THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 16-25

Jiahan Li
University of Notre Dame, USA

Ilias Tsiakas
University of Guelph, Canada
The Rimini Centre for Economic Analysis, Italy

EQUITY PREMIUM PREDICTION: THE ROLE OF ECONOMIC AND STATISTICAL CONSTRAINTS

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 – 47921 Rimini (RN) – Italy
www.rcfea.org – secretary@rcfea.org

Equity Premium Prediction: The Role of Economic and Statistical Constraints[1]

Jiahan Li
University of Notre Dame
jiahan.li0@gmail.com

Ilias Tsiakas
University of Guelph
itsiakas@uoguelph.ca

July 2016

Abstract

This paper shows that the equity premium is predictable out of sample when we use a predictive regression that conditions on a large set of economic fundamentals, subject to: (i) economic constraints on the sign of coefficients and return forecasts, and (ii) statistical constraints imposed by shrinkage estimation. Equity premium predictability delivers a certainty equivalent return of about 2.7% per year over the benchmark for a mean-variance investor. Our predictive framework outperforms a large group of competing models that also condition on economic fundamentals as well as models that condition on technical indicators.

Keywords: Equity Premium; Out-of-Sample Prediction; Economic Fundamentals; Technical Indicators; Shrinkage Estimation.

JEL Classification: G11; G14; G17.

[1]This paper is forthcoming in the *Journal of Financial Markets*. The authors are grateful for useful comments to Amit Goyal (editor), an anonymous referee, Zhi Da, Nikola Gradojevic, Davide Pettenuzzo, Jun Tu, Guofu Zhou and seminar participants at McMaster University, Brock University, the 2014 RCEF Conference in Rimini, the 2015 CEA Conference in Toronto and the 2016 Financial Econometrics and Empirical Asset Pricing Conference in Lancaster. *Corresponding author*: Ilias Tsiakas, Department of Economics and Finance, University of Guelph, Guelph, Ontario N1G 2W1, Canada. Tel: 519-824-4120 ext. 53054. Fax: 519-763-8497. Email: itsiakas@uoguelph.ca.

1 Introduction

This paper shows that the equity premium is predictable out of sample when we condition on a large set of economic fundamentals. This is an important result in empirical asset pricing for the following reasons. From a conceptual point of view, return predictability is crucial for rationalizing the observed variation in stock prices (Cochrane, 2008, 2011).[2] From a practical point of view, equity premium predictability has important implications for financial decisions involving capital budgeting and asset allocation. From an empirical point of view, equity premium predictability contributes to a large literature in asset pricing, which tends to be more successful when focussing on in-sample analysis, long predictive horizons or latent variable models. Our paper builds on this literature by establishing that the equity premium is predictable out of sample for short predictive horizons when we directly condition on economic fundamentals.[3]

The key to establishing equity premium predictability is implementing a predictive framework based on three aspects: (i) using a single predictive regression that conditions on a large number of predictors (a "kitchen-sink" regression); (ii) imposing economic constraints on the sign of coefficients and return forecasts as in Campbell and Thompson (2008); and (iii) using a shrinkage estimator designed to improve performance by reducing the effect of less informative predictors in out-of-sample forecasting. Our empirical analysis uses monthly returns and monthly economic fundamentals for a sample period ranging from January 1927 to December 2014. The analysis is performed purely out of sample in order to inform real-time investment decisions. We also assess performance around business cycles by dividing the full sample into expansions and recessions dated by the NBER.

Our main empirical finding is that equity premium forecasts based on our predictive framework consistently outperform the historical mean benchmark, especially during recessions. The out-of-sample R^2 of the monthly equity premium forecasts is about 1.7% and is statistically significant. This translates into a certainty equivalent return of about 2.7% per year over and above the benchmark for a mean-variance investor. For recessions, when it matters the most, equity premium predictability is stronger as the certainty equivalent return rises to over 10% above the benchmark. More importantly, our predictive framework outperforms a large group of competing models that use different approaches to condition on economic fundamentals. It also outperforms models that condition on technical indicators.

[2]In the context of the Campbell and Shiller (1988) present value relation, changes in stock prices are either due to news about future returns or news about future dividends. Hence the relative predictability of returns and dividends forms the rational paradigm to interpret asset pricing variation.

[3]For recent studies that discover predictability using latent factors, see Ludvigson and Ng (2007), van Binsbergen and Koijen (2010), Kelly and Pruitt (2013), and Neely, Rapach, Tu and Zhou (2014). For a review of the large literature on the predictability of returns, see Koijen and Van Nieuwerburgh (2011).

Therefore, the highlight of our empirical findings is that using our framework to condition on economic fundamentals can be effective in establishing predictability of the equity premium.

The first aspect of our approach is to specify a kitchen-sink regression that provides a direct way of pooling information by constructing a "super" model, which nests each of the models that condition on a single predictor. This approach is distinct from but comparable to combined forecasts, which are designed to pool forecasts rather than pool information. Note that commonly used forecast combinations involve two stages: first, estimate several predictive regressions each conditioning on one predictor; and second, combine the individual forecasts into one forecast combination (see, e.g., Rapach, Strauss and Zhou, 2010). As the two-stage process introduces an efficiency loss and ignores the correlations between the predictors, it is often argued that pooling information is optimal relative to pooling forecasts (e.g., Timmermann, 2006). This is an argument that supports the kitchen-sink regression relative to combined forecasts.

The second aspect of our approach is to impose economic constraints on the sign of coefficients and return forecasts. These constraints impose economic theory on the predictive regressions and, as a result, almost universally improve performance. Campbell and Thompson (2008) show that the economic constraints improve the performance of predictive regressions that condition on a single predictor. In this paper, we show that the economic constraints perform even better when applied to a kitchen-sink regression that conditions on a large number of predictors as long as the regression is estimated with shrinkage methods.

The third aspect of our approach is to implement shrinkage estimation of the kitchen-sink regression. We find that the empirical performance of shrinkage estimation is far superior to that of ordinary least squares (OLS). Indeed, the kitchen-sink regression transforms from being the worst model when estimated with OLS to the being the best model when estimated with a shrinkage estimator. We use three prominent shrinkage estimators (ridge regression, lasso and the elastic net), which deliver similar results. Shrinkage estimation produces biased parameter estimates by shrinking all estimates towards zero, which is the value implied by the benchmark historical mean model. This is done in a way that improves the bias-variance tradeoff in estimation and reduces the mean squared error of the forecasts thus leading to a higher degree of predictive accuracy. In conclusion, our empirical analysis shows that the combination of the three aspects above in our predictive framework leads to superior out-of-sample prediction of the equity premium.

It is important to note that our predictive framework significantly outperforms all other models in recessions. This is a critical finding because the predictive information of economic fundamentals is more valuable to an investor during recessions. To begin with, this is true because during recessions the equity premium is on average negative (-6.5% annually in our sample) and its volatility is high (29% annually in our sample). Therefore, simply using the historical mean forecast (i.e., the benchmark) is a poor way of forecasting the equity premium in recessions.

More importantly, however, the state of the economy changes the way investors process information. In a recent paper, Kacperczyk, Van Nieuwerburgh and Veldkamp (2016) show that in recessions fund managers care more about aggregate shocks, whereas in expansions they care more about idiosyncratic shocks. This is because aggregate risk is significantly higher in recessions than in expansions. In contrast, idiosyncratic risk is essentially the same in expansions and recessions. For these reasons, the information in economic fundamentals for predicting aggregate stock returns matters the most in recessions.

This paper is especially related to two recent studies. First, Dangl and Halling (2012) estimate predictive regressions for the equity premium, which explicitly allow for time-variation in the regression coefficients. Then, they use Bayesian Model Averaging to form combined forecasts for a large set of models that condition on economic fundamentals. Therefore, they provide a Bayesian approach to pooling forecasts for the equity premium, and find that this approach performs well out of sample.[4]

Second, Pettenuzzo, Timmermann and Valkanov (2014) use economic constraints in the context of a Bayesian approach for computing the predictive density of the equity premium. The constraints require the equity premium to be positive and the Sharpe ratio to lie between zero and one. Our predictive framework differs from both Dangl and Halling (2012) and Pettenuzzo, Timmermann and Valkanov (2014) primarily for two reasons: we follow a distinct approach based on shrinkage estimation rather than Bayesian analysis; and our kitchen-sink approach directly conditions on a large number of observed economic fundamentals.

The remainder of the paper is organized as follows. In the next section we describe the data on economic fundamentals. Section 3 discusses the predictive framework and the empirical results. In Section 4, we present the dynamic mean-variance strategy for evaluating the portfolio performance of equity premium predictability. Section 5 evaluates the predictive content of technical indicators. Finally, Section 6 concludes.

2 Data on Economic Fundamentals

We use a set of monthly economic fundamentals for predicting the monthly equity premium for the period of January 1927 to December 2014. All data are taken from Amit Goyal's website. These are the same data used in Welch and Goyal (2008), Campbell and Thompson

[4]See also Della Corte, Sarno and Tsiakas (2009) for similar work on exchange rate prediction using Bayesian Model Averaging.

(2008), Dangl and Halling (2012), Neely, Rapach, Tu and Zhou (2014), and Pettenuzzo, Timmermann and Valkanov (2014) extended to 2014.

The equity premium is the continuously compounded return on the S&P 500 index including dividends obtained from CRSP minus the treasury bill rate (defined below). For the full sample, the equity premium in annualized terms exhibits a mean return of 7.8%, volatility of 19% and a Sharpe ratio of 0.41. For NBER-dated expansions, the annualized Sharpe ratio of the equity premium rises to 0.73 and for recessions it falls to -0.22. Recessions correspond to about 20% of the full sample and are known ex post. The full set of descriptive statistics is reported in Table C1 of the Internet Appendix.

