A Mixed Frequency Approach to Return and Dividend Growth Predictability

Anessa Custovic*

This Draft: October 1, 2019

This paper evaluates the predictive ability of a high frequency dividend price ratio. I show that we can uncover both dividend growth and return predictability across equity markets by weighting the dividends in each month differently through the application of MIDAS regressions. For US equity markets, I find strong predictability of long horizon real dividend growth with the theoretically correct sign and show that there is some predictability with nominal dividend growth as well. Out-of-sample I find that the high frequency constructed dividend price ratio can result in superior predictability for returns over the annual single frequency dividend price ratio. When combined with other predictors, the high frequency constructed dividend price ratio results in even better out-of-sample predictability for both returns and the equity premium. The out-of-sample performance is robust to horizons as well.

*I would like to thank Eric Ghysels for his extensive and very helpful comments. I also want to thank Andrii Babii and Mike Aguilar for their valuable comments. I have also benefited greatly from the many helpful comments and suggestions from seminar participants at UNC Chapel Hill.
1 Introduction

The return predictability literature has accumulated a large body of evidence documenting both the existence and lack thereof of market return predictability. Despite the contradictory findings, the general impression is that market returns are predictable (see Lettau and Ludvigson (2001) and Welch and Goyal (2007)) by a number of financial variables, you just have to find the right combination of variables. This paper poses the question: what if it is not just about the right combination of variables, but also the choice of sampling frequency in estimation?

One branch of the literature is mainly concerned with finding possible predictors while another analyzes return predictability in conjunction with dividend growth predictability. This paper considers the latter and provides a comprehensive study, examining the in-sample and out-of-sample predictability of market returns and dividend growth through the application of the mixed frequency data sampling (MIDAS regressions) approach of Ghysels et al. (2006). I examine formally whether systematic differences exist in the predictive ability of the mixed frequency regression approach and the single frequency approach commonly used in the literature for both dividend growth and market returns. Then, I consider the performance of the mixed frequency approach with the inclusion of control variables and at longer horizons.

The time series models used in the literature generally involve data sampled at the same frequency. However, some models might predict long-term market returns better than short-term returns, or vice-versa. Given this, it is not uncommon to look at multiple frequencies when exploring return predictability (see Welch and Goyal (2007), Ang (2011), Engsted and Pedersen (2009), Chen (2009), and others).

By only estimating models involving data sampled at the same frequency, we are not fully exploiting all of the available information. Consider the case of a simple annual return predictability regression model, which is the most common one used within the literature:

\[ r_{t+1} = \beta_0 + \beta_1 x_t + \epsilon_{t+1} \]  (1)
where $r$ is market return and $x$ is some predictor. In this case, we are only using some annual aggregate predictor. Though, this predictor may be available at a much higher frequency, say monthly. By estimating equation (1) at the annual aggregate frequency, we are implicitly forcing each month to have the same effect on returns. If this assumption does not hold true then the model might potentially be misspecified and potentially useful information may be destroyed.

Recently, econometric models that take into account information in unbalanced frequencies have attracted substantial attention. In particular, the application of the MIDAS regression approach has amassed a substantial literature. It is particularly attractive given that it exploits a much larger information set and is more flexible than say equation (1). Many studies have uncovered relationships between the data that was previously undetectable due to data aggregation. For example, Ghysels et al. (2005) used MIDAS volatility models and uncovered a significantly positive relation between risk and return. Clements and Galvo (2008) find that MIDAS regressions can lead to improvements in forecasting current and next quarter output growth. For the sake of brevity I omit many other examples from the literature.\footnote{For other applications of MIDAS regressions see for example Ghysels et al. (2006), Engle et al. (2013), Ghysels (2016), Andreou et al. (2013), Clements and Galvo (2009), among many others}

This paper explores whether or not relationships between the data was possibly undetectable due to data aggregation. Specifically, MIDAS regression weights are only applied to the monthly dividends in the dividend price ratio, the main predictor under consideration here. I allow dividends paid out in different months to be weighted differently and keep the end of year price constant when constructing the annual mixed frequency dividend price ratio. In many cases I uncover predictability (for dividend growth, market returns and the equity premium) and this finding is robust to horizon. This paper shows that we can uncover dividend growth predictability and that we can improve out-of-sample return predictability by leveraging higher frequency data.

Furthermore, I uncover statistically significant long horizon predictability for dividend growth for domestic aggregate equity markets. Here, I am able to estimate all coefficients with the theoretically correct sign and uncover statistical significance at long horizons.
for the both the SP500 and CRSP. In this paper, I also show that using MIDAS regressions can result in superior out-of-sample return predictability when compared to the conventional annual frequency.

The findings of this paper have implications for both the return predictability literature and for other applications in finance, such as portfolio management. Most applications of portfolio management require a robust method out-of-sample to forecast expected market returns. The mixed frequency dividend price ratio predictor proposed here meets much of that criteria, especially when combined with other predictors. There are important implications for the asset pricing literature in general here as well, especially in applications where the dividend price ratio is assumed to be a satisfactory proxy for expected stock returns.

1.1 Motivation and Related Literature

The present value decomposition from Campbell and Shiller (1988) suggests that the dividend price ratio must be related to either future returns or future dividend growth. It is possible for both future returns and future dividend growth to be predictable by the dividend price ratio, but at least one must be significantly related to it. Put differently, any variation in the dividend price ratio must be caused by either movement in expected returns, dividend growth, or both.

If the dividend price ratio were constant, then neither expected returns nor dividend growth would be forecastable by it. We know that the dividend price ratio is not constant over time: it does move. The present value identity then implies that at least one of either expected returns or dividend growth should be forecastable by the dividend price ratio. This relation is why the dividend price ratio is almost always chosen as a possible predictor of returns or dividend growth while other market ratios have assumed a secondary role.

Most of the research within the literature centers around return predictability, as the most common finding is that dividend price ratio is only significantly related to returns (see Lettau and Ludvigson (2005) Cochrane (2007), Lettau and Van Nieuwerburgh (2007), among others). Some studies found that small adjustments to the classic dividend
price ratio can even improve the relation with market returns. Lacerda and Santa-Clara (2010) are able to improve forecasts of both returns and dividend growth by incorporating the moving averages of past dividend growth into the dividend price ratio. Lettau and Nieuwerburgh (2008) adjust the dividend price ratio to accommodate for shifts in the steady state of the economy and find strong evidence of in-sample return predictability. A few studies that looked at international equity markets have uncovered a significant relationship between the dividend price ratio and dividend growth like Engsted and Pedersen (2009) and Rangvid et al. (2014). Maio and Santa-Clara (2015) find that there is strong evidence of dividend growth predictability by the dividend price ratio when one looks at the cross-section of stock returns. Kelly and Pruitt (2013) showed that by using cross-sectional firm disaggregated data one can successfully uncover a significant relationship between the dividend price ratio and dividend growth. Chen (2009) finds that the lack of detection of a significant relationship between dividend growth and the dividend price ratio for aggregate markets is in part due to postwar dividend smoothing by firms.

