from the overlapping regression in (1). Since this equivalence is exact it does not depend on the assumed data generating process. Of course, inference does depend on the assumed data generating process, and the next section develops inference procedures for the OLS parameter estimate based on various stationary data generating processes.

2.1 Inference on β

This section shows that several standard inference procedures applied to the transformed regression are valid asymptotically. We first derive a general expression for the asymptotic sampling distribution of the OLS parameter estimate assuming only stationarity, and we then derive the properties of specific estimators under more restrictive assumptions.

Consider the transformed regression (5) which we write in scalar notation as

$$r_{t+1} = z_t{}'\beta + u_{t+1} \qquad t = 1, 2, \ldots T-1, \tag{7}$$

where z_t' is the t^{th} row of the matrix **Z** defined above. Now suppose that r_t and u_t are stationary and ergodic random variables, that x_t is an l-dimensional stationary and ergodic random vector, and that $E[x_t'x_t]$ is non-singular. Note we do not make direct assumptions about z_t since it is a transformation of **X**. Now suppose we observe a sample of $\{r_{t+1}, x_t\}$ of size $T-1$. The OLS estimate of β for the transformed regression (5) is

$$\begin{aligned} \hat{\beta}_{T-1} &= \left(\sum_{t=1}^{T-1} z_t{}'z_t\right)^{-1} \sum_{t=1}^{T-1} z_t{}'r_{t+1} \\ &= \left(\sum_{t=1}^{T-1} z_t{}'z_t\right)^{-1} \sum_{t=1}^{T-1} z_t{}'\left(z_t\beta + u_t\right) \\ &= \beta + \left(\sum_{t=1}^{T-1} z_t{}'z_t\right)^{-1} \sum_{t=1}^{T-1} z_t{}'u_t \\ &= \beta + \left(\sum_{t=1}^{T-k} x_t{}'x_t\right)^{-1} \sum_{t=1}^{T-1} \hat{x}_t{}'u_t \end{aligned} \tag{8}$$

To derive the asymptotic distribution consider:

$$\frac{T-k}{\sqrt{T-1}}\left(\hat{\beta}_{T-1} - \beta\right) = \left(\frac{1}{T-k}\sum_{t=1}^{T-k} x_t{}'x_t\right)^{-1} \frac{1}{\sqrt{T-1}}\sum_{t=1}^{T-1} \hat{x}_t{}'u_t \ . \tag{9}$$

Given our assumptions of stationarity and ergodicity, it follows that

$$\frac{1}{T-k}\sum_{t=1}^{T-k} x_t{}'x_t \xrightarrow{\ p\ } Q \equiv E[x_t{}'x_t], \tag{10}$$

and that

$$\frac{1}{\sqrt{T-1}}\sum_{t=1}^{T-1} \hat{x}_t u_t \xrightarrow{\ D\ } N(0, \mathbf{S}), \tag{11}$$

where **S** is the long-run covariance matrix of $\hat{x}_t u_t$:

$$\mathbf{S} = \lim_{N \to \infty} \sum_{j=-N+1}^{N-1} E\left[\hat{x}_t ' u_t u_{t-j} \hat{x}_{t-j} \right]. \tag{12}$$

Standard asymptotic theory then tells us that the OLS estimate is distributed asymptotically normal:

$$\frac{T-k}{\sqrt{T-1}} \left(\hat{\beta}_{T-1} - \beta \right) \xrightarrow{D} N\left(0, Q^{-1} \mathbf{S} Q^{-1}\right). \tag{13}$$

Furthermore asymptotic normality holds when we replace the population moments with consistent estimators of them. So the asymptotic distribution is normal with mean β and an estimated asymptotic covariance matrix:

$$\text{Est. Asy. Var}\left[\hat{\beta} \right] = \frac{T-1}{(T-k)^2} \hat{Q}^{-1} \hat{S} \hat{Q}^{-1}, \tag{14}$$

where $\hat{S}$ is a consistent estimator of **S**, and

$$\hat{Q} = \frac{1}{T-k} \sum_{t=1}^{T-k} x_t ' x_t \tag{15}$$

The main advantage of the transformed regression is that estimation of **S** is more straightforward than in the usual case with overlapping observations. By making more specific assumptions we can examine specific estimators for **S**.

2.2 IID Returns

When one-period returns are serially uncorrelated and conditionally homoscedastic then conventional OLS standard errors from the transformed regression provide asymptotically correct inference. To see this note that if returns are serially uncorrelated and conditionally homoscedastic then a consistent estimate of **S** is

$$S_1 = \frac{1}{T-1} \left(\sum_{t=1}^{T-1} \hat{x}_t ' \hat{x}_t \right) \times \frac{1}{T-1} \sum_{t=1}^{T-1} \hat{u}_t^2 . \tag{16}$$

It follows that an asymptotically valid expression for the sampling variance of the coefficient estimate is obtained from inserting this expression for **S** into (14):

$$\text{Est. Asy. Var}\left[\hat{\beta}\right] = \frac{T-1}{(T-k)^2}\left[\frac{1}{T-k}\sum_{t=1}^{T-k} x_t\,'x_t\right]^{-1} S_1 \left[\frac{1}{T-k}\sum_{t=1}^{T-k} x_t\,'x_t\right]^{-1},$$

but this expression simplifies to the conventional covariance matrix of OLS parameter estimates obtained from the transformed regression,

$$\text{Est. Asy. Var}\left[\hat{\beta}\right] = \hat{\sigma}^2\left[\sum_{t=1}^{T-1} z_t\,'z_t\right]^{-1} \qquad (17)$$

where $\hat{\sigma}^2 = (T-1)^{-1}\sum_{t=1}^{T-1}\hat{u}_t^2$.

Note that the assumptions required for this result are weaker than those required by Richardson and Smith (1991). They require that the explanatory variables as well as the dependent variable be constructed from one-period returns that are serially independent and conditionally homoscedastic, whereas our only requirement of the explanatory variables is that they be stationary.

2.3 Heteroscedasticity

Similarly, when one-period returns are serially uncorrelated but not necessarily homoscedastic, the standard White heteroscedasticity-consistent standard errors from the transformed regression are correct asymptotically. Note that under these conditions we can estimate **S** consistently by

$$S_2 = \frac{1}{T-1}\sum_{t=1}^{T-1}\hat{u}_t^2\hat{x}_t\,'\hat{x}_t\ . \qquad (18)$$

Insertion of this estimator into (14) gives the following expression

$$\text{Est. Asy. Var}\left[\hat{\beta}\right] = \frac{1}{(T-k)^2}\left[\frac{1}{T-k}\sum_{t=1}^{T-k} x_t\,'x_t\right]^{-1}\sum_{t=1}^{T-1}\hat{u}_t^2\hat{x}_t\,'\hat{x}_t\left[\frac{1}{T-k}\sum_{t=1}^{T-k} x_t\,'x_t\right]^{-1}, \quad (19)$$

which can be rearranged to give White's (1980) heteroscedasticity-consistent covariance matrix of OLS parameter estimates from the transformed regression:

$$\text{Est. Asy. Var}\left[\hat{\beta}\right] = \left(\sum_{t=1}^{T-1} z_t\,' z_t\right)^{-1} \left(\sum_{t=1}^{T-1} \hat{u}_t^2 z_t\,' z_t\right)\left(\sum_{t=1}^{T-1} z_t\,' z_t\right)^{-1}. \tag{20}$$

This covariance estimator in the form of (19) was presented in Hodrick (1992, p.362) in equation (8). Hodrick (1992) finds this estimator has much better small sample properties than the Hansen and Hodrick covariance estimator, but he notes that it is only valid under the null of no predictability.

2.4 Heteroscedasticity and Autocorrelation

If returns are predictable then in general the errors from the transformed regression are not serially uncorrelated. This is the case even when the original regression model (1) is correctly specified so that the expectation of the *k*-period return conditional upon the time *t* information set I_t is a linear function of the explanatory variables:

$$E\left[y_{t,k} \,|\, I_t\right] = x_t \beta\,. \tag{21}$$

Thus when returns are predictable, consistency requires that we account for serial correlation in one-period returns.