There is a long list of available economic fundamentals. Our analysis follows Welch and Goyal (2008), Campbell and Thompson (2008), Dangl and Halling (2012), Neely, Rapach, Tu and Zhou (2014), and Pettenuzzo, Timmermann and Valkanov (2014) in using a set of 11 monthly predictors, which are primarily based on stock characteristics and interest rates:[5]

1. *Dividend Yield* (*dy*) is the difference between the log of dividends and the log of lagged prices.

2. *Earnings-Price Ratio* (*epr*) is the difference between the log of earnings and the log of prices.

3. *Book-to-Market Ratio* (*bm*) is the ratio of book value to market value for the Dow Jones Industrial Average.

4. *Net Equity Expansion* (*ntis*) is the ratio of twelve-month moving sums of net issues by NYSE-listed stocks divided by the total market capitalization of NYSE stocks.

5. *Stock Variance* (*svar*) is the sum of squared daily returns on the S&P 500.

6. *Treasury Bill Rate* (*tbl*) is the 3-month *tbl*.

7. *Term Spread* (*tms*) is the difference between the long-term yield on government bonds and the treasury bill rate.

8. *Long-Term Rate of Return* (*ltr*) for government bonds.

9. *Default Yield Spread* (*dfy*) is the difference between BAA- and AAA-rated corporate bond yields.

[5]Note that the Welch and Goyal (2008) data contains 14 monthly predictors. However, we have to exclude three predictors in order to avoid perfect multicollinearity: the dividend-price ratio, which is almost perfectly correlated with the dividend yield; the payout ratio, which is a combination of the dividend-price ratio and the earnings-price ratio; and the long-term yield, which is a combination of the treasury bill rate and the term spread.

10. *Default Return Spread* (*dfr*) is the difference between the return on long-term corporate bonds and the return on long-term government bonds.

11. *Inflation* (*infl*). Note that since inflation information is released only in the following month, we lag inflation by an additional month in the predictive regressions.

The cross-correlations between the predictors are reported in Table C2 of the Internet Appendix. The values of the cross-correlations range from -0.461 to 0.831 with an average value of 0.047.

3 A Model for Predicting the Equity Premium

3.1 The Kitchen-Sink Regression

The model we use for predicting the equity premium is based on a kitchen-sink (KS) predictive regression. The KS regression conditions on a large set of predictive variables and has the following linear structure:

$$r^e_{t+1} = \alpha + \sum_{j=1}^{N} \beta_j x_{j,t} + \varepsilon_{t+1}, \tag{1}$$

where $r^e_{t+1} = r_{t+1} - r_f$ is the equity premium at time $t+1$, r_{t+1} is the total return on the S&P 500 index at time $t+1$, r_f is the treasury bill rate, $x_{j,t}$ is the $j \leq N$ predictor at time t, α and $\beta = \{\beta_j\}$ are constant parameters to be estimated, and ε_{t+1} is a normal error term. The KS regression is an attractive framework because by design it captures all available information in a single regression. This provides a direct way of pooling information into a single forecast. The key to making sure that the KS regression delivers reliable forecasts is to impose statistical and economic constraints on the slope coefficients and the forecasts. We turn to this next.

3.1.1 Statistical Constraints due to Shrinkage Estimation

The main feature of our approach is that we estimate the KS regression with a shrinkage estimator. This is motivated by the fact that the out-of-sample performance of the KS regression estimated with OLS is very poor relative to the historical mean benchmark (see Welch and Goyal, 2008). Our empirical analysis will show that the KS regression will transform from being the worst performing model when estimated with OLS to the best performing model when estimated with a shrinkage estimator.

We implement shrinkage estimation because it is designed to shrink the regression coefficients towards zero (i.e., the value implied by the historical mean benchmark) in a way that directly minimizes the out-of-sample mean squared error (MSE). In contrast to the OLS estimator that is unbiased, a shrinkage estimator is biased but may have lower variance and lower MSE than OLS. In other words, a shrinkage estimator may be more efficient than OLS because it improves the variance-bias tradeoff thus leading to more accurate forecasts, which is critically important in equity premium prediction.[6]

To better clarify the role of shrinkage consider the following example. Suppose that according to the true model, which is unknown to us, economic fundamentals are somehow related to the future equity premium. Then, by setting the regression coefficients to zero, the benchmark model produces biased estimates with zero variance. In contrast, OLS estimation of the KS regression produces unbiased estimates with potentially high variance. Then, a shrinkage estimator of the KS regression shrinks the regression coefficients towards zero and hence can be thought of as a combination of the two cases above leading to a biased estimator with a better variance-bias tradeoff.[7]

We shrink the regression coefficients by estimating the KS regression with the elastic-net estimator of Zou and Hastie (2005), which solves the following system:

$$\begin{aligned} & \min_{\beta} \tfrac{1}{2} \sum_{t=1}^{T-1} \left(r_{t+1}^{e} - \alpha - \sum_{j=1}^{N} \beta_j x_{j,t} \right)^2 \\ \text{s.t.} \quad & \sum_{j=1}^{N} |\beta_j| < s_1 \\ \text{and} \quad & \sum_{j=1}^{N} \beta_j^2 < s_2, \end{aligned} \tag{2}$$

where s_1 and s_2 are positive constants, which are estimated in a way that minimizes the MSE of the forecasts. The estimation algorithm for the elastic net is described in detail in the Internet Appendix.

The elastic net (henceforth e-net) is a general estimator that encompasses two well-known special cases. When $s_1 = \infty$, i.e., the first constraint is unbounded, then Eq. (2) reduces to the ridge regression (Hoerl and Kennard, 1970). When $s_2 = \infty$, i.e., the second constraint is unbounded, then Eq. (2) reduces to the lasso regression (Tibshirani, 1996).[8] Both ridge

[6] For applications of shrinkage methods in finance, see Jorion (1985, 1986), Jagannathan and Ma (2003), Ledoit and Wolf (2003), DeMiguel, Garlappi, Nogales and Uppal (2009), Rapach, Strauss and Zhou (2013), Li (2015), and Li, Tsiakas and Wang (2015).

[7] Tu and Zhou (2011), page 205, provide an economic motivation for shrinkage estimation: "... a concave utility investor will prefer a suitable average of good and bad performances to either a good or a bad performance randomly, similar to the diversification over two assets."

[8] Lasso stands for "least absolute shrinkage and selection operator" since it shrinks the absolute value of

regression and lasso regression shrink regression coefficients towards zero but with different shrinkage intensities. Ridge regression shrinks more the large regression coefficients, whose value may be large due to multicollinearity, but keeps all predictors in the model. In contrast, lasso regression shrinks more the small regression coefficients, but sets some coefficients to exactly zero, thus producing a parsimonious model. Then, the elastic net is an estimator that combines the benefits from both ridge regression and lasso regression.[9]

Our empirical analysis reports results on the elastic net and the two special cases of ridge and lasso regression. As all three models implement shrinkage estimation, a comparison of their performance will allow us to determine whether it is shrinkage that drives the results or the particular type of shrinkage that implements one or both statistical constraints.

3.1.2 Economic Constraints

Following Campbell and Thompson (2008), we also impose two constraints motivated by economic theory. First, we constrain the equity premium forecast to be positive in every time period. We do so by replacing negative forecasts with zero. Campbell and Thompson (2008) argue that a reasonable investor would not have used a model to forecast a negative equity premium. The positive forecast constraint is also implemented more recently by Pettenuzzo, Timmermann and Valkanov (2014), who motivate this constraint by arguing that risk-averse investors would not hold stocks if their expected excess return was negative.

Second, we constrain the sign of the slope coefficients to be consistent with economic theory. Campbell and Thompson (2008) explain that "[a] regression estimated over a short sample period can easily generate perverse results, such as a negative coefficient when theory suggests that the coefficient should be positive... In practice, an investor would not use a perverse coefficient but would likely conclude that the coefficient is zero, in effect imposing prior knowledge on the output of the regression."[10] We implement this constraint by setting a value of zero for a coefficient that does not have the theoretically motivated sign of Campbell and Thompson (2008). To be more specific, the slope constraint is positive for all predictors except for *ntis*, *tbl* and *infl*. As we report later, these economic constraints considerably improve the performance of the models.[11]

---

regression coefficients and performs variable selection.

[9] In more technical terms, the first constraint (i.e., the L1 norm) introduces sparsity and promotes variable selection and model interpretation. In contrast, the second constraint (i.e., the L2 norm) is equivalent to using a shrinkage estimator of the covariance matrix in minimizing least squares. This addresses the multicollinearity problem among correlated predictors. For a more detailed discussion, see the Internet Appendix.

[10] See Campbell and Thompson (2008), page 1516.

[11] Note that for models estimated with shrinkage, the economic constraints are imposed in addition to the statistical constraints. Hence a coefficient may be zero because, for example, the lasso sets it to zero or

3.2 Other Predictive Regressions

3.2.1 The benchmark

The benchmark against which we compare the KS regression is the historical mean for the equity premium. This is the prevalent benchmark in the literature and corresponds to the case of $\beta_j = 0$ for all j. In other words, the historical mean benchmark reflects the view that the expected equity premium is constant and hence it is not predictable when conditioning on economic fundamentals.

3.2.2 Individual Predictors

The majority of the literature uses predictive regressions that typically condition on a single predictor, such as the dividend yield. Thus we report results from 11 regressions each conditioning on one predictor. This approach effectively ignores all information except that captured by the selected predictor. It is worth noting that this approach has not been empirically successful as Welch and Goyal (2008) demonstrate that most of the individual predictors fail to consistently forecast the equity premium out of sample.

