A Practitioner's Defense of Return Predictability

Blair Hull and Xiao Qiao

Blair Hull
is the founder and a managing partner of Ketchum Trading, LLC and the founder and chairman of Hull Investments, LLC in Chicago, IL.
blairh@hullinvestmentsllc.com

Xiao Qiao
is a research analyst at SummerHaven Investment Management, LLC in Stamford, CT.
xqiao@uchicago.edu

A central question in financial economics is the estimation of expected market returns. Financial claims on real assets bear non-zero returns for two reasons. First, one dollar received tomorrow is not equal to one dollar received one year from today, since investors demand compensation for non-immediacy. The second source of returns comes from the fact that many financial assets are risky, and investors are compensated for holding these risky assets. For the aggregate equities market, this adjustment for risk is known as the equity premium.

It is well known that the equity premium is difficult to estimate. Merton [1980] called attempts to estimate the equity premium a "fool's errand":

> Indeed, even if the expected return on the market were known to be a constant for all time, it would take a very long history of returns to obtain an accurate estimate. And, of course, if this expected return is believed to be changing through time, then estimating these changes is still more difficult. (Merton [1980], p. 326).

Much of the empirical asset-pricing literature up until Merton [1980] assumed a constant rate of return for the market, while Merton anticipated the possibility of a non-constant equity premium. Indeed, the equity premium may be time varying and move around depending on prevailing business conditions.

If the equity premium is time varying, then presumably we can forecast this quantity given the appropriate information set. Early evidence from Fama and French [1988] and Campbell and Shiller [1988a, 1988b], among others, showed that market returns can be predicted using dividend yields. However, evidence both for and against return predictability cropped up in the years following these pioneering works. In an influential study, Welch and Goyal [2008] examined 14 different forecasting variables proposed by academics and found that the predictors are unstable both in-sample and out-of-sample. They concluded the variables would not have helped investors profitably time the market. On the other hand, Cochrane [2008] made sound theoretical arguments in favor of return predictability by jointly examining the forecastability of returns and dividend growth. Subsequently, Rapach, Strauss, and Zhou [2010] provided strong evidence that the stock market can be consistently predicted out-of-sample.

There appears to be evidence for predictability over both the long and short term. At the one-month frequency, Moskowitz, Ooi, and Pedersen [2012] documented that past 12-month market excess return is a positive indicator of the next-month market return. Dividend yield (see Campbell and Shiller [1988a, 1988b]; Cochrane [2005]) also has some forecasting power for next month's market returns, which becomes stronger at longer horizons of one to five years, as the R-squared of the forecasting regression rises with the extension of the forecasting horizon. We include a variety of variables that the literature has demonstrated to work at various frequencies, and we combine them to extract more information than is generated by univariate forecasting regressions.

We study return predictability along several novel dimensions. We utilize many predictors from the predictability literature and combine them to produce a better forecast. Many previous studies address predictability in isolation, running univariate forecasting regressions. We know many candidate variables that may forecast the equity premium, but it is unclear if they all carry different amounts of information or if they approximate some small set of state variables that govern future investment opportunities. We show that different predictor variables contain different information about future returns at various horizons.

By combining predictors with diverse characteristics, we can produce a superior return forecast. Like Welch and Goyal [2008], we look at a number of different forecasting variables. Unlike Welch and Goyal [2008], we examine the joint forecasting power of all of these variables, and find multiple predictors outperform univariate forecasting regressions. Rapach, Strauss, and Zhou [2010] argued that forecast combination using multiple predictors outperforms the historical average. Our article is similar in that we also combine the information contained in multiple variables, but we look at a broader set of variables (including technical indicators, macroeconomic variables, return-based predictors, price ratios, commodity prices, etc.) and we combine them using correlation screening (Hero and Rajaratnam [2011]).

Correlation screening is a simple way of combining multiple predictor variables that does not depend on the number of predictors. Because correlation screening treats predictors separately and does not require estimating the predictor covariance matrix, it continues to be a feasible technique even as the number of predictors grows large.

In our forecasting setting, we run into the overlapping data problem of Hodrick [1992] because we are forecasting future six-month returns, but our data are at the daily frequency. Our forecasting target contains overlapping time periods for adjacent observations. This problem causes many standard variable selection techniques such as information criterion or stepwise selection to perform poorly. Hodrick [1992] offered standard error computations, but his derivation under the null of no predictability is somewhat restrictive. Although we could have used LASSO or Elastic Net, cross-validation for overlapping data is nontrivial. Under these considerations, we use correlation screening to get around the overlapping data problem.

We show it is possible to forecast medium-term market returns. The return predictability literature has put much focus on predicting returns one or more years into the future. There is also a large literature on the short-term forecastability of market returns at the daily or weekly frequency. We find that we are able to predict market returns for the next six months—between the long-term and short-term horizons. The focus on six months is unique to our work. We find that we can forecast returns well enough to implement our statistical results as an investment strategy.

We illustrate the economic magnitude of return predictability through simulation of trading strategies based on expected returns forecasts. A good yardstick to measure return predictability is to ask the question "Can investors make a profit trading on the predictability?" If the answer is yes, then return predictability is economically important, at least for those who have the resources to implement a market-timing strategy. A simulation from June 8, 2001, through May 4, 2015, shows that taking daily positions in the SPDR S&P 500 ETF Trust (SPY) proportional to the estimated expected risk premium results in an annual return of over 12%, with a Sharpe ratio of 0.85. The annual return is more than twice that of the buy-and-hold strategy, with a Sharpe ratio four times as high in the same period. Through combining variables and using daily data, we can forecast market returns well enough to earn excess risk-adjusted returns. Using our return-forecasting model, we obtain a slight advantage in predicting market returns, and we systematically bet many times to realize this edge.

Most studies on return forecasting stop at statistical results, and their authors do not touch on real-world issues that may prevent investors from fully capturing the benefit of predictability. Through our implementation of the market-timing strategy, we stress the importance of taxes, transaction costs, and other implementation difficulties, which can erode the profitability of the strategy. Among practitioners, many "smart beta" products create some alpha, but the alpha is typically eroded by taxes and sometimes by transaction costs. Our market-timing strategy faces the same problem, so it is important to carefully consider the impact of taxes and trading costs.

There are several shortcomings to the current state of literature on return predictability. Previous studies often restrict the return series to monthly data. Although higher-frequency data have been available for many years, it is messy to deal with data at different frequencies. Previous work preferred to obtain clean statistical results rather than to sacrifice some rigor to create a system that works well in practice. Our primary focus is to create a system that is implementable, so we willingly deal with predictors designed to capture different frequency returns. Many studies examine return predictors in isolation. Some studies, such as Rapach, Strauss, and Zhou [2010], have attempted to combine information across predictors, but they used only a small set of predictors restricted to a similar time horizon. Instead, we look at a relatively large set of predictors, and combine them in sensible ways to produce better forecasts than they do separately. Previous studies often rely exclusively on ordinary least squares (OLS) in forecasting regressions.

Many economic decisions require the input of an estimated equity premium. Superior decisions can be made based on a better forecast of future market returns. Individual and institutional investors both face the problem of asset allocation, for which a good estimate of the equity premium is strongly desired. Traditional investment advice is that market timing is hopeless and investors should seek to keep a constant split between stocks and bonds instead of strategically changing the proportions. At the 2013 Rebalance IRA Conference (Center for Retirement Investing), Burton Malkiel stated "Don't try to time the market. No one can do it. It's dangerous."

Market timing is also related to active management. Passive funds often beat active ones, and mutual fund managers who do well in one year are no more likely to do well in the following year (Berk [2005]; Carhart [1997]). During the recent financial crisis, our investment fund adjusted our portfolio by investing more in equities as the market declined, but our overall performance was less than stellar. To time the market, we need sufficient evidence that actively managing the portfolio will beat passively investing in the index.

Return predictability does not necessarily imply inefficient markets. In general equilibrium models in which the market is perfectly efficient, asset returns may still be predictable (Bansal and Yaron [2004]; Campbell and Cochrane [1999]; and Zhou and Zhu [2015]). Indeed, predictability is consistent with time-varying expected returns driven by changing risk quantities or changing compensation per unit of risk, both of which are possible under efficient markets.

