LET'S DO IT AGAIN: BAGGING EQUITY PREMIUM PREDICTORS

ERIC HILLEBRAND, TAE-HWY LEE, AND MARCELO C. MEDEIROS

ABSTRACT. The literature on excess return prediction has considered a wide array of estimation schemes, among them unrestricted and restricted regression coefficients. We propose bootstrap aggregation (bagging) as a means of imposing parameter restrictions. In this context, bagging results in a soft threshold as opposed to the hard threshold that is implied by a simple restricted estimation. We show analytically that the resulting forecast has lower variance than the forecast that results from a simple restricted estimator. In an empirical application using the same data set as in Campbell and Thompson (2008), "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?" forthcoming in the Review of Financial Studies, we show that the resulting forecasts have more predictive power than those resulting from simple parameter restrictions.

KEYWORDS: bagging, equity premium, return prediction, restricted estimation, restricted forecasting

1. INTRODUCTION

Excess returns prediction has attracted academics and practitioners for many decades since the early 1920s, when Dow (1920) studied the role of dividend ratios as a possible predictor for returns. In the 1980s, a number of authors presented empirical evidence of ex-post (in-sample) return predictability. Fama and Schwert (1977), Fama and Schwert (1981), Rozeff (1984), Keim and Stambaugh (1986), Campbell (1987), Campbell and Shiller (1988a,b) and Fama and French (1988, 1989) showed that excess returns could be successfully predicted based on lagged values of variables such as dividend-price ratio and dividend yield, earnings-price ratio and dividend-earnings ratio, interest rates and spreads, inflation rates, book-to-market ratio, volatility, investment-capital ratio, consumption, wealth, and income ratio, and aggregate net or equity issuing activity.

Subsequent work, however, demonstrated that these results do not hold during the bull market period of the 1990s; see Lettau and Ludvigson (2001) or Schwert (2002). For example, during this period when stock prices soared, the dividend yield systematically drifted downwards, thus generating negative sample correlation between returns and dividend yield, contrary to the positive historical correlation. Furthermore, since early results concerned only ex-post predictability, they were of little practical interest. Studies of ex-ante (out-of-sample) return predictability have found either that previous successful results were restricted to particular sub-samples (Pesaran and Timmermann 1995) or

Date: December 14, 2008.

that return predictability was a statistical illusion; see Bossaerts and Hillion (1999). In addition, several authors pointed out that the apparent predictability of stock returns might be spurious as many of the predictor variables were highly persistent, leading to possibly biased coefficients and incorrect t-tests in predictive regressions; see, for example, Nelson and Kim (1993), Cavanagh, Elliot, and Stock (1995), and Stambaugh (1999). These problems are exacerbated when large numbers of variables are considered and only results that are apparently statistically significant are reported; see Foster, Smith, and R.E.Whaley (1997) and Ferson, Sarkissian, and Simin (2003).

The inconclusive evidence has inspired the use of time-varying regression models. As pointed out by Pesaran and Timmermann (2002) and Timmermann (2007) "forecasters of stock returns face a moving target that is constantly changing over time. Just when a forecaster may think that he has figured out how to predict returns, the dynamics of market prices will, in all likelihood, have moved on, possibly as a consequence of the forecaster's own efforts." On the other hand, alternative econometric methods were advocated for correcting the above mentioned bias and conducting valid inference; Cavanagh, Elliot, and Stock (1995), Mark (1995), Kilian (1999), Ang and Bekaert (2006), Jansson and Moreira (2006), Lewellen (2004), Torous, Valkanov, and Yan (2004), Campbell and Yogo (2006), and Polk, Thompson, and Vuolteenaho (2006).

More recently, Goyal and Welch (2008) argued that none of the conventional predictor variables proposed in the literature seems capable of systematically predicting stock returns out-of-sample. Their empirical evidence suggests that most models were unstable or spurious, and most models are no longer significant even in-sample. The authors show that the earlier apparent statistical significance was especially confined to the years of the Oil Shock of 1973–1975; see also Butler, Grullon, and Weston (2006).

Campbell and Thompson (2008), on the other hand, show that many predictive regressions outperform the historical average return once weak restrictions are imposed on the signs of coefficients and return forecasts. The out-of-sample explanatory advantage over the historical mean is small and usually statistically not significant, but nonetheless economically meaningful for mean-variance investors.

Our contribution is a new application of bagging as a means of imposing parameter restrictions. Bagging in this context results in a soft threshold as opposed to the hard threshold that is implied by a simple restricted estimation. We show that the resulting forecast has lower variance than the forecast that results from a simple restricted estimator. In an empirical application using the same data set as in Campbell and Thompson (2008), we show that the resulting forecasts have more predictive power than those resulting from simple parameter restrictions.

The paper is organized as follows. Section 2 specifies the forecast equation, defines the estimators and forecasts and presents our bagging approach to restricted parameter estimation. Section 3 presents our main theoretical result of variance reduction compared to simple restricted estimation. Section 5 describes the data set. Section 6 presents the main empirical results. Section 7 concludes.

2. FORECASTING WITH PARAMETER RESTRICTIONS

2.1. **Forecast Equation, Estimators, and Forecasts.**

The forecast equation is a univariate regression

$$y_{t+1} = \alpha + \beta x_t + \varepsilon_t, \quad t = 1, \ldots, T-1, \tag{1}$$

where y_t is the excess return, x_t is the predictor variable, and ε_t is the error term. Following Campbell and Thompson (2008), we will consider monthly and annualized returns. For the independent variable x_t, we use the predictor list in Section 5.

To fix ideas, we define the following estimators of a parameter θ:

(1) Unrestricted OLS estimator $\tilde{\theta}$, for example,

$$\tilde{\theta} = \tilde{\alpha}, \text{ or } \tilde{\theta} = \tilde{\beta}.$$

(2) Restricted estimator for θ subject to a lower bound $\theta_1 \leq \theta$:

$$\bar{\theta} := \max\{\tilde{\theta}, \theta_1\}.$$

(3) Gordon and Hall (2008) propose the estimator

$$\hat{\theta} := \frac{1}{B}\sum_{j=1}^{B} \max\{\tilde{\theta}^{*(j)}, \theta_1\} = \frac{1}{B}\sum_{j=1}^{B} \bar{\theta}^{*(j)} = \hat{\mathbb{E}}(\max\{\tilde{\theta}^*, \theta_1\}|\mathcal{X}) \tag{2}$$

for the situation where a lower bound θ_1 is known. Here, $\mathcal{X}$ is the available data set, $\mathcal{X}^*$ is a bootstrap sample, and $\tilde{\theta}^*$ is a bootstrap replication of $\tilde{\theta}$ from $\mathcal{X}^*$. There are B such bootstrap replications. Gordon and Hall (2008) show that subject to regularity conditions (see proof of Proposition 1), the asymptotic variance of $\hat{\theta}$ is smaller than that of $\bar{\theta}$ if the population parameter θ_0 coincides with the boundary θ_1.

Based on these estimators, we define the following forecast schemes:

(1) Historical mean forecast:

$$\check{y}_{t+1} := \frac{1}{t}\sum_{i=1}^{t} y_i.$$

(2) **UF** (unrestricted forecast) using $\tilde{\theta}$:

$$\tilde{y}_{t+1} := \tilde{\alpha} + \tilde{\beta} x_t.$$

UF is used in Goyal and Welch (2008).

(3) **PC** (forecast with positive coefficient restriction for β) using $\bar{\theta}$:

$$\bar{y}_{t+1} := \bar{\alpha} + \bar{\beta} x_t.$$

PC is used in Campbell and Thompson (2008).

(4) **PC-GH** (forecast with positive coefficient restriction for β) using the Gordon and Hall (2008) estimator $\hat{\theta}$:

$$\hat{y}_{t+1} := \hat{\alpha} + \hat{\beta} x_t.$$

(5) **PF** (forecast with positivity restriction):

$$y_{t+1}^{PF} := \mathbb{1}_{\{\tilde{y}_{t+1} > 0\}} \tilde{y}_{t+1}.$$

(6) **PCF** (forecast with joint positivity restriction and positive coefficient restriction)

$$y_{t+1}^{PCF} := \mathbb{1}_{\{\bar{y}_{t+1} > 0\}} \bar{y}_{t+1}.$$

(7) **PF-GH** (forecast with positivity restriction) using the Gordon and Hall (2008) estimator $\hat{\theta}$:

$$y_{t+1}^{PF-GH} := \mathbb{1}_{\{\hat{y}_{t+1} > 0\}} \hat{y}_{t+1}.$$

2.2. **Bagging Scheme.** The idea of bagging was introduced in Breiman (1996)and studied more rigorously in Bühlmann and Yu (2002). It has been shown in a number of studies that bagging can reduce the mean squared error of forecasts considerably by averaging over the randomness of variable selection (Inoue and Kilian 2008, Lee and Yang 2006). Applications include financial volatility (Huang and Lee 2007b, Hillebrand and Medeiros forthcoming), equity premia (Huang and Lee 2007a, Rapach, Strauss, and Zhou 2008), short-term interest rates (Audrino and Medeiros 2008), and employment data (Rapach and Strauss 2007).

We bootstrap-average over the forecast schemes in the next step of our proposal:

(1) Compute the historical mean forecast $\check{y}_{t+1}$ using all available observations on the excess return.
(2) Run the unrestricted forecast regression (1) and compute the unrestricted forecast $\tilde{\theta}$.
(3) Apply a bootstrap scheme to obtain B bootstrap replications of the estimated unrestricted parameters $\tilde{\alpha}^*$, $\tilde{\beta}^*$ and the forecast $\tilde{y}_{t+1}^*$.
(4) For bootstrap replications $\tilde{\beta}^*$ of the estimated parameter that have the correct sign, store the forecast $\tilde{y}_{t+1}^*$. For those that do not have the correct sign, replace the forecast by the historical mean forecast $\check{y}_{t+1}$. (This is equivalent to computing the restricted forecast $\bar{y}_{t+1}^*$.)
(5) Compute the bagged restricted coefficient forecast **B-PC** as the mean over the B obtained forecasts $\bar{y}_{t+1}^*$.
(6) Analogously, compute the bagged forecast with positivity restriction **B-PF** by averaging over B bootstrap replications of y_{t+1}^{PF}. Bootstrap replications of the forecast that are negative are replaced by zero, not by the historical mean.
(7) Analogously, compute the forecast with positivity restriction *and* with sign restriction on the coefficient **B-PCF** by averaging over B bootstrap replications of y_{t+1}^{PCF}. Forecasts that violate either restriction are replaced by zero.