3.2.3 Combinations of Predictors

Next we consider the "model selection" (MS) approach of Welch and Goyal (2008). The MS approach estimates regressions with all possible combinations of predictors. Then, at each time period, it selects the one forecast that has performed the best by displaying the lowest cumulative MSE up to that point. The MS approach corresponds to choosing one among 2^N models, which in our case is 2048 models.

We also consider predictive regressions based on Principal Component Analysis (PCA), which is recently implemented in equity premium prediction by Neely, Rapach, Tu and Zhou (2014). This approach involves estimating a set of principal components that parsimoniously incorporate information from the 11 predictors. As in Neely, Rapach, Tu and Zhou (2014), at each time period we select the number of principal components that give the highest adjusted R^2 using data up to that point.

Finally, we estimate the kitchen-sink regression with OLS. This makes it straightforward to assess the effect of shrinkage estimation on the KS regression.

because it has the wrong sign.

3.2.4 Forecast Combinations

We also form forecast combinations, which are designed to combine the forecasts of several predictive regressions that condition on one predictor (see, e.g., Timmermann, 2006). Following Rapach, Strauss and Zhou (2010), we implement two approaches to forecast combination. First, we simply compute the equally-weighted average of all forecasts at each point in time. We refer to this as the "mean" combination.[12] Second, we compute a weighted average of the individual forecasts at each point in time using as weights the inverse of the discounted MSE of each model up to that point. The discount factor is set to 0.90 as in Rapach, Strauss and Zhou (2010). We refer to this as the "MSE" combination.[13]

3.2.5 Grand Combinations of Pooling Information and Pooling Forecasts

Our predictive framework provides a way of pooling information since it is based on a kitchen-sink regression that directly conditions on a large number of predictors. In contrast, combined forecasts are designed to pool forecasts rather than pool information. Pooling forecasts involves two stages: first, estimate several predictive regressions each conditioning on one predictor, and then, second, combine the individual forecasts into one forecast combination. In theory, this two-stage process introduces an efficiency loss and ignores the correlations between the predictors. Therefore, it is often argued that pooling information is optimal relative to pooling forecasts (e.g., Timmermann, 2006).

In addition to just pooling information or just pooling forecasts, we also form a "grand" combination of the two approaches. This provides a framework for assessing whether pooling forecasts adds to the predictive ability of pooling information and vice versa. The grand combination is an interesting addition to the model set because, as we will see later, the combined forecasts are the closest competitor to our predictive framework. In particular, we form equally-weighted grand combinations of forecasts by combining: (i) the e-net KS forecasts with the mean combined forecasts; (ii) the e-net KS forecasts with the MSE combined forecasts; (iii) the lasso KS forecasts with the mean combined forecasts; and (iv) the lasso KS forecasts with the MSE combined forecasts. We focus on the e-net and lasso because, as we will see later, there are the best performing shrinkage models. In all cases we impose the economic constraints since they seem to universally improve prediction.

[12] Note that the mean combination is equivalent to a kitchen-sink regression with fixed slope coefficients equal to $\frac{1}{N}\beta_j$, where β_j is the slope from the regressions with only one predictor. See Rapach, Strauss and Zhou (2010) for more details.

[13] Note that we have also implemented other forecast combinations used in Rapach, Strauss and Zhou (2010), but we find that their performance is very similar to the mean and MSE combinations. These other combinations include the median and trimmed mean as well as MSE combinations with discount factors equal to 0.95 and 1.0.

3.3 Out-of-Sample Analysis

All empirical models are evaluated out of sample (OOS) relative to the historical mean benchmark. We generate OOS forecasts with rolling predictive regressions using a 20-year estimation window such that the first forecast is for January 1947 and the last for December 2014. We adopt a rolling window approach to be consistent with Welch and Goyal (2008) and the ensuing literature. We also show in the Internet Appendix that a 20-year estimation window delivers better results than alternative estimation windows.

The main statistical criterion for evaluating the OOS predictive ability of the models is the Campbell and Thompson (2008) and Welch and Goyal (2008) OOS R^2 statistic, R^2_{oos}. The R^2_{oos} compares the unconditional one-month ahead forecasts $\overline{r}^e_{t+1|t}$ of the historical mean benchmark to the conditional forecasts $\widehat{r}^e_{t+1|t}$ of the alternative model, and is defined as follows:

$$R^2_{oos} = 1 - \frac{MSE\left(\widehat{r}^e_{t+1|t}\right)}{MSE\left(\overline{r}^e_{t+1|t}\right)} = 1 - \frac{\sum_{t=1}^{T-1}\left(r^e_{t+1} - \widehat{r}^e_{t+1|t}\right)^2}{\sum_{t=1}^{T-1}\left(r^e_{t+1} - \overline{r}^e_{t+1|t}\right)^2}. \quad (3)$$

A positive R^2_{oos} implies that the alternative model outperforms the benchmark by means of lower MSE.

We assess the statistical significance of the R^2_{oos} statistic by applying the Clark and West (2006, 2007) testing procedure. This is a test of the null hypothesis of equal predictive ability between the benchmark and the alternative model. The Clark and West (2006, 2007) procedure accounts for the fact that, under the null, the MSE of the benchmark is expected to be lower. This is because the alternative models estimate a parameter vector that, under the null, is not helpful in prediction thus introducing noise into the forecasting process. Clark and West (2006, 2007) propose to adjust the MSE as follows:

$$MSE_{adj} = \frac{1}{T-1}\sum_{t=1}^{T-1}(r^e_{t+1} - \widehat{r}^e_{t+1|t})^2 - \frac{1}{T-1}\sum_{t=1}^{T-1}(\overline{r}^e_{t+1|t} - \widehat{r}^e_{t+1|t})^2. \quad (4)$$

Then, we define:

$$\widehat{test}_{t+1} = (r^e_{t+1} - \overline{r}^e_{t+1|t})^2 - [(r^e_{t+1} - \widehat{r}^e_{t+1|t})^2 - (\overline{r}^e_{t+1|t} - \widehat{r}^e_{t+1|t})^2], \quad (5)$$

and regress $\widehat{test}_{t+1}$ on a constant, using the t-statistic for a zero coefficient. Even though the asymptotic distribution of this test is non-standard (e.g., McCracken, 2007), Clark and West (2006, 2007) show that standard normal critical values provide a good approximation, and therefore recommend to reject the null of equal predictive ability if the statistic is greater than +1.645 (for a one-sided 0.05 test) or +2.326 (for a one-sided 0.01 test).

3.4 Empirical Results

3.4.1 Main Results

We assess the OOS performance of the empirical models by reporting the R^2_{oos} in Table 1. The models include: three KS regressions using shrinkage estimation based on the e-net, lasso and ridge regression; the mean and MSE forecast combinations of Rapach, Strauss and Zhou (2010); four grand combinations of the e-net+mean, e-net+MSE, lasso+mean and lasso+MSE; the model selection (MS) approach of Welch and Goyal (2008); the principal component analysis (PCA) approach of Neely, Rapach, Tu and Zhou (2014); a KS regression estimated with OLS; and, finally, 11 OLS predictive regressions that condition on one predictor at a time.

Table 1 reports results for regressions that apply the two economic constraints of positive forecasts and bounded slopes. Our discussion focusses on the constrained regressions because the two constraints improve performance almost universally. In addition to the full sample, we also report results for the two subsamples of expansions and recessions. Note that the expansion and recession results are based on estimating the models each month over the full forecasting period and then separating the forecasting errors ex post across the two subsamples as in Neely, Rapach, Tu and Zhou (2014).

Our main result is that the three shrinkage estimators of the KS regression deliver an R^2_{oos} that is positive, significant and higher than all other models. Indeed, the shrinkage models are the only models that have a positive and significant R^2_{oos} in both expansions and recessions. The results also indicate that the e-net and the lasso perform the best. For the full sample, the lasso R^2_{oos} is slightly better than the e-net (1.77% vs. 1.64%). The lasso is also better in recessions (5.05% vs. 3.44%), but the e-net is better in expansions (0.89% vs. 0.40%).

The closest competitors to the shrinkage models are the combined forecasts. However, the shrinkage models exhibit a much higher R^2_{oos} for the full sample, which becomes even higher in recessions. For example, the mean combination delivers an R^2_{oos} of 0.82% compared to 1.64% for the e-net and 1.77% for the lasso. This evidence, therefore, strongly favors our approach of pooling information to the standard approach of pooling forecasts.

Turning to the grand combinations of pooling information and pooling forecasts, we find that they deliver similar performance to the plain shrinkage models. For example, the lasso+mean combination produces an R^2_{oos} of 1.70% and the lasso+MSE combination produces an R^2_{oos} of 1.67%, but the plain lasso delivers a higher R^2_{oos} of 1.77%. Hence the grand combinations do not add to the predictive ability of our approach.

Next are the standard models for combining predictors, which do not perform well. The R^2_{oos} of the MS is -1.73%, for the PCA it is -0.06% and for the OLS KS it is -55.84%. Notably, the OLS KS model is by far the worst performing model, which indicates that the way we estimate the KS regression is of critical importance.

Among the single predictors, the best performing model is *dy* (the dividend yield) with an R^2_{oos} of 0.72, which is significant at 1%. However, *dy* has an R^2_{oos} that is less than half of the value of the e-net or lasso and it is only positive in expansions, while being negative in recessions. In fact, among the 11 predictors, three have a significantly positive R^2_{oos} in expansions (*dy*, *epr* and *bm*), and three do so in recessions (*tbl*, *ltr* and *tms*). None of the 11 predictors have a significantly positive R^2_{oos} in both expansions and recessions. Only the three shrinkage estimators achieve a significantly positive R^2_{oos} for the full sample as well as both expansions and recessions.