DATA AND VARIABLES

This section describes our forecasting variables and data sources. We draw heavily on the previous work on return predictability. The literature on return predictability is voluminous yet controversial. There are many voices on both sides of the argument. Detractors of return predictability commonly cite Welch and Goyal [2008]; supporters frequently cite Cochrane [2008]. We include well-known variable proposed in the literature by supporters. We also include variables that have previously worked but do not work now. The goal is to have an accurate picture of performance in real time.

For some of the variables, we use their raw values in forecasting returns. For others, we transform the variables into an exponential moving average (EMA) or the log of the raw values minus their EMAs. The EMA of a raw variable creates a persistent series that captures a slow-moving component of market returns. Log of the raw value minus its EMA is similar to a statistical innovation, which may capture a short-term component in market returns. For all the variables, we examine the forecasting performance of the raw values and various transformations, staying true to the form proposed in the original studies whenever possible. We consider the following variables:

1. **Dividend–Price Ratio (DP).** Campbell and Shiller [1988a, 1988b] have shown the dividend–price ratio can be used to forecast future market returns. If the current dividend–price ratio is high, future returns are also likely to be high. We use the log of a 12-month moving sum of dividends paid on the S&P 500 Index minus the log of S&P 500 prices.

2. **Price-to-Earnings Ratio (PE).** Graham and Dodd [1934] used PE as an indicator of value. Campbell and Shiller [1988b] reported that the PE ratio explains as much as 40% of future returns. A high price-to-earnings ratio today indicates a low equity premium. We use the price divided by earnings over the last 12 months.
3. **Book-to-Market Ratio (BM).** Pontiff and Schall [1998] proposed using the book-to-market ratio of the Dow Jones Industrial Average (DJIA) to predict market returns. A high current book-to-market ratio indicates high future market returns. We use the book value of the S&P 500 divided by the S&P 500 Index, SPX.
4. **Cyclically Adjusted Price to Earnings Ratio (CAPE).** This is also known as the Shiller PE. Shiller [2000] used CAPE, price divided by the average inflation-adjusted earnings over the last 10 years, as a predictor of future returns. We use the same definition as Shiller [2000].
5. **Principal Component of Price Ratios (PCA-price).** Since the four price ratios DP, PE, BM, and CAPE all involve prices and are highly correlated, we take the largest principal component of these variables as a predictor to avoid multicollinearity.
6. **Bond Yield (BY).** Pastor and Stambaugh [2009] suggested using the negative value of the difference between the 30-year Treasury bond yield and its 12-month moving average as a return predictor. A high value of BY forecasts lower future returns. We use the 10-year Treasury bond yield divided by the bond yield EMA.
7. **Default Spread (DEF).** Fama and French [1989] proposed using the difference between the Baa and Aaa corporate bond yields as a measure of short-term business conditions. DEF is related to discount rates effects at the business cycle frequency. If DEF is high, expected returns are also high. We use the difference between Baa yield and Aaa yield.
8. **Term Spread (TERM).** Fama and French [1989] also put forward using the difference between the yield on Aaa bond portfolio and the one-month Treasury bill rate as a variable to track the business cycle. They found TERM tracks time-varying stock returns. If TERM is high today, future discount rates are high and the equity premium is also high. We use the yield difference between the 10-year Treasury note and the three-month Treasury bill.
9. **Cointegrating Residual of Consumption, Assets, and Wealth (CAY).** Lettau and Ludvigson [2001] proposed using the cointegrating residual of log consumption, assets, and wealth as a return predictor. The idea is that the cointegrating residual is stationary, and the information they contain may be correlated with discount rates. They find a larger CAY value today indicates that future returns are high, and CAY outperforms the dividend yield at the one-year horizon. We use the original definition of CAY in our exercise.
10. **Sell in May and Go Away (SIM).** Bouman and Jacobsen [2002] and Doeswijk [2009] believed that vacation timing and optimism for the upcoming year create lower returns during the summer months and higher returns moving into the coming year. They find market returns are, on average, lower from May to October and higher from November to April. We use our version of SIM = d/130, in which d is the number of days in the next 130 business days that lie between the second business day in May and the 15th business day of October.
11. **Variance Risk Premium (VRP).** Bollerslev, Tauchen, and Zhou [2009] showed that short-term to intermediate-term returns can be predicted by the VIX squared minus the five-minute realized variance. A high-variance risk premium is associated with high future returns. We use VIX minus the volatility forecast from a GARCH-style model incorporating the Yang and Zhang [2000] estimator using the open, high, low, and close data.
12. **Implied Correlation (IC).** Driessen, Maenhout, and Vilkov [2013] found the average equity options-implied correlation is able to forecast the equity premium. A high IC leads high future returns. We use the CBOE S&P 500 Implied Correlation Index, which measures the expected average correlation of price returns of the 50 largest components of SPY.
13. **Baltic Dry Index (BDI).** Bakshi, Panayotov, and Skoulakis [2011] showed that the three-month change in the BDI predicts intermediate

returns in global stock markets, both in-sample and out-of-sample. Higher BDI growth rates indicate more robust macroeconomic activities and point to higher future stock returns.

14. **New Orders/Shipments (NOS).** Jones and Tuzel [2012] found that high levels of the ratio between new orders and shipments of durable goods are able to forecast excess market returns. Higher levels of NOS are associated with business cycle peaks and forecast lower excess returns on equities. Both new orders and shipments are subject to revision. To see how this variable would have performed in real time, we have gone back to get the originally reported numbers. Our variable is the log of the originally reported new orders divided by the original shipments.
15. **Consumer Price Index (CPI).** Campbell and Vuolteenaho [2004] argued that stock mispricing can be explained by inflation. We use the change in CPI over the last 12 months as the measure of inflation.
16. **Ratio of Stock Price to Commodity Price (PCR).** Black et al. [2014] showed they are able to forecast future returns using the log of the ratio between the stock price and commodity price, measured using the S&P GSCI. PCR is essentially another price ratio, which has commodity price in place of the usual fundamental variable. If PCR is high, expected returns are low. We follow their approach and use log of the ratio between SPY and GSCI.
17. **Moving Average (MA).** Faber [2007] proposed buy-and-sell rules based on the relative levels of the current price versus the past 10-month simple moving average. If the current monthly price is higher than the trailing 10-month moving average, it is a buy signal, and future market returns are expected to be high. We follow Faber [2007] in constructing our MA measure.
18. **Principal Component of Technical Indicators (PCA-tech).** Neely et al. [2014] used principal component analysis to show that macroeconomic variables best identify a rising equity premium near business-cycle troughs, and technical indicators best identify a declining equity premium near business-cycle peaks. If the current principal component value is high, expected returns are also high. We follow their approach and use the first principal component of a set of technical indicators to forecast future returns.
19. **Oil Price Shocks (OIL).** Casassus and Higuera [2011] found that oil price changes are a strong predictor of excess stock returns at short horizons. If OIL is high, future returns are expected to be low. OIL is constructed as the log of the current front oil futures price (CL1) minus the log of the fourth futures price (CL4) with a three-month lag.
20. **Short Interest (SI).** Rapach, Ringgenberg, and Zhou [2015] proposed using the average of short interest divided by total shares outstanding of individual stocks as a return predictor. High current SI indicates the equity premium is low. We use our definition of SI, which uses the sum of all shares short on the NYSE divided by the average daily trading volume over the past 30 days.

The bulk of the data we use comes from publicly available sources. We obtain the necessary data to construct DP, PE, BM, BY, DEF, TERM, CAY, SIM, VRP, IC, BDI, PCR, MA, PCA-tech, OIL, and SI. CAPE is constructed with data from Bloomberg and the Federal Reserve Bank of St. Louis. NOS is from the U.S. Census Bureau, and CPI is from the Federal Reserve Bank of St. Louis. Short interest data from Rapach, Ringgenberg, and Zhou [2015] are kindly provided by Matt Ringgenberg, although in our results we use our own definition of SI. We use the difference between the realized returns on SPX from Bloomberg and 90-day Treasury bill as our forecasting target.