We employ and compare five different bootstrap methods in the bootstrap aggregation scheme.

(1) **I.I.D. Bootstrap:** In the simplest bootstrap scheme, we draw B bootstrap samples from the pairs (y_{t+1}, x_t) with replacement and estimate equation (1) on each bootstrap sample.
(2) **Moving Block Bootstrap:** We apply the moving block bootstrap (Künsch 1989, Politis and Romano 1994, Hall, Horowitz, and Jing 1995) with a block length of 12 months. Experimenting with different block sizes did not lead to substantially different results.
(3) **Parametric Bootstrap:** For the parametric bootstrap, we estimate the forecast equation (1) and obtain the estimated error series $\hat{\varepsilon}_t$. We then draw bootstrap samples $\hat{\varepsilon}_t^*$ with replacement from this time series and compute the bootstrap sample $y_{t+1}^* = \hat{\alpha} + \hat{\beta} x_t + \hat{\varepsilon}_t^*$.
(4) **Wild Bootstrap:** We use a two-point wild bootstrap based on the divine proportion following Li and Wang (1998). The forecast equation (1) is estimated and the estimated error series $\hat{\varepsilon}_t$ obtained. The bootstrap sample $\hat{\varepsilon}_t^*$ is generated by setting

$$\hat{\varepsilon}_t^* = \begin{cases} \frac{1-\sqrt{5}}{2}\hat{\varepsilon}_t \text{ with probability } \frac{\sqrt{5}+1}{2\sqrt{5}}, \\ \frac{1+\sqrt{5}}{2}\hat{\varepsilon}_t \text{ with probability } \frac{\sqrt{5}-1}{2\sqrt{5}}. \end{cases}$$

The bootstrap sample $y_{t+1}^* = \hat{\alpha} + \hat{\beta} x_t + \hat{\varepsilon}_t^*$ is then generated as in the parametric bootstrap.

2.3. **Shrinkage.** The sign restrictions on the coefficients of the predictors in equation (1) can be understood as a shrinkage procedure with critical value $c = 0$. We avoid the term pre-testing because in the return prediction problem the regressand has low persistence and the regressors have high persistence. This renders standard inference invalid (Jansson and Moreira 2006, Torous, Valkanov, and Yan 2004). We relax this somewhat restrictive critical value $c = 0$ and consider two additional critical values, $c = \sqrt{2}$ and $c = \sqrt{\log(T)}$. The choice of these critical values is motivated by the shrinkage representation of forecast regression models proposed in Stock and Watson (2005). $c = \sqrt{2}$ corresponds to selecting regressors according to AIC; $c = \sqrt{\log(T)}$ corresponds to selecting regressors according to BIC.

3. MAIN RESULTS

The Gordon and Hall (2008) estimator is a bagging approach by construction. Both i.i.d. and moving block bootstrap methods can be used for this estimator. We show the applicability of their result to the coefficient of interest β and to the forecast $\mathbb{E}(y_{t+1}|x_t)$ in Proposition 1. In Proposition 2, we establish the equivalence of the PC-GH forecast $\hat{y}_{t+1}$ with B-PC, which averages the positive coefficient forecast $\bar{y}_{t+1}$. By Proposition 2, the bagged positive coefficient forecast B-PC is more efficient than the forecast $\bar{y}_{t+1}$ from simple restricted estimation.

ASSUMPTION 1. *Consider Equation* (1). *Make the following assumptions on the error process and on the regressor series.*

(1) *The* ε_t *have mean zero and variance* $\sigma^2 = Var(\varepsilon_t) < \infty$.

(2) *The ε_t satisfy the Lyapunov-condition $\mathbb{E}|\varepsilon_t|^{2+\epsilon} \leq C$ for some $C, \epsilon > 0$.*
(3) *The regressor time series has finite mean $\mathbb{E}x$ and variance $Var(x)$.*

The third assumption is critical in the presence of highly persistent regressors. We assume that despite their slow mean-reversion, which possibly even includes long memory, the regressors are covariance-stationary. Under these assumptions, the following Proposition shows the applicability of the Gordon and Hall (2008) bagging estimator to the parameter β and to the forecast $\mathbb{E}(y_{t+1}|x_t)$. The proof is provided in Appendix A.

PROPOSITION 1. *[Gordon and Hall (2008)]*
Let $\tilde{\theta}$ denote the least-squares estimator of θ. Let Z be a standard normal random variable with probability density function $\phi(z)$ and cumulative distribution function $\Phi(z)$. Consider the parameters $\theta = \beta$ and $\theta = \mathbb{E}(y_{T+1}|x_T)$ subject to the positivity restriction $\theta > \theta_1 = 0$. Let θ_0 denote the population parameter.

(1) *Case $\theta_0 > 0$. Then, the estimator $\hat{\theta}$ defined in Equation (2) follows $\hat{\theta} = \tilde{\theta} + O(T^{-1})$.*
(2) *Case $\theta_0 = 0$. Then, $T^{\frac{1}{2}}\tau^{-1}(\hat{\theta} - \theta_0)$ converges in distribution to the random variable $Z\Phi(Z) + \phi(Z)$, where τ^2 is the variance of the estimator $\tilde{\theta}$.*

The case where the constraint $\theta \geq 0$ is binding is the interesting case in terms of variance reduction. The asymptotic distribution of the simple constrained estimator $\bar{\theta}$ is a standard normal truncated to the positive half-line and thus has variance $(1 - 1/\pi)/2 \approx 0.3408$. The distribution of $Z\Phi(Z) + \phi(Z)$ has variance $1/3 + \sqrt{3}/(2\pi) - 1/\pi \approx 0.2907$. Thus, in the binding case, $\hat{\theta}$ has about 15% less variance than $\bar{\theta}$.

PROPOSITION 2. *Bagging the positive coefficient forecast $\bar{y}_{t+1}$ (B-PC) is equivalent to computing the forecast $\hat{y}_{t+1}$ from the Gordon and Hall (2008) estimator $\hat{\theta}$ (PC-GH).*

The proof is provided in Appendix B. Proposition 2 shows how and why bagging improves the positivity-restricted forecast $\bar{y}_{t+1}$, since the variance reduction result from Proposition 1 carries over.

4. SIMULATION

In order to evaluate the performance of the GH and restricted estimators, we consider the simulation experiment described bellow.

(1) For $T = 100, \ldots, 200$ do the following:
   (a) generate T observations of
$$\begin{aligned} \log(x_t) &= \gamma \log(x_{t-1}) + e_t, \;\; e_t \sim \mathsf{NID}(0, 0.04) \\ y_t &= 0.01 + \beta x_t + u_t, \;\; u_t \sim \mathsf{NID}(0,1), \text{ and } \mathbb{E}(u_t e_s) = 0, \, \forall t, s; \end{aligned} \tag{3}$$
   (b) estimate β using the unrestricted, restricted, and the Gordon-Hall estimators. The Gordon-Hall estimator is computed over 200 bootstrap samples;

(c) using each of the above estimators, compute the unrestricted, restricted, and Gordon-Hall forecasts of y_{T+1}.

(2) Repeat the steps above over 1000 Monte Carlo replications.

We consider the following values for γ and β:

$$\gamma = 0,\ 0.3,\ 0.5,\ 0.8,\ 0.9$$

and

$$\beta = 2^0,\ 2^{-1},\ 2^{-2},\ 2^{-3},\ 2^{-4},\ 2^{-5},\ 2^{-6},\ 2^{-7}.$$

In this manner we consider different signal-to-noise ratios as well as distinct levels of persistence of the regressor.

In Tables 1 and 2 we report the mean and the standard deviation of 100 times the out-of-sample R^2 over the 1000 Monte Carlo replications.

5. DATA

We use the data set of Campbell and Thompson (2008), which was kindly provided by Sam Thompson. The data frequency is monthly. Excess returns on the S&P 500 are calculated from the returns time series (1871M2 through 2005M12, CRSP since 1927) and the 3-month Treasury-Bill interest rate (1920M1 through 2005M12, 1870M2 through 1919M12 calculated following Goyal and Welch (2008)). The predictor variables are the dividend yield ***d/p*** (1872M2 through 2005M12), earnings yield ***e/p*** (1872M2 through 2005M12), smoothed earnings yield ***se/p*** following Campbell and Shiller (1988b), Campbell and Shiller (1998) (1881M1 through 2005M12), book-to-market ratio ***b/m*** (1926M6 through 2005M12), smoothed return on equity ***roe*** as described in Campbell and Thompson (2008) (1936M6 through 2005M12), the 3-month Treasury-Bill ***tbl*** (1920M1 through 2005M12), long-term government bond yield ***lty*** (1870M1 through 2005M12), the term spread ***ts***, i.e. the difference between long-term and short-term treasury yields (1920M1 through 2005M12), the default spread ***ds***, i.e. the difference between corporate and Treasury bond yields (1919M1 through 2005M12), the lagged inflation rate ***inf*** (1871M5 through 2005M12), the equity share of new issues ***nei*** proposed by Baker and Wurgler (2000), and the consumption-wealth ratio ***cay*** proposed by Lettau and Ludvigson (2001). As sample and forecast periods we report the same as in Campbell and Thompson (2008). Additionally, we consider the sample period 1960M1 through 2005M12 with forecasts starting in 1980M1.

We apply sign restrictions on the coefficients β depending on the predictor, a positivity restriction on the forecast y_{t+1} of the risk premium, and the intersection of these two. The coefficient restrictions are listed for the different predictors in the table below.

