

Unconventional Stock Market Predictors

And the Unstable Performance of Traditional Stock Market Predictors

Daniel Hniopek

22108@student.hhs.se

Abstract

Stock market predictability is a contradictory topic within research in finance. Despite the fact that numerous articles presenting evidence on stock market predictability have been published in the last two decades, Goyal and Welch (2008) argue that most previously suggested models have performed poorly out-of-sample during the last 35 years or so. I confirm these results, and highlight the years around the IT-bubble as a particularly problematic period for traditionally used predictors. Furthermore, I have in this paper examined the performance of some unconventional predictors which measure aspects such as investor sentiment, consumer confidence, development in the manufacturing sector, and early indications of business cycle fluctuations. Analyzing the period of 1978-2010 in search of predictive ability of 1-month, 1-year, and 5-year returns I find that the unconventional predictors in general provide better results both in-sample and out-of-sample, compared to the traditional predictors. Two particular variables stand out – a leading indicator of the business cycle and a consumer confidence index.

Keywords: Predictive regressions, Stock returns, Investor sentiment, Consumer confidence, Leading indicators

Tutor: Francesco Sangiorgi
Date: May 15, 2013

Table of Contents

1. Introduction	3
2. Literature Review	5
2.1 Traditional predictors	5
2.2 Unconventional predictors.....	6
3. Data sources, data construction and data description	7
3.1 Publication lags.....	10
4. Method	11
4.1 Estimation periods	11
4.2 In-sample tests	12
4.2.1 Limitations of in-sample testing	12
4.3 Out-of-sample tests.....	13
4.4 Overlapping observations	14
5. Results	15
5.1 In-sample results.....	16
5.1.1 In-sample, traditional predictors	16
5.1.2 In-sample, unconventional predictors	18
5.2 Out-of-sample results	18
5.2.1 Out-of-sample, traditional predictors.....	19
5.2.2 Out-of-sample, unconventional predictors.....	19
5.3 Further discussion.....	20
6. Implications and conclusions	22
6.1 Future research	24
7. References	25
7.1 Website Data Sources	27
8. Appendix	28

1. Introduction

Stock market predictability is a well-researched topic although the implications of this research are difficult to interpret. Early exploration of the subject dates back to 1920, where Dow (1920) examined dividend yields in a paper that is well-cited in the literature on stock market predictability. The dividend yield is still today one of the most well-tested variables where evidence supporting the possibility of successful forecasting have been found in for example Fama and French (1988) and Hodrick (1992) among others. The list of other variables where there is evidence of predictive ability can be made long, and includes measures such as earnings yields (Campbell and Shiller, 1988), the book-to-market ratio (Pontiff and Schall, 1998), the short-term interest rate and inflation (Fama and Schwert, 1977), the consumption-wealth ratio (Lettau and Ludvigson, 2001), and the share of equity issues in total new equity and debt issues (Baker and Wurgler, 2000).

The theories behind why stock returns may be predictable can mainly be attributed either to rational explanations or to ideas of investor irrationality. The rational view suggests that the equity risk premium varies over time, for example due to time-varying risk aversion (Campbell and Cochrane, 1999). The tested variables then predict returns simply because they capture information about the equity risk premium or the risk aversion. The irrational view relies instead on the assumption that some investors are irrational which creates systematic mispricings that are not fully eliminated due to limits to arbitrage (Shleifer and Vishny, 1997). The underlying causes may be different types of biases related to information processing (e.g. extrapolation bias or overconfidence) or decision-making biases (for example that decisions seem to be affected by how choices are framed, or different variants of mental accounting whereby certain decisions may be mentally segregated from each other and treated in irrational ways). The resulting under- or overpricing may then be predicted by using these variables as means of determining when the stock market deviates from what is normal.

That stock market prediction actually works at all has been questioned however. Data mining is a possibility that should always be considered when a large amount of variables are tested, or even when consulting past research for guidance on which variables to choose (Foster et al., 1997). Some of the variables, mainly those where information about stock prices are included in the regressor, have also been criticized since changes in prices will affect both the predictor variable and the predicted return, violating the OLS assumption of independence at all leads and lags (Lewellen, 2004). Put more generally, the slope

coefficient will be biased if the predictor variable is endogenous in the system that generates returns (Nelson and Kim, 1993). Many of the predictors are also highly persistent over time, which can cause econometric problems. More recent critique of the evidence of stock market predictability can be found in Goyal and Welch (2008) which examines the out-of-sample performance of a large range of previously suggested predictors, concluding that most models seem unstable and have performed poorly during the last 35 years or so.

The purpose of this paper is mainly to test some unconventional predictors in search of more reliable performance, especially during the last decades, than what has been recorded previously. I have however included (yet another) analysis of the performance of some traditional stock market predictors, even though this analysis should be viewed more as means of creating a comparable base from where to evaluate and discuss the performance of the unconventional predictors. This paper will inevitably deal with the question of whether or not stock market returns are predictable, although it should not be seen as an addition to the econometric debate on the subject. There are other articles using more advanced econometric techniques, such as Goyal and Welch (2008) or Campbell and Thompson (2008).

In total, I have tested six traditional predictor variables and five unconventional ones. The traditional variables consist of the dividend yield, the share of equity issues in total new equity and debt issues, and four variations of earnings yields. The unconventional variables on the other hand measure aspects such as investor sentiment, consumer confidence, development in the manufacturing sector, and early indications of business cycle fluctuations. All variables are tested with respect to how well they predict excess returns one month, one year and five years into the future, using monthly data. The problem of overlapping observations that is created when using this approach is corrected for using two alternative methods. In the first method I will use the Newey-West heteroskedasticity and autocovariance consistent (HAC) estimator, and in the second method I will instead rescale the OLS standard error as suggested by Harri and Brorsen (2009). All variables are tested during the period 1978-2010. In addition to that the traditional variables have also been tested during the period 1938-2010, mainly because data availability allows it. Furthermore, all variables have been tested both in-sample and out-of-sample, where the out-of-sample test is based on a set of rolling regressions with an expanding window, designed so that each forecast is made using only data that was available at the time in which the forecast was made.

During the later period (1978-2010), which is the main period of interest in this study, the tests show that there are predictor variables with significant forecasting power in-sample for all tested forecasting lengths. Interestingly, the performance of the unconventional

predictors is better than that of the traditional predictors, both in terms of the number of significant models and in terms of having the models with the strongest significance. Out-of-sample, performance drops drastically however, with only one single variable being significant, and only when forecasting five-year returns. Again, this variable turned out to be one of the unconventional predictors, namely a measure of consumer confidence. These findings suggest that there are measures such as consumer confidence and other leading indicators that should be considered as possible substitutes or complements to the more commonly used models.

2. Literature Review

2.1 Traditional predictors

The previous literature on traditional predictors goes hand in hand with the academic debate on stock market predictability, to which the reader has already been introduced in Section 1. As described, this question is still not resolved and is very much a current topic of research in finance. Reviewing the most recent research of the subject, the influential paper by Goyal and Welch (2008), which criticized out-of-sample performance, introduced an important challenge from where researchers now try to find methods of improving the performance of previous models. Campbell and Thompson (2008) for example argue that the out-of-sample performance of predictive regressions can be substantially improved by imposing weak restrictions on the regressions. They explore two alternative restrictions – that the regression coefficient has the theoretically expected sign, and that the fitted value of the equity risk premium is positive – and conclude that both restrictions improve performance. Ferreira and Santa-Clara (2011) use a fundamentally different approach where three components of stock market returns – the dividend yield, the earnings growth and the price-earnings ratio growth – are instead forecasted independently. They argue that this approach produces better out-of-sample results compared to using the historical mean as the best estimate of future returns, and that these results are economically significant.

Cochrane (2008) uses another line of reasoning by referring to the present value relationship. If both price growth and dividend growth are unpredictable, dividend yield must be constant. Since this does not hold, one of dividend growth or price growth *must* be predictable. Cochrane concludes that the *absence* of dividend growth predictability is strong evidence that stock market price growth is instead predictable.

As already mentioned however, the main purpose of this paper will not be to settle the question of stock market predictability by improving on previous models, but instead to look at the predictive performance of new variables.

2.2 Unconventional predictors

In this paper “unconventional” will be the title for variables such as investor sentiment, consumer confidence and other leading indicators. Overall the research on these types of stock market predictors is patchy, and consists mainly of papers where the authors have examined the performance of one particular variable. This makes the relative performance between different variables hard to evaluate, since different authors use different methods. The amount of high quality papers is also substantially lower, with the exception of research on investor sentiment which has become a popular topic in the wake of behavioral finance.

A main problem concerning unconventional predictors is that there are no universal definitions of them, and no consensus on how to measure them. This is especially apparent when it comes to investor sentiment, which can be measured both directly using surveys or, as popular in previous research, using different proxies such as the level of discounts on closed-end funds, or net mutual fund redemptions. Brown and Cliff (2004) investigate a range of commonly used sentiment measures and conclude that many proxies for sentiment are however strongly correlated with survey-based measures of sentiment. They however find little evidence that investor sentiment can predict short-term stock market returns. Neal and Wheatley (1998) on the other hand test three common proxies for investor sentiment and find that two of them predict the size premium, which is the difference in return between small and large firms. Baker and Wurgler (2006) achieves similar results by creating a composite sentiment index based on six different proxies for investor sentiment and finding that when sentiment is low, subsequent returns are relatively high for certain groups of stocks, such as small stocks, extreme growth stocks and distressed stocks.

Even though consumer confidence and its relationship to the level of household spending have been thoroughly examined, its connection to the stock market is less researched. As with investor sentiment, measurement is a problem and there are several different survey-based indices of consumer sentiment. The most well-established indices that have been studied in previous research are the Thomson Reuters/University of Michigan Index of Consumer Sentiment and The Conference Board Consumer Confidence Index, which monitor the same population (American households) but ask different questions. Fisher and Statman (2003) test both these measures for predictive ability of one, six and twelve months’

return, and find a negative relationship between consumer confidence and subsequent returns. The results were only significant for both indices when examining six-month returns, and the index by University of Michigan generally performed better than the one from The Conference Board. Kalotay et al. (2007) on the other hand tested the predictive ability of three-month returns of consumer confidence as measured by University of Michigan, but found no stand-alone predictive ability.