Recently, Asimakopoulos et al. (2017) use MIDAS regressions and regress high frequency (monthly) dividend price ratio growth data on low frequency (annual) dividend growth data. They decompose the dividend price ratio into two components: dividend price ratio growth, which is the high frequency component, and lagged annual dividend price ratio. Their findings suggest that it is the dividend price ratio growth that is the predictable component of the dividend price ratio.

This paper builds upon their findings and shows that this may not necessarily be the case. Specifically, here MIDAS regression weights are only applied to the monthly dividends in the dividend price ratio, as opposed to dividend price ratio growth as in Asimakopoulos et al. (2017). Asimakopoulos et al. (2017) only consider the relationship between the dividend price ratio and dividend growth. While I consider dividend growth, I also examine the relationship between the dividend price ratio and returns. To the best of my knowledge, this is the first paper to apply a mixed frequency approach to examine the relationship between the dividend price ratio and returns.
The general consensus within the literature continues to be that the dividend price ratio is mainly related to expected returns for aggregate equity markets, especially at longer horizons. The empirical findings for US equity markets within the literature are counter intuitive to the classic asset pricing theories. With the application of MIDAS regressions, I show that this is not the case.

This finding that the dividend price ratio is strongly related to market returns can be heavily impacted by the data. Ang (2011), Ang and Bekaert (2006) and Goyal and Welch (2003) find that returns are not predictable by the dividend price ratio when the 1990’s are included in the estimation period. Chen (2009) concluded that returns are strongly related to the dividend price ratio, but mainly in the post war period (1946-2005). This finding of predictability is also rarely validated out-of-sample, aside from that of Chen (2009), Welch and Goyal (2007), Lacerda and Santa-Clara (2010) and Goyal and Welch (2003). In this paper, I show that using MIDAS regressions can result in superior out-of-sample return predictability when compared to the conventional annual frequency.

The paper is organized as follows: section 2 describes the data and methodology used in the analysis. It will also provide a brief description of MIDAS regressions and the Almon polynomial. Section 3 presents empirical findings at the annual horizon. Section 4 presents empirical findings at long horizons. Finally, section 5 concludes the paper.

2 Data and Methodology

In this section, I discuss the Campbell and Shiller (1988) present value identity in more detail. Then I present a more in formal overview of the econometric methodology utilized within this paper. Finally, this section will detail the source and construction of the data used in empirical estimation.

2.1 The Present Value Model

To derive the present value model we first start with the definition of returns, where asset returns are comprised of capital gains yield and dividend yield. Specifically, we
write returns as $R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}$, where $P$ denotes the price and $D$ denotes dividends. Let $r_t = \log(1 + R_t)$, $p_t = \log(P_t)$, and $d_t = \log(D_t)$. Taking a first-order Taylor approximation yields the original Campbell and Shiller (1988) log linearized model.

$$dp_t = -\kappa + r_{t+1} - \Delta d_{t+1} + \rho(dp_{t+1})$$

Where $\rho = \exp(p - d)/(1 + \exp(p - d)) = (P/D)/(1 - (P/D)) \in (0, 1)$ is the log linearization discount coefficient and $\kappa = -\ln(1 - \rho) - \rho\ln((1/1 - \rho) - 1)$ is defined as a constant.

Following Cochrane (2007) and iterating the Campbell and Shiller (1988) equation $k$ times,

$$dp_t = \frac{-\kappa(1 - \rho^k)}{1 - \rho} + \sum_{j=1}^{k} \rho^{j-1}r_{t+j} - \sum_{j=1}^{k} \rho^{j-1}\Delta d_{t+j} + \rho^k(dp_{t+k})$$

Equation (2) gives us the present-value relation, which says that the current log dividend price ratio is positively related with future discounted log returns and future discounted log dividend price ratio at time $t + k$. From the relation, the log dividend price ratio is negatively related with future discounted log dividend growth. This gives us the theoretically correct sign for dividend price ratio when we estimate later. In the limit, if we assume no bubbles the present value decomposition states that the dividend price is equal to a constant plus the difference between the future discounted returns and the future discounted dividend growth.

Methods commonly used in the literature to estimate the model are long horizon regressions and VAR’s. One convention is to follow Cochrane (2007) and to directly estimate weighted long-horizon regressions of future log returns, log dividend growth and log dividend to price ratio on the current dividend to price ratio. The second method is to run a first-order VAR as in Cochrane (2007), Engsted and Pedersen (2009), Maio and Santa-Clara (2015), and others.

### 2.2 The Almon Polynomial and MIDAS Regressions

First I briefly discuss the Almon polynomial and distributed lag models, then I move onto MIDAS regressions. If we have reason to believe that our dependant variable, $y_t$, is
affected by more than one lag of our explanatory variables, $x_{t-(M-1)},...,x_{t-1}, x_t$, we can write a distributed lag model in the following form:

$$y_t = \sum_{k=0}^{K-1} w_k x_{t-k} + \epsilon_t \tag{3}$$

where $K$ is equal to the lag length.

The Almon polynomial distributed lag structure was put forth by Almon (1965) and has since become one of the most popular lag structures implemented in distributed lag models. The polynomial is defined as follows:

$$w_k = \sum_{j=0}^{p} \theta_j i^j \tag{4}$$

where $k = 0, 1, 2, ..., K - 1$ and $p$ is the degree of the polynomial. Given that the order of the polynomial is much lower than that of the lag length, $K$, the resulting Almon distributed lag model can be estimated parsimoniously via OLS.

Define a MIDAS regression model as follows with a basic single high frequency regressor (h-step ahead).

$$y_{t+1}^L = \beta_0 + \beta_1 \sum_{k=0}^{K-1} \omega_k(\theta) x_{t-k/K}^L + \epsilon_{t+h}^L$$

The MIDAS regression weights are governed by polynomial specifications and the parameters $\theta$ govern the weighting scheme. The weighting scheme is purely data driven, there are no assumptions required for estimation and MIDAS does not suffer from parameter proliferation.

