Macroeconomic Fundamentals and Stock Return Predictability

Ryan Compton*
University of Manitoba
compton@cc.umanitoba.ca

James Morley
Washington University
morley@wustl.edu

Draft: December 2006
Please do not cite; Comments welcome

Abstract

Does the real economy possess predictive power for stock returns? While a sizeable literature exists on this question, the evidence is mixed and continues to draw debate. In this paper, we examine the issue of stock return predictability over a range of time horizons using business condition measures and a Bayesian time-varying parameter model for the United States. Not only does this model allow for testing of stock predictability, the time-varying coefficients allow for investigation of how this predictability has changed over time. Our findings indicate both discrete structural breaks as well as gradual changes in the predictive ability of a number of the variables. We also find more evidence of predictability over fixed coefficient prediction models. Importantly, these results reinforce the need to consider time variation in forecasting relationships.

Keywords: Predictive regression models, real stock returns, MCMC, time-varying parameters

JEL Codes: C11, C53, E44, G12

* Compton: Department of Economics, University of Manitoba, 501 Fletcher Argue Bldg, Winnipeg, MB, R3T 5V5. Morley: Department of Economics, Campus Box 1208, Washington University, St. Louis MO, USA, 63130. We would like to thank Gaetano Antinolfi and Steve Fazzari for helpful comments. The usual disclaimer applies.
1. Introduction

A large literature exists which examines, over a range of horizons, whether macroeconomic variables are useful predictors of stock returns. The results have proven to be mixed, with some papers finding evidence that macroeconomic variables predict stock returns and others finding little evidence of predictability.

These results, however, come largely from work based on fixed coefficient models that assume parameter stability and thus do not consider the possibility of structural change in the underlying predictive model. As Pesaran and Timmerman (2002) suggest, numerous sources of parameter instability exist, which makes this issue an important area of study. In particular, the economic value from exploitation of predictability may be reduced if the relationships may break down. Possible sources of instability include large changes in market sentiments, bubbles, changes in the conduct of monetary or debt management policies, economic shocks, market volatility, or learning on the part of investors. All of course which are possibilities over the US post-war era.

Increasingly, research in this area has addressed parameter instability. Recent papers have usually approached this question using discrete structural break tests such as those developed by Andrews (1993), Bai (1997), or Bai and Perron (1998) to test the structural stability of the underlying predictive model, and incorporate any breaks in their estimations.

This paper continues in this spirit, using a Bayesian time-varying parameter model to investigate the stability of in-sample stock return predictability over time. The key advantage of this approach is that it allows not only for the sort of discrete breaks captured with traditional structural break tests, but also is general enough to capture more gradual changes in predictability. We consider the case of the United States from 1955 to

---

2 These are suggestions raised by Pesaran and Timmerman in their 1995 and 2002 articles.
2004 using monthly data, and draw on standard macroeconomic variables as well as a handful of financial variables which are commonly used in this literature and have been found to have predictive power for stock returns.³

Previewing our results, there is much evidence of instability in the stock prediction model over all horizons and across many possible forecasting variables. The time-varying parameters indicate large and interesting time variation which can be both gradual as well as more discrete, and by allowing for the coefficients to vary obtain evidence of macro and financial variables being more useful to predict stock returns. Overall, these results demonstrate the usefulness of the time-varying parameter model to investigate stock predictability and they call into question the practice of using fixed coefficient models for predicting stock returns.

The remainder of the paper proceeds as follows. Section 2 provides an overview of the literature on model instability in stock return predictability. Section 3 details the data and empirical approach. Section 4 presents the results along with discussion. Section 5 concludes by summarizing the main findings and suggesting avenues for future research.

2. Background

The literature on stock predictability is extensive. Our survey intentionally focuses on papers which directly or indirectly consider the idea of structural stability of the prediction regression. Rapach, Wohar, and Rangvid (2005) provide an extensive list of papers which have investigated the performance of macro variables for stock predictability more generally, and interested readers are encouraged to consult their paper for more on this literature.

In addition to their survey of the literature, Rapach, Wohar and Rangvid (2005), represents one of the most extensive recently published papers in the stock predictability literature, and is worth discussing briefly. Their study uses the mixed findings of the earlier literature as motivation to reexamine the issue of stock return predictability across 12 industrialized countries over 1 to 24 month horizons. The authors investigate in-

³ See Pesaran and Timmerman (1995), and Rapach, Wohar and Rangvid (2005).
sample predictability and use new procedures developed by Clark and McCracken (2001), and McCracken (2004) to consider the out-of-sample performance of macro variables in turn. Clark’s (2004) general-to-specific model selection procedure is also used to identify and test the “best” forecasting model for each country. The evidence proves strongest for interest rates as predictors of real stock returns, particularly over shorter horizons, with weaker evidence in favour of the inflation rate, monetary aggregates, and the term spread. The strong performance of interest rates found for short-horizon predictability is consistent with the findings of Ang and Bekaert (2001), and unlike many studies, the in-sample and out-of-sample results of this study are often in agreement.  