The last group of unconventional variables consists of leading indicators which are normally designed to help predict GDP and provide early signals of business cycle turning points. Not surprisingly, previous literature is mainly focused on finding new predictors of GDP and evaluating whether or not established indicators actually manage to provide reliable forecasts of GDP. Research on the link between leading indicators and stock market returns is virtually inexistent. Umstead (1977) examines data of the Leading Composite Index that was published by the Department of Commerce (today The Conference Board Leading Economic Index). Analyzing quarterly data between 1948 and 1974 using an out-of-sample approach where the model parameters are estimated in the first half of the period and tested in the second, Umstead finds that the leading indicator does predict stock market returns. The conclusion is that the stock market appears to be overvalued during economic expansions and undervalued during contractions – an inefficiency that is interpreted as being related to how information about the business cycle unfolds over time, and how the stock market use this information. It should however be noted that the composition of the index has been changed since then, possibly rendering Umstead's findings not only out-of-date but also irrelevant.

3. Data sources, data construction and data description

The study will be conducted using monthly observations, and all return figures are expressed in terms of monthly percentages unless stated otherwise. All website data sources are provided in Section 7.1 for the convenience of anyone who wishes to reproduce or extend my work.

Stock returns: the value-weighted return for S&P 500 including distributions, extracted from The Center for Research in Security Prices (CRSP).

Stock prices: monthly averages of daily closing prices for the S&P 500 adjusted to real values using CPI. The data are from Robert Shiller of Yale University, and can be retrieved from his webpage.

Risk-free interest rate: the 10-Year Treasury Constant Maturity Rate (GS10), expressed as a monthly percentage. The data are from Robert Shiller's webpage.

Excess return: Stock returns – Risk-free interest rates.

Earnings: real earnings for S&P 500, where linear interpolation is used to estimate monthly earnings between the quarterly reports. The data are from Robert Shiller's webpage.

Dividends: real dividends for S&P 500, where linear interpolation is used to estimate monthly dividends between the quarterly reports. The data are from Robert Shiller's webpage.

Earnings yield (E_P): earnings divided by stock prices.

Multiyear earnings yields (EX_P): moving average of earnings using X years of past data, divided by stock prices.

Dividend yield (D_P): dividends divided by stock prices.

Percent equity issuing (eq_{is}): a measure that is proposed by Baker and Wurgler (2000) and is calculated as the amount of equity issuing divided by the sum of equity issuing plus debt issuing. The reason for including this particular variable is that it is one of few measures that seems to exhibit some forecasting potential even out-of-sample, according to Goyal and Welch (2008). Data up to April 2008 are from Jeffrey Wurgler of New York University Stern School of Business, and can be retrieved from his webpage. I have gathered the rest of the data from the Board of Governors of the Federal Reserve System and their statistical release New Security Issues, U.S. Corporations (1.46), which is issued monthly.

Investor sentiment ($SENT$): an index based on the common variation in six proxies for sentiment: the closed-end fund discount, NYSE share turnover, the number of IPOs, the average first-day returns on IPOs, the equity share in new issues, and the dividend premium. The data was constructed and used in Baker-Wurgler (2006) and Baker-Wurgler (2007). It should however be noted that the time series has been constructed by standardizing the coefficients for each sentiment proxy over the period 1962-2005. This means that all

observations inherently contain information that was not available until after 2005, making a true out-of-sample test impossible to perform. The data are again from Jeffrey Wurgler's webpage.

The Conference Board Consumer Confidence Index (CCI): The Conference Board is an independent, nonprofit business membership and research association founded in 1916. They have conducted a consumer confidence survey on a monthly basis since 1977 using consistent definitions and questions since its inception. Respondents are asked to appraise present business conditions and employment conditions, and state their expectations regarding future business conditions, employment conditions and personal income. The horizon of the forward-looking questions is six months. The survey is directed towards residential households in the United States, which are first stratified geographically before selecting a random sample that ultimately leads to a sample size of approximately 3000 completed questionnaires. The data can be acquired from The Conference Board.

Thomson Reuters/University of Michigan Index of Consumer Sentiment (ICS): the Surveys of Consumers has been conducted by University of Michigan since 1946, initially once each quarter. Since 1978, the survey is performed monthly by talking to 500 individuals selected randomly from the contiguous United States. The survey asks both for an evaluation of present conditions and for expectations of future conditions, and cover the respondents' personal finances, future buying plans, and their assessment of business conditions. The index is normally used as an indicator of the future economy, and it is included as a component in the OECD Composite Leading Indicators. A subset of the index is also included as one of ten components in The Conference Board Leading Economic Index (originally the Leading Composite Index, compiled and published by the U.S. Department of Commerce). The data are made available at the Federal Reserve Economic Data (FRED) database at Federal Reserve Bank of St. Louis with six months delay. More recent data are issued monthly by Thomson Reuters and can be accessed from their history of press releases related to this index.

Purchasing Managers' Index (PMI): the PMI is a composite index based on data from the manufacturing sector of: new orders, production, employment, supplier deliveries, and inventories. The data are published monthly by the Institute for Supply Management (ISM), which is a nonprofit association founded in 1915 that provides education and research

information to enhance the performance of procurement and supply chain management of practitioners and their organizations worldwide.

OECD Composite Leading Indicators, normalized (CLI): the Composite Leading Indicators (CLI) for USA is published by the OECD and is based on: the number of dwellings started, the net new orders for durable goods, share prices (NYSE), the Index of Consumer Sentiment, weekly hours of work in manufacturing, Purchasing Managers' Index and finally interest rates spreads. The Index of Consumer Sentiment and Purchasing Managers' Index that are included in CLI are the exact same components which I will also investigate individually, as mentioned above. In order to isolate cyclical patterns, all historical data of CLI are revised once a month as soon as new data are published. The revisions are rather small but still make true out-of-sample testing impossible since the time series is altered ex post.¹ The data are from the OECD.Stat Extracts.

3.1 Publication lags

A considerable drawback of using variables that are not based on quickly available financial data is that such measures are often subject to noticeable publication lags. Publication lag is here defined as being present when the release of information occurs after the month to which the information actually relates. Percent equity issuing (eq_is) is subject to a publication lag of around 20-30 days, and will therefore be lagged one month. The OECD Composite Leading Indicators (CLI) has a publication lag of around 40 days, and will be lagged two months. There are two reasons to correct for this. The first relates to the question of practical usability – a measure may be attributable to whichever period in theory, but it is not possible for an investor to use the information until it has been released. The second reason is that the release of information may itself affect the stock market. The eventual short-term profit associated with such an event is not what I am trying to catch in this paper. Still, even when correcting for this problem, there is an element of error involved since publication lag times may have varied over time. This type of information is however very hard to find.

I will not address the question of publication lag for investor sentiment (SENT), since it is not clear exactly how this issue is dealt with by Baker and Wurgler (2006) when constructing the time series.

¹ The particular time series that I have been working with was extracted on March 13, 2013.

4. Method

4.1 Estimation periods

There are two important criteria when choosing estimation period: the first is to have a sample that is long enough to be representative of whatever the future may hold; the second is to have a sample that consists of recent enough data to provide a view of the stock market that coincides with how the market functions today. A sample that violates the first criteria could be one that only covers a period with mainly positive and stable stock market development. The second criteria is important since stock market prices are based on expectations that change over time – simply learning that one strategy will make money could cause this strategy to be utilized to such extent that the opportunity disappears. That means, in this particular context, that measures that may have predicted future excess returns in the past may no longer do so. Joining these two criteria into the selection of a single sample period is difficult, since the two criteria to some extent are mutually exclusive – a period that only consists of recent data will perhaps not be representative of the future. Using a longer period on the other hand may instead produce misleading results if the behavior of the stock market has changed significantly over time.

Another issue that also relates to the subject of having a sample that well describes reality concerns the Great Depression. Most available measures show extreme observations especially during 1932 when the stock market hit rock bottom. One example is percent equity issuing where extreme observations are recorded as late as 1934, with a few values being about 5-10 times larger than any other values observed during the entire sample (1926-2013). Thus, I decided to begin the analysis after the Great Depression, excluding both the extreme downturn as well as the rapid climb during the subsequent years, and instead use data from 1938 and onwards.

Data availability, and the difference in the amount of historical data that is available for the different measures, is also a problem. The data required to analyze earnings yields, the dividend yield and percent equity issuing are available roughly from the beginning of 1927 to the end of 2012. For some of the unconventional measures however, data is not available until the beginning of 1978 and ends at the end of 2010. Analyzing some of the measures using 75 years of data (1938-2012) and some using only 33 years of data (1978-2010) would not give comparable results.

Taking all these concerns into account, I decided to analyze the predictive power of all measures for the period 1978-2010, henceforth mainly referred to as the *later sample*

subset. In addition to that, the measures where a lot more data is available (the earnings yields, the dividend yield and percent equity issuing) will also be analyzed for the period 1938-2010, henceforth mainly referred to as the *full sample*, to provide a long-run view on stock market prediction. Still, any choice of estimation period is necessarily arbitrary and does have the potential for a sample selection bias. Depending on the estimation period results can vary a lot, which this paper will be an example of.

4.2 In-sample tests

All tested measures will be evaluated based on how well they predict the excess returns of three different forecasting lengths – next period's return (one month), next 12 periods' return (one year) and next 60 periods' return (five years).² In other words, the following predictive regressions will be run:

$$r_{t+1} = \alpha + \beta * x_t + u_{t+1} \quad (1)$$

$$r_{t+12}(12) = \alpha + \beta * x_t + u_{t+12}(12) \quad (2)$$

$$r_{t+60}(60) = \alpha + \beta * x_t + u_{t+60}(60) \quad (3)$$

where x_t is the tested predictor at time t .