To summarize, evidence based on the R^2_{oos} indicates that the best model for predicting the equity premium out of sample is a kitchen-sink regression that conditions on a large set of predictors and imposes statistical constraints through shrinkage estimation together with economic constraints on slope coefficients and forecasts.

3.4.2 The Effect of Economic Constraints on Performance

Given the prominence of statistical and economic constraints in our analysis, it is useful to assess their relative importance. To this end, Table 2 demonstrates the effect of constraints on the R^2_{oos} of the models. Our discussion focusses on the main results by first isolating the effect of statistical constraints in the absence of economic constraints. Specifically, the unconstrained OLS KS delivers an R^2_{oos} of -11.86% (insignificant) but when imposing the lasso constraint the R^2_{oos} rises to -0.63% (significant).[14] Hence the statistical constraints due to shrinkage estimation deliver a massive improvement on predictability relative to OLS, but alone they fall short of producing a positive R^2_{oos}.

Then, imposing the economic constraints raises the lasso R^2_{oos} from -0.63% to 1.77%, which is now positive, significant at 1% and the highest of all models. Therefore, the economic constraints seem to be very effective when applied to many predictors in a single regression estimated with shrinkage. In short, it is the combination of economic and statis-

[14]Note that the Clark and West (2006, 2007) statistic is testing the one-sided null hypothesis of equal predictive accuracy in population, while the reported R^2_{oos} values reflect finite-sample performance. For this reason, a rejection of the null hypothesis may occasionally be associated with a negative R^2_{oos}. This is the case here for the shrinkage estimators in the absence of economic constraints, where although the R^2_{oos} is slightly negative, the test statistic rejects the null in favor of higher predictive accuracy for the shrinkage estimators.

tical constraints in the context of the kitchen-sink regression that delivers the most powerful results.

3.4.3 Assessing Performance over Time

To illustrate the results, Figures 1 and 2 plot the OOS performance of each model over time. Following Welch and Goyal (2008), the figures plot the difference of the cumulative squared error of the null (historical mean) minus the cumulative squared error of the alternative. All plots are for forecasts subject to the two economic constraints. The figures show that the shrinkage estimators display the most pronounced upward trend in their OOS performance over time. The closest competitors are again the combined forecasts, while the majority of the remaining models do not exhibit a sustained positive trend over time. Hence the figures clearly demonstrate that the good performance of shrinkage estimators is not due to a particular subsample but is systematic over a long sample spanning the full postwar period.

To add to this evidence, Table 3 reports the R^2_{oos} of the shrinkage estimators for various subsamples. The results are particularly good for lasso, the best model so far, which delivers a positive R^2_{oos} for every subsample since the 1950s. Overall, the lasso seems to be the most stable model and the one that consistently beats the benchmark from the 1960s onwards.

3.4.4 The Effect of Shrinkage on Slope Coefficients

Next we assess the effect of shrinkage on slope coefficients. Table 4 shows the average slope estimates of (i) the e-net KS regression, and (ii) the OLS KS regression, where in both cases the economic constraints are applied. This allows to determine the magnitude of shrinkage imposed by the most general shrinkage estimator on each of the predictors. Overall, the average slope estimate of the e-net averaged across time and across predictors is 0.042, which is one third of 0.127, the value for the OLS KS. We can see, therefore, that shrinkage has a substantial effect in reducing the slope estimates.

In addition to the actual slopes, we also report standardized slopes, which are the slopes for predictors divided by their standard deviation up to that point. The standardized slopes allow us to determine the relative importance of each predictor as they capture the effect on the one-month ahead equity premium of a one standard deviation change in the predictor. Note that the reported standardized slopes have been multiplied by 100. We find that for the e-net, dy has the greatest positive impact as a one standard deviation increase in dy on average raises the monthly equity premium by 20.8 basis points. This is consistent with the popularity of the dividend yield as a predictor of the equity premium in the literature (e.g., Cochrane, 2008, 2011). In contrast, *tbl* has the greatest negative impact as a one standard deviation increase in *tbl* lowers the monthly equity premium by 32.3 basis points.

Finally, Figure 3 illustrates the time-variation of the slope estimates of the e-net and hence shows when the slopes are set to zero by the constraints.

4 Predictability and Asset Allocation

4.1 A Mean-Variance Trading Strategy

Following Campbell and Thompson (2008), we assess the economic value of equity premium predictability using a dynamic asset allocation strategy. The strategy involves monthly rebalancing of a portfolio that invests in the S&P 500 index (the risky asset) and the treasury bill (the riskless asset). We consider a mean-variance investor with a one-month ahead horizon, who determines the optimal weights by implementing a maximum expected utility rule as follows:

$$\begin{aligned} \max_{w_t} \quad & E_t\left[U\left(r_{p,t+1}\right)\right] && = r_{p,t+1|t} - \tfrac{\gamma}{2}\sigma^2_{p,t+1|t} \\ \text{s.t.} \quad & r_{p,t+1|t} && = w_t r_{t+1|t} + \left(1-w_t\right) r_f, \\ & \sigma^2_{p,t+1|t} && = w_t^2 \sigma^2_{t+1|t}, \end{aligned} \tag{6}$$

where $r_{p,t+1|t}$ is the $t+1$ forecast of the portfolio return conditional on time t information, γ is the investor's degree of relative risk aversion, $\sigma^2_{p,t+1|t}$ is the $t+1$ forecast of the portfolio variance made at time t, $r_{t+1|t}$ is the $t+1$ forecast of the S&P 500 index return made at time t, r_f is the risk-free rate of return, and $\sigma^2_{t+1|t}$ is the $t+1$ forecast of the variance to the S&P 500 index return made at time t. Note that we forecast $\sigma^2_{t+1|t}$ using a 5-year rolling average of the variance of past monthly returns as in Campbell and Thompson (2008). We also set $\gamma = 5$ as in Neely, Rapach, Tu and Zhou (2014).

The solution to the maximum expected utility rule delivers the risky asset weight:

$$w_t = \frac{1}{\gamma}\frac{r_{t+1|t} - r_f}{\sigma^2_{t+1|t}}. \tag{7}$$

Consistent with the literature (e.g., Campbell and Thompson, 2008; and Neely, Rapach, Tu and Zhou, 2014) we constrain the weight on the risky asset by imposing $w_t \in [0, 1.5]$. In other words, we do not allow short-selling and leverage is limited to no more than 50%.

We evaluate the performance of portfolios generated by a given set of equity premium forecasts using the Sharpe ratio (SR) and the certainty equivalent return (CER). The Sharpe ratio is perhaps the most commonly used performance measure and is defined as the average excess return of a portfolio divided by the standard deviation of the portfolio returns. We assess statistical significance using the Ledoit and Wolf (2008) bootstrap two-sided test of whether the Sharpe ratio of the alternative model is different from the benchmark.

The certainty equivalent return is defined as:

$$CER = \left(\overline{r}_p - \frac{\gamma}{2}\overline{\sigma}_p^2\right), \tag{8}$$

where $\overline{r}_p$ is the mean portfolio return and $\overline{\sigma}_p^2$ is the portfolio variance over the forecast evaluation period. The CER can be interpreted as the performance fee the risk-averse investor is willing to pay for switching from the riskless asset to the risky portfolio. We focus on the difference in CER (ΔCER), which is equal to the CER of the portfolio generated by the forecasts of the alternative model minus the CER of the portfolio generated by the historical mean benchmark. ΔCER measures the performance fee the risk-averse investor is willing to pay for switching from the risky portfolio generated by the benchmark model to the risky portfolio generated by the alternative model. To provide a realistic assessment of the profitability of dynamic trading strategies, we also take into account the effect of transaction costs. In particular, we compute the ΔCER net of proportional transaction costs equal to 50 basis points per month as in Neely, Rapach, Tu and Zhou (2014).[15]

Finally, we compute the average turnover of each trading strategy, which is defined as follows:

$$Turnover = \frac{1}{T-1}\sum_{t=1}^{T-1}\left(|w_{t+1} - w_{t+1}^{-}|\right), \tag{9}$$

where $T-1$ is the number of trading periods, w_{t+1} is the weight on the risky asset at time $t+1$, and $w_{t+1}^{-} = w_t \frac{1+r_{t+1}}{1+r_{p,t+1}}$ is the weight on the risky asset right before rebalancing at time $t+1$. This turnover measure represents the average monthly trading volume. We report the average relative turnover, which is the ratio of the average turnover of the alternative model divided by the average turnover of the benchmark.

4.2 Portfolio Performance

We assess the performance of dynamically rebalanced portfolios generated by the monthly forecasts of the predictive models. Table 5 reports the empirical findings. The first result to note is that the historical mean benchmark delivers a $CER = 5.82\%$ per year relative to riskless investing. The CER rises to 8.33% in expansions but falls to −6.81% in recessions.

[15] Note that the statistical significance of the CER can be assessed with the McCracken and Valente (2012) bootstrap test. However, this approach is not feasible in our context since estimation with the elastic net is computationally intensive and we cannot repeat out-of-sample estimation hundreds or thousands of times as would be required by the McCracken and Valente (2012) test. Hence we focus on assessing the statistical significance of Sharpe ratios.

Clearly, using the historical mean is a poor predictor of the equity premium in recessions.

Similar to our statistical findings, our main result here is that any of the three shrinkage KS regressions performs better than any of the other models. For example, the lasso delivers a ΔCER relative to the historical mean benchmark of 2.71% per year, which becomes 0.71% in expansions and 12.71% in recessions. Net of transaction costs, the lasso retains the highest ΔCER, which is 1.36% per year. Moreover, the annualized Sharpe ratio of the lasso strategy is 0.66, which is significantly higher than the 0.45 of the benchmark according to the Ledoit and Wolf (2008) bootstrap two-sided test. The results for the e-net are similar. It is also interesting to note that the shrinkage strategies substantially outperform the combined forecasts (the closest competitor according to R^2_{oos}), which display little economic value above the benchmark.