Focusing on the issue of structural breaks in stock predictability, a number of papers find evidence of model instability. Pesaran and Timmermann (1995) examine the robustness of US stock return predictability and whether investors could have made profits beyond a buy-and-hold strategy in the market index. The authors simulate investors’ decisions in real time using publicly available macro and financial data, and take into account the statistical methods and technology available at the time to make one-step ahead forecasts. The authors find only a single variable, the lagged one-month T-bill rate, is included in their forecasting models throughout the entire sample (1960-1992). Further, by plotting the inclusion frequency of the various variables considered, many prove useful for predictive purposes but at varying times over the sample period. Thus the determinants of stock return predictability in the US appear to have undergone important changes over the course of the sample period. Pesaran and Timmerman point out this outcome may be related to macroeconomic events such as the oil price shocks, changing Federal Reserve operating procedures (1979-82), or due to market volatility.

Qi (1997) follows the approach and uses the same variables as Pesaran and Timmerman (1995). However Qi uses a neural network model which allows for flexible nonlinear functional approximation, thus extending Pesaran and Timmerman (1995) by allowing

---

4 Ang and Bekaert (2001) test for in-sample stock predictability for France, Germany, Japan, the UK, and US, using the short rate, the dividend yield, and earnings yield as potential predictors. Interest rates prove to be the sole useful predictor of stock returns.
for choice not only among linear forecasting models but nonlinear ones as well. A recursive procedure is again used to simulate the decisions of an investor in real time, with the nonlinear neural network model proving to better fit the data in-sample and forecast excess stock returns out-of-sample than the linear alternative. As in the Pesaran and Timmerman paper, Qi finds that good out-of-sample forecasting performance occurs primarily during the 1970s and 1980s when markets were rather volatile.

Neely and Weller (2000) investigate a claim in the literature that long-horizon stock return predictability in a VAR framework can be inferred using relatively short spans of data. This claim relies critically on the assumption of parameter stability in the model (basically in-sample predictability implies out-of-sample predictability if the structural relationships are stable), which the authors argue has failed to receive much investigation. Examining the out-of-sample forecasts with a VAR setup in line with Bekaert and Hodrick (1992), the forecasts are usually found to underperform a simple benchmark model. Neely and Weller consider a number of explanations in turn for this lack of predictability, and using Monte Carlo analysis and structural break tests, determine that structural breaks in the data which undermine the assumption of parameter stability are the cause of this poor performance.

In their 2002 paper, Pesaran and Timmerman employ a two-step procedure for identifying structural breaks in real time, with the first step using a reversed ordered Cusum (ROC) test to identify structural breaks, and the second using post-break data to estimate the parameters of a US stock return forecasting model. With this procedure, the authors find evidence of three major breaks (1969, 1974, 1990) in the forecasting model, which uses both financial as well as economic variables.

Goyal and Welch (2004) are interested in the general question of the ability of financial and macro variables to predict the equity premium out-of-sample over a range of horizons. Using monthly, quarterly, and annual data, the authors find not one of their variables prove to be a stable, useful out-of-sample predictor (though they do have good in-sample properties). An important point relevant for our purposes is the notion of
“stable”. Using annual data, the authors plot the cumulative squared prediction errors of the prevailing mean minus the cumulative squared prediction error of the predictive variable from the linear historical regression and find some variables do have predictive power over some periods of the sample.

Lee (2004) considers whether there is a constant linear relationship between real stock returns and the price-dividend ratio. Examining the log real price-dividend ratio from 1872-2001, and using a two-regime LSTAR model to test for linearity in the ratio, Lee finds in one regime that the valuation ratio is stationary and not strongly mean reverting. This suggests that the forecasting ability of this variable may not be as strong as previous studies suggest. In the other regime (which has existed since the late 1950s), the ratio is non-stationary, which Lee argues suggests the possibility of a spurious fit for stock return predictability.

Finally, Rapach and Wohar (2004) investigate the stability of post-war US real stock return predictability based on a number of financial variables and using a range of structural break tests (Andrews (1993) SupF statistic, Bai (1997) procedure, and Bai and Perron (1998) procedure). The authors find evidence of structural breaks in seven of eight bivariate prediction models of S&P 500 returns, and three of eight prediction models of CRSP equal-weighted returns.\(^5\)

A few papers in this literature fail to find parameter instability in the forecasting model. For example, Viciera (1997) examines stability of the stock return forecasting equation (with the dividend-price ratio as the variable of interest) in the context of recursive estimation, and fails to reject the null of stability of the model parameters. As well the author is also unable to reject the null that the dividend yield does not forecast one-month stock returns once the persistence in the dividend yield is taken into account. Further, Ferson, Heuson, and Su (2004) examine time variation in the expected returns of stocks by comparing unconditional sample variances of returns to estimates of expected

\(^5\) Rapach and Wohar suggest a time-varying parameter model as an alternative approach for future research possibilities.
conditional variances. The authors use semi-strong form tests based on lagged macro and financial variables, finding evidence which supports predictive power for stock returns, but generally does not support the notion of declining predictability over time.\(^6\)

Despite these two examples, clearly there is growing evidence which supports the notion of parameter instability in forecasting stock returns. Our paper investigates the stability of in-sample stock predictability using a Bayesian time-varying parameter model to further investigate this issue and provide additional information about whether there is evidence of breaks or changes over time in the predictability of stock returns based on a range of variables.