After having estimated a relationship between the tested variable x_t and subsequent excess returns, we can construct a time series of expected returns as predicted by the model:

$$E_t(r_{t+q}(q)) = \hat{\alpha} + \hat{\beta} * x_t \quad (4)$$

where q is the forecasting length, i.e. 1, 12 or 60. Described in other words, the expectation of the future returns that will materialize during the forecasting horizon will be computed by using the relationship between x_t and subsequent returns, which is estimated in regression (1), (2) or (3). Computing a time series of expected returns is useful, since it makes it possible to directly compare the predicted returns against the realized returns.

4.2.1 Limitations of in-sample testing

Goyal and Welch (2008) argue that in-sample tests are mainly to be relied upon when the underlying model is stable and well-specified. Under this assumption, in-sample tests are more efficient, and can provide valuable information to a researcher who is confident in the

² Note that for 5-year returns, the analysis will only be extended up to 2007 since no more data are currently available.

underlying model specification but uncertain about the underlying model parameters. They continue however by stressing the importance of out-of-sample tests as a diagnostic tool to help determine whether a model actually is stable and well specified, or changing over time. Given that their final conclusion was that there is reason to be highly skeptical about the performance of many previously accepted predictors, especially during the last 35 years, in-sample tests alone do not provide enough robustness to the analyses.

4.3 Out-of-sample tests

An out-of-sample test will give an assessment that to larger extent evaluates the performance that an actual real-time investor would have been able to achieve. The expected return will be calculated in the following way:

$$E_t(r_{t+q}(q)) = \hat{\alpha}_t + \hat{\beta}_t * x_t \quad (5)$$

which is the same method used in (4) but with one adjustment, namely that $\hat{\alpha}$ and $\hat{\beta}$ now have a subscript for time. The model parameters will now be re-estimated *each period* using regression (1), (2) or (3), using only data that was available at time t .³ In other words I will be using a rolling regression technique with an expanding window, so that each forecast is made using only data that was actually available at the time in which the forecast was made. Thus, the investor will update his belief of the true model parameters as new data becomes available. The estimates of expected returns will begin 11 years after the first observation, to make sure that the earliest regressions contain enough data to provide reasonably reliable estimations.⁴

Since one regression is run each period, this method will provide time series of β 's, t -statistics and R^2 's. A final regression has to be run where the realized excess return is regressed on the estimated excess return that was given by equation (5):

$$r_{t+1} = \alpha + \beta * E_t(r_{t+1}) + u_{t+1} \quad (6)$$

$$r_{t+12}(12) = \alpha + \beta * E_t(r_{t+12}(12)) + u_{t+12}(12) \quad (7)$$

$$r_{t+60}(60) = \alpha + \beta * E_t(r_{t+60}(60)) + u_{t+60}(60) \quad (8)$$

³ Available data, in the case of regression (2) and (3), means that the regressions can only use observations up to $t-12$ and $t-60$ respectively, since data for the left hand side, $r_{t+12}(12)$ and $r_{t+60}(60)$, are not known to a real-time investor until 12 and 60 periods has passed.

⁴ 5 years are required due to the issue described in footnote 1. On top of that, the statistics package Stata requires another 6 years of data to be able to perform regressions using the Newey-West estimator.

These regression equations will be the out-of-sample equivalent to regressions (1), (2) and (3). A major difference between the coefficients reported in the in-sample approach versus the coefficients reported in the out-of-sample approach should however be noted. In the in-sample approach, future excess return is regressed on the current value of the predictor variable. The interpretation of a significant negative coefficient is thus that a high value of the variable predicts lower returns in the future – but the variable can still effectively predict future excess returns. In the out-of-sample approach however, future excess return is not regressed on the value of the predictor variable but on the *expected return* that was predicted by the predictor variable. The interpretation of a negative coefficient in this case is thus that the model actually provided high estimates of expected returns when subsequent realized returns turned out to be low, and vice versa.

A fundamentally different method of testing out-of-sample performance is to estimate the model parameters during one time period and test them in another time period. Using this approach, the model parameters will remain stable over the testing period which can of course have an impact on the results. I do however believe that the method I have chosen provides a more realistic view that better describes how actual investors would use the information.

4.4 Overlapping observations

It should be noted that the methods described in Section 4.2 and Section 4.3 will cause observations of excess returns to overlap in those regressions where the forecasting length is longer than one period, i.e. in-sample regression (2) and (3) and out-of-sample regression (7) and (8). This gives rise to some econometric obstacles. Taking regression (2) as an example, the left hand side will always consist of the 12 subsequent observations of excess returns which means that two values following each other in time will always have 11 observations in common. This overlap of observations creates a moving average (MA) error term (Hansen and Hodrick, 1980). When errors are serially correlated, ordinary least squares (OLS) is no longer the best linear unbiased estimator (BLUE) and the OLS standard errors and test statistics are not valid (Wooldridge, 2008). Harri and Brorsen (2009) states a number of various procedures that can be used overcome these problems, such as transforming the data or the regression in different ways to eliminate the autocorrelation in the error term, or using heteroskedasticity and autocovariance consistent (HAC) estimators. According to their research, one of the most commonly used methods in published finance articles is to adjust for autocorrelation using the HAC estimator based on Newey and West (1987). I will report

Newey-West (henceforth NW) corrected standard errors where the lag has been selected to match the level of overlap.

Harri and Brorsen (2009) do however also criticize NW and claim that this procedure can in some cases be very inefficient and provide too high t-statistics. Similar critique is put forth by Goetzmann and Jorion (1993) and Valkanov (2003), where the latter author instead proposes a rescaled standard error and t-statistic specifically designed for the case when overlapping observations are faced in long-horizon regressions, as is the case here. Harri and Brorsen (2009) evaluate several alternative rescaling methods suggested since Valkanov (2003), and propose a minor change to previous literature by rescaling the OLS standard errors by multiplying them with the following factor:

$$\text{If } \frac{k}{n} < 0.10, \quad SE^{HB} = SE^{OLS} * \sqrt{\frac{2}{3}(0.9k - 1)} \quad (9)$$

$$\text{If } \frac{k}{n} > 0.10, \quad SE^{HB} = SE^{OLS} * \sqrt{\frac{2}{3}k \left(0.9 + \frac{k}{n}\right)} \quad (10)$$

where k is the level of aggregation (i.e. 12 for 1-year returns and 60 for 5-year returns), n is the sample size and SE^{HB} is the resulting rescaled standard error according to Harri and Brorsen (2009). This procedure will provide higher standard errors and lower t-statistics compared to OLS for all scenarios that are considered in this paper, which is not surprising. Furthermore, the rescaling factor will be larger when the level of aggregation is large compared to the sample size. The t-statistics associated with these rescaled standard errors will be reported, in addition to the NW corrected standard errors, for all regressions where overlapping observations have been used.

For in-sample regression (1) and out-of-sample regression (6) the forecasting length is one month, which means that there are no overlapping observations. OLS estimates will be used, but with heteroskedasticity-robust standard errors and t-statistics.

5. Results

All results are presented in Table 1 – Table 6. Some general results that can be seen when comparing the results of the different tests are:

- It is clear that predictive power and R^2 normally increases with the forecasting horizon, as emphasized by Fama and French (1988). However, some measures do yield better results for shorter horizons than for longer horizons.

- Only a fraction of the significant models for the full sample (1938-2010) produce reasonable, significant results during the later sample subset (1978-2010).
- Only a few of the in-sample significant models produce reasonable, significant out-of-sample results.
- HB generally provides higher t-statistics than NW when predicting 1-year returns and lower t-statistics than NW when predicting 5-year returns. I will mainly consider a variable to be significant only when demonstrating significant results for both methods.

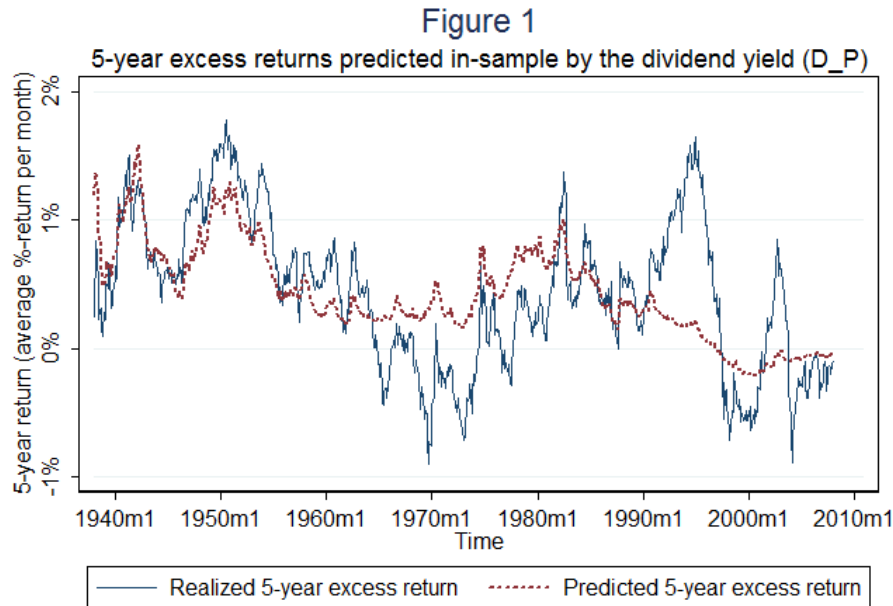
These findings will be commented on in greater detail in the following paragraphs. The presentation of the results will continue as follows. Section 5.1 will discuss the in-sample results which are displayed in Table 1, Table 2 and Table 3. Section 5.2 will then discuss out-of-sample performance which is presented in Table 4, Table 5 and Table 6. In Section 5.3 I will further discuss the stability of the tested models, especially during 1978-2010.

5.1 In-sample results

Table 1 presents in-sample results when testing for predictive ability of 1-month returns. Table 2 and Table 3 present in-sample results when for testing predictive ability of 1-year returns and 5-year returns, respectively. Furthermore, Figure 8 – Figure 40 present scatterplots of all in-sample estimations for all forecasting lengths, during the period 1978-2010 which is the main period of interest.