The Newey-West estimator of long-run covariance is probably the most widely used heteroscedasticity and autocorrelation consistent (HAC) estimator. The Newey-West covariance estimate with *J* lags is

$$\begin{gathered} S_3 = \hat{\Gamma}(0) + \sum_{j=1}^{J} w(j,J)\left(\hat{\Gamma}(j) + \hat{\Gamma}(j)'\right), \quad \text{where} \\ \hat{\Gamma}(j) = \sum_{t=j+1}^{T-1} \hat{x}_t\,'\hat{u}_t\hat{u}_{t-j}\hat{x}_{t-j} \ \text{ and } \ w(j,J) = 1 - \frac{j}{J+1}. \end{gathered} \tag{22}$$

Inserting this estimate of long-run covariance into (14) gives an expression for the asymptotic covariance matrix of the parameter estimates:

$$\frac{T-1}{(T-k)^2}\left(\frac{1}{T-k}\sum_{t=1}^{T-k}x_t\,'x_t\right)^{-1}\left(\hat{\Gamma}(0)+\sum_{j=1}^{J}w(j,J)\left(\hat{\Gamma}(j)+\hat{\Gamma}(j)'\right)\right)\left(\frac{1}{T-k}\sum_{t=1}^{T-k}x_t\,'x_t\right)^{-1} \quad (23)$$

This expression can be rearranged to give the standard Newey-West HAC covariance matrix of the OLS parameter estimates in the transformed regression. To see this, denote an autocovariance term analogous to $\hat{\Gamma}(j)$ but in terms of *z*:

$$\hat{\Lambda}(j)=\sum_{t=j+1}^{T-1}z_t\,'\hat{u}_t\hat{u}_{t-j}z_{t-j}\ , \quad (24)$$

and note that the following relation holds between the two autocovariance terms:

$$\hat{\Gamma}(j)=\left(\sum\hat{x}_t\,'\hat{x}_t\right)\left(\sum x_t\,'x_t\right)^{-1}\hat{\Lambda}(j)\left(\sum x_t\,'x_t\right)^{-1}\left(\sum\hat{x}_t\,'\hat{x}_t\right).$$

Inserting this expression into (23) then gives an expression for the asymptotic sampling variance that is identical to the Newey-West HAC covariance matrix of the OLS parameter estimates in the transformed regression:

$$\frac{1}{T-1}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}z_t\,'z_t\right)^{-1}\left(\hat{\Lambda}(0)+\sum_{j=1}^{J}w(j,J)\left(\hat{\Lambda}(j)+\hat{\Lambda}(j)'\right)\right)\left(\frac{1}{T-1}\sum_{t=1}^{T-1}z_t\,'z_t\right)^{-1} \quad (25)$$

This estimator is simple, it is consistent, and it is guaranteed to be positive definite. Of course the same properties hold for the standard approach where the Newey-West covariance estimator is used in conjunction with the original regression specification (1). The advantage of our approach lies in its finite sample properties which are explored in the next section.

3. MONTE CARLO ANALYSIS

We have seen that inference based on the transformed regression (1) is valid asymptotically, but conventional methods are also valid asymptotically. This section shows that inference based on the transformed regression has better finite-sample properties than the conventional approaches.

We run Monte Carlo simulations from two sets of data-generating processes. The first set starts with a pair of uncorrelated persistent processes $\left\{x_t^1, x_t^2\right\}$ and then generates the one period returns $\{r_t\}$ and the multi-period returns $\{y_{t,k}\}$ using a linear model as in equation (1). The second is intended to capture the spirit of regressions of k-period returns on lagged k-period returns as in Fama and French (1988). We use a variety of values for k and for the length of the data, and we compare the performance of our procedures based on the transformed regression with the more conventional approaches based on Newey-West and Hansen-Hodrick estimators of long-run covariance applied to the overlapping data regression.

Andrews (1991) examines and compares a variety of other HAC estimators, which differ from Newey-West in their weighting function $w(j, J)$. We have done the simulations using the four other HAC estimators he considers (with kernel functions that are Truncated, Parzen, Tukey-Hanning and Quadratic Spectral) and using a comparable bandwidth to ensure they have the same asymptotic variance as the Newey-West estimator (see Andrews 1991 p 829). The results are substantially unaltered, and are not reported[1].

Our main finding from both types of simulation is that conventional OLS standard errors obtained from the transformed regression provide the most accurate small-sample inference for homoscedastic data generating processes. In the presence of heteroscedasticity, White's (1980) heteroscedasticity-consistent standard errors from the transformed regression provides the most accurate inference in small samples. When the forecast return horizon is long in comparison to the sample period, and when the regressors are strongly positively autocorrelated, the Newey-West and Hansen-Hodrick procedures produce standard errors that are severely biased downwards. To help understand the source of this bias we derive a simple expression for this bias (under restrictive conditions).

[1] The full tables with all the HAC estimators are available [on authors' website].

3.1 Simulations with two exogenous regressors

The results are for inference concerning the OLS parameter estimates in an overlapping regression as in (1). The models include an intercept and two exogenous regressors. The regressors are homoscedastic, stationary, mutually uncorrelated, AR(1) processes with unit variance and AR parameter 0.8. Table 1 presents results from simulations from a model where the errors are homoscedastic. In Panel A returns are unpredictable (at all horizons), whereas in Panel B returns are predictable (we choose coefficients such that the R^2 for one period returns is five percent). The model in the absence of predictability has one-period returns that are mean-zero independent normal random variables with variance one.

For each data generating process, and for each sample length and return horizon we present results for four conventional covariance estimators applied to the overlapping regression: 'OLS' the standard OLS covariance estimator, 'White' which is White's (1980) heteroscedasticity-consistent covariance estimator (both these estimators fail to account for induced serial correlation), 'NW' which uses the Newey-West HAC estimator of long-run covariance, and 'HH' which is the heteroscedasticity-consistent version of Hansen and Hodrick (1980). For NW we show results for lag lengths of k - the number of periods in the long-horizon return - and $2k$. Of course any fixed lag length is not consistent, but in any given sample a lag length must be chosen and the above choices correspond to those frequently reported in the empirical finance literature.

We then present results for covariance estimators based on the transformed regression. We consider the three estimators presented in the previous section: OLS, White, and Newey-West HAC covariance estimators applied to the transformed regression.

For each covariance estimator and each scenario we report the bias, standard deviation, and RMSE (root mean square error), as well as the true confidence levels of the nominal 99%, 95%, and 90% regression coefficient confidence intervals. 5000 simulations are used for each scenario.

Table 1 Panel A shows that the OLS and White estimators are severely biased down, as expected, since they fail to account for induced serial correlation. However the Newey-

West and Hansen-Hodrick estimators also exhibit a downward bias, which is particularly strong when the return horizon is long and the sample length short. For example with a forecast return horizon of 12 and 100 observations the downward bias in the NW and HH estimators is sufficiently large to result in the 99% confidence intervals from these estimators having coverage frequencies below 87%.

In contrast, the estimators based on the transformed regression have much better properties. In particular, the standard OLS estimator of covariance obtained from the transformed regression performs very well in this situation, exhibiting low bias and coverage frequencies that are close to their nominal levels. Note however that the Newey-West estimator applied to the transformed regression produces estimates that are also biased down quite substantially, though not by as much as the Newey-West estimator applied to the overlapping regression.

Table 1 Panel B reports results from simulations where the regressors and errors follow the same processes as in Panel A, but the actual returns are predictable with a one-period ahead R^2- of five percent. The results are broadly similar to those in Panel A. Again the procedures based on the transformed regression perform best, and the best performing estimator is again OLS applied to the transformed regression. Note however that when returns are predictable the OLS covariance estimator is biased down, in some cases substantially. We would expect some bias as OLS applied to the transformed regression ignores the serial correlation in one-period returns due to the predictability of returns. In such cases we might expect a HAC estimator such as Newey-West applied to the transformed regression to work better, but this is not the case.

Table 2 reports results from simulations where returns are predictable and where the errors are conditionally heteroscedastic. The standard deviation of the error is proportional to the deviation of the lead regressor from its mean. The presence of heteroscedasticity has a clear effect, significantly worsening the performance of the OLS covariance estimates obtained from the transformed regression. The White (heteroscedasticity consistent) covariance estimate obtained from the transformed regression performs quite well. As in Table 1 Panel B the presence of return predictability produces some bias in the White estimator as it does not account for serial correlation in errors. However the use of Newey-West to account for this serial correlation actually increases the bias, with the result that the best performing estimates in this case are the White covariance estimates obtained from the transformed regression.

The broad conclusions seem to be robust to the choice of parameters. In particular, if the regressors are less persistent (AR parameter of 0.1 rather than 0.8) Monte-Carlo simulations that are not reported here also show that the procedures performed on the transformed regressions work best, with the White estimate being the best in the presence of heteroscedasticity and the OLS estimate being best otherwise. As one would expect however, the difference between the estimates from the transformed and untransformed regressions is less marked than in the case of persistent regressors.