It is also important to point out that while this is an in-sample exercise, our findings have important implications for results based on out-of-sample studies. As is evident in the literature, many variables which possess predictive power for stock returns in-sample, lose their significance once the analysis is moved to out-of-sample. The results of our paper prove useful for out-of-sample analysis as a likely reason that some variables perform well in-sample but not out-of-sample is due to the fact that structural breaks or changes in the predictive relationship invalidate the fixed coefficient models used to make out-of-sample forecasts. By improving our understanding of whether and why structural changes in the predictive relationship occur, this will allow for forecasters to better predict out-of-sample as these breaks can be taken into account when constructing a forecast. This may be of use especially for those using Bayesian models where forecasters can account for breaks in real time rather than having to wait for ex post evidence of a break.

### 3. Data and the Model

The primary focus of this paper is the use of macroeconomic variables to predict stock returns, and to examine whether there is time-variation in a variable’s ability to predict

---

\(^6\) There is some evidence of a small reduction in predictability in recent periods when only macroeconomic variables are included as the lagged regressors.
returns. However given the pervasive use of financial variables in the general stock predictability literature, a number of financial variables are also examined.

The choice of variables is largely driven by those used in the existing predictability literature.\(^7\) The macro variables we consider are:

- three-month Treasury bill rate (levels);
- ten-year government bond yield (levels);
- term spread (difference between the 10 and 3 month Treasury rates);
- inflation rate (log difference of CPI);
- industrial production growth (log difference of the industrial production index);
- M1 money growth (log difference of M1);
- M2 money growth (log difference of M2);
- change in the unemployment rate.

As well, we consider the following financial variables:

- price-earnings ratio (log ratio; earnings based on a 12-month moving sum);
- price-dividend ratio (log ratio; dividends based on a 12-month moving sum)\(^8\);
- default spread (difference of Moody’s Baa and Moody’s Aaa).

Stock returns are expressed as excess returns, measured as the value-weighted return on the S&P 500 Composite index (including dividends) less the risk free rate. Throughout the paper, we will refer to excess returns as simply ‘returns’.

The data are monthly observations from 1955.1-2004.12 from CRSP, the Federal Reserve Bank of St. Louis FRED II data archive, and Robert Shiller’s website. The data appendix

---


\(^8\) Note: The P/E and P/D ratios cover the period 1955.1 – 2004.5, and required interpolation based on COMPUSTAT data from March 2004 to May 2004 and October 2003 to May 2004 respectively.
provides further details. Table 1 below provides summary statistics of the variables used in the study.

Table 1: Summary Statistics*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Month Treasury Bill Rate</td>
<td>5.30</td>
<td>2.80</td>
</tr>
<tr>
<td>10 Year Treasury Bond Yield</td>
<td>6.71</td>
<td>2.65</td>
</tr>
<tr>
<td>Term Spread</td>
<td>1.40</td>
<td>1.20</td>
</tr>
<tr>
<td>Inflation Rate</td>
<td>3.92</td>
<td>3.92</td>
</tr>
<tr>
<td>Industrial Production Growth</td>
<td>3.22</td>
<td>10.70</td>
</tr>
<tr>
<td>M1 Money Growth</td>
<td>4.71</td>
<td>6.45</td>
</tr>
<tr>
<td>M2 Money Growth</td>
<td>6.66</td>
<td>4.00</td>
</tr>
<tr>
<td>Unemployment Rate Change</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>Price-Earnings Ratio</td>
<td>2.76</td>
<td>0.38</td>
</tr>
<tr>
<td>Price-Dividend Ratio</td>
<td>3.47</td>
<td>0.38</td>
</tr>
<tr>
<td>Default Spread</td>
<td>0.95</td>
<td>0.42</td>
</tr>
<tr>
<td>S&amp;P 500 Excess Return</td>
<td>5.39</td>
<td>51.24</td>
</tr>
</tbody>
</table>

*Note: Statistics are based on the transformed data and are stated in annualized terms.

We examine the predictability of stock returns over 1, 6, and 12-month horizons (non-overlapping) using a Bayesian time-varying parameter model. This model can be expressed in state space form as follows

**Measurement Equation**

\[ r_{t+k,t} = \beta_t Z_t + \varepsilon_{t+k} \]  
(1)

**State Equation**

\[ \beta_t = \beta_{t-1} + v_t \]  
(2)

\[ \varepsilon_{t+k} \sim \text{i.i.d. } N(0,R), \ v_t \sim \text{i.i.d. } N(0,Q) \]

where \( r_{t+k,t} \) is the cumulative S&P 500 excess return from period \( t \) to \( t+k \), \( \beta_t \) is a time-varying coefficient, \( Z_t \) is a macro or financial variable, and \( \varepsilon_{t+k} \) is a forecast error.\(^9\) The

---

\(^9\) Cogley and Sargent (2002) provide a detailed example of a Bayesian time-varying parameter model.
model coefficient is treated as an unobserved state variable which is specified in equation 2, follows a random walk process. $\varepsilon_{t+k}$ is assumed i.i.d. normal with mean zero and variance $R$, while $v_t$ is also assumed i.i.d. normal with mean 0 and variance $Q$. Additionally, $\varepsilon_{t+k}$ and $v_t$ are allowed to have a nonzero covariance.