Exhibit 1 presents the pairwise correlations of the forecasting variables. This table highlights the diversity of our variables. Many variables apparently carry information that is weakly correlated with other variables. For example, OIL is positively correlated to BY and BDI, but only mildly correlated with most of the other variables. BDI is positively correlated with OIL and has low correlations with the rest of the variables. BY is also not highly correlated with most of the variables, except OIL. The price ratios at the upper-left corner of the table are all highly correlated or highly negatively correlated, depending on whether the price is in the numerator or the denominator. Since the four price ratios—DP, PE, BM, and CAPE—contain similar information, in our forecasting models we have tried including all four

Exhibit 1

Correlation Matrix of Predictor Variables

	DP	PE	BM	CAPE	PCA-Price	BY	DEF	TERM	CAY	SIM	VRP	IC	BDI	NOS	CPI	PCR	MA	PCA-Tech	OIL
DP																			
PE	–0.38																		
BM	0.48	–0.76																	
CAPE	–0.59	0.75	–0.96																
PCA-Price	–0.63	0.92	–0.93	0.97															
BY	–0.03	0.08	–0.12	0.13	0.10														
DEF	0.15	–0.39	0.53	–0.48	–0.41	–0.16													
TERM	0.22	–0.12	0.49	–0.52	–0.44	0.17	0.25												
CAY	0.42	0.15	–0.02	–0.07	0.03	0.07	–0.08	0.12											
SIM	–0.15	0.04	–0.07	0.05	0.07	0.21	–0.03	0.06	0.07										
VRP	0.05	–0.08	0.22	–0.17	–0.07	–0.19	0.54	0.04	0.16	–0.13									
IC	0.12	–0.16	0.07	–0.14	–0.15	–0.23	0.36	–0.06	0.12	0.01	0.38								
BDI	–0.09	0.06	–0.07	0.06	0.08	0.11	–0.12	–0.03	0.05	–0.03	0.11	–0.09							
NOS	–0.14	–0.19	–0.15	0.15	0.00	–0.01	–0.32	–0.32	–0.32	–0.05	–0.39	–0.04	–0.12						
CPI	0.08	0.06	–0.20	0.16	0.04	–0.09	–0.21	–0.18	–0.13	–0.02	–0.39	–0.04	–0.15	0.35					
PCR	–0.65	0.60	–0.84	0.86	0.87	0.05	–0.21	–0.36	–0.16	0.03	0.02	0.03	0.02	0.02	–0.05				
MA	0.00	0.11	–0.21	0.25	0.12	0.17	–0.54	–0.16	–0.09	0.00	–0.41	–0.41	0.03	0.23	0.11	0.05			
PCA-Tech	0.02	–0.05	–0.07	0.13	0.02	0.25	–0.48	–0.15	–0.06	0.05	–0.38	–0.38	0.00	0.22	–0.05	–0.06	0.80		
OIL	–0.19	0.08	–0.09	0.11	0.14	0.35	–0.21	0.00	–0.01	0.14	–0.09	–0.11	0.29	0.00	–0.06	–0.09	0.04	0.06	
SI	0.14	–0.15	0.18	–0.22	–0.17	–0.10	0.34	–0.05	0.03	–0.03	0.13	0.21	–0.01	0.05	0.23	–0.34	–0.31	–0.25	0.05

Notes: This table shows the pairwise full-sample correlations among the forecasting variables. Darker cells indicate stronger correlations.

separately or including the first principal component of these series, PCA-price, in place of the four series.

MA and PCA-tech are both technical indicators and are 0.80 correlated. Interestingly, PCR (ratio of stock price to commodity price) is highly correlated with the price ratios. Commodity price in the denominator appears to serve a similar role as the fundamental variables in the price ratios—dividends, earnings, or book value. VRP and IC appear to contain information about credit markets, as they are 0.54 and 0.36 correlated with DEF. The technical indicators MA and PCA-tech are negatively correlated with these variables.

Exhibit 2 shows correlations among the predictor variables and future market returns. Correlation is one measure of how well each predictor variable would do in univariate forecasting regressions. Two key observations are evident. First, predictor variables are related to future returns in different ways. Some variables, including DP, BM, CAY, VRP, BDI, MA, and PCA, have positive correlations with future market returns at all horizons. Other variables, including PE, CAPE, SIM, NOS, CPI, PCR, and SI, are negatively correlated with future market returns at all horizons. Still, some other variables, such as BY, DEF, TERM, IC, and OIL, may have positive or negative correlation with future returns depending on the horizon. Second, different variables forecast returns at different horizons. The slow-moving price ratios DP, PE, BM, and CAPE all have stronger correlations at longer horizons. Other predictors such as CAY, NOS, MA, PCA, OIL, and SI exhibit the same pattern. However, some predictors appear to work better for shorter horizons, and their forecasting power weakens at longer horizons: DEF, SIM, VRP, IC, and BDI. Combining predictors that forecast different horizon returns should give us a superior forecast compared with using predictors that all forecast the same horizon.

Many previous studies on return predictability tend to focus on expected returns at the business-cycle frequency—one to five years—and forecasting variables are designed to capture this variation. For example, DP, PE, BM, and CAPE forecast one-year returns more strongly compared with one-month returns. DEF and TERM are specifically chosen to coincide with business cycle peaks and troughs and the equity premium variation associated with those. Less focus has been put on short-term variables that attempt to capture expected return variation in the next days or weeks, because short-term returns contain much more noise, and reliable statistical evidence is harder to establish.

EXHIBIT 2
Correlations between Predictors and Future Returns

	R_1M	R_3M	R_6M	R_12M
DP	0.07	0.14	0.21	0.32
PE	–0.08	–0.15	–0.21	–0.25
BM	0.07	0.11	0.19	0.26
CAPE	–0.06	–0.09	–0.16	–0.24
PCA-price	–0.08	–0.13	–0.19	–0.28
BY	–0.05	–0.05	–0.04	0.06
DEF	–0.06	–0.09	–0.04	0.02
TERM	–0.03	–0.06	–0.04	0.08
CAY	0.11	0.19	0.30	0.45
SIM	–0.04	–0.13	–0.15	–0.02
VRP	0.17	0.32	0.29	0.24
IC	0.09	0.12	0.08	–0.03
BDI	0.10	0.22	0.14	0.03
NOS	–0.08	–0.17	–0.20	–0.25
CPI	–0.15	–0.27	–0.32	–0.29
PCR	–0.02	–0.03	–0.08	–0.16
MA	0.11	0.20	0.21	0.21
PCA-tech	0.11	0.18	0.24	0.27
OIL	0.03	0.04	–0.04	–0.13
SI	–0.14	–0.24	–0.28	–0.30

Notes: Pairwise correlations between predictor variables and future one-month (R_1M), three-month (R_3M), six-month (R_6M), and 12-month (R_12M) market returns. Darker cells indicate stronger correlations.

We seek to combine differential information regarding financial markets. We do so by jointly examining variables that likely contain distinct information sets. Many of the predictor variables we use contain information about the macroeconomy: DP, PE, BM, CAPE, BY, DEF, TERM, CAY, BDI, NOS, and OIL. This should not be a surprise, as the macroeconomy and asset returns are intimately linked. Although all of these variables contain information about the macroeconomy, they do not all reflect the same information. DP, PE, BM, CAPE are classic price ratios often used to gauge where the economy is in the business cycle. Since they contain similar information and are highly correlated, we replace the price ratios with PCA-price in our model. BY and TERM contain information about the bond market. DEF contains information about the credit cycle. CAY gauges how closely key macroeconomic variables are moving together. BDI and NOS are more direct measures of the real economy. OIL is a measure of oil price shocks.

Information contained in inflation measures is also helpful in forecasting equity returns, but this is not necessarily so for information contained in the variables measuring the macroeconomy. We augment our information set by including a transformation of CPI as a measure of inflation. Of course, the real economy and inflation are not independent, so it is possible CPI may contain information about the macroeconomy as well, and the macro variables contain information about inflation. Variables that contain direct information about financial markets will enlarge the forecasting information set. We include SIM, VRP, IC, PCR, MA, PCA, and SI to gauge the performance of future returns from a different perspective than the macroeconomy. VRP and IC use information in derivative markets, and MA and PCA-tech are technical indicators. PCR incorporates information from commodity markets. SI examines investor behavior by looking at how bearish they are. Finally, we include information about international trade with BDI.

