Predictable Downturns

Carter Davis*

June 28, 2018

Abstract

Eugene Fama stated in his Nobel Prize lecture that "there is no statistically reliable evidence that expected stock returns are sometimes negative" (2013). However, various theoretical models such as Barberis et al. (2015) and Barlevy and Veronesi (2003) imply that expected stock returns are sometimes negative. This paper provides evidence that expected excess aggregate stock market returns are sometimes negative, and that portfolios composed of the most liquid stocks have predictable downturns as well. This paper presents a forecasting model that relies exclusively on ex-ante information to predict stock market downturns only when the day-prior confidence of a downturn is relatively high, and shows that the average excess return on days which are predicted to be downturns by the forecasting model is -13.9 basis points. Volatility and classic factor return variables alone are sufficient to predict downturns in the sample and are the most powerful downturn predictors. A market timing portfolio using these ex-ante predictions generates a risk-adjusted return of 3.5 basis points per day, annualized to an average 8.8% risk-adjusted return.

*University of Chicago Booth School of Business. I want to thank Lubos Pastor, Stefan Nagel, Eugene Fama, Dacheng Xiu, Pietro Veronesi, Zhiguo He, and Eric Zwick for their comments and guidance. This is a preliminary draft with early results. All remaining errors are my own. Please direct all questions or comment to carterdavis@chicagobooth.edu.

There is one remaining result in the literature on return predictability that warrants mention. The available evidence says that stock returns are somewhat predictable from dividend yields and interest rates, but there is no statistically reliable evidence that expected stock returns are sometimes negative. Fama and French (1987) find that predictions from dividend yields of negative returns for market portfolios of U.S stocks are never more than two standard errors below zero. Fama and Schwert (1977) find no evidence of reliable predictions of negative market returns when the forecast variable is the short-term bill rate.

Eugene F. Fama (2013), Nobel Prize Lecture

Although there is a large literature that predicts returns in the time series, there is surprisingly little focus on predicting *negative* excess stock returns. This is despite the fact that researchers continue to produce behavioral and rational asset pricing models with the primary prediction that expected stock returns are sometimes negative. For example, some behavioral models (e.g. Cutler and Poterba (1990), Barberis et al. (2016), and Barberis et al. (2015)) have extrapolative agents that generate periods of negative expected excess returns. Barlevy and Veronesi (2003) present a rational model with uninformed and informed agents which also generates negative expected returns. In this model, the uninformed agents observe price drops and correctly realize there is negative future cash flows news. They therefore reduce demand, leading to a further price drop. However, as referenced in Eugene Fama's quote above, there remains very little evidence of negative expected excess aggregate stock market returns in the empirical literature.

The very large predictability literature has typically focused on statistical significance of predictors in predictive regressions and marginal R^2 (e.g. Martin Lettau and Sydney Ludvigson (2001)), simple correlations (e.g. Owen A. Lamont and Jeremy C. Stein (2004)), or root mean squared error (e.g. Welch and Goyal (2008)). These methods of course do not answer the question of whether there exists predictable downturns.

In this paper, I employ a simple two-step process to evaluate whether expected returns are sometimes negative as follows: 1) estimate a forecasting model on a rolling window time period to generate rolling out-of-sample predictions of negative or non-negative excess returns and 2) test whether the average return of all days with negative ex-ante predictions is actually statistically less than zero. I show that the average return, conditional on a negative day-prior prediction, is strongly statistically negative.

Many finance practitioners and academics hold the view that predicting downturns is a losing endeavor that "diminishes long-term returns" because it causes investors to "lose out in rising markets."[1] Thus the simplest test for sensitivity (i.e. whether a model correctly predicts downturns) and specificity (i.e. whether a model incorrectly predicts upturns as downturns) is to simply take the average excess return over all the days that were predicted to be downturns. If the average is positive (negative), then pulling out of the market based on the forecasting model hurts (boosts) returns relative to the market.

To avoid making false inference due to stale stock prices at the daily level, I predict downturns in the value weight portfolio of the largest 10 stocks (selected at the end of the year for the following year). I show that the daily average excess return across all days with ex-ante negative excess return predictions is -13.9 basis points (bps) with a standard error of 1.5 bps.

I show a battery of robustness checks. I show that this result is robust to changing the number of stocks in the portfolio. I also show that there are predictable downturns in the Center for Research in Security Prices (CRSP) value weight market excess return. I decompose the return on the days with negative predictions into intraday and overnight returns according to Dong et al. (2017), and show that both intraday and overnight returns are statistically negative and economically large. I also show that there are still days with expected negative excess returns even in the last few decades as well as earlier periods.

Whether or not predictable downturns exist in the market is a very different question than whether downturns are predictable. For example, the latter question is concerned with both correctly predicting upturns (true positives) and downturns (true negatives), and both incorrectly predicting positive excess returns (false positives) and incorrectly predicting downturns (false negatives). However, showing the existence of negative expected stock returns (predictable downturns) is concerned only with a predictor that is accurate on average conditional on a negative prediction. In other words, the question of the existence of predictable downturns is concerned only with true negatives and false negatives, but is not concerned with true positives or false positives. Figure 3 illustrates this point.

I also examine which predictor variables are the most powerful downturn forecasting variables. I find that volatility and classic asset pricing factor variables are both enough to predict downturns and the most powerful predictors. I also find most of the other variables in the model help very little or hurt the ability to predict downturns in the main model on av-

[1] Goodbye to Market Timing, February 6 2012, Wall Street Journal

erage, at least at the end of the sample period where the predictive strength of the variables can be effectively compared.

I also construct a market timing portfolio based on the ex-ante predictions, which either invests completely in the value weight portfolio of the largest 10 stocks or completely in the risk-free asset, depending on the downturn forecast. This portfolio, which never has a short position, yields a risk-adjusted return (alpha) of about 3.5 bps per day.

I show that given a very simple mean-reverting process of returns, that the probability of predicting downturns in the aggregate market over periods longer than a couple of weeks should be zero given the estimated parameters. Consistent with this, the forecasting model's ability to predict downturns steadily declines over longer periods, and fails at periods longer than two weeks.

The rest of the paper is as follows. In section I, I give a brief review of the related literature and discuss this paper's contribution. In section II, I discuss the empirical strategy I employ to test the hypothesis of negative expected returns. In section III, I present the data I use. In section IV I present the main results. In section V, I evaluate how the market timing portfolio, based the ex-ante downturn prediction, performs. In section VI, I examine how the results vary over longer horizons than daily returns. Finally, in section VII, I conclude.

I Literature Review

As discussed above, there are is a multitude of theoretical models that have negative expected returns. For example, Barberis et al. (2015) present a behavioral model with extrapolative agents that generate negative expected returns, while Barlevy and Veronesi (2003) show that rational but uninformed investors can also generate negative expected returns.

Although this is the first paper that gives evidence of negative expected returns in the aggregate stock market, there are other papers that give evidence of negative expected returns in other assets. For example, Greenwood and Hanson (2013) give evidence of predictable negative returns on high yield debt, and Baron and Wei (2017) predict negative returns on banking sector equity.

This paper is connected to the very large literature on market return predictability. For example, aggregate short interest (Owen A. Lamont and Jeremy C. Stein (2004)), the consumption / wealth ratio (Martin Lettau and Sydney Ludvigson (2001)) new stock issuance

(Malcolm Baker and Jeffrey Wurgler (2000)), aggregate dividend yields (Fama and French (1988)), stock volatility (French et al. (1987)) all predict variation in returns in the time series. Welch and Goyal (2008) argued that historical averages were actually better for forecasting aggregate market returns than out-of-sample regressions. However, Campbell and Thompson (2008) showed that by using economically motivated sign restrictions on estimates to decrease estimation noise, out-of-sample regressions were better at forecasting aggregate returns than historical averages. Importantly, the key tests of these papers is whether the predictive variables statistically predict *variation* in returns and not whether they predict *negative* returns.

Greenwood et al. (2017) seek to address, in challenge a statement from Eugene Fama, the question of whether downturns are predictable. They try to predict downturns at the industry level at a monthly frequency, instead of the daily frequency at the market level like this paper. In their paper, they use a variety of predictors in different models, but with most of their predictors, the strategy of shorting the industry when it is predicted to decline usually results in a loss. In other words, according to their different models, when an industry is predicted to have a decline, it tends to go up more than down at most horizons across many of their models. They do create portfolios that switch in and out of industries and into the market according to decline predictions, which can beat the market. However, they ultimately do not have a parsimonious model that shows that the expected industry returns are sometimes negative.

It is important to note that I am not claiming that my method predicts bubbles. Fama (2013) defines a bubble to be "an irrational strong price increase that implies a predictable strong decline." I give substantial evidence that there are predictable declines at the aggregate level. However, I do not show that these predictable declines (or the preceding run-ups) are driven by irrationality, which is of course more difficult to show.

II Empirical Strategy

I first write down the hypothesis in formal terms, and then I discuss the hypothesis, assumptions, and possible misconceptions. Then I discuss the prediction method that I employ to generate ex-ante downturn predictions.

A. Formal Hypothesis Development

In this subsection, I show formally that if the average over all the days that are predicted to be negative by some forecasting model is actually negative statistically, where the forecasting model uses only ex-ante information, then there is evidence of periods with negative expected returns.

Let r_t be the value weight market return in excess of the risk-free rate. Define

$$z_t = \mathbb{E}_t[r_{t+1}] = \mathbb{E}[r_{t+1}|\mathcal{F}_t] \tag{1}$$

where each $\mathcal{F}_t$ contains all *public* information about the stock market at the market close on day t.

If it is observed that $z_t < 0$, then I can conclude that expected stock returns are sometimes negative. However, since z_t is of course not observed, I use some prediction of r_{t+1} using only information at time t, denoted as $\hat{r}_{t+1|t}$. Define $\epsilon_{t+1} = r_{t+1} - z_t$. Also, let $\hat{\pi}_{t+1|t} = \widehat{\text{Prob}_t}(\hat{r}_{t+1|t} < 0)$ be some ex-ante prediction of the probability of a downturn. Thus if $\hat{\pi}_{t+1|t}$ is close to 1 (0), then the ex-ante probability of r_{t+1} being below zero is high (low). I refer to $\hat{\pi}_{t+1|t}$ as the negative probability prediction. If $\hat{\pi}_{t+1|t} > c$ for some probability cutoff c, then I refer to this as a negative prediction. Formally, assume $\hat{\pi}_{t+1|t}$ is a random variable on the probability space associated with $\mathcal{F}_t$, which I denote, as an abuse of notation, as $\hat{\pi}_{t+1|t} \in \mathcal{F}_t$. This leads to the very simple but important proposition below:

Proposition: If $\hat{\pi}_{t+1|t} \in \mathcal{F}_t$ and for any $c \in (0, 1)$, $\mathbb{E}[r_{t+1}|\hat{\pi}_{t+1|t} > c] < 0$, then $z_t < 0$.

The following proof is quite simple. By way of contradiction, assume $z_t \geq 0$, but $\hat{\pi}_{t+1|t} \in \mathcal{F}_t$ and for any $c \in (0, 1)$, $\mathbb{E}[r_{t+1}|\hat{\pi}_{t+1|t} > c] < 0$. Then:

$$\mathbb{E}[r_{t+1}|\hat{\pi}_{t+1|t} > c] = \mathbb{E}[z_t|\hat{\pi}_{t+1|t} > c] + \mathbb{E}[\epsilon_{t+1}|\hat{\pi}_{t+1|t} > c] \geq 0 \tag{2}$$

since $\hat{\pi}_{t+1|t} \in \mathcal{F}_t$ implies that $\mathbb{E}[\epsilon_{t+1}|\hat{\pi}_{t+1|t} > c] = 0$ and $z_t \geq 0$ implies $\mathbb{E}[z_t|\hat{\pi}_{t+1|t} > c] \geq 0$.

In words, this proposition states that if the negative probability prediction contains only information available at time t ($\hat{\pi}_{t+1|t} \in \mathcal{F}_t$) and the expected returns conditional on a forecasted negative return is actually negative ($\mathbb{E}[\epsilon_{t+1}|\hat{\pi}_{t+1|t} > c] < 0$), then the expected return is negative ($\mathbb{E}_t[r_{t+1}] = z_t < 0$). Thus the key to the paper is to show that a prediction model that uses ex-ante information can actually predict downturns. If this model is successful, in a formal statistical sense, then this proposition states that expected returns are sometimes negative.

Under standard ergodicity assumptions, the average return over all days where $\hat{\pi}_{t+1|t} > c$, denoted as $\hat{\gamma}^- = A(r_{t+1}|\hat{\pi}_{t+1|t} > c)$ converges in probability to the expectation of $\mathbb{E}[r_{t+1}|\hat{\pi}_{t+1|t} > c]$ over the distribution of $\hat{\pi}_{t+1|t}$, which I denote as γ^-. Thus in this case, rejecting the following null hypothesis:

$$H_0 : \hat{\gamma}^- = A(r_{t+1}|\hat{\pi}_{t+1|t} > c) \geq 0 \tag{3}$$

provides evidence that expected excess market returns are sometimes negative. This is the key hypothesis tested in this paper. Thus, for this key hypothesis, a type I error is false rejection of the null hypothesis that expected returns are always non-negative. A type II error is of course retaining the false null hypothesis of always non-negative expected returns.