The choice of modeling the regression coefficient as a random walk process is due to its attractive property of allowing for changes in the regression coefficient while also allowing for the special case of a constant coefficient (where the variance of $v_t$ collapses to zero). As well, it has been shown that the random walk specification is fairly robust to misspecification.¹¹

Fairly diffuse priors are used to ensure that the likelihood function dominates the prior in our posterior. Further, in order to simulate draws from the posterior, this paper uses Markov-Chain Monte Carlo methods in the form of multi-move Gibbs-sampling. For our results, 10,000 draws are drawn from the Gibbs-sampler, with 3,000 used for the “burn-in period” to ensure the Gibbs-sampler has converged. The remaining 7,000 draws are used to estimate the marginal posterior distribution of the time-varying parameter. More detail on the priors used, as well as the Bayesian setup, can be found in the appendix.

¹⁰ Note that the regressions use demeaned data, assuming a fixed constant. As well, the data is nonoverlapping.
4. Results

In order to provide a starting point, Table 2 provides the results for forecasts of 1-month, 6-month, and 1-year excess returns using a constant coefficient bivariate model.

Table 2: Fixed Coefficient Model (1, 6, 12-month horizons)

<table>
<thead>
<tr>
<th>Variable</th>
<th>1 Month Return</th>
<th>6 Month Return</th>
<th>12 Month Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Month T-Bill Rate</td>
<td>-1.43*</td>
<td>-5.68</td>
<td>-6.39</td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(5.10)</td>
<td>(10.26)</td>
</tr>
<tr>
<td>10 Year T-Bond Yield</td>
<td>-0.80</td>
<td>-2.16</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.78)</td>
<td>(5.64)</td>
<td>(9.81)</td>
</tr>
<tr>
<td>Term Spread</td>
<td>3.86**</td>
<td>17.86*</td>
<td>26.4</td>
</tr>
<tr>
<td></td>
<td>(1.18)</td>
<td>(9.84)</td>
<td>(21.66)</td>
</tr>
<tr>
<td>Inflation Rate</td>
<td>-1.54**</td>
<td>-4.80</td>
<td>-8.44</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(3.95)</td>
<td>(9.07)</td>
</tr>
<tr>
<td>Industrial Prod. Growth</td>
<td>-0.09</td>
<td>-1.78</td>
<td>-3.72**</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(1.27)</td>
<td>(1.47)</td>
</tr>
<tr>
<td>M1 Money Growth</td>
<td>-0.21</td>
<td>-0.41</td>
<td>-7.55**</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(2.70)</td>
<td>(4.11)</td>
</tr>
<tr>
<td>M2 Money Growth</td>
<td>0.19</td>
<td>0.58</td>
<td>-15.22**</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(3.52)</td>
<td>(7.11)</td>
</tr>
<tr>
<td>UE Rate Change</td>
<td>18.19*</td>
<td>101.18</td>
<td>319.40***</td>
</tr>
<tr>
<td></td>
<td>(11.21)</td>
<td>(92.77)</td>
<td>(83.34)</td>
</tr>
<tr>
<td>Price-Earnings Ratio</td>
<td>-6.12</td>
<td>-44.21</td>
<td>-81.76</td>
</tr>
<tr>
<td></td>
<td>(6.23)</td>
<td>(36.85)</td>
<td>(74.39)</td>
</tr>
<tr>
<td>Price-Dividend Ratio</td>
<td>-7.55</td>
<td>-49.11</td>
<td>-91.25</td>
</tr>
<tr>
<td></td>
<td>(6.29)</td>
<td>(34.43)</td>
<td>(73.05)</td>
</tr>
<tr>
<td>Default Spread</td>
<td>7.43</td>
<td>31.19</td>
<td>66.11</td>
</tr>
<tr>
<td></td>
<td>(5.06)</td>
<td>(36.66)</td>
<td>(52.94)</td>
</tr>
<tr>
<td>S&amp;P 500 Excess Return</td>
<td>0.03</td>
<td>-0.01</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.13)</td>
<td>(0.16)</td>
</tr>
</tbody>
</table>

Note: Estimates are based on OLS using White’s standard errors.
*, **, *** significant at 10%, 5%, 1%.

From Table 2, there is mixed evidence of predictive power, which changes over variables and forecast horizons. Looking at the relatively short horizon of 1-month, many of the variables with (as well as without) predictive power are what we would expect given the findings of earlier literature for short horizon predictability.\(^{12}\) The 3-month T-Bill rate, term spread, and inflation often appear to possess predictive power at very short horizons. Somewhat surprisingly, the change in the unemployment rate also proves significant. As well, the price-earnings ratio and price-dividend ratio, popular financial variables in the

\(^{12}\) For example, Goval and Welch (2004), Rapach and Wohar (2004), Rapach, Wohar, and Rangvid (2005)
predictability literature, prove to be insignificant (which is in line with Rapach and Wohar (2004) who use quarterly data for a similar time period).