FORECASTING RESULTS

Many studies of return predictability focus on using individual variables in univariate forecasting settings to predict future market returns. As we have stressed, we combine information using variables that are likely to contain different information sets, in an attempt to use as much information as possible to produce the most accurate signal for future returns. Many return predictors that have been proposed in the previous literature have similar predictive accuracy, and it becomes difficult to identify a single best forecast from a set of candidate forecasts. Combining forecasting variables creates diversification gains and model stability in sample and out of sample (Timmermann [2006]).

We run simulations of the portfolio performance based on our market-timing model.[1] There is potentially a look-ahead bias, since some variables were only discovered after the simulation start date. Including those variables at the beginning of the simulation would assume prescient knowledge of these return predictors. We repeat our analysis including a return predictor only after its discovery, to alleviate any look-ahead bias.

In combining return predictors, we uncover better forecasting results, and larger economic significance, compared with using return predictors individually. Of course, we are not the first to combine return predictors. Rapach, Strauss, and Zhou [2010] combined individual return forecasts and found the combination delivers statistically and economically large out-of-sample gains compared with the historical average. We use a larger set of return predictors that likely cover a broader information set and illustrate the large economic gains from timing the market.

We look for medium-term return forecasts. The forecast target is the upcoming 130-day market return. We first determine the best transformation of each forecasting variable by maximizing the correlation between the transformed variable and the forecasting target. Transformations include the raw value, an exponentially weighted moving average, and log value minus its exponentially weighted moving average. Specific transformations are determined by the maximal correlation, using previous published work as a guideline. Every 20 days beginning June 2001, we use a training period of 10 years to estimate model coefficients, either with fixed-variable transformations or transformations that maximize correlations with 130-day future returns subject to sign constraints (Campbell and Thompson [2008]). For the next 20 days, we calculate expected returns using the estimated coefficients, and take a position eight times the expected equity premium. The parameters we use (20 days, 10 years, 130 days, and eight times expected returns) are robust: Other combinations that we have tried give us similar results.[2]

Our first forecasting model is a simple kitchen-sink regression, which includes all of the return predictors except the price ratios, which we replace with PCA-price, for a total of 16 variables.

$$R^{e}_{m,t\to t+130} = \alpha_{KS} + \beta'_{KS}\mathbf{x_t} + \epsilon_{KS,t\to t+130} \tag{1}$$

where

$$\mathbf{x_t} = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \cdots \\ x_{16,t} \end{bmatrix} = \begin{bmatrix} PCA-price_t \\ BY_t \\ \cdots \\ SI_t \end{bmatrix} \tag{2}$$

We fit our model every 20 days to obtain parameter estimates, which we use with updated return predictors each day for the following 20 days to produce expected equity premium forecasts. We then take positions in SPY proportional to our return forecasts. This process is repeated every 20 days.

Exhibit 3 plots the wealth evolution of $1 invested in a market-timing strategy based on the kitchen-sink model, cash, or buy-and-hold SPY. At the end of our sample, kitchen sink and buy-and-hold SPY have similar cumulative returns. It is notable that through the two large market downturns we experience during this period, in 2002 and 2008, the kitchen-sink model would have kept us from large drawdowns as the overall market experienced. In fact, during the large downturns the kitchen-sink-based strategy adjusts the position to be negative as the six-month forecast implies medium-term future returns are likely to be low or negative.

The lower panel in Exhibit 3 displays the positions taken by the kitchen-sink model. One hundred percent indicates a buy-and-hold strategy. As expected, the market exposure is generally positive, as the market tends to go up on average. During the recent financial crisis there was an extended period in which the position taken by the model was negative, indicating our model was able to capture the falling market as it was happening. In our implementation, we do not adjust our position each time the return forecast changes, but only when the changes exceed 10%.

The kitchen-sink model does not outperform buy-and-hold in this period in terms of returns. This is probably because naively dumping all of the variables into a linear model increases the likelihood of overfitting in sample, such that out-of-sample forecasting power actually deteriorates. To remove some of the noise in the forecasting variables, our second model uses correlation screening in selecting individual forecasting variables: Using a look-back period of 10 years, we keep only those variables that have at least a 10% correlation with the upcoming 130-day returns. The correlation-screening model is

$$R^{e}_{m,t\to t+130} = \alpha_{CS} + \beta'_{CS}\tilde{\mathbf{x}}_{\mathbf{t}} + \epsilon_{CS,t\to t+130} \tag{3}$$

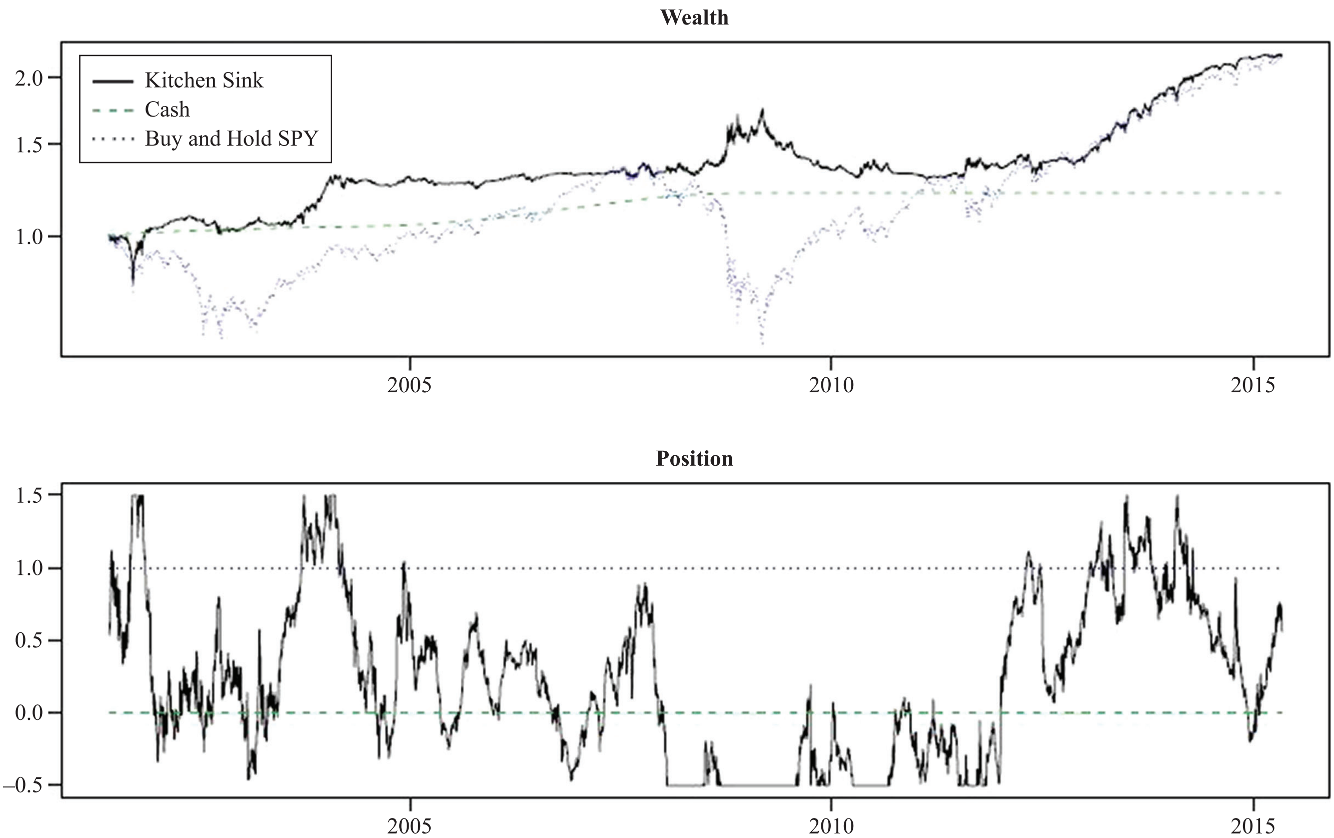
where

$$\tilde{\mathbf{x}}_{\mathbf{t}} = \begin{bmatrix} x_{1,t} I_{|\rho_{1,m}|>0.1} \\ x_{2,t} I_{|\rho_{2,m}|>0.1} \\ \cdots \\ x_{16,t} I_{|\rho_{16,m}|>0.1} \end{bmatrix} \tag{4}$$