The average $\hat{\gamma}^-$ can be obtained by estimating the regression

$$r_{t+1} = \gamma^+ I(\hat{\pi}_{t+1|t} \leq c) + \gamma^- I(\hat{\pi}_{t+1|t} > c) + e_{t+1} \tag{4}$$

where $I(\cdot)$ is simply the indicator function. Of course, from this regression, $\hat{\gamma}^+$ is simply the estimated mean of returns on days with an estimated ex-ante probability of being positive greater than or equal to $1 - c$. Estimating γ^- with the regression or just by a simple average of course yields identical numerical estimates, but using the regressions allows me to easily vary the way the standard error is calculated to ensure the results are robust.

B. Hypothesis Development Discussion

I highlight the three main insights below from this proposition.

The first insight from the proposition above is that although there could potentially be large differences between the actual conditional expectation of returns, z_t, and the return forecast $\hat{r}_{t+1|t}$, a statistically negative average return on days with forecasted downturns provides evidence that the conditional expectation is sometimes negative. One possible misconception about the empirical strategy of this paper is that a negative estimated average over days with forecasted downturns does not actually inform researchers that the expected return is sometimes negative if the forecasting model differs from the true expected return by a large amount. This proposition shows that even if $|\hat{r}_{t+1|t} - z_t|$ is large, but $\mathbb{E}[r_{t+1}|\hat{\pi}_{t+1|t} > c]$ is negative, then the conditional expected return z_t is negative.

For example, a forecasting model that underfits or overfits increases the average $|\hat{r}_{t+1|t} - z_t|$, but does not actually increase the type I error rate (false rejection of always non-negative expected returns). Thus if a forecaster accurately predicts downturns, the potential underfitting or overfitting does not in any way diminish the validity of the results that there are negative expected stock returns. In fact, underfitting and overfitting both increase the average $|\hat{r}_{t+1|t} - z_t|$, thus increasing the type II error rate (making it more difficult to find evidence of negative expected returns).

The second insight is that for *any* ex-ante negative return probability forecaster $\hat{\pi}_{t+1|t}$ and for *any* cutoff c, if the forecasting model predicts negative returns ($\mathbb{E}[r_{t+1}|\hat{\pi}_{t+1|t}] < 0$), then it must be the case that the expected return is negative ($\mathbb{E}_t[r_{t+1}] = z_t < 0$). Thus it takes only a single predictor and a single cutoff value c to find evidence that expected returns are sometimes negative. However, if multiple different cutoff values c and different forecasting models $\hat{\pi}_{t+1|t}$ are considered, then this becomes a classic multiple comparisons problem. By using Bonferroni's correction for multiple comparisons, if even a single model with a single cutoff value c finds statistically significant evidence of predicting negative returns, then this provides evidence of negative expected stock returns.

In practice, as I describe below, I picked one main forecasting model with a single cutoff probability of 55% ($c = 0.55$). However, I also calculate with Bonferroni's correction the number of other hypotheses that would need to be tested in order to lose statistical significance of the main model with this cutoff value. I find that the number of other hypotheses that would need to be tested in order for the main model to lose statistical significance is far too large for any researcher to have feasibly tested.

If many forecasting models are selected, and the one with the most statistically significant results are presented without a multiple comparisons correction, then surely the type I error is much higher than reported. Thus model selection, and not model complexity, matters for determining the validity of the results. Thus I outline in the next subsection transparently how I selected the forecasting model. In the rest of the paper, I show that the results are robust through a variety of ways:

1. All predictor variables are known at time t.

2. I compute the number of other predictive models that would need to be tested in order

to lose statistical significance at the 99.9% confidence level, using Bonferroni's correction. I find that the number of other models that would need to be tested in order to lose statistical significance is large enough that it would have been unfeasible for any researcher to have possibly tested.

3. Forecasting model hyperparameters are robust to variation.

4. Alternative forecasting models are used and the results are robust.

5. All downloaded variables are used in the model, even though I show that simple models with few predictive variables can also predict negative returns. In other words, the predictive variables are not selected based on ex-post information.

6. I estimate the conditional averages in different periods.

7. I decompose the returns into intraday and overnight returns, in order to ensure that the intraday component is statistically negative.

It is important to note that there is a trade-off between a high and relatively low cutoff probability c. As shown in Figure 3, a negative prediction, defined as $\hat{\pi}_{t+1|t} > c$, can either result in a true negative (return is actually negative) or a false negative (return is not negative). A higher value of c (more conservative negative predictions) should increase the true negative rate and decrease the false negative rate, resulting in a more negative γ^-. This channel therefore decreases the type II error rate as c increases. However, as c increases, the sample size of the estimate $\hat{\gamma}^- = A(r_{t+1}|\hat{\pi}_{t+1|t} > c)$ decreases, which works to increase the type II error rate. Thus being relatively conservative about negative predictions ($c > 0.5$) while being careful so c is not so high as to decrease the sample size substantially gives the best chance of finding evidence that expected returns are sometimes negative.

However, if c is chosen ex-post in order to find evidence of negative expected returns without correcting for testing multiple hypotheses (different values of c), then of course the type I error rate is higher than stated. Thus in practice, I chose $c = 0.55$ (at least a predicted 55% chance of being negative), but I show that the results are robust to varying c over a relatively large range.

C. Forecasting Model Selection

In this subsection, I explain the out-of-sample forecasting model that generates the return predictions $\hat{r}_{t+1|t}$ and the predicted probability of a negative excess return $\hat{\pi}_{t+1|t}$. Denote X_t as a large vector of variables known at time t. I seek to forecast r_{t+1} with X_t using some function $g_t(X_t)$. In other words, with an estimated function $\hat{g}_t(X_t)$ I have a forecast

$$\hat{r}_{t+1|t} = \hat{g}_t(X_t) \tag{5}$$

To estimate $g_t(X_t)$, I use the rolling window approach shown in Figure 2. In particular, at the end of the first trading day of each month, I fit $g_t(X_t)$ using the previous 10 years of daily trading data. This fitted function g_t is used to predict returns for the entire month, once the information becomes available. The next month, the process is repeated. Each day, at the end of the trading day, X_t is available, and the forecast for the next day is generated as $\hat{r}_{t+1|t} = \hat{g}_t(X_t)$. This creates an out-of-sample time series of forecasts, that I then subsequently use to evaluate downturn predictions. For example, consider forecasting the January 7, 1983 ($t+1$) aggregate excess market return. The method uses data from January 4, 1973 through January 3, 1983 (the first trading day of the month) to fit g_t. Then using January 6, 1983 (t) information X_t, plugged into g_t, I forecast January 7, 1983 ($t+1$) aggregate excess market returns with $\hat{r}_{t+1|t} = \hat{g}_t(X_t)$.

I use a random forest regression with 100 regression trees as the main forecasting model function g_t. I chose to use a random forest because it is flexible, does not assume linearity, and is easy to understand and interpret (Breiman (2001)). I also use linear regression as a robustness check. I use a host of predictive variables X_t, discussed in the data section below. Finally, as discussed above, I predict negative returns when the ex-ante negative return probability is at least 55% ($c = 0.55$), and I show how the results vary as this cutoff varies. I refer to this model, with the $c = 0.55$ cutoff as the main model throughout the paper.

An obvious concern is that random forests will overfit the training data, especially with a large amount of predictive variables. However, as discussed above, overfitting will only increase $|\hat{r}_{t+1|t} - z_t|$ but will not increase the type I error rate. In other words, overfitting will not increase the chance of finding evidence of negative expected returns.

In order to compute $\hat{\pi}_{t+1|t}$, I assume

$$r_{t+1} \sim N(\hat{r}_{t+1|t}, \sigma_t^2) \tag{6}$$

I estimate σ_t^2 as

$$\hat{\sigma}_t^2 = \widehat{\mathrm{Var}}(r_{\tau+1} - \hat{r}^I_{\tau+1|\tau}) \tag{7}$$

where $\hat{r}^I_{\tau+1|\tau}$ is the in-sample out-of-bag return predictions and of course $\widehat{\mathrm{Var}}(\cdot)$ is the typical estimator of variance across the 10-year sample. I use this estimate to compute

$$\hat{\pi}_{t+1|t} = \Phi\left(-\frac{\hat{r}_{t+1|t}}{\hat{\sigma}_t}\right) \tag{8}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

This ex-ante estimate has the property that

$$\hat{\pi}_{t+1|t} > 0.5 \text{ iff } \hat{r}_{t+1|t} < 0 \tag{9}$$

While one common way to estimate the distribution of the target variables with a random forest is to use the individual regression trees to generate an empirical distribution, I eschew this method for the one described above because the estimate of $\hat{\pi}_{t+1|t}$ has this property, while the empirical distribution does not.

I also fit a simple linear model that uses the same predictive variables as the random forest model, and also the same 10 year rolling window approach fitted monthly. The linear model, like the main model, incorporates the predictive variables into the model as they become available in the time series. I calculate the ex-ante negative return probability in two ways. First, I use the same normality assumption as above, but instead of using $\hat{r}^I_{\tau+1|\tau}$, I use the in-sample linear fit of the model. The second way I use to compute the ex-ante negative return probability prediction is to calculate the prediction confidence of a negative return based on the standard ordinary least squares (OLS) prediction interval. That is, the standard $1-\alpha$ prediction confidence interval for a given estimate $\hat{r}_{t+1|t}$ is

$$[\hat{r}_{t+1|t} - t_d(1-\alpha/2)\hat{\sigma}_{p,t},\ \hat{r}_{t+1|t} + t_d(1-\alpha/2)\hat{\sigma}_{p,t}] \tag{10}$$

where $t_d(1-\alpha/2)$ is the cumulative density function of the t distribution with d degrees of freedom evaluated at $1-\alpha/2$, and $\sigma_{p,t}$ is the standard deviation of the prediction (which of course is larger than the standard error of the estimate). The degrees of freedom equals the number of observations minus the number of covariates. From this, a natural negative return probability forecast is

$$t_d\left(-\frac{\hat{r}_{t+1|t}}{\hat{\sigma}_{p,t}}\right) \tag{11}$$

In practice, unsurprisingly, these two methods yield very similar results. I simply use both in order to show the typical OLS prediction confidence is similar to just the method that assumes normality described above.

III Data

I use a host of predictive variables. In particular, I downloaded a total of 126 variables from four sources: Chicago Board Options Exchange (CBOE) data from Wharton Research Data Services (WRDS), Center for Research in Security Prices (CRSP), Federal Reserve Economic Data (FRED), and Ken French's website[2]. I also use the month and weekday indicators as predictor variables. I use only day t information to predict $r_{i,t+1}$. The earliest data available is from July 1926 through the end of July 2017. Although there are 126 variables, since I also month and weekday indicators, there are 143 variables total.

I group these variables into six different groups: CBOE, CRSP, FRED, French portfolios, French industry portfolios, and calendar indicator variables. The CBOE group consists of only 4 variables: the VIX, VXO, VXD, and VXN. The CRSP group consists of 19 predictor variables that include market capitalization decile return indices and various other stock market variables, as well as 9 different portfolios of the largest N stocks (with 9 different values of N as described below) and their associated overnight and intraday return. Thus there are 46 total CRSP variables. There are 12 variables downloaded from FRED, such as the Bank of America high yield spread and federal funds rate. I use 15 classic asset pricing factor variables, such as the risk-free rate, excess market returns, Small Minus Big (SMB) and High Minus Low (HML) returns from Ken French's website. I refer to these as French portfolio variables. There are 49 French stock market industry portfolios. Finally, there are 17 calendar indicator variables (12 months and 5 weekdays).

Table 1 shows summary statistics of these variables, the group of the variables, and a de-

[2]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

scription of each variable. It also gives the date that each variable begins. Note that many of the variables are not available until later in the sample period. This is an important point. Instead of restricting my sample period to the intersection of when all the variables are available, I simply incorporate new predictors when the data become available, similar to Welch and Goyal (2008). The predictors are generated at the end of the trading day or before. Also, none of the predictor data has been modified after the fact, except CRSP which has done historic return backfilling. This backfilling of course is not an issue, since the information being backfilled was available at the time.

Although the portfolio returns are available in 1926, since I use a 10-year rolling window to predict future returns, the generated time series of $\hat{r}_{j,t+1}$ is not available until 1936. Thus all of the return averages on days with negative predictions are estimated using the sample period from July 1936 through July 2017.

