Online Appendix to
A Fresh Look at Return Predictability
Using a More Efficient Estimator

Travis L. Johnson*
McCombs School of Business
The University of Texas at Austin

July 2018

Abstract

I present additional results on return predictability using weighted least squares with ex ante variance (WLS-EV). Online Appendices A and B provide further evidence of the out-of-sample forecasting improvements afforded by WLS-EV relative to OLS. Online Appendix C shows OLS and WLS-EV estimates of the predictability afforded by the variance risk premium are weaker using alternative proxies that reduce the influence of outliers. Online Appendix D shows WLS-EV estimates of the predictability afforded by the variance risk premium are also insignificant in non-US countries. Online Appendix E shows how my results are affected by different methodologies for addressing overlapping observations in WLS-EV regressions.

*Thanks to Aydoğan Altı, Svetlana Bryzgalova, Jonathan Cohn, Andres Donangelo, John Griffin, Michael Halling (Imperial discussant), Yufeng Han (SFA discussant), Michael Johannes (TFF discussant), Bryan Kelly, Arthur Korteweg, Xiang (Nicole) Liu, Zack Liu, Laura Starks, Sheridan Titman, Rasmus Varneskov (AFA discussant), James Weston, and seminar participants at the American Finance Association Meeting, the 11th Annual Hedge Fund Conference at Imperial College, the University of California San Diego, Massachusetts Institute of Technology, the Southern Finance Association Meeting, the Texas Finance Festival, The University of North Carolina at Chapel Hill, and The University of Texas at Austin for their helpful comments. I conducted part of this research while visiting the MIT Sloan School of Management. Data and code are available on my website, http://travislakejohnson.com. Contact: travis.johnson@mccombs.utexas.edu or 2110 Speedway Stop B6600, Austin, TX 78712.

Appendix A. Evolution of out-of-sample coefficients and forecast improvements

In this Online Appendix, I discuss the evolution of out-of-sample predictability coefficients, $\hat{b}_\tau$, for both OLS and WLS-EV, and cumulative forecast error improvements from using WLS-EV relative to OLS. I compute the $\hat{b}_\tau$ for each predictor following the procedure described in Section 3.2 of the paper. For each predictor and forecast horizon, I compute the cumulative forecast improvement through month τ, CFI_τ, following Goyal and Welch (2008) as:

$$\text{CFI}_\tau = \sum_{m=1}^{\tau-h} \left(e_{\text{OLS}}(m,x)^2 - e_{\text{WLS-EV}}(m,x)^2 \right), \tag{1}$$

$$e_{\text{OLS}}(m,x) \equiv r_{m+1,h} - \mathbb{E}_m^{\text{OLS}}\left(r_{m+1,h}|x\right), \tag{2}$$

$$e_{\text{WLS-EV}}(m,x) \equiv r_{m+1,h} - \mathbb{E}_m^{\text{WLS-EV}}\left(r_{m+1,h}|x\right). \tag{3}$$

Higher values of CFI_τ indicate OLS forecast had larger errors relative to WLS-EV forecasts, meaning WLS-EV performed better out of sample.

Appendix Figure 1 shows the evolution of $\hat{b}_\tau$ and CFI_τ for each of the 16 predictors for predicting both next-month and next-year returns. The rolling coefficient plots indicate that, in general, the WLS-EV predictability coefficients are more stable than their OLS counterparts, and closer to their full-sample values earlier for earlier τ. This stability and faster convergence both stem from the efficiency gains associated with WLS-EV.

The CFI_τ plots illustrate the accumulation of the out-of-sample performance gains afforded by WLS-EV. Echoing the evidence in Table 3 of the paper, by the end of the sample WLS-EV has lower cumulative forecast errors than OLS, meaning CFI is positive, for 11 (12) predictors of next-month (next-year) returns. Appendix Figure 1 shows that these out-of-sample forecast improvements are disproportionately concentrated in the early and middle parts of the sample when, as the $\hat{b}_\tau$ plots show, WLS-EV point estimates are closer to the full-sample values than the OLS point estimates. The plots for the term spread (tms) and default spread (dfy), both on Page 20, exemplify these patterns.

Appendix B. Out-of-sample certainty equivalents

In Section 3.2 of the paper, I show that using WLS-EV results in higher out-of-sample R^2 (OOS R^2) than OLS for most traditional return predictors. In this Online Appendix, I estimate the economic magnitude of these out-of-sample performance gains by computing their impact on certainty equivalents for an investor optimizing their portfolio in each period conditional on time-varying forecasts of both volatility and expected returns. I find WLS-EV results in certainty equivalent gains between 20bp and 50bp per year relative to OLS, which are substantial compared to baseline certainty equivalents around 6%.

B.1. Computing certainty equivalents

I compute optimal portfolio choices for an investor with one-period mean-variance preferences allocating capital between the risk-free asset and the US equity market. Each month m, for holding periods $h = 1$ and $h = 12$, the investor chooses risky asset weight $w_{m,h}$, resulting in next-period portfolio return:

$$r_{p,m+1,h}(w_{m,h}) = r_{f,m+1,h} + w_{m,h} \left(r_{mkt,m+1,h} - r_{f,m+1,h} \right), \tag{4}$$

where $r_{f,m+1,h}$ is the risk-free rate in months $m+1$ through $m+h$, $r_{mkt,m+1,h}$ is the equity market porfolio's return in months $m+1$ through $m+h$, and all returns are net of the initial investment but not logged. I assume next period's risk-free rate $r_{f,m+1,h}$ is observable to the investor at the end of month m, meaning the conditional mean and variance of the portfolio return satisfy:

$$\mathbb{E}_m \left(r_{p,m+1,h} \right) = r_{f,m+1,h} + w_{m,h} \left(\mathbb{E}_m(r_{mkt,m+1,h}) - r_{f,m+1,h} \right), \tag{5}$$

$$\mathrm{Var}_m \left(r_{p,m+1,h} \right) = w_{m,h}^2 \mathrm{Var}_m(r_{mkt,m+1,h}). \tag{6}$$

I follow the approach in Johannes, Korteweg, and Polson (2014) (JKP) and assume the investor chooses optimal weight $w^*_{m,h}$ based on the following mean-variance preferences:

$$w^*_{m,h} = \arg\max_{w} \mathbb{E}_m\left(r_{p,m+1,h}(w)\right) - \frac{\gamma}{2}\text{Var}_m\left(r_{p,m+1,h}(w)\right), \tag{7}$$

$$\Rightarrow w^*_{m,h} = \frac{1}{\gamma} \cdot \frac{\mathbb{E}_m\left(r_{mkt,m+1,h} - r_{f,m+1,h}\right)}{\text{Var}_m\left(r_{mkt,m+1,h}\right)}. \tag{8}$$

For conditional expected excess returns, I consider seven possible techniques investors could use to obtain fully out-of-sample estimates of future excess returns:

(1) The average past excess return observed prior to τ.

(2–3) Out-of-sample forecast $a_\tau + b_\tau \cdot x_\tau$, where a_τ and b_τ are estimated using an OLS or WLS-EV regression of $r_{m,s+1} - r_{f,s+1}$ on x_s with only observations satisfying $s < \tau - h$.

(4–5) Out-of-sample forecast $a_\tau + b_\tau \cdot x_\tau$ using OLS or WLS-EV, adjusted using the Campbell and Thompson (2008) (CT) approach.

(6–7) Out-of-sample forecast $a_\tau + b_\tau \cdot x_\tau$ using OLS or WLS-EV, adjusted using the Pettenuzzo, Timmermann, and Valkanov (2014) (PTV) approach.

I pair these out-of-sample excess return forecasts with the out-of-sample variance forecasts I use to estimate WLS-EV. Specifically, I use the fitted value from a regression of $\text{RV}_{s+1,s+h}$ on RV_s and $\text{RV}_{s-11,s}$, where RV_s is the realized daily return variance in month s and $\text{RV}_{s+k_1,s+k_2}$ is the realized variance in months $s + k_1$ through $s + k_2$, using only observations $s < \tau - 1$. JKP shows that allowing investors to adjust portfolio weights as a function of time-varying volatility results in substantial certainty equivalent games. I control for the role of out-of-sample variance forecasts by using the same sample conditional variance forecast across all expected return methodologies, including the rolling average baseline, allowing me to focus on the role of expected return forecasts.

I substitute out-of-sample expected return forecasts and variance forecasts into Equation (12) to compute the optimal portfolio weight $w^*_{m,h}$ for each time period in the post-training sample. Using these, I compute the time series of portfolio returns these weights imply $r_{p,m+1,h}(w^*_{m,h})$. From these returns, I arrive at the investor's certainty equivalent:

$$CE = \mathbb{E}(r_{p,m+1,h}(w^*_{m,h})) - \frac{\gamma}{2}\mathrm{Var}\left(r_{p,m+1,h}(w^*_{m,h})\right). \tag{9}$$

Note that I estimate certainty equivalents from the unconditional moments of the full sample of realized portfolio returns, rather than the average period-by-period conditional certainty equivalent the investor expects when making their portfolio choice. The approach based on unconditional moments, also followed in CT and JKP, is more conservative because it requires the out-of-sample conditional moments match the subsequent return distributions.

B.2. Results

For each predictor I study in Section 3 of the paper, I compute the out-of-sample certainty equivalent CE for all seven return forecasting methodologies listed above. Following the approach I use in Section 3, I estimate out-of-sample forecasts in a monthly sample from 1927-2015 starting after a 20-year training period. For each month, I compute both next-month and next-year return moments. Based on these, I compute optimal portfolio weights using Equation (12) and γ set so the unconditionally optimal weight in the sample is $\frac{1}{2}$.[1] Following CT and JKP, I impose that due to leverage constraints, $w^*_{m,h}$ cannot exceed 1.5.

For each predictor-methodology pair, Appendix Table 1 presents ΔCE, the improvement in annualized certainty equivalent relative to the baseline certainty equivalent the investor receives when ignoring the predictor and using average past returns as their out-of-sample forecast. The baseline approach results in annualized certainty equivalents of 5.88% for next-month returns and 5.61% for next-year returns.

The first result in Appendix Table 1 is that using OOS forecasts based on OLS produces *worse* average CEs for these predictors, replicating the pessimistic conclusion about OLS estimates in Goyal and Welch (2008), CT, and Table 3 of the paper. I also find that both

[1] This implies $\gamma = 2\frac{\mathbb{E}_m(r_{mkt,m+1,h} - r_{f,m+1,h})}{\mathrm{Var}_T(r_{mkt,m+1,h})}$, where both moments are computed in the full sample. I find $\gamma = 6.92$ ($\gamma = 5.61$) for next-month (next-year) returns, similar to the values used in CT and JKP.

the CT and PTV approaches result in certainty equivalent gains for the average predictor between 15bp and 45bp, echoing the evidence in those papers.

The main result in Appendix Table 1 is that WLS-EV offers substantial CE improvements when compared to OLS. Without the CT or PTV adjustments, the average WLS-EV improvement is particularly substantial, 36bp annualized for next-month returns and 58bp for next-year returns. These improvements are smaller for the CT approach and absent for the PTV approach because these adjustments already mitigate the influence of high-volatility observations. As discussed in the paper, the advantage of WLS-EV relative to other approaches to improving out-of-sample return predictability is its ease of implementation and clear econometric motivation.

In addition to confirming the patterns we find in OOS R^2, the changes in certainty equivalents in Appendix Table 1 provide a sense of the economic magnitude of these improvements. The natural point of comparison is the certainty equivalent from ignoring the return predictors, 5.88% or 5.61%. Most of this certainty equivalent is from the risk-free rate alone, which averaged 3.50% on average in my sample. The remainder, 2.38% or 2.11%, is the value added by investing in equities. Relative to these values, the improvement offered by WLS-EV is substantial, between 15% and 27% of the value the investor gets from equity markets.

Appendix C. Alternative proxies for variance risk premia

In this Online Appendix, I revisit the evidence on the variance risk premium as a return predictor with alternative proxies ($\widehat{\text{VRP}}$) designed to reduce the frequency of extreme or implausible $\widehat{\text{VRP}}$. I show that these alternative proxies are weaker predictors of future returns than the proxies used in Drechsler and Yaron (2011) (DY) and Bollerslev, Tauchen, and Zhou (2009) (BTZ), that the OLS estimates are insignificant in many cases, and that the WLS-EV estimates remain insignificant in all cases.