3.2 Analysis of bias in HAC variance estimators

Tables 1 and 2 show that both the Newey-West and Hansen-Hodrick estimators of variance are biased down, in some cases quite severely. The bias in the Newey-West estimator is also present when applied to the transformed regression. We can distinguish two distinct sources of bias in these estimators. The first source of bias relates to the non-unitary weights that are applied to sample autocorrelations. The true long-run covariance matrix is an unweighted sum of all autocovariance terms. The Newey-West estimator uses weights that are strictly less than one, and if the true autocovariances are positive then this understates the long-run covariance. Note that this argument does not apply to the Hansen-Hodrick estimator under the null hypothesis since it uses a truncated kernel with unit weights up to the truncation point. The other source of bias results from the use of estimated residuals rather than true errors in forming the sample autocovariances.

We can analyze the bias induced by the use of estimated residuals in more detail. Consider the difference between the 'true' error terms in a regression model and the (OLS) estimated errors or residuals in a generic regression of a T-vector $\mathbf{y}$ onto a T by l matrix $\mathbf{X}$:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u} . \qquad (26)$$

The estimated residuals are given by $\hat{\mathbf{u}} = \mathbf{M}\mathbf{u}$ where $\mathbf{M}$ is the idempotent error projection matrix:

$$\mathbf{M} \equiv \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' . \tag{27}$$

'True' errors are transformed into estimated residuals by **M**. This transformation also alters the covariance structure of errors. Denote the covariance matrix of true errors by Ω. Then the covariance of the estimated residuals is given by:

$$\operatorname{cov}(\hat{\mathbf{u}}) = \mathbf{M}\Omega\mathbf{M}. \tag{28}$$

When the covariance matrix of the true errors is proportional to the identity matrix:

$$\operatorname{cov}(\mathbf{u}) = \sigma^2\mathbf{I} \tag{29}$$

then the covariance matrix of the estimated residuals is simply proportional to the error projection matrix:

$$\operatorname{cov}(\hat{\mathbf{u}}) = \sigma^2\mathbf{M} . \tag{30}$$

The elements on the main diagonal of the error projection matrix **M** sum to $T - l$. This fact lies behind the commonly used degrees of freedom adjustment for estimation of the error variance in a linear regression. Unfortunately, no general rules (that we know of) apply to the summation of elements along the other diagonals. However, we can use the specific error projection matrix **M** associated with the regression under consideration, to obtain an estimate of the bias of the HAC covariance estimators, under the assumption that there is zero autocorrelation in true errors. All we need do is replace the terms $u_t u_{t-j}$ that appear in the expression for the HAC estimator with an estimate of the expectation of this term. Denote the element in row i column j of the matrix M by $m_{i,j}$. Then the estimate of bias in the Newey-West covariance estimator presented in (22) is

$$\text{Bias}\left(S_3\right) = \sum_{j=1}^{J} w(j,J)\left(\hat{\Gamma}(j) + \hat{\Gamma}(j)'\right), \text{ where}$$

$$\hat{\Gamma}(j) = \hat{\sigma}^2 \sum_{t=j+1}^{T-1} \hat{x}_t' m_{t,t-j} \hat{x}_{t-j} \ , \ \ w(j,J) = 1 - \frac{j}{J+1} \tag{31}$$

$$\text{and } \ \hat{\sigma}^2 = \left(T - j - 1\right)^{-1} \sum_{t=j+1}^{T-1} \hat{u}_t^2$$

Applying this approach to the Hansen-Hodrick estimator of long-run covariance results in estimates of bias that are close to those reported in Table 1. The bias in each of these autocovariance terms may be small, but the Newey-West and Hansen-Hodrick procedures involve the summation of many autocovariances and this can result in a substantial overall bias. Indeed, the bias observed in Table 1 for the Hansen-Hodrick procedure is removed if the true errors are used in place of the estimated residuals.

This analysis and the Monte Carlo studies shows the need for caution in using HAC covariance estimators in small samples. A desire to be conservative and account for positive autocorrelation leads many to use HAC covariance estimators, but in the absence of strong positive autocorrelation this can backfire and produce covariance estimates that are actually more downwardly biased (i.e. less conservative) than estimators that ignore autocorrelation. The advantage of the transformed regression approach is that it removes the autocorrelation induced by overlapping observations, thus reducing the cost of ignoring autocorrelation. We leave to future research whether it is possible to extend the above analysis to make finite-sample adjustments to HAC estimators to remove or ameliorate this finite-sample bias.

3.3 Fama-French style regressions

In this second set of simulations, one period log returns $\{r_t\}$ are generated as independent normal random variables. The k-period return at time t, $y_{t,k}$, is then regressed on the lagged k-period return, $y_{t-k,k}$. The experiment is in the spirit of Fama and French (1988).

It is natural to think of the period length as annual. For the simulations, we use $k = 5$ and 10 years, and the data length is 60 and 120 years. The results are reported in Table 3 which is set out in a similar way to Tables 1 and 2. We also report the analytic estimator of the standard error as proposed in Richardson and Smith (1991). In panel A, the volatility of returns is constant, while in panel B, volatility σ_t follows a mean reverting process. More specifically:

$$\ln \sigma_{t+1} = (1-\kappa)\alpha + \kappa \ln \sigma_t + \tilde{u}_t. \qquad (32)$$

In the simulation we take κ = 0.8, and the volatility of volatility is chosen so that the unconditional standard deviation of log volatility is 0.5.

The conclusions are very similar to the previous regressions. The OLS and White estimators are severely biased down, as expected, since they fail to account for induced serial correlation. The Newey-West and Hansen-Hodrick estimators also exhibit a downward bias, which is particularly strong when the return horizon is long and the sample length short. For example with regressions of ten year returns on lagged ten year returns using sixty years of data, the downward bias in the NW and HH estimators is sufficiently large to result in the 99% confidence intervals from these estimators having coverage frequencies below 72%.

In contrast, the estimators based on the transformed regression have much better properties. In particular, the standard OLS estimator of covariance obtained from the transformed regression performs very well in this situation, exhibiting low bias and coverage frequencies that are close to their nominal levels. Note however that the Newey-West estimator applied to the transformed regression produces estimates that are also biased down quite substantially, though not by as much as the Newey-West estimator applied to the overlapping regression.

The analytic estimator from Richardson and Smith also has little bias and has the advantage that it has a far lower root mean square error than the other estimators. Its coverage frequencies are close to their nominal levels. Panel B suggests that it may be slightly downward biased when the error term in the regression is heteroscedastic, but its coverage frequencies are at least as close to their nominal levels as is the White's estimator from the transformed regression, despite the assumption of homoscedasticity being violated.

4. FAMA-MACBETH AND PANEL REGRESSIONS

4.1 Fama-MacBeth Regressions

Regression analysis of panel data with overlapping observations is complicated by the need to account for both correlation between contemporaneous errors in the cross-section and the autocorrelation induced by overlapping observations. The Fama-MacBeth methodology is a simple approach that neatly accounts for cross-sectional correlation in errors. We can use our transformed regression approach in combination with the Fama-MacBeth methodology to also remove the autocorrelation induced by overlapping observations.

Consider a set (indexed by $t = 1,...,T$) of cross-sectional regressions of k-period log returns on N assets denoted by the N-vector $\mathbf{y}_{t,k}$ onto a matrix of explanatory variables $\mathbf{X}_t$:

$$\mathbf{y}_{t,k} = \mathbf{X}_t \beta_t + v_{t,k} . \tag{33}$$

In the Fama-MacBeth approach the cross-sectional OLS parameter estimates

$$\hat{\beta}_t = \left(\mathbf{X}_t{}'\mathbf{X}_t\right)^{-1} \mathbf{X}_t{}'\mathbf{y}_{t,k} \tag{34}$$

are averaged over time to give an estimate of the effect of the explanatory variables:

$$\bar{\beta} = T^{-1}\sum_{t=1}^{T}\hat{\beta}_t . \tag{35}$$

Inference is based on the long-run variance of $\hat{\beta}_t$. With non-overlapping observations, the conventional time-series *t* and *F* statistics are commonly used. In the case examined here, with overlapping observations, adjustments for autocorrelation are required and typically Newey-West standard errors are used for inference (for example, see Hong, Lee, and Swaminathan 2003). Given the results in the previous section and from Monte Carlo studies in other papers this is likely to be misleading.

The approach we propose, again focuses on the one-period returns that comprise the *k*-period returns. We first show how such a focus relates to the 'rolling-portfolio' approach popularised by Jegadeesh and Titman (1993), and we then show how the approach can be implemented using a set of transformed regressions.