Overall, this is evidence that the equity premium is predictable out of sample and, in the context of a dynamic mean-variance strategy, there is high economic value in using a KS regression with both statistical and economic constraints. For example, the economic gains of the lasso approach can be summarized into a performance fee of 2.71% per year before transaction costs and 1.36% per year after transaction costs, together with an increase in the Sharpe ratio from 0.45 to 0.66.

The superior performance of the shrinkage estimators in recessions is an important finding because the predictive information of economic fundamentals is more valuable to an investor during recessions. This is true because during recessions the equity premium is on average negative with high volatility, which makes the historical mean a poor forecast. Furthermore, Kacperczyk, Van Nieuwerburgh and Veldkamp (2016) show that the state of the economy changes the way investors process information. Fund managers care more about aggregate shocks in recessions because these are periods in which stocks contain more aggregate risk. In short, therefore, the information in economic fundamentals for predicting the equity premium matters the most in recessions, and we find that this is when our predictive framework performs the best.

4.3 Comparing Statistical and Economic Gains

Our statistical analysis has established that the shrinkage models perform well in out-of-sample prediction of the equity premium. For example, the lasso delivers an R^2_{oos} of 1.77%, which is positive, significant and the highest of all other models. In this section, we relate the R^2 of the models with the Sharpe ratio of the strategies. Following Campbell and Thompson (2008), "...the correct way to judge the magnitude of R^2 is to compare it with the squared

Sharpe ratio..."[16] since the proportional increase in the expected return is approximately equal to the ratio of $\frac{R^2_{oos}}{SR^2}$, where SR is the unconditional Sharpe ratio of the risky asset.[17]

For example, recall that the lasso delivers an R^2_{oos} of 1.77%. Over the same forecasting period (1947-2014), the risky asset has a squared monthly Sharpe ratio of $0.1521^2 = 0.0231$. Then, the proportional increase in the expected return is $0.0177/0.0231 = 0.7662$. In other words, the lasso forecasts can increase the average monthly portfolio return by a factor of 76.62%. Campbell and Thompson (2008) show that this corresponds to an actual increase in the expected return of $\left(\frac{1}{\gamma}\right)\left(\frac{R^2_{oos}}{1-R^2_{oos}}\right)(1+SR^2) = 0.37\%$ per month or 4.44% per year. In conclusion, modest predictive ability for the equity premium can plausibly generate large economic gains.

5 Prediction with Technical Indicators

In assessing the predictability of the equity premium, our empirical analysis focuses primarily on the information content of economic fundamentals. At the same time, however, there is a series of technical indicators that have been used by the academic literature for the same purpose. In this section, we compare the empirical performance of models that condition on economic fundamentals to models that condition on technical indicators. We use the 14 technical indicators recently used by Neely, Rapach, Tu and Zhou (2014), which are based on three popular trend-following strategies.

The first strategy is based on the moving average (MA) rule that generates a buy or sell signal as follows:

$$S_{i,t} = \begin{Bmatrix} 1 & if\ MA_{s,t} \geq MA_{l,t} \\ 0 & if\ MA_{s,t} < MA_{l,t} \end{Bmatrix}, \tag{10}$$

where

$$MA_{j,t} = \frac{1}{j}\sum_{i=0}^{j-1} P_{t-i} \qquad for\ j = s, l. \tag{11}$$

Note that P_t is the price of the S&P 500 index and s (l) is the length of the short (long) MA, where $s < l$. We use six MA specifications: MA(1,9), MA(2,9), MA(3,9), MA(1,12), MA(2,12) and MA(3,12).

The second strategy is based on the momentum (MOM) rule that generates a buy or sell

[16] See Campbell and Thompson (2008), page 1525.

[17] Specifically, when moving from the unconditional forecast of the expected return to a conditional forecast, the proportional increase in the expected return is $\left(\frac{R^2_{oos}}{1-R^2_{oos}}\right)\left(\frac{1+SR^2}{SR^2}\right)$, which is approximately equal to $\frac{R^2_{oos}}{SR^2}$, when R^2_{oos} and SR^2 are both small.

signal as follows:

$$S_{i,t} = \begin{Bmatrix} 1 & if\ P_t \geq P_{t-m} \\ 0 & if\ P_t < P_{t-m} \end{Bmatrix}, \quad (12)$$

where m is the number of lagged periods. We use two MOM specifications: MOM(9) and MOM(12).

The third strategy is based on the on-balance volume rule that generates a buy or sell signal as follows:

$$S_{i,t} = \begin{Bmatrix} 1 & if\ MA_{s,t}^{OBV} \geq MA_{l,t}^{OBV} \\ 0 & if\ MA_{s,t}^{OBV} < MA_{l,t}^{OBV} \end{Bmatrix}, \quad (13)$$

where

$$MA_{j,t}^{OBV} = \frac{1}{j}\sum_{i=0}^{j-1} OBV_{t-i} \qquad for\ j = s, l, \quad (14)$$

and

$$OBV_t = \sum_{k=1}^{t} VOL_k D_k. \quad (15)$$

Note that VOL_k is a measure of trading volume in period k, and D_k is a binary variable that takes a value of 1 if $P_k - P_{k-1} \geq 0$ and -1 otherwise. We use six VOL specifications: VOL(1,9), VOL(2,9), VOL(3,9), VOL(1,12), VOL(2,12) and VOL(3,12).

The R^2_{oos} of this exercise are reported in Table 6. In addition to individual technical indicators, we also report on combinations of all technical indicators (TECH), all economic fundamentals (ECON) and grand combinations of both TECH and ECON. These combinations are obtained using four methods: PCA, mean combined forecasts, e-net and lasso. The monthly forecasts are obtained using the same 20-year rolling window but for the shorter sample period of January 1951 to December 2014 due to data availability. Note that we do not impose economic constraints on the technical predictors. However, we do impose the economic constraints on the economic fundamentals whenever these are used in the analysis.

The results overwhelmingly show that there is little or no predictive ability in technical indicators. Although half of the indicators have a positive R^2_{oos} for the full sample, none of them has a significant R^2_{oos}. Notably all indicators have a negative R^2_{oos} in expansions and a positive R^2_{oos} in recessions, which is consistent with the evidence reported in Neely, Rapach, Tu and Zhou (2014).[18]

More importantly, the technical indicators do not add to the R^2_{oos} when combined with

[18]Note that, although technical indicators perform well in recessions, the kitchen-sink regression that conditions on technical indicators does not perform well in recessions. This is likely due to the absence of economic constraints because there is no economic rationale for imposing economic constraints on kitchen-sink regressions with technical indicators. In unreported results, we find that, for example, the positive forecast constraint improves the performance of the kitchen-sink regression with technical indicators. However, this performance is still quite inferior to kitchen-sink regressions with economic fundamentals.

economic fundamentals using some of the techniques implemented in this paper: PCA (as in Neely, Rapach, Tu and Zhou, 2014), mean forecast combinations, e-net and lasso. Once again, the best performing models are the e-net with an R^2_{oos} of 1.39% and the lasso with an R^2_{oos} of 1.38%. When these models also incorporate technical indicators (enet-ALL and lasso-ALL), the R^2_{oos} falls to 0.45% and 1.13% respectively. We conclude, therefore, that technical indicators do not add to the predictive ability of our framework that conditions on economic fundamentals.

6 Conclusion

Return predictability is conceptually important in asset pricing because it helps us understand what drives stock price variation. It is also practically important because out-of-sample predictability allows investors and firms to make better real-time investment decisions. This paper investigates the predictability of the equity premium using a kitchen-sink regression that directly conditions on a large set of economic fundamentals. The regression is estimated with a shrinkage methodology designed to maximize predictive performance. We then follow Campbell and Thompson (2008) in imposing a sign constraint on the slope coefficients and a positivity constraint on the forecasts, which are motivated by economic theory.

We implement this framework using a long sample of monthly data and arrive at three main empirical findings. First, the equity premium is predictable out of sample as our predictive framework consistently outperforms the historical mean benchmark as well as all competing models. For example, the R^2_{oos} of the best shrinkage model (the lasso) is 1.77%, which is significant at 1%. Second, a dynamic mean-variance strategy that conditions on the shrinkage forecasts can generate a certainty equivalent return of about 2.7% per year over and above the benchmark, which is the highest among all models. Third, conditioning on technical indicators adds little or no value to the predictive content of economic fundamentals in equity premium prediction. Overall, our analysis highlights the role of economic and statistical constraints in the context of kitchen-sink regressions for assessing the predictability of the equity premium, and suggests that economic fundamentals contain important information about future returns.

References

Campbell, J.Y., and R.J. Shiller (1988). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies* **1**, 195–227.

Campbell, J.Y., and S.B. Thompson (2008). "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies* **21**, 1509–1531.

Clark, T.E., and K.D. West (2006). "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis," *Journal of Econometrics* **135**, 155–186.

Clark, T.E., and K.D. West (2007). "Approximately Normal Tests for Equal Predictive Accuracy," *Journal of Econometrics* **138**, 291–311.

Cochrane, J. (2008). "The Dog that Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies* **21**, 1533–1575.

Cochrane, J. (2011). "Presidential Address: Discount Rates," *Journal of Finance* **66**, 1047–1108.

Dangl, T., and M. Halling (2012). "Predictive Regressions with Time-Varying Coefficients," *Journal of Financial Economics* **106**, 157–181.

Della Corte, P., L. Sarno, and I. Tsiakas (2009). "An Economic Evaluation of Empirical Exchange Rate Models," *Review of Financial Studies* **22**, 3491–3530.