The objective of this paper is to test whether the aggregate stock market has periods of negative expected excess returns. However, as described in Ahn et al. (2002), stale stock prices induce positive autocorrelation in portfolio returns. This illiquidity creates an illusion of high potential returns due to this predictability by trading the component stocks. Thus a forecasting model that predicts downturns in the CRSP market portfolio excess returns may actually just be due to stale prices. Thus, in order to avoid this, for the entirety of the paper, the key variable predicted is the excess return of the value weight portfolio of the largest 10 stocks chosen at the end of the previous year. To construct this portfolio, I simply take all common shares of CRSP, and at the end of each calendar year, the largest 10 stocks by market capitalization are chosen to be in the portfolio. As usual, I subtract the treasury bill rate from CRSP as the risk-free rate to create a daily excess return. I refer to this excess return as simply the big 10 portfolio return. I use this as the key portfolio return in the paper in order to alleviate the concerns that any predictability is due to the illiquidity of smaller stocks in the CRSP value weight market index. As a robustness check, I also create other portfolios in a similar way but with the largest N stocks, where N is either 50, 100, 200, 300, 400, 500, 1,000, or 2,000. I show that there is evidence of periods of negative expected returns in each of these portfolios, as well as the CRSP value weight market index. In Table 2 I show the correlations of these excess portfolio returns. Note that the correlation of the excess big 10 portfolio return and the CRSP value weight market return is 93.4%. Table 3 shows the Capital Asset Pricing Model (CAPM) alphas and betas of these same portfolios, using the excess return on the CRSP value weight market return on the right-hand side of the regression. Note that the CAPM betas are all close to 1, and the CAPM beta of the big 10 portfolio in particular is 0.99.

IV Results

A. Main Estimates

With the main prediction model explained above, 21.2% of trading days in the sample are predicted to be have negative excess returns. Figure 4 shows the number of trading days in each calendar year in the sample that are predicted to be negative. As this plot shows, between 1% to nearly 50% of trading days in any given year are predicted to be negative. Figure 5 shows the total number of downturn predictions by weekday, and Figure 6 shows the total number of downturn predictions by month.

Although this paper is focused on the mean of excess returns on days that are predicted to be negative, Figure 7 shows how the the entire distribution of excess returns on days with downturn predictions differs from the distribution of returns on days that lack a downturn prediction. The beige and purple colors together represent the histogram of returns in the sample which had ex-ante downturn predictions. The purple and blue colors represent the histogram of excess returns on days which did not have downturns predictions. From this plot, although most of the plot is overlap between the two empirical distributions, the distribution of returns with negative predictions is slightly to the left of the distribution of returns that lack negative predictions.

The average excess big 10 return conditional on a negative prediction is -13.9 bps, which is the key estimate of this paper. This result is shown in the first column of Table 4. The standard error is 1.6 bps even when using Newey and West (1987) standard errors with 3 lags. Table 5 shows that the standard error estimates of this average using the typical OLS estimate is 15 bps, Newey West standard error estimates calculated using 2, 3, 5, and 21 (one month of trading days) lags are all 1.6 bps. I use Newey West standard errors with 3 lags in order to be relatively conservative throughout the paper, but the difference in standard errors using typical OLS standard errors or Newey West standard errors with various lags is always small.

As discussed above, under the null of always non-negative expected returns, $\hat{\gamma}^-$ is greater than or equal to zero. The estimated t-statistic is -8.57, which means that a one-sided test gives a rejection of this null hypothesis with a p-value of $5.33(10)^{-18}$. A natural worry about forecasting estimates is that the results are just a result of data drudging. That is, it could perhaps be the case that a researcher runs many different forecasting regressions, and reports only the results are the statistically significant, without taking into account the problem of multiple comparisons. However, I would need to estimate approximately 187 trillion $(0.001/(5.33(10)^{-18}))$ models that use only ex-ante information in order to lose statistical significance of this model at the 99.9% significance level under Bonferroni's correction for multiple comparisons. Bonferroni's correction is the most conservative correction for multiple comparisons, and using other more statistically powerful corrections results in an even larger number of statistical tests that would need to be conducted in order to lose statistical significance at the 99.9% confidence level. In other words, since the predictor variables contain only ex-ante information, a result with this p-value is highly unlikely to be data mined or occur by chance and gives strong evidence that expected excess stock returns are sometimes negative.

Since most of the predictive variables I use are available at market close, one obvious concern is that the predictable component is mostly the overnight return. If it is the case that most of the predictable movement occurs overnight, this would likely decrease or perhaps even eliminate potential profits made from trading on the strategy. In order to test this, I decompose the return on the days that are predicted to be negative into intraday and overnight returns following the same methodology as Dong et al. (2017). Column 3 of Table 4 shows that the average intraday return on days with negative predictions is an average -5 bps. This estimate is statistically significantly less than zero at the 99.9% level. Column 2 shows that the average overnight return on these days is -2.2 bps, and is also statistically significant at the 99.9% level. Note that these are not log returns, and the intraday and overnight returns are not in excess of the risk-free rate, thus the average intraday and overnight returns do not sum up the excess close-to-close big 10 portfolio return.

Another obvious concern is that perhaps this predictability is because of early sample issues, and is eliminated in the later end of the sample. In order to test this, I take the average return conditional on a negative prediction from the beginning of of 1980 through the end of the sample. Column 4 of Table 4 shows that this average return is -10.5 bps. This is statistically significant at the 99.9% confidence level. Column 6 shows the average intraday return on these days is -7.6 bps, and also statistically significant at the 99.9% confidence level. Interestingly, the point estimate of the overnight return is 0.5 bps, and is not statistically different from zero.

Although stale prices should likely make portfolios with less liquid stocks more predictable as described in Ahn et al. (2002), I test whether the portfolios with more stocks, discussed in the data section above, exhibit periods of negative expected stock returns. In Table 6, I show the average excess returns of different portfolios conditional on a negative return prediction. In each case, the random forest is fit to the portfolio excess return every month, just as described above. As in the main model, I use a 55% ex-ante negative return probability cutoff. Note that all of these portfolios have statistically significant negative return averages conditional on downturn predictions, with p-values well below 0.001. Column 10 of this table shows the average excess return of the CRSP value weight market index, conditional on a negative return prediction.

I also estimate the average return on days with negative predictions by decade. Figure 8 shows these results, with the black curve showing the estimates, and the grey region showing the 95% confidence interval. The point estimates are negative for every decade, but the estimates for the 1940's, 1980's, and 1990's are not statistically significant at the 95% confidence level. However, as Table 4 shows, the average return on days with negative predictions on all days from the 1980's through the end of the sample is negative and statistically significant at the 99% confidence level.

I also show that the average return on days with negative predictions is statistically significant at the 95% confidence level for a wide range of cutoff probabilities c. Figure 9 plots the average return on days with negative predictions across the entire sample, as a function of the cutoff probability c. The dark grey region represents the 95% confidence interval. Note that for cutoff probabilities anywhere in the region between 46% - 83%, the result is statistically significant at the 95% confidence level.

It is important to note from the proposition above that statistical significance with even a single cutoff probability c, given that there is an appropriate correction for multiple hypotheses being tested if multiple cutoff values are used, provides evidence that expected returns are sometimes negative. In the discussion above of multiple hypotheses and Bonferroni's correction, it is clear that it is infeasible for any researcher to have possibly tested enough hypotheses to have lost statistical significance of the main model I use in this paper.

Column 5 of Table 7 shows that the average returns conditional on a downturn prediction of a linear model is -8.4 bps, which is also statistically less than zero at the 99.9% confidence level. This model uses a cutoff probability of 55% and a negative return probability prediction using the normality assumption described above. Column 4 shows the the result is similar using the OLS prediction probability described above. Column 3 shows that the random forest model with a cutoff probability of 50%, which is the same as predicting a downturn precisely when $\hat{r}_{t+1|t} < 0$, has an average return on days with negative predictions of -5.5 bps. This is also statistically significant at the 99.9% confidence level. Columns 1 and 2 also show that the linear model with a cutoff probability of 50% yields an average -4.1 bps return on days with negative predictions. Column 1 uses the OLS prediction probability method, while column 2 uses the normality approach.

Linear regressions for prediction are classically known for underfitting the training data, while random forests can easily overfit training data. The fact that both types of models find statistically significant evidence of negative expected returns indicates that the results are not model dependent. Also, the random forest model yields a 65% higher average return in absolute value on days with negative predictions than the linear model, with the 55% probability cutoff. Thus, perhaps unsurprisingly, it appears that the linear model underfitting is more severe than the overfitting or perhaps even underfitting of the random forest model.

B. What are the Predictor Variables that Most Effectively Predict Downturns?

In order to understand what variables are the best downturn predictors, I take the average importances across all of the random forests that are fitted monthly. An importance of 0 means that the predictor variable explained 0% of the variation in the target variable, while an importance of 100 means that the variable explained 100% of the variation in the target variable that is explained by all predictor variables. Since not all predictor variables are available at the beginning of the sample period but are incorporated into the random forest models when they are available, the average importance for each predictor variable is just the average over all the random forest models which use the predictor. Table 8 shows the ten predictors with the largest importances. The volatility indices VXO, VIX, and VXN are among the most important, and the precious metals stock index and petroleum and natural gas stock index rank in the top 10 as well. As the table shows, the volatility indices correlate positively with next day returns, while the precious metals and petroleum and natural gas indices correlate negatively with next day returns.

In order to test which groups of variables were most influential, I compared the various predictive abilities of the main model using all combinations of the six groups of predictive variables described in the data section above. There are 63 ($2^6 - 1 = 63$) such combinations, so I generated 63 different times series of predictions using every different combination of these groups of predictor variables. I then estimate the 63 different average returns on days with downturn predictions from the various models. I do this exercise in order to compare the predictive power of the groups of variables, but the CBOE variables in particular are not available until later in the sample. For example, the first CBOE variables that are available are the VXO in 1986 and the VIX in 1990. In order to compare the groups during a period where the data is available, I estimate the average from 2000 onward, since this is 10 years after the VIX creation (because the model uses a 10 year rolling window prediction method). Like the main model, I use a random forest model with a 55% negative return probability cutoff.

Some of the notable estimation results, from among these 63 estimates, are in Table 9. Since there are 63 different tests, I use Bonferroni's correction in order to assess statistical significance. In order to account for these multiple comparisons, this table uses significance asterisks differently than the rest of the tables in the paper. In this table, three asterisks next to an average return conditional on a downturn prediction (γ^-) estimate signifies that the estimate is statistically less than zero at the 95% confidence level using a one sided test with Bonferroni's correction (it has a p-value less than $0.05/63 \approx 0.00079$). The asterisks on the γ^+ estimates are similar, but use a one-sided test with a null hypothesis that the average on these days is equal to or less than zero. Two asterisks means that the p-value is less than 0.01, and one asterisk signifies that the p-value is less than 0.05.

Although there are 63 tests estimated over this short sample period and I use Bonferroni's correction (which is well-known to reduce the power of the individual tests), there are still 3 models that reject the null hypothesis of always non-negative expected returns at the 95% significance level. Columns 2, 3, and 4 of Table 9 show these estimates. Column 2 in particular shows that the model with only CBOE and French portfolio variables are enough to reject the null hypothesis of always non-negative expected returns at the 95% confidence level even after Bonferroni's correction.

Note that the estimate using the model with only CBOE variables to predict downturns would be statistically significant at the 99% confidence level without the multiple comparisons correction, as shown in Column 1. However, the model that uses CBOE variables alone loses statistical significance at the 95% confidence level since 63 tests are simultaneously performed. Even with Bonferroni's correction, this estimate is still statistically significant at the 90% confidence level. Column 5 shows the full model results during this period with all six groups, and like the model with only CBOE variables, γ^- has a p-value is less than 0.01 but loses statistical significance at the 95% confidence level when Bonferroni's correction is applied.

In all three cases where the models reject the null hypothesis of always non-negative expected returns even after Bonferroni's correction (Columns 2, 3, and 4 of Table 9), the CBOE and French portfolio groups are included in the models, and the French industry portfolios and CRSP variables are not included. In fact, the full model loses statistical significance with Bonferroni's correction when these are included. This seems to suggest, at least in this period, that the CBOE and French portfolio variables tend to help predict downturns, while including the CRSP and French industry portfolios tends to lead to overfitting and worse downturn predictions.

In order to investigate further which predictor groups tended to help or hinder the prediction during the sample, I compare models excluding and including different groups. More precisely, consider predictor variable group j, and consider the value $t_j - t_j^b$, where t_j is the estimated t-statistic on γ^- from some model that includes group j, and t_j^b is the estimated t-statistic from a model with the same predictors but excludes group j (baseline model). If $t_j - t_j^b$ is very negative, then adding the predictor group j to the baseline model helped predict downturns better statistically. If $t_j - t_j^b$ is close to zero, then adding group j to the baseline model did very little, and if $t_j - t_j^b$ is positive, then adding group j to the baseline model likely caused overfitting and actually eroded the quality of downturn predictions. Note that the t statistics can be compared, because the degrees of freedom are equivalent across models. For each j, there are of course 31 ($2^5 - 1$) different possible baseline model comparisons. I plot these t-statistic differences in Figure 10 for each of the six predictor groups.

Note that by this difference of t-statistic metric, the CBOE, FRED, and French portfolio variables tend to improve the downturn predictions, while the CRSP, calendar, and French industry portfolios tend to decrease or have little affect on the ability to predict downturns during this period on average. Note that the grey points represent estimates t_j that are not statistically significant at the 95% confidence level, and the yellow points have p-values between 0.05/63 (Bonferroni 95% confidence level cutoff) and 0.05. The red points represent t_j values negative enough be statistically significant at the 95% confidence level with Bonferroni's correction.