Combining equation (3) with (4) is a specific case of MIDAS regressions, which can be estimated via OLS. Throughout this paper, I will use the Almon polynomial for estimation with degree set to two. A quadratic Almon polynomial is able to take on many shapes, they can be very similar to the Beta weighting function and the Exponential Almon weighting function. Estimating via OLS will allow us to stay very close to the Campbell and Shiller (1988) framework and utilize evaluation methods which will ease

\[\text{For a more detailed treatment of MIDAS weighting schemes see Ghysels et al. (2007)}\]
any comparisons within the literature.

The Almon polynomial, as specified above, does not necessarily result in MIDAS regression weights that sum to one. The estimation of Almon lags in MIDAS regressions via OLS requires properly transformed high frequency data regressors. Once the weights are estimated, given that they do not sum to zero, they can be re-scaled to obtain the slope coefficients and normalized weights. I recover the slope coefficient and normalize the weights so that they do sum to one throughout this paper. Hence, the equations specified in the following sections will be written under this assumption.\footnote{If we reported results where the MIDAS regression weights were not normalized to 1 and the slope coefficient was not recovered, then the slope and weights would be specified as \( \beta w_i(\theta_0, \theta_1, \ldots, \theta_p) = w^i_m = \sum_{j=0}^{p} \theta_j i^j \).}

By normalizing the weights after estimation I ensure that the weights sum to one, then our high-frequency slope, \( \beta_1 \), is identified. This is what allows me to weigh each monthly observation (dividends) differently. Motivated by the findings of Asimakopoulos et al. (2017), I focus on monthly data rather than quarterly. Asimakopoulos et al. (2017) demonstrated that even aggregating at the quarterly frequency resulted in a loss of useful information.

All of the empirical analysis that follows was also estimated via MIDAS profiling (put forth by Ghysels and Qian (2019)) with the exponential Almon and Beta weighting schemes. The results and conclusions are largely the same. I choose to report results estimated with the Almon polynomial MIDAS regressions technique rather than the MIDAS profiling method so as to avoid any possible generated regressor issues.

### 2.3 Data

The bulk of this analysis is concerned with domestic equity market data. The main domestic data are from the CRSP value-weighted portfolio and from the SP500 value-weighted portfolio. Both domestic data sets span from January 1927 to December 2017. These are the two most commonly used data sets within the literature, see Cochrane (2007), Maio and Santa-Clara (2015), Engsted and Pedersen (2009), Campbell and Shiller (1988), and others. The CRSP data is a total market measures, it consists of all firms...
listed on US stock exchanges. CRSP consists of small, medium, and large cap stocks while SP500 only has large caps. I also consider the Russell 2000 Index obtained from Factset. The data set spans from 1980 to 2018.

Some of the analysis will consider international equity market data. I follow Asimakopoulos et al. (2017) and Engsted and Pedersen (2009) who look at international equity markets alongside domestic. The international data are from Factset and consist of three different large equity market indices. I consider the FTSE All Shares Index, the FTSE Euro First 300 Index, and the Canada SP/TSX Composite Index. All of the international data sets span from 1987 to 2018.

Within the literature, there are two common ways of constructing the dividend price ratio. Cochrane (2007), Maio and Santa-Clara (2015) and many others do assume that monthly dividends are reinvested in the stock market. Under reinvestment, the dividend price ratio is constructed from the difference between the value-weighted returns with dividends and the value-weighted returns excluding dividends.\footnote{See Cochrane (2007) for an in depth explanation of calculations.}

Campbell and Shiller (1988), Chen (2009), Ang (2011), Ang and Bekaert (2006) and others construct the dividend price ratio assuming that monthly dividends are not reinvested in the stock market. Under this assumption, monthly dividends are recovered from the value-weighted returns with dividends and the value-weighted returns excluding dividends. The monthly dividends are then summed to get annual dividends.\footnote{See Chen (2009) or Ang (2011) for an in depth explanation of calculations without re-investment.} Here, I assume no reinvestment and apply MIDAS regression weights to only the monthly dividends.

3 Annual Empirical Results

3.1 One Year Ahead Predictability

In this section, the empirical results for equity markets are presented. All empirical results that follow in this section are for one step ahead (annual) horizons. Before we look at the empirical results of the models we will first briefly look at the data.
3.2 Properties of the data

Table 1 displays the annual sample statistics for both the annual dividend growth rate and the returns. I report the annual mean, maximum, minimum, standard deviation and auto-correlation ($\phi$). There are no issues of persistence with either dividend growth or returns. We can also see that the sample statistics for both CRSP and the SP500 are very similar.

Figure 1 plots the annually constructed dividend price ratios\(^6\) for both CRSP and the SP500. We see that they track one another rather closely. Maio and Santa-Clara (2015) demonstrated that there are in fact differences between what drives small cap vs. large cap stocks. CRSP consists of small, medium, and large cap stocks while the SP500 is only tilted towards large cap stocks. As such, I expect to see differences in the monthly weights.

Table 1: Sample Statistics for Annual Dividend Growth and Returns
This table provides sample mean, max, min, standard deviation and $\phi$ for the annual dividend growth, annual return and annual dividend price ratio. Panel A is for the CRSP data set. Panel B is for the SP500 data set.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Std</th>
<th>$\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: CRSP Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta d^L$</td>
<td>0.05</td>
<td>0.43</td>
<td>-0.51</td>
<td>0.12</td>
<td>0.28</td>
</tr>
<tr>
<td>$r^L$</td>
<td>0.09</td>
<td>0.45</td>
<td>-0.56</td>
<td>0.19</td>
<td>0.04</td>
</tr>
<tr>
<td>$dp^L$</td>
<td>-3.41</td>
<td>-2.35</td>
<td>-4.48</td>
<td>0.45</td>
<td>0.89</td>
</tr>
<tr>
<td><strong>Panel B: SP500 Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta d^L$</td>
<td>0.05</td>
<td>0.45</td>
<td>-0.63</td>
<td>0.13</td>
<td>0.21</td>
</tr>
<tr>
<td>$r^L$</td>
<td>0.09</td>
<td>0.42</td>
<td>-0.58</td>
<td>0.19</td>
<td>0.05</td>
</tr>
<tr>
<td>$dp^L$</td>
<td>-3.38</td>
<td>-2.29</td>
<td>-4.45</td>
<td>0.46</td>
<td>0.88</td>
</tr>
</tbody>
</table>