DeMiguel, V., L. Garlappi, F.J. Nogales, and R. Uppal (2009). "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms," *Management Science* **55**, 798–812.

Hoerl, A.E., and R.W. Kennard (1970). "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics* **12**, 55–67.

Jagannathan, R., and T. Ma (2003). "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance* **58**, 1651–1684.

Jorion, P. (1985). "International Portfolio Diversification with Estimation Risk," *Journal of Business* **58**, 259–278.

Jorion, P. (1986). "Bayes-Stein Estimation for Portfolio Analysis," *Journal of Financial and Quantitative Analysis* **21**, 279–292.

Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp (2016). "A Rational Theory of Mutual Funds' Attention Allocation," *Econometrica* **84**, 571–626.

Kelly, B., and S. Pruitt (2013). "Market Expectations in the Cross-Section of Present Values," *Journal of Finance* **68**, 1721–1756.

Koijen, R., and S. Van Nieuwerburgh (2011). "Predictability of Returns and Cash Flows," *Annual Review of Financial Economics* **3**, 467–491.

Ledoit, O., and M. Wolf (2003). "Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection," *Journal of Empirical Finance* **10**, 603–621.

Ledoit, O., and M. Wolf (2008). "Robust Performance Hypothesis Testing with the Sharpe Ratio," *Journal of Empirical Finance* **15**, 850–859.

Li, J. (2015). "Sparse and Stable Portfolio Selection with Parameter Uncertainty," *Journal of Business and Economic Statistics* **33**, 381–392.

Li, J., I. Tsiakas, and W. Wang (2015). "Predicting Exchange Rates Out of Sample: Can Economic Fundamentals Beat the Random Walk?" *Journal of Financial Econometrics* **13**, 293–341.

Ludvigson, S.C., and S. Ng. (2007). "The Empirical Risk-Return Relation: A Factor Analysis Approach," *Journal of Financial Economics* **83**, 171–222.

McCracken, M.W. (2007). "Asymptotics for Out of Sample Tests of Granger Causality," *Journal of Econometrics* **140**, 719–752.

McCracken, M.W., and G. Valente (2012). "Asymptotic Inference for Performance Fees and the Predictability of Asset Returns," Federal Reserve Bank of St. Louis Working Paper 2012-049B.

Neely, C.J., D.E. Rapach, J. Tu, and G. Zhou (2014). "Forecasting the Equity Risk Premium: The Role of Technical Indicators," *Management Science* **60**, 1772–1791.

Pettenuzzo, D., A. Timmermann, and R. Valkanov (2014). "Forecasting Stock Returns under Economic Constraints," *Journal of Financial Economics* **114**, 517–553.

Rapach, D.E., J.K. Strauss, and G. Zhou (2010). "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy," *Review of Financial Studies* **23**, 821–862.

Rapach, D.E., J.K. Strauss, and G. Zhou (2013). "International Stock Return Predictability: What is the Role of the United States?" *Journal of Finance* **68**, 1633–1662.

Tibshirani, R. (1996). "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)* **58**, 267–288.

Timmermann, A. (2006). "Forecast Combinations," in Elliott, G., C.W.J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Volume 1, Amsterdam: North-Holland.

Tu, J., and G. Zhou (2011). "Markowitz Meets Talmud: A Combination of Sophisticated and Naive Diversification Strategies," *Journal of Financial Economics* **99**, 204–215.

van Binsbergen, J.H., and R. Koijen (2010). "Predictive Regressions: A Present-Value Approach," *Journal of Finance* **65**, 1439–1471.

Welch, I., and A. Goyal (2008). "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies* **21**, 1455–1508.

Zou, H., and T. Hastie (2005). "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **67**, 301–320.

Table 1. Equity Premium Prediction

The table displays the out-of-sample R^2_{oos} in percent for predictive models of the monthly equity premium against the null of the historical mean. The R^2_{oos} is for models that impose a sign constraint on the slope coefficients and a positivity constraint on the forecasts. The out-of-sample monthly forecasts are obtained using a 20-year rolling window for the sample period of January 1927 to December 2014. Expansions and recessions are according to the NBER. **, and *** denote statistical significance at the 5%, and 1% level, respectively, using the Clark and West (2006, 2007) one-sided t-statistic.

	R^2_{oos} (%)		
Predictor	*Full Sample*	*Expansion*	*Recession*
E-net KS	1.64***	0.89***	3.44**
Lasso KS	1.77***	0.40**	5.05***
Ridge KS	1.17***	0.90***	1.83**
Mean	0.82***	0.86***	0.72
MSE	0.87***	0.86***	0.91
E-net KS + Mean	1.69***	1.28***	2.65**
E-net KS + MSE	1.66***	1.28***	2.56**
Lasso KS + Mean	1.70***	1.01***	3.34***
Lasso KS + MSE	1.67***	1.02***	3.24***
MS	−1.73	−3.27	1.94**
PCA	−0.06	0.92***	−2.39
OLS KS	−55.84	−72.17	−16.90
dy	0.72***	1.07***	−0.11
epr	0.03**	1.13***	−2.58
bm	0.05	0.10**	−0.08
ntis	−0.26	−0.32	−0.11
svar	−2.38	−2.62	−1.83
tbl	0.35**	−0.43	2.20**
ltr	0.39**	−0.57	2.67**
tms	0.08***	−1.12	2.94***
dfy	−1.57	−1.04	−2.84
dfr	−0.76	−0.17	−2.16
infl	−0.28	−0.61	0.51

Table 2. Equity Premium Prediction – The Effect of Constraints

The table displays the out-of-sample R^2_{oos} in percent for predictive models of the monthly equity premium against the null of the historical mean. The R^2_{oos} is for four types of models: models that impose no constraints; models that impose only the sign constraint on the slope coefficients; models that impose only the positivity constraint on the forecasts; and models that impose both constraints. The out-of-sample monthly forecasts are obtained using a 20-year rolling window for the sample period of January 1927 to December 2014. **, and *** denote statistical significance at the 5%, and 1% level, respectively, using the Clark and West (2006, 2007) one-sided t-statistic.

	R^2_{oos} (%)			
Predictor	*No Constraint*	*Slope Constraint*	*Forecast Constraint*	*Both Constraints*
E-net KS	−0.46**	0.97***	1.22***	1.64***
Lasso KS	−0.63**	0.65***	1.15***	1.77***
Ridge KS	−0.15**	0.42***	1.29***	1.17***
Mean	0.67**	0.79***	0.64***	0.82***
MSE	0.60**	0.75***	0.64***	0.87***
E-net KS + Mean	0.83**	1.57***	1.36***	1.69***
E-net KS + MSE	0.88**	1.59***	1.36***	1.66***
Lasso KS + Mean	0.62**	1.35***	1.21***	1.70***
Lasso KS + MSE	0.66**	1.37***	1.21***	1.67***
MS	−3.60	−3.33	−1.18	−1.73
PCA	−0.32	−0.32	−0.06	−0.06
OLS KS	−11.86	−479.91	−5.28	−55.84
dy	0.44**	0.91***	0.51**	0.72***
epr	−1.65	−2.76	−0.24	0.03**
bm	−1.42	−0.21	−0.49	0.05
ntis	−0.72	−0.74	−0.14	−0.26
svar	−3.16	−2.39	−2.67	−2.38
tbl	−1.75	−0.98	−0.47	0.35**
ltr	−0.20**	0.14**	0.05**	0.39**
tms	−0.62**	−0.46**	−0.09**	0.08***
dfy	−1.84	−2.07	−0.64	−1.57
dfr	−1.48	−1.10	−1.06	−0.76
infl	−0.30	−0.70	−0.10	−0.28

Table 3. Equity Premium Prediction – Subsamples

The table displays the out-of-sample R^2_{oos} in percent for different subsamples for the three shrinkage models: the elastic-net kitchen-sink regression, the lasso kitchen-sink regression and the ridge kitchen-sink regression. The subsample results are based on out-of-sample monthly forecasts obtained using a 20-year rolling window for the sample period of January 1927 to December 2014. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively, using the Clark and West (2006, 2007) one-sided t-statistic.

	R^2_{oos} (%)		
	E-net KS	*Lasso KS*	*Ridge KS*
Full sample			
1947–2014	1.64***	1.77***	1.17***
10-year subsamples			
1947–49	3.31*	2.34***	0.24
1950–59	−1.32	−1.91	−1.71
1960–69	4.26**	3.12**	5.12***
1970–79	1.53*	0.05	1.59*
1980–89	5.71***	6.29***	4.69***
1990–99	−0.65	0.37	−0.97
2000–09	−0.75	1.30*	−2.23
2010–14	0.30*	0.21**	0.31
15-year subsamples			
1947–54	0.39	0.27	−1.06
1955–69	2.58**	1.26***	3.62***
1970–84	4.00***	3.41***	3.22***
1985–99	1.12*	1.67**	1.01*
2000–14	−0.49	1.04*	−1.61
20-year subsamples			
1947–59	−0.05	−0.74	−0.45
1960–79	2.56***	1.21**	2.92***
1980–99	3.23***	3.98***	2.48***
2000–14	−0.49	1.04*	−1.61

Table 4. The Effect of Shrinkage on Slope Coefficients

The table illustrates the effect of shrinkage on the slope coefficients of the kitchen-sink regression. The table reports the average out-of-sample slopes of the elastic-net kitchen-sink regression and the OLS kitchen-sink regression. *Standardized* are the slopes for predictors which are divided by their standard deviation, whereas *actual* are the slopes for non-standardized predictors. The standardized slopes are multiplied by 100. Both estimators impose the positivity constraint on the slope of all predictors, except for *ntis*, *tbl* and *infl* which are constrained to be negative. % Zero reports the percentage of slopes set to be equal to zero. The out-of-sample forecasts are obtained using a 20-year rolling window for the sample period of January 1927 to December 2014.