TABLE 1. SIMULATION RESULTS: AVERAGE

	$\beta = 1$					$\beta = 0.5$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	3.4074	4.0077	4.6158	10.5224	20.2218	0.2147	0.4330	0.5870	2.3659	5.4493
PC	3.4220	4.0193	4.6252	10.5226	20.2229	0.3323	0.5602	0.6848	2.4097	5.4683
PC-GH	3.4634	4.0547	4.6515	10.5238	20.2232	0.4065	0.6525	0.7726	2.4859	5.5112
PF	3.4074	4.0085	4.6166	10.5225	20.2237	0.2252	0.4418	0.5989	2.3823	5.4880
PF-GH	3.3980	4.0048	4.6134	10.5157	20.2250	0.2476	0.4654	0.6177	2.4046	5.5163
	$\beta = 0.25$					$\beta = 0.125$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	-0.4734	-0.4861	-0.2177	-0.1571	0.7566	-0.6075	-0.6761	-0.6788	-0.6255	-0.5323
PC	-0.1988	-0.1922	0.0313	0.0154	0.9247	-0.2492	-0.3086	-0.3221	-0.2519	-0.1885
PC-GH	-0.1659	-0.1364	0.0689	0.0862	1.0222	-0.2545	-0.3307	-0.3360	-0.2321	-0.1538
PF	-0.3607	-0.3788	-0.1339	-0.0493	0.8688	-0.3702	-0.3638	-0.3591	-0.2576	-0.2138
PF-GH	-0.2503	-0.2610	-0.0435	0.0559	0.9936	-0.2269	-0.2219	-0.2227	-0.1162	-0.0452
	$\beta = 0.0625$					$\beta = 0.0313$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	-0.7901	-0.7435	-0.7446	-0.8628	-0.8549	-0.7134	-0.7568	-0.7645	-0.9021	-0.9805
PC	-0.4001	-0.3867	-0.3661	-0.4615	-0.4261	-0.3409	-0.4082	-0.4264	-0.4925	-0.4953
PC-GH	-0.4723	-0.4511	-0.4089	-0.4976	-0.4455	-0.4151	-0.4728	-0.4920	-0.5749	-0.5467
PF	-0.2427	-0.2163	-0.2200	-0.2862	-0.2371	-0.1319	-0.1082	-0.0843	-0.2362	-0.2480
PF-GH	-0.1700	-0.1578	-0.1667	-0.2305	-0.1730	-0.1828	-0.1301	-0.1069	-0.2669	-0.2638
	$\beta = 0.0156$					$\beta = 0.0078$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	-0.7716	-0.7932	-0.8084	-0.7993	-0.8592	-0.7007	-0.7035	-0.7204	-0.7711	-0.9403
PC	-0.3952	-0.3772	-0.4243	-0.4412	-0.4346	-0.3390	-0.3191	-0.3091	-0.3844	-0.4737
PC-GH	-0.4531	-0.4416	-0.4967	-0.5241	-0.5284	-0.4145	-0.3936	-0.3977	-0.4783	-0.5907
PF	-0.1028	-0.0543	-0.1039	-0.1451	-0.0951	-0.0883	0.0008	-0.0011	-0.0506	-0.0758
PF-GH	-0.1831	-0.1389	-0.1734	-0.2208	-0.1810	-0.1855	-0.0950	-0.0917	-0.1592	-0.1923

The table shows the average of the out-of-sample R^2 over 1000 Monte Carlo replications.

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint.

Variable	Sign β	Variable	Sign β
d/p	$+$	***lty***	$-$
e/p	$+$	***ts***	$+$
se/p	$+$	***ds***	$+$
b/m	$+$	***inf***	$-$
roe	$+$	***nei***	$-$
tbl	$-$		

We follow Campbell and Thompson (2008) in the case of the consumption-wealth ratio ***cay*** and use consumption, assets, and income as three regressors with sign restriction $(+, -, -)$ instead of generating one fitted regressor as proposed in Lettau and Ludvigson (2001).

TABLE 2. SIMULATION RESULTS: STANDARD DEVIATION

	$\beta = 1$					$\beta = 0.5$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	3.9749	4.4555	4.4955	7.4016	12.2176	2.2901	2.7641	2.5726	3.8885	6.1906
PC	3.9624	4.4464	4.4881	7.4015	12.2166	2.2344	2.7028	2.5282	3.8634	6.1834
PC-GH	3.9632	4.4416	4.4831	7.3927	12.2091	2.2910	2.7490	2.5830	3.8766	6.1876
PF	3.9748	4.4560	4.4952	7.4008	12.2189	2.2819	2.7640	2.5736	3.8836	6.1770
PF-GH	3.9775	4.4552	4.4960	7.3942	12.2129	2.2624	2.7407	2.5655	3.8769	6.1558
	$\beta = 0.25$					$\beta = 0.125$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	1.7379	1.8637	1.8928	2.2622	3.1454	1.5594	1.6739	1.5128	1.6874	2.0767
PC	1.6296	1.7397	1.7910	2.2032	3.0739	1.3748	1.4612	1.3502	1.4493	1.9201
PC-GH	1.7441	1.8394	1.8872	2.3000	3.1286	1.4861	1.5714	1.4550	1.5919	2.0229
PF	1.6921	1.7795	1.8884	2.2031	3.1035	1.4158	1.5037	1.4498	1.5541	1.9649
PF-GH	1.6500	1.7271	1.8376	2.1611	3.0723	1.3956	1.4555	1.4127	1.5478	1.9307
	$\beta = 0.0625$					$\beta = 0.0313$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	1.4750	1.5323	1.4650	1.7653	1.7951	1.4459	1.4785	1.5379	1.5660	1.6635
PC	1.1799	1.2438	1.2645	1.4735	1.4980	1.1471	1.1725	1.2172	1.3267	1.3810
PC-GH	1.3140	1.3513	1.3888	1.5835	1.6102	1.2445	1.2898	1.3250	1.4482	1.5182
PF	1.3229	1.4358	1.3654	1.5851	1.7020	1.3588	1.4536	1.5204	1.5110	1.5882
PF-GH	1.3664	1.4664	1.4243	1.6124	1.7273	1.4626	1.5413	1.6159	1.6075	1.6698
	$\beta = 0.0156$					$\beta = 0.0078$				
γ	0	0.3	0.5	0.8	0.9	0	0.3	0.5	0.8	0.9
UF	1.4144	1.4831	1.5558	1.5425	1.7217	1.4329	1.3289	1.4960	1.5270	1.6153
PC	1.0549	1.0918	1.2872	1.2027	1.3063	1.0260	0.9907	1.1250	1.1632	1.2450
PC-GH	1.1782	1.1988	1.3918	1.3382	1.4514	1.1434	1.1161	1.2420	1.2890	1.3842
PF	1.4424	1.5259	1.4446	1.5014	1.6192	1.3741	1.3792	1.4464	1.4637	1.6280
PF-GH	1.5526	1.6580	1.5536	1.6431	1.7393	1.5037	1.5067	1.5756	1.5947	1.7722

The table shows the standard deviation of the out-of-sample R^2 over 1000 Monte Carlo replications.

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint.

6. EMPIRICAL RESULTS

The estimation results are presented in Tables 1 through 4. Each table has two panels, Panel A for monthly returns and Panel B for annual returns. The tables are directly related to Tables 1 in Goyal and Welch (2008) and Campbell and Thompson (2008). The reported numbers are out-of-sample R^2 statistics R^2_{OS} multiplied by 100.

$$100R^2_{OS} = 100\left(1 - \frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{T}(y_t - \bar{y}_t)^2}\right), \tag{4}$$

where $\hat{y}_t$ is the fitted value from the predictive regression (1) estimated through $t-1$ and $\bar{y}_t$ is the historical average return estimated through period $t-1$. When we use the same sample and forecast periods as Campbell and Thompson (2008), we follow them in starting the calculation of the average return at the beginning of the sample in 1871, irrespective of possibly later availability of the predictor variables.

The columns of the tables represent the different predictor variables and the rows present the different methods applied. In Table 1, we report bagging results for the same sample and forecast periods as in Campbell and Thompson (2008). In Table 2, we report bagging results for the estimation sample starting in 1960M1 and one-period ahead forecasts beginning in 1980M1. In Table 3, we report bagging results with shrinkage at different critical values for the same sample and forecast periods as in Campbell and Thompson (2008). In Table 4, we report bagging results with shrinkage at different critical values for the estimation sample starting in 1960M1 and one-period ahead forecasts beginning in 1980M1.

The first block in each table reports the results without bagging. For Table 1 Panels A and B, where the same estimation and forecast samples are used as in Campbell and Thompson (2008), this block reports our replication of their results.

In Tables 1 and 2, the next blocks report the results for bagging using the i.i.d. bootstrap, the moving block bootstrap, the parametric bootstrap, the wild bootstrap, and the Gordon and Hall (2008) bagging estimation, sequentially. In Tables 3 and 4, the first block reports the results without bagging, which are repeated from Tables 1 and 2, respectively. The second block reports bagging results with shrinkage at critical value $c = 0$, which are again repeated from Tables 1 and 2. The third and fourth blocks report bagging results with shrinkage at critical values $c = \sqrt{2}$ and $c = \sqrt{\log T}$, respectively.

In Tables 1 and 2 we report the results for B-PC and PC-GH separately, despite their equivalence shown in Proposition 2. The numerical difference between the two stems from the computation of the historical mean used in B-PC from 1871 onwards. The equivalence established in Proposition 2 implies computation of the historical mean from the beginning of the estimation sample. In this sense, PC-GH can be understood as B-PC with a different historical mean, namely the one computed from the beginning of the estimation sample.

The main findings that emerge from the tables are summarized in the following list.

(1) For monthly returns and the sample and forecast periods considered in Campbell and Thompson (2008), for every single predictor variable there is a bagging procedure that results in improved forecast performance. As outlined in Campbell and Thompson (2008) Section 2, these differences are not statistically significant but nevertheless economically meaningful for a mean-variance investor.

(2) For annual returns, the same statement holds with the single exception of the smoothed price/earnings ratio ***se/p***.

(3) In the case of the estimation sample starting in 1960M1 and forecasts beginning in 1980M1, for both monthly and annual returns and all predictor variables, there is a bagging procedure that improves the forecast performance. In most cases, however, here an improvement means that

the out-of-sample R^2 is a smaller negative number, i.e. for the majority of predictor variables, the historical mean outperforms the forecast regression. Bagging only reduces this advantage of the historical mean.

(4) We see a sharp decline in predictive power of the regressors as we move from the Campbell and Thompson (2008) estimation and forecast periods to the 1960M1/1980M1 period.

(5) When we apply different critical values for the pre-test in the bagging procedure, on the Campbell and Thompson (2008) monthly sample, the BIC selection improves the forecast for 7 of 11 regressors. On the Campbell and Thompson (2008) annual sample, the AIC selection improves the forecast for 2 regressors and BIC selection improves the forecast for 2 regressors.

(6) On the 1960M1/1980M1 monthly sample, BIC selection improves the forecast for 8 of 11 regressors. On the 1960M1/1980M1 annual sample, BIC selection improves the forecast for 6 regressors and AIC selection improves the forecast for 1 regressor. These are again for the most part reductions of the disadvantage compared to the mean forecast.