In summary, the only variables needed to predict downturns during this period are the CBOE volatility variables and classic asset pricing factor variables, referred to as French factor portfolio variables. Furthermore, it appears that these groups of variables, along with the FRED group helped to predict downturns during this period, while the calendar, CRSP and French industry portfolio variables tend to hurt or have little affect on the ability to successfully predict downturns.

V Market Timing Portfolio

I form a simple market timing portfolio using the results from the main model. The portfolio invests completely in the big 10 portfolio when there is no downturn prediction, and invests completely in the risk-free asset (treasury bill) when there is a negative return prediction.

Figure 11 shows the cumulative return of both this market timing portfolio, and the big 10 portfolio, on a log scale. The market timing portfolio results in a cumulative return that is over one hundred times larger at the end of the sample period than just investing in the big 10 portfolio.

This market timing portfolio has an unconditional market beta and alpha, using the excess return on the CRSP market index, of 0.73 and 3.5 bps respectively, as shown in column 1 of Table 10. This translates into an 8.8% annual alpha (252 * 0.035). Columns 2, 3, 4, and 5 correspond to, respectively, the Fama and French (1993) 3 factor model, Carhart (1997) 4 factor model, Fama and French (2015) 5 factor model, and a model with these same 5 factors with momentum. The alpha stays around 3.5 bps or slightly larger, and is statistically significant at the 99.9% confidence level in every model.

In order to see how this portfolio performs each year, I estimate the average difference of the market timing portfolio return and the big 10 portfolio return every year in the sample. Figure 12 plots these yearly differences, and the grey region represents the 95% confidence region. For most years, the market timing portfolio outperforms the big 10 portfolio, although the difference is rarely statistically significant, at least in part due to a small sample size for each estimate.

Perhaps a better benchmark for the market timing portfolio is a portfolio with similar market exposure on average. The big 10 portfolio has a CAPM beta, using the CRSP value weight market portfolio, of 0.99, as shown in Table 3. Since the market timing portfolio consists of both the risk-free asset and the big 10 portfolio, the market timing portfolio has a CAPM beta of 0.73. Thus I estimate the CAPM alpha yearly, using the CRSP value weight market return, and plot these estimates through time in Table 13. For most years, the market timing portfolio has positive alpha, but once again is rarely statistically significant.

VI Prediction Horizon Results

In this section, I answer the following question: given the predictive power of the empirical model, what is the time horizon over which downturns can be predicted?

I first discuss a basic but classic prediction horizon model. Let r_{t+1} be a log excess return. Define $a = E[r_{t+1}]$ and $x_t = E_t[r_{t+1}] - a$. Also, assume

$$\epsilon_{t+1} = r_{t+1} - a - x_t,\ x_{t+1} = \rho x_t + e_{t+1} \tag{12}$$

$$\epsilon_t \sim N(0, \sigma_\epsilon^2)\ iid,\ e_t \sim N(0, \sigma_e^2)\ iid,\ \sigma_{\epsilon,e} = \mathrm{Cov}(e_t, \epsilon_t) \tag{13}$$

The classic prediction horizon result is that the log of the cumulative returns from period $t+1$ to $t+T$, regressed on x_t results in the equation:

$$\sum_{s=1}^{T} r_{t+s} = Ta + \left(\sum_{s=1}^{T} \rho^{s-1}\right) x_t + \eta_{t+T} \tag{14}$$

where

$$\eta_{t+T} = \sum_{s=1}^{T} \left[\epsilon_{t+s} + e_{t+s}\left(\sum_{l=0}^{T-s-1} \rho^l\right)\right] \tag{15}$$

Thus the intercept and the coefficient on x_t increase over longer horizons. If the predictor x_t is persistent, i.e. $\rho \approx 1$, then the coefficient on x_t grows at approximately a linear rate as a function of the time horizon. The regression R^2 tends to increase as the prediction horizon T increases, until ρ^T decays sufficiently. Thus, in summary, return predictability tends to improve over longer horizons.

However, in the case of predicting downturns, the opposite is true. Define a true negative over the return horizon $t+1$ to $t+T$, denoted as p_{t+T}, as the unconditional probability of observing both a negative prediction and an actually negative return. Thus if $a > 0$ and $0 < \rho < 1$, then

$$p_{t+T} = \text{Prob}\left(Ta + \left(\sum_{s=1}^{T} \rho^{s-1}\right) x_t < 0 \text{ and } \sum_{s=1}^{T} r_{t+s} < 0\right) \to 0 \text{ as } T \to \infty \tag{16}$$

In words, the probability of experiencing a true negative converges to zero as the prediction horizon expands. In summary, although predictability tends to improve over longer horizons, the ability to predict *downturns* decreases over longer horizons.

I estimate the parameters of this very simple model using the main prediction model, in order to see how fast this convergence occurs. Tables 11 contain these point estimates. I use the mean of the log excess big 10 portfolio return as an estimate for a. The table shows 100 times the estimate of a, in order to be roughly in terms of percentage points. For x_t, I set

$$x_t = \log(\hat{r}_{t+1|t}) - \overline{\log(\hat{r}_{t+1|t})} \tag{17}$$

where the second term on the right hand side is just the mean of the log predicted return. I demean the predictions because in this very simple model, $\mathbb{E}[x_t] = 0$. I estimate ρ with the typical AR(1) regression using x_t. I then set ϵ_t and e_t as the respective residuals, and estimate the variance and covariance terms with the typical consistent estimators.

Figure 14 plots the model generated probability of a true negative as a function of the prediction horizon. The probability of a true negative is approximately zero at any prediction horizon longer than seven days. Prediction horizons of seven or fewer days have positive true negative probabilities. Thus this basic model, with these parameter values, implies that predicting downturns at horizons longer than a week is difficult, and longer than two weeks is not possible.

I fit the main model, using the same monthly estimation with a 10 year rolling window and 55% prediction probability cutoff as described above, to predict the log of the cumulative excess return of the big 10 portfolio at different horizons. Figure 15 plots the average log of the cumulative excess returns during periods that were predicted ex-ante to be negative, as a function of the prediction horizon. The grey area shows the 95% confidence region of the estimates, using Hodrick (1992) standard errors to correct for the overlapping sample. Similar to the basic model, it appears that predicting downturns with the predictive power of this model is possible at anything less than a one week horizon, becomes difficult between one and two weeks, and fails at horizons longer than two weeks.

VII Conclusion

This paper finds evidence that expected returns are sometimes negative. I present a parsimonious forecasting model where, conditional on a downturn prediction, the average excess return is -13.9 bps on portfolio of only the most liquid stocks. The same model works to predict stock market wide downturns as well, with a similar magnitude. The degree of statistical significance suggests that the null hypothesis of always non-negative expected stock returns is overwhelmingly rejected. Although I selected the main predictive model without regard to its actual empirical success, the estimate would lose statistical significance, using Bonferroni's correction, only if literally trillions of other predictive models are also tested.

I show that the main model predicts downturns during the entire sample period, and still predicts downturns in the last few decades as well. I show that the results are robust to varying different model parameters, such as the negative probability prediction cutoff. The main model predicts downturns even in the intraday return over the next day alleviating any concerns about using information available at market close on day t to predict day $t+1$ close-to-close returns. I also show that a simple linear prediction model also predicts market downturns.

I perform a formal test of multiple predictive models, correcting for the multiple comparisons, in order to determine the power of the predictive variables in forecasting downturns. I find that the CBOE volatility variables and classic asset pricing factor variables are the most powerful predictors, and these two groups are sufficient to reject the null hypothesis of always non-negative expected returns at the 95% confidence level even with Bonferroni's correction.

I construct a simple market timing portfolio based on the main model downturn predictions, and find that it earns an approximately daily 3.5 bps alpha, which translates into an 8.8% annual alpha. I show that this return typically outperforms the market during the sample period, and is not focused at the beginning or end of the sample.

I also show, with a simplistic model, that predicting downturns over periods longer than two weeks is essentially impossible with the estimated parameters. This paper suggests that even with a flexible model that leverages many variables to predict downturns, predictions of market downturns that are longer than two weeks are extremely difficult. This section suggests that claims of impending declines over a long time horizon, even using powerful forecasting methods, are heavily suspect.

As discussed in the literature review, much of the asset pricing prediction literature has focused on overall variation that is predictable, but not actually predicting downturns. Greenwood et al. (2017) discuss predicting downturns, but do not have a parsimonious model that predicts downturns, even in industry portfolios. The lack of empirical evidence of predictable downturns, before this paper, is especially striking given both the behavioral models (e.g. Barberis et al. (2015)) and at least one rational model (Barlevy and Veronesi (2003)) predicts that expected returns are sometimes negative. Although other papers, such as Greenwood and Hanson (2013) and Baron and Wei (2017) give evidence of predictable downturns in other assets, this is the first paper that gives evidence of negative expected excess aggregate stock market returns.

References

Ahn, Dong-Hyun, Jacob Boudoukh, Matthew Richardson, and Robert F. Whitelaw, 2002, Partial adjustment or stale prices? implications from stock index and futures return autocorrelations., *Review of Financial Studies* 15, 655 – 689.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-capm: An extrapolative capital asset pricing model., *Journal of Financial Economics* 115, 1 – 24.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2016, Extrapolation and bubbles, *working paper* .

Barlevy, Gadi, and Pietro Veronesi, 2003, Rational panics and stock market crashes., *Journal of Economic Theory* 110, 234.

Baron, Matthew, and Xiong Wei, 2017, Credit expansion and neglected crash risk. 132, 713 – 764.

Breiman, Leo, 2001, Random forests., *Machine Learning* 45, 5 – 32.

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

Carhart, Mark M., 1997, On persistence in mutual fund performance., *The Journal of Finance* 57.

Cutler, David M., and Lawrence H. Poterba, James M. Summers, 1990, Speculative dynamics and the role of feedback traders., *The American Economic Review* 63.

Dong, Lou, Christopher Polk, and Spyros Skouras, 2017, A tug of war: Overnight versus intraday expected returns., *working paper* .

Fama, Eugene F., 2013, Two pillars of asset pricing., *Nobel Prize Lecture* .

Fama, Eugene F., and Kenneth R. French, 1987, Commodity future prices: Some evidence on forecast power, premiums, and the theory of storage., *Journal of Business* 60, 55 – 73.

Fama, Eugene F, and Kenneth R French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3 – 25.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds., *Journal of Financial Economics* 33, 3 – 56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model., *Journal of Financial Economics* 116, 1 – 22.

Fama, Eugene F., and O. William Schwert, 1977, Asset returns and inflation., *Journal of Financial Economics* 5, 115 – 146.

French, Kenneth, G. Schwert, and Robert Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Greenwood, Robin, and Samuel G. Hanson, 2013, Issuer quality and corporate bond returns., *Review of Financial Studies* 26, 1483 – 1525.

Greenwood, Robin, Andrei Shleifer, and Yang You, 2017, Bubbles for fama., *Journal of Financial Economics* Forthcoming.

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

Malcolm Baker, author, and author Jeffrey Wurgler, 2000, The equity share in new issues and aggregate stock returns., *The Journal of Finance* 2219.

Martin Lettau, author, and author Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns., *The Journal of Finance* 815.

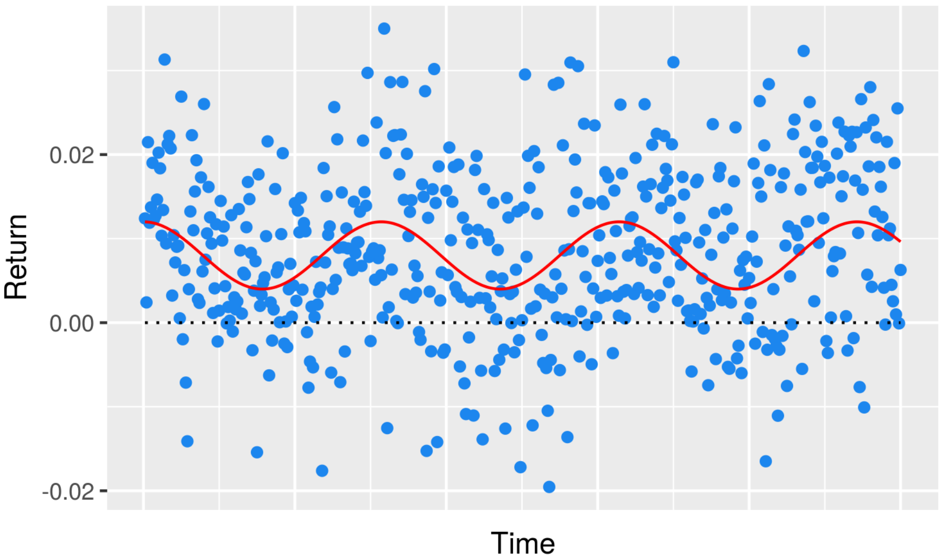
Newey, Whitney, and Kenneth West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix., *Econometrica* 55, 703 – 708.