3.3 The Benefits of High Frequency Data

Within this subsection I explore the potential gains from utilizing high-frequency data by looking at large equity market data. I only allow the monthly dividends to be weighted differently in the construction of the dividend price ratio. In contrast, Asimakopoulos et al. (2017) allowed the monthly dividend price ratio growth rate to be weighted differ-

---

\(^6\)While I am unable to reject the null hypothesis of a unit root with the Augmented Dickey-Fuller test for the annual dividend price ratio (both CRSP and SP500), I am able to reject it for the high frequency dividend price ratio data that is used in estimating the MIDAS regression.
Figure 1: CRSP and SP500 Dividend Price Ratio
This figure depicts the annually constructed dividend price ratio for CRSP and the SP500.

ently. To test the possibility that dividends paid in certain months matter more than other months of the year, we will first explore annual one year ahead regressions for large equity markets. I then compare the results using the mixed frequency constructed dividend price ratio to that of the conventional annually constructed dividend price ratio.

Table 2 presents the empirical results of the one year ahead regressions for both the CRSP and SP500 data sets. The following models are estimated,

\[ \Delta d_{t+1}^L = \beta_0 + \beta_1 dp_t^L(\omega) + \epsilon_{t+1}^d \]  
\[ r_{t+1}^L = \beta_0 + \beta_1 dp_t^L(\omega) + \epsilon_{t+1}^r \]

where, \( dp_t^L(\omega) = \frac{\sum_{j=1}^{12} \omega_j d_{j,t}^H}{p_{12,t}} \) is the MIDAS regression weighted high frequency data.

The model is estimated via OLS using the Almon polynomial specification. Recall that estimation of Almon lags in MIDAS regressions via OLS requires properly transformed high frequency data regressors. Once the weights are estimated via OLS, I rescale them to obtain the slope coefficient and normalized weights.

We can see from Table 2 that there are benefits to using high frequency data. Conventional estimation involves equal weights, which implicitly forces each month to be as important as the rest. For the annual CRSP data, there is only a statistically significant
Table 2: Annual Regressions
This table provides empirical results for one-year ahead predictability. The dependent variables are the annual log dividend growth and annual log returns. The models are estimated via OLS. The resulting weights are normalized and the slope coefficient is recovered. T-NW reports the calculated Newey-West t-statistic for each estimate with one lag. T-H reports the calculated Hodrick (1992) t-statistic for each estimate with one lag. As usual, *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>t-NW</th>
<th>t-H</th>
<th>R²</th>
<th>Coef.</th>
<th>t-NW</th>
<th>t-H</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Almon</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRSP ∆d</td>
<td>-0.14</td>
<td>-2.67**</td>
<td>-2.41***</td>
<td>0.23</td>
<td>-0.08</td>
<td>-1.68*</td>
<td>-1.49</td>
<td>0.09</td>
</tr>
<tr>
<td>SP500 ∆d</td>
<td>-0.15</td>
<td>-2.2**</td>
<td>-2**</td>
<td>0.18</td>
<td>-0.09</td>
<td>-1.62</td>
<td>-1.46</td>
<td>0.09</td>
</tr>
<tr>
<td>CRSP r</td>
<td>0.08</td>
<td>0.93</td>
<td>0.84</td>
<td>0.03</td>
<td>0.06</td>
<td>1.05</td>
<td>0.96</td>
<td>0.02</td>
</tr>
<tr>
<td>SP500 r</td>
<td>0.07</td>
<td>0.9</td>
<td>0.82</td>
<td>0.03</td>
<td>0.06</td>
<td>1.12</td>
<td>1.03</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Annual</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

relationship with dividend growth at the 10% level, when looking at the Newey and West (1987) t-statistics. When estimating with high frequency data, there is now a statistically significant relationship with dividend growth at the 1% level, judged by both the Newey and West (1987) t-statistic and the Hodrick (1992) t-statistic. For the SP500, we only see a statistically significant relationship between the dividend price ratio and dividend growth when we estimate with Almon weights.

Given that neither the annual, nor mixed frequency approach resulted in a statistically significant relationship between the dividend price ratio and returns, it is difficult to draw any conclusions in-sample. However, it is important to note that there does exist a statistically significant relationship between returns and the mixed frequency regressor if we consider different sampling periods.\textsuperscript{7}

We can note that the mixed frequency approach resulted in slightly larger estimated coefficients and $R^2$.\textsuperscript{8} From Table 2 we can see that this result of the mixed frequency approach having slightly larger estimated coefficients holds true when either dividend growth or returns are the dependent variables. In regards to the $R^2$ coefficients for dividend growth, the mixed frequency approach resulted in an $R^2$ that is almost 150% larger than that of the annual approach. This is true for both CRSP and SP500 data. For returns, we see a more modest increase in the $R^2$ of 50% for both data sets.

\textsuperscript{7}When the same mixed frequency model is estimated with the data spanning from 1950-2017, I find it is highly statistically significant. I also find it to be highly statistically significant when I only include the past 30 years of data.

\textsuperscript{8}Note, this result holds for the different sampling periods as well. The mixed frequency model results in slightly larger coefficients and $R^2$'s when data from 1950-2017 is used and when only the past 30 years is included.
For robustness, I also conduct an F-test test. The estimated mixed frequency coefficients are a linear combination of the estimated monthly weights. The null hypothesis is that all weights are jointly equal to zero. The resulting p-values when the dependant variable is returns, coincides with the results of Table 2 for both CRSP and SP500 data. For dividend growth, the CRSP data has statistical significance at well below the 5% level. For the SP500 however, the resulting p-value from the F-test is 15%.

Figure 2: CRSP and SP500 Dividend Price Ratio Weighting Schemes
This figure depicts the MIDAS regressions Almon weighting schemes for equation (5).

![Figure 2](image)

Figure 3: CRSP and SP500 Dividend Price Ratio Weighting Schemes
This figure depicts the MIDAS regressions Almon weighting schemes for equation (6).

![Figure 3](image)

Figures 2 and 3 plot the optimal weighting schemes from estimating equations (5)
and (6). The two weighting schemes differ for dividend growth and for returns, even though the sample statistics looked very similar. Figure 4 plots the dividend price ratios constructed from each of the four weighting schemes from estimation. While the weighting schemes from estimating equation (6) for both the SP500 and CRSP are rather similar, the weighting schemes from estimating equation (5) differ. The SP500 weighting scheme is almost linearly increasing from January through December. For CRSP, it is slightly more curved, with January being weighted almost 1.5 times more than in the SP500 weighting scheme from Figure 2.