The results in Section 4 of the paper document that WLS-EV estimates of the return predictability afforded by proxies for the variance risk premium are statistically insignificant.

This indicates the OLS evidence of predictability is concentrated among a few observations with extreme estimates of variance risk premia and high ex ante volatility. The scatterplots in Figure 3 show that the extreme estimates of variance risk premia are enormously positive or negative, taking values that are economically implausible. The WLS-EV approach downweights these observations because they have a low signal-to-noise ratio, but another robustness check is to discretize or rescale the estimates of variance risk premia themselves.

The first alternative approach I employ is discretizing the variance risk premia estimates into deciles. This approach has the advantage of putting minimal weight on extreme variance risk premia estimates, but suffers from the disadvantage that the relation between expected returns and variance risk premia deciles is unlikely to be linear.

Appendix Table 2 shows that for both the DY and BTZ variance risk premia estimates, discretizing into deciles further weakens the evidence for return predictability for both OLS and WLS-EV estimates. For the DY approach, the WLS-EV estimates are near-zero and far from significant, and even the OLS estimates are statistically insignificant. For the BTZ approach, both the OLS and WLS-EV estimates have higher p-values than their continuous counterparts, though the OLS estimates are marginally significant.

The second alternative approach is winsorizing the variance risk premia estimates below at zero. The models in DY and BTZ both predict that the true variance risk premia is always positive.[2] Both papers' empirical proxies are mostly, but not always, positive. Worryingly, these negative variance risk premia estimates tend to be concentrated in bad times (e.g. in October, 2008), which is counter-intuitive, and constitute some of the most extreme values. As a simple robustness check to address these observations, I replace the negative observations of the DY and BTZ variance risk premia with zero.

Appendix Table 2 shows that estimates of variance risk premia are worse predictors when forced to be positive. Without negative values, in OLS regressions the DY estimates are insignificant predictors of next-month and next-quarter market returns. The BTZ estimates

[2] These papers define the variance risk premia as $\text{Var}^{\mathbb{Q}}(r_{t+1}) - \text{Var}^{\mathbb{P}}(r_{t+1})$, making it positive. Others define it as the risk premia associated with long-variance assets, making it negative.

are are also much weaker predictors without negative values, though remain statistically significant in OLS regressions. For both DY and BTZ approaches, WLS-EV estimates are weaker without negative values and remain statistically insignificant.

Appendix D. International evidence on variance risk premia's predictability

Bollerslev et al. (2014) shows the return predictability afforded by proxies for variance risk premia in Bollerslev, Tauchen, and Zhou (2009) and Drechsler and Yaron (2011) is significant in six of seven non-US countries with available data. In this Online Appendix, I replicate their results in an extended 2000-2013 sample period. However, I also show that the results are mostly insignificant when using alternative weighting schemes, sampling frequency, or standard error computations.

D.1. Data

The stock and volatility index data for these analyses are from Datastream. I compute index returns directly from each country's index level. I follow Bollerslev et al. (2014) and estimate the conditional variance risk premia on day d using:

$$\mathrm{V\hat{R}P}_d \equiv \mathrm{IV}_d - \mathrm{RV}_{d-20,d} \tag{10}$$

$$\mathrm{RV}^2_{d-20,d} \equiv \sum_{s=0}^{20} r^2_{d-s}, \tag{11}$$

where r_{d-s} is the percent log index return s trading days prior to d and IV_d is the volatility index squared and divided by twelve to make it an monthly percent variance. The sample period is limited by the availability of the country-specific volatility indices, which go back to 2000 for all seven countries and end in 2011 for Belgium.

The risk-free rate data are the shortest-duration yields available on the respective central banks' websites. For the Eurozone countries (France, Germany, Netherlands, Belgium), I use the yield for 3-month AAA-rated euro area central government bonds from the ECB yield curve. For the UK, I use the 1-month yield from the Bank of England. For Japan, I use the

1-year yield from Ministry of Finance Japan. For Switzerland, I use the 1-year yield from the Swiss National Bank. As in Bollerslev et al. (2014), my results are robust to using log index returns without subtracting the log risk-free rate.

D.2. Estimation

Bollerslev et al. (2014) estimates the regression equation:

$$r_{t+1,t+h} = a + b \cdot \hat{\mathrm{VRP}}_t + \epsilon_{t+h}, \tag{12}$$

using OLS on a monthly sample for $h = 1$ though $h = 12$, overlapping when necessary, with standard errors computing using Newey and West (1987). I repeat this analysis on my longer sample, and also consider alternative estimation weights, sampling frequencies, and standard error calculations.

The two weighting schemes I use are equal-weighting (OLS) and ex ante variance (WLS-EV), as described in Section 2 of the paper. For WLS-EV, the ex ante variance estimates I use are fitted values from a first stage regression of RV_{t+1} on IV_t. I estimate (12) using both overlapping monthly and overlapping daily samples. In the latter-case, I predict next–$21h$ day returns, where h is the monthly forecast horizon. Provided the standard errors properly account for overlap, because $\hat{\mathrm{VRP}}_t$ has meaningful daily variations, daily sampling should result in a larger sample and therefore a more a powerful test of the null.

Finally, I compute standard errors using the asymptotic Newey and West (1987) approach used in Bollerslev et al. (2014) and using the heteroskedastic simulation approach described in Appendix B of the paper, the latter of which uses the Hodrick (1992) mapping between overlapping regressions and non-overlapping regressions with a rolling average of the predictor.

D.3. Results

I estimate Equation (12) for a eight different combinations of estimator, sampling period, and standard error calculation, summarized here:

Standard Errors	Sampling Frequency	Estimator	# countries 5% sig $h=3$	$h=6$	Any h
Asymptotic Newey and West (1987)	Monthly	OLS	3	5	5
Asymptotic Newey and West (1987)	Monthly	WLS-EV	1	1	1
Asymptotic Newey and West (1987)	Daily	OLS	2	4	5
Asymptotic Newey and West (1987)	Daily	WLS-EV	0	0	0
Simulated Hodrick (1992)	Monthly	OLS	0	0	0
Simulated Hodrick (1992)	Monthly	WLS-EV	0	1	1
Simulated Hodrick (1992)	Daily	OLS	0	0	0
Simulated Hodrick (1992)	Daily	WLS-EV	0	0	0

The first combination is the approach used in Bollerslev et al. (2014), OLS with monthly sampling and Newey and West (1987) standard errors. The results are summarized in Panel A of Appendix Table 3. Like Bollerslev et al. (2014), I find evidence of predictability for every non-US country except for Japan, with t-stats and R^2 peaking for h between 3 and 6 months using this specification. I find slightly weaker evidence than Table 3 in Bollerslev et al. (2014) due to the addition of 2012 and 2013 to the relatively-short 2000-2011 sample.

Panel B of Appendix Table 3 shows that WLS-EV estimates, by contrast, are significant for only one of the seven indices (BEL) despite using the same standard error approach and sampling frequency as Bollerslev et al. (2014). Like the US-based results in Section 5, the decrease in predictability is primarily due to WLS-EV point estimates being much smaller than OLS estimates. For example, the AEX VRP prediction for a 5-month horizon has an OLS coefficient of 0.17 with a t-stat of 2.22, while WLS-EV results in a coefficient of 0.04 with a 0.32 t-stat. This decline indicates OLS estimates are driven by a few ex ante volatile observations which WLS-EV downweights because their signal-to-noise ratio is low.

I repeat the analysis using a daily overlapping sample that uses all available observations of $\hat{\text{VRP}}_t$ rather than discarding intra-month observations. Panel C of Appendix Table 3 shows that including the additional intra-month observations generally weakens the magnitude of the OLS coefficients while increasing the standard errors. For example, the same AEX $h = 5$ coefficient is 0.11 in Panel C, substantially smaller than the corresponding monthly sampling OLS estimate, 0.17. As a result, only two of the seven countries have significant predictability for $h = 3$, three for $h = 6$, and five for any h.

Like OLS, the daily sampling frequency weakens the already-sparse WLS-EV evidence for predictability. Out of the 56 country-h pairs tested, only one is significant at the 10% level (BEL 20 for $h = 12$) and none are significant at the 5% level.

An important feature of the 2000-2013 sample for all countries is the financial crisis and subsequent worldwide turmoil. During such incidents, equity volatility spikes enormously and, as Bollerslev et al. (2014) shows, $\hat{\text{VRP}}_t$ tend to be extremely positive or extremely negative. A few observations with the combination of extreme x_t and extremely volatile r_{t+1} have a large influence on OLS coefficients. WLS-EV adjusts for this automatically, but another approach is to compute standard errors using heteroskedastic simulations that retain the observed x_t and conditional variances $\hat{\sigma}_t^2$, as I detail in Appendix B of the paper.

Appendix Table 4 shows that standard errors based on these simulations result in very little evidence of return predictability even using OLS. Across all four panels, VRP_t is a significant predictor at the 5% level in only 1 of the 168 methodology-country-h pairings tested (monthly sampling, WLS-EV, BEL 20, $h = 6$). These results indicate the failure of WLS-EV to reject the no-predictability null using Newey and West (1987) is due to the predictability evidence being concentrated in a few observations with extreme x_t values and volatile future returns that serve as poor proxies for expected returns.

The results in Tables 4 and 5 of the paper, Figure 3 of the paper, and Online Appendices C and D all indicate that there is no evidence for a linear relation between conditional equity and variance risk premia when using more-efficient WLS-EV regressions. While this non-result could indicate the OLS evidence of a relation is spurious, there are many other potential interpretations. One is the relation between equity and variance risk premia is positive but weaker than indicated by OLS, meaning that we cannot detect it in the relatively-short 1990–2015 sample periods. Another is the relation between equity and variance risk premia is time-varying or nonlinear, meaning linear regressions are misspecified.

Given these possible interpretations, I view my variance premia results as indicating the existing linear regression analysis is inconclusive. We therefore need further analysis with additional data, or alternative nonlinear or time-varying specifications, to reach a conclusion about the predictive value of the variance risk premium for future equity returns.

Appendix E. Overlapping regressions

Throughout the paper, I address overlapping samples using a trick from Hodrick (1992) whereby I map the overlapping regression into an equivalent non-overlapping regression of one-period returns on a rolling sum of past predictors, and then scale the resulting coefficients appropriately. I describe this approach in Section 2.2 and Appendix A.1 of the paper.

An alternative approach would be to directly predict future multi-period returns, and correct the standard errors for overlapping returns using Newey and West (1987) or the Hodrick (1992) error correction (as opposed to the mapping in Hodrick (1992) I use). This is the standard approach in the literature when using OLS.

In this Online Appendix, I develop and apply a WLS-EV equivalent to the standard overlapping regression approach, illustrate its shortcomings relative to the Hodrick (1992) mapping I use in the paper, and show that my main results are mostly unchanged when using this alternative approach. I also show that, as argued in Appendix A.1 of the paper, OLS point estimates from the two approaches are almost identical in the settings I study.

The standard overlapping approach consists of a regression of next-h period returns $r_{t+1,t+h}$ on today's predictors X_t:

$$r_{t+1,t+h} = X_t \cdot \beta + \epsilon_{t+1,t+h}. \tag{13}$$

The OLS estimator equally-weights each observation and corrects standard errors for the autocorrelation in $\epsilon_{t+1,t+h}$ driven by overlapping left-hand side observations. I implement this correction using Newey and West (1987) with $1.5h$ lags.

I consider a WLS-EV version of this approach that estimates Equation (13) weighting observations by $\frac{1}{\hat{\sigma}^2_{t,h}}$, where $\hat{\sigma}^2_{t,h}$ is an estimate of next-h period return variance. I use $\hat{\sigma}^2_{t,h}$

from a first-stage regression of next-h period realized variance on past-month and past-year variance, the same predictors I use for RV $\hat{\sigma}_m^2$ and $\hat{\sigma}_d^2$ in the paper.