Owen A. Lamont, author, and author Jeremy C. Stein, 2004, Aggregate short interest and market valuations., *The American Economic Review* 29.

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

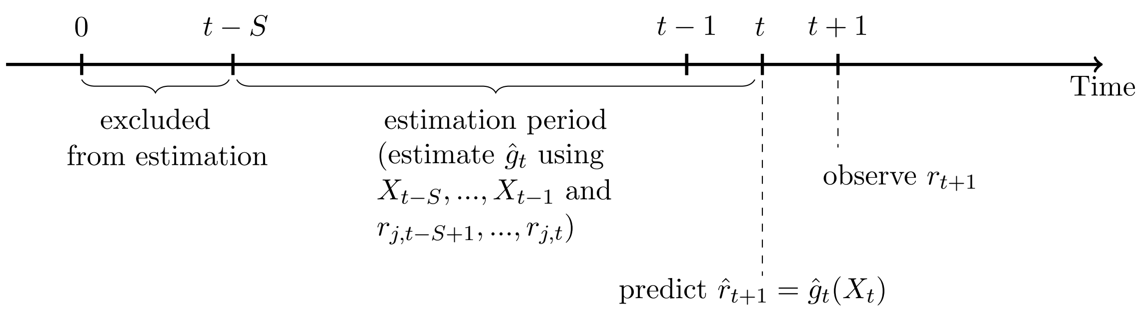
Figures

Figure 1: Variation in $\mathbb{E}_t[r_{t+1}]$ versus Negative $\mathbb{E}_t[r_{t+1}]$



This stylized plot shows hypothetical $\mathbb{E}_t[r_{t+1}]$ values (red curve) and actual return realizations r_{t+1} (blue dots). Note that in this hypothetical case, while there is substantial variation in $\mathbb{E}_t[r_{t+1}]$, it is still the case that $\mathbb{E}_t[r_{t+1}] > 0$ for all t.

Figure 2: Estimation Rolling Window Timeline



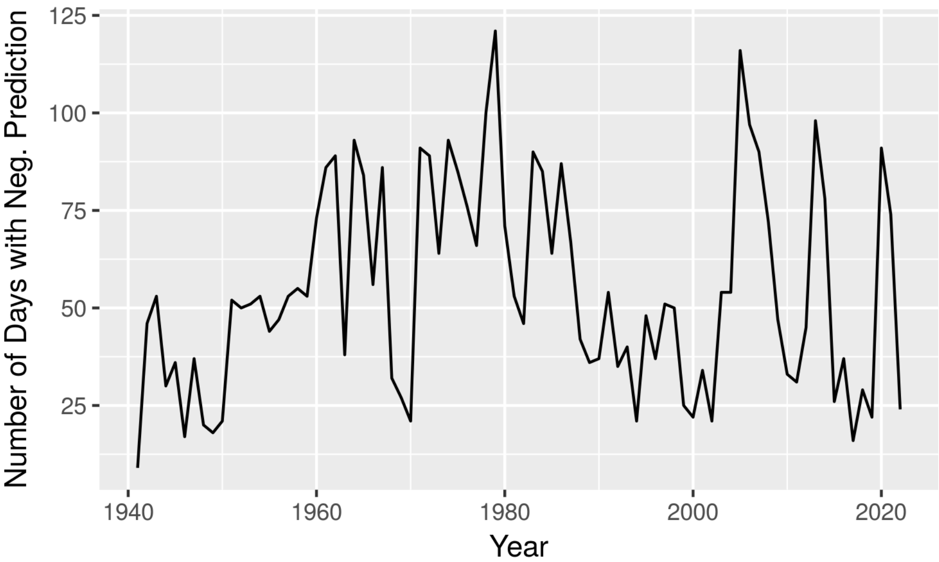
This figure shows how the forecasts of future returns are generated. The $S + 1$ periods from $t - S$ to t are used to forecast the $t + 1$ return. X_t is a vector of predictor variables available at time t, and $\hat{r}_{t+1}$ is the forecast of r_{t+1} generated at time t.

Figure 3: Confusion Matrix

		Prediction	
		$\hat{\pi}_{t+1\|t} < c$	$\hat{\pi}_{t+1\|t} > c$
Actual Return	$r_{t+1} > 0$	True Positive	False Negative
	$r_{t+1} < 0$	False Positive	True Negative

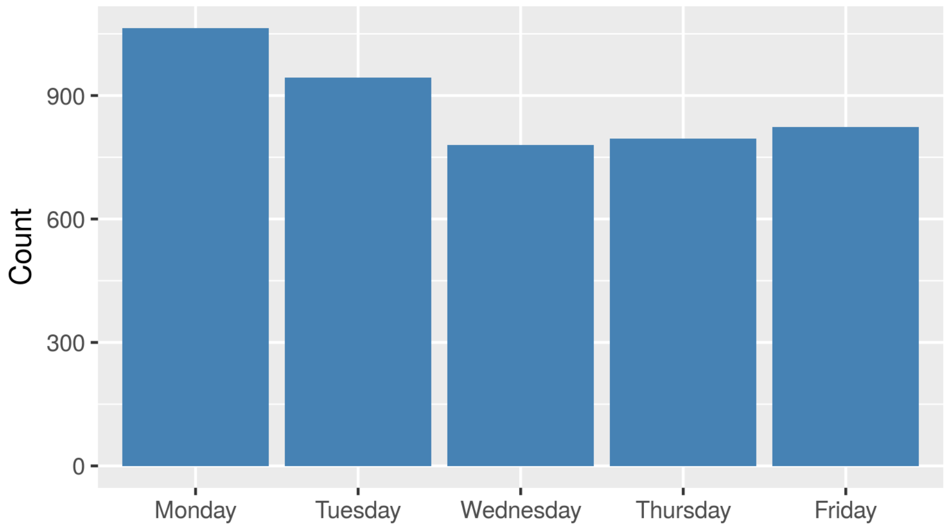
This table shows the two kinds of successful predictions (true positive and true negatives) and the two kinds of errors (false positives and false negatives). I discuss in the paper how the crucial error is false negatives, and that for the sake of testing whether a prediction method can predict negative returns, false positives matter very little. Here, $\hat{\pi}_{t+1|t}$ is the forecast of the probability of r_{t+1} being negative generated at time t.

Figure 4: Number of Days with Negative Predictions by Year



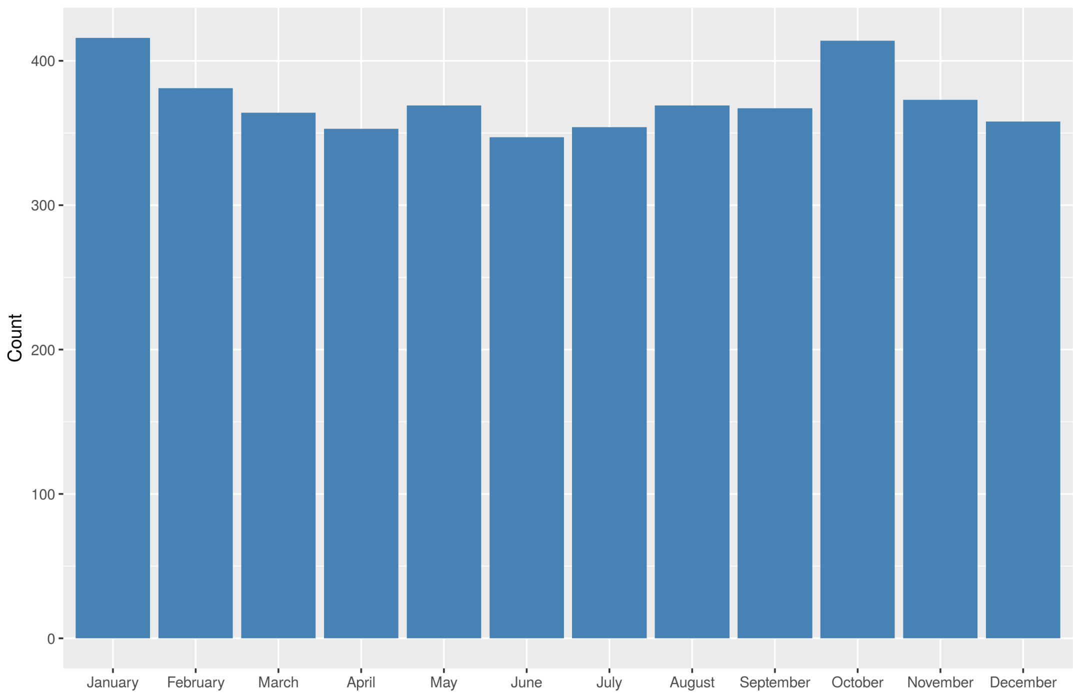
This plot shows the number of days with negative excess return big 10 predictions by year.

Figure 5: Number of Days with Negative Predictions by Day of the Week



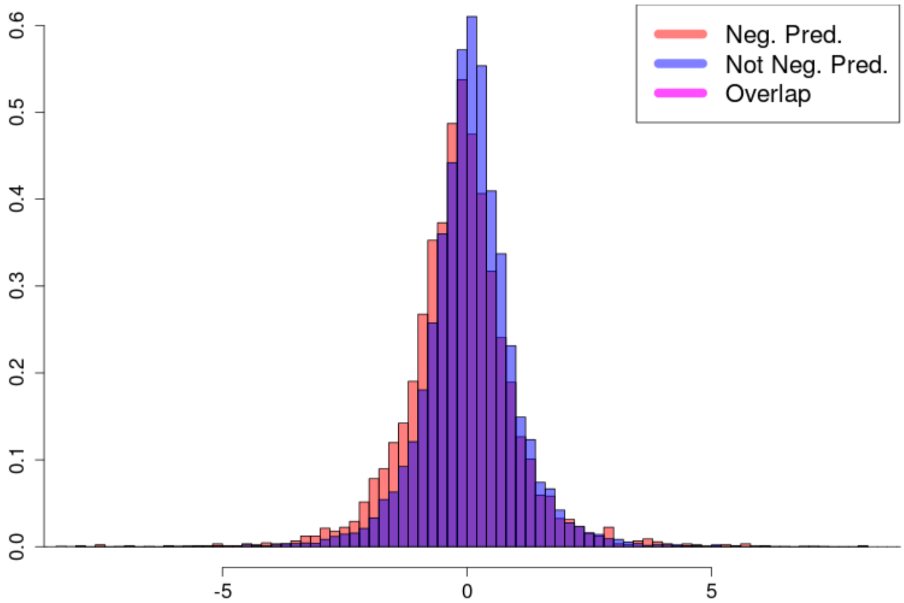
This shows the number of days with negative excess return big 10 predictions by day of the week.

Figure 6: Number of Days with Negative Predictions by Month



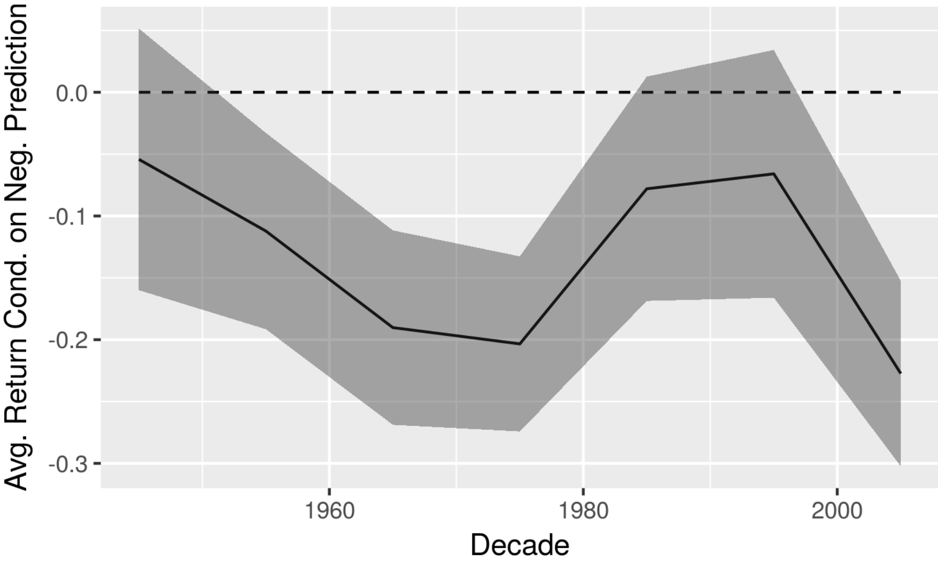
This shows the number of days with negative excess return big 10 predictions by month.