Figure 4: CRSP and SP500 Dividend Price Ratios
This figure depicts the mixed frequency constructed dividend price ratio for CRSP and the SP500, both returns and dividend growth weighting schemes.

The results that dividends paid in the month farthest out in time are weighted the most is consistent with the findings of Ball and Easton (2013). Ball and Easton (2013) examine the intra-year relationship between earnings of a firm and news for that firm, where news is annual returns calculated as the sum of daily price change plus daily dividend payments divided by beginning of year price. They find that news at the beginning of year \( t \) is incorporated into the earnings of year \( t \). They also find that news at the beginning of year \( t \) is incorporated into the earnings of year \( t \) and the earnings of year \( t + 1 \). Ball and Easton (2013) explain that there is less time for news at the end of the year \( t \) to be incorporated into the earnings. For example, news on trading day 250 would then only have 1 trading day to be incorporated into the earnings.
If dividends are a fraction of earnings, then it is possible that dividends in year $t$ and $t+1$ are affected by news at the beginning of the year. Dividend growth is defined as $d_{t+1} - d_t$. Hence, it is also possible that dividend growth is affected mainly by news at the beginning of the year since it is calculated from dividends in year $t$ and $t+1$. This effect might be what we are seeing in Figure 2.

Asimakopoulos et al. (2017) conclude that it is only the monthly dividend price growth which is the predictable component of dividend growth. Here, it is demonstrated that this might not necessarily be the case. The application of MIDAS regressions above shows that the monthly dividend price ratio can still produce the theoretically correct sign and significance. The results above also suggest that it matters how we construct our dividend price ratio when exploiting the high frequency data. If I allow the monthly dividends paid out to vary across the year and use the year end price, the dividend price ratio is still a predictable component of dividend growth.

### 3.4 Dividend Growth and Returns Over Time

There is evidence within the literature of a reversal of return and dividend growth predictability. Chen (2009) explores this in great detail, Lettau and Ludvigson (2005) and Engsted and Pedersen (2009) also examine the post war US data. Chen (2009) shows that there does exist a strong relationship between dividend growth and the dividend price ratio, but only if you use pre war data. In the post war period Chen (2009) finds that this dividend growth predictability tends to disappear and that instead returns become strongly related to the dividend price ratio.

Engsted and Pedersen (2009) demonstrate that this reversal does not necessarily hold true when you look at both nominal and real returns and dividend growth. Their results show that real dividend growth is unpredictable in the pre war period, but strongly predictable in the post war period. It is important to mention that this significant predictability in the post war period was in the wrong direction, that is the estimated coefficient had the theoretically incorrect sign.

In this section, I explore whether there is a reversal in predictability over time and
whether or not it holds true with real or nominal data. I estimate in-sample expanding MIDAS regressions, so that we can see how the estimated coefficients, t-statistics and $R^2$ change over time. This approach is similar to that of Goyal and Welch (2003), where they reported time varying coefficients for their dividend models. Similarly to Welch and Goyal (2007), I begin the expanding window⁹ regressions 20 years after data is available. This brings us close to the post war period. If there is a reversal in predictability we should see this reflected in the estimation results.

Throughout this section I will continue comparing the mixed frequency regressions to the conventional annual regressions. The equations are specified as follows:

$$y_{t+1}^L = \beta_0 + \beta_1 dp_t^L(\omega) + \epsilon_{t+1} \quad (7)$$

$$y_{t+1} = \beta_0 + \beta_1 dp_t + \epsilon_{t+1} \quad (8)$$

where $y_{t+1}^L$ in equation (7) and (8) is either returns or dividend growth (nominal and real) and $dp_t^L(\omega)$ is the MIDAS regression weighted dividend price ratio. Equation (8) is the usual regression run in the literature, all are annual data.

We can see from Figure 5 that the estimated slopes from the CRSP data for returns generally start higher near the post war period. There is also an increase in the coefficients from the 1970’s through the early 1990’s. This pattern is reflected in the t-statistics sub plot in Figure 5. We see the t-statistics peak in the early 1990’s. This is in line with what Ang (2011) found, that estimating with the 1990’s destroys return predictability. Ang and Bekaert (2006) and Goyal and Welch (2003) also find that return predictability is not detectable when the 1990’s are included in the estimation period. Figure 7 with the SP500 data set has the same pattern for returns.

We can see from Figure 6, however, that we also have benefits from using the mixed frequency approach when it comes to dividend growth predictability over time. The estimated t-statistics of the CRSP slopes are statistically significant at or below the 5% level starting in the post-war period and remain so until the end of the data period. This

---

⁹By implementing expanding windows we are able to utilize a much larger sample size than in rolling window regressions, and thereby increasing the power of the tests.
Figure 5: Expanding Window Returns Regressions (CRSP Data)

This figure plots the expanding window estimated return coefficients, their respective t-statistics and R² for the mixed frequency $dp_t$ and the annual $dp_t$. Each is estimated with real and nominal returns. The sample is 1928 to 2017.

(a) CRSP Slopes

(b) CRSP t-statistics

(c) CRSP R²
Figure 6: Expanding Window Dividend Growth Regressions (CRSP Data)
This figure plots the expanding window estimated dividend growth coefficients, their respective t-statistics and R2 for the mixed frequency $d_{pt}$ and the annual $d_{pt}$. Each is estimated with real and nominal dividend growth. The sample is 1928 to 2017.

((a)) CRSP Slopes

((b)) CRSP t-statistics

((c)) CRSP R2
is only true when we estimate using high frequency data. Using annual data, we see that in 1997 we lose statistically significant predictability with nominal dividend growth. For real dividend growth we lose statistically significant predictability in 2000. Figure 8 echos the results of the CRSP data set for the SP500.

The $\beta$ slope coefficient is a linear combination of the estimated 12 monthly weights. Hence, I also test whether or not the 12 monthly weights estimated are all equal to zero via a F-test. For returns, the p-values from the F-test test coincide with the associated p-values from the reported t-statistics. This is true for both the SP500 and CRSP. Figures 12 and 13 in the Appendix plot the p-values from the F-test.