For each method of adjusting for overlap, I compute both asymptotic and simulated standard errors. For the Hodrick (1992) mapping approach, standard errors are computed as described in the paper. For the overlapping alternative, the asymptotic standard errors use Newey and West (1987) with $1.5h$ lags, a standard practice in the literature. The simulated standard errors are based on the same simulations I use throughout the paper, which retain observed x_t and $\hat{\sigma}_t^2$ but re-draw returns heteroskedastically. Within each simulated sample, I estimate coefficients and standard errors using overlapping OLS and WLS-EV. Finally, I compute simulated standard errors as the cross-simulation standard deviation of point estimates, and simulated p-values as the fraction of simulated samples with larger-magnitude point estimates ($|\hat{b}_{\text{sim}}| > |b|$).

I apply this alternative overlapping estimator to four important predictors from the paper, and compare the results to what I find with the results I find in the paper with using the Hodrick (1992) mapping. The four predictors are the dividend to price ratio (dp), perhaps the most-studied predictor in finance; the term spread (tms), one of the predictors with consistently significant in- and out-of-sample WLS-EV tests; the consumption to wealth ratio (cay), another significant predictor but with a shorter sample period; and the Bollerslev, Tauchen, and Zhou (2009) $\hat{\text{VRP}}$, an example of a predictor that becomes insignificant when using WLS-EV instead of OLS.[3]

The results of this analysis, presented in Appendix Table 5, yield thre main conclusions. The first is that the two approaches to addressing overlap produce OLS estimates that are approximately equal, but not identical, as argued analytically in Appendix A.1 of the paper. The biggest difference in point estimates is for cay, and even that is only 16% of a standard error. The two approaches result in noticeably different coefficients when using WLS-EV because they use different weights, based on next-period variance for the Hodrick (1992)

[3]In untabulated tests, I find the Drechsler and Yaron (2011) $\hat{\text{VRP}}$ is even less sensitive to the alternative overlapping specification, with WLS-EV results remaining far from statistically significant.

mapping and next-h period for the overlapping approach.

The second conclusion from Appendix Table 5 is that overlapping WLS-EV suffers from two important drawbacks relative to the Hodrick (1992) mapping I use in the paper. The first is that asymptotic standard errors for the overlapping regressions approach are understated relative to their bootstrapped counterparts for both OLS and WLS-EV, indicating the Newey and West (1987) correction suffers from size distortions, as expected given the evidence in Hodrick (1992) and Ang and Bekaert (2006). The asymptotic standard errors for the Hodrick (1992) mapping, by contrast, are sometimes slightly over- or under-stated but overall are much more reliable. The second drawback is that the overlapping WLS-EV implementation has higher simulated standard errors, particularly for the variance risk premia predictor, meaning it is less efficient than the Hodrick (1992)-style implementation I use. My approach is more efficient because next-period variance forecasts are more accurate than next-h period forecasts, which are required in traditional overlapping regressions. These drawbacks lead me to use the Hodrick (1992) mapping throughout the paper.

Finally, and most importantly, Appendix Table 5 shows that in all four applications, both approaches yield very similar results when using simulated p-values. The biggest difference is for VRP, which becomes marginally significant when using the overlapping approach (p-value goes from 19.4% to 10.0%). The overall conclusions about return predictability remain that WLS-EV results in stronger evidence for predictability by traditional predictors and insignificant evidence for predictability by variance risk premia. Other results in the paper either do not use overlapping samples, or also yield the same conclusions in untabulated tests using the alternatives overlapping version of WLS-EV.

References

Ang, Andrew and Geert Bekaert, 2006, Stock return predictability: Is it there?, *The Review of Financial Studies* 20, 651–707.

Bollerslev, Tim, James Marrone, Lai Xu, and Hao Zhou, 2014, Stock return predictability and variance risk premia: statistical inference and international evidence, *Journal of Financial and Quantitative Analysis* 49, 633–661.

Bollerslev, Tim, George Tauchen, and Hau Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463.

Campbell, John Y and Samuel B Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average?, *Review of Financial Studies* 21, 1509–1531.

Drechsler, Itamar and Amir Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1.

Goyal, Amit and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Hodrick, Robert J, 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

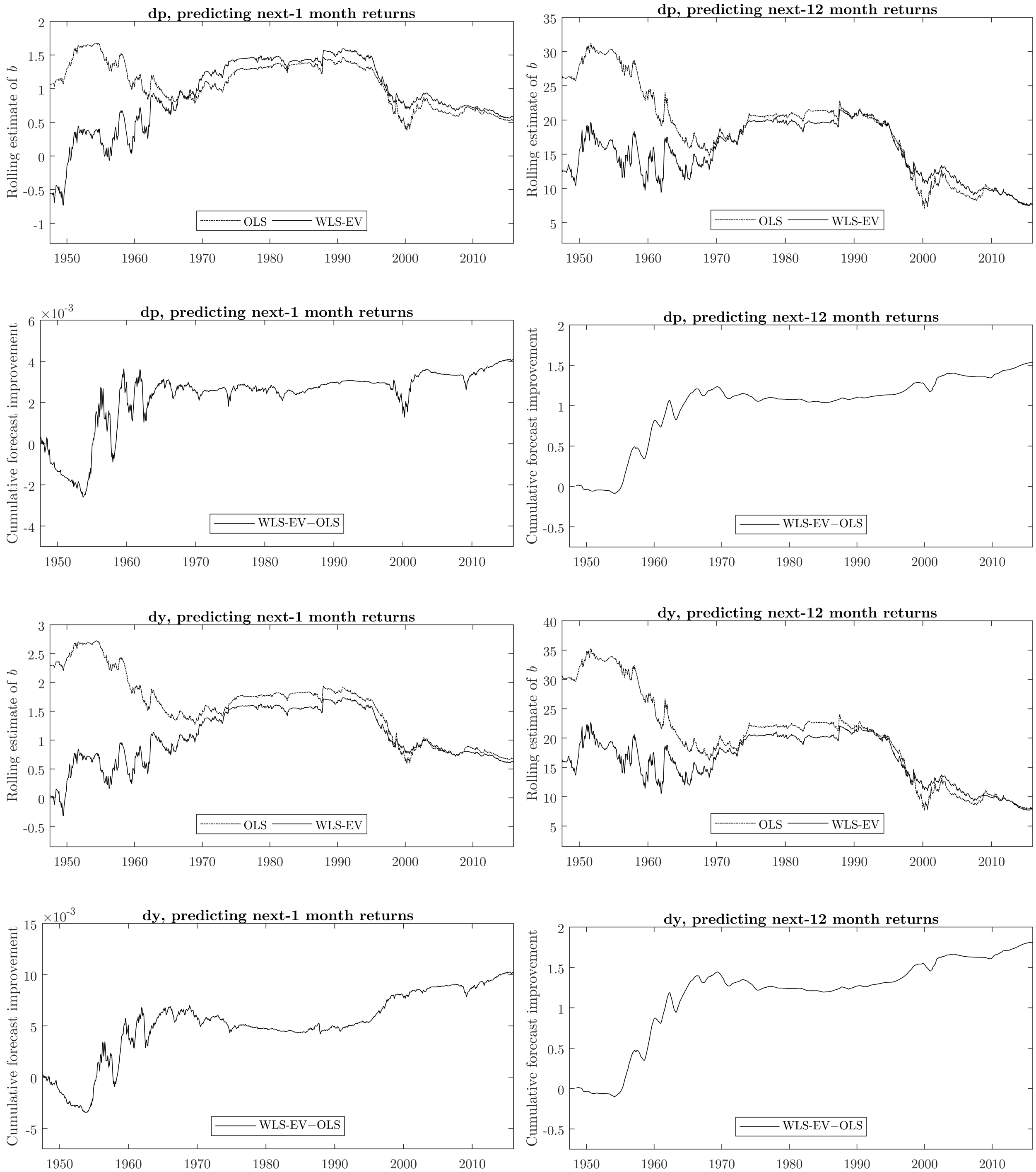
Johannes, Michael, Arthur Korteweg, and Nicholas Polson, 2014, Sequential learning, predictability, and optimal portfolio returns, *The Journal of Finance* 3, 161–174.

Newey, Whitney K and Kenneth D West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–08.

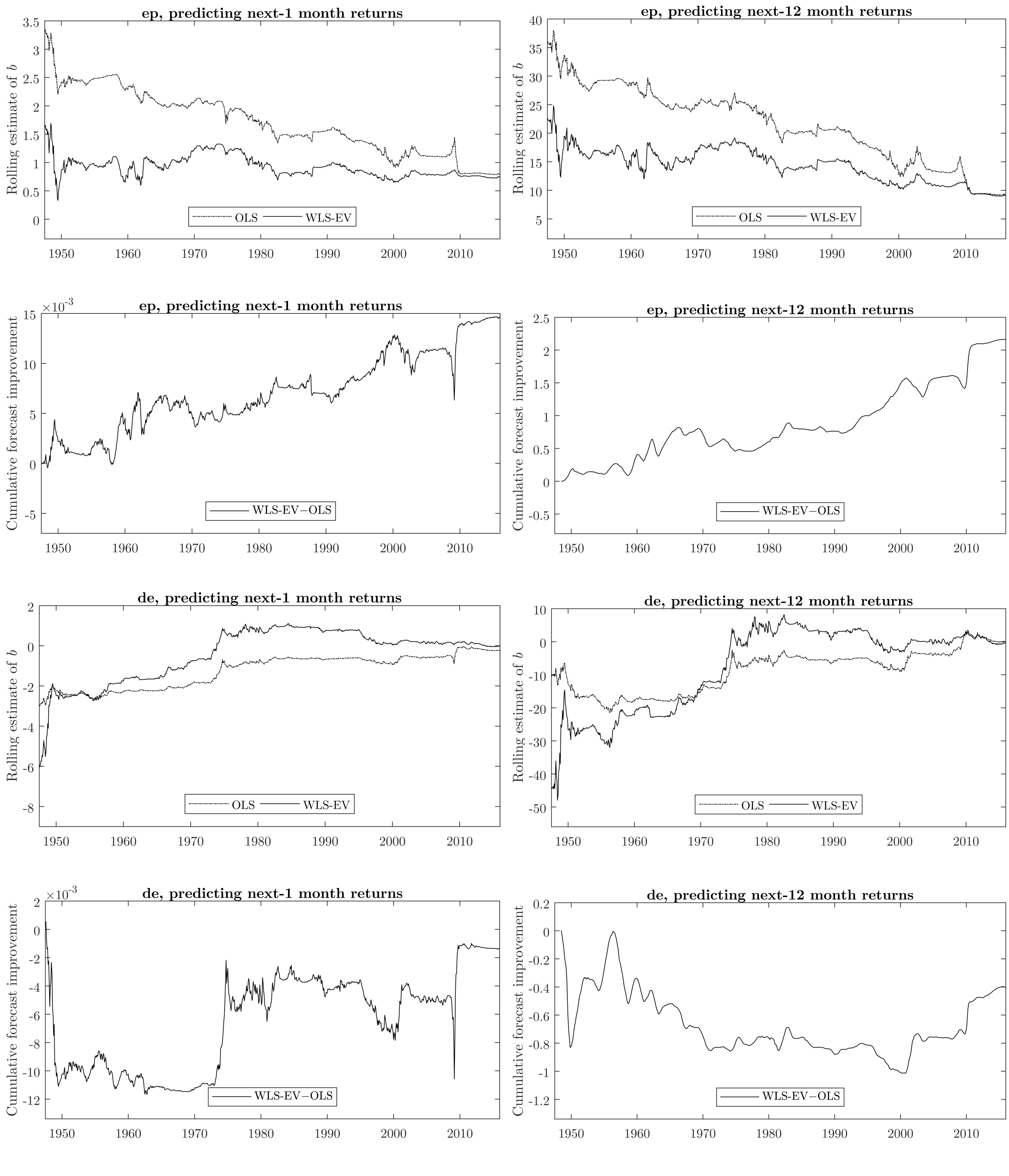
Pettenuzzo, Davide, Allan Timmermann, and Rossen Valkanov, 2014, Forecasting stock returns under economic constraints, *Journal of Financial Economics* 114, 517–553.

Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements

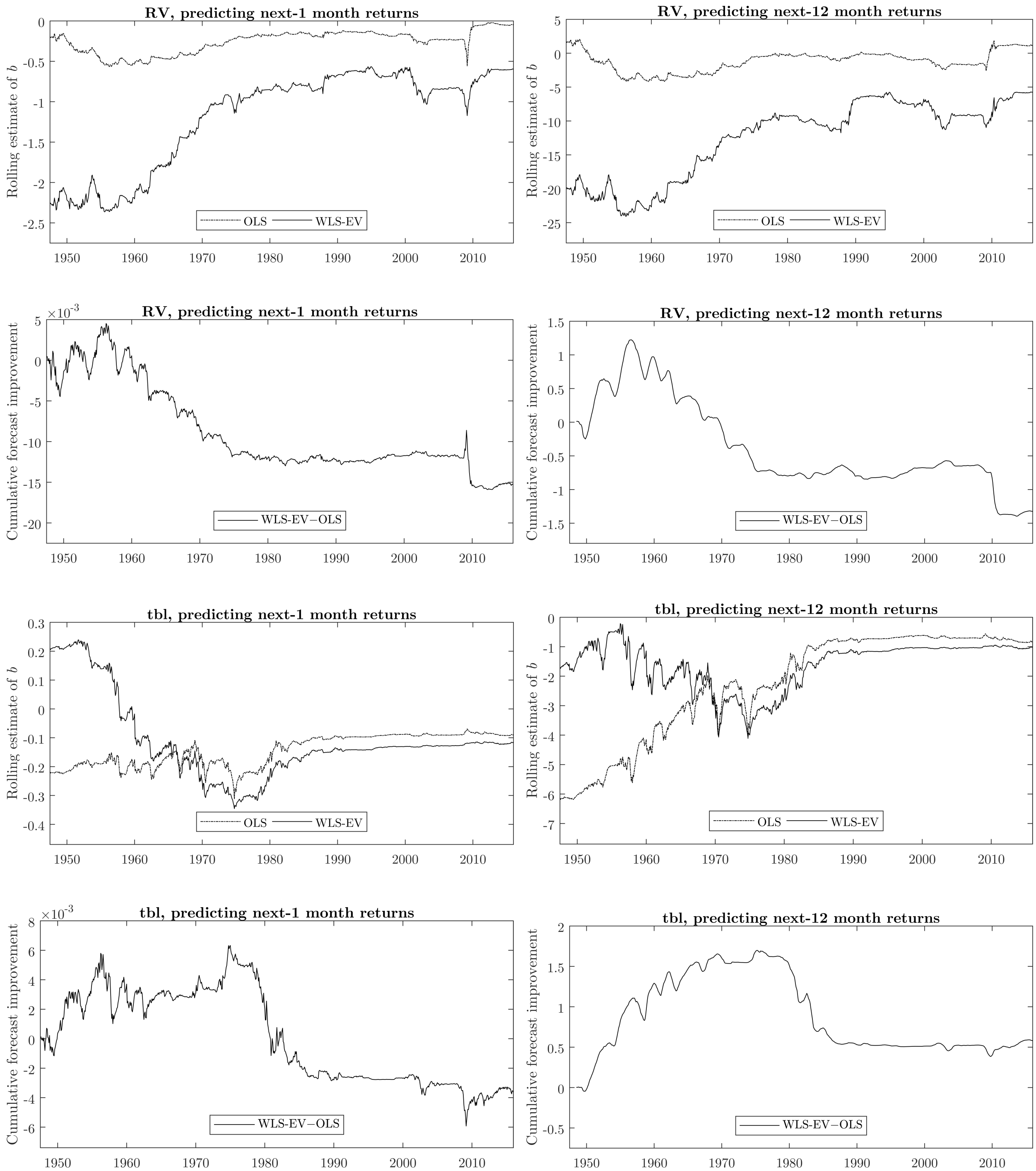
This figure presents rolling point estimates $\hat{b}_\tau$ of return predictability coefficients using only data available prior to τ. From these point estimates, I compute and plot the cumulative forecast improvement from using WLS-EV instead of OLS, defined as the cumulative difference between squared forecast errors from OLS and WLS-EV for all observations prior to τ. I include plots for all 16 predictors from Goyal and Welch (2008) for both next-month and next-year returns.



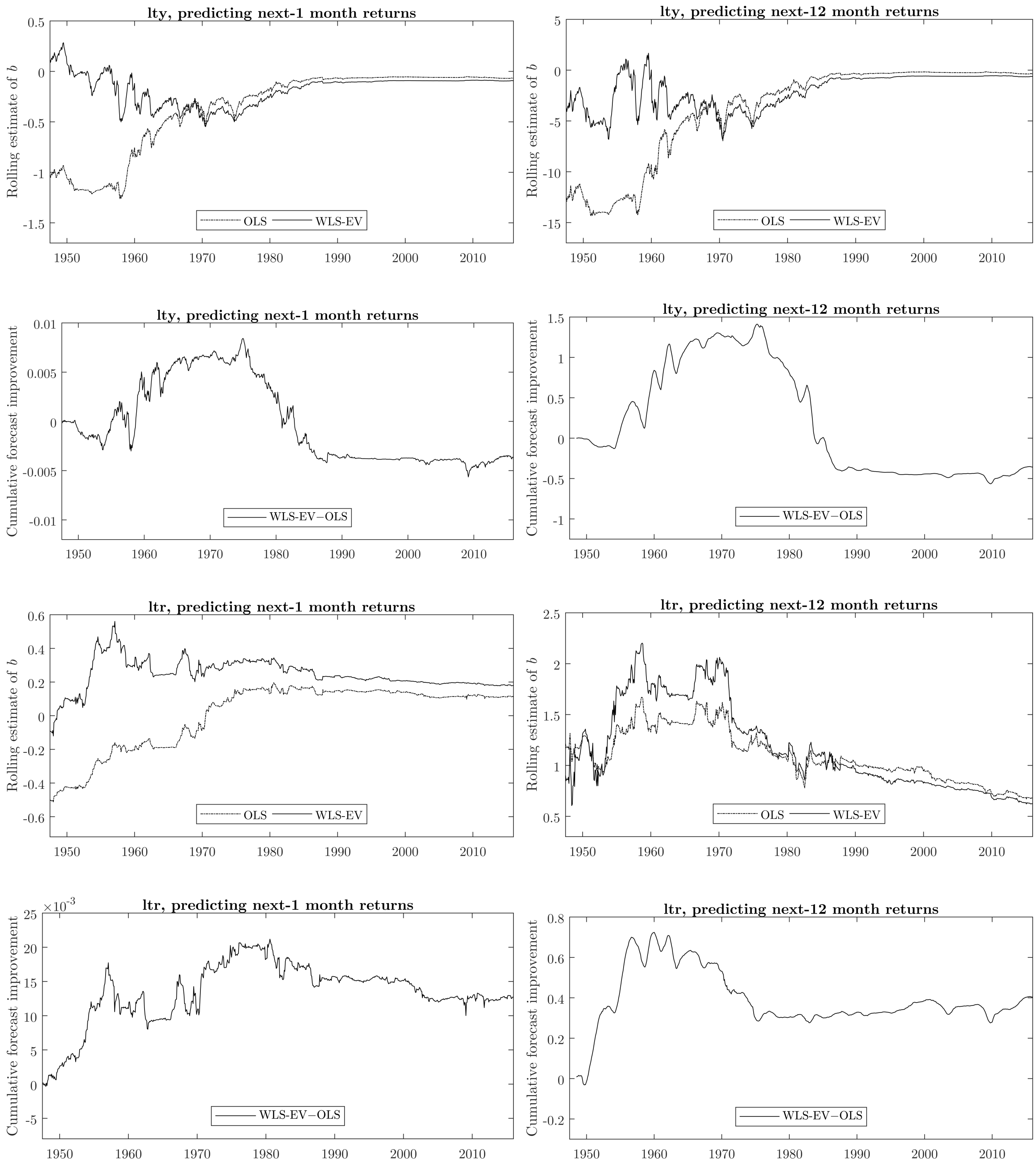
Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



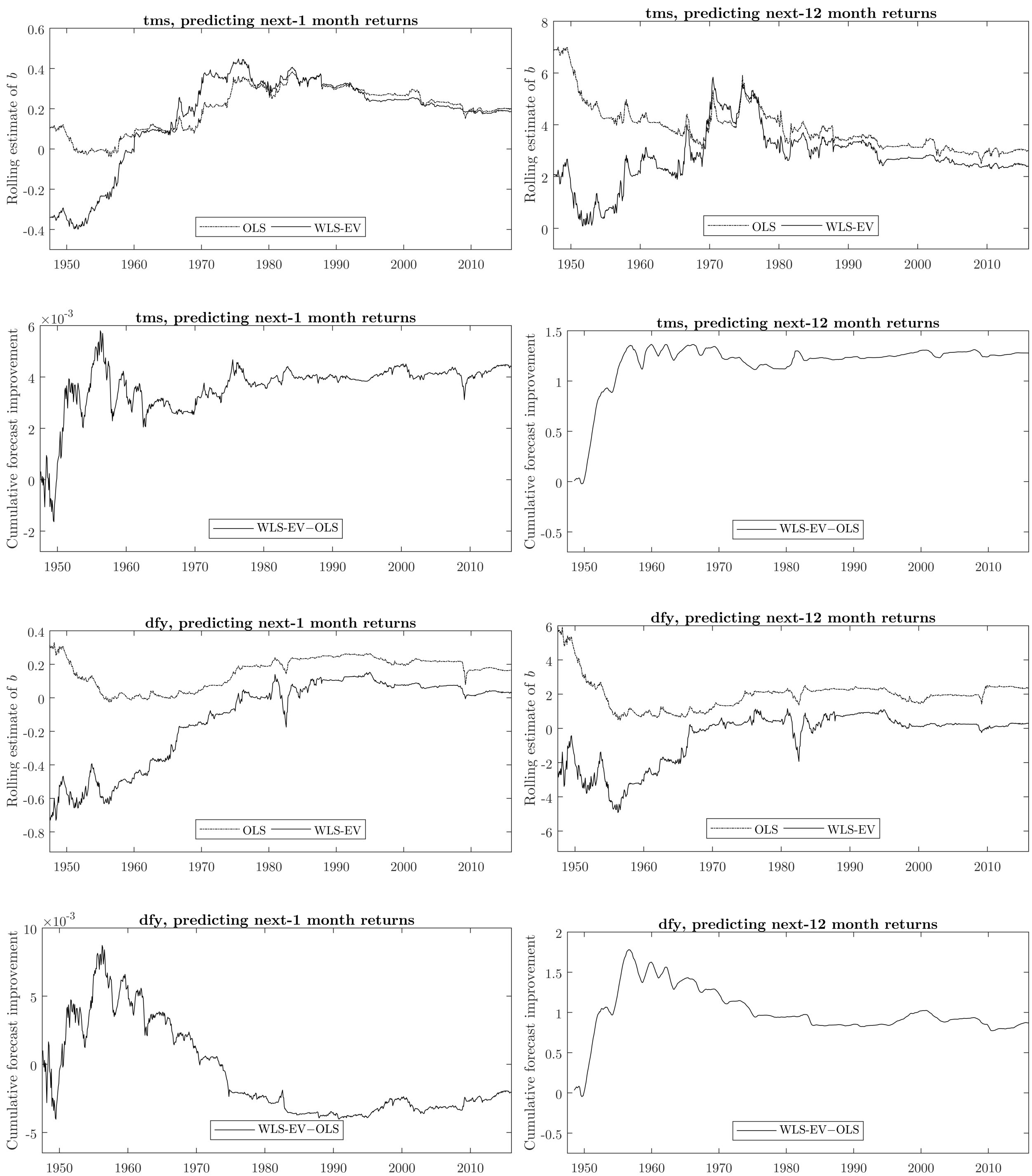
Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