(7) Comparing the different bagging techniques, on the Campbell and Thompson (2008) monthly sample, the Gordon and Hall (2008) method works best for 8 of 11 regressors. The wild bootstrap works best for 2 regressors. On the Campbell and Thompson (2008) annual sample, the picture is scattered: Gordon and Hall (2008) is best for only 1 regressor, the wild bootstrap is best for 4 regressors, the i.i.d. bootstrap is best for 3 regresors.

(8) On the 1960M1/1980M1 sample, for both monthly and annual returns, the i.i.d. bootstrap performs best.

7. CONCLUSION

In this paper, we propose a new application of bagging as a means of imposing parameter restrictions. Bagging imposes a soft threshold at the boundary as opposed to the hard threshold that corresponds to simple restricted estimation. We show that the resulting forecast has lower variance than the forecast that results from a simple restricted estimator. The main result of variance reduction is the consequence of a result from Gordon and Hall (2008). In Proposition 1, we show the applicability of their estimator to the return prediction problem. In Proposition 2, we show that bagging forecasts from simple restricted estimations is equivalent to applying the Gordon and Hall (2008) estimator. In an empirical application using the same data set as in Campbell and Thompson (2008), we show that the resulting forecasts have more predictive power than those resulting from simple parameter restrictions.

ACKNOWLEDGMENTS

Sam Thompson kindly provided the data set used in Campbell and Thompson (2008). We gratefully acknowledge comments from the participants of *Forecasting in Rio*, Getulio Vargas Foundation, Rio De Janeiro, Brazil, where this paper was presented on 31 July 2008.

REFERENCES

ANG, A., AND G. BEKAERT (2006): "Stock Return Predictability: Is It There?," *Review of Financial Studies*, forthcoming.

AUDRINO, F., AND M. MEDEIROS (2008): "Smooth Regimes, Macroeconomic Variables, and Bagging for the Short-Term Interest Rate Process," Discussion paper, Pontifical Catholic University of Rio de Janeiro.

BAKER, M., AND J. WURGLER (2000): "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance*, 55, 2219–2257.

BÜHLMANN, P., AND B. YU (2002): "Analyzing Bagging," *Annals of Statistics*, 30, 927–961.

BOSSAERTS, P., AND P. HILLION (1999): "Implementing statistical criteria to select return forecasting models: what do we learn?," *Review of Financial Studies*, 12, 405–428.

BREIMAN, L. (1996): "Bagging Predictors," *Machine Learning*, 36, 105–139.

BUTLER, A., G. GRULLON, AND J. WESTON (2006): "Can Managers Forecast Aggregate Market Returns?," *Journal of Finance*, 60, 963–986.

CAMPBELL, J. (1987): "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373–399.

CAMPBELL, J., AND R. SHILLER (1988a): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

——— (1988b): "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

——— (1998): "Valuation Ratios and the Long-Run Stock Market Outlook," *Journal of Portfolio Management*, 24, 11–26.

CAMPBELL, J., AND S. THOMPSON (2008): "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies*, forthcoming.

CAMPBELL, J., AND M. YOGO (2006): "Efficient Tests of Stock Return Predictability," *Journal of Financial Economics*, 81, 27–60.

CAVANAGH, C., G. ELLIOT, AND J. STOCK (1995): "Inference in Models with Nearly Integrated Regressors," *Econometric Theory*, 11, 1131–1147.

FAMA, E., AND K. FRENCH (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–25.

——— (1989): "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25, 23–49.

FAMA, E., AND G. SCHWERT (1977): "Asset Returns and Inflation," *Journal of Financial Economics*, 5, 115–146.

——— (1981): "Stock Returns, Real Activity, Inflation and Money," *American Economic Review*, 71, 545–565.

FERSON, W., S. SARKISSIAN, AND T. SIMIN (2003): "Spurious Regressions in Financial Economics?," *Journal of Finance*, 58, 1393–1413.

FOSTER, F., T. SMITH, AND R.E.WHALEY (1997): "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R2," *Journal of Finance*, 52, 591–607.

GORDON, I., AND P. HALL (2008): "Estimating a Parameter When It Is Known that the Parameter Exceeds a Given Value," Discussion paper, University of Melbourne.

GOYAL, A., AND I. WELCH (2008): "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, forthcoming.

HALL, P., J. HOROWITZ, AND B.-Y. JING (1995): "On Blocking Rules for the Bootstrap with Dependent Data," *Biometrika*, 82, 561–574.

HILLEBRAND, E., AND M. MEDEIROS (forthcoming): "The Benefits of Bagging for Forecast Models of Realized Volatility," *Econometric Reviews*.

HUANG, H., AND T.-H. LEE (2007a): "Forecasting Using High-Frequency Financial Time Series," Working paper, University of California at Riverside.

——— (2007b): "To Combine Forecasts or to Combine Information," Working paper, University of California at Riverside.

INOUE, A., AND L. KILIAN (2008): "How Useful is Bagging in Forecasting Economic Time Series? A Case Study of U.S. CPI Inflation," *Journal of the American Statistical Association*, forthcoming.

JANSSON, M., AND M. J. MOREIRA (2006): "Optimal Inference in Regression Models with Nearly Integrated Regressors," *Econometrica*, 74, 681–715.

KEIM, D., AND R. F. STAMBAUGH (1986): "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics*, 17, 357–390.

KILIAN, L. (1999): "Exchange Rates and Monetary Fundamentals: What DoWe Learn from Long-Horizon Regressions?," *Journal of Applied Econometrics*, 14, 491–510.

KÜNSCH, H. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217–1241.

LEE, T.-H., AND Y. YANG (2006): "Bagging Binary and Quantile Predictors for Time Series," *Journal of Econometrics*, 135, 465–497.

LETTAU, M., AND S. LUDVIGSON (2001): "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56, 815–849.

LEWELLEN, J. (2004): "Predicting Returns with Financial Ratios," *Journal of Financial Economics*, 74, 209–235.

LI, Q., AND S. WANG (1998): "A Simple Consistent Bootstrap Test for a Parametric Regression Function," *Journal of Econometrics*, 87, 145–165.

MARK, N. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85, 201–218.

NELSON, C., AND M. KIM (1993): "Predictable Stock Returns: The Role of Small Sample Bias," *Journal of Finance*, 48, 641–661.

PESARAN, M., AND A. TIMMERMANN (2002): "Market Timing and Return Prediction under Model Instability," *Journal of Empirical Finance*, 9, 495–510.

PESARAN, M. H., AND A. TIMMERMANN (1995): "Predictability of stock returns: robustness and economic significance," *Journal of Finance*, 50, 1201–1228.

POLITIS, D., AND J. ROMANO (1994): "The Stationary Bootstrap," *Journal of the American Statistical Association*, 89, 1303–1313.

POLK, C., S. THOMPSON, AND T. VUOLTEENAHO (2006): "Cross-Sectional Forecasts of the Equity Premium," *Journal of Financial Economics*, 81, 101–141.

RAPACH, D., AND J. STRAUSS (2007): "Bagging or Combining (or Both)? An Analysis Based on Forecasting U.S. Employment Growth," Working paper, Saint Louis University.

RAPACH, D., J. STRAUSS, AND G. ZHOU (2008): "Out-of-Sample Equity Premium Prediction: Consistently Beating the Historical Average," Working paper, Saint Louis University.

ROZEFF, M. (1984): "Dividend Yields are Equity Risk Premiums," *Journal of Portfolio Management*, 11, 68–75.

SCHWERT, G. (2002): "Anomalies and market efficiency," in *Handbook of the Economics of Finance*, ed. by M. Harris, and R. Stulz. North Holland: Amsterdam.

STAMBAUGH, R. (1999): "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

STOCK, J., AND M. WATSON (2005): "An Empirical Comparison of Methods for Forecasting Using Many Predictors," Working paper, Harvard University.

TIMMERMANN, A. (2007): "Elusive Return Predictability," *International Journal of Forecasting*, forthcoming.

TOROUS, W., R. VALKANOV, AND S. YAN (2004): "On Predicting Stock Returns with Nearly Integrated Explanatory Variables," *Journal of Business*, 77, 937–966.

APPENDIX A. PROOF OF PROPOSITION 1

We need to establish the validity of the assumptions for the theorem in Gordon and Hall (2008) for the special cases $\theta = \beta$ and $\theta = \mathbb{E}(y_{T+1}|x_T)$. Then, the statement of the theorem follows from the proof in Gordon and Hall (2008). Write F_T for the empirical distribution function of the data set $\mathcal{X} = \{(x_t, y_{t+1})\}_{t=0,\dots,T-1}$ and F_T^* for its resample $\mathcal{X}^* = \{(x_t, y_{t+1})^*\}_{t=0,\dots,T-1}$. Then, the assumptions are

$$\mathbb{E}\left[\theta(F_T) - \left\{\theta_0 + \frac{1}{T}\sum_{t=0}^{T-1} a(y_t, F)\right\}\right]^2 \leq T^{-2}\int b(x, F)\mathrm{d}F(x), \tag{5}$$

$$\mathbb{E}\left(\left[\theta(F_T^*) - \left\{\tilde{\theta} + \frac{1}{T}\sum_{t=0}^{T-1} a(y_t^*, F_T)\right\}\right]^2 \middle| \mathcal{X}\right) \leq T^{-2}\int b(x, F_T)\mathrm{d}F_T(x), \tag{6}$$

where, for distribution function G, the functionals a and $b \geq 0$ satisfy

$$\int a(x, G)\mathrm{d}G(x) = 0, \tag{7}$$

$$\int \left\{a(x, F)^2 + b(x, F)\right\} \mathrm{d}F(x) < \infty, \tag{8}$$

$$\sup_T \mathbb{E}\left[a(y_1, F_T)^2 + b(y_1, F_T)\right] < \infty, \tag{9}$$

$$\lim_{T\to\infty} \mathbb{P}\left[\int |a(x, F_T)|^{2+\epsilon}\mathrm{d}F_T(x) \leq C\right] = 1 \tag{10}$$

for some $C, \epsilon > 0$, and

$$\int a(x, F_T)\mathrm{d}F_T(x) \to \tau^2 = \int a(x, F)^2\mathrm{d}F(x) > 0, \tag{11}$$

where the convergence is in probability.