For CRSP dividend growth, the p-values from the F-test generally coincide with the p-values reported in Figure 6. Figure 10 in the Appendix plots the p-values from the F-test over time. We can see that the mixed frequency approach with real dividend growth associated p-values are near zero throughout the entire time period. Figure 10 also shows that for nominal dividend growth the associated p-values are statistically significant below the 5% level for all but eight years reported. Put differently, the p-values are statistically significant below the 5% level for the mixed frequency approach about 89% of the sample period compared to the 70% of the sample period when using the annual approach.

Figure 11 in the Appendix echos the results of the CRSP data set for the SP500, with slight differences. For the SP500, we can see that the mixed frequency approach with real dividend growth associated p-values are near zero throughout the entire time period. There are only six years in the sample period where the p-value goes slightly above 5%, but remains well below 10%. Figure 11 also shows that for nominal dividend growth the associated p-values are statistically significant below the 5% level for all years until 2000, when it goes slightly above the 10% significance level.

Not only did the mixed frequency approach result in larger estimated coefficients for both the SP500 and CRSP data, but in general, higher $R^2$'s as well. For dividend growth (nominal and real) the mixed frequency estimated $R^2$'s surpass the annual model $R^2$ throughout the entire sample period. For returns (nominal and real), the mixed frequency estimated $R^2$'s surpass the annual model $R^2$ starting in the 1990's. This is
Figure 7: Expanding Window Returns Regressions (SP500 Data)
This figure plots the expanding window estimated return coefficients, their respective t-statistics and R2 for the mixed frequency $dp_t$ and the annual $dp_t$. Each is estimated with real and nominal returns. The sample is 1928 to 2017.

((a)) SP500 Slopes

((b)) SP500 t-statistics

((c)) SP500 R2
This figure plots the expanding window estimated dividend growth coefficients, their respective t-statistics and R2 for the mixed frequency $d_{pt}$ and the annual $d_{pt}$. Each is estimated with real and nominal dividend growth. The sample is 1928 to 2017.

((a)) SP500 Slopes

((b)) SP500 t-statistics

((c)) SP500 R2
particularly interesting since Ang (2011), Ang and Bekaert (2006) and Goyal and Welch (2003) found that return predictability is destroyed if the 1990’s are included in the estimation sample. Here, we can see that the mixed frequency estimation approach loses less of it’s explanatory power for return predictability than the annual approach.\footnote{I conducted the same analysis, but only considered data from 1950-2017. I find much larger return coefficients and $R^2$’s for market returns. The finding that the mixed frequency approach results in larger estimated coefficients for both the SP500 and CRSP data and in higher $R^2$’s was also true during this data period.}

### 3.5 Out-Of-Sample Performance

In sections 3.3 and 3.4 we saw that the gains of utilizing mixed frequency data are maintained throughout the sample. In the return predictability literature it is commonplace to test predictability out-of-sample since the in-sample significance is sometimes undetectable. In this section I explore whether or not the mixed frequency approach maintains superiority out-of-sample. I estimate out-of-sample regressions similar to Welch and Goyal (2007). As per Welch and Goyal (2007), I estimate the expanding out-of-sample MIDAS regressions 20 years after the sample starts and also starting in 1965.

Following Welch and Goyal (2007) and Chen (2009), I estimate equations (7) and (8). I also estimate the following,

$$y_{t+1} = \beta_0 + \epsilon_{t+1}$$

where $y_{t+1}$ is $r_{t+1}$ or $\Delta d_{t+1}$. Equation (9) assumes no predictability, it is the benchmark historical mean model. Equations (7) and (8) assume there exists a relationship between the variables.

First, denote the mean-square error (MSE) calculated from equation (9) as $MSE_N$ and denote the MSE calculated from either equations (7) and (8) as $MSE_A$. I then calculate the following out-of-sample $R^2$ (OOS R2) based off of the error terms:

$$OOSR^2 = R^2 - (1 - R^2) \times \frac{T - k}{T - 1}$$

where $R^2 = 1 - \frac{MSE_A}{MSE_N}$. As per Welch and Goyal (2007) and Chen (2009) I also report
the $\Delta RMSE = \sqrt{MSE_N} - \sqrt{MSE_A}$ and the MSE-F statistic from McCracken (2007).
A positive $\Delta RMSE$ implies that equations (7) or (8) are out performing the benchmark historical mean model out-of-sample. Put differently, a positive $\Delta RMSE$ suggests that the dividend price ratio is a better predictor than the historical mean.

The MSE-F statistic is calculated as follows:

$$MSE - F = (T - h + 1) \times \frac{MSE_N - MSE_A}{MSE_A}. \quad (11)$$

Following Welch and Goyal (2007) and Chen (2009), significance is based off of bootstrapped\(^{11}\) MSE-F statistics. One advantage of the MSE-F statistic is that it is scaled, unlike $MSE$, and can be compared across horizons.

I do not have much success in predicting dividend growth out-of-sample with neither annual nor mixed frequency data, so I omit those results. Within this subsection, I mainly focus on predicting market returns out-of-sample. Engsted and Pedersen (2009) found that there are differences in-sample when you use real or nominal data. I will explore both nominal and real returns to see if differences also exist out-of-sample.

During the expanding window out-of-sample evaluation period, I am not only re-estimating the coefficients, but also the high frequency weighting scheme for the dividend price ratio. If I conducted the out-of-sample evaluation with the optimal full sample weighting scheme applied to the dividend price ratio at each time $t$, and only re-estimated the coefficient, there would be consistently positive and statistically significant out-of-sample performance. This is true for both the CRSP and SP500 data sets at the 1% level. However, this would result in look ahead bias so I do not report these results, but only the ones where I am re-estimating the high frequency weighting scheme.

Tables 3 and 4 report the out-of-sample results for returns. We see from Table 3 that the annual dividend price ratio and the mixed frequency dividend price ratio cannot successfully predict returns when we start forecasts 20 years after the sample begins.

Though, from Table 4 we can also see that when the mixed frequency approach is used, it results in a consistently positive $\Delta RMSE$ and OOS R2 for the SP500. While

\(^{11}\)For a more in depth treatment see Welch and Goyal (2007).
Table 3: Out-Of-Sample Performance- 20 Years After the Sample
This table provides the out-of-sample results for both the mixed frequency and annual methods, denoted as HF and Ann respectively. The dependent variables are returns, both nominal and real. All numbers reported are in percentage terms. \( \Delta RMSE \) is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in the section.