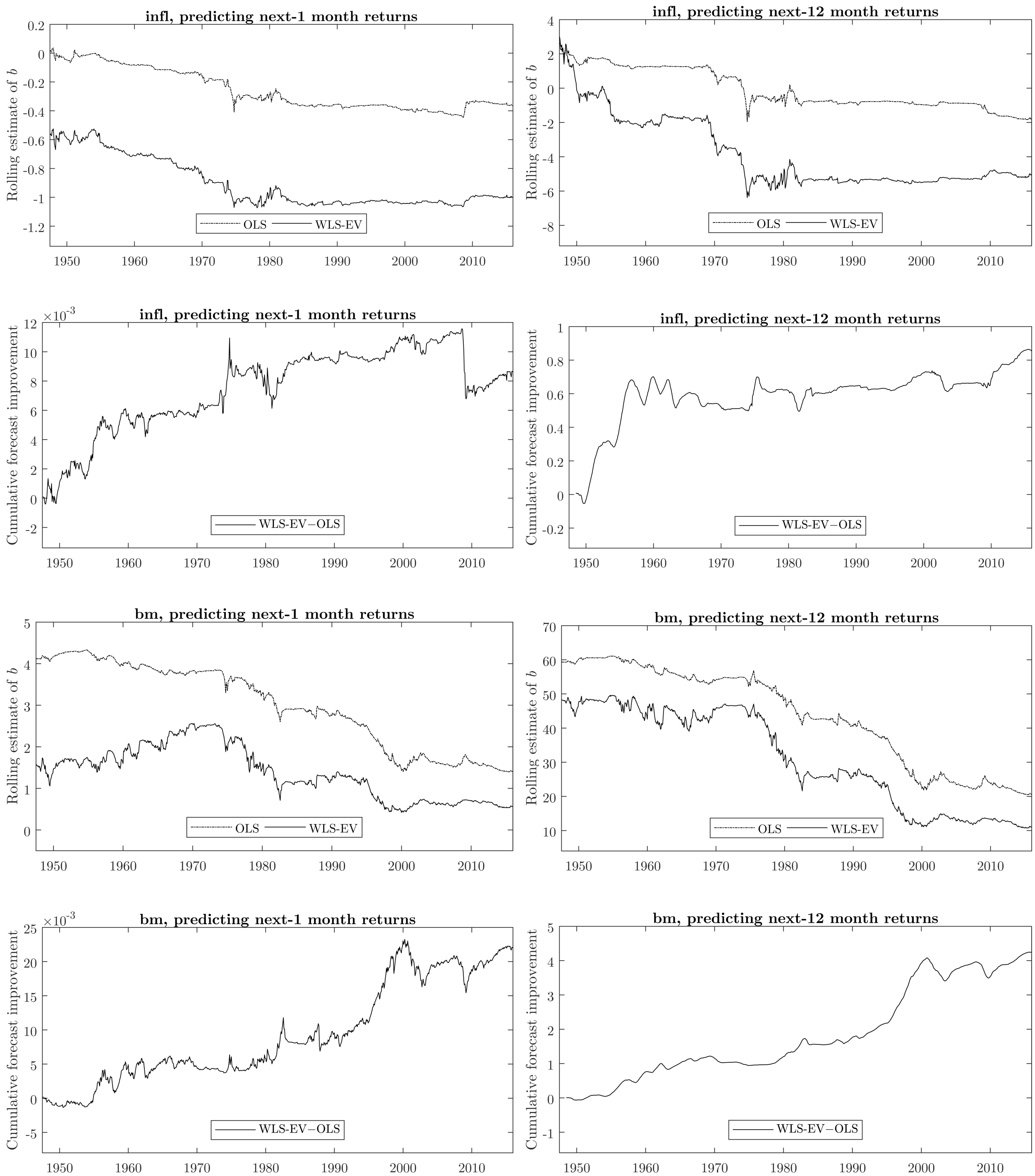
Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



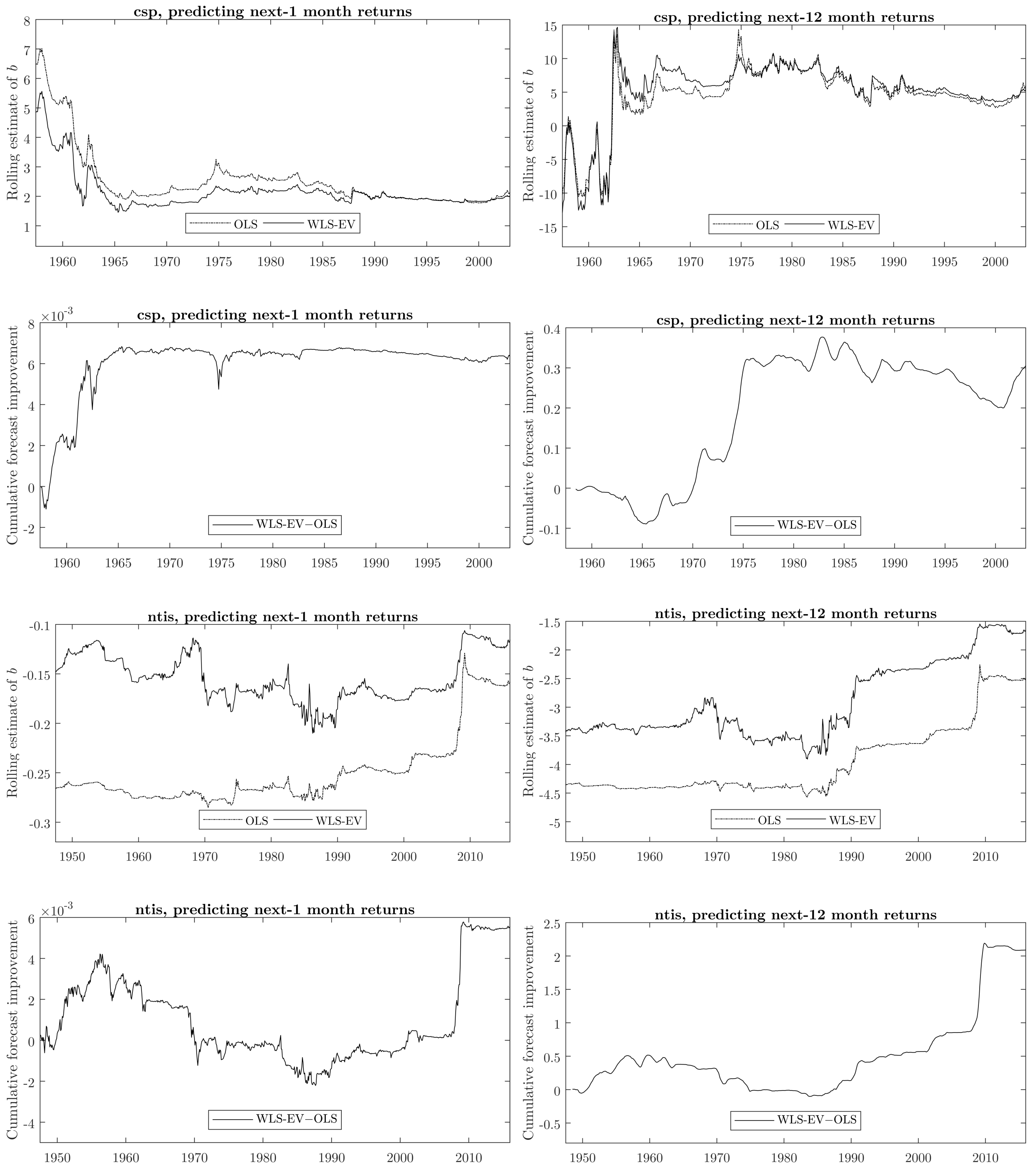
Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



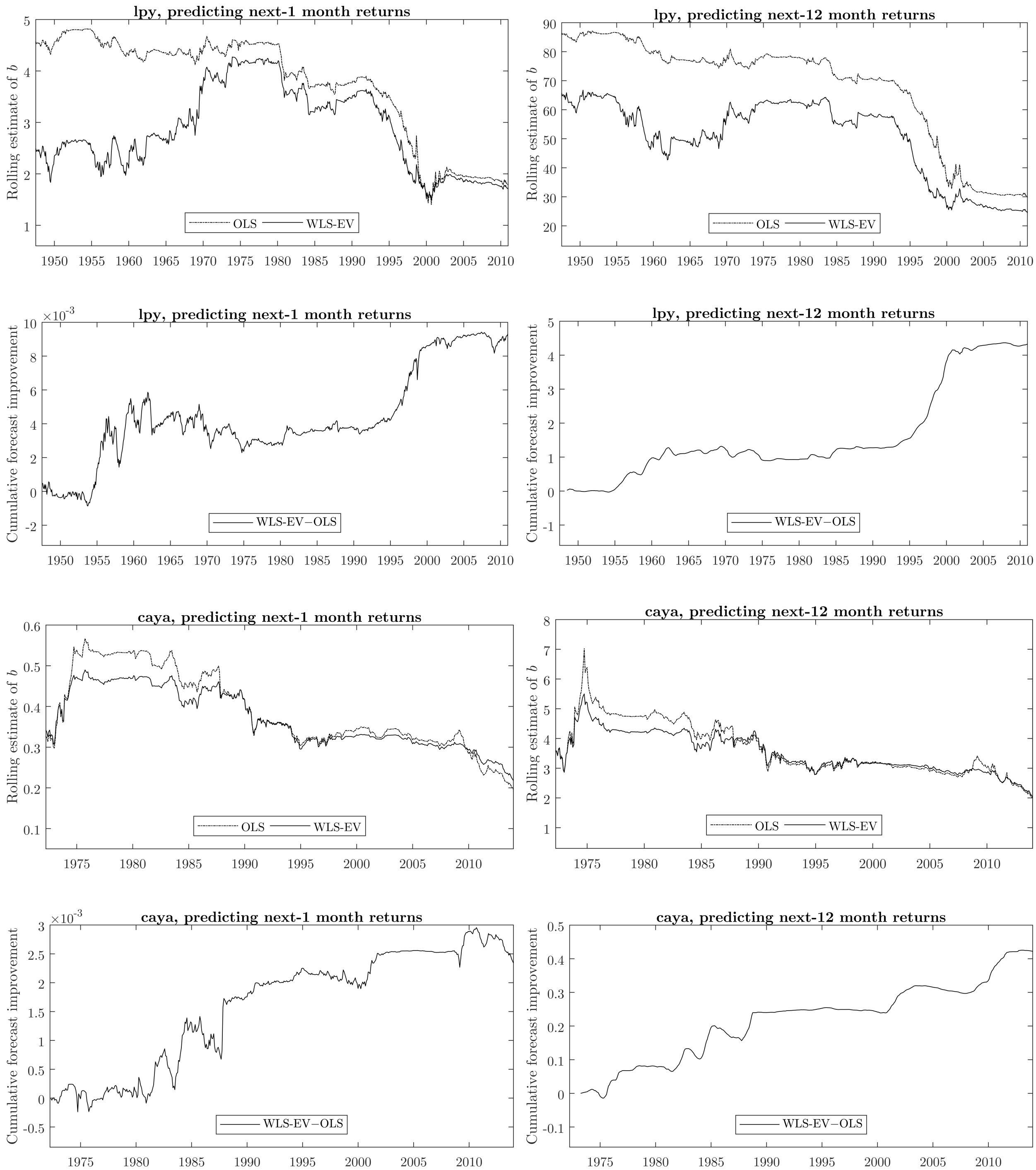
Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



Appendix Figure 1: OOS Coefficient Estimates and Forecast Improvements (cont'd)



Appendix Table 1: Out-of-Sample Certainty Equivalents

This table presents statistics on the out-of-sample predictability afforded by 16 candidate predictors x_t. In Panel A, I predict next-month log dividend-inclusive excess returns of the CRSP value-weighted index. In Panel B, the forecast horizon is one year. The predictors are identical to those in Table 3 of the paper. For each predictor, I compute out-of-sample return forecasts starting 20 years after the sample begins, using both OLS and weighted least squares with ex ante return variance (WLS-EV), detailed in Section 2 of the paper. Given these out-of-sample forecasts, I compute certainty equivalents for an investor with mean-variance preferences who optimizes their portfolio each period using out-of-sample forecasts of both the mean and variance of future returns. For each out-of-sample forecast technique used in Table 4, I compute the change in certainty equivalent ΔCE that results from using the predictor variable instead of a simple average of past returns. In both panels, I use 1,063 monthly observations from 1927-2015.