When convenient, we will consider the parameter vector $b := (\alpha, \beta)^T$ instead of β only. Let X denote the regressor matrix that contains a columns of ones and the single regressor time series $\{x_t\}_{t=0,\dots,T-1}$, $y = \{y_t\}_{t=1,\dots,T}$, and $\varepsilon = \{\varepsilon_t\}_{t=1,\dots,T}$. To establish Equation (5) in the case $\theta = \beta$, expand the OLS estimator $\tilde{b} = (X^TX)^{-1}X^Ty$ in a Taylor series around b_0;

$$\begin{aligned}\tilde{b} &= b_0 + \tilde{b}'(y)(y - Xb), \\ &= b_0 + (X^TX)^{-1}X^T\varepsilon.\end{aligned} \tag{12}$$

In (5), set $\theta(F_T) = \tilde{b}$, $\theta_0 = b_0$, $1/T\sum_{t=1}^{T} a(y_t, F) = (X^TX)^{-1}X^t\varepsilon$. Since $\tilde{b}$ is a linear estimator, there are no terms of order higher than one, and therefore

$$\mathbb{E}\left[\tilde{b} - \{b_0 + (X^TX)^{-1}X^T\varepsilon\}\right]^2 = 0.$$

Similarly, for a resample $\mathcal{X}^*$, expand the bootstrap estimator $\tilde{b}^*$ around $\tilde{b}$:

$$\begin{aligned}\tilde{b}^* &= \tilde{b} + (\tilde{b}^*)'(y)(y^* - X^*\tilde{b}),\\ &= \tilde{b} + ((X^*)^T X^*)^{-1}(X^*)^T\varepsilon^*,\end{aligned}$$

and therefore we have (6) in the form of

$$\mathbb{E}\left(\left[\tilde{b}^* - \{\tilde{b} + ((X^*)^T X^*)^{-1}(X^*)^T\varepsilon^*\}\right]^2 \middle| \mathcal{X}\right) = 0.$$

To establish (7) for $\theta = \beta$, consider (12) element-wise:

$$\tilde{b} = \begin{bmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \frac{1}{T}\begin{bmatrix} \frac{\frac{1}{T}\sum_{t=1}^T x_{t-1}^2 \sum_{t=1}^T \varepsilon_t - \frac{1}{T}\sum_{t=1}^T x_{t-1}\sum_{t=1}^T x_{t-1}\varepsilon_t}{\frac{1}{T}\sum_{t=1}^T x_{t-1}^2 - \left(\frac{1}{T}\sum_{t=1}^T x_{t-1}\right)^2} \\ \frac{\sum_{t=1}^T x_{t-1}\varepsilon_t - \frac{1}{T}\sum_{t=1}^T x_{t-1}^2\sum_{t=1}^T \varepsilon_t}{\frac{1}{T}\sum_{t=1}^T x_{t-1}^2 - \left(\frac{1}{T}\sum_{t=1}^T x_{t-1}\right)^2} \end{bmatrix}$$

Thus, for $\theta = \beta$, set in (5)

$$a(y_t, F) = \frac{x_t(y_t - \alpha - \beta x_{t-1}) - \frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}(y_t - \alpha - \beta x_{t-1})}{\frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}^2 - \left(\frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}\right)^2}.$$

For $\theta = \alpha$, the functional a can be set as

$$a(y_t, F) = \frac{\frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}(y_t - \alpha - \beta x_{t-1}) - \frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}^2(y_t - \alpha - \beta x_{t-1})}{\frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}^2 - \left(\frac{1}{T}\sum_{\tau=1}^T x_{\tau-1}\right)^2}.$$

Then, (7) follows from the zero-mean assumption on ε_t:

$$\int a(x, G)\mathrm{d}G(x) = \mathbb{E}_G a(x, G) = 0.$$

Since the Taylor expansion (12) ends at the first order term, $b(x, F) = 0$ and (8) reduces to

$$\int a(x, F)^2 \mathrm{d}F(x) = \mathbb{E}_F a(y, F)^2 = \frac{\sigma^2}{\mathrm{Var}(x)} < \infty.$$

Note that on the left-hand side, x is just an indicator variable, whereas on the right-hand side, $\mathrm{Var}(x)$ means the variance of the regressor series $\{x_t\}_{t=0,\dots,T-1}$. By the same argument, we obtain condition (11), since

$$\int a(x, F_T)^2 \mathrm{d}F_T(x) = \frac{\frac{1}{T}\sum_{t=1}^T \varepsilon_t^2}{\frac{1}{T}\sum_{t=1}^T x_{t-1}^2 - \left(\frac{1}{T}\sum_{t=1}^T x_{t-1}\right)^2} = \frac{1}{T}\sum_{t=1}^T \varepsilon_t^2 (X^T X)^{-1}_{22},$$

where A_{ij} denotes the entry in row i and column j of matrix A, converges in probability to

$$\tau^2 = \int a(x, F)^2 \mathrm{d}F(x) = \frac{\sigma^2}{\mathrm{Var}(x)}.$$

Equation (9) reduces to

$$\sup_T \mathbb{E}a(y_1, F_T)^2 = \sup_T \mathbb{E}\left[x_0\varepsilon_1 - \frac{1}{T}\sum_{t=1}^{T} x_{t-1}\varepsilon_1\right]^2$$
$$= \sup_T \sigma^2 \left(\frac{1}{T}\sum_{t=2}^{T} x_{t-1}\right)^2 < \infty$$

for $\mathbb{E}x < \infty$. Equation (10) can be written as $\mathbb{E}|a(x, F)|^{2+\epsilon} \leq C$ for some $\epsilon > 0$, $C > 0$, and follows from the assumption of the Lyapunov-condition for ε_t. The variance τ^2 of the estimator is given by $\sigma^2/\operatorname{Var}(x)$ and the finite sample version $\hat{\sigma}^2(X^TX)^{-1}_{22}$ can be used for studentization in finite samples.

Since for the case $\theta = \mathbb{E}(y_{T+1}|x_T) = \tilde{\alpha} + \tilde{\beta}x_T$ the parameter is a linear combination of the estimators of the parameters α and β, the validity of the assumption can be shown analogously. The variance term τ^2 can become more complicated depending on further assumptions on the error process. For i.i.d. errors, we have $\tau^2 = \sigma^2[(X^TX)^{-1}_{11} + x_T^2(X^TX)^{-1}_{22}]$. For autocorrelated error processes with covariance matrix $\mathbb{E}\varepsilon\varepsilon^T = \Omega$, we have $\tau^2 = [(X^TX)^{-1}X^T\Omega X(X^TX)^{-1}]_{11} + x_T^2[(X^TX)^{-1}X^T\Omega X(X^TX)^{-1}]_{22} + 2x_T[(X^TX)^{-1}X^T\Omega X(X^TX)^{-1}]_{12}$. For more general, heteroskedastic and autocorrelated processes, the expressions are accordingly more involved. □

APPENDIX B. PROOF OF PROPOSITION 2

The proof is a very short application of the linearity of expectations.

$$\frac{1}{B}\sum_{j=1}^{B} \bar{y}_{t+1}^{(j)} = \frac{1}{B}\sum_{j=1}^{B}(\bar{\alpha}^{*(j)} + \bar{\beta}^{*(j)}x_t)$$
$$= \left(\frac{1}{B}\sum_{j=1}^{B}\bar{\alpha}^{*(j)}\right) + \left(\frac{1}{B}\sum_{j=1}^{B}\bar{\beta}^{*(j)}\right)x_t$$
$$= \hat{\alpha} + \hat{\beta}x_t = \hat{y}_{t+1}.$$

□

TABLE 3. PANEL A: MONTHLY RETURNS, CAMPBELL AND THOMPSON (2008) SAMPLE SIZES

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*	*cay*
Begin												
Sample	1872:1	1872:2	1881:2	1926:6	1936:6	1920:1	1870:1	1920:1	1919:1	1871:5	1927:12	1951:12
Forecast	1927:1	1927:1	1927:1	1946:1	1956:6	1940:1	1927:1	1940:1	1939:1	1927:1	1947:12	1971:12
Campbell and Thompson (2008)												
In-Sample R^2	1.1208	0.7081	1.3521	0.6078	0.0225	0.8691	0.1911	0.6561	0.0956	0.0549	0.4828	2.5983
In-Sample t	1.2519	2.2829	1.8495	1.9566	0.3587	2.4586	1.4730	2.1804	0.7248	0.3898	1.7510	3.5924
UF	-0.6563	0.1159	0.3239	-0.3968	-0.9259	0.5048	-0.1893	0.4449	-0.2391	-0.2210	0.3218	-1.5393
PC	0.0483	0.1770	0.4175	-0.3968	-0.0778	0.4899	-0.1893	0.4565	-0.2391	-0.2092	0.3218	-1.5393
PF	0.0723	0.1321	0.3777	0.0168	-0.9259	0.5401	0.2047	0.4392	-0.2391	-0.1834	0.4843	0.1596
PCF	0.0798	0.1770	0.4297	0.0168	-0.0778	0.5251	0.2047	0.4507	-0.2391	-0.1717	0.4843	0.1596
i.i.d. bootstrap												
B-PC	0.2644	0.2908	0.4319	-0.5697	-0.5852	0.5723	-0.1783	0.5558	-0.5579	-0.1296	0.3388	-1.2483
B-PF	0.5060	0.2433	0.4529	0.2728	-0.7925	0.5468	0.2178	0.2548	0.0802	-0.0329	0.3921	0.4458
B-PCF	0.3558	0.2992	0.4894	0.0943	-0.2446	0.3640	0.1676	0.2530	-0.4871	-0.1430	0.4070	0.4490
moving block bootstrap												
B-PC	0.1277	0.1703	0.5128	-1.1542	-0.7427	0.4207	-0.2970	0.3320	-0.4614	-0.3859	0.1342	-2.9017
B-PF	0.2649	0.1696	0.5494	0.2259	-1.0368	0.3923	0.1788	0.0665	0.0202	-0.2413	0.2762	-0.7165
B-PCF	0.2677	0.2174	0.5853	-0.2123	-0.3855	0.2372	0.1495	0.0266	-0.3868	-0.3567	0.3146	-0.5658
parametric bootstrap												
B-PC	0.1345	0.2044	0.3972	-0.5069	-0.6769	0.5307	-0.1578	0.5256	-0.1625	-0.2176	0.3815	-1.5576
B-PF	0.2210	0.1481	0.4224	0.1000	-0.8596	0.4643	0.2316	0.3147	0.0859	-0.1475	0.4344	0.2143
B-PCF	0.2739	0.1934	0.4497	0.0891	-0.3690	0.1572	0.1714	0.0863	-0.0476	-0.2141	0.4301	0.2223
wild bootstrap												
B-PC	0.3510	0.2075	0.5725	-0.2828	-0.7369	0.4475	-0.1691	0.5838	-0.3732	-0.1803	0.4025	-1.2818
B-PF	0.4975	0.1445	0.5542	0.3317	-0.8746	0.5077	0.2456	0.3207	0.1510	-0.1389	0.4760	0.5025
B-PCF	0.4142	0.1938	0.6038	0.0965	-0.4053	0.3416	0.1949	0.1921	-0.4719	-0.2257	0.4737	0.5090
Gordon and Hall (2008)												
PC-GH iid	0.7388	0.3829	0.6704	-0.7287	-1.1179	0.5999	-0.1544	0.7707	-0.9398	-0.0728	0.3276	-1.3136
PC-GH mbb	0.3172	0.2655	0.5433	-2.0418	-0.7579	0.5547	-0.2229	0.6153	-0.6771	-0.2314	0.2210	-2.9282
PF-GH iid	0.8865	0.3850	0.6884	-0.0454	-1.1179	0.6046	0.2520	0.7130	-0.6272	0.0146	0.4901	0.2914
PF-GH mbb	0.4134	0.2902	0.5649	-0.1343	-0.7579	0.5809	0.2388	0.5902	-0.4807	-0.0368	0.3998	-0.9646