Figure 7: Overlapping Histograms of Returns Conditional on Predictions



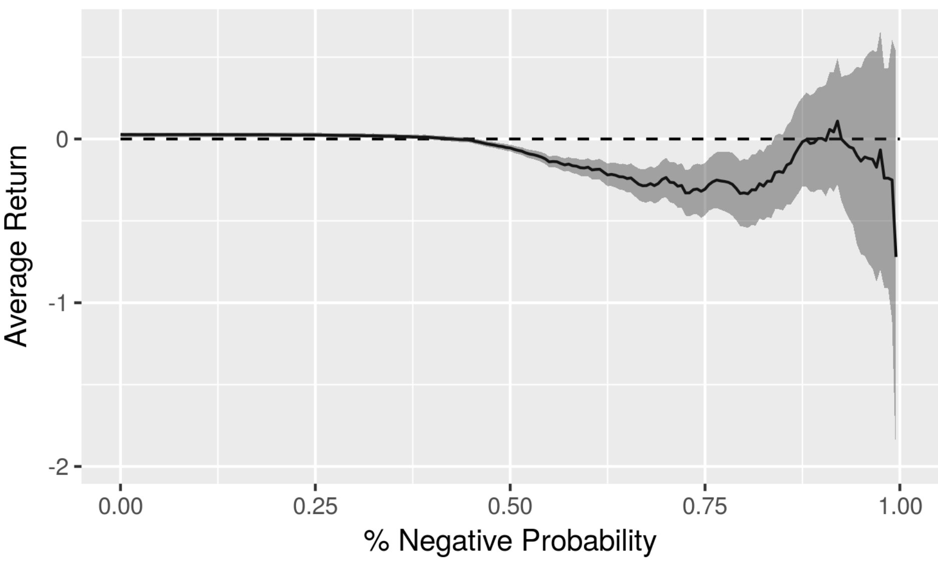
This shows the empirical distribution of two types of returns, those with negative return predictions and those that are not predicted to be negative. The beige and purple colors represent the distribution of returns conditional on a negative prediction. The blue and purple colors represent the distribution of returns conditional on not being predicted to be negative. The purple color is the overlap between the two distributions. Note that 21.2% percent of days in the sample are predicted to be negative, so both distributions are scaled in order to integrate to 1.

Figure 8: Estimated Average of Returns Conditional on a Negative Prediction by Decade



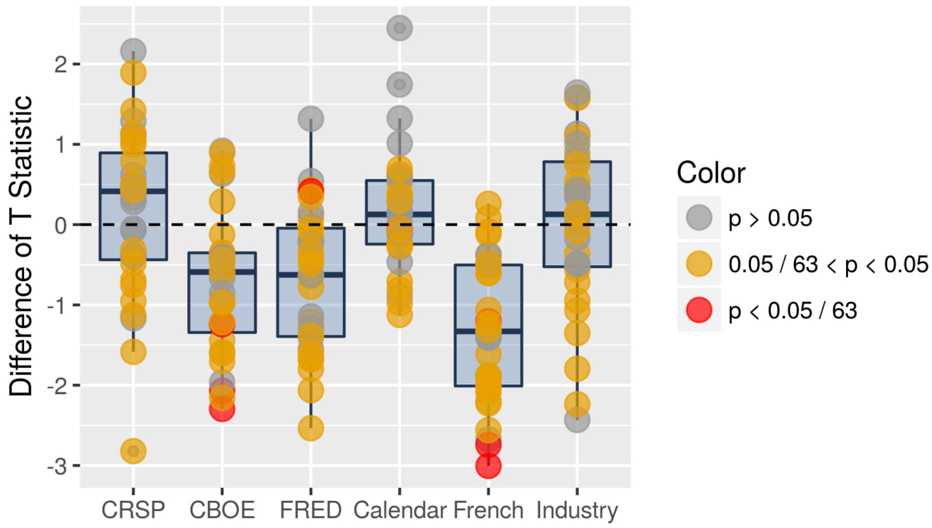
This shows the average return of days that are predicted to have a negative excess return, estimated by each decade. The grey region show the estimated 95% confidence interval for the estimated averages.

Figure 9: Varying the Cutoff Probability



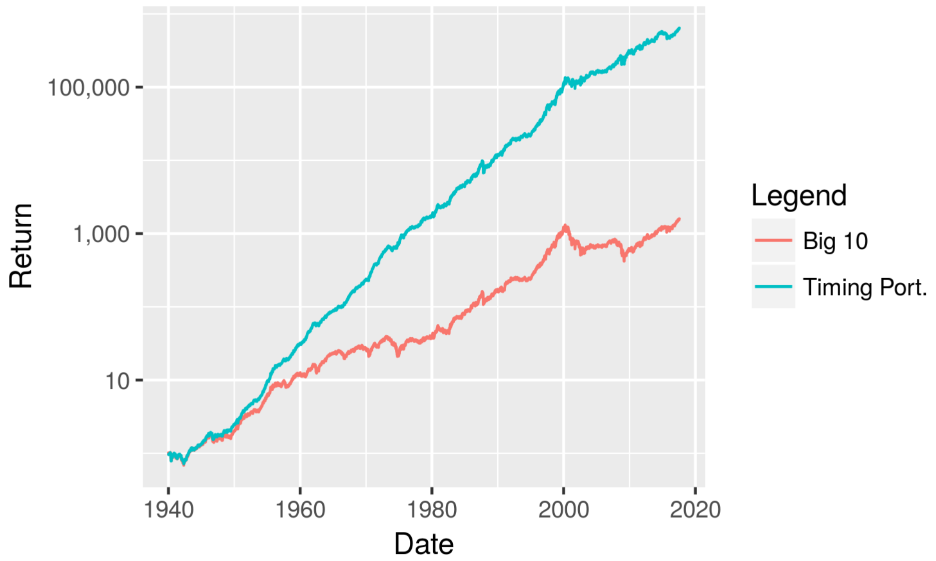
This shows how the average return conditional on a negative prediction $\hat{\pi}_{t+1|t} > c$ varies as the cutoff probability c varies. The cutoff c is on the x-axis, and the average return given a negative prediction with different cutoff values c are on the y-axis, with the shaded area representing the 95% confidence interval of the mean for different values of c.

Figure 10: Comparing Six Groups of Data Predictors



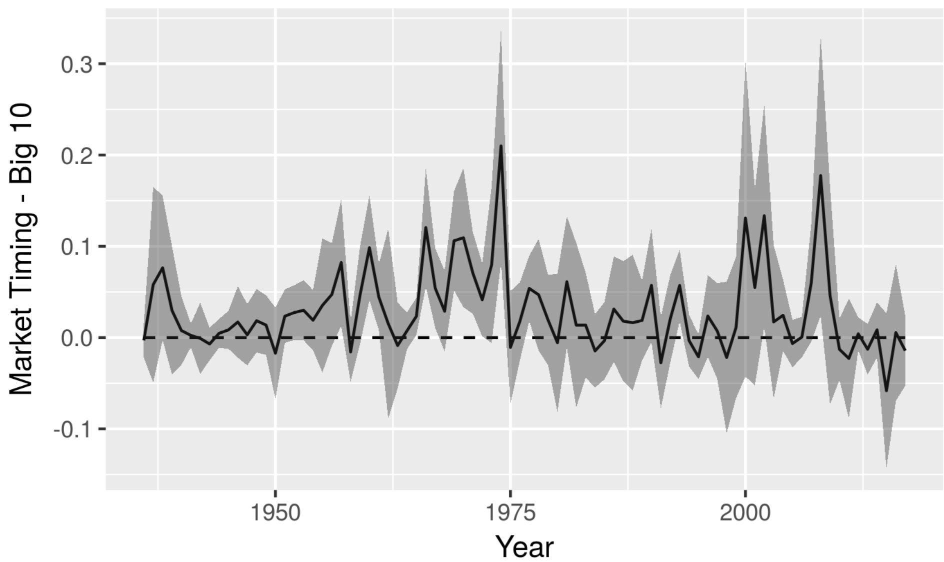
This box plot seeks to determine the strength of the six predictor variable groups described in the data section of the text. Consider predictor variables group j, and consider the value $t_j - t_j^b$, where t_j is the estimated t-statistic on γ^- from some model that includes group j, and t_j^b is the estimated t-statistic from a model with the same predictors but excludes group j (baseline model). If $t_j - t_j^b$ is very negative, then adding the predictor group j to the baseline model helped predict downturns better statistically. For each j, there are 31 $(2^5 - 1)$ different possible baseline model comparisons. This plots each of the 31 comparisons, for each of the six groups. Note that Industry refers to the 49 French Industry portfolio returns, and French refer to the 15 classic asset pricing factor variables. The sample period for these estimates is from January 2000 to July 2017.

Figure 11: Cumulative Return of Market Timing Portfolio



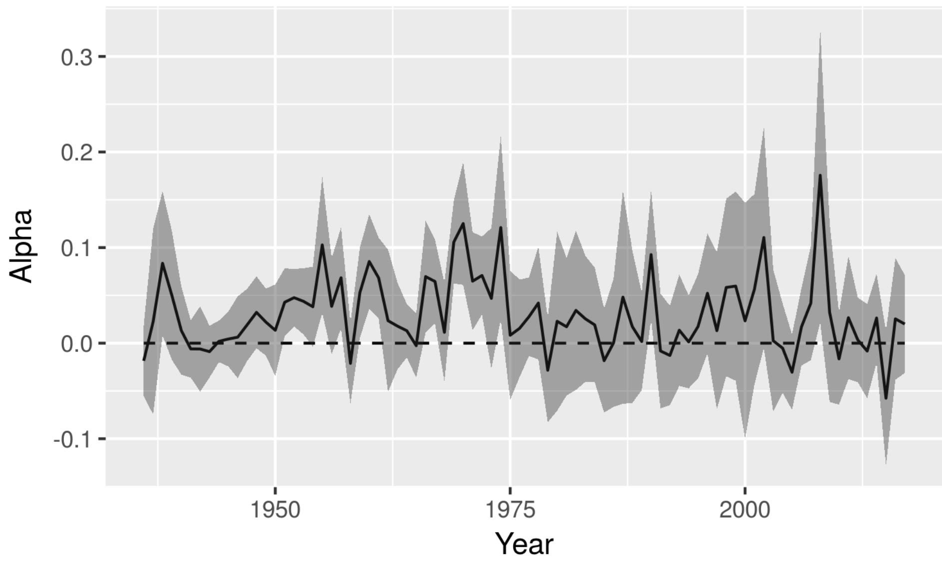
This graph shows the cumulative return of two portfolios. One is the big 10 portfolio discussed in the text, and the other is the market timing portfolio, which is 100% long the big 10 portfolio on days lacking downturn predictions, and holds on the risk-free asset on days with downturn predictions.

Figure 12: Market Timing Portfolio Performance by Year



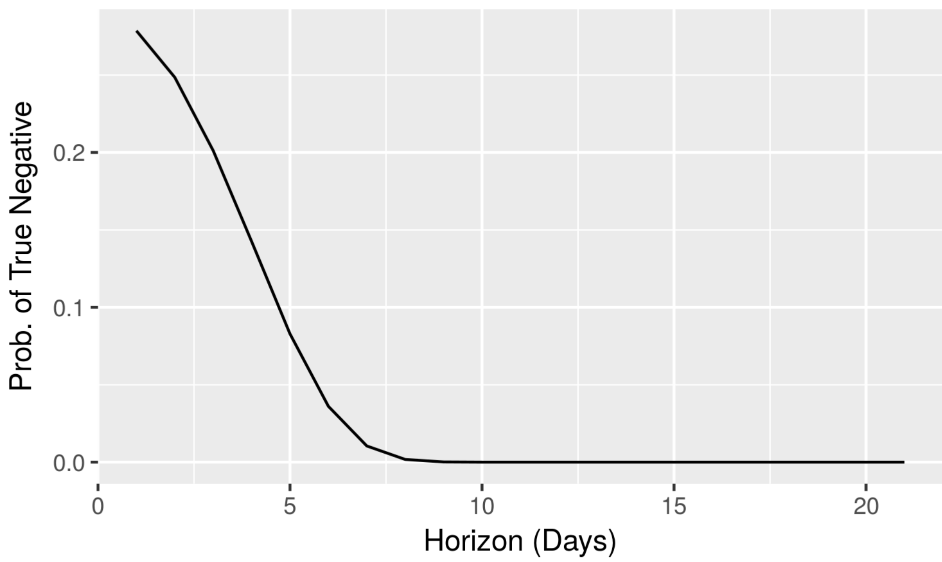
This graphs shows the estimated average market timing portfolio return minus the big 10 portfolio return every year in the sample. The shaded region represents the 95% confidence interval of the average difference.

Figure 13: CAPM α of Market Timing Portfolio by Year



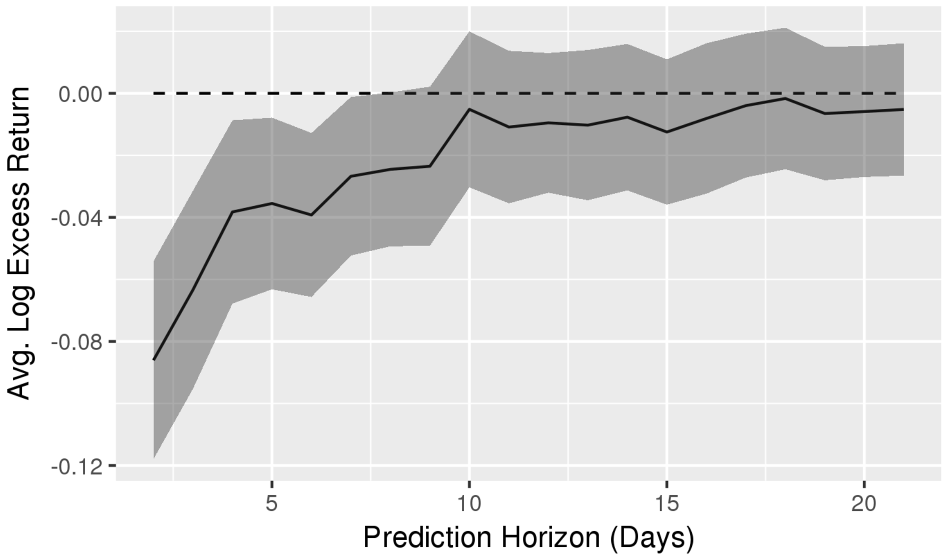
This graphs shows the estimated CAPM alpha estimated every year in the sample, with the shaded region representing the corresponding 95% confidence interval.