Panel A: Predicting Next-Month Returns

	OLS	WLS-EV	OLS	WLS-EV	OLS	WLS-EV	OLS	WLS-EV
Predictor:	dp		dy		ep		de	
OOS ΔCE	-0.21%	-0.12%	-1.02%	-0.06%	-0.18%	-0.18%	-0.48%	0.07%
CT ΔCE	0.16%	0.09%	-0.07%	0.20%	-0.08%	-0.14%	-0.06%	0.00%
PTV ΔCE	0.37%	0.09%	0.58%	0.17%	0.09%	0.00%	-0.22%	-0.05%
Predictor:	$\text{RV}_{t-252,t}$		tbl		lty		ltr	
OOS ΔCE	0.09%	-0.11%	1.22%	0.23%	0.57%	-0.78%	0.02%	1.43%
CT ΔCE	0.09%	0.00%	1.71%	1.49%	1.50%	1.33%	0.50%	1.46%
PTV ΔCE	0.14%	0.06%	1.55%	1.36%	1.52%	1.41%	0.21%	0.61%
Predictor:	tms		dfy		infl		bm	
OOS ΔCE	0.55%	0.66%	-0.27%	0.11%	0.70%	1.37%	-2.49%	-0.53%
CT ΔCE	0.58%	0.74%	-0.26%	0.02%	0.71%	1.51%	-1.68%	-0.53%
PTV ΔCE	0.57%	0.48%	-0.18%	0.20%	0.19%	0.20%	-0.24%	-0.18%
Predictor:	csp		ntis		lpy		caya	
OOS ΔCE	-1.30%	0.44%	-0.22%	-0.12%	-0.06%	0.30%	0.66%	0.59%
CT ΔCE	1.40%	1.42%	-0.22%	-0.12%	0.59%	0.60%	-0.21%	-0.22%
PTV ΔCE	1.52%	1.57%	0.09%	0.07%	0.33%	0.35%	0.60%	0.62%

Summary Statistics

	OOS ΔCE			CT ΔCE			PTV ΔCE		
	Mean	#>0	#>OLS	Mean	#>0	%>OLS	Mean	#>0	#>OLS
OLS	-0.15%	7		0.29%	9		0.44%	13	
WLS-EV	0.21%	9	12	0.49%	11	10	0.44%	14	8

Appendix Table 1: Out-of-Sample Certainty Equivalents (cont'd)

Panel B: Predicting Next-Year Returns

	OLS	WLS-EV	OLS	WLS-EV	OLS	WLS-EV	OLS	WLS-EV
Predictor:	dp		dy		ep		de	
OOS ΔCE	-1.57%	-0.22%	-1.80%	-0.26%	-0.30%	-0.26%	-0.32%	-0.24%
CT ΔCE	-0.04%	0.42%	-0.06%	0.41%	-0.08%	-0.18%	-0.14%	0.07%
PTV ΔCE	0.68%	0.53%	0.70%	0.55%	0.04%	-0.03%	-0.32%	-0.24%
Predictor:	$\mathrm{RV}_{t-252,t}$		tbl		lty		ltr	
OOS ΔCE	0.30%	-0.57%	0.42%	-0.32%	0.38%	-0.96%	-0.21%	-0.17%
CT ΔCE	0.30%	0.00%	1.74%	1.78%	1.64%	1.59%	-0.20%	-0.17%
PTV ΔCE	0.30%	-0.32%	2.04%	1.90%	1.72%	1.67%	-0.09%	-0.08%
Predictor:	tms		dfy		infl		bm	
OOS ΔCE	0.70%	1.04%	-0.37%	0.03%	0.25%	0.46%	-4.51%	-1.12%
CT ΔCE	0.79%	1.05%	-0.35%	0.06%	0.25%	0.47%	-1.88%	-1.02%
PTV ΔCE	0.98%	1.03%	-0.21%	0.03%	0.25%	0.31%	-1.12%	-0.68%
Predictor:	csp		ntis		lpy		caya	
OOS ΔCE	0.32%	0.33%	-1.54%	-1.07%	-4.95%	-0.67%	1.27%	1.35%
CT ΔCE	0.32%	0.33%	-1.54%	-1.07%	0.47%	0.81%	1.19%	1.23%
PTV ΔCE	0.32%	0.33%	-0.32%	-0.28%	0.65%	0.58%	1.59%	1.56%

Summary Statistics

	OOS ΔCE			CT ΔCE			PTV ΔCE		
	Mean	#>0	#>OLS	Mean	#>0	#>OLS	Mean	#>0	#>OLS
OLS	-0.75%	7		0.15%	8		0.45%	11	
WLS-EV	-0.17%	5	13	0.36%	11	13	0.43%	10	8

Appendix Table 2: Predicting Returns Using Alternative $\hat{\text{VRP}}_t$

This table presents estimates of return predictability regressions of the form:

$$r_{t+1,t+h} = a + b \cdot \hat{\text{VRP}}_t + \epsilon_{t+h},$$

where $r_{t+1,t+h}$ is the log dividend-inclusive excess return of the CRSP value-weighted index over the h days starting with t+1, annualized and in percent. $\hat{\text{VRP}}_t$ is one of four proxies for the variance risk premium, each expressed as monthly percents squared. Panel A presents the Drechsler and Yaron (2011) estimate of $\hat{\text{VRP}}_t$ discretized into deciles or winsorized below at 0. Panel B presents the same modifications for the Bollerslev, Tauchen, and Zhou (2009) $\hat{\text{VRP}}_t$. For each predictor, I compute point estimates of b using OLS and weighted least squares with ex ante return variance (WLS-EV), detailed in Section 2 of the paper, using both VIX and VIXF $\hat{\sigma}_t^2$. I also compute $\hat{b}$ adjusted for the Stambaugh bias using a simulation procedure. I compute standard errors and p-values for the bias-adjusted coefficients using Hodrick (1992) (Asy) and the heteroskedastic simulation procedure (Sim) described in Section 2. The sample is 6,552 daily observations from 1990-2015. ***, **, and * indicate Sim p-values below 1%, 5%, and 10% levels, respectively.

Panel A: Drechsler and Yaron (2011) Approach

$\hat{\text{VRP}}_d = \text{VIX}_d^2 - \hat{\mathbb{E}}_d(\text{FutRV}_{d+1,d+21}^2)$

Forecast horizon:	*One month ($h=12$)*			*Three months ($h=63$)*		
	OLS	WLS-RV	WLS-VIXF	OLS	WLS-RV	WLS-VIXF
Stambaugh $\hat{b}_{\text{adj}}$	1.590	0.633	0.582	1.156	0.357	0.373
Unadjusted $\hat{b}$	1.654	0.697	0.643	1.206	0.406	0.429
SE (Asy)	(1.003)	(0.832)	(0.796)	(0.859)	(0.764)	(0.735)
p-value (Asy)	11.3%	44.6%	46.5%	17.8%	64.0%	61.2%
SE (Sim)	(1.145)	(0.912)	(0.845)	(0.988)	(0.838)	(0.780)
p-value (Sim)	16.5%	48.8%	48.9%	24.1%	67.0%	63.1%

$\hat{\text{VRP}}_t = \text{Max}[\text{VIX}_t^2 - \hat{\mathbb{E}}_t(\text{FutRV}_{t+1,t+21}^2), 0]$

Forecast horizon:	*One month ($h=12$)*			*Three months ($h=63$)*		
	OLS	WLS-RV	WLS-VIXF	OLS	WLS-RV	WLS-VIXF
Stambaugh $\hat{b}_{\text{adj}}$	0.380	0.253	0.210	0.344	0.215	0.193
Unadjusted $\hat{b}$	0.390	0.263	0.219	0.352	0.222	0.201
SE (Asy)	(0.235)	(0.191)	(0.185)	(0.190)	(0.147)	(0.141)
p-value (Asy)	10.5%	18.6%	25.7%	7.1%	14.3%	17.1%
SE (Sim)	(0.296)	(0.212)	(0.192)	(0.224)	(0.168)	(0.156)
p-value (Sim)	(0.198)	(0.232)	(0.276)	(0.124)	(0.203)	(0.217)

Appendix Table 2: Predicting Returns Using Alternative $\hat{\text{VRP}}_t$ (cont'd)

Panel B: Bollerslev, Tauchen, and Zhou (2009) Approach

$\hat{\text{VRP}}_d = \text{VIX}_d^2 - \text{IndRV}_{d-20,d}^2$

Forecast horizon:	*One month (h = 12)*			*Three months (h = 63)*		
	OLS	WLS-RV	WLS-VIXF	OLS	WLS-RV	WLS-VIXF
Stambaugh $\hat{b}_{\text{adj}}$	1.853*	0.655	0.671	1.544*	0.491	0.579
Unadjusted $\hat{b}$	1.911	0.713	0.724	1.593	0.540	0.624
SE (Asy)	(0.971)	(0.802)	(0.766)	(0.780)	(0.710)	(0.680)
p-value (Asy)	5.6%	41.4%	38.1%	4.8%	48.9%	39.4%
SE (Sim)	(1.081)	(0.875)	(0.805)	(0.902)	(0.785)	(0.729)
p-value (Sim)	(0.087)	(0.454)	(0.404)	(0.087)	(0.531)	(0.425)

$\hat{\text{VRP}}_t = \text{Max}[\text{VIX}_t^2 - \text{IndRV}_{t-20,t}^2, 0]$

Forecast horizon:	*One month (h = 12)*			*Three months (h = 63)*		
	OLS	WLS-RV	WLS-VIXF	OLS	WLS-RV	WLS-VIXF
Stambaugh $\hat{b}_{\text{adj}}$	0.457*	0.249	0.227	0.435**	0.237	0.232
Unadjusted $\hat{b}$	0.467	0.259	0.237	0.443	0.246	0.241
SE (Asy)	(0.248)	(0.188)	(0.182)	(0.174)	(0.139)	(0.134)
p-value (Asy)	6.5%	18.7%	21.2%	1.2%	8.8%	8.4%
SE (Sim)	(0.275)	(0.208)	(0.186)	(0.205)	(0.164)	(0.150)
p-value (Sim)	(0.097)	(0.233)	(0.222)	(0.034)	(0.148)	(0.120)

Appendix Table 3: Country-Specific VRP Regressions, NW Standard Errors

This table presents estimates of return predictability regressions of the form:

$$r_{t+1,t+h} = a + b \cdot \hat{\mathrm{VRP}}_t + \epsilon_{t+h}, \tag{14}$$

where $r_{t+1,t+h}$ is the next-h month excess log return and $\hat{\mathrm{VRP}}_t$ is an estimate of the conditional variance risk premium at t. I estimate (14) for each of seven country-specific equity indices: Netherlands' AEX, Belgium's BEL 20, Canada's CAC 40, Germany's DAX30, London's FTSE 100, Japan's Nikkei 225, and Switzerland's SMI 20. Following Bollerslev et al. (2014), $\hat{\mathrm{VRP}}_t$ is the difference between the index's option-implied volatility index and the rolling sum of past-month squared log daily returns, both expressed as monthly percent variances squared. In Panels A and C, I estimate (14) using OLS. In Panels B and D, I estimate (14) using the WLS-EV procedure described in Section 2 of the paper. In Panels A and B, I use overlapping monthly sampling to predict next-h month returns. In Panels C and D, I use overlapping daily sampling to predict next-$21 \cdot h$ trading day returns, where h is the monthly horizon. The sample period is 2000-2013 for every index except BEL 20, which ends in 2011 because its volatility index was discontinued. Newey and West (1987) t-statistics with lags equal to one and a half times the forecast horizon are in parenthesis. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Monthly Sampling, OLS