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance, *cay* consumption-wealth ratio

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, X-GH simple forecast of type X with bagged coefficient estimate according to Gordon and Hall (2008)

TABLE 1. PANEL B: ANNUAL RETURNS, CAMPBELL AND THOMPSON (2008) SAMPLE SIZES

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*	*cay*
Begin												
Sample	1872:1	1872:2	1881:2	1926:6	1936:6	1920:1	1870:1	1920:1	1919:1	1871:5	1927:12	1951:12
Forecast	1927:1	1927:1	1927:1	1946:1	1956:6	1940:1	1927:1	1940:1	1939:1	1927:1	1947:12	1971:12
Campbell and Thompson (2008)												
In-Sample R^2	10.8938	6.7784	13.5664	8.2560	0.3181	4.2606	0.7710	3.0962	0.0069	0.0695	0.3508	19.8661
In-Sample t	2.6863	2.8431	3.0138	1.9809	0.3511	1.7718	0.9072	1.7208	0.0722	0.1727	0.5357	3.7554
UF	5.5097	4.9371	7.8745	-3.0922	-8.6058	5.4356	-0.1541	4.6836	-3.8433	-0.6954	-4.2734	-7.9187
PC	5.5097	4.9371	7.8745	-3.0922	-0.0257	5.4356	-0.1541	4.6836	-3.8433	-0.6954	-4.2734	-7.9187
PF	5.6139	4.9419	7.8333	1.5456	-8.3601	7.3628	2.2612	4.6336	-3.8433	-0.6954	-2.3796	-1.6680
PCF	5.6139	4.9419	7.8333	1.5456	-0.0257	7.3628	2.2612	4.6336	-3.8433	-0.6954	-2.3796	-1.6680
i.i.d. bootstrap												
B-PC	5.3571	4.9291	7.8280	-2.7367	-4.1589	5.5399	-0.1647	4.7963	-3.5163	-0.7443	-4.2597	-7.9210
B-PF	5.5990	4.9635	7.8056	1.7966	-8.0279	7.4829	2.2976	4.6604	-3.0442	-0.6455	-2.3881	-0.9136
B-PCF	5.5969	4.9634	7.8062	1.7966	-11.9815	6.6743	2.1833	3.7147	-3.6887	-0.9503	-2.3881	-0.9136
moving block bootstrap												
B-PC	1.9165	3.5838	6.5631	-5.3136	-6.8061	0.9646	-3.7241	1.9863	-6.8233	-3.1526	-6.8641	-22.5497
B-PF	2.3590	3.7919	6.6253	1.4475	-11.5294	4.4786	0.5247	1.2838	-3.5188	-2.6369	-4.2866	-3.9640
B-PCF	2.4220	3.9039	6.7807	1.3865	-8.1528	1.9567	-0.4112	0.4068	-6.4227	-4.4039	-4.1011	-3.9640
parametric bootstrap												
B-PC	5.5013	4.9292	7.8153	-3.1185	-4.1922	5.4298	-0.2077	4.6298	-3.9420	-0.7116	-4.3374	-7.7007
B-PF	5.6607	4.9731	7.7790	1.6560	-8.1409	7.3747	2.2992	4.5333	-3.6519	-0.6346	-2.4196	-0.9274
B-PCF	5.6611	4.9739	7.7781	1.6560	-12.1203	6.2446	2.1009	3.8714	-3.7019	-0.7829	-2.4196	-0.9274
wild bootstrap												
B-PC	5.4566	4.9571	7.8360	-3.2391	-4.1806	5.4691	-0.1777	4.7245	-3.8810	-0.7095	-4.2643	-7.7968
B-PF	5.7628	4.9866	7.8310	1.5354	-8.1313	7.3998	2.3083	4.5217	-3.3864	-0.6342	-2.3434	-0.8519
B-PCF	5.7613	4.9866	7.8320	1.5354	-11.9751	6.6654	2.1575	3.6576	-3.9830	-0.9251	-2.3434	-0.8519
Gordon and Hall (2008)												
PC-GH iid	5.4897	4.9408	7.8602	-2.7540	-4.5023	5.4778	-0.1662	4.8011	-3.9124	-0.6821	-4.2250	-7.7409
PC-GH mbb	3.0843	3.7113	6.6793	-5.3304	-10.1075	1.5500	-3.3582	2.5868	-8.7983	-2.7429	-6.4397	-23.4325
PF-GH iid	5.5770	4.9469	7.8208	1.6484	-4.5023	7.4149	2.2678	4.7497	-3.9124	-0.6821	-2.3529	-1.6576
PF-GH mbb	3.4168	3.7375	6.7497	0.3029	-10.1075	5.6677	0.9859	2.5852	-8.3492	-2.4071	-4.0384	-5.9999

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance, *cay* consumption-wealth ratio

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, X-GH simple forecast of type X with bagged coefficient estimate according to Gordon and Hall (2008)

TABLE 2. PANEL A: MONTHLY RETURNS, SAMPLE BEGIN 1960M1, FORECAST BEGIN 1980M1

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*	*cay*
Without Bagging (Campbell and Thompson 2008)												
In-Sample R^2	0.7909	0.6399	0.5846	0.5989	-0.4923	1.0079	0.6896	0.8725	0.6284	0.5376	0.4438	1.3091
In-Sample t	0.2285	0.1555	0.1273	0.1357	0.0809	0.3251	0.1632	0.2737	0.1385	0.1046	0.0558	1.3574
UF	-1.1409	-0.2858	-0.3635	-1.8618	-0.7030	-1.0722	-0.7357	-0.7341	-0.4805	-0.1243	-1.1336	-3.2744
PC	-1.1409	-0.2858	-0.3635	-1.8618	-0.6202	-1.0722	-0.7357	-0.7341	-0.4805	-0.0625	-1.1336	-3.2744
PF	-0.9988	-0.2858	-0.3635	-1.3007	-0.6910	-0.4594	-0.1863	-0.8430	-0.4498	-0.0404	0.0626	-1.1130
PCF	-0.9988	-0.2858	-0.3635	-1.3007	-0.6083	-0.4594	-0.1863	-0.8430	-0.4498	0.0214	0.0626	-1.1130
i.i.d. bootstrap												
B-PC	-1.1051	-0.2078	-0.1829	-1.5888	-0.7957	-1.2028	-0.6509	-0.6411	-0.3505	-0.1746	-1.0737	-2.9020
B-PF	-0.6943	-0.1788	-0.1596	-0.6639	-0.6450	-0.5281	-0.0772	-0.7665	0.1681	0.0303	0.1414	-0.7212
B-PCF	-0.5372	-0.2964	-0.2292	-0.7126	-0.8091	-0.4659	-0.1333	-0.7757	-0.1921	-0.2460	0.1276	-0.6907
moving block bootstrap												
B-PC	-1.2345	-0.2662	-0.3342	-2.0488	-0.7779	-1.0640	-0.5139	-1.1229	-1.2810	-0.1469	-1.3664	-3.3959
B-PF	-0.6143	-0.1831	-0.2584	-0.9279	-0.7082	-0.4176	-0.0558	-1.1344	-1.0399	0.0360	0.0630	-0.7559
B-PCF	-0.6646	-0.1936	-0.2119	-0.8831	-0.8406	-0.4421	-0.2258	-1.1404	-1.0756	-0.2173	0.0565	-0.7626
parametric bootstrap												
B-PC	-1.0009	-0.2612	-0.3182	-1.8997	-0.5682	-1.1979	-0.6245	-0.8581	-0.5807	-0.0730	-1.1535	-3.1214
B-PF	-0.5882	-0.1793	-0.2121	-0.8615	-0.5427	-0.3895	-0.0682	-0.8778	-0.3984	-0.0287	0.0994	-0.9175
B-PCF	-0.5150	-0.2363	-0.2419	-0.9178	-0.6430	-0.3816	-0.3585	-0.8702	-0.3726	-0.1205	0.1110	-0.9313
wild bootstrap												
B-PC	-1.2462	-0.2406	-0.3575	-1.6692	-0.7537	-0.9067	-0.6426	-0.7640	-0.4902	-0.0882	-1.1156	-3.0902
B-PF	-0.7947	-0.1742	-0.2718	-0.7543	-0.6297	-0.2574	-0.0802	-0.8594	-0.3189	0.0431	0.0128	-0.7130
B-PCF	-0.6761	-0.3484	-0.4002	-0.8035	-0.8350	-0.2557	-0.2187	-0.8613	-0.3083	-0.1645	0.0043	-0.7199
Gordon and Hall (2008)												
PC-GH iid	-1.1502	-0.1896	-0.2691	-1.9333	-0.8918	-1.1086	-0.8768	-0.7312	-0.4071	-0.2155	-1.2075	-3.8267
PC-GH mbb	-1.2537	-0.1729	-0.3163	-1.9521	-0.8830	-1.1149	-0.5860	-1.0617	-1.1809	-0.2173	-1.4428	-3.9719
PF-GH iid	-0.8630	-0.1857	-0.2691	-1.2501	-0.8218	-0.4787	-0.2043	-0.8408	-0.3736	0.1363	0.0481	-1.0719
PF-GH mbb	-0.8512	-0.1417	-0.3131	-1.2249	-0.8177	-0.5503	-0.1393	-1.1586	-1.0900	0.1407	0.0838	-0.9267

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance, *cay* consumption-wealth ratio

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, X-GH simple forecast of type X with bagged coefficient estimate according to Gordon and Hall (2008)