$$\rho_{i,m} = Corr(x_{i,t}, R^{e}_{m,t\to t+130}) \tag{5}$$

EXHIBIT 3

Wealth Accumulation and Positions of the Kitchen-Sink Model



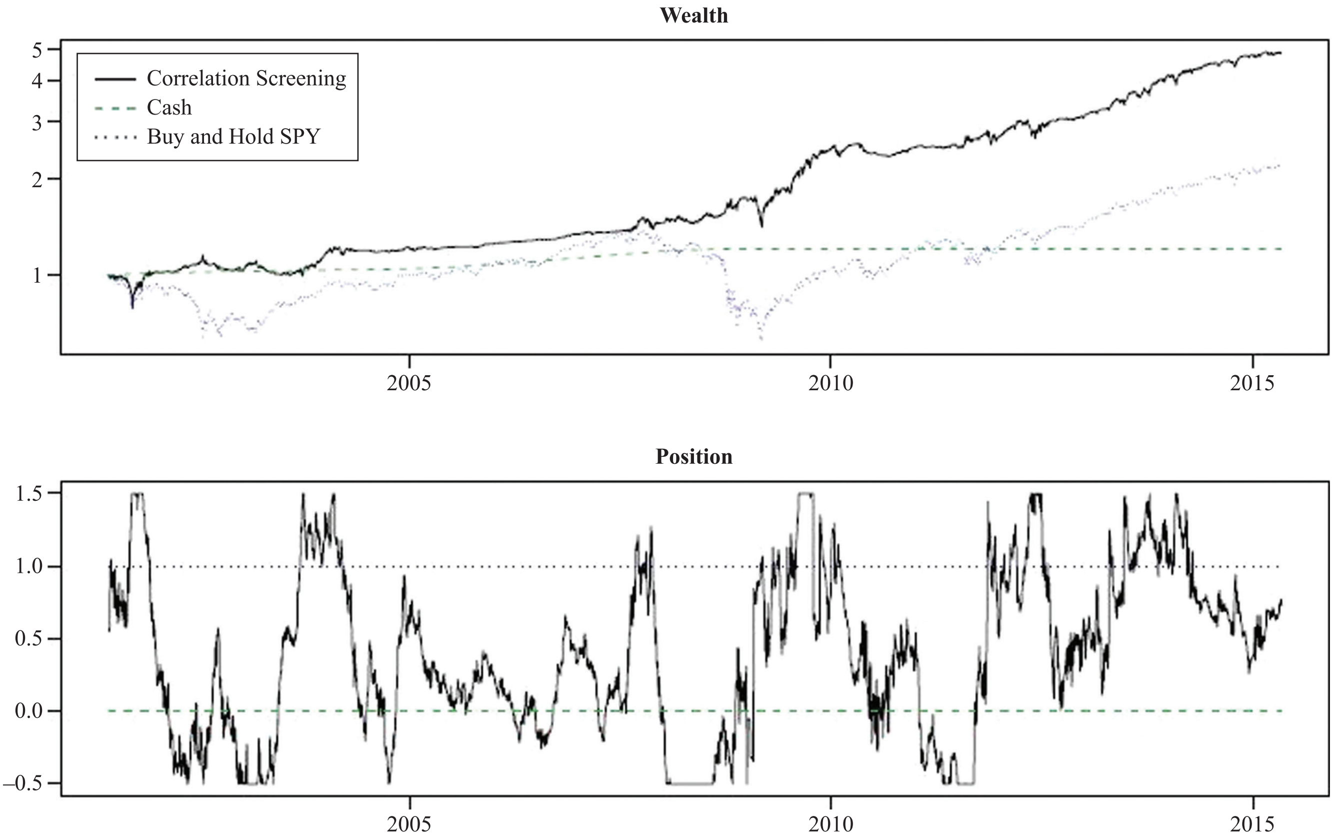
Notes: The top panel plots the cumulative returns ($1 compounded) of the market-timing strategy (solid line) from the kitchen-sink model, buy-and-hold SPY (dotted line), and cash (dashed line). The bottom panel plots the strategy positions, capped at 150% long and 50% short SPY.

Correlation screening is a simple way to build a parsimonious model using the variables with the highest predictive power (Hero and Rajaratnam [2011]). We use a threshold of 10% to select variables that have the highest predictive power for future returns. Other values of the threshold are available in the online appendix at www.iijpm.com.

Exhibit 4 shows the cumulative wealth of $1 invested in the correlation-screening model, cash, or buy-and-hold SPY. The market-timing strategy based on the correlation-screening model outperforms the buy-and-hold strategy. The cumulative returns of correlation screening from 2001 to 2015 are more than twice that of the buy-and-hold strategy. We do not suffer large negative returns in the two large market downturns in 2002 and 2008. The lower panel shows the position in SPY of the correlation-screening strategy. Positions undergo large changes through time and are negative in 2002 and 2008 when the overall market had large negative returns. On the whole, the correlation-screening model performs much better compared with the buy-and-hold strategy as well as the kitchen-sink strategy.

Most of these variables were discovered before our simulation start date of June 8, 2001, but some were discovered afterwards. To guard against look-ahead bias, we repeat our simulation using only variables that are known at the time and add variables after they have been discovered. BDI has been known at least since January 2011. NOS was discovered in December 2008 (private correspondence with Chris Jones). OIL was first mentioned as a predictor of equity returns in 2005. PCR was found in late 2014. PCA-tech was first used as a return predictor in 2010.

EXHIBIT 4
Wealth Accumulation and Positions of the Correlation Screening Model



Notes: The top panel plots the cumulative returns ($1 compounded) of the market-timing strategy (solid line) from the correlation-screening model, buy-and-hold SPY (dotted line), and cash (dashed line). The bottom panel plots the changing positions of the strategy. The strategy is capped at 150% long and 50% short SPY.

We repeat our correlation-screening model to include variables only as they are discovered, called the real-time correlation-screening model.

$$R^{e}_{m,t\to t+130} = \alpha_{RTCS} + \beta'_{RTCS}\,\breve{\mathbf{x}}_{\mathbf{t}} + \epsilon_{RTCS,t\to t+130} \quad (6)$$

where

$$\breve{\mathbf{x}}_{\mathbf{t}} = \tilde{\mathbf{x}}_{\mathbf{t}} \,\big|\, x_{i,t} \text{ has been discovered} \quad (7)$$

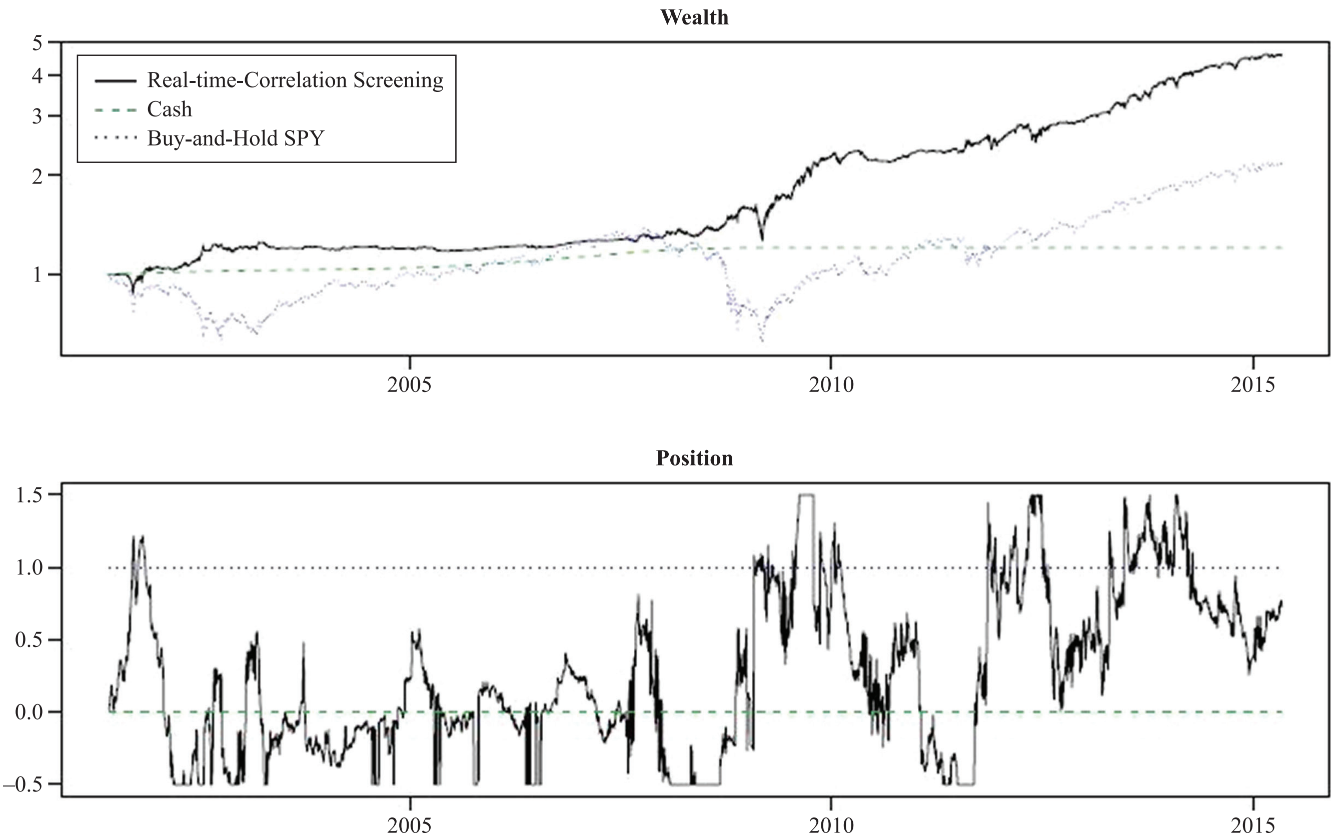
Exhibit 5 shows the wealth accumulation of the real-time correlation-screening model and its positions in SPY. The wealth accumulation process for the real-time correlation-screening model is highly similar to that of the correlation-screening model (Exhibit 4), except the positions taken prior to 2005 for the real-time model are more conservative compared with the correlation-screening model. This indicates the look-ahead bias for the correlation-screening model is small.