5.1.1 In-sample, traditional predictors

All tested earnings yields and the dividend yield managed to demonstrate statistically significant forecasting power for all forecasting lengths when analyzing the full sample of 1938-2010. During the later sample subset, only 5-year returns could be predicted with reasonable accuracy, and only by the dividend yield. The predicted 5-year returns as estimated by the dividend yield for the full sample are compared with the realized 5-year returns in Figure 1.



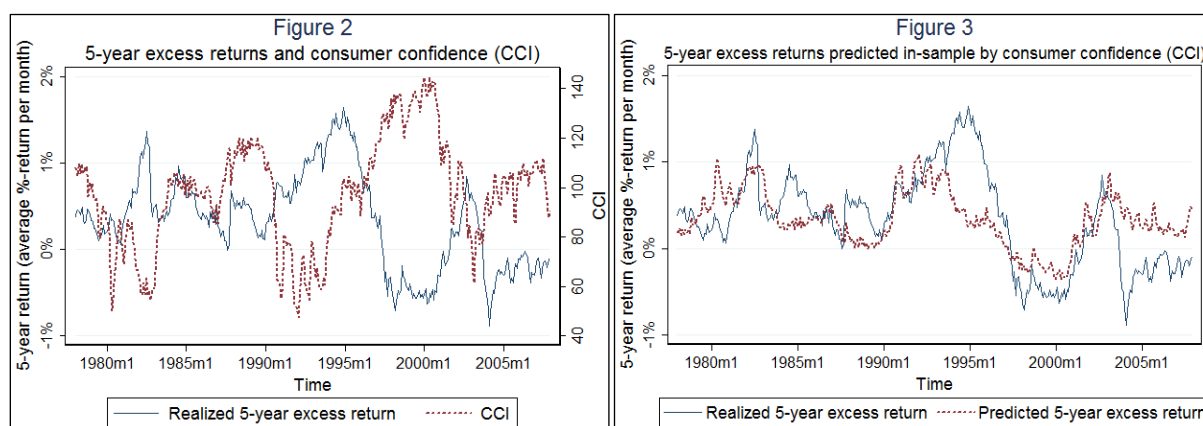
As can be seen, the predicted 5-year returns matches the returns that actually materialized quite well up to 1990. After that, the forecasting power was very limited. This reflects a trend that can be seen for all earnings yields as well, which explains why most variables demonstrate very significant predictive results for the full sample but less so when only analyzing the later period.

Percent equity issuing (eq_is) managed to significantly predict 1-month returns when analyzing the full sample. If the time series would *not* have been adjusted for publication lag, these results would have been significant during 1978-2010 as well. Goyal and Welch (2008) highlight the difference between using eq_is to study the fund-raising behavior of firms or using it as a stock market predictor, and argue that in-sample performance is more important in the first case while out-of-sample performance is more important in the latter. Using a similar argument, I believe that publication lag adjustments should of course not be made if the objective is to study the financing of corporations. However, in a study on stock market prediction, the data should be adjusted according to the information that is actually available to an investor. Thus, I find it strange that Goyal and Welch (2008) have not corrected this measure for the publication lag, even though their work clearly takes an investor perspective with stock market prediction as the main objective. Despite using different methods, I am still able to reconcile my findings with the main conclusion of Goyal and Welch (2008), which is that there is a negative relationship between eq_is and subsequent returns, but that it would not have been possible to profit from this as an external investor during the most recent 35 years or so. Goyal and Welch (2008) are however

more optimistic about eq_is compared to many of the other traditional variables, which I find little reason to be.

5.1.2 In-sample, unconventional predictors

For the unconventional predictors, there were a few different variables that displayed significant forecasting power – Investor sentiment (SENT) for 1-month returns, the OECD Composite Leading Indicators (CLI) for 1-month and 1-year returns, and The Conference Board Consumer Confidence Index (CCI) for 5-year returns. CLI was the best in-sample predictor for 1-month and 1-year returns out of all tested variables during the later sample subset, and CCI was the best predictor for 5-year returns. Interestingly, the coefficients for all these three variables – which inherently measure something positive in the economy – were negative, indicating a clear contrarian pattern. Taking 5-year returns as an example realized excess returns are on average lower when consumer confidence is high, as shown in Figure 2, consistent with stock market overreaction and reversal. Figure 3 shows the corresponding expected 5-year returns as estimated by CCI. A few other unconventional predictors were significant as well, but only for one of the two methods of calculating the standard error.



5.2 Out-of-sample results

Table 4 presents out-of-sample results when testing for predictive ability of 1-month returns. Table 5 and Table 6 present out-of-sample results when for testing predictive ability of 1-year returns and 5-year returns, respectively. Scatterplots of the main period of 1978-2010 similar to those presented for the in-sample results are not included in this report, since there was only one significant variable during this period which is presented in Section 5.2.2.

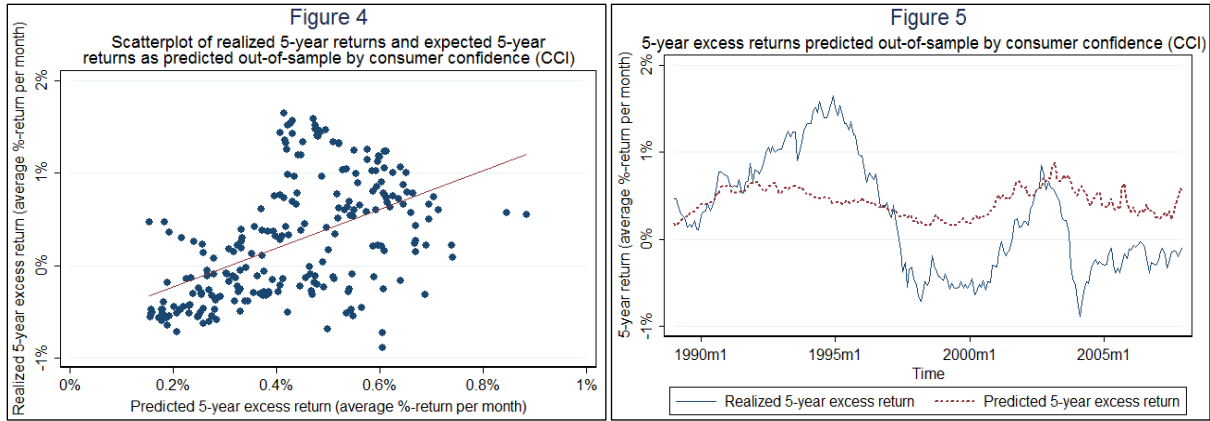
5.2.1 Out-of-sample, traditional predictors

Overall the out-of-sample results are far less impressive. However, if we begin by looking at the full sample of 1938-2010, 2 out of 6 tested variables significantly predicted 1-month returns, 4 out of 6 were successful in predicting 1-year returns, and 2 out of 6 variables predicted 5-year returns. Moreover, the dividend yield was a common denominator in the sense of being a statistically significant predictor for all tested forecasting lengths, which is notable.

For the later sample subset, all tested earnings yields and the dividend yield had negative coefficients for all forecasting lengths. As explained in Section 4.3 on the out-of-sample method, since the realized excess return is regressed on the predicted excess return, a negative coefficient means that the predictor on average predicted high returns when subsequent realized returns turned out to be low, and vice versa. This is however only partly due to poor performance during the period – another contributing factor is how the out-of-sample test is constructed, as will be discussed further in 6.3. Still, regardless of exactly how the coefficients should be interpreted, it is clear that the performance was very poor during later sample subset.

5.2.2 Out-of-sample, unconventional predictors

The out-of-sample performance for the unconventional predictors, which are only analyzed during 1978-2010, is better than the performance of the traditional predictors during the same period, but not a lot better. During this period there was actually only one significant variable out of all tested measures, both traditional and unconventional, and that was yet again the 5-year predictive power of The Conference Board Consumer Confidence Index (CCI) as shown in Figure 4. It should be noted however that CCI was significant on the 1% level with a wide margin when using Newey-West corrected standard errors, while barely being significant at all at the 5% level when using corrected standard errors as suggested by Harri-Brorsen.



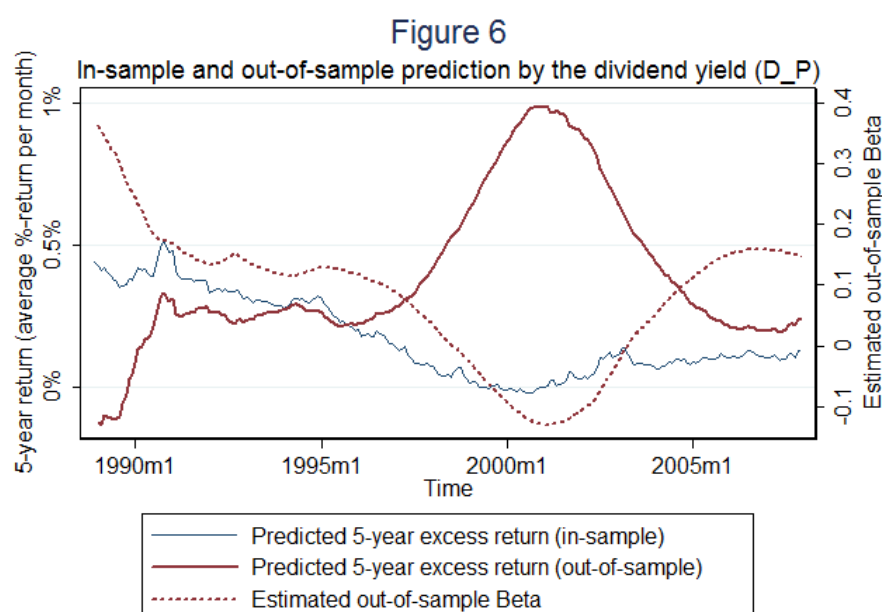
While Figure 4 clearly shows the relationship between expected 5-year returns as estimated by CCI and realized returns during the subsequent 5 years, Figure 5 reveals that the magnitude in realized returns is poorly captured by the model. However, it is not reasonable to expect even 5-year returns to be predictable with too much accuracy out-of-sample. The R^2 of this regression is 0.2585 which is relatively high.