<table>
<thead>
<tr>
<th></th>
<th>( \Delta RMSE )</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>( \Delta RMSE )</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nominal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRSP( ^{HF} ) ( r_{t+1} )</td>
<td>-0.22</td>
<td>-2.7</td>
<td>-1.87</td>
<td>-0.37</td>
<td>-4.34</td>
<td>-2.95</td>
</tr>
<tr>
<td>CRSP( ^{Ann} ) ( r_{t+1} )</td>
<td>-0.06</td>
<td>-0.68</td>
<td>-0.48</td>
<td>-0.14</td>
<td>-1.69</td>
<td>-1.12</td>
</tr>
<tr>
<td>SP( ^{HF} ) ( r_{t+1} )</td>
<td>-0.12</td>
<td>-1.52</td>
<td>-1.06</td>
<td>-0.3</td>
<td>-3.64</td>
<td>-2.49</td>
</tr>
<tr>
<td>SP( ^{Ann} ) ( r_{t+1} )</td>
<td>-0.06</td>
<td>-0.68</td>
<td>-0.48</td>
<td>-0.14</td>
<td>-1.69</td>
<td>-1.12</td>
</tr>
<tr>
<td><strong>Real</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

the annually constructed dividend price ratio for the SP500 has poor out-of-sample performance, the mixed frequency approach has statistical significance at the 5% level when it comes to nominal returns.

The mixed frequency approach also results in a statistically significant out-of-sample performance for nominal returns for both the SP500 and the CRSP data sets, while the annual dividend price ratio has poor performance for nominal returns. Thus, we can see that there are benefits to the mixed frequency approach when it comes to predicting returns. It is important to note this out performance only exists when the out-of-sample evaluation period begins in 1965.

Table 4: Out-Of-Sample Performance- 1965-2017
This table provides the out-of-sample results for both the the mixed frequency and annual methods, denoted as HF and Ann respectively. The dependent variables are returns, both nominal and real. All numbers reported are in percentage terms. \( \Delta RMSE \) is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in the section.

<table>
<thead>
<tr>
<th></th>
<th>( \Delta RMSE )</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>( \Delta RMSE )</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nominal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRSP( ^{HF} ) ( r_{t+1} )</td>
<td>0.12</td>
<td>1.38</td>
<td>0.74*</td>
<td>-0.04</td>
<td>-0.52</td>
<td>-0.27</td>
</tr>
<tr>
<td>CRSP( ^{Ann} ) ( r_{t+1} )</td>
<td>-0.06</td>
<td>-0.72</td>
<td>-0.38</td>
<td>-0.15</td>
<td>-1.72</td>
<td>-0.89</td>
</tr>
<tr>
<td>SP( ^{HF} ) ( r_{t+1} )</td>
<td>0.25</td>
<td>3.03</td>
<td>1.65**</td>
<td>0.04</td>
<td>0.47</td>
<td>0.25</td>
</tr>
<tr>
<td>SP( ^{Ann} ) ( r_{t+1} )</td>
<td>-0.06</td>
<td>-0.72</td>
<td>-0.38</td>
<td>-0.15</td>
<td>-1.72</td>
<td>-0.89</td>
</tr>
<tr>
<td><strong>Real</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For robustness, I now consider four other equity markets, three of which are international. I will consider the Russell 2000 Index, the FTSE All Shares Index, the FTSE Euro First 300 Index, and the Canada SP/TSX Composite Index. From Table 5 we see
that for three out of the four indices reported, the mixed frequency approach yields very similar results to that of the annual approach. The mixed frequency approach appears to have better predictive power for the FTSE Euro First 300 Index, resulting in a $\Delta RMSE$ and an OOS R2 almost twice as high as with the annual approach.

Though unreported, I also explore the out-of-sample performance of dividend growth for the Russell 2000 Index, the FTSE All Shares Index, the FTSE Euro First 300 Index, and the Canada SP/TSX Composite Index. The Canada SP/TSX Composite Index is the only one with a positive and statistically significant dividend growth out-of-sample performance. To see the out-of-sample performance over time, I construct the Welch and Goyal (2007) out-of-sample plots. These can be seen in the Appendix.

Table 5: Out-Of-Sample Performance- 20 Years After the Sample
This table provides the out-of-sample results for both the mixed frequency and annual methods, denoted as HF and Ann respectively. The dependent variables are nominal returns. All numbers reported are in percentage terms. $\Delta RMSE$ is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in the section.

<table>
<thead>
<tr>
<th></th>
<th>$\Delta RMSE$</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>$\Delta RMSE$</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russell</td>
<td>0.3</td>
<td>3.12</td>
<td>0.61</td>
<td>0.49</td>
<td>4.6</td>
<td>0.63</td>
</tr>
<tr>
<td>FTSE All</td>
<td>1.67</td>
<td>20.57</td>
<td>3.11**</td>
<td>1.72</td>
<td>21.17</td>
<td>3.22**</td>
</tr>
<tr>
<td>FTSE EUR</td>
<td>0.43</td>
<td>3.57</td>
<td>0.48*</td>
<td>0.21</td>
<td>1.72</td>
<td>0.23</td>
</tr>
<tr>
<td>Canada</td>
<td>-3.75</td>
<td>-29.37</td>
<td>-2.95</td>
<td>-0.17</td>
<td>-1.25</td>
<td>-0.16</td>
</tr>
</tbody>
</table>

3.6 Out-Of-Sample Performance with Controls
Now that we have seen that there are gains in predictability when the mixed frequency data is utilized, I consider adding predictors from Welch and Goyal (2007) to see if that improves out-of-sample predictability. Within this subsection I mainly focus on predicting market returns (real and nominal) and the equity premium (real and nominal) out-of-sample.

I first only consider predictors at the annual frequency that have data available since 1927 or earlier, so as to not sacrifice sample length. Then, I further limit the possibilities by only considering predictors that are not highly correlated with the dividend price ratio. This immediately excludes a number of price ratio predictors.
Additionally, I check that potential predictors have a somewhat poor out-of-sample performance themselves. I do this to ensure that any improvement in the out-of-sample performance is not simply a result of the predictor itself having a strong out-of-sample performance. For example, had I included the consumption, wealth, income ratio (cay) of Lettau and Ludvigson (2001) the out-of-sample performance would have improved dramatically, but only because cay itself is a strong out-of-sample predictor of returns. Note, I also do not consider cay as it has forward looking information that would not be available at the time of the forecast as mentioned by Welch and Goyal (2007).