To further examine the performance of our model, Exhibit 6 plots the actual returns against the predicted returns. If the forecast were perfect, all of the data points would lie on the 45-degree line originating from the origin (solid line). The left panel is the result for the kitchen-sink model. We see the forecast returns are in a cloud, somewhat correlated with actual returns. The dashed line is the least squares line. The slope of the dashed line is much smaller than one, indicating the expected return forecasts are not close to realized returns.

In contrast, the expected return forecasts from the correlation-screening model do capture considerable

Exhibit 5

Wealth Accumulation and Positions of the Real-Time Correlation-Screening Model



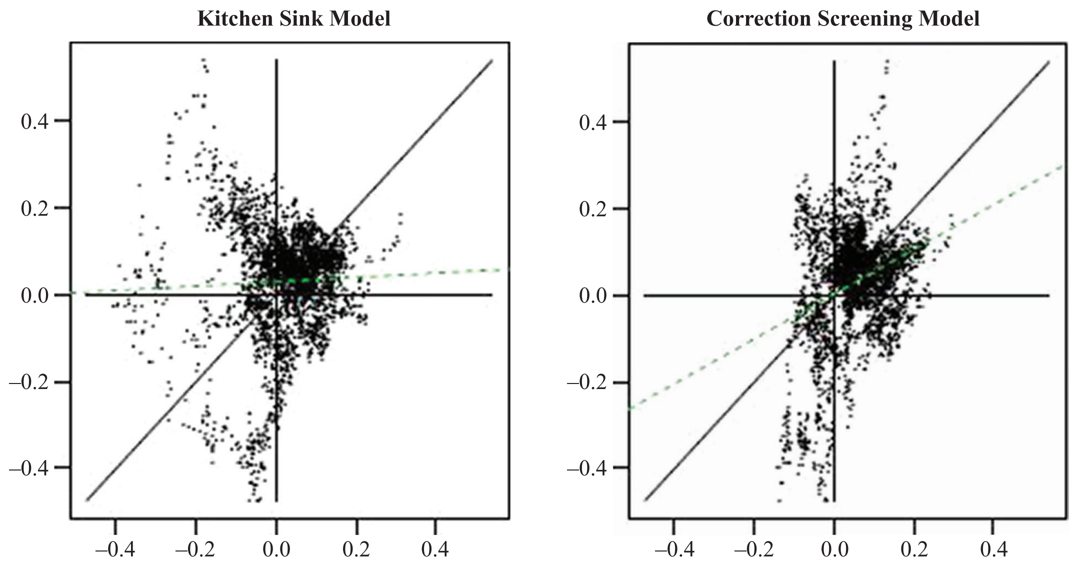
Notes: The top panel plots the cumulative returns ($1 compounded) of the market-timing strategy (solid line) from the real-time correlation-screening model, buy-and-hold SPY (dotted line), and cash (dashed line). The bottom panel plots the changing positions of the strategy. The strategy is capped at 150% long and 50% short SPY.

variation in actual returns, shown on the right panel in Exhibit 6. We see forecast returns are positively correlated with actual returns and the dashed line has a large positive slope. The dashed line is much closer to the solid line compared with the left panel, indicating expected returns from this model are much closer to a perfect forecast compared with the kitchen-sink model. It is evident the correlation-screening model is able to pick up important information about future returns. Data points tend to bunch up in the middle because actual returns are more volatile compared with forecast returns.

Correlation screening is able to produce superior forecasts compared with kitchen sink because it reduces noise and stabilizes the forecasts. By effectively penalizing the least informative variables, correlation screening builds parsimonious models that outperform kitchen sink out-of-sample. Kitchen sink keeps the noisy predictors that do not add much value, whereas correlation screening drops those that add more noise than signal.

As an additional way to understand the gains coming from correlation screening, in the online appendix, we compare equal-weight univariate forecasts and the associated market-timing strategy with those of correlation screening. Rapach, Strauss, and Zhou [2010] and Timmermann [2006] have demonstrated that equal-weight forecasts often perform well out of sample. We find the equal-weight strategy does not sufficiently reduce the noise in the forecasts and performs similarly to the kitchen-sink model.

EXHIBIT 6
Actual vs. Forecast Returns



Notes: Actual returns are on the vertical axis, and forecast returns are on the horizontal axis. The solid line is the 45-degree line on which the data points would lie if forecast returns exactly coincided with realized returns. The dashed line is the best fit for actual data.

IMPLEMENTATION DETAILS

We invest in two assets: SPY and cash. Every day, we take positions in SPY based on our expected market return forecasts. Every 20 days, we refit our model and keep its parameters constant for the next 20 days, at which time the procedure is repeated. Each day at 3:55 p.m. EST, we download the data and use our fitted model to produce a forecast of expected returns. Our desired position is proportional to the expected return forecast, with a cap of 150% (long) and −50% (short) SPY. Orders are submitted to the closing auction at NYSE Arca. Orders for the closing auction must be submitted by 3:59 p.m. EST. As a result of using the closing auction, we receive the settlement price. Our assumption is that we are able to get the closing price at 4:00 p.m. EST and can execute our trade on market close.

Although we could use a number of equity indexes to calculate market excess returns, we implement our strategies with SPY because the S&P 500 market including futures contracts is the most liquid equity market in the world. In 2014, the SPX futures traded \$145 billion and SPY traded \$21 billion per day. The closing auction at NYSE Arca alone averaged more than \$422 million per day. At such depth, it is unlikely that slippage will significantly degrade the returns from our strategies.

Exhibit 7 shows the performance of the three market-timing strategies we consider, along with the buy-and-hold strategy. The correlation-screening model, our benchmark, yields an annual return of 12.11% from 2001 to 2015, compared with 5.79% for the buy-and-hold SPY. The Sharpe ratio 0.85 is four times higher compared with that of SPY. The max drawdown is also much smaller compared with SPY. In our simulations, the average equity exposure is around 60%, and 40% is in cash. Therefore, the volatility of market-timing strategies is less than that of the buy-and-hold strategy.

The real-time correlation-screening model, which adds variables only as they are discovered in the literature, performs almost equally well: 11.66% annual returns, 0.88 Sharpe, and a slightly smaller drawdown. The similarity in performance between the correlation screening and real-time correlation-screening model provides evidence that the look-ahead bias in our correlation-screening model was small to start with. The kitchen-sink model, which naively includes all of the forecasting models, has returns similar to that of the buy-and-hold strategy but a Sharpe ratio about twice as high and a smaller drawdown. It is clear that

EXHIBIT 7

Performance of Market-Timing Strategies, June 8, 2001–May 4, 2015

	KS	CS	RTCS	SPY
Return	5.89%	12.11%	11.66%	5.79%
Sharpe Ratio	0.41	0.85	0.88	0.21
Max Drawdown	26.44%	21.12%	21.83%	55.20%

the market-timing strategies give superior performance compared with the buy-and-hold strategy.