The 1-month predictive ability of investor sentiment (SENT) suggested by Baker and Wurgler (2006) was almost significant at the 5% level with a t-statistic of 1.47. However, as mentioned in Section 3, this time series is constructed by assuming knowledge of some ex post information, making a true out-of-sample test impossible to perform using this data.

5.3 Further discussion

There is a very notable difference in performance between in-sample tests and out-of-sample regressions. Considering all performed tests of all variables during all periods for all forecasting lengths, there are 51 in-sample tests and 51 out-of-sample tests in total. Out of these, 21 models are significant in-sample while only 8-9 models are significant out-of-sample. Furthermore, when analyzing the later period of 1978-2010 most models actually provide high out-of-sample estimates of expected returns when subsequent realized returns are low, and vice versa. The out-of-sample test is designed to simulate the performance that would be achieved if an investor would have used the model in real-time. In this case, the investor will use all historical information available at time t when estimating the expected return that is believed to materialize after time t , and more information will gradually become available to the investor as we move forward in time. The main implication of this is that the *best estimate of the true model coefficients changes gradually over time*. This of course causes problems especially when the performance of the underlying model is unstable and when the

sample size is small – two criteria that both are met when analyzing the period 1978-2010. The 5-year predictive power of the dividend yield during this period is a good example of this since performance was good and significant in-sample but negative when testing out-of-sample. As seen in Figure 6, the difference between the predicted return in-sample and out-of-sample is substantial. The reason is that the out-of-sample estimation of the true β varies from near 0.4 when performing the regression in 1989 to roughly -0.1 when running the same regression using all information available up to 2001. During the same time period, the t-statistic of the regression drop from 4.16 to -1.05, as measured using Newey-West. How an actual investor would behave when experiencing such a drop in significance is not clear.



As demonstrated previously, for example in Figure 1, performance is especially poor during the build-up and bust of the IT-bubble. The reason why many traditional predictors, such as earnings yields and the dividend yield, performed so poorly during this period is that the stock market valued future growth prospects extremely high. This resulted in very low earnings yields and dividend yields which caused these models to predict low future returns. What actually happened though was that the stock market kept climbing for years, in exact opposite of what was predicted by the models. Performance was poor for so long time that the best estimates of the true underlying relationship changed substantially over time. These results are shown in a more general way in Figure 7 by looking at how the t-statistics change over time.

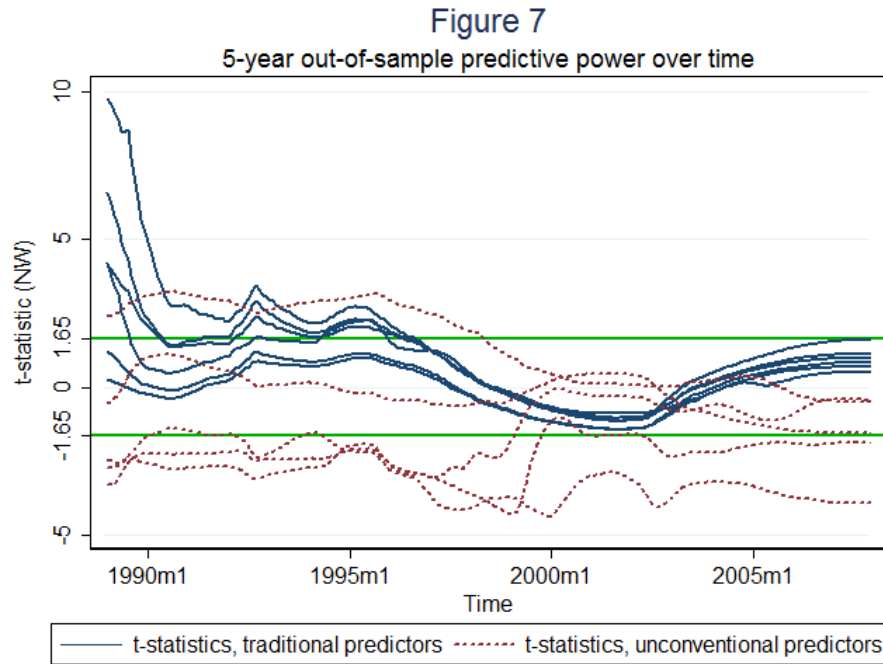


Figure 7

The graph shows how the predictive power changes over time as new information gradually becomes available to the investor. The plotted time series should not be confused with the actual out-of-sample results. The horizontal lines denote statistical significance at the 5% level.

The large t-statistics for the traditional predictors in the beginning of the sample is mainly due to the exceptional performance in the starting sample, which is relatively small. Running the regressions using data available up to 1995, some of the traditional predictors are still significant. After 1995 performance drops substantially however and all traditional predictors are insignificant for the remainder of the sample, even though performance gets better towards the end of the sample. The main message is that the performance of all traditional predictors is very unstable during this time frame, while there are unconventional predictors that demonstrate significant results during almost the entire period despite new data gradually becoming available to the investor. Looking at 1-month or 1-year returns provides a similar picture of the performance of the traditional predictors during this period.

6. Implications and conclusions

The findings presented in this paper show that it is clear that stock market returns have been possible to predict *in the past*, by using certain predictor variables. When analyzing data for 1938-2010 the dividend yield demonstrates significant forecasting ability for all tested forecasting lengths, in-sample as well as out-of-sample. Several other variables also produce significant results during this period.

The million dollar question is how to interpret the poor performance during recent times, especially since 1990. I have demonstrated that the abrupt shifts in stock market valuations related to the IT-bubble did hurt the performance of most predictors, and that especially the traditional predictors performed poorly. The IT-bubble was different from many other crises in the sense that it was created and driven primarily by stock market speculation that was manifested through skyrocketing valuation multiples. This of course affects traditional predictors such as earnings yields and the dividend yield to a larger extent, since they are valuation based measures.

If the IT-bubble is considered an exception in the history of stock markets, the results from the analysis of the period 1938-2010 should probably be reliable even in the future, meaning that predictions can be made with statistical significance for all forecasting lengths tested in this paper by using the traditional predictors. Determining whether it can be considered an exception or not is however beyond the scope of this paper, and considering that traditional stock market predictors have performed poorly for more than 20 years, it may be justified to consider using variables that have performed better in recent times.

Out of the five unconventional predictors that were tested in this paper, only one lacked forecasting power completely for the period 1978-2010 and that was the Purchasing Managers' Index (PMI). When examining the measure of investor sentiment suggested by Baker and Wurgler (2006) and Baker and Wurgler (2007), I find similar results as they do which confirm that low investor sentiment is followed by high subsequent returns, and vice versa, when examining 1-month in-sample predictions. Among the most promising measures are the Composite Leading Indicators (CLI) published by the OECD and the Consumer Confidence Index (CCI) published by The Conference Board. CLI was the best in-sample predictor of 1-year returns during this period out of all 11 variables that were tested, being significant at the 1% level. CCI was the best in-sample predictor of 5-year returns, also being significant at the 1% level and with a wide margin. The negative relationship between consumer confidence and subsequent returns found in Fisher and Statman (2003) is present in this study as well, although I find no predictive power of CCI on such short forecasting horizons as 6 months. The 5-year predictive ability of CCI was also the only significant variable out-of-sample during the later sample subset. Overall, these findings suggest that CLI and CCI should be considered as being possible substitutes or complements to the more commonly used models.

6.1 Further research

Given the econometric difficulties surrounding the topic of stock market prediction, I encourage further research on unconventional predictors. This study tries to answer the question of whether stock returns are predictable by certain variables – one could however instead examine performance relative to the alternative, which is to use the historical mean as estimate of future returns. This is for example done in Goyal and Welch (2008) and Campbell and Thompson (2008), where cumulative R^2 relative to the historical mean is instead used as the measure of out-of-sample performance.

A related question which I have left unanswered is that of economic significance, which could be investigated by looking at the certainty equivalent gains associated with a trading strategy based on the optimal portfolio choice that is implied by these models.

Another potential direction is to explore the predictive performance of the unconventional predictors on subgroups of stocks to see if the results differ between, for example, growth stocks and value stocks or small-cap stocks and large-cap stocks. This seems to be a more common approach in research on investor sentiment [see, e.g., Baker and Wurgler (2006), Brown and Cliff (2004), Neal and Wheatley (1998), and Fisher and Statman (2003)] but is less frequently applied in studies of other variables.

7. References

Baker, Malcolm, and Jeffrey Wurgler, 2000, The equity share in new issues and aggregate stock returns, *Journal of Finance* 55, 2219-2257.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645-1680.

Baker, Malcolm, and Jeffrey Wurgler, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129-151.

Brown, Gregory W., and Michael T. Cliff, 2004, Investor sentiment and the near-term stock market, *Journal of Empirical Finance* 11, 1-27.

Campbell, John, and John Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.

Campbell, John, and Robert Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661-676.

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

Cochrane, John, 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533-1575.

Dow, Charles, 1920, Scientific Stock Speculation (The Magazine of Wall Street, New York).

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

Fama, Eugene, and G. William Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115-146.

Ferreira, Miguel, and Pedro Santa-Clara, 2011, Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514-537.

Fisher, Kenneth, and Meir Statman, 2003, Consumer confidence and stock returns, *Journal of Portfolio Management* 30, 115-127.

Foster, Douglas, Tom Smith, and Robert E. Whaley, 1997, Assessing goodness-of-fit of asset pricing models: The distribution of the maximal R^2 , *Journal of Finance* 52, 591-607.

Goetzmann, William, and Philippe Jorion, 1993, Testing the predictive power of dividend yields, *Journal of Finance* 48, 663-679.