Similar to the time-series case, we can express the average of the cross-sectional OLS parameter estimates in terms of one-period returns:

$$\bar{\beta} = T^{-1}\sum_{t=1}^{T}\left(\mathbf{X}_t{}'\mathbf{X}_t\right)^{-1}\mathbf{X}_t{}'(\mathbf{r}_{t+1} + \ldots + \mathbf{r}_{t+k}). \qquad (36)$$

Now define $\mathbf{U}_t{}' \equiv \left(\mathbf{X}_t{}'\mathbf{X}_t\right)^{-1}\mathbf{X}_t{}'$ for t = 1, …, T and equal to zero for t outside this range. Then we obtain:

$$\begin{aligned}\bar{\beta} &= T^{-1}\sum_{t=1}^{T}\mathbf{U}_t{}'(\mathbf{r}_{t+1} + \ldots + \mathbf{r}_{t+k}) \\ &= T^{-1}\sum_{t=2}^{T+k}(\mathbf{U}_{t-1}{}' + \mathbf{U}_{t-2}{}' + \ldots + \mathbf{U}_{t-k}{}')\mathbf{r}_t.\end{aligned} \qquad (37)$$

This approach is used by George and Hwang (2003) to study momentum based forecasts of annual and semi-annual returns. The expression in thesecond line of (37) has an interesting interpretation. Each column of $\mathbf{U}_t$ can be viewed as a set of portfolio weights. The entire term in the summation is thus the sum of the return at time t to k lagged sets of portfolio weights. This is exactly the ‘rolling-portfolio’ methodology as popularized by Jegadeesh and Titman (1993) and advocated in different contexts by Fama (1998).

Inference can be based on direct analysis of these rolling-portfolio returns. Alternatively we can implement this approach by application of standard Fama-MacBeth methodology to a set of transformed cros-sectional regressions. Define

$$\begin{aligned}\mathbf{A}_t &= \mathbf{U}_{t-1} + \mathbf{U}_{t-2} + \ldots + \mathbf{U}_{t-K}, \text{ and} \\ \mathbf{Z}_t &= \mathbf{A}_t\left(\mathbf{A}_t{}'\mathbf{A}_t\right)^{-1}.\end{aligned} \qquad (38)$$

The cross-sectional OLS regression of one-period returns onto Z gives an OLS parameter estimate that is identical to the ‘rolling portfolio return’ term inside the summation in (37):

$$\left(\mathbf{Z}_t'\mathbf{Z}_t\right)^{-1}\mathbf{Z}_t'\mathbf{r}_t = \mathbf{A}_t'\mathbf{r}_t. \tag{39}$$

Conventional time-series t and F statistics can then be used on these cross-sectional OLS parameter estimates.

4.2 Panel Regression

Now consider estimation of the panel regression

$$\mathbf{y}_{t,k} = \mathbf{X}_t\beta + v_{t,k} \qquad t = 1,2,...,T-k \tag{40}$$

where the notation is the same as in (34). The OLS estimate of the panel regression coefficient is

$$\hat{\beta} = \left(\sum_{t=1}^{T-k}\mathbf{X}_t'\mathbf{X}_t\right)^{-1}\sum_{t=1}^{T-k}\mathbf{X}_t'\mathbf{y}_{t,k}. \tag{41}$$

This can be rearranged to show the role of one-period returns, where the N-vector $\mathbf{r}_t$ denotes the one-period log returns on the N assets:

$$\hat{\beta} = \left(\sum_{t=1}^{T-k}\mathbf{X}_t'\mathbf{X}_t\right)^{-1}\sum_{t=1}^{T-k}\mathbf{X}_t'\left(\mathbf{r}_{t+1}+...+\mathbf{r}_{t+k}\right).$$

Define a variable $\hat{\mathbf{X}}_t = \sum_{i=0}^{k-1}\mathbf{X}_{t-i}$, where $\mathbf{X}_t = 0$, for $t \le 0$ and for $t > T - k$. The OLS panel coefficient estimate can be expressed in terms of one-period returns as

$$\hat{\beta} = \left(\sum_{t=1}^{T-k}\mathbf{X}_t'\mathbf{X}_t\right)^{-1}\sum_{t=1}^{T-1}\hat{\mathbf{X}}_t'\mathbf{r}_{t+1}. \tag{42}$$

Define the large block matrices:

$$\mathbf{X} = \begin{bmatrix}\mathbf{X}_1\\ \vdots \\ \mathbf{X}_{T-k}\end{bmatrix}, \hat{\mathbf{X}} = \begin{bmatrix}\hat{\mathbf{X}}_1\\ \vdots \\ \hat{\mathbf{X}}_{T-1}\end{bmatrix}, \text{ and } \mathbf{r} = \begin{bmatrix}\mathbf{r}_2\\ \vdots \\ \mathbf{r}_T\end{bmatrix}. \tag{43}$$

The OLS panel coefficient estimate is

$$\hat{\beta} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\hat{\mathbf{X}}'\mathbf{r}. \tag{44}$$

We can obtain this estimate by regressing one-period returns onto a set of transformed regressors **Z:**

$$\mathbf{r} = \mathbf{Z}\beta + \mathbf{u}, \tag{45}$$

where **Z** is defined as $\mathbf{Z} = \hat{\mathbf{X}}\left(\hat{\mathbf{X}}'\hat{\mathbf{X}}\right)^{-1}\mathbf{X}'\mathbf{X}$. The OLS coefficient estimate in the transformed regression is $\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{r}$. The matrix algebra detailed in (6) shows this is identical to the expression in (44), showing that the regression of one-period returns on transformed regressors gives the same estimate as OLS estimation of the original panel regression in (40). Inference can then be based on this transformed regression using the approaches detailed in Section 2. Of course it is likely that errors across stocks will be contemporaneously correlated and one might wish to use GLS techniques to estimate and account for this contemporaneous correlation. But this will be easier to do in the case of the transformed regression (45) since the autocorrelation induced by overlapping observations has been removed.

5. IMPROVED ESTIMATOR FOR β IN PREDICTIVE REGRESSIONS

The regression parameter β in an overlapping regression is a linear function of one-period returns. We can exploit this insight to increase the efficiency of the standard OLS estimation procedure by not discarding the ending data points. To see this, note that the regression parameter β is the sum of regression parameters from regressions of one-period returns at leads of 1,2,…, k onto the explanatory variables:

$$\beta \equiv \left(E\left[x_t{}'x_t\right]\right)^{-1} E\left[x_t{}'y_{t,k}\right]$$
$$= \left(E\left[x_t{}'x_t\right]\right)^{-1} \sum_{i=1}^{k} E\left[x_t{}'r_{t+i}\right] \tag{46}$$
$$= \sum_{i=1}^{k} \beta_i,$$

where β_i is the regression parameter for forecasting a one-period return i periods ahead:

$$\beta_i \equiv \left(E\left[x_t{}'x_t\right]\right)^{-1} E\left[x_t{}'r_{t+i}\right]. \tag{47}$$

Wwhen we run a standard OLS regression with overlapping observations, we implicitly estimate the sum of regression coefficients, as equation (3) makes clear. But consider estimating each of the coefficients with separate regressions. Start with estimation of β_1:

$$r_{t+1} = x_t\beta_1 + e_{t+1}. \tag{48}$$

The implicit estimation of this regression in equation (3) uses observations $t = 1,\ldots,T-k$. But there is no reason to exclude the ending observations $t = T-k+1,\ldots,T-1$. Typically the explanatory variables at these dates are known. The standard procedure uses $T-k$ observations for estimation of $\beta_1, \beta_2,\ldots, \beta_K$. The number of observations available to estimate β_i is $T-i$, so for all parameters apart from β_k the standard procedure throws data away.