Several comments are in order. Welch and Goyal [2008] argued that the kitchen-sink model including all of the predictors does not perform well. Indeed, kitchen sink without any penalization does not make an attractive market-timing strategy. In Exhibit 6, we see the kitchen-sink forecasts are only weakly correlated with realized returns. After making use of 20 return predictors, we get about the same annualized returns as buy-and-hold SPY. Our results still appear more attractive than those of Welch and Goyal [2008], because we use different return predictors and our sample periods differ. The key missing period in Welch and Goyal [2008] that we include is the recent Great Recession, during which the kitchen-sink model does a good job forecasting the persistently poor market returns.

In fact, because the kitchen-sink model performed well in 2008 and 2009, the associated investment strategy has reduced volatility compared with buy-and-hold and doubles the buy-and-hold Sharpe ratio. As a result, the kitchen-sink model is not a simple straw man. Our comparison between the correlation screening and the kitchen-sink model is useful to help understand the role of parsimony for return predictability.

One issue that we have to address while implementing market-timing strategies is taxes. Past work on return predictability often does not consider the effect of taxes on a market-timing strategy. In practice, most smart beta and tactical asset allocation (TAA) products are able to create some alpha, but the outperformance is often eroded by taxes. Our market-timing strategy suffers from the same problem, and we recommend this strategy be used only by retirement accounts or foundations. We try to keep our transactions to a minimum and do not make small adjustments to our portfolio positions. We also consider transactions costs and assume we pay two cents per share to buy or sell SPY. We assume cash earns daily interest at the three-month T-bill rate minus 30 basis points. We also assume that we pay interest on the shares borrowed at the Fed funds rate plus 30 basis points.

There are additional difficulties in implementing a market-timing strategy. Surely it would be nice to avoid the financial crisis of 2008 or predict the next market boom, but the resources required to implement a market-timing model are greater than one would think. Reliable and timely data sources are a necessity. For many data series, errors exist that need to be corrected before the data can be used. Some series are subject to revisions after they have been made public, which may introduce biases in the forecasting or backtesting results. For our purposes, CPI, NOS, CAY, and earnings are revised. To remove any biases introduced by the data revision, we were able to track down the original reports and use the originally reported values in our forecasting exercise (for CPI and NOS) so that we only use information that would have been known at the time of the forecast. In some cases, we even sent the unrevised numbers back to the original authors to assist them in producing more accurate estimates of obtainable returns.

There is a lot of noise in return forecasts. Aside from data issues, information that may impact expected returns arrives at irregular frequencies. One needs to continuously monitor a large number of factors that may or may not provide information about future returns. A forecasting variable that has been proposed in the past may have worked for some specific time periods but not for other periods. It is inherently difficult to assess whether that means that the result was spurious and the variable does not have any forecasting power or that the result was genuine but the data just had a bad run and the variable may work again in the future.

Investors who wish to time markets must maintain strict discipline and keep emotions out of the investment process. Optimally, the game is to find the right mix of indicators, appropriately assess them, and trade immediately when an opportunity presents itself. A traditional investment committee may meet on the third Thursday following the end of a quarter. Such a structure would be much too slow to react to the much faster pace of market timing. One needs to continually track the market and effectively execute on the tiny signals that sometimes present themselves in a sea of noise. Few retail investors or even professionals have the discipline to act continuously in an unbiased manner.

A market-timing strategy requires a contrarian spirit—selling in hot booms and buying in market downturns. Furthermore, the strategy may not always work, and one must maintain complete faith and continue to trade even if it is currently losing money. The uncertainty may partially explain why so few institutions have tried to build this type of system.

There are other miscellaneous difficulties for investors who want to carry out a market-timing strategy. Many money managers have to worry about what their investors think, and they naturally place more focus on the near term. If the market-timing strategy fails to work for a period of time, these delegated managers may just abandon the strategy because they find it increasingly burdensome to explain the poor results to investors. Another problem is that investors may become very risk averse at the exact time that the market-timing strategy is the most valuable. The last quarter of 2008 and the first quarter of 2009 were great times to increase equity market exposure, but many funds liquidated because they did not want to be invested in risky assets in those periods. The high cost of information acquisition and hiring staff to perform the necessary analysis may be yet another reason why market timing is not more common.

A final comment about implementation: Investors should determine how much risk they are willing to bear and adjust their risk exposures accordingly. The correlation-screening model has a maximum drawdown of 21%. That is a significant loss. An investor should not implement such a strategy unless he is prepared to accept such a drawdown. If our risk tolerance were lower, we could implement the same strategy but scale back the overall exposure.

ECONOMIC SIGNIFICANCE

The acid test for return predictability is if investors can make a profit through timing the market. We have shown it is indeed possible to construct profitable strategies based on market return forecasts. But how well are we doing? Returns of 200 to 300 basis points above the market over a long period of time are exceptional. We have shown that it was possible to gain more than that during the 14-year period from 2001 to 2015. The historical Sharpe ratio during our test period is 0.21, while the long-term Sharpe ratio from 1926 to 2015 is around 0.4. Our market-timing strategy produces a Sharpe ratio of 0.85, so it would seem that investors can significantly time the market. We need a more precise way of measuring our performance, though.

We examine the maximum potential to time the market through estimating the theoretical Sharpe ratio from a six-month forecast. Grinold and Kahn [2000], in their book *Active Portfolio Management*, provide one way to evaluate that question. They provide a calculation for the maximum possible information ratio, assuming investors know with certainty stock returns over the next six months and they trade to maximize their wealth. In our case, the maximum information ratio is 1.59. Since we benchmark against cash, this is also the theoretical maximum Sharpe ratio.

Another way to compute the maximum possible Sharpe ratio of a strategy is to assume perfect knowledge of future returns (private correspondence with Rick Anderson).[3] Anderson uses daily data on the Dow Jones Industrial Average from 1926 to 1996 and assumes investors have perfect information. Each day he calculates the return in the next 252 days. He takes a long position if the return is positive and a short position if the return is negative. He finds the Sharpe ratio of such a strategy is 1.5. We repeat Anderson's exercise for the CRSP value-weighted index with a look-forward period of 130 days. For the period 2011–2014, the maximum Sharpe ratio is 1.15.

Considering the maximum possible Sharpe ratios range from 1.15 to 1.59, our market-timing strategy has a Sharpe ratio that's about two-thirds of the theoretical maximum. Our strategy is capturing a significant amount of the time-varying expected returns that can possibly be captured even with perfect knowledge about future returns. Although we have not employed more sophisticated statistical forecasting techniques, the potential to use such techniques appears limited because the room for improvement in the Sharpe ratio is small.

Although our simulation worked well, it is important to recognize that two significant downturns in 2002 and 2008 contributed to the outperformance of the market-timing strategies over the buy-and-hold strategy. In fact, those two events were two of the three largest cumulative negative returns in the last 100 years (the third being the Great Depression). Our market-timing strategies are designed to outperform in periods of persistently low returns, as we adjust our positions to changing conditions while the market continues to underperform. The rate of outperformance in our simulation must be interpreted with caution because downturns of the magnitude observed in 2002 and 2008 do not occur a mere six years apart.

Another caveat in interpreting our results is an inherent publication bias in academic finance. The publication process favors positive results over negative results, so variables showing predictive power are more likely to be published. The proposed predictors may work well precisely because they have been published—many poor predictors may have been tested and never made public. This data-dredging concern is difficult to address, as we do not observe how many other variables have been tried before we found these 20.[4] Only time will tell if our market-timing models truly outperform: New data are the best out-of-sample test.

CONCLUDING REMARKS

In this article, we revisit return predictability. We examine 20 prominent return predictors proposed in the literature and combine them using correlation screening. We find that we can forecast market returns six months into the future. A market-timing strategy taking positions in SPY proportional to our model estimates outperforms buy-and-hold SPY returns. We illustrate the economic significance of return predictability by simulating a market-timing strategy that makes large economic profits. Furthermore, we discuss the execution details of our strategy, emphasizing various implementation difficulties.