TABLE 2. PANEL B: ANNUAL RETURNS, SAMPLE BEGIN 1960M1, FORECAST BEGIN 1980M1

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*	*cay*
Without Bagging (Campbell and Thompson 2008)												
In-Sample R^2	0.8198	0.5730	0.6393	0.6966	-0.5539	0.3262	-0.1571	1.4186	0.0392	0.9692	-0.1211	12.2903
In-Sample t	3.0782	1.4507	1.8207	2.2429	1.4368	0.3611	0.0999	5.0697	0.0062	1.9298	0.0257	1.3901
UF	-4.3136	-0.5010	0.3993	-15.4237	-3.5269	-1.7396	-0.9190	7.1785	-1.2170	0.4148	-0.5880	-8.8274
PC	-4.3136	-0.5010	0.3993	-15.4237	-3.4753	-1.7396	-1.2369	7.1785	-1.2170	0.6379	-0.5433	-8.8274
PF	-4.5819	-0.5010	0.3993	-7.6233	-3.2032	-1.0312	-0.2809	6.8561	-1.2170	0.4148	-0.0255	-2.3109
PCF	-4.5819	-0.5010	0.3993	-7.6233	-3.1516	-1.0312	-0.5988	6.8561	-1.2170	0.6379	0.0191	-2.3109
i.i.d. bootstrap												
B-PC	-4.3197	-0.4995	0.3139	-15.4379	-3.4825	-1.8801	-1.7670	7.3900	-1.0159	0.7707	-0.9246	-8.1701
B-PF	-4.1265	-0.5291	0.3176	-7.2206	-3.0138	-0.9676	0.0109	6.8470	-1.0352	0.4860	0.2450	-1.6670
B-PCF	-4.2626	-0.6142	0.2534	-7.2707	-3.0723	-2.8319	-16.3027	6.8617	-0.9678	-2.9616	-3.7161	-1.6670
moving block bootstrap												
B-PC	-11.3655	-5.9102	-5.6628	-21.1407	-12.6312	-8.3206	-9.9206	2.6758	-7.7161	-1.4252	-7.0377	-25.6145
B-PF	-8.2608	-5.8967	-6.0762	-8.5638	-10.3951	-3.8694	-4.2070	1.9091	-8.8657	-3.0789	-3.8645	-10.7910
B-PCF	-8.2904	-7.0983	-7.2053	-8.5970	-11.0049	-7.4200	-13.2409	2.3896	-10.2604	-6.2221	-6.2123	-10.7910
parametric bootstrap												
B-PC	-4.0600	-0.5025	0.3980	-15.4223	-3.5899	-1.7872	-1.4392	7.2711	-1.2062	0.8453	-0.4205	-8.4574
B-PF	-3.9074	-0.5264	0.4236	-7.2999	-3.2290	-0.9222	-0.0390	6.8125	-1.2136	0.4260	0.2519	-1.8066
B-PCF	-3.8751	-0.6058	0.3523	-7.3429	-3.3995	-2.1978	-16.6117	6.8172	-1.2072	-2.7483	-3.0556	-1.8066
wild bootstrap												
B-PC	-4.2941	-0.3783	0.4155	-15.4183	-3.6969	-1.8216	-1.5513	7.1831	-1.1895	0.7711	-0.7787	-8.2015
B-PF	-4.0881	-0.4168	0.4073	-7.2451	-3.1945	-0.9428	0.0894	6.6462	-1.2167	0.4299	0.2817	-1.6618
B-PCF	-4.1196	-0.5060	0.3473	-7.3226	-3.3514	-2.7327	-15.8251	6.6716	-1.2502	-3.0058	-3.1977	-1.6618
Gordon and Hall (2008)												
PC-GH iid	-4.3259	-0.5016	0.3847	-15.4260	-3.8480	-1.9835	-2.9920	7.2028	-1.0857	0.9533	-1.2270	-8.8870
PC-GH mbb	-11.9478	-5.6223	-5.3842	-21.6567	-14.5263	-10.0778	-13.5205	3.2326	-7.5565	-1.9933	-9.1914	-28.6215
PF-GH iid	-4.6040	-0.5016	0.3847	-7.6585	-3.5247	-1.2617	-2.1754	6.8773	-1.0857	0.9533	-0.5452	-2.2748
PF-GH mbb	-11.2908	-5.6223	-5.3842	-11.5090	-12.3341	-4.9272	-6.1310	3.3877	-7.5565	-1.3520	-2.2403	-13.5375

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance, *cay* consumption-wealth ratio

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, X-GH simple forecast of type X with bagged coefficient estimate according to Gordon and Hall (2008)

TABLE 3. PANEL A: PRETESTING, MONTHLY RETURNS, CAMPBELL AND THOMPSON (2008) SAMPLE SIZES

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*
Begin											
Sample	1872:1	1872:2	1881:2	1926:6	1936:6	1920:1	1870:1	1920:1	1919:1	1871:5	1927:12
Forecast	1927:1	1927:1	1927:1	1946:1	1956:6	1940:1	1927:1	1940:1	1939:1	1927:1	1947:12
UF	-0.6563	0.1159	0.3239	-0.3968	-0.9259	0.5048	-0.1893	0.4449	-0.2391	-0.2210	0.3218
PC	0.0483	0.1770	0.4175	-0.3968	-0.0778	0.4899	-0.1893	0.4565	-0.2391	-0.2092	0.3218
PF	0.0723	0.1321	0.3777	0.0168	-0.9259	0.5401	0.2047	0.4392	-0.2391	-0.1834	0.4843
PCF	0.0798	0.1770	0.4297	0.0168	-0.0778	0.5251	0.2047	0.4507	-0.2391	-0.1717	0.4843
Table 1A											
Best	0.8865	0.3850	0.6884	0.3317	-0.0778	0.6046	0.2520	0.7707	0.1510	0.0146	0.4901
$c = 0$ (Table 1A)											
PC	0.0483	0.1770	0.4175	-0.3968	-0.0778	0.4899	-0.1893	0.4565	-0.2391	-0.2092	0.3218
B-PC iid	0.2644	0.2908	0.4319	-0.5697	-0.5852	0.5723	-0.1783	0.5558	-0.5579	-0.1296	0.3388
B-PC mbb	0.1277	0.1703	0.5128	-1.1542	-0.7427	0.4207	-0.2970	0.3320	-0.4614	-0.3859	0.1342
B-PC par	0.1345	0.2044	0.3972	-0.5069	-0.6769	0.5307	-0.1578	0.5256	-0.1625	-0.2176	0.3815
B-PC wild	0.3510	0.2075	0.5725	-0.2828	-0.7369	0.4475	-0.1691	0.5838	-0.3732	-0.1803	0.4025
$c = \sqrt{2}$											
PC	0	0.0008	0.0411	0	0	-0.1302	-0.2937	-0.4090	0	-0.1177	0.3218
B-PC iid	0.5015	0.2392	0.5461	-0.0000	-0.0799	0.5072	-0.1941	0.5852	-0.5616	-0.0807	0.3354
B-PC mbb	0.1417	0.1762	0.4357	-0.0000	-0.1426	0.4722	-0.2213	0.5681	-0.5393	-0.2253	0.2691
B-PC par	0.1064	0.2013	0.3643	-0.4783	-0.0215	0.2420	-0.1939	0.4268	-0.3071	-0.1686	0.3579
B-PC wild	0.1388	0.1508	0.4226	-0.0766	-0.0367	0.2147	-0.1254	0.3823	-0.1964	-0.1197	0.3436
$c = \sqrt{\log(T)}$											
PC	0	0	0	0	0	-0.3480	-0.1615	0	0	0	-0.2961
B-PC iid	0.5046	0.1932	0.5526	-0.0000	-0.0814	0.4831	-0.1404	0.6275	-0.4465	-0.0817	0.5311
B-PC mbb	0.1750	0.1165	0.4674	-0.0000	-0.1431	0.2789	-0.1448	0.3547	-0.3532	-0.1623	0.6183
B-PC par	0.0386	0.0742	0.2455	-0.2900	-0.0002	0.1220	-0.0701	0.1519	-0.1434	-0.0925	0.4102
B-PC wild	-0.0047	0.0321	0.0730	0.0504	-0.0101	0.1653	0.1026	0.1330	-0.0173	-0.0763	0.5048

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, iid bootstrap, mbb moving block bootstrap, par parametric bootstrap, wild bootstrap

TABLE 3. PANEL B: PRETESTING, ANNUAL RETURNS, CAMPBELL AND THOMPSON (2008) SAMPLE SIZES

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*
Begin											
Sample	1872:1	1872:2	1881:2	1926:6	1936:6	1920:1	1870:1	1920:1	1919:1	1871:5	1927:12
Forecast	1927:1	1927:1	1927:1	1946:1	1956:6	1940:1	1927:1	1940:1	1939:1	1927:1	1947:12
UF	5.5097	4.9371	7.8745	-3.0922	-8.6058	5.4356	-0.1541	4.6836	-3.8433	-0.6954	-4.2734
PC	5.5097	4.9371	7.8745	-3.0922	-0.0257	5.4356	-0.1541	4.6836	-3.8433	-0.6954	-4.2734
PF	5.6139	4.9419	7.8333	1.5456	-8.3601	7.3628	2.2612	4.6336	-3.8433	-0.6954	-2.3796
PCF	5.6139	4.9419	7.8333	1.5456	-0.0257	7.3628	2.2612	4.6336	-3.8433	-0.6954	-2.3796
Table 1B											
Best	5.7628	4.9866	7.8745	1.7966	-0.0257	7.4829	2.2992	4.8011	-3.0442	-0.6342	-2.3434
$c = 0$ (Table 1B)											
PC	5.5097	4.9371	7.8745	-3.0922	-0.0257	5.4356	-0.1541	4.6836	-3.8433	-0.6954	-4.2734
B-PC iid	5.3571	4.9291	7.8280	-2.7367	-4.1589	5.5399	-0.1647	4.7963	-3.5163	-0.7443	-4.2597
B-PC mbb	1.9165	3.5838	6.5631	-5.3136	-6.8061	0.9646	-3.7241	1.9863	-6.8233	-3.1526	-6.8641
B-PC par	5.5013	4.9292	7.8153	-3.1185	-4.1922	5.4298	-0.2077	4.6298	-3.9420	-0.7116	-4.3374
B-PC wild	5.4566	4.9571	7.8360	-3.2391	-4.1806	5.4691	-0.1777	4.7245	-3.8810	-0.7095	-4.2643
$c = \sqrt{2}$											
PC	4.1888	4.3358	6.5022	-3.0922	0	-2.8677	-0.1382	-0.7091	0	0.0044	-4.2734
B-PC iid	4.9669	4.9163	7.8620	-2.6553	0.0364	3.3680	-0.1565	2.8665	-3.8307	-0.5796	-4.1940
B-PC mbb	1.9115	3.4947	5.9894	-5.2833	-0.2092	-1.6451	-2.6713	1.6354	-5.1156	-1.5391	-5.4715
B-PC par	5.4814	4.9112	7.8513	-2.9829	0.0068	2.9367	-0.1079	2.9481	-4.0332	-0.6997	-4.2561
B-PC wild	5.5658	4.9548	7.8618	-1.8894	0.0159	3.6214	-0.0966	2.9466	-3.5154	-0.5281	-4.1588
$c = \sqrt{\log(T)}$											
PC	0.0051	-0.7080	0	-3.0922	0	0	0	0	0	0	-3.1217
B-PC iid	3.7183	4.8666	7.8077	-2.8018	0.0013	0.8163	-0.0601	1.5048	-2.5888	-0.4947	-4.2081
B-PC mbb	1.8529	2.1063	4.1781	-4.6217	-0.0481	-1.4862	-1.5924	0.9834	-3.1886	-0.5538	-2.3645
B-PC par	5.3101	4.8559	7.7387	-2.9255	-0.0017	0.5590	-0.1202	1.4386	-3.9984	-0.5255	-4.2724
B-PC wild	3.9617	4.8707	7.6689	-1.5246	0.0017	1.0883	0.0253	1.8403	-2.2485	-0.4773	-4.1645