Our proposed estimator is the sum of OLS estimates of $\beta_1, \beta_2,\ldots, \beta_k$, where those estimates use all the available data:

$$\breve{\beta} = \sum_{i=1}^{k} \hat{\beta}_i,$$
$$\hat{\beta}_i = \left[\sum_{t=1}^{T-i} x_t{}'x_t\right]^{-1} \sum_{t=1}^{T-i} x_t{}'r_{t+i}. \tag{49}$$

The available data varies for each of the components, but overall we are using more data. We can obtain this estimate by running a regression of one-period returns on transformed regressors. As in the previous section this approach simplifies inference. We first write $\hat{\beta}_i$ as

$$\hat{\beta}_i = \left(\sum_{t=i+1}^{T} x_{t-i}' x_{t-i}\right) \sum_{t=i+1}^{T} x_{t-i}' r_t . \quad (50)$$

Define a (T-1)-dimensional column vector **r** with the t^{th} element r_{t+1}. Define k matrices $\mathbf{D}_i$ (i = 1,…, k) of explanatory variables lagged i periods. Each matrix has T-1 rows with the t^{th} row given by x_{t-I} where $x_t = \mathbf{0}$ for $t \leq 0$. We can then write

$$\hat{\beta}_i = \left(\mathbf{D}_i ' \mathbf{D}_i\right)^{-1} \mathbf{D}_i ' \mathbf{r} . \quad (51)$$

Now define

$$\mathbf{U}' = \sum_{i=1}^{k} \left(\mathbf{D}_i ' \mathbf{D}_i\right)^{-1} \mathbf{D}_i ' . \quad (52)$$

It follows that we can write the proposed estimator in (49) as $\breve{\beta} = \mathbf{U}'\mathbf{r}$ and that we can obtain this estimate by regressing **r** onto a transformed matrix of regressors **Z** where

$$\mathbf{Z} = \mathbf{U}\left(\mathbf{U}'\mathbf{U}\right)^{-1} . \quad (53)$$

Inference can then be based on this transformed regression using the approaches described in Section II.

6. EMPIRICAL EXAMPLES OF THE METHODOLOGY

Overlapping regressions have been central to the debate over the predictability of stock market returns (Fama and French 1988; Campbell and Shiller 1988). To illustrate the relevance of our approach we conduct two empirical analyses. In the first we re-examine the issue of the predictability of long-horizon US stock market returns using Robert Shiller's data on stock market returns and earnings. In the second we illustrate our approach to Fama-MacBeth regressions by looking at the predictability of relative country stock returns.

It should be emphasised that the purpose of this analysis is to illustrate the approach to overlapping regressions we have developed in this paper rather than to cast new light on the debate over the predictability of stock prices. In particular, we have not attempted to allow for other econometric issues raised by the use of a highly persistent regressor, or the joint endogeneity of the dependent and independent variables.

6.1 US Stock Market Predictability

Taking data from Robert Shiller's website (http://www.econ.yale.edu/~shiller/data.htm) for annual US stock returns and price-earnings ratios from 1871 to 2004, we estimate the regression:

$$y_{t,k} = y_{t-k,k}\beta + u_{t,k} \,. \tag{54}$$

where $y_{t,k}$ is the log real return over years t+1 to t+k. The results are set out in Table 4.

Where the dependent variable is the ten-year return (k = 10), the standard approach, with either Newey-West or Hansen-Hodrick estimates of the covariance matrix, leads to severe underestimates of the standard error, with corresponding over-estimates of the t-statistics, in comparison to the analysis based on the transformed regression. This bias is observable also in each of the sub-periods. The coefficient on the lagged return, which appears to be significantly negative over the whole period and in the first half period, is indistinguishable from zero when using the transformed regression. For the five year return, the position is broadly similar except that the coefficient is not significantly different from zero in either the standard or the transformed regression except when looking at the first half of the period. The table also shows that making better use of the end-data, using the improved methodology described in Section 5, has a measurable effect on the results, though the data from the end of the period is not sufficiently different in this case from the rest as to alter the broad conclusions.

We now examine the predictability of long-period returns from the price earnings ratio. We consider the regression:

$$y_{t,k} = x_t\beta_1 + y_{t-k,k}\beta_2 + u_{t,k} \,. \tag{55}$$

where $y_{t,k}$ is the log real return over years t+1 to t+k, and x_t is the year t ratio of price to smoothed earnings. The results are set out in Table 5.

The results are similar to the previous regression in that the standard approach, with either Newey-West or Hansen-Hodrick estimates of the covariance matrix, leads to severe underestimates of the standard error, with corresponding over-estimates of the t-statistics, in comparison to the analysis based on the transformed regression. This bias is observable also in each of the sub-periods. The coefficient on the price earnings ratio is significantly negative at conventional significance levels both over the period as a whole, and in the second half, under both the standard approach and the transformed regression, but the standard errors are roughly doubled. This holds both for five and ten year rolling returns. The use of the full data set in the improved transformed regression brings the results to close to conventional significance levels.

The coefficient on lagged returns, which appears to be significantly positive in the second half of the period for both 5 and 10 year returns, and to be significantly negative in the first half for 10 year returns, turns out, to be insignificantly different from zero when using the transformed variables,.

6.2 Country Stock Returns

To illustrate our approach to Fama-MacBeth regressions, we analyze return predictability in international equity indices. Richards (1997) documents reversal in the relative returns of international equity indices. Countries that have done relatively well in the past period tend to under-perform their peers in the future. The reversal is strongest at the three year horizon. The finding is confirmed by Balvers, Wu and Gilliland (2000).

A natural way of exploring the predictability of relative country returns at different horizons is to follow the Fama-MacBeth procedure using country stock indices as the assets. Specifically we run the cross-sectional regression:

$$\mathbf{y}_{t,k} = \mathbf{y}_{t-k,k}\beta_t + v_{t,k} .$$

where $\mathbf{y}_{t,k}$ is the vector of k-month returns across different countries at month t, and then test whether the estimated slope coefficient differs from 0. The individual country returns are equity index returns less the return on the US market, over the period January 1982 to May 2007. The countries are Austria, Australia, Belgium, Canada, Denmark, Germany, Hong Kong, Ireland, Italy, Japan, Netherlands, Norway, Singapore, Sweden, Switzerland, and the UK. The data come from Datastream. The results are shown in Table 6.

The point estimate of beta is positive at the one year horizon. This is consistent with the findings of Bhojraj and Swaminathan (2003), and suggests some momentum in returns at shorter horizons. The value of beta goes negative at longer horizons, taking its largest negative values at around the six year horizon.

According to the untransformed regression, the beta is significantly less than zero for horizons of four years or more (the reason that no Hansen-Hodrick t-statistics are available beyond 6 years is that the covariance matrix is not positive definite). According to the transformed regression, however, the positive beta at the one year horizon is just significant, but at all other horizons the beta does not significantly differ from zero.

7. CONCLUSION

The main contribution of this paper is to show how a long-horizon overlapping regression can be transformed into a short-horizon non-overlapping regression, greatly simplifying inference. While the paper is written in terms of predictive (forecasting) regressions, nothing material changes when considering a regression in which the regressors are contemporaneous with the long-horizon return: we simply redefine the explanatory variables in (1) from being variables observed at time t to variables observed at time $t + k$. The transformation then proceeds exactly as in the predictive regression.

The transformation of a long-horizon regression into a short-horizon regression casts new light on the desirability of using long-horizon regressions. It strongly suggests that long-horizon regressions cannot have more statistical power than short-horizon regressions, since the former can be transformed into the latter without changing the estimated coefficients. Nonetheless there are strong economic reasons for considering longer horizons such as measurement error, adjustment costs, etc. Statistical inference is then made considerably easier and simpler by formulating the analysis in terms of long-horizon returns and then transforming to short-horizon returns. It is not immediately clear how to formulate a short horizon return regression so as to be informative about longer period returns, but the transformation presented here does just that.

References

Amihud, Y. and Hurvich, C.M. (2004), “Predictive Regressions: A Reduced-Bias Estimation Method”, *Journal of Financial and Quantitative Analysis,* 39(4), 813-841.

Andrews, D. (1991), “Heteroskedasticity and Autocorrelation Consistence Covariance Matrix Estimation”, *Econometrica*, 59(3), 817-858.

Ang, A., and Bekaert, G. (2007), “Stock Return Predictability: Is it there?”, *Review of Financial Studies*, 20(3), 651-707.

Balvers, R., Wu Y., and Gilliland, E. (2000), “Mean Reversion across National Stock Markets and Parametric Contrarian Investment Strategies”, *Journal of Finance*, 55(2), 745-772.

Bhojraj, S. and Swaminathan, B. (2006), “Macromomentum: Returns and Predictability in International Equity Indices”, *Journal of Business*, 76(3), 455-475.

Boudoukh, J., Richardson, M. and Whitelaw R.F. (forthcoming), “The Myth of Long-Horizon Return Predictability”, *Review of Financial Studies.*

Campbell, J., Lo, A., and MacKinlay, A. (1997), “The Econometrics of Financial Markets”, Princeton University Press, New Jersey.

Campbell, J., and Shiller, R. (1988), “The Dividend-price Ratio and Expectations of Future Dividends and Discount Factors”, *Review of Financial Studies*, 1, 195-227.

Campbell, J. and Yogo, M. (2006), “Efficient Tests of Stock Return Predictability”, *Journal of Financial Economics*, 81(1), 27-60.

Cochrane, J. (1991), “Volatility Tests and Efficient Markets: A Review Essay”, *Journal of Monetary Economics*, 27, 463-487.