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

- Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829-853.
- Harri, Ardian, and Wade B. Brorsen, 2009, The overlapping data problem, *Quantitative and Qualitative Analysis in Social Sciences* 3, 78-115.
- Hodrick, Robert J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357-386.
- Kalotay, Egon, Philip Gray, and Samantha Sin, 2007, Consumer expectations and short-horizon return predictability, *Journal of Banking & Finance* 31, 3102-3124.
- Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815-849.
- Lewellen, Jonathan, 2004, Predicting returns with financial ratios, *Journal of Financial Economics* 74, 209-235.
- Neal, Robert, and Simon M. Wheatley, 1998, Do measures of investor sentiment predict returns?, *Journal of Financial and Quantitative Analysis* 33, 523-547.
- Nelson, Charles R., and Myung J. Kim, 1993, Predictable stock returns: The role of small sample bias, *Journal of Finance* 48, 641-661.
- Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.
- Pontiff, Jeffrey, and Lawrence Schall, 1998, Book-to-market ratios as predictors of market returns, *Journal of Financial Economics* 49, 141-160.
- Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.
- Umstead, David A., 1977, Forecasting stock market prices, *Journal of Finance* 32, 427-441.
- Valkanov, Rossen, 2003, Long-horizon regressions: Theoretical results and applications, *Journal of Financial Economics* 68, 201-232.
- Wooldridge, Jeffrey M., 2008, Introductory Econometrics (South-Western, Cengage Learning)

7.1 Website Data Sources

Robert Shiller's webpage: <http://www.econ.yale.edu/~shiller/data.htm>

Jeffrey Wurgler's webpage: <http://people.stern.nyu.edu/jwurgler/>

Federal Reserve System: <http://www.federalreserve.gov/econresdata/releases/corpsecure/current.htm>

The Conference Board: <http://www.conference-board.org/data/>

FRED: <http://research.stlouisfed.org/fred2/series/UMCSENT/downloaddata>

OECD.Stat Extracts: http://stats.oecd.org/Index.aspx?DataSetCode=MEI_CLI#

ISM: <http://www.ism.ws/ISMReport/content.cfm?ItemNumber=13339&navItemNumber=12958>

8. Appendix

Table 1: In-sample regressions, predictive power for 1 period.

Variable	1938-2010			1978-2010		
	β (OLS)	t-stat	R-squared	β (OLS)	t-stat	R-squared
E_P	0.1075	2.06*	0.0052	0.0829	1.08	0.0034
E1_P	0.1166	2.22*	0.0060	0.0596	0.82	0.0018
E5_P	0.1235	2.25*	0.0054	0.0519	0.73	0.0014
E10_P	0.1347	2.25*	0.0058	0.0712	0.98	0.0025
D_P	0.2835	2.51**	0.0094	0.2026	1.19	0.0036
eq_is	-0.0314	-2.29*	0.0061	-0.0239	-1.26	0.0036
PMI				-0.0002	-0.56	0.0010
ICS				0.0000	-0.21	0.0001
CLI				-0.0038	-2.14*	0.0121
CCI				-0.0001	-0.74	0.0018
SENT				-0.0056	-1.77*	0.0079

Table 2: In-sample regressions, predictive power for 12 periods.

Variable	1938-2010				1978-2010			
	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared
E_P	0.1193	3.02**	3.25**	0.0715	0.0689	1.19	1.24	0.0246
E1_P	0.1204	3.06**	3.36**	0.0716	0.0439	0.81	0.85	0.0102
E5_P	0.1357	2.77**	3.13**	0.0728	0.0587	1.00	1.07	0.0185
E10_P	0.1395	2.75**	3.11**	0.0696	0.0753	1.24	1.35	0.0295
D_P	0.2786	3.33**	3.74**	0.1021	0.2163	1.53	1.69*	0.0438
eq_is	-0.0130	-1.47	-1.23	0.0117	-0.0067	-0.49	-0.42	0.0031
PMI					-0.0002	-0.79	-0.73	0.0073
ICS					0.0000	-0.05	-0.05	0.0001
CLI					-0.0029	-2.32*	-2.24*	0.0717
CCI					-0.0001	-0.79	-0.77	0.0104
SENT					-0.0043	-1.59	-1.70*	0.0483

Table 3: In-sample regressions, predictive power for 60 periods.

Variable	1938-2007				1978-2007			
	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared
E_P	0.0971	2.96**	2.75**	0.2560	0.0484	1.40	0.97	0.0762
E1_P	0.0941	2.76**	2.58**	0.2404	0.0459	1.32	0.91	0.0720
E5_P	0.1002	2.67**	2.68**	0.2290	0.0592	1.83*	1.35	0.1299
E10_P	0.1002	2.59**	2.60**	0.2061	0.0649	2.04*	1.50	0.1507
D_P	0.2378	5.55**	4.52**	0.4184	0.1922	2.77**	1.93*	0.2343
eq_is	0.0024	0.35	0.24	0.0023	0.0105	1.82*	0.82	0.0445
PMI					-0.0001	-1.41	-0.48	0.0232
ICS					-0.0002	-2.06*	-1.28	0.1468
CLI					-0.0015	-2.82**	-1.14	0.0998
CCI					-0.0001	-5.46**	-2.88**	0.3477
SENT					0.0000	-0.03	-0.02	0.0000

Table 4: Out-of-sample regressions, predictive power for 1 period.

Variable	1949-2010			1989-2010		
	β (OLS)	t-stat	R-squared	β (OLS)	t-stat	R-squared
E_P	0.5935	1.12	0.0023	-0.6638	neg.	0.0010
E1_P	0.4911	1.40	0.0029	-1.9270	neg.	0.0062
E5_P	0.3541	1.20	0.0020	-1.7670	neg.	0.0052
E10_P	0.2154	0.72	0.0008	-1.3675	neg.	0.0044
D_P	0.7512	2.12*	0.0057	-1.1379	neg.	0.0038
eq_is	0.9507	2.25*	0.0064	0.5745	0.44	0.0010
PMI				-0.8874	neg.	0.0043
ICS				-2.9130	neg.	0.0137
CLI				0.7396	1.24	0.0048
CCI				-0.7269	neg.	0.0031
SENT				2.7302	1.47	0.0124

Table 5: Out-of-sample regressions, predictive power for 12 periods.

Variable	1949-2010				1989-2010			
	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared
E_P	0.5914	2.28*	2.45**	0.0389	-1.7324	neg.	neg.	0.0591
E1_P	0.4824	2.37**	2.60**	0.0372	-2.2014	neg.	neg.	0.0846
E5_P	0.3530	1.85*	1.99*	0.0265	-1.6129	neg.	neg.	0.0540
E10_P	0.2675	1.40	1.51	0.0153	-1.1414	neg.	neg.	0.0414
D_P	0.7987	2.63**	2.94**	0.0530	-1.2201	neg.	neg.	0.0541
eq_is	0.2926	0.52	0.49	0.0022	-1.1567	neg.	neg.	0.0185
PMI					-1.6040	neg.	neg.	0.0608
ICS					-2.4152	neg.	neg.	0.0861
CLI					0.6505	0.64	0.68	0.0134
CCI					-1.4304	neg.	neg.	0.0357
SENT					0.7576	0.37	0.40	0.0031

Table 6: Out-of-sample regressions, predictive power for 60 periods.

Variable	1949-2007				1989-2007			
	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared	β (OLS)	t-stat (NW)	t-stat (HB)	R-squared
E_P	0.5450	1.69*	1.65*	0.1417	-1.7433	neg.	neg.	0.1942
E1_P	0.5074	1.42	1.37	0.1079	-1.8203	neg.	neg.	0.2302
E5_P	0.3948	1.47	1.41	0.0799	-0.9923	neg.	neg.	0.1630
E10_P	0.3303	1.33	1.24	0.0578	-0.8772	neg.	neg.	0.2085
D_P	0.7310	2.11*	2.04*	0.1912	-0.8867	neg.	neg.	0.1429
eq_is	0.0856	0.16	0.16	0.0008	-2.2671	neg.	neg.	0.1362
PMI					-2.1209	neg.	neg.	0.0687
ICS					-1.4281	neg.	neg.	0.0378
CLI					0.4961	0.53	0.29	0.0078
CCI					2.0802	2.95**	1.56	0.2585
SENT					-3.8161	neg.	neg.	0.3834

Figure 8 – Figure 13
Forecasting length: 1 month
Estimation period: 1978-2010
Estimation method: in-sample