At the 6-month horizon, only the term spread continues to have significant predictive power, however at the longer 12-month horizon, a number of variables begin to show evidence of predicting stock returns. The changes in industrial production, as well as both measures of money are significant, and the change in the unemployment rate once again exhibits significance. While these results serve as a starting point and are interesting in their own right, they are based on a stable prediction model. The papers discussed in section 2 have raised serious doubts on the stability of stock prediction models, and it is this issue that we turn to next.

Figures 1a, 1b, and 2, provide the results of the bivariate Bayesian time-varying parameter model discussed in section 3. Figures 1a and 1b provide results for the ability of macroeconomic variables to forecast stock returns. Figure 1a provides the results for the bivariate models which include in turn as independent variables, the log change of industrial production, inflation, log change of M1, and log change of M2. Figure 1b includes the 3-month T-Bill rate (levels), 10-year T-Bond yield (levels), term spread, and log change of the unemployment rate. These results are divided between two figures simply due to the number of macroeconomic variables investigated in this study. Figure 2 provides results based on the financial variables - log price-earnings ratio, log price-dividend ratio, default spread, and log excess returns as predictors.

Each figure is organized such that each row contains the results for a given macro/financial variable, while the three columns present results for 1-month, 6-month, and 12-month stock return horizons. Each graph within the figure details the posterior median as well as posterior bands, which indicate the 25th and 75th percentile of the posterior distribution for the time-varying parameter of a given macroeconomic/financial variable. The posterior median provides an estimate of the value of the time-varying parameter, while the posterior bands indicate the tightness of the posterior distribution.13

13 The use of 25th and 75th percentile posterior bands is seen in Cogley and Sargent (2002).
So for example, in Figure 1a, the top row details the posterior median and posterior bands for the time-varying coefficient of industrial production for predicting 1-month returns (left column), 6-month returns (middle column), and 12-month returns (right column).

Figure 1a: Posterior Distribution of Macro Variable Time-Varying Parameters

Note: IP-Industrial Production; INF-Inflation; M1-M1 Money; M2-M2 Money. In each graph, the solid line indicates the posterior mean of the time-varying coefficient, while the dotted lines indicate interquartile bands. The shaded regions coincide with NBER dated recessions.

Note that in the case of predicting 1-month returns, shaded regions coinciding with NBER dated recessions are included for use when analyzing the movement of a given coefficient over time.
Figure 1b: Posterior Distribution of Macro Variable Time-Varying Parameters

Note: T3-3 Month Treasury Bill Rate; T10-10 Year Treasury Yield; TERM-Term Spread; UE-Unemployment Rate. In each graph, the solid line indicates the posterior mean of the time-varying coefficient, while the dotted lines indicate inter-quartile bands. The shaded regions coincide with NBER dated recessions.
Figure 2: Posterior Distribution of Financial Variable Time-Varying Parameters

Note: PE-Price Earnings Ratio; PD- Price Dividend Ratio; DEF-Default Spread; XS-Excess Return. In each graph, the solid line indicates the posterior mean of the time-varying coefficient, while the dotted lines indicate inter-quartile bands. The shaded regions coincide with NBER dated recessions.

Much of the following discussion will focus on the results for predicting stock returns over 1-month horizons. That said, an important result emerging from many of the 6 and 12-month horizon forecasts is the extent that the exhibited time variation of the coefficients tend to pick up many of the major movements seen in the monthly results, suggesting that some factors driving the short-horizon results likely also have an impact
on longer horizon predictability. As well, by allowing for time variation in the coefficients, it is clear that a number of variables pick up predictive power for stock returns over parts of the sample that are not present in the fixed coefficient results detailed in Table 2, and that the dynamics certainly change over time for a number of cases.

Turning to the monthly results, we first briefly survey the results before investigating some of the cases more in depth. In Figure 1a, the coefficient for industrial production suggests little predictability for 1-month stock returns, however there is evidence of some predictability over the later half of the 1990s and into the early 2000s. This is interesting, as a point sometimes raised in the literature is that during the bull market in the US during the 1990s, many predictive models which relied on financial variables performed rather poorly, and some have suggested that perhaps macro variables may have possessed more predictive content during this unusual period.

In the case of inflation, not surprisingly, given the significant results of the fixed coefficient model, there are a number of periods where inflation appears to be a useful predictor of stock returns. M1 and M2 both exhibit changes in predictability with largely positive coefficients over the first half of the sample and evidence of predictability which largely vanishes post 1980.

In Figure 1b, the coefficient for the 3-month T-Bill rate suggests periods of predictability over the first half of the sample, with little predictability post 1975, while the 10-year T-Bill yield exhibits similar movements to that of the 3-month T-Bill coefficient, however with posterior bands that largely contain zero throughout the sample. The term spread provides very nice results, being a useful predictor of stock returns over a number of periods in the sample, while the unemployment rate also proves to have predictive power, largely during the 1970s.