Figure 14: Theoretical Probability of True Negative over Different Time Horizons



This graph shows the unconditional probability of experiencing a true negative return (predicting to be negative and is actually negative) as a function of the number of days over which the cumulative return is predicted. Thus, at least with this basic model and estimated parameters, the probability of predicting a downturn correctly over a period of one week (5 trading days) is small, while the probability of predicting a downturn beyond two weeks is essentially zero.

Figure 15: Prediction Horizon Results



This graph shows the average log excess return over periods that are predicted to be negative, as a function of the number of days over which the prediction is made. The shaded region represents the 95% confidence interval of the estimates, where the standard errors are calculated according to a Hodrick (1992) overlapping cumulative return sample technique.

Tables

Table 1: **Description of Predictors**

First Date	Variable Name	Group	Description	Mean	Std. Dev.
1986-01-02	VXO	CBOE	S & P 100 Volatility Index	20.32	8.96
1990-01-02	VIX	CBOE	S & P 500 Volatility Index	19.51	7.86
1997-10-07	VXD	CBOE	DJIA Volatility Index	19.52	7.97
2001-02-02	VXN	CBOE	NASDAQ-100 Volatility Index	25.17	12.45
1926-11-01	CAP1RET	CRSP	Lowest Decile Market Cap Return	0.07%	1.26
1926-11-01	CAP2RET	CRSP	Decile 2 Market Cap Return	0.05%	1.17
1926-11-01	CAP3RET	CRSP	Decile 3 Market Cap Return	0.05%	1.12
1926-11-01	CAP4RET	CRSP	Decile 4 Market Cap Return	0.05%	1.11
1926-11-01	CAP5RET	CRSP	Decile 5 Market Cap Return	0.05%	1.14
1926-11-01	CAP6RET	CRSP	Decile 6 Market Cap Return	0.05%	1.14
1926-11-01	CAP7RET	CRSP	Decile 7 Market Cap Return	0.05%	1.15
1926-11-01	CAP8RET	CRSP	Decile 8 Market Cap Return	0.04%	1.12
1926-11-01	CAP9RET	CRSP	Decile 9 Market Cap Return	0.05%	1.10
1926-11-01	CAP10RET	CRSP	Largest Decile Market Cap Return	0.04%	1.07
1926-11-01	EWRETD	CRSP	Equal Weight Stock Return	0.09%	1.06
1926-11-01	EWRETX	CRSP	Equal Weight Stock Return (Excluding Dividends)	0.07%	1.06
1926-11-01	TOTCNT	CRSP	% Chg. in Number of Listed Stocks	0.01%	1.06
1926-11-01	TOTVAL	CRSP	% Chg. in Value of Listed Stocks	0.04%	1.11
1926-11-01	USDCNT	CRSP	% Chg. in Number of Listed Stocks in CRSP Market Index	0.01%	1.05
1926-11-01	USDVAL	CRSP	% Chg. in Value of Listed Stocks in CRSP Mkt. Index	0.04%	1.11
1926-11-01	VWRETD	CRSP	Value Weight Stock Return	0.04%	1.07
1926-11-01	VWRETX	CRSP	Value Weight Stock Return (Excluding Dividends)	0.03%	1.07
1962-07-03	SPRTRN	CRSP	S&P 500 Return	0.03%	1.01
1954-07-01	DFF	FRED	Federal Funds Rate	4.89	3.64
1962-01-02	DGS1	FRED	1-Year Treasury Yield	5.23	3.41
1962-01-02	DGS10	FRED	10-Year Treasury Yield	6.28	2.86
1976-06-01	T10Y2Y	FRED	10-Year Treasury Yield Minus 2-Year Treasury Yield	0.97	0.93
1986-01-02	DBAA	FRED	Moody's Baa Corporate Bond Yield	7.51	1.87
1986-01-02	USD1MTD156N	FRED	1-Month LIBOR	3.76	2.80
1987-05-20	DCOILBRENTEU	FRED	Brent Crude Oil Price	45.09	33.46
1995-01-04	DTWEXB	FRED	Trade Weighted US Dollar Index	109.69	10.24
1996-12-31	BAMLH0A0HYM2EY	FRED	BofA High Yield Index	9.15	2.85
1999-01-04	DEXUSEU	FRED	US/Euro Exchange Rate	1.21	0.17
2003-01-02	T10YIE	FRED	10-Year Inflation Rate	2.09	0.41
2003-01-02	T5YIFR	FRED	5-Year Forward Inflation Expectation Rate	2.33	0.35
1926-11-01	MKTRF	French	Excess Value Weight Market Return	0.03%	1.07
1926-11-01	SMB	French	Small Minus Big	0.01%	0.59
1926-11-01	HML	French	High Minus Low	0.02%	0.59
1926-11-03	UMD	French	Up Minus Down (Momentum Portfolio)	0.03%	0.75
1926-11-01	RF	French	risk-free Rate (Treasury)	0.01%	0.01
1963-07-01	CMA	French	Conservative Minus Aggressive (Investment Portfolio)	0.01%	0.36

1963-07-01	RMW	French	Robust Minus Weak (Profitability Portfolio)	0.01%	0.37
1926-07-01	HIGH	French	Long End of HML Return	0.03%	1.67
1926-07-01	LOW	French	Short End of HML Return	0.03%	0.44
1963-07-01	SMALL	French	Long End of SMB Return	0.04%	1.38
1927-07-01	BIG	French	Short End of SMB Return	0.03%	2.02
1926-07-01	UP	French	Long End of UMD Return	0.08%	0.28
1926-07-01	DOWN	French	Short End of UMD Return	0.08%	0.28
1930-03-20	LTREV	French	Long Run Reversals Index	0.01%	0.60
1926-11-01	STREV	French	Short Run Reversals	0.02%	0.92
1926-11-01	AERO	French Industry	Aircraft Index	0.05%	1.78
1926-11-01	AGRIC	French Industry	Agriculture Index	0.03%	1.50
1926-11-01	AUTOS	French Industry	Automobiles and Trucks Index	0.03%	1.57
1926-11-01	BANKS	French Industry	Banking Index	0.04%	1.47
1926-11-01	BEER	French Industry	Beer & Liquor Index	0.04%	1.46
1926-11-01	BLDMT	French Industry	Construction Materials Index	0.03%	1.25
1926-11-01	BOOKS	French Industry	Printing and Publishing Index	0.03%	1.55
1926-11-01	BOXES	French Industry	Shipping Containers Index	0.04%	1.25
1926-11-01	BUSSV	French Industry	Business Services Index	0.04%	1.97
1926-11-01	CHEMS	French Industry	Chemicals Index	0.03%	1.27
1926-11-01	CHIPS	French Industry	Electronic Equipment Index	0.04%	1.75
1926-11-01	CLTHS	French Industry	Apparel Index	0.03%	1.14
1926-11-01	CNSTR	French Industry	Construction Index	0.04%	2.00
1926-11-01	COAL	French Industry	Coal Index	0.03%	2.12
1926-11-01	DRUGS	French Industry	Pharmaceutical Products Index	0.04%	1.14
1926-11-01	ELCEQ	French Industry	Electrical Equipment Index	0.04%	1.56
1963-07-01	FABPR	French Industry	Textiles Index	0.01%	1.49
1926-11-01	FIN	French Industry	Fabricated Products Index	0.04%	1.58
1926-11-01	FOOD	French Industry	Food Products Index	0.03%	0.92
1926-11-01	FUN	French Industry	Entertainment Index	0.04%	1.80
1963-07-01	GOLD	French Industry	Precious Metals Index	0.03%	2.28
1963-07-01	GUNS	French Industry	Defense Index	0.04%	1.38
1926-11-01	HARDW	French Industry	Computers Index	0.04%	1.53
1969-07-01	HLTH	French Industry	Healthcare Index	0.03%	1.53
1926-11-01	HSHLD	French Industry	Consumer Goods Index	0.03%	1.16
1926-11-01	INSUR	French Industry	Insurance Index	0.03%	1.37

1926-11-01	LABEQ	French Industry	Measuring and Control Equipment Index	0.04%	1.43
1926-11-01	MACH	French Industry	Machinery Index	0.03%	1.37
1926-11-01	MEALS	French Industry	Restaurants, Hotels, and Motels Index	0.04%	1.34
1926-11-01	MEDEQ	French Industry	Medical Equipment Index	0.04%	1.59
1926-11-01	MINES	French Industry	Non-Metallic and Industrial Metal Mining Index	0.03%	1.53
1926-11-01	OIL	French Industry	Petroleum and Natural Gas Index	0.03%	1.28
1926-11-01	OTHER	French Industry	Almost Nothing Index	0.02%	1.48
1929-07-01	PAPER	French Industry	Business Supplies Index	0.06%	3.28
1927-07-01	PERSV	French Industry	Personal Services Index	0.03%	2.02
1926-11-01	RLEST	French Industry	Real Estate Index	0.03%	2.14
1926-11-01	RTAIL	French Industry	Retail Index	0.03%	1.13
1930-07-01	RUBBR	French Industry	Rubber and Plastic Products Index	0.04%	1.67
1926-11-01	SHIPS	French Industry	Shipbuilding and Railroad Equipment Index	0.03%	1.51
1926-11-01	SMOKE	French Industry	Tobacco Products Index	0.04%	1.19
1963-07-01	SODA	French Industry	Candy & Soda Index	0.04%	1.39
1965-07-01	SOFTW	French Industry	Computer Software Index	0.03%	2.37
1926-11-01	STEEL	French Industry	Steel Works Etc Index	0.03%	1.67
1926-11-01	TELCM	French Industry	Communication Index	0.03%	1.03
1926-11-01	TOYS	French Industry	Recreation Index	0.03%	2.14
1926-11-01	TRANS	French Industry	Transportation Index	0.03%	1.35
1926-11-01	TXTLS	French Industry	Textiles Index	0.03%	1.31
1926-11-01	UTIL	French Industry	Utilities Index	0.03%	1.09
1926-11-01	WHLSL	French Industry	Wholesale Index	0.03%	1.64
1926-11-01	MONTH	Calendar	month		
1926-11-01	WEEKDAY	Calendar	weekday		
1926-07-01	BIG10	CRSP	VW Portfolio of Largest 10 Stocks	0.03%	1.12
1926-07-01	BIG10INTRADAY	CRSP	Intraday Return of Big10	0.00%	0.85
1926-07-01	BIG10OVERNIGHT	CRSP	Overnight Return of Big10	0.02%	0.45
1926-07-01	BIG50	CRSP	VW Portfolio of Largest 50 Stocks	0.03%	1.08
1926-07-01	BIG50INTRADAY	CRSP	Intraday Return of Big50	0.00%	0.81
1926-07-01	BIG50OVERNIGHT	CRSP	Overnight Return of Big50	0.02%	0.43
1926-07-01	BIG100	CRSP	VW Portfolio of Largest 100 Stocks	0.03%	1.08
1926-07-01	BIG100INTRADAY	CRSP	Intraday Return of Big100	0.00%	0.8
1926-07-01	BIG100OVERNIGHT	CRSP	Overnight Return of Big100	0.02%	0.43
1926-07-01	BIG200	CRSP	VW Portfolio of Largest 200 Stocks	0.03%	1.08
1926-07-01	BIG200INTRADAY	CRSP	Intraday Return of Big200	0.00%	0.8
1926-07-01	BIG200OVERNIGHT	CRSP	Overnight Return of Big200	0.03%	0.44

1926-07-01	BIG300	CRSP	VW Portfolio of Largest 300 Stocks	0.03%	1.08
1926-07-01	BIG300INTRADAY	CRSP	Intraday Return of Big300	0.00%	0.8
1926-07-01	BIG300OVERNIGHT	CRSP	Overnight Return of Big300	0.03%	0.44
1926-07-01	BIG400	CRSP	VW Portfolio of Largest 400 Stocks	0.03%	1.07
1926-07-01	BIG400INTRADAY	CRSP	Intraday Return of Big400	0.00%	0.79
1926-07-01	BIG400OVERNIGHT	CRSP	Overnight Return of Big400	0.03%	0.44
1926-07-01	BIG500	CRSP	VW Portfolio of Largest 500 Stocks	0.03%	1.07
1926-07-01	BIG500INTRADAY	CRSP	Intraday Return of Big500	0.00%	0.79
1926-07-01	BIG500OVERNIGHT	CRSP	Overnight Return of Big500	0.03%	0.44
1926-07-01	BIG1000	CRSP	VW Portfolio of Largest 1000 Stocks	0.03%	1.07
1926-07-01	BIG1000INTRADAY	CRSP	Intraday Return of Big1000	0.00%	0.79
1926-07-01	BIG1000OVERNIGHT	CRSP	Overnight Return of Big1000	0.03%	0.44
1926-07-01	BIG2000	CRSP	VW Portfolio of Largest 2000 Stocks	0.03%	1.07
1926-07-01	BIG2000INTRADAY	CRSP	Intraday Return of Big2000	0.00%	0.79
1926-07-01	BIG2000OVERNIGHT	CRSP	Overnight Return of Big2000	0.03%	0.44

This table shows the time t predictor variables used to predict time $t+1$ returns. It gives the variables name, the day the data series is available, the group variable belongs to as described in the text (6 group total), the description, and the mean and standard deviation. Both the French industry and French portfolio group variables were downloaded from Ken French's website, as discussed in the text.