Cochrane, J. and Piazzesi, M. (2007), “Bond Risk Premia”, *The American Economic Review*, 95(1), 138-160.

Daniel, K. and Marshall, D. (1997), "Equity-Premium and Risk-Free-Rate Puzzles at Long Horizons", *Macroeconomic Dynamics*, 1, 452-484.

Fama, E. (1998), "Market Efficiency, Long-term Returns, and Behavioral Finance", *Journal of Financial Economics*, 49, 283-306.

Fama, E., and French, K. (1988), "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics*, 22, 3-27.

George, T., and Hwang, C-Y., (2004), "The 52-week High and Momentum Investing", *Journal of Finance*, 59, 2145-2176.

Hansen, L., and Hodrick, R. (1980) "Forward Exchange Rates as Optimal Predictors of Future Spot Rates", *Journal of Political Economy*, 88, 829-853.

Hjmarlsson, E. (2006), "New Methods for Inference in Long-Run Predictive Regressions", *International Finance Discussion Papers No. 853,* Board of Governors of the Federal Reserve System.

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

Hong, D., Lee, C., and Swaminathan, B. (2003), "Earnings Momentum in International Markets", SSRN working paper http://ssrn.com/abstract=390107.

Jagannathan, R. and Wang, Y. (2005), "Consumption Risk and the Cost of Equity Capital", *NBER Working Paper No. 11026.*

Jagannathan, R. and Wang, Y. (2007), "Lazy Investors, Discretionary Consumption, and the Cross Section of Stock Returns", *Journal of Finance,* forthcoming.

Jegadeesh, N. (1991), "Seasonality in Stock Price Mean Reversion: Evidence from the US and the UK", *Journal of Finance*, 46, 1427-1444.

Jegadeesh, N., and Titman, S. (1993), "Returns to Buying Winners and Selling Losers: Implications for Market Efficiency", *Journal of Finance*, 45, 65-91.

Julliard, C. and Parker, J.A. (2005), "Consumption Risk and the Cross-Section of Expected Returns", *Journal of Political Economy*, 113(1), 185-222.

Malloy, C.J., Moskowitz, T.J. and Vissing-Jorgensen, A. (2005), "Long-run Stockholder Consumption and Asset Returns", working paper, GSB, University of Chicago.

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

Nelson, C., and Kim, M. (1993), "Predictable Stock Returns: the Role of Small Sample Bias", *Journal of Finance*, 48, 641-661.

Newey, W., and West, K. (1987), "A Simple, Positive Semi-definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55, 703-708.

Parker, J.A. (2001), "The Consumption Risk of the Stock Market", *Brookings Papers on Economic Activity*, 179-348.

Richards, A. (1997), "Winner-Loser Reversals in National Stock Market Indices: Can They Be Explained?", *Journal of Finance*, 52(5), 2129-2144.

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

Stambaugh, R., (1999), "Predictive Regressions", *Journal of Financial Economics*, 54, 375-421.

Valkanov, R. (2003), "Long Horizon Regressions: Theoretical Results and Applications", *Journal of Financial Economics*, 68, 201-232.

White, H. (1980), "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, 817-838.

Table 1 – Panel A: Monte Carlo Simulations. Regressors follow an AR1 process with AR parameter 0.8. No return predictability, homoscedastic errors.

No. Obs.	Forecast Horizon	Variance Estimator	Bias	Std. Dev.	RMSE	99%	95%	90%
			Overlapping Regression					
250	**3**	OLS	-0.364	0.058	0.369	89.5%	78.6%	69.6%
		White	-0.368	0.068	0.375	89.0%	77.7%	69.4%
		NW(k)	-0.148	0.148	0.210	96.9%	90.3%	83.6%
		NW(2k)	-0.116	0.175	0.210	97.3%	90.9%	84.6%
		HH	-0.055	0.204	0.212	97.9%	92.4%	86.7%
			Transformed Regression					
		OLS	-0.001	0.110	0.110	98.8%	95.0%	89.6%
		White	-0.006	0.132	0.132	98.7%	94.8%	89.2%
		NW(k)	-0.032	0.164	0.167	98.6%	93.7%	88.3%
			Overlapping Regression					
100	**3**	OLS	-0.994	0.263	1.028	89.5%	77.8%	69.5%
		White	-1.021	0.297	1.064	87.9%	76.3%	67.8%
		NW(k)	-0.521	0.588	0.785	94.7%	87.3%	80.3%
		NW(2k)	-0.484	0.674	0.829	94.5%	87.0%	80.3%
		HH	-0.345	0.786	0.858	95.1%	88.3%	81.9%
			Transformed Regression					
		OLS	-0.007	0.472	0.472	98.8%	94.7%	89.8%
		White	-0.041	0.566	0.567	98.6%	94.0%	88.7%
		NW(k)	-0.211	0.650	0.683	97.6%	91.4%	85.7%
			Overlapping Regression					
250	**12**	OLS	-4.631	0.315	4.642	71.0%	57.9%	49.8%
		White	-4.671	0.357	4.685	69.3%	56.5%	49.3%
		NW(k)	-1.964	1.822	2.679	94.1%	85.7%	78.6%
		NW(2k)	-1.842	2.169	2.846	93.8%	85.1%	78.5%
		HH	-1.469	2.526	2.922	94.4%	86.6%	79.6%
			Transformed Regression					
		OLS	-0.108	0.963	0.969	98.8%	94.9%	89.8%
		White	-0.155	1.152	1.162	98.6%	94.3%	89.3%
		NW(k)	-1.024	1.885	2.145	96.6%	90.7%	84.9%
			Overlapping Regression					
100	**12**	OLS	-11.244	1.478	11.341	70.8%	57.9%	50.3%
		White	-11.463	1.524	11.564	67.6%	55.4%	48.1%
		NW(k)	-7.176	5.271	8.904	86.6%	76.2%	68.4%
		NW(2k)	-7.746	5.532	9.518	83.2%	72.9%	65.1%
		HH	-7.334	6.591	9.861	80.3%	69.9%	63.0%
			Transformed Regression					
		OLS	0.006	4.206	4.206	98.9%	94.6%	89.0%
		White	-0.200	4.863	4.867	98.5%	93.9%	88.8%
		NW(k)	-5.167	5.426	7.493	92.5%	83.9%	77.1%

Table 1 – Panel B: Monte Carlo Simulations. Predictable Returns

No. Obs.	Forecast Horizon	Variance Estimator	Bias	Std. Dev.	RMSE	99%	95%	90%
			Overlapping Regression					
250	**3**	OLS	-0.346	0.058	0.351	90.1%	79.2%	71.2%
		White	-0.351	0.069	0.357	89.6%	78.6%	70.4%
		NW(k)	-0.130	0.150	0.198	97.2%	90.5%	84.1%
		NW(2k)	-0.097	0.179	0.204	97.4%	91.2%	84.8%
		HH	-0.037	0.208	0.211	97.9%	92.8%	86.7%
			Transformed Regression					
		OLS	0.015	0.110	0.111	99.0%	95.3%	90.2%
		White	0.010	0.134	0.134	98.8%	94.9%	89.8%
		NW(k)	-0.018	0.164	0.165	98.6%	94.1%	88.4%
			Overlapping Regression					
100	**3**	OLS	-1.064	0.261	1.095	88.4%	76.9%	68.3%
		White	-1.092	0.298	1.132	86.8%	75.4%	67.0%
		NW(k)	-0.594	0.585	0.834	94.3%	85.9%	79.5%
		NW(2k)	-0.563	0.673	0.877	93.7%	85.6%	79.3%
		HH	-0.423	0.781	0.889	94.4%	87.0%	80.8%
			Transformed Regression					
		OLS	-0.086	0.468	0.475	98.8%	94.0%	88.5%
		White	-0.124	0.553	0.567	98.5%	93.2%	87.6%
		NW(k)	-0.298	0.638	0.704	96.9%	90.6%	84.6%
			Overlapping Regression					
250	**12**	OLS	-5.171	0.349	5.183	71.3%	58.8%	50.3%
		White	-5.228	0.397	5.243	69.6%	57.2%	48.8%
		NW(k)	-2.082	2.045	2.918	93.2%	85.9%	79.0%
		NW(2k)	-1.973	2.431	3.131	92.7%	85.3%	78.6%
		HH	-1.556	2.880	3.273	92.7%	85.8%	79.4%
			Transformed Regression					
		OLS	-0.670	0.960	1.171	98.2%	92.9%	86.9%
		White	-0.727	1.163	1.372	97.9%	92.7%	86.6%
		NW(k)	-1.382	1.992	2.424	95.8%	89.3%	83.1%
			Overlapping Regression					
100	**12**	OLS	-13.749	1.661	13.849	68.9%	56.0%	48.4%
		White	-14.057	1.714	14.161	65.8%	53.1%	45.6%
		NW(k)	-8.981	5.996	10.798	84.2%	74.1%	66.5%
		NW(2k)	-9.785	6.187	11.577	79.9%	70.1%	62.6%
		HH	-9.446	7.462	12.038	75.9%	66.9%	60.4%
			Transformed Regression					
		OLS	-2.520	4.282	4.968	97.5%	91.0%	84.8%
		White	-2.789	5.022	5.744	96.8%	90.3%	83.8%
		NW(k)	-7.180	5.952	9.327	89.6%	80.7%	73.1%