Figure 2 details the results of the financial variables considered. For both the price-earnings ratio as well as price-dividend ratio, the coefficients seem to exhibit similar
movements, which move quite quickly and widely. The default spread coefficient exhibits movement very similar to that of the term spread results of Figure 1b, with a lot of predictability especially over the 1970s. Finally, lagged excess returns interestingly appear to be useful as a predictor but for the most part only during periods when the economy is in a recession.

Clearly from the previous discussion, the time-varying parameter model appears to be useful for better understanding the extent to which variables possess predictive power for stock returns. By capturing the estimated coefficient at each point in time, evidence of shifts or gradual change in coefficients as well as basic predictability is captured in much more detail than would be possible using standard structural break tests. However with time-varying parameter results, the question that begs to be answered is what does this technique give us that other approaches do not, why do we see the changes we do, and how do these changes compare with the results of earlier work? To answer this question we consider some of the monthly results more closely.

Consider first the results for the 3-month T-bill rate. The 3-month T-bill rate seems to demonstrate a large break in 1975, with the median of the distribution jumping towards zero and remaining in the vicinity largely for the remainder of the sample (a similar pattern is seen in the 6-month results as well). This proves interesting as Rapach and Wohar (2004) using the Andrews (1993) procedure as well as Bai and Perron (1998) procedure find evidence from both procedures of a structural break in the second quarter of 1975 for predicting 1-quarter real S&P 500 returns, however they argue the break does not meet statistical significance requirements.\(^{15}\) Based on the results of our time-varying parameter graphs there however does appear to be a shift which occurred around that time.

---

\(^{15}\) Rapach and Wohar (2004) consider real S&P 500 returns, which is slightly different from the excess returns used in our study, however these two measures are highly correlated. As well the data is quarterly, however for the case of short horizon predictability is still interesting for comparison sake with our monthly results.
Another interesting result is the default spread. Rapach and Wohar (2004) consider the default spread in their prediction of S&P 500 returns. Their Bai and Perron results indicate structural breaks in the third quarter of 1967, second quarter of 1968 and second quarter of 1975, with the estimated coefficient insignificant in regime 1 and 3, but much larger and significant in the second regime. Again, our time-varying parameter results seem to coincide with their results as the coefficient exhibits movements generally along the lines of what they find with their structural break estimates. However with our approach the entire path of the coefficient is available rather than just the estimated coefficients over the three regimes. Also of interest is that Pesaran and Timmerman (2002) estimate the coefficient for the default spread for predicting monthly NYSE excess returns using a rolling window estimation procedure with results similar to ours.

Now consider the term spread results. Rapach and Wohar (2004) consider structural breaks in their analysis of the term spread, finding evidence of a structural break in the third quarter of 1974, with the coefficient moving from positive and significant in the first regime to much smaller and insignificant for the post-break regime. Our results are similar, however as is evident from the graph, more dynamics are present. The time-varying coefficient is generally much larger over the pre-1975 part of the sample, however what the graph also indicates is the affect recessions appear to have on the coefficient as it tends to spike around many of the recessions, indicating the effect stages of the business cycle may have on the term spread as a predictor.

A similar business cycle dynamic appears to be occurring with the monthly excess return results. As mentioned before, lagged returns does not appear to have much ability to predict future returns at monthly horizons however looking closely at the shaded recessions, the series does possess predictive power for future returns during many of the downturns, especially those of the 1970s and 1980s.
5. Summary and Conclusions

This paper reconsiders stock return predictability using both macroeconomic and financial variables to predict 1-month, 6-month, and 12-month excess S&P 500 returns from the beginning of 1955 to the end of 2004. Stock predictability is reconsidered in the sense that we no longer assume a fixed-coefficient forecasting model, but rather use a bivariate Bayesian time-varying parameter model to test for predictability as well as determine whether this predictability has changed.

By allowing the parameters to vary, a number of important results emerge. First, for a number of variables and over a range of horizons, the coefficients exhibit a good degree of time variation, where this time variation proves to capture gradual changes as well as discrete changes in the parameters. This is an important point as it indicates the advantages of a time-varying parameter model over the discrete structural break models commonly used. Not only are more discrete breaks detailed, gradual change is also captured, and the results provide a fuller picture of how predictability has changed over time by showing the estimated parameter at every point in time. In addition, while discrete structural break tests generally require a trimming of the beginning and end of the sample, this is not required for the time-varying parameter model.

Second, the results indicate increased occurrences of predictability over the fixed coefficient case, and third, these results call into question the common practice of assuming a stable prediction model, and may explain part of the reason many variables fail to perform well in predicting stock returns out-of-sample.

In addition to the implications for modeling stock return predictions, there are a number of interesting results that inform us about predictability itself over the 1955-2004 period. First is the interesting finding is that in some cases it is evident that business cycles, especially recessions, appear to affect the degree of predictability. As well, a striking result is the degree that predictability seems to be strongest for the first half of the sample, with predictability having decreased in many cases over the last half. This is consistent with findings in other papers and may be suggestive of learning on the part of
investors. However it is also consistent with another story. Predictability seems strongest during the 1970s (and part of the 1980s) which coincides with a period of sizeable volatility in US markets, and which papers have argued may explain this period of predictability.