Table 2: Correlation Matrix of Key Portfolios

	10	50	100	200	300	400	500	1000	2000	market
10	1	0.975	0.966	0.958	0.953	0.949	0.947	0.941	0.938	0.934
50		1	0.997	0.992	0.989	0.987	0.985	0.981	0.978	0.975
100			1	0.998	0.996	0.995	0.993	0.990	0.987	0.985
200				1	0.999	0.999	0.998	0.996	0.994	0.992
300					1	1.000	0.999	0.998	0.996	0.995
400						1	1.000	0.999	0.998	0.997
500							1	0.999	0.998	0.997
1000								1	1.000	0.999
2000									1	1.000
market										1

This table shows the correlation matrix of the excess returns of the various portfolios I predict in this paper. For example, 10 represents the value weight portfolio excess return of the 10 stocks with the largest market capitalization at the end of the previous year. This portfolio is referred to as the big 10 portfolio in the text. The portfolio labeled "market" is the excess return on the CRSP value weight portfolio.

Table 3: CAPM Alpha and Beta of Key Portfolios

	Dependent variable:								
	10	50	100	200	300	400	500	1000	2000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
α	−0.001	−0.001	−0.002	−0.001	−0.001	−0.001	−0.0004	0.0001	0.0003*
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0005)	(0.0003)	(0.0002)
Excess Market	0.987***	0.993***	1.000***	1.003***	1.005***	1.005***	1.005***	1.003***	1.002***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0005)	(0.0003)	(0.0002)
Obs.	24,034	24,034	24,034	24,034	24,034	24,034	24,034	24,034	24,034
R^2	0.873	0.951	0.971	0.985	0.990	0.993	0.995	0.998	0.999
Note:									*p<0.1; **p<0.05; ***p<0.01

This table shows the CAPM alpha and beta, using the excess return on the CRSP value weight market portfolio on the right-hand side of the regression, of the excess returns of the various portfolios I predict in this paper. For example, 10 represents the value weight portfolio excess return of the 10 stocks with the largest market capitalization at the end of the previous year. This portfolio is referred to as the big 10 portfolio in the text.

Table 4: Conditional Averages

	Dependent variable:					
	Big 10	Big 10 Overnight	Big 10 Intraday	Big 10	Big 10 Overnight	Big 10 Intraday
	July 1936 - July 2017			January 1980 - July 2017		
	(1)	(2)	(3)	(4)	(5)	(6)
γ^+	0.072***	0.027***	0.025***	0.063***	0.026***	0.024***
	(0.007)	(0.003)	(0.005)	(0.012)	(0.005)	(0.009)
γ^-	−0.139***	−0.022***	−0.050***	−0.105***	0.005	−0.076***
	(0.016)	(0.007)	(0.011)	(0.029)	(0.014)	(0.020)
Observations	21,058	21,058	21,058	9,478	9,478	9,478
Note:						*p<0.1; **p<0.05; ***p<0.01

This table shows estimates of the average excess return of the big 10 portfolio conditional on a negative prediction, γ^-, using the main model described in the text. It also shows the average return conditional on lacking a negative prediction, γ^+. Note that the units are given in percentage points, and the Newey and West (1987) standard errors with 3 lags are in parentheses. Column 1 is the average on days with downturns predicted by the main model described in the text, while columns 2 and 3 are the average return on these same days of just the overnight and intraday return respectively. Columns 1, 2, and 3 are estimated over the entire sample period. Columns 4, 5, 6 are the same as columns 1, 2, and 3 respectively except these statistics are estimated using the period from the beginning of 1980 through the end of the sample (July 2017).

Table 5: Different Standard Error Calculations of Conditional Averages

	Dependent variable:				
	Big 10 Excess Return				
Std. Err. Calculation:	OLS	NW2	NW3	NW5	NW21
	(1)	(2)	(3)	(4)	(5)
γ^+	0.072***	0.072***	0.072***	0.072***	0.072***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
γ^-	−0.139***	−0.139***	−0.139***	−0.139***	−0.139***
	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)
Observations	21,058	21,058	21,058	21,058	21,058
Note:					*p<0.1; **p<0.05; ***p<0.01

This table shows estimates of the average excess return of the big 10 portfolio conditional on a negative prediction, γ^-, with the main forecasting model discussed in the text. It also shows the average return conditional on lacking a negative prediction, γ^+. Note that the units are given in percentage points. Column 1 shows the typical OLS standard errors in parentheses. Columns 2, 3, 4, and 5 show the Newey and West (1987) standard errors with 2, 3, 5, and 21 lags respectively in parentheses.

Table 6: Predicting Various Excess Portfolio Returns

	Dependent variable:									
	10	50	100	200	300	400	500	1000	2000	mkt
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
γ^+	0.072***	0.068***	0.065***	0.077***	0.078***	0.077***	0.081***	0.089***	0.086***	0.086***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
γ^-	−0.139***	−0.115***	−0.101***	−0.138***	−0.138***	−0.131***	−0.142***	−0.158***	−0.147***	−0.150***
	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Obs	21,058	21,058	21,058	21,058	21,058	21,058	21,058	21,058	21,058	21,058
Note:										*p<0.1; **p<0.05; ***p<0.01

This table shows estimates of the average excess return of various portfolios conditional on a negative prediction, γ^-, using the main model described in the text. It also shows the average return conditional on lacking a negative prediction, γ^+. Note that the units are given in percentage points, and the Newey and West (1987) standard errors with 3 lags are in parentheses. The column labeled 10 represents the value weight portfolio excess return of the 10 stocks with the largest market capitalization at the end of the previous year. This portfolio is referred to as the big 10 portfolio in the text. The other columns are similar, except with the labeled number of stocks instead of 10. The portfolio labeled “mkt” is the excess return on the CRSP value weight portfolio.

Table 7: Different Prediction Methods

Dep. Variable:	Big 10 Excess Return					
Model:	Linear		RF	Linear		RF
Cutoff:	$c = 0.5$			$c = 0.55$		
Prob. Measure:	OLS	Normal		OLS	Normal	
	(1)	(2)	(3)	(4)	(5)	(6)
γ^+	0.082***	0.082***	0.094***	0.059***	0.059***	0.072***
	(0.009)	(0.009)	(0.009)	(0.007)	(0.007)	(0.007)
γ^-	−0.041***	−0.041***	−0.055***	−0.086***	−0.084***	−0.139***
	(0.010)	(0.010)	(0.010)	(0.016)	(0.016)	(0.016)
Observations	21,058	21,058	21,058	21,058	21,058	21,058
Note:				*p<0.1; **p<0.05; ***p<0.01		

This table shows estimates of the average excess return of the big 10 portfolio conditional on a negative prediction, γ^-, using various prediction models. It also shows the average return conditional on lacking a negative prediction, γ^+. Note that the units are given in percentage points, and the Newey and West (1987) standard errors with 3 lags are in parentheses. Columns 1, 2, 4, and 5 use the linear prediction model described in the text, while columns 3 and 6 use the random forest model described in the text. Columns 1, 2, and 3 use a cutoff probability value of 0.5, while columns 4, 5, and 6 use a probabilty cutoff value of 0.55. Columns 1 and 4 use the OLS negative probability prediction, while columns 2, 3, 5, and 6 use the normality negative probability prediction.

Table 8: Top 10 Predictor Importances

First Date	Variable Name	Source	Description	Avg. Imp.	Corr.
1986-01-02	VXO	CBOE	S&P 100 Volatility Index	1.5	0.14
1926-11-01	USDVAL	CRSP	% Chg. in Value of Listed Stocks in CRSP Mkt. Index	1.24	-0.05
1926-11-01	STREV	French	Short Run Reversals	1.09	-0.03
1990-01-02	VIX	CBOE	S&P 500 Volatility Index	1.09	0.15
1926-11-01	OIL	French	Petroleum and Natural Gas Index	1.07	-0.01
1963-07-01	GOLD	French	Precious Metals Index	1.02	-0.05
1997-10-07	VXD	CBOE	DJIA Volatility Index	1.01	0.16
1963-07-01	RMW	French	Robust Minus Weak (Profitability Portfolio)	0.96	0.13
1926-11-01	SMB	French	Small Minus Big	0.95	-0.08
2001-02-02	VXN	CBOE	NASDAQ-100 Volatility Index	0.95	0.11

This table shows the 10 predictor variables with the largest random forest importances, averaged across time. Note that not all predictor variables are available during the entire sample, thus each average is actually just the average of all the models in which the predictor is used. An importance of 100 means the predictor variable explained 100% of the variation in next day returns explained by all predictor variables. The columns correspond, from left to right, to the day the data series is available, the name of the variable, the source of the data, the variable description, the variable importance, and the correlation of this variable at time t with day $t+1$ big 10 excess returns.

Table 9: Strength of Predictor Variable Groups

	Dependent variable:				
	big10				
	(1)	(2)	(3)	(4)	(5)
γ^+	0.048^{*}	0.055^{**}	0.056^{**}	0.058^{**}	0.050^{**}
	(0.021)	(0.019)	(0.019)	(0.020)	(0.019)
γ^-	-0.079^{**}	-0.149^{***}	-0.153^{***}	-0.127^{***}	-0.133^{**}
	(0.034)	(0.042)	(0.042)	(0.040)	(0.043)
CRSP	No	No	No	No	Yes
CBOE	Yes	Yes	Yes	Yes	Yes
FRED	No	No	No	Yes	Yes
Calendar	No	No	Yes	Yes	Yes
French Portfolios	No	Yes	Yes	Yes	Yes
French Industry	No	No	No	No	Yes
Observations	4,422	4,422	4,422	4,422	4,422

Note: *p<0.05; **p<0.01; ***p< 0.05/63

This table shows estimates of the average excess return of the big 10 portfolio conditional on a negative prediction, γ^-, using various prediction models. It also shows the average return conditional on lacking a negative prediction, γ^+. Note that the units are given in percentage points, and the Newey and West (1987) standard errors with 3 lags are in parentheses. Note that the sample period is from January 2000 through July 2017. The cutoff probability c used to make the predictions, like the main model, is 0.55. For each of the six predictor variable groups described in the data section of the text, a "Yes" indicates the predictor variable group was used in the model, and "No" of course means the opposite. Note that, as the text explains, there were 63 models total, and thus I use a Bonferroni's correction with one-sided tests to calculate statistical significance. The note immediately below the table shows what the asterisks represent.

Table 10: Factor Regressions with Market Timing Portfolio

	Dependent variable:				
	Market Timing Portfolio Excess Return				
	(1)	(2)	(3)	(4)	(5)
α	0.035***	0.039***	0.037***	0.036***	0.035***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Excess Market	0.734***	0.730***	0.735***	0.732***	0.735***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
HML		−0.092***	−0.077***	−0.190***	−0.168***
		(0.007)	(0.007)	(0.011)	(0.012)
SMB		−0.324***	−0.325***	−0.329***	−0.333***
		(0.007)	(0.007)	(0.010)	(0.010)
Momentum			0.061***		0.043***
			(0.005)		(0.007)
CMA				0.176***	0.159***
				(0.016)	(0.017)
RMW				0.135***	0.127***
				(0.014)	(0.014)
Observations	21,058	21,058	21,058	13,614	13,614
R^2	0.625	0.661	0.663	0.653	0.654
Note:					*p<0.1; **p<0.05; ***p<0.01

This table shows the alpha and betas of the market timing portfolio described in the text. The excess market variable is the excess return on the CRSP value weight market index return. Columns 1, 2, 3, 4, and 5 correspond to, respectively, the CAPM model, Fama and French (1993) 3 factor model, Carhart (1997) 4 factor model, Fama and French (2015) 5 factor model, and a model with these same 5 factors with momentum. The units of alpha are in terms of percentage points, and the estimated standard errors are in parentheses.

Table 11: Time Horizon Model Parameter Estimates

a	ρ	σ_ϵ	σ_e	$\text{cor}(\epsilon_t, e_t)$
0.021	0.363	1.045	0.290	−0.010

This table shows the estimated parameters of the basic prediction horizon model described in Section VI.