	Elastic-net Kitchen-Sink Regression			*OLS Kitchen-Sink Regression*		
	Actual	*Standardized*	% Zero	*Actual*	*Standardized*	% Zero
dy	0.010	0.208	42.9	0.029	0.695	36.3
epr	0.003	0.078	59.3	0.021	0.687	35.0
bm	0.011	0.161	54.1	0.046	0.805	40.0
ntis	−0.123	−0.154	39.3	−0.262	−0.313	23.8
svar	0.145	0.033	79.6	0.532	0.137	69.4
tbl	−0.147	−0.323	41.8	−0.924	−1.587	22.6
ltr	0.094	0.202	36.9	0.147	0.313	15.6
tms	0.128	0.079	68.5	0.448	0.260	70.9
dfy	0.415	0.169	49.8	1.443	0.650	15.8
dfr	0.093	0.110	70.1	0.230	0.269	43.7
infl	−0.166	−0.085	63.4	−0.310	−0.177	52.4
Average	0.042	0.044	55.1	0.127	0.158	38.7

Table 5. Portfolio Performance

The table shows the out-of-sample portfolio performance for a mean-variance investor, who each month rebalances her portfolio by investing in one risky asset (S&P 500) and the riskless rate (T-bill). The investor has a degree of relative risk aversion equal to 5 and follows a maximum utility strategy. ΔCER is the gain in the percent annualized Certainty Equivalent Return (CER) for switching from the forecasts of the benchmark to the forecasts generated by the alternative model. SR is the annualized Sharpe ratio. Turnover is the ratio of the average turnover for the portfolio generated by the forecasts of the alternative model divided by the average turnover for the portfolio generated by the benchmark. For the historical mean benchmark, we report the level of CER and the average turnover for that portfolio. For the alternative models, we impose a sign constraint on slopes and a positivity constraint on forecasts. The superscripts *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively, using the Ledoit and Wolf (2008) bootstrap two-sided test of whether the Sharpe ratio of a model is different from that of the benchmark. The last column is the annual percent ΔCER assuming a proportional transaction cost of 50 basis points per month. The out-of-sample monthly forecasts are obtained using a 20-year rolling window for the sample period of January 1927 to December 2014.

	ΔCER (%)			*SR*	*Turnover*	ΔCER (%)
	Full Sample	*Expansion*	*Recession*	*Full Sample*	*Full Sample*	$c = 50$ *bps* *Full Sample*
Historical mean	5.82	8.33	−6.81	0.45	0.04	5.69
E-net KS	2.52	0.90	10.56	0.64**	6.02	1.32
Lasso KS	2.71	0.71	12.71	0.66**	6.71	1.36
Ridge KS	1.71	0.80	6.16	0.58	5.64	0.60
Mean	0.41	0.50	−0.07	0.48	1.83	0.18
MSE	0.50	0.54	0.24	0.49	2.11	0.20
E-net KS + Mean	1.97	1.05	6.44	0.59**	3.98	1.23
E-net KS + MSE	1.91	1.02	6.24	0.59**	3.90	1.19
Lasso KS + Mean	1.94	0.86	7.28	0.59**	4.25	1.14
Lasso KS + MSE	1.88	0.82	7.08	0.59**	4.20	1.09
MS	1.46	−0.84	13.00	0.47	6.18	0.21
PCA	1.36	1.05	2.78	0.55	3.60	0.70
OLS KS	0.64	−0.12	4.37	0.47	3.32	0.11
dy	0.53	1.13	−2.70	0.48	2.44	0.15
epr	1.20	1.01	2.15	0.53	1.89	1.01
bm	−1.05	−0.44	−4.22	0.38	2.17	−1.32
ntis	−0.02	−0.29	1.32	0.45	2.13	−0.32
svar	−0.94	−0.35	−3.90	0.40	1.57	−1.06
tbl	1.32	−0.62	11.06	0.54	1.46	1.16
ltr	0.81	0.03	4.55	0.52	7.99	−0.80
tms	1.09	0.08	6.06	0.54	2.47	0.75
dfy	−0.79	−0.49	−2.47	0.43	1.83	−0.97
dfr	−0.35	0.10	−2.54	0.43	4.50	−1.19
infl	0.11	−0.35	2.39	0.45	3.09	−0.39

Table 6. Equity Premium Prediction with Technical Indicators

The table displays the out-of-sample R^2_{oos} in percent for models that condition on technical indicators to predict the monthly equity premium. The out-of-sample monthly forecasts are obtained using a 20-year rolling window for the sample period of January 1951 to December 2014. Expansions and recessions are according to the NBER. **, and *** denote statistical significance at the 5%, and 1% level, respectively, using the Clark and West (2006, 2007) one-sided t-statistic.

	R^2_{oos} (%)		
Predictor	*Full Sample*	*Expansion*	*Recession*
MA(1,9)	0.03	−0.60	1.41**
MA(2,9)	−0.06	−0.71	1.38
MA(3,9)	−0.22	−0.77	1.00
MA(1,12)	0.12	−1.11	2.84
MA(2,12)	0.45	−0.83	3.26
MA(3,12)	−0.64	−1.18	0.56
MOM(9)	−0.36	−1.47	2.09
MOM(12)	−0.22	−0.93	1.33
VOL(1,9)	0.38	−0.75	2.88**
VOL(2,9)	0.47	−0.26	2.10
VOL(3,9)	−0.72	−1.50	1.02
VOL(1,12)	0.26	−0.97	2.97
VOL(2,12)	0.31	−0.24	1.54
VOL(3,12)	0.06	−0.75	1.85
PCA-TECH	0.01	−1.65	3.67**
PCA-ECON	−1.20	−0.40	−2.96
PCA-ALL	−1.56	−1.20	−2.34
Mean-TECH	0.18	−0.64	2.00
Mean-ECON	0.48	0.52	0.38
Mean-ALL	0.45***	−0.01	1.56***
E-net KS - TECH	−0.93	−1.10	−0.55
E-net KS - ECON	1.39***	0.37**	3.64**
E-net KS - ALL	0.45**	−0.54	2.64**
Lasso KS - TECH	−0.41	−0.38	−0.46
Lasso KS - ECON	1.38***	0.74**	2.80**
Lasso KS - ALL	1.13***	0.08	3.47**

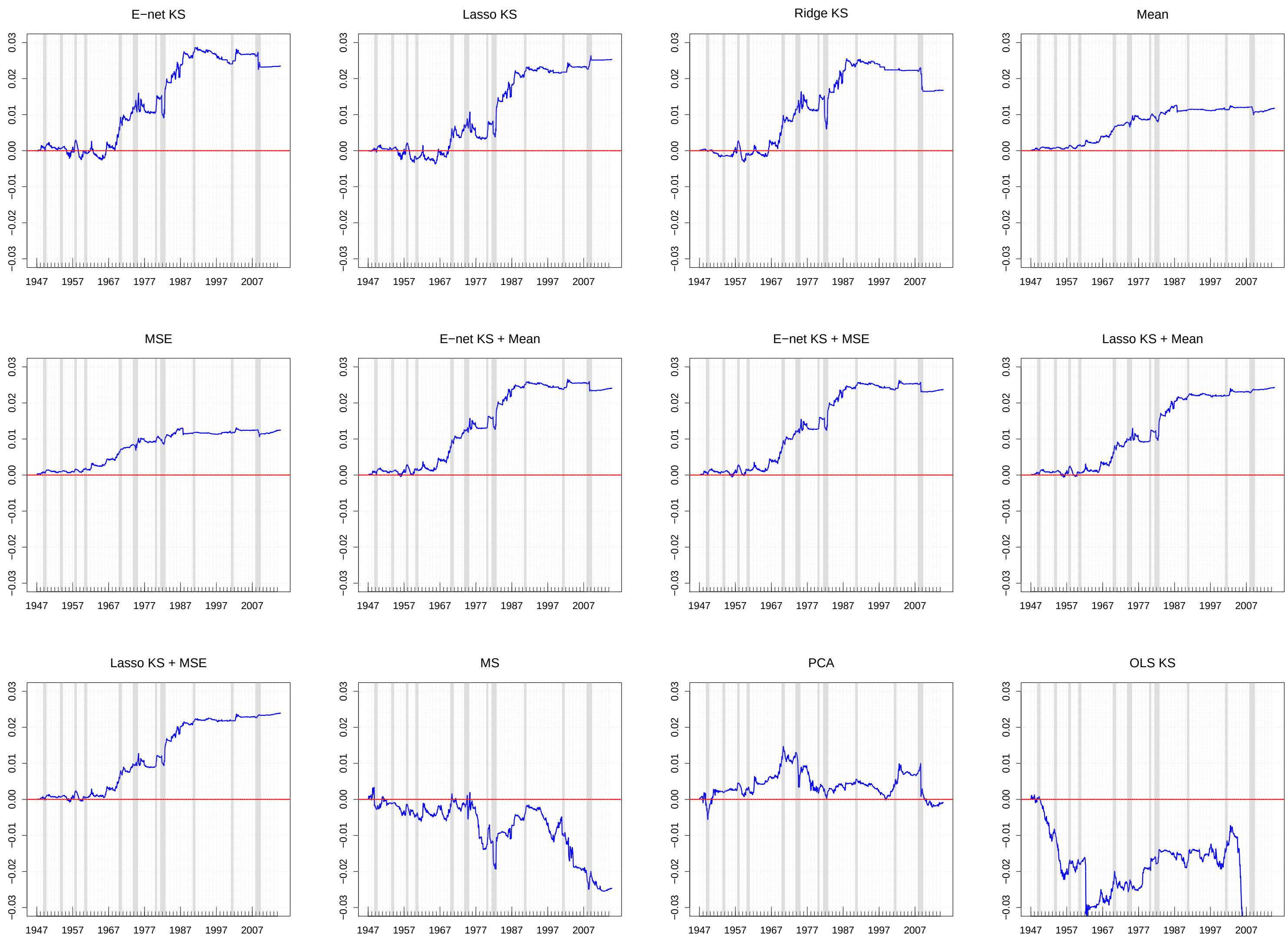

Figure 1. Equity Premium Prediction – multiple predictors

The figure displays the performance of empirical models conditioning on multiple predictors in out-of-sample equity premium prediction. Each graph plots the difference of the cumulative squared error of the null (historical mean) minus the cumulative squared error of the alternative. All forecasts are constrained to be positive and all slopes are bounded by the sign constraints. The grey areas indicate NBER recessions. The out-of-sample monthly forecasts are obtained with rolling regressions using a window of 20 years for the sample period of January 1927 to December 2014.