Results from a Monte-Carlo simulation of an overlapping regression on two regressors. The two regressors are homoscedastic, stationary, mutually uncorrelated, first order autoregressive processes with unit variance and autoregressive parameter 0.8. The error term is homoscedastic. The dependent variable is *k*-period returns, where *k*, the forecast horizon is 3 or 12. The sample length is 100 or 250 periods. In Panel A the dependent variable is unpredictable, while in Panel B returns are predictable with a one-period ahead R^2 of five percent. The simulation is based on 5,000 runs. Four conventional covariance estimators (Ordinary Least Squares; White 1980; Newey-West 1987; and Hansen-Hodrick 1980) are used for the overlapping regression, and the first three are also used for the transformed regression. For each estimator, the bias in the estimate, its standard error and its root mean square error are shown, as well as the true confidence levels of 99%, 95% and 90% regression coefficient confidence intervals.

Table 2: Monte Carlo Simulations. Heteroscedastic errors and predictable returns.

No. Obs	Forecast Horizon	Variance Estimator	Bias	Std. Dev.	RMSE	99%	95%	90%
			Overlapping Regression					
250	**3**	OLS	-1.172	0.047	1.172	70.1%	57.0%	49.2%
		White	-0.922	0.187	0.941	85.8%	73.2%	65.2%
		NW(k)	-0.435	0.460	0.633	95.8%	88.2%	80.6%
		NW(2k)	-0.371	0.536	0.652	96.0%	89.0%	81.6%
		HH	-0.234	0.637	0.678	96.8%	90.7%	84.1%
			Transformed Regression					
		OLS	-0.835	0.090	0.839	89.7%	78.3%	70.1%
		White	-0.053	0.461	0.464	98.4%	93.8%	88.6%
		NW(k)	-0.165	0.525	0.551	97.9%	92.4%	86.5%
			Overlapping Regression					
100	**3**	OLS	-2.676	0.180	2.682	70.8%	58.5%	50.9%
		White	-2.223	0.533	2.286	82.7%	70.3%	62.2%
		NW(k)	-1.384	1.184	1.821	92.3%	82.8%	74.6%
		NW(2k)	-1.360	1.299	1.881	91.7%	82.0%	74.5%
		HH	-1.130	1.558	1.924	92.0%	83.3%	76.2%
			Transformed Regression					
		OLS	-1.849	0.353	1.882	90.1%	78.5%	71.2%
		White	-0.261	1.451	1.474	98.0%	92.4%	86.6%
		NW(k)	-0.794	1.381	1.593	95.9%	88.8%	81.9%
			Overlapping Regression					
250	**12**	OLS	-7.358	0.368	7.367	62.4%	49.6%	42.6%
		White	-7.159	0.550	7.180	65.5%	52.9%	45.5%
		NW(k)	-3.383	3.197	4.655	91.7%	82.7%	74.5%
		NW(2k)	-3.272	3.668	4.916	90.9%	81.9%	74.2%
		HH	-2.765	4.385	5.184	90.7%	82.2%	75.0%
			Transformed Regression					
		OLS	-3.055	1.196	3.280	95.2%	87.0%	80.0%
		White	-0.841	3.106	3.218	97.5%	92.2%	86.1%
		NW(k)	-2.346	3.482	4.199	94.6%	86.5%	79.6%
			Overlapping Regression					
100	**12**	OLS	-15.203	1.441	15.271	62.4%	50.1%	43.1%
		White	-15.126	1.642	15.215	63.0%	50.2%	43.5%
		NW(k)	-10.603	5.945	12.156	80.5%	69.4%	61.7%
		NW(2k)	-11.415	5.924	12.860	76.9%	65.2%	57.8%
		HH	-11.271	6.982	13.258	71.9%	61.7%	55.2%
			Transformed Regression					
		OLS	-5.349	4.734	7.143	95.2%	86.8%	80.1%
		White	-2.817	8.713	9.157	95.7%	88.7%	81.3%
		NW(k)	-8.474	6.899	10.927	87.5%	77.6%	69.2%

Results from a Monte-Carlo simulation of an overlapping regression on two regressors. The two regressors are homoscedastic, stationary, mutually uncorrelated, first order autoregressive processes with unit variance and autoregressive parameter 0.8. The error term is heteroscedastic, with the standard deviation of the error proportional to the deviation of the lead regressor from its mean. The dependent variable is *k*-period returns, where *k*, the forecast horizon is 3 or 12. The sample length is 100 or 250 periods. Returns are predictable with a one-period ahead R^2 of five percent. The simulation is based on 5,000 runs. Four conventional covariance estimators (Ordinary Least Squares; White 1980; Newey-West 1987; and Hansen-Hodrick 1980) are used for the overlapping regression, and the first three are also used for the transformed regression. For each estimator, the bias in the estimate, its standard error and its root mean square error are shown, as well as the true confidence levels of 99%, 95% and 90% regression coefficient confidence intervals.

Table 3 – Monte Carlo Simulations of Fama-French Regressions.

Panel A: Overlapping regression of long period returns on previous long period return where log prices are iid normal

No. Obs.	Forecast Horizon	Variance Estimator	Bias	Std. Dev.	RMSE	99%	95%	90%
			Overlapping Regression					
60	**10**	OLS	-0.151	0.012	0.152	59.2%	45.1%	37.8%
		White	-0.156	0.013	0.157	52.0%	39.1%	32.4%
		NW(k)	-0.130	0.035	0.135	71.6%	57.8%	49.5%
		NW(2k)	-0.139	0.033	0.143	64.3%	50.6%	43.0%
		HH	-0.143	0.030	0.146	61.0%	47.5%	40.1%
			Transformed Regression					
		OLS	-0.001	0.074	0.074	98.7%	93.7%	87.6%
		White	-0.007	0.083	0.083	98.2%	92.3%	85.5%
		NW(k)	-0.092	0.047	0.103	88.8%	76.8%	67.8%
			Richardson Smith Analytic Estimator					
			-0.010	0.000	0.010	99.7%	97.0%	91.1%
			Overlapping Regression					
120	**5**	OLS	-0.0222	0.0007	0.0222	82.8%	69.7%	61.1%
		White	-0.0230	0.0018	0.0231	79.8%	66.7%	58.2%
		NW(k)	-0.0125	0.0064	0.0141	92.7%	83.9%	77.1%
		NW(2k)	-0.0118	0.0084	0.0144	92.1%	83.7%	76.8%
		HH	-0.0132	0.0082	0.0156	91.0%	82.0%	75.0%
			Transformed Regression					
		OLS	0.0001	0.0053	0.0053	99.0%	94.5%	89.1%
		White	-0.0008	0.0075	0.0076	98.6%	93.7%	88.2%
		NW(k)	-0.0060	0.0074	0.0096	96.9%	90.5%	84.2%
			Richardson Smith Analytic Estimator					
			-0.0003	0.0000	0.0003	99.3%	95.2%	90.0%

Panel B: Overlapping regression of long period returns on previous long period return where log prices are normal with a standard deviation that is stochastic

No. Obs.	Forecast Horizon	Variance Estimator	Bias	Std. Dev.	RMSE	99%	95%	90%
			Overlapping Regression					
60	**10**	OLS	-0.160	0.017	0.161	58.8%	45.4%	38.5%
		White	-0.164	0.018	0.165	51.9%	39.9%	33.0%
		NW(k)	-0.137	0.044	0.144	71.5%	58.3%	50.0%
		NW(2k)	-0.145	0.043	0.151	65.5%	51.9%	44.1%
		HH	-0.149	0.040	0.155	62.2%	48.7%	41.5%
			Transformed Regression					
		OLS	0.005	0.107	0.107	98.9%	94.3%	88.0%
		White	-0.005	0.131	0.131	98.4%	92.3%	85.5%
		NW(k)	-0.095	0.116	0.116	90.5%	78.4%	68.6%
			Richardson Smith Analytic Estimator					
			-0.020	0.000	0.020	99.2%	96.0%	90.8%