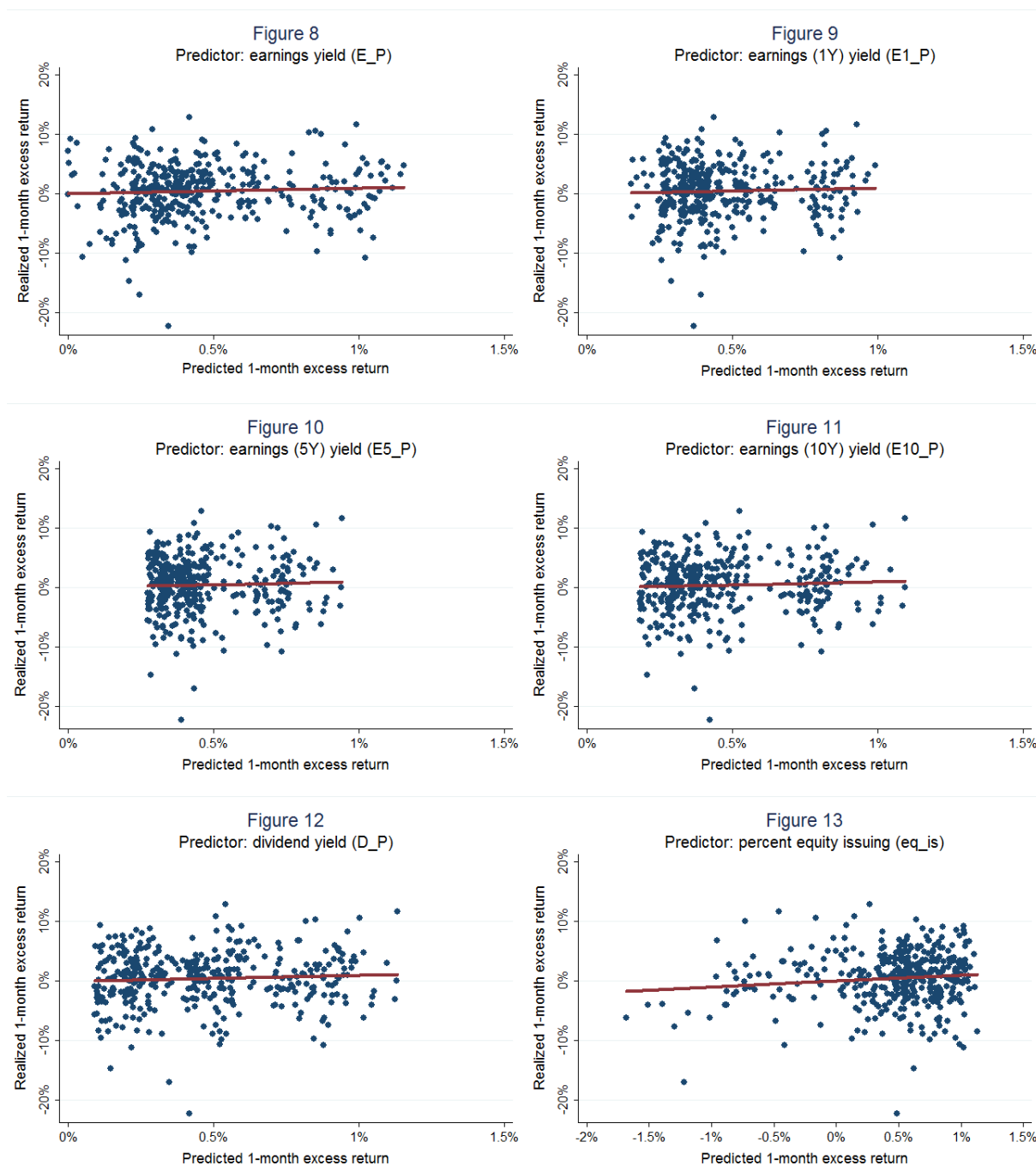


Figure 14 – Figure 18
Forecasting length: 1 month
Estimation period: 1978-2010
Estimation method: in-sample

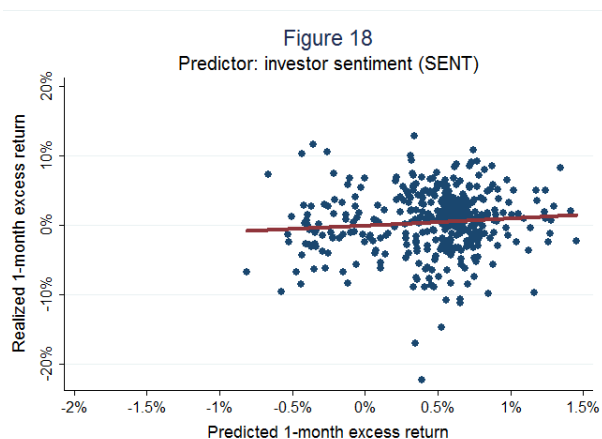
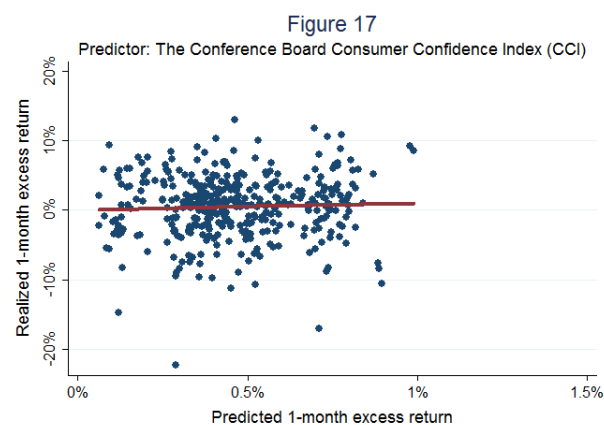
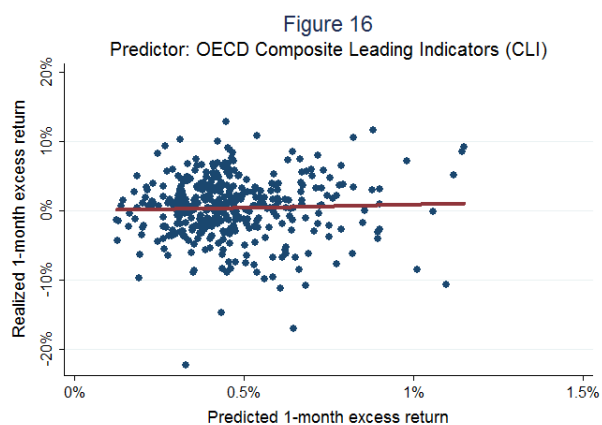
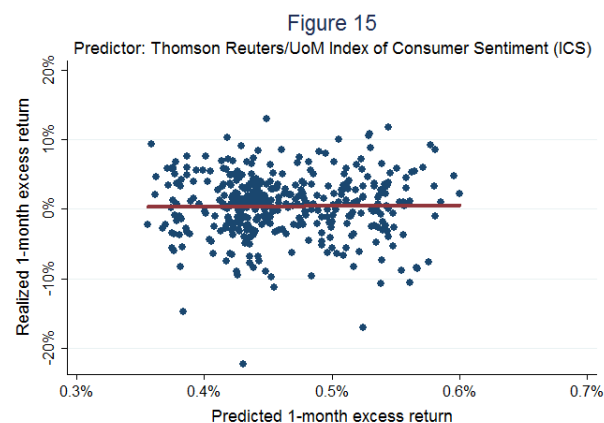
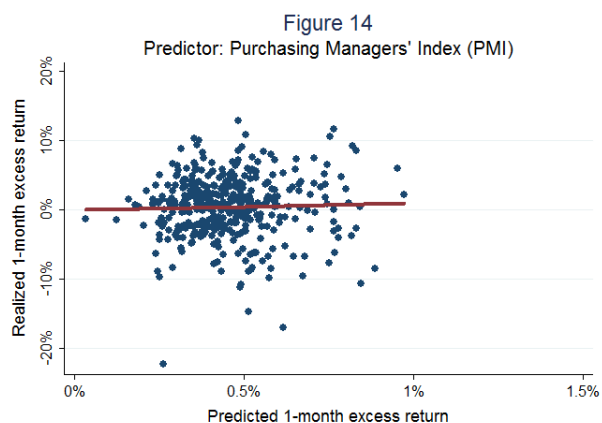


Figure 19 – Figure 24
Forecasting length: 1 year
Estimation period: 1978-2010
Estimation method: in-sample

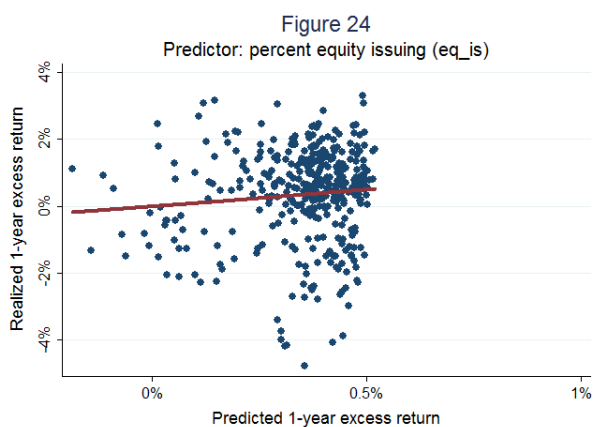
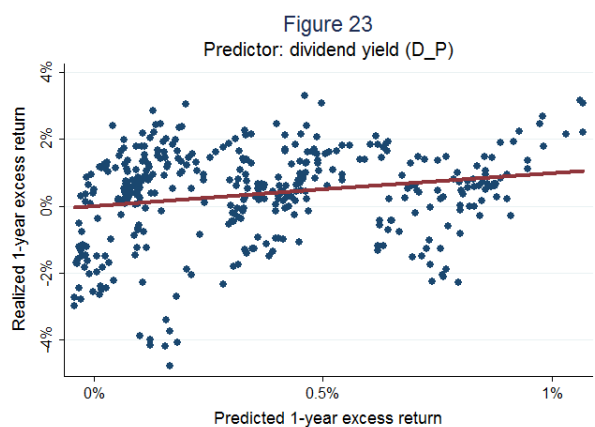
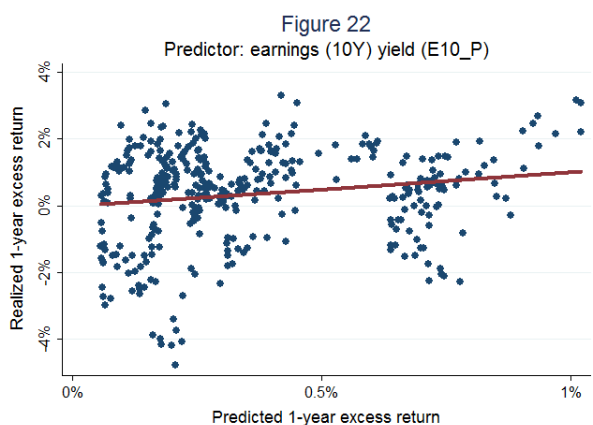
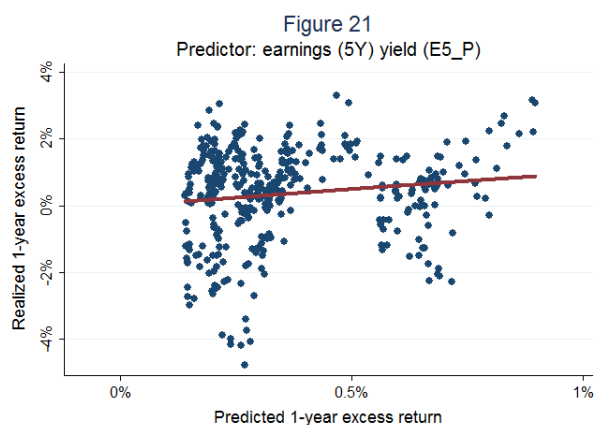
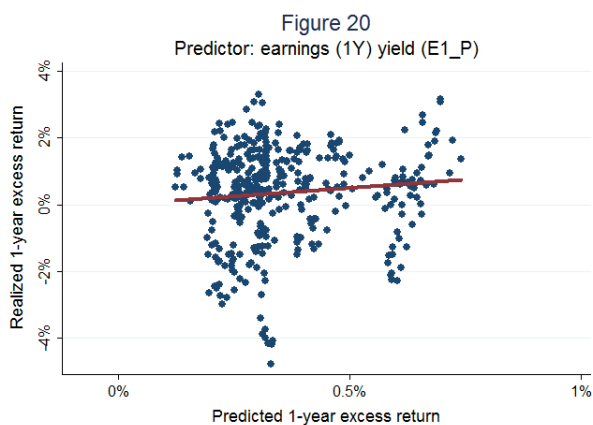
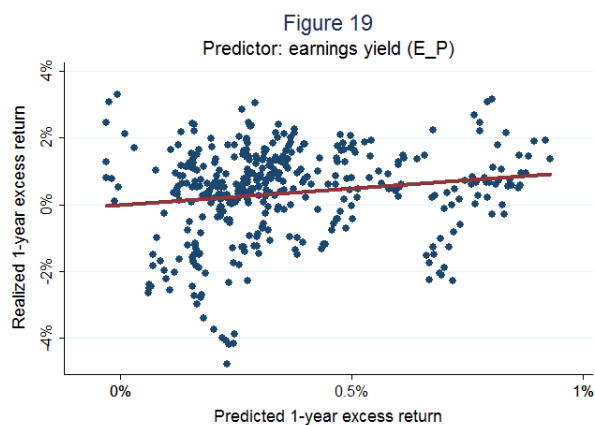