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, iid bootstrap, mbb moving block bootstrap, par parametric bootstrap, wild bootstrap

TABLE 4. PANEL A: PRETESTING, MONTHLY RETURNS, SAMPLE BEGIN 1960M1, FORECAST BEGIN 1980M1

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*
UF	-1.1409	-0.2858	-0.3635	-1.8618	-0.7030	-1.0722	-0.7357	-0.7341	-0.4805	-0.1243	-1.1336
PC	-1.1409	-0.2858	-0.3635	-1.8618	-0.6202	-1.0722	-0.7357	-0.7341	-0.4805	-0.0625	-1.1336
PF	-0.9988	-0.2858	-0.3635	-1.3007	-0.6910	-0.4594	-0.1863	-0.8430	-0.4498	-0.0404	0.0626
PCF	-0.9988	-0.2858	-0.3635	-1.3007	-0.6083	-0.4594	-0.1863	-0.8430	-0.4498	0.0214	0.0626
Table 2A											
Best	-0.5150	-0.1417	-0.1829	-0.6639	-0.5427	-0.2557	-0.0558	-0.6411	0.1681	0.1407	0.1414
$c = 0$ (Table 2A)											
PC	-1.1409	-0.2858	-0.3635	-1.8618	-0.6202	-1.0722	-0.7357	-0.7341	-0.4805	-0.0625	-1.1336
B-PC iid	-1.1051	-0.2078	-0.1829	-1.5888	-0.7957	-1.2028	-0.6509	-0.6411	-0.3505	-0.1746	-1.0737
B-PC mbb	-1.2345	-0.2662	-0.3342	-2.0488	-0.7779	-1.0640	-0.5139	-1.1229	-1.2810	-0.1469	-1.3664
B-PC par	-1.0009	-0.2612	-0.3182	-1.8997	-0.5682	-1.1979	-0.6245	-0.8581	-0.5807	-0.0730	-1.1535
B-PC wild	-1.2462	-0.2406	-0.3575	-1.6692	-0.7537	-0.9067	-0.6426	-0.7640	-0.4902	-0.0882	-1.1156
$c = \sqrt{2}$											
PC	-1.2190	0	-0.0139	-1.3888	-0.6675	-1.8831	-0.4534	-0.7341	-0.6186	0	-1.1336
B-PC iid	-0.9080	-0.0329	-0.2956	0	-0.6214	-1.1634	-0.8181	-0.7133	-0.5478	-0.2050	-1.1967
B-PC mbb	-1.1917	-0.2763	-0.4936	0	-0.6469	-0.9684	-0.5861	-0.9976	-1.3268	-0.1125	-1.4660
B-PC par	-0.9877	-0.1169	-0.3403	-1.5484	-0.4716	-0.9697	-0.5235	-0.7889	-0.4814	-0.0700	-0.9043
B-PC wild	-0.6660	-0.0170	-0.1718	-1.4573	-0.4335	-0.9158	-0.6149	-0.7244	-0.5761	-0.0605	-1.0287
$c = \sqrt{\log(T)}$											
PC	0	0	0	-0.2311	0	-0.2276	0	-0.6187	-0.0863	0	-0.3842
B-PC iid	-0.9977	0.0037	-0.0163	0	-0.4878	-1.1456	-0.6555	-0.5721	-0.2797	-0.1620	-0.9354
B-PC mbb	-1.2659	0.0541	-0.2056	0	-0.4526	-1.1959	-0.5429	-0.8838	-1.1936	-0.0826	-1.1693
B-PC par	-0.4999	-0.0550	-0.0167	-0.8998	-0.2109	-0.5751	-0.2443	-0.4960	-0.0187	-0.0130	-0.5354
B-PC wild	-0.3357	0.1264	0.0131	-0.4045	-0.1493	-0.4101	-0.2482	-0.3391	0.0443	-0.0747	-0.4207

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, iid bootstrap, mbb moving block bootstrap, par parametric bootstrap, wild bootstrap

TABLE 4. PANEL B: PRETESTING, ANNUAL RETURNS, SAMPLE BEGIN 1960M1, FORECAST BEGIN 1980M1

	d/p	*e/p*	*se/p*	*b/m*	*roe*	*tbl*	*lty*	*ts*	*ds*	*inf*	*nei*
UF	-4.3136	-0.5010	0.3993	-15.4237	-3.5269	-1.7396	-0.9190	7.1785	-1.2170	0.4148	-0.5880
PC	-4.3136	-0.5010	0.3993	-15.4237	-3.4753	-1.7396	-1.2369	7.1785	-1.2170	0.6379	-0.5433
PF	-4.5819	-0.5010	0.3993	-7.6233	-3.2032	-1.0312	-0.2809	6.8561	-1.2170	0.4148	-0.0255
PCF	-4.5819	-0.5010	0.3993	-7.6233	-3.1516	-1.0312	-0.5988	6.8561	-1.2170	0.6379	0.0191
Table 2B											
Best	-3.8751	-0.3783	0.4236	-7.2206	-3.0138	-0.9222	0.0894	7.3900	-0.9678	0.9533	0.2817
$c = 0$ (Table 2B)											
PC	-4.3136	-0.5010	0.3993	-15.4237	-3.4753	-1.7396	-1.2369	7.1785	-1.2170	0.6379	-0.5433
B-PC iid	-4.3197	-0.4995	0.3139	-15.4379	-3.4825	-1.8801	-1.7670	7.3900	-1.0159	0.7707	-0.9246
B-PC mbb	-11.3655	-5.9102	-5.6628	-21.1407	-12.6312	-8.3206	-9.9206	2.6758	-7.7161	-1.4252	-7.0377
B-PC par	-4.0600	-0.5025	0.3980	-15.4223	-3.5899	-1.7872	-1.4392	7.2711	-1.2062	0.8453	-0.4205
B-PC wild	-4.2941	-0.3783	0.4155	-15.4183	-3.6969	-1.8216	-1.5513	7.1831	-1.1895	0.7711	-0.7787
$c = \sqrt{2}$											
PC	-3.5660	-1.2005	-0.7730	-14.3487	1.1685	-1.8731	-1.5323	4.0093	-1.3207	0	0
B-PC iid	-4.0416	-0.5752	0.4094	-15.5353	-3.2620	-1.7759	-1.2964	7.0441	-1.2530	0.4279	-0.6606
B-PC mbb	-10.6838	-4.3308	-4.0017	-19.7776	-8.9961	-7.7191	-8.3312	2.7906	-5.9344	-0.5218	-4.3312
B-PC par	-4.4628	-0.4587	0.2880	-15.3856	-3.0867	-1.5177	-1.2586	7.1513	-1.0925	0.5796	-0.4437
B-PC wild	-4.3557	-0.2795	0.2366	-15.1903	-3.1779	-1.4602	-1.2367	7.1954	-1.1339	0.5051	-0.5505
$c = \sqrt{\log(T)}$											
PC	-0.5856	0	0	-7.4699	-0.1978	0	0	0	0	0	0
B-PC iid	-3.9864	-0.3348	0.3853	-15.4303	-1.7450	-1.3947	-0.5979	6.9187	-1.1328	0.1808	-0.1660
B-PC mbb	-6.7599	-1.6642	-1.9351	-14.7411	-4.0987	-5.3838	-5.7042	2.3074	-2.6443	-0.2536	-2.0288
B-PC par	-4.0257	-0.4510	0.3136	-15.3386	-1.8249	-1.0201	-0.4974	6.7379	-0.9582	0.2216	-0.1947
B-PC wild	-3.9929	-0.3850	0.3872	-15.2339	-1.9632	-0.9748	-0.5849	6.8055	-1.1899	0.2600	-0.3153

Data: *d/p* dividend-price ratio, *e/p* earnings-price ratio, *se/p* smooth earnings-price ratio, *b/m* book-to-market, *roe* return on equity, *tbl* T-Bill rate, *lty* long-term yield, *ts* term spread, *ds* default spread, *inf* CPI-inflation, *nei* net equity issuance

Forecast types: UF unconstrained forecast, PC positive coefficient constraint, PF positive forecast constraint, PCF positive coefficient and positive forecast constraint

Bagging types: B-X bagged forecast of type X, iid bootstrap, mbb moving block bootstrap, par parametric bootstrap, wild bootstrap

(E. Hillebrand) DEPARTMENT OF ECONOMICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, USA.
E-mail address: erhil@lsu.edu

(T.-H. Lee) DEPARTMENT OF ECONOMICS, UNIVERSITY OF CALIFORNIA AT RIVERSIDE, RIVERSIDE, USA.
E-mail address: taelee@ucr.edu

(M. C. Medeiros) DEPARTMENT OF ECONOMICS, PONTIFICAL CATHOLIC UNIVERSITY OF RIO DE JANEIRO, RIO DE JANEIRO, RJ, BRAZIL.
E-mail address: mcm@econ.puc-rio.br