			Overlapping Regression					
120	**5**	OLS	-0.0264	0.0011	0.0264	79.7%	66.8%	58.8%
		White	-0.0262	0.0034	0.0264	79.2%	66.3%	58.0%
		NW(k)	-0.0160	0.0090	0.0183	91.5%	82.4%	74.5%
		NW(2k)	-0.0160	0.0101	0.0190	90.4%	81.3%	73.8%
		HH	-0.0177	0.0096	0.0201	89.0%	79.2%	71.8%
			Transformed Regression					
		OLS	-0.0037	0.0068	0.0077	98.0%	92.9%	87.2%
		White	0.0003	0.0166	0.0166	98.4%	93.7%	88.4%
		NW(k)	-0.0082	0.0117	0.0143	96.7%	90.0%	83.2%
			Richardson Smith Analytic Estimator					
			-0.0045	0.0000	0.0045	98.7%	93.7%	87.6%

Results from a Monte-Carlo simulation of the regression is $y_{t,k} = \alpha + y_{t-k,k}\beta$ where $y_{t,k} = \sum_{u=0}^{k-1} \varepsilon_{t-u}\sigma_{t-u}$, and the ε_t are iid normal. In Panel A, the volatility σ_t is constant, while in Panel B log volatility follows an AR-1 process with auto-regressive parameter 0.8, and where the unconditional standard deviation of log volatility is 0.5. The forecast horizon, *k*, is 5 or 10 periods. The sample length is 60 or 120 periods. The simulation is based on 10,000 runs. Four conventional covariance estimators (Ordinary Least Squares; White 1980; Newey-West 1987; and Hansen-Hodrick 1980) are used for the overlapping regression, and the first three are also used for the transformed regression. The Richardson Smith (1991) analytic estimator ($\left(2k^2+1\right)\Big/3k\left(T-2k\right)$) is also computed. For each estimator, the bias in the estimate, its standard error and its root mean square error are shown, as well as the true confidence levels of 99%, 95% and 90% regression coefficient confidence intervals.

Table 4: Regression of US stock market returns on lagged returns

Dependent variable	Variance Estimator	1881-2004		1881-1942		1943-2004	
		β	t-stat	β	t-stat	β	t-stat
10-year log		*Overlapping Regression*					
Real return	NW(k)	-.272	-2.32	-.489	-3.36	-.155	-1.10
	NW(2k)		-2.65		-4.14		-1.61
	HH		-5.80		-4.88		-
		Transformed Regression					
	White	-.272	-1.22		-1.26		-0.62
		Improved Transformed Regression					
	White	-.313	-1.39	-.451	-1.14	-.260	-0.97
5-year log		*Overlapping Regression*					
Real return	NW(k)	-.132	-0.93	-.366	-2.66	.141	0.65
	NW(2k)		-0.85		-2.71		0.65
	HH		-0.86		-3.49		0.60
		Transformed Regression					
	White	-.132	-0.67		-1.32		0.53
		Improved Transformed Regression					
	White	-.141	-0.73	-.360	-1.31	.112	0.42
1-year log		*Standard Regression*					
Real return	White	.022	0.23	-.008	-.07	.053	0.34

The table estimates the regression $y_{t,k} = y_{t-k,k}\beta + u_{t,k}$ where *y* is the *k*-year log real return on the S&P index. The estimate of beta and its t-statistic are shown for 10, 5 and 1 year returns, for a variety of different time periods. The estimates based on overlapping regressions apply OLS to the data as is, and use Newey-West and Hansen-Hodrick to estimate the covariance matrix. The transformed regression uses the methodology described in this paper and uses White to estimate the covariance matrix. The improved transformed regression uses the methodology described in section 5 of this paper, which makes fuller of all the annual return data. The last row of each panel shows the non-overlapping case where annual returns are regressed annually. The data are from http://www.econ.yale.edu/~shiller/data.htm.

Table 5: Regression of US stock returns on price earnings ratio and lagged return.

Panel A: Coefficient on the price earnings ratio

Independent variable	Variance Estimator	1881-2004		1881-1942		1943-2004	
		β	t-stat	β	t-stat	β	t-stat
10-year log				*Overlapping Regression*			
Real return	NW(k)	-.075	-3.20	-.012	-0.80	-.133	-4.83
	NW(2k)		-2.78		-0.90		-5.17
	HH		-2.72		-0.81		-4.60
				Transformed Regression			
	White	-.075	-2.12		-0.20		-2.37
				Improved Transformed Regression			
	White	-.055	-1.68	-.022	-0.41	-.091	-2.02
5-year log				*Overlapping Regression*			
Real return	NW(k)	-.030	-3.02	-.036	-1.50	-.035	-4.18
	NW(2k)		-2.79		-1.46		-3.89
	HH		-2.54		-1.34		-3.69
				Transformed Regression			
	White	-.030	-2.28		-1.25		-2.26
				Improved Transformed Regression			
	White	-.029	-2.02	-.038	-1.30	-.033	-2.04
1-year log				*Standard Regression*			
Real return	White	-.0056	-2.48	-.0121	-2.31	-.0043	-1.81

Panel B: coefficient on lagged returns

Dependent variable	Variance Estimator	1881-2004		1881-1942		1943-2004	
		β	t-stat	β	t-stat	β	t-stat
10-year log				*Overlapping Regression*			
Real return	NW(k)	.246	1.30	-.390	-2.34	.675	2.91
	NW(2k)		1.14		-3.28		3.07
	HH		1.22		-2.86		2.80
				Transformed Regression			
	White	.246	0.74		-0.58		1.60
				Improved Transformed Regression			
	White	.137	0.43	-.291	-0.46	.467	1.19
5-year log				*Overlapping Regression*			
Real return	NW(k)	.177	1.10	-.006	-0.27	.573	3.09
	NW(2k)		0.96		-0.26		2.73
	HH		0.96		-0.29		2.53
				Transformed Regression			
	White	.177	0.69		-0.14		1.70
				Improved Transformed Regression			
	White	.152	0.60	-.040	-0.10	.524	1.65
1-year log				*Standard Regression*			
Real return	White	.073	0.75	.111	0.82	.084	0.54

The table estimates the regression $y_{t,k} = x_t\beta_1 + y_{t-k,k}\beta_2 + u_{t,k}$ where y is the k-year log real return on the S&P index, and x is the ten year rolling price earnings ratio. The estimate of beta and its t-statistic are shown for 10, 5 and 1 year returns, for a variety of different time periods. The estimates based on overlapping regressions apply OLS to the data as is, and use Newey-West and Hansen-Hodrick to estimate the covariance matrix. The transformed regression uses the methodology described in this paper and uses White to estimate the covariance matrix. The improved transformed regression uses the methodology described in section V of this paper, which makes fuller of all the annual return data. The last row of each panel shows the non-overlapping case where annual returns are regressed annually. The data are from http://www.econ.yale.edu/~shiller/data.htm.

Table 6: Cross-country estimates of auto-correlation in stock returns

Horizon (years)	Mean beta	t-statistic using: Untransformed regression			Transformed regression
		NW(k)	NW(2k)	HH	
1	0.123	1.72	1.59	1.44	1.98
2	0.041	0.39	0.36	0.32	0.52
3	-0.083	-1.01	-1.27	-1.10	-0.84
4	-0.179	-3.12	-3.83	-4.69	-1.59
5	-0.172	-3.20	-4.28	-3.41	-1.25
6	-0.240	-5.36	-7.30	-6.03	-1.61
7	-0.212	-5.54	-9.05	n/a	-1.47
8	-0.174	-6.85	-9.73	n/a	-1.22
9	-0.139	-2.78	-3.92	n/a	-0.84
10	-0.096	-7.32	-10.33	n/a	-0.50

The table shows the estimate of the regression coefficient of long horizon country index returns lagged returns. The basic regression is $\mathbf{y}_{t,k} = \mathbf{y}_{t-k,k}\beta_t + v_{t,k}$ .where $\mathbf{y}_{t,k}$ is the vector of *k*-month log returns (in excess of the US log return) across 22 different countries at month *t*, and the slope estimates are then pooled and tested for whether the mean ('mean beta') differs from zero. Since the regressions are done each month, the data are overlapping, so the t-statistics are adjusted for autocorrelation using a Newey-West (NW) or Hansen-Hodrick (HH) procedure. As an alternative the regression is transformed as described in the text and the standard error calculated from the transformed regression ('transformed regression'). The data are from Datastream.