Figure 25 – Figure 29
Forecasting length: 1 year
Estimation period: 1978-2010
Estimation method: in-sample

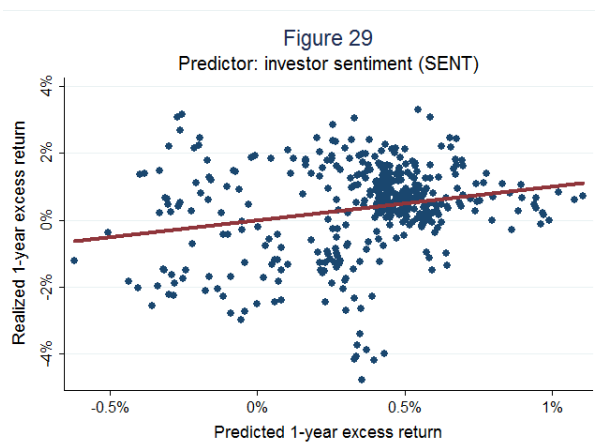
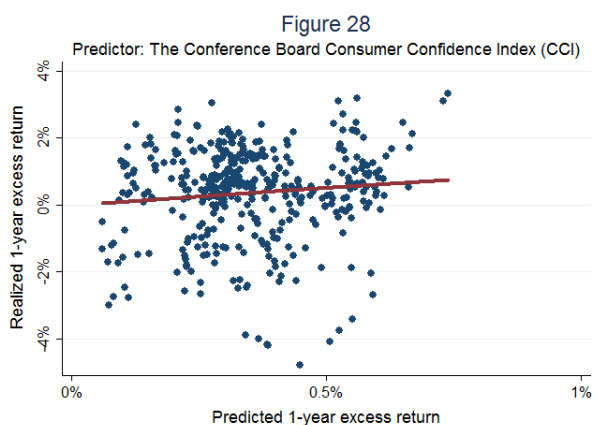
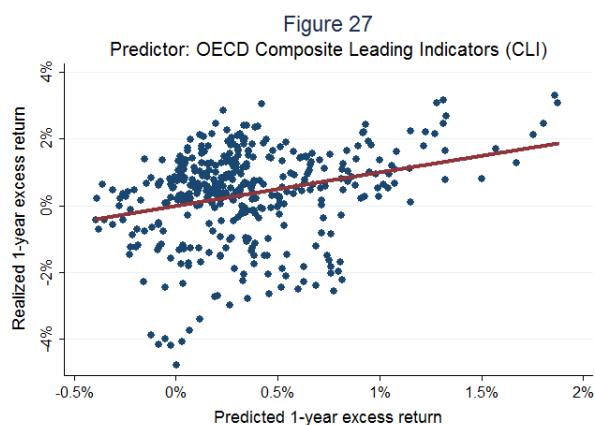
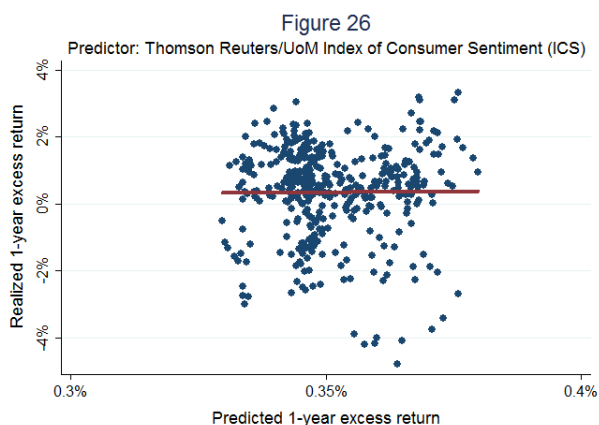
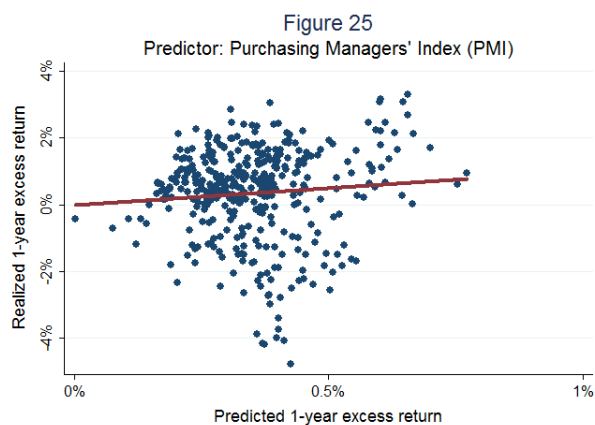


Figure 30 – Figure 35
 Forecasting length: 5 years
 Estimation period: 1978-2007
 Estimation method: in-sample

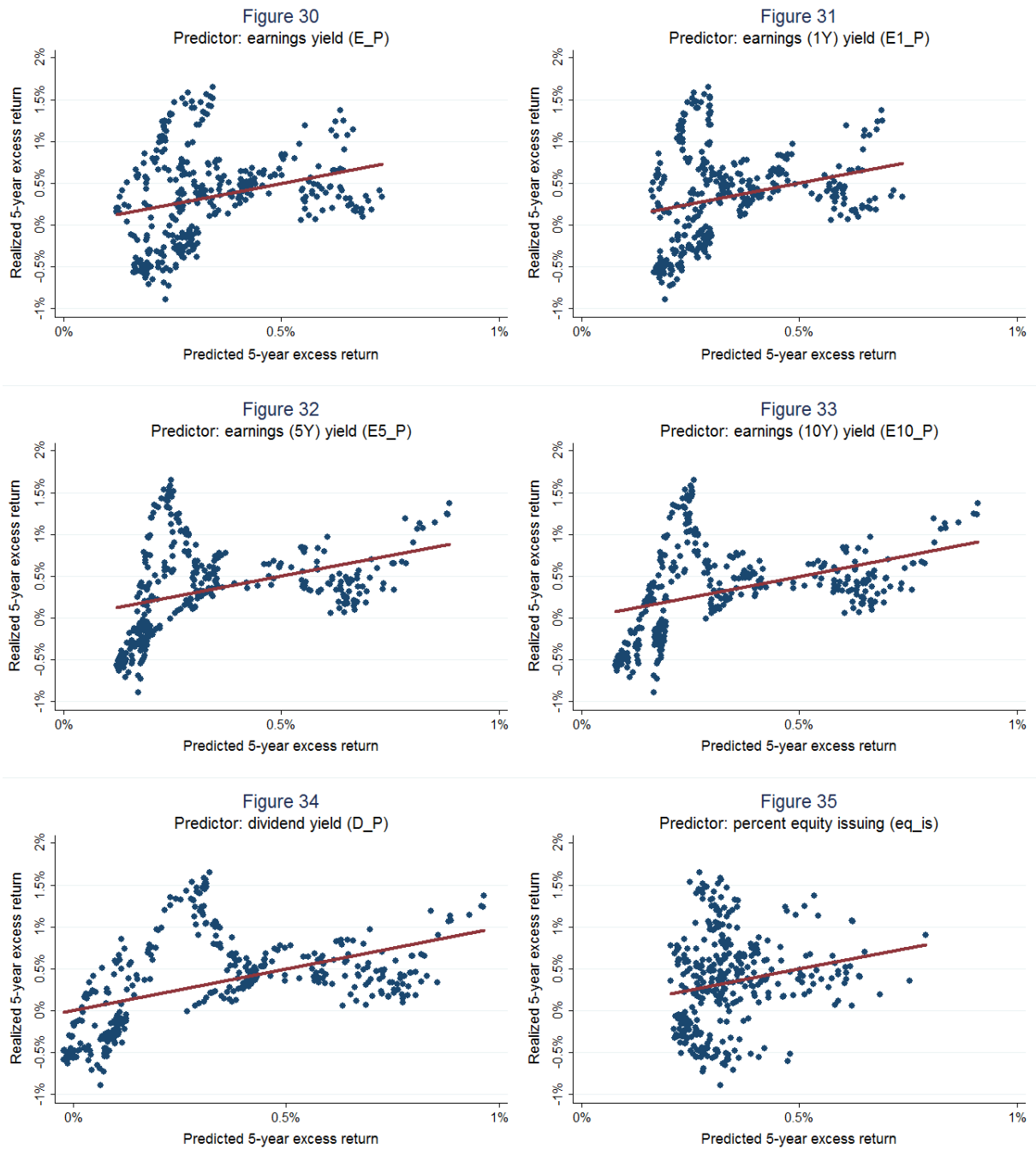


Figure 36 – Figure 40
Forecasting length: 5 years
Estimation period: 1978-2007
Estimation method: in-sample