		Horizon							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	0.12	0.17	0.12	0.14*	0.16***	0.13***	0.06*	0.03
		(0.63)	(1.18)	(1.38)	(1.72)	(2.83)	(2.78)	(1.79)	(1.10)
	Adj. R^2 (%)	-0.20	0.80	0.42	1.09	2.23	1.48	0.05	-0.39
BEL 20	VRP_t	0.42**	0.32***	0.21***	0.21***	0.20***	0.15***	0.08**	0.03
		(2.40)	(2.62)	(3.05)	(3.28)	(3.75)	(3.20)	(2.45)	(1.06)
	Adj. R^2 (%)	6.53	5.58	3.19	4.12	4.23	2.28	0.27	-0.64
CAC 40	VRP_t	0.25**	0.23**	0.20***	0.19***	0.15***	0.11***	0.04	0.02
		(2.08)	(2.26)	(2.95)	(3.08)	(3.25)	(3.01)	(1.13)	(0.48)
	Adj. R^2 (%)	1.50	2.53	3.16	3.54	2.34	1.13	-0.35	-0.57
DAX 30	VRP_t	0.09	0.17	0.15	0.17	0.13	0.06	0.01	0.02
		(0.51)	(0.97)	(1.16)	(1.51)	(1.46)	(0.73)	(0.31)	(0.57)
	Adj. R^2 (%)	-0.44	0.58	0.63	1.40	0.92	-0.29	-0.62	-0.54
FTSE 100	VRP_t	0.04	0.08	0.15***	0.16***	0.14***	0.10**	0.03	0.00
		(0.36)	(0.98)	(3.59)	(4.08)	(4.07)	(2.39)	(0.78)	(0.13)
	Adj. R^2 (%)	-0.54	-0.18	1.53	2.77	2.12	1.21	-0.48	-0.64
Nikkei 225	VRP_t	-0.04	-0.01	0.04	0.08	0.09	0.08	0.05	0.01
		(0.33)	(0.07)	(0.32)	(0.66)	(0.96)	(0.98)	(0.86)	(0.34)
	Adj. R^2 (%)	-0.57	-0.61	-0.55	-0.27	-0.04	-0.06	-0.40	-0.62
SMI 20	VRP_t	0.03	0.14	0.13	0.20**	0.19***	0.15**	0.10**	0.08***
		(0.18)	(1.17)	(1.41)	(2.24)	(2.87)	(2.38)	(2.45)	(2.86)
	Adj. R^2 (%)	-0.58	0.27	0.42	2.37	2.60	1.78	0.69	0.59

Appendix Table 3: Country-Specific VRP Regressions, NW SEs (cont'd)

Panel B: Monthly Sampling, WLS-EV

		Horizon							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	-0.18	-0.05	-0.04	-0.08	0.04	0.00	-0.02	-0.08
		(0.67)	(0.23)	(0.21)	(0.41)	(0.32)	(0.00)	(0.25)	(0.67)
	Adj. R^2 (%)	-2.97	-1.89	-1.80	-3.91	0.17	-1.36	-2.12	-4.05
BEL 20	VRP_t	0.39*	0.34**	0.23**	0.21**	0.21**	0.17**	0.11	0.07
		(1.92)	(2.17)	(2.03)	(2.04)	(2.30)	(2.17)	(1.55)	(1.05)
	Adj. R^2 (%)	5.98	5.06	2.44	3.27	3.53	1.48	-0.82	-1.97
CAC 40	VRP_t	0.19	0.21	0.18	0.16	0.14	0.13	0.03	0.02
		(1.31)	(1.63)	(1.48)	(1.43)	(1.38)	(1.62)	(0.48)	(0.27)
	Adj. R^2 (%)	1.33	2.43	2.99	3.25	2.09	0.69	-1.11	-1.38
DAX 30	VRP_t	0.02	0.15	0.08	0.12	0.13	0.09	0.05	0.07
		(0.12)	(0.88)	(0.57)	(1.00)	(1.20)	(0.92)	(0.96)	(1.00)
	Adj. R^2 (%)	-0.68	0.49	0.26	1.15	0.82	-0.57	-1.33	-1.17
FTSE 100	VRP_t	-0.04	0.03	0.07	0.08	0.06	0.04	-0.05	-0.06
		(0.20)	(0.20)	(0.55)	(0.59)	(0.50)	(0.35)	(0.50)	(0.61)
	Adj. R^2 (%)	-0.89	-0.60	0.59	1.18	0.49	-0.60	-3.29	-2.85
Nikkei 225	VRP_t	-0.17	-0.09	-0.12	-0.08	-0.02	0.00	0.01	-0.03
		(0.90)	(0.61)	(0.77)	(0.57)	(0.19)	(0.02)	(0.14)	(0.40)
	Adj. R^2 (%)	-0.91	-0.89	-1.76	-1.85	-1.14	-0.73	-0.80	-1.23
SMI 20	VRP_t	-0.07	0.10	0.10	0.13	0.16	0.13	0.07	0.07
		(0.36)	(0.66)	(0.86)	(1.28)	(1.52)	(1.19)	(0.70)	(0.80)
	Adj. R^2 (%)	-1.11	0.02	0.19	1.85	2.24	1.40	0.25	0.29

Appendix Table 3: Country-Specific VRP Regressions, NW SEs (cont'd)

Panel C: Daily Sampling, OLS

Index		Horizon 1	2	3	4	5	6	9	12
AEX	VRP_t	0.18	0.07	0.11	0.13	0.11	0.10**	0.08**	0.04
		(1.16)	(0.66)	(1.44)	(1.62)	(1.64)	(2.01)	(2.04)	(1.14)
	Adj. R^2 (%)	0.60	0.16	0.62	1.09	1.06	0.91	0.76	0.19
BEL 20	VRP_t	0.29	0.23	0.19*	0.20**	0.19**	0.17***	0.11**	0.03
		(1.45)	(1.61)	(1.90)	(2.03)	(2.29)	(2.72)	(2.49)	(0.93)
	Adj. R^2 (%)	1.90	2.38	2.24	3.10	3.32	2.59	1.43	0.13
CAC 40	VRP_t	0.22*	0.11	0.17**	0.16**	0.11**	0.08**	0.06*	0.02
		(1.68)	(1.24)	(2.38)	(2.26)	(2.08)	(2.20)	(1.79)	(0.52)
	Adj. R^2 (%)	0.97	0.52	1.87	2.10	1.22	0.79	0.48	0.03
DAX 30	VRP_t	0.06	0.06	0.12	0.13	0.09	0.05	0.03	0.01
		(0.38)	(0.42)	(1.09)	(1.27)	(1.24)	(0.62)	(0.73)	(0.30)
	Adj. R^2 (%)	0.03	0.11	0.65	0.99	0.66	0.18	0.05	-0.01
FTSE 100	VRP_t	0.22*	0.07	0.15***	0.16***	0.12***	0.11***	0.07*	0.02
		(1.81)	(0.85)	(3.02)	(3.98)	(3.18)	(2.87)	(1.77)	(0.61)
	Adj. R^2 (%)	0.96	0.21	1.47	2.51	1.71	1.60	0.81	0.07
Nikkei 225	VRP_t	0.07	0.03	0.06	0.07	0.05	0.05	0.01	0.01
		(0.60)	(0.44)	(0.67)	(0.97)	(0.67)	(0.81)	(0.37)	(0.42)
	Adj. R^2 (%)	0.08	0.03	0.19	0.43	0.18	0.24	0.01	0.01
SMI 20	VRP_t	0.02	0.01	0.07	0.12*	0.12**	0.11*	0.11**	0.10***
		(0.12)	(0.10)	(0.86)	(1.91)	(2.11)	(1.90)	(2.49)	(2.90)
	Adj. R^2 (%)	-0.02	-0.02	0.20	0.92	1.03	1.02	1.42	1.32

Appendix Table 3: Country-Specific VRP Regressions, NW SEs (cont'd)

Panel D: Daily Sampling, WLS-EV

		Horizon							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	0.03	-0.06	-0.01	-0.02	0.01	-0.01	-0.01	-0.04
		(0.15)	(0.45)	(0.09)	(0.17)	(0.11)	(0.13)	(0.06)	(0.45)
	Adj. R^2 (%)	0.13	-0.55	-0.32	-0.71	-0.16	-0.73	-0.83	-1.22
BEL 20	VRP_t	0.17	0.20	0.17	0.16	0.17	0.16	0.11	0.03
		(0.74)	(1.02)	(1.11)	(1.12)	(1.45)	(1.59)	(1.31)	(0.40)
	Adj. R^2 (%)	0.94	1.34	0.76	1.41	1.75	1.20	-0.21	-1.82
CAC 40	VRP_t	0.08	0.01	0.08	0.05	0.02	0.02	0.00	-0.03
		(0.49)	(0.08)	(0.66)	(0.44)	(0.18)	(0.20)	(0.01)	(0.45)
	Adj. R^2 (%)	0.51	-0.15	1.00	0.72	-0.25	-0.35	-1.03	-1.37
DAX 30	VRP_t	0.00	-0.03	0.00	0.03	0.04	0.02	0.02	0.01
		(0.03)	(0.19)	(0.01)	(0.26)	(0.46)	(0.21)	(0.45)	(0.28)
	Adj. R^2 (%)	-0.12	-0.42	-0.34	0.06	0.06	-0.36	-0.75	-0.58
FTSE 100	VRP_t	0.07	-0.05	0.03	0.03	0.00	-0.01	-0.05	-0.07
		(0.42)	(0.39)	(0.21)	(0.28)	(0.01)	(0.11)	(0.43)	(0.67)
	Adj. R^2 (%)	0.51	-0.60	0.13	0.42	-0.70	-1.23	-2.69	-2.78
Nikkei 225	VRP_t	-0.24	-0.22	-0.26	-0.21	-0.18	-0.15	-0.12	-0.11
		(1.48)	(1.48)	(1.61)	(1.41)	(1.18)	(1.09)	(1.01)	(1.24)
	Adj. R^2 (%)	-2.24	-3.22	-6.90	-7.51	-6.30	-6.12	-5.42	-5.62
SMI 20	VRP_t	-0.15	-0.13	-0.03	-0.03	0.00	-0.02	-0.03	-0.01
		(0.83)	(0.85)	(0.20)	(0.24)	(0.03)	(0.14)	(0.23)	(0.11)
	Adj. R^2 (%)	-0.58	-0.85	-0.58	-1.09	-0.81	-1.41	-2.04	-1.44

Appendix Table 4: Country-Specific VRP Regressions, Simulation SEs

This table presents estimates of return predictability regressions of the form:

$$r_{t+1,t+h} = a + b \cdot \hat{\mathrm{VRP}}_t + \epsilon_{t+h}, \tag{15}$$

where $r_{t+1,t+h}$ is the next-h month excess log return and $\hat{\mathrm{VRP}}_t$ is an estimate of the conditional variance risk premium at t. I estimate (15) for each of seven country-specific equity indices: Netherlands' AEX, Belgium's BEL 20, Canada's CAC 40, Germany's DAX30, London's FTSE 100, Japan's Nikkei 225, and Switzerland's SMI 20. Following Bollerslev et al. (2014), $\hat{\mathrm{VRP}}_t$ is the difference between the index's option-implied volatility index and the rolling sum of past-month squared log daily returns, both expressed as monthly percent variances squared. In Panels A and C, I estimate (14) using OLS. In Panels B and D, I estimate (15) using the WLS-EV procedure described in Section 2 of the paper. In Panels A and B, I use overlapping monthly sampling to predict next-h month returns. In Panels C and D, I use overlapping daily sampling to predict next-$21 \cdot h$ trading day returns, where h is the monthly horizon. The sample period is 2000-2013 for every index except BEL 20, which ends in 2011 because its volatility index was discontinued. I use the Hodrick (1992) mapping to non-overlapping return regressions and compute small-sample t-statistics (in parenthesis) using the heteroskedastic simulation procedure described in Section 3. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Monthly Sampling, OLS

		Horizon (months)							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	0.12	0.17	0.12	0.14	0.17	0.14	0.07	0.04
		(0.33)	(0.62)	(0.55)	(0.79)	(1.12)	(1.04)	(0.71)	(0.46)
	Adj. R^2 (%)	-0.20	0.64	0.29	0.95	2.21	1.63	0.24	-0.31
BEL 20	VRP_t	0.42	0.32	0.22	0.22	0.21*	0.17	0.08	0.03
		(1.41)	(1.48)	(1.26)	(1.55)	(1.82)	(1.62)	(1.02)	(0.53)
	Adj. R^2 (%)	6.53	6.01	3.60	5.07	6.29	4.49	1.12	-0.39
CAC 40	VRP_t	0.25	0.23	0.21	0.19	0.15	0.11	0.04	0.02
		(0.73)	(0.89)	(1.00)	(1.14)	(1.03)	(0.83)	(0.39)	(0.23)
	Adj. R^2 (%)	1.50	2.23	2.47	2.74	1.93	0.94	-0.36	-0.56
DAX 30	VRP_t	0.09	0.18	0.15	0.17	0.14	0.06	0.02	0.03
		(0.26)	(0.75)	(0.80)	(1.10)	(1.02)	(0.53)	(0.19)	(0.42)
	Adj. R^2 (%)	-0.44	0.51	0.51	1.27	0.94	-0.24	-0.59	-0.46
FTSE 100	VRP_t	0.04	0.08	0.15	0.17	0.14	0.11	0.03	0.00
		(0.10)	(0.27)	(0.60)	(0.84)	(0.84)	(0.74)	(0.26)	(0.04)
	Adj. R^2 (%)	-0.54	-0.27	0.87	1.93	1.70	1.06	-0.46	-0.64
Nikkei 225	VRP_t	-0.04	-0.01	0.04	0.08	0.09	0.08	0.02	-0.01
		(0.11)	(0.04)	(0.22)	(0.56)	(0.74)	(0.88)	(0.31)	(0.23)
	Adj. R^2 (%)	-0.57	-0.61	-0.54	-0.18	0.01	0.14	-0.55	-0.60
SMI 20	VRP_t	0.03	0.14	0.13	0.20	0.19	0.16	0.10	0.08
		(0.09)	(0.57)	(0.69)	(1.33)	(1.63)	(1.52)	(1.36)	(1.31)
	Adj. R^2 (%)	-0.58	0.42	0.75	3.93	4.97	4.01	2.49	2.25