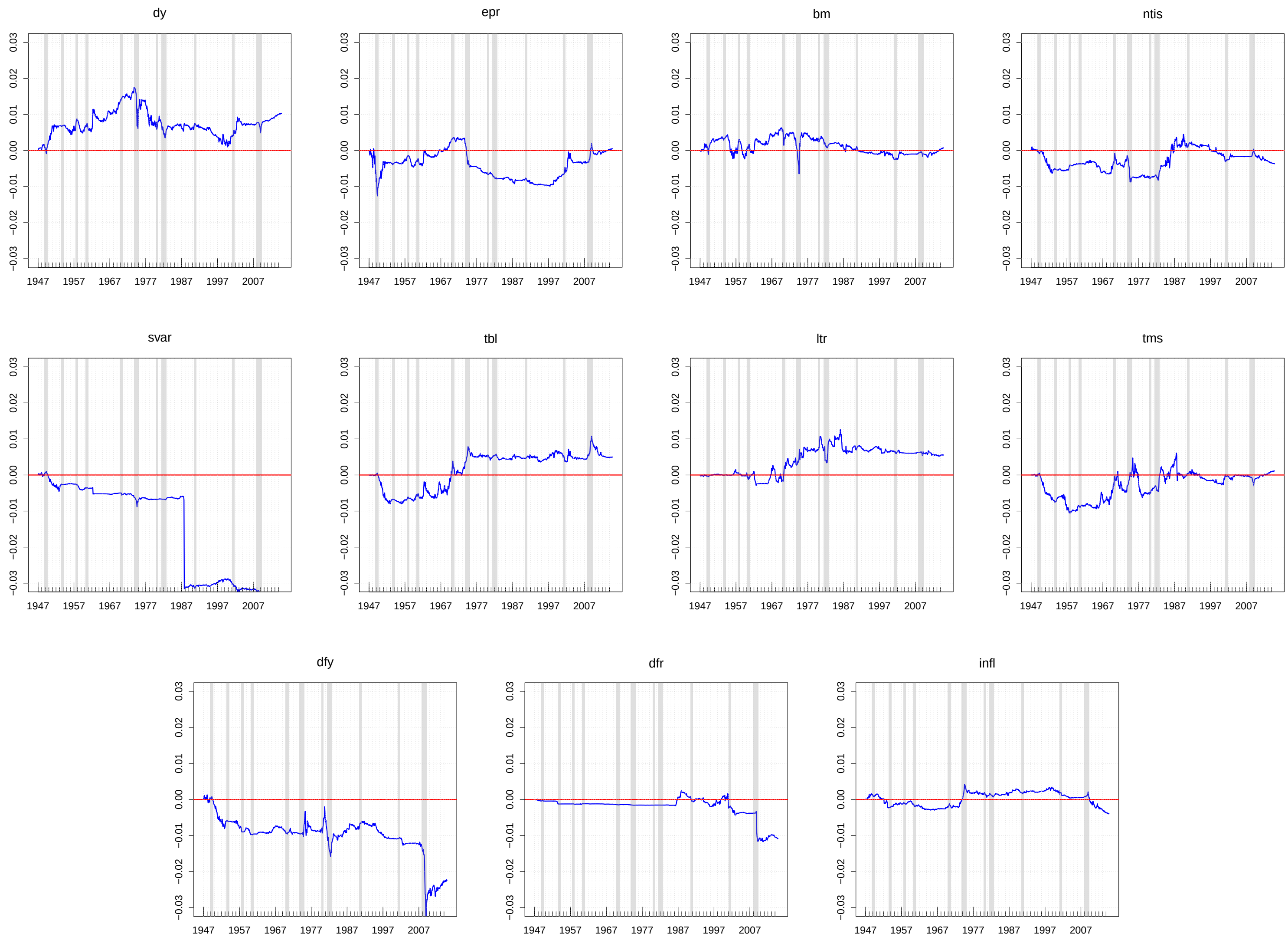

Figure 2. Equity Premium Prediction – individual predictors

The figure displays the performance of empirical models conditioning on individual predictors in out-of-sample equity premium prediction. Each graph plots the difference of the cumulative squared error of the null (historical mean) minus the cumulative squared error of the alternative. All forecasts are constrained to be positive and all slopes are bounded by the sign constraints. The grey areas indicate NBER recessions. The out-of-sample monthly forecasts are obtained with rolling regressions using a window of 20 years for the sample period of January 1927 to December 2014.

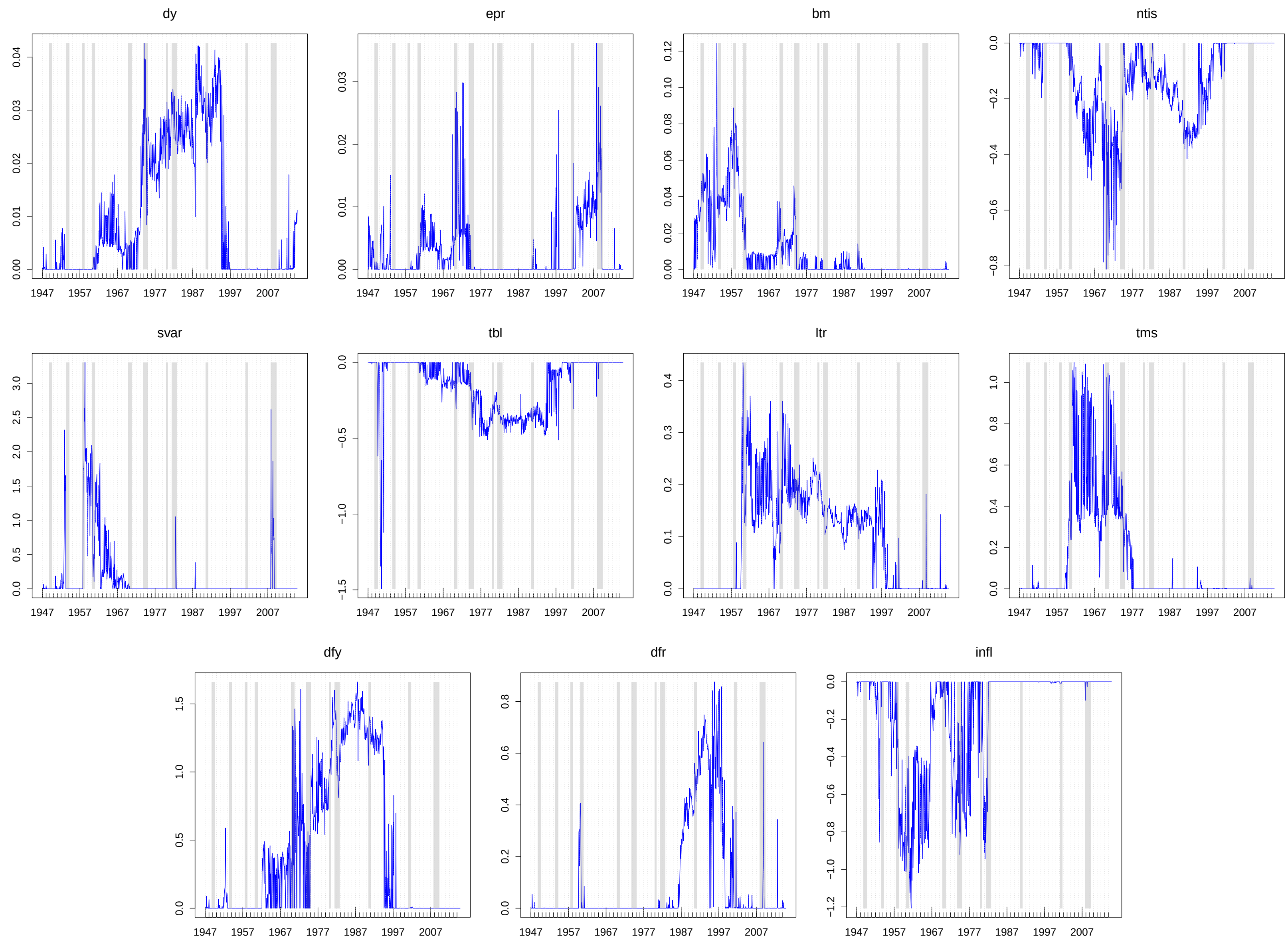

Figure 3. Equity Premium Prediction – out-of-sample slope estimates

The figure displays the out-of-sample slope estimates for each equity premium predictor used by the elastic-net kitchen-sink regression. The slopes are constrained to have the sign implied by economic theory. This is based on rolling regressions using a window of 20 years for the sample period of January 1927 to December 2014. The grey areas indicate NBER recessions.