Finally, it is important to point out that the changing predictability does not necessarily rule out a market efficiency story. Under market efficiency it is feasible that opportunities may arise where a given variable is useful for predicting returns, and this information is taken advantage of until the opportunity disappears. As well, we may see predictability arise again using that variable, but for different reasons. This reoccurrence would satisfy market predictability as well as the reason for the predictability has changed. This may explain predictability that reoccurs but with a different sign or sized slope than the previous episode.

In terms of future extensions to this paper, a number of possible extensions come to mind. One is including marginal likelihood analysis to test statistically whether allowing for time-varying parameters better explains the data than a fixed coefficient model. Evidence in support of this would further substantiate the results found in this paper. As well, the forecasts for 6 and 12-month horizons could be extended to allow for overlapping observations which would bring more data to bear on this issue. Extending the model from bivariate to multivariate would also be interesting to see which variables perform well when included with others. Finally, given the interest on the issue of long horizon stock predictability, the horizons considered could be extended to include 2, 5, and 10-year horizons to see how the results change from those found in the fixed coefficient literature.

---

16 See for example, Kim, Morley and Nelson (2001).
17 In the next version of this paper, we intend to incorporate two-regime Markov-switching of residual errors to account for heteroskedasticity.
References


Appendix: Data

St. Louis Fed FRED II Database (http://research.stlouisfed.org/fred2):

- 3 Month T-Bill Rate: 3-Month Treasury Bill: Secondary Market Rate, Code TBSMS
- 10 Year T-Bill Yield: 10-Year Treasury Constant Maturity Rate, Code GS10
- Consumer Price Index: CPI for All Urban Consumers: All Items, Code CPIAUCNS
- Industrial Production: Industrial Production Index, Code INDPRO
- M1: M1 Money Stock, Code M1SL
- M2: M2 Money Stock, Code M2SL
- Unemployment: Civilian Unemployment Rate, Code UNRATE
- Moody’s Baa: Moody’s Seasoned Baa Corporate Bond Yield, Code BAA
- Moody’s Aaa: Moody’s Seasoned Aaa Corporate Bond Yield, Code AAA

Wharton Research Data Service (http://wrds.wharton.upenn.edu):

- S&P 500: Standard and Poor’s 500 Composite Index Value Weighted Return (Dividends included), Code SPINDX (VWRETD)
- Risk Free Rate: 1-month Treasury Bill Rate (Fama and French)


- Price-Earnings Ratio: Price and Earnings, See ie_data.xls
- Price-Dividend Ratio: Price and Dividends, See ie_data.xls
Appendix: Monthly Time Series Plots

- S&P 500 Excess Return (log)
- Inflation (log-difference)
- Industrial Production (log-difference)
- M1 (log-difference)
- M2 (log-difference)
- Unemployment Rate (change)
- 3 Month T-Bill Rate
- 10 YR T-Bond Yield
- Term Spread
- Default Spread
- Price-Earnings Ratio (log)
- Price-Dividend Ratio (log)
Appendix: Bayesian Approach\textsuperscript{18}

General Model Setup
Recall from Section 3, the following:

\textit{Measurement Equation}
\[ r_{t+k,t} = \beta_t Z_t + \epsilon_{t+k} \]

\textit{State Equation}
\[ \beta_t = \beta_{t-1} + v_t \]

where \( r_{t+k,t} \) is the cumulative S&P 500 excess return from period \( t \) to \( t+k \), \( \beta_t \) is a time-varying coefficient, \( Z_t \) is a macro or financial variable, and \( \epsilon_{t+k} \) is a forecast error. \( \epsilon_{t+k} \) is assumed i.i.d. normal with mean zero and variance \( R \), while \( v_t \) is assumed i.i.d. normal with mean 0 and variance \( Q \). Additionally, \( \epsilon_{t+k} \) and \( v_t \) are assumed to have nonzero covariance.

Now, expanding on the above, the unobserved state variable \( \beta_{t+1} \) evolves as follows

\[ p(\beta_{t+1} | \beta_t, V) \propto I(\beta_{t+1})f(\beta_{t+1} | \beta_0, V) \]

where rejection sampling is employed such that \( I(\beta_{t+1})=0 \) when the eigenvalue of the associated time-varying coefficient is outside the unit circle and 1 otherwise.\textsuperscript{19} \( V \) is a variance-covariance matrix described below and \( f(\beta_{t+1} | \beta_0, V) \) is normally distributed with mean \( \beta_{t+1} \) and variance-covariance \( Q \). \( f(\beta_{t+1} | \beta_0, V) \) follows a random walk process as described above.

We assume the innovations \( (\epsilon_{t+k}, v_t) \) are i.i.d. normal random variables with mean zero and variance-covariance matrix \( V \)

\textsuperscript{18} Note: subscript \( t \) represents the observation at time \( t \), while superscript \( t \) represents the history of observations up to and including time \( t \). \( f \) represents a normal distribution, while \( p \) represents a more general distribution.