Appendix Table 4: Country-Specific VRP Regressions, Sim. SEs (cont'd)

Panel B: Monthly Sampling, WLS-EV

		Horizon (months)							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	-0.18	0.07	0.05	0.04	0.17	0.11	0.04	0.00
		(0.50)	(0.28)	(0.28)	(0.28)	(1.46)	(1.09)	(0.55)	(0.02)
	Adj. R^2 (%)	-2.97	-0.09	-0.29	-0.10	2.07	1.38	-0.12	-1.00
BEL 20	VRP_t	0.39	0.35*	0.22	0.20	0.21*	0.18**	0.10	0.03
		(1.42)	(1.66)	(1.34)	(1.54)	(1.95)	(2.00)	(1.33)	(0.46)
	Adj. R^2 (%)	5.98	5.66	3.22	4.69	6.07	4.25	0.71	-1.21
CAC 40	VRP_t	0.19	0.25	0.18	0.17	0.16	0.13	0.03	0.01
		(0.67)	(1.12)	(1.04)	(1.17)	(1.25)	(1.16)	(0.33)	(0.11)
	Adj. R^2 (%)	1.33	2.15	2.37	2.64	1.88	0.85	-0.51	-0.75
DAX 30	VRP_t	0.02	0.23	0.14	0.15	0.17	0.10	0.06	0.05
		(0.08)	(1.13)	(0.85)	(1.17)	(1.42)	(0.92)	(0.69)	(0.70)
	Adj. R^2 (%)	-0.68	0.38	0.45	1.23	0.89	-0.35	-0.85	-0.53
FTSE 100	VRP_t	-0.04	0.07	0.11	0.11	0.10	0.06	-0.01	-0.04
		(0.11)	(0.28)	(0.52)	(0.68)	(0.74)	(0.56)	(0.12)	(0.67)
	Adj. R^2 (%)	-0.89	-0.38	0.63	1.52	1.39	0.69	-0.93	-1.31
Nikkei 225	VRP_t	-0.17	-0.07	-0.06	0.00	0.04	0.05	-0.01	-0.04
		(0.62)	(0.42)	(0.44)	(0.04)	(0.43)	(0.55)	(0.09)	(0.69)
	Adj. R^2 (%)	-0.91	-0.79	-1.06	-0.62	-0.22	-0.10	-0.76	-0.80
SMI 20	VRP_t	-0.07	0.10	0.09	0.14	0.14	0.10	0.07	0.03
		(0.22)	(0.52)	(0.62)	(1.19)	(1.53)	(1.20)	(1.12)	(0.68)
	Adj. R^2 (%)	-1.11	0.13	0.41	3.34	4.42	3.25	1.88	0.96

Appendix Table 4: Country-Specific VRP Regressions, Sim. SEs (cont'd)

Panel C: Daily Sampling, OLS

		Horizon (months)							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	0.18	0.07	0.11	0.13	0.12	0.10	0.08	0.03
		(0.58)	(0.28)	(0.54)	(0.74)	(0.81)	(0.80)	(0.86)	(0.43)
	Adj. R^2 (%)	0.02	-0.02	0.00	0.03	0.03	0.03	0.03	-0.02
BEL 20	VRP_t	0.29	0.24	0.19	0.21	0.20	0.16	0.10	0.03
		(0.96)	(0.96)	(0.95)	(1.20)	(1.40)	(1.30)	(1.18)	(0.46)
	Adj. R^2 (%)	0.10	0.10	0.08	0.15	0.20	0.15	0.10	-0.02
CAC 40	VRP_t	0.22	0.11	0.17	0.16	0.11	0.09	0.06	0.02
		(0.75)	(0.49)	(0.89)	(0.97)	(0.79)	(0.67)	(0.61)	(0.19)
	Adj. R^2 (%)	0.03	-0.01	0.04	0.05	0.02	0.01	-0.01	-0.03
DAX 30	VRP_t	0.06	0.07	0.12	0.13	0.10	0.05	0.03	0.02
		(0.22)	(0.29)	(0.64)	(0.83)	(0.73)	(0.42)	(0.33)	(0.20)
	Adj. R^2 (%)	-0.02	-0.02	0.00	0.02	0.01	-0.02	-0.02	-0.03
FTSE 100	VRP_t	0.22	0.06	0.15	0.17	0.13	0.11	0.07	0.02
		(0.69)	(0.26)	(0.72)	(0.99)	(0.88)	(0.93)	(0.77)	(0.26)
	Adj. R^2 (%)	0.04	-0.02	0.03	0.07	0.04	0.04	0.01	-0.03
Nikkei 225	VRP_t	0.07	0.03	0.06	0.07	0.04	0.04	0.00	-0.01
		(0.25)	(0.17)	(0.39)	(0.64)	(0.48)	(0.53)	(0.05)	(0.15)
	Adj. R^2 (%)	-0.02	-0.03	-0.02	0.00	-0.02	-0.02	-0.03	-0.03
SMI 20	VRP_t	0.02	0.01	0.07	0.12	0.12	0.11	0.11	0.10
		(0.08)	(0.06)	(0.39)	(0.85)	(1.03)	(1.11)	(1.46)	(1.45)
	Adj. R^2 (%)	-0.03	-0.03	-0.02	0.02	0.03	0.04	0.06	0.05

Appendix Table 4: Country-Specific VRP Regressions, Sim. SEs (cont'd)

Panel D: Daily Sampling, WLS-EV

		Horizon (months)							
Index		1	2	3	4	5	6	9	12
AEX	VRP_t	0.06	-0.05	-0.01	0.01	0.03	0.03	0.00	-0.02
		(0.23)	(0.23)	(0.07)	(0.06)	(0.26)	(0.28)	(0.04)	(0.31)
	Adj. R^2 (%)	-0.01	-0.05	-0.04	-0.03	-0.01	-0.01	-0.03	-0.06
BEL 20	VRP_t	0.20	0.18	0.13	0.11	0.13	0.11	0.07	-0.01
		(0.82)	(0.92)	(0.80)	(0.78)	(1.16)	(1.17)	(1.00)	(0.24)
	Adj. R^2 (%)	0.05	0.05	0.03	0.07	0.14	0.10	0.05	-0.13
CAC 40	VRP_t	0.10	0.05	0.10	0.08	0.05	0.04	0.02	-0.02
		(0.43)	(0.26)	(0.61)	(0.60)	(0.39)	(0.40)	(0.33)	(0.30)
	Adj. R^2 (%)	0.01	-0.01	0.03	0.03	0.00	0.00	-0.01	-0.04
DAX 30	VRP_t	0.07	0.07	0.09	0.11	0.09	0.05	0.06	0.03
		(0.27)	(0.34)	(0.59)	(0.81)	(0.81)	(0.49)	(0.78)	(0.49)
	Adj. R^2 (%)	-0.03	-0.02	0.00	0.02	0.01	-0.02	-0.03	-0.03
FTSE 100	VRP_t	0.06	-0.07	0.02	0.03	0.01	0.01	0.00	-0.02
		(0.25)	(0.38)	(0.16)	(0.24)	(0.14)	(0.12)	(0.02)	(0.44)
	Adj. R^2 (%)	0.00	-0.06	-0.01	0.00	-0.01	-0.02	-0.03	-0.05
Nikkei 225	VRP_t	-0.24	-0.15	-0.12	-0.06	-0.04	-0.03	-0.04	-0.03
		(1.18)	(1.03)	(1.10)	(0.70)	(0.51)	(0.43)	(0.96)	(0.87)
	Adj. R^2 (%)	-0.19	-0.13	-0.16	-0.11	-0.07	-0.06	-0.06	-0.05
SMI 20	VRP_t	-0.16	-0.10	0.03	0.03	0.04	0.03	0.06	0.03
		(0.69)	(0.56)	(0.22)	(0.27)	(0.43)	(0.35)	(0.93)	(0.51)
	Adj. R^2 (%)	-0.07	-0.05	-0.02	-0.01	0.00	0.00	0.03	0.00

Appendix Table 5: Methods for Addressing Overlapping Observations

This table compares two methods for addressing overlapping observations in four different settings from the paper. The first method, which I use throughout the main paper, is the Hodrick (1992) 'de-overlapping' method (Hod.) whereby I first estimate a non-overlapping regression of returns r_t on a rolling sum of past values of the predictor $\overline{x}_t \equiv \sum_{s=0}^{h} x_{t-s}$, and then scale the resulting coefficients by $\frac{\mathrm{Var}(\overline{x}_t)}{\mathrm{Var}(x_t)}$. The alternative method is to estimate overlapping regressions (Over.) of $r_{t+1,t+h}$ on x_t and adjusting the standard errors using Newey and West (1987) with $1.5h$ lags. For the Hod. approach, I apply WLS-EV to the non-overlapping regressions using estimates of next-period return variance. For the Over. approach, I apply WLS-EV using estimates of next-h period return variance. The four predictors are the dividend-to-price ratio (dp), the term spread (tms), the consumption-to-wealth ratio (cay), and the Bollerslev, Tauchen, and Zhou (2009) variance risk premia proxy (VRP). For each predictor and estimator, I compute point estimates, asymptotic standard errors and p-values, and simulated standard errors and p-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Predictor:	dp				tms			
Sample period:	1927–2015				1927–2015			
Sample interval:	Monthly				Monthly			
Forecast horizon:	12 months				12 months			
	OLS		WLS-EV		OLS		WLS-EV	
	Hod.	Over.	Hod.	Over.	Hod.	Over.	Hod.	Over.
Stambaugh $\hat{b}_{\text{adj}}$	3.167	3.316	3.192	4.526	3.025**	2.899**	2.406**	2.401**
Unadjusted $\hat{b}$	7.687	7.783	7.712	8.993	3.015	2.891	2.397	2.393
SE (Asy)	(5.370)	(4.613)	(4.005)	(3.913)	(1.357)	(1.108)	(1.172)	(1.081)
p-value (Asy)	55.5%	53.7%	42.6%	24.7%	2.6%	3.3%	4.0%	2.6%
SE (Sim)	(5.081)	(5.059)	(3.419)	(3.788)	(1.381)	(1.369)	(1.030)	(1.188)
p-value (Sim)	53.3%	51.5%	35.1%	23.3%	2.9%	3.4%	1.9%	4.3%

Predictor:	cay				VRP			
Sample period:	1953–2013				1990–2015			
Sample interval:	Monthly				Daily			
Forecast horizon:	12 months				63 days			
	OLS		WLS-EV		OLS		WLS-EV	
	Hod.	Over.	Hod.	Over.	Hod.	Over.	Hod.	Over.
Stambaugh $\hat{b}_{\text{adj}}$	1.898*	2.073**	2.046**	2.431**	0.396**	0.393**	0.170	0.302*
Unadjusted $\hat{b}$	1.958	2.140	2.106	2.498	0.404	0.400	0.177	0.309
SE (Asy)	(1.144)	(1.022)	(1.013)	(0.970)	(0.136)	(0.095)	(0.118)	(0.114)
p-value (Asy)	9.7%	7.0%	4.3%	1.2%	0.4%	0.4%	15.0%	0.8%
SE (Sim)	(1.036)	(1.052)	(0.866)	(0.992)	(0.157)	(0.157)	(0.131)	(0.183)
p-value (Sim)	6.6%	4.9%	1.8%	1.4%	1.2%	1.2%	19.4%	10.0%