We have addressed return predictability only in a limited setting. By focusing on a previously proposed set of return predictors and a return forecast horizon of roughly six months, we have shrunk the large universe of potential return-forecasting models to a much smaller one. An interesting extension would be to examine return predictability at alternative forecast horizons, especially at one year and five years, for which many of these predictors were first proposed. Such an exercise will readily illustrate the importance of combining information in different return predictors. Another interesting extension is to examine alternate methods of combining forecasting variables. We have used correlation screening. Other potential methods, including stepwise selection, elastic net (Zou and Hastie [2005]), least-angle regression (Efron et al. [2004]), or ensemble methods, may improve forecasting results further.

Throughout this article, we have included what we consider the most prominent return predictors. With the 20 variables we use, we uncover economically significant outperformance compared with the buy-and-hold market. Clearly, there may be other variables with strong predictive power that we have not covered. In fact, we have discovered some proprietary predictors for the equity risk premium that perform well in sample and out of sample. The forecasting performance is even better if we include our proprietary predictors.

If an investor has the ability to reliably forecast market excess returns, then having a constant exposure to different asset classes surely is suboptimal. The investor should increase his exposure to equity when its expected returns are high and decrease his exposure when the expected returns are low. Such practice has been termed tactical asset allocation (TAA). TAA has become pervasive in industry practices, and the academic community is growing increasingly more interested (see Campbell and Viceira [2002]).

As our understanding of return predictability changes, so will the stigma associated with market-timing strategies. Anybody who claimed to implement a market-timing strategy in the past 30 years would have been considered irresponsible, as such a strategy was thought to underperform buy-and-hold returns. In the next 30 years, it is likely that it will be considered irresponsible not to engage in informed market timing. Investors should change their asset allocation as estimates for expected returns change, in order to maximize the long-run growth rate of their investment.

ENDNOTES

We would like to thank Rick Anderson, Petra Bakosova, Mike Fearon, and Jerome Pansera for data analysis, and John Halter and Brian Von Dohlen for contributing to the original project. We thank Rick Anderson, Petra Bakosova, Dirk Eddelbuettel, Petri Fast, Alexios Galanos, Jiahan Li, Jim Lodas, Jerome Pansera, John Rizner, and an anonymous referee for helpful comments. We are grateful to Chris Jones and Matt Ringgenberg for sharing their data and Martin Lettau and Robert Shiller for sharing their data on their website.

[1]Details on how to replicate the results of this paper are available at http://www.ullinvest.com/HI/wp-content/uploads/2015/06/How-to-replicate-A-Practitioners-Defense.pdf.

[2]Robustness results are available upon request.

[3]Rick Anderson, the chief investment officer of Hull Investments, LLC, is the author of *Market Timing Models*. See Anderson [1996].

[4]Even these 20 variables behave differently before and after their discovery. The signs on BM, CAPM, and CPI as predictors changed after the papers were published. The forecasting coefficients on BY and MA differ in magnitude before and after publication. CAY, SIM, VRP, and BDI were robust before and after discovery.

REFERENCES

Anderson, R. *Market Timing Models: Constructing, Implementing & Optimizing a Market Timing Based Investment Strategy*. McGraw-Hill, 1996.

Bakshi, G., G. Panayotov, and G. Skoulakis. "The Baltic Dry Index as a Predictor of Global Stock Returns, Commodity Returns, and Global Economic Activity." AFA 2012 Chicago Meeting Paper, 2011.

Bansal, R., and A. Yaron. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *The Journal of Finance*, Vol. 59, No. 4 (2004), pp. 1481-1509.

Berk, J.B. "Five Myths of Active Portfolio Management." *The Journal of Portfolio Management*, 31 (2005), pp. 27-31.

Black, A.J., O. Klinkowska, D.G. McMillan, and F.J. McMillan. "Forecasting Stock Returns: Do Commodity Prices Help?" *Journal of Forecasting*, 33 (2014), pp. 627-639.

Bollerslev, T., G. Tauchen, and H. Zhou. "Expected Stock Returns and Variance Risk Premia." *The Review of Financial Studies*, 22 (2009), pp. 4463-4492.

Bouman, S., and B. Jacobsen. "The Halloween Indicator, 'Sell in May and Go Away': Another Puzzle." *The American Economic Review*, 92 (2002), pp. 1618-1635.

Campbell, J.Y., and R.J. Shiller. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *The Review of Financial Studies*, 1 (1988a), pp. 195-228.

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

Campbell, J.Y., and J.H. Cochrane. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, Vol. 107, No. 2 (1999), pp. 205-251.

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

Campbell, J.Y., and L.M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, 2002.

Campbell, J.Y., and T. Vuolteenaho. "Inflation Illusion and Stock Prices." National Bureau of Economic Research, Working Paper No. w10263, 2004.

Carhart, M.M. "On Persistence in Mutual Fund Performance." *The Journal of Finance*, 52 (1997), pp. 57-82.

Casassus, J., and F. Higuera. "Stock Return Predictability and Oil Prices." Pontificia Universidad Catolica de Chile, Instituto de Economia, Working Paper No. 406. 2011.

Cochrane, J.H. *Asset Pricing*. Princeton University Press, 2005.

——. "The Dog That Did Not Bark: A Defense of Return Predictability." *The Review of Financial Studies*, 21 (2008), pp. 1533-1575.

Doeswijk, R.Q. "The Halloween Effect in US Sectors: Comment." 2009. Available at https://ssrn.com/abstract=1456357.

Driessen, J., P. Maenhout, and G. Vilkov. "Option-Implied Correlations and the Price of Correlation Risk." 2013. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2359380.

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani. "Least Angle Regression." *Annals of Statistics*, 32 (2004), pp. 407-499.

Faber, M.T. "A Quantitative Approach to Tactical Asset Allocation." *The Journal of Wealth Management*, 9 (2007), pp. 69-79.

Fama, E.F., and K.R. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22 (1988), pp. 3-25.

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

Graham, B., and D.L. Dodd. *Security Analysis: Principles and Technique*. McGraw-Hill, 1934.

Grinold, R.C., and R.N. Kahn. *Active Portfolio Management*. Probus, 2000.

Hero, A., and B. Rajaratnam. "Large-Scale Correlation Screening." *Journal of the American Statistical Association*, 106 (2011), pp. 1540-1552.

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

Jones, C.S., and S. Tuzel. "New Orders and Asset Prices." *The Review of Financial Studies*, 26 (2012), pp. 115-157.

Lettau, M., and S. Ludvigson. "Consumption, Aggregate Wealth, and Expected Stock Returns." *The Journal of Finance*, 56 (2001), pp. 815-849.

Merton, R.C. "On Estimating the Expected Return on the Market: An Exploratory Investigation." *Journal of Financial Economics*, 8 (1980), pp. 323-361.

Moskowitz, T.J., Y.H. Ooi, and L.H. Pedersen. "Time Series Momentum." *Journal of Financial Economics*, 104 (2012), pp. 228-250.

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

Pastor, L., and R.F. Stambaugh. "Predictive Systems: Living with Imperfect Predictors." *The Journal of Finance*, 64 (2009), pp. 1583-1628.

Pontiff, J., and L.D. Schall. "Book-to-Market Ratios as Predictors of Market Returns." *Journal of Financial Economics*, 49 (1998), pp. 141-160.

Rapach, D.E., M.C. Ringgenberg, and G. Zhou. "Short Interest and Aggregate Stock Returns". WFA-Center for Finance and Accounting Research Working Paper, 2015.

Rapach, D.E., J.K. Strauss, and G. Zhou. "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy." *The Review of Financial Studies*, Vol. 23, No. 2 (2010), pp. 821-862.

Shiller, R.J. *Irrational Exuberance*. Princeton University Press, 2000.

Timmermann, A. "Forecast Combinations." *Handbook of Economic Forecasting*, Vol. 1 (2006), pp. 135-196.

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

Yang, D., and Q. Zhang. "Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices." *Journal of Business*, 73 (2000), pp. 477-492.

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

Zhou, G., and Y. Zhu. "Macroeconomic Volatilities and Long-Run Risks of Asset Prices." *Management Science*, Vol. 61, No. 2 (2014), pp. 413-430.