\textsuperscript{19} Note: Rejection sampling is really only required when lagged values of the dependent variable are included in the regression, and so only in the case of lagged returns was it used.
\[ E_i \begin{bmatrix} \epsilon_{t+k} \\ V_i \end{bmatrix} \begin{bmatrix} \epsilon_t \\ V_t \end{bmatrix} = V = \begin{pmatrix} R & C \\ C & Q \end{pmatrix}, \]

where \( R \) is the variance of the measurement equation innovation, \( Q \) is the variance of the state equation innovation, and \( C \) is the covariance between the two innovations.

\( \beta \) is referred to as a parameter, while we refer to \( R, Q, \) and \( C \) as hyperparameters. The initial state \( \beta_0 \) is assumed to be a truncated Gaussian random variable, while the hyperparameters come from an inverse-Wishart distribution. These assumptions rely on their properties as natural conjugates.

Letting \( f(\beta_0) = N(\bar{\beta}, \bar{P}) \) represent a normal prior with mean \( \bar{\beta} \) and variance \( \bar{P} \), the prior for the initial state is

\[ p(\beta_0) \propto I(\beta) N(\bar{\beta}, \bar{P}), \]

while the prior for the hyperparameters is

\[ p(V) = IW(V^{-1}, T_0), \]

where \( IW(S, df) \) is an inverse-Wishart distribution with scale matrix \( S \) and degrees of freedom \( df \).

Together, the joint prior can be expressed as

\[ p(\beta_0, V) \propto I(\beta) N(\bar{\beta}, \bar{P}) IW(V^{-1}, T_0), \]

which combined with the likelihood function, provides the following joint posterior

\[ p(\beta^T, V | Y^T) \propto p(Y^T | \beta^T, V) p(\beta^T, V) \]

In order to simulate draws from this posterior, we employ Gibbs-sampling, which involves two steps. First, conditional on the data and hyperparameters, a history of states is drawn from \( p(\beta^T | Y^T, V) \). Then, conditional on the data and states, the hyperparameters are drawn from \( p(V | Y^T, \beta^T) \). This sequence of draws continues, undergoing a “burn-in period” (recall, 10,000 draws are used in our analysis, with 3,000 devoted to the burn-in period) until eventually converging such that the draws are equivalent to draws from the joint posterior, providing us with our parameters of interest.
**Discussion of Priors**

Priors are required for the starting value of the time-varying coefficient $\beta_0$ as well as the variance-covariance matrix $V$. Those chosen for this paper are best explained by first detailing the state space setup used more specifically.

*Measurement Equation*

$$r_{t+k,t} = [Z_t \ 1] \begin{bmatrix} \beta_t \\ \epsilon_{t+k} \end{bmatrix}$$

*Transition Equation*

$$\begin{bmatrix} \beta_t \\ \epsilon_{t+k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{t-1} \\ \epsilon_{t+k-1} \end{bmatrix} + \begin{bmatrix} v_t \\ \epsilon_{t+k} \end{bmatrix}$$

In the case of the state vector, the prior for $\begin{bmatrix} \beta_0 \\ \epsilon_0 \end{bmatrix}$ is based on a mean vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and a variance-covariance matrix $\begin{bmatrix} 2^{*} (\text{var}(r) / \text{var}(Z)) & 0 \\ 0 & \text{var}(r) \end{bmatrix}$.

This sets the starting value for the time-varying parameter in the Kalman filter at zero, however it incorporates a good deal of uncertainty, ensuring the Kalman gain puts little weight on the initial value (and relatively heavy weight on the new information) when updating the estimate of the parameter. Note the forecast error is part of the state vector, however receives zero weight in the transition equation, and so is constructed as a residual in the estimation.

In terms of the prior for the hyperparameters, this is based on a scale matrix $\begin{bmatrix} 0.05^{*} (\text{var}(r) / \text{var}(Z)) & 0 \\ 0 & \text{var}(r) \end{bmatrix}$, with 5 degrees of freedom. While the choice of 5 degrees of freedom is to allow for a fairly diffuse prior, the choice of scale matrix is
based on normalizing the distribution of our time-varying parameter such that one standard deviation for $\beta$ is equal to $(\text{std dev}(r)/\text{std dev}(Z))$.\textsuperscript{20}

We assume 5 percent of the variance is due to time variation.\textsuperscript{21} The prior also assumes the covariance of the time-varying parameter ($v_t$) and $\epsilon_{t+k}$ is zero. Finally the prior variance of $\epsilon_{t+k}$ is set as the full variance of the excess stock return series. This choice is due largely to the fact that previous studies show even with evidence of predictability, little of the variance of stock returns is actually explained by macroeconomic or financial variables.

\textsuperscript{20} Transforming this into variances we get $(\text{var}(r)/\text{var}(Z))$ as seen above.

\textsuperscript{21} In order to test the sensitivity of our results to our choice of prior for time variation of the time-varying parameter, a range of priors were considered and estimated which assumed 1%, 10%, 25%, and 100% time variation, as well as a range of degrees of freedom. Larger priors for time variation proved to pick up more noise in the dynamics of the time-varying parameter, but did not alter the lower frequency movements of the coefficient over time.