I will consider two different Welch and Goyal (2007) predictors along with the mixed frequency dividend price ratio: percent equity issuing (eqis) and the default return spread (dfr). The percent equity issuing ratio contains information that the dividend price ratio may not necessarily have like IPO’s. Likewise, the default return spread is a measure of risk that the dividend price ratio might not contain.

It is important to note that both percent equity issuing and default return spread had negative out-of-sample performance individually, regardless of what year the forecast evaluation began. When combined, the two had even worse out-of-sample performance. Hence, if the addition of the mixed frequency dividend price ratio resulted in positive out-of-sample performance, we can be confident it is due to the inclusion of the mixed frequency predictor and not being driven by the performance of default return spread or percent equity issuing.

Both the measure of corporate issuing activity and default return spread could improve the out-of-sample performance of the mixed frequency dividend price ratio. To test this, I estimate the following two expanding window equations:

\[ y_{t+1} = \beta_0 + \beta_1 dp_t^L(\omega) + \beta_2 eqis_t^L \epsilon_{t+1} \]  

Equations (12) and 13 are estimated via OLS using the Almon polynomial specification. Once the weights are estimated via OLS, I re-scale them to obtain the slope coefficient and normalized weights. For this reason, I am able to write the equations with the normalized weights and recovered slope coefficients.

\[ y_{t+1} = \beta_0 + \beta_1 dp_t^L(\omega) + \beta_2 eqis_t^L \epsilon_{t+1} \]  

\[ y_{t+1} = \beta_0 + \beta_1 dp_t^L(\omega) + \beta_2 eqis_t^L \epsilon_{t+1} \]
\[ y_{t+1}^L = \beta_0 + \beta_1 dp_t^L(\omega) + \beta_2 eqis_t^L + \beta_3 df r_t^L + \epsilon_{t+1} \]  

where \( y^L \) is either returns \( (r_{t+1}) \) or the equity premium \( (r_{t+1}^{ep}) \) (nominal and real). I do not bootstrap the MSE-F statistics, but rather obtain the p-values from the tables in McCracken (2007). Notice that I do not allow for more than three predictors at a time, this is to avoid over-fitting.

Table 6: Out-Of-Sample Performance- CRSP Data
This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are either returns or the equity premium, both nominal and real. Panel A reports results for equation 12 while Panel B reports results for equation 13. All numbers reported are in percentage terms. \( \Delta RMSE \) is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the McCracken (2007) asymptotic values.

<table>
<thead>
<tr>
<th></th>
<th>( \Delta RMSE )</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>( \Delta RMSE )</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: equation 12</td>
<td>20 yrs</td>
<td>1965</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_{t+1}^{nom} )</td>
<td>0.45</td>
<td>5.32</td>
<td>3.99***</td>
<td>0.54</td>
<td>6.35</td>
<td>3.67**</td>
</tr>
<tr>
<td>( r_{t+1}^{real} )</td>
<td>0.6</td>
<td>6.94</td>
<td>5.29***</td>
<td>0.53</td>
<td>6.06</td>
<td>3.42**</td>
</tr>
<tr>
<td>( r_{t+1}^{eqnom} )</td>
<td>0.51</td>
<td>6.07</td>
<td>4.59***</td>
<td>0.64</td>
<td>7.42</td>
<td>4.25**</td>
</tr>
<tr>
<td>( r_{t+1}^{eqreal} )</td>
<td>0.64</td>
<td>7.27</td>
<td>5.57***</td>
<td>0.6</td>
<td>6.63</td>
<td>3.70**</td>
</tr>
<tr>
<td>Panel B: equation 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_{t+1}^{nom} )</td>
<td>0.06</td>
<td>0.67</td>
<td>0.48*</td>
<td>0.26</td>
<td>3.03</td>
<td>1.65*</td>
</tr>
<tr>
<td>( r_{t+1}^{real} )</td>
<td>0.2</td>
<td>2.37</td>
<td>1.72**</td>
<td>0.25</td>
<td>2.91</td>
<td>1.59*</td>
</tr>
<tr>
<td>( r_{t+1}^{eqnom} )</td>
<td>0.08</td>
<td>0.94</td>
<td>0.67**</td>
<td>-0.03</td>
<td>-0.39</td>
<td>-0.21</td>
</tr>
<tr>
<td>( r_{t+1}^{eqreal} )</td>
<td>0.2</td>
<td>2.33</td>
<td>1.69**</td>
<td>0.27</td>
<td>3.02</td>
<td>1.65*</td>
</tr>
</tbody>
</table>

From Tables 6 and 7 we can see that the mixed frequency approach combined with additional predictors results in positive and significant out-of-sample performance for both CRSP and the SP500. The results are particularly strong for equation (12), where with the CRSP data we have statistical significance at the 1% level for all dependent variables when forecasts begin 20 years after the sample starts.

Though unreported, I do run equations (12) and (13) with the annual dividend price ratio to compare. When it comes to predicting market returns (both nominal and real) out-of-sample, it is the mixed frequency dividend price ratio that consistently outperforms the annual dividend price ratio. Overall, we can see that there are gains in out-of-sample predictability when mixed frequency data is utilized. This can be seen in Tables 4 and 5 above. However, it is the inclusion of additional predictors (corporate issuing activ-
ity and default return spread) with the mixed frequency dividend price ratio that results in consistent, statistically significant out-of-sample predictability of returns regardless of when the evaluation period begins.

To evaluate the out-of-sample performance over time for both equations (12) and (13), I construct the Welch and Goyal (2007) out-of-sample plots. These can be seen in the Appendix. From the plots we can see that equation (12) has consistent positive performance out-of-sample since the 1970’s. There is a slight dip around the mid 2000’s in out-of-sample performance, but there appears to be a turn around since 2009.

Table 7: Out-Of-Sample Performance- SP500 Data
This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are either returns or the equity premium, both nominal and real. Panel A reports results for equation 12 while Panel B reports results for equation 13. All numbers reported are in percentage terms. ∆RMSE is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the McCracken (2007) asymptotic values.

<table>
<thead>
<tr>
<th></th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: equation 12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 yrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_{t+1}^{\text{nom}} )</td>
<td>0.25</td>
<td>3.05</td>
<td>2.23**</td>
<td>0.53</td>
<td>6.51</td>
<td>3.69**</td>
</tr>
<tr>
<td>( r_{t+1}^{\text{real}} )</td>
<td>0.42</td>
<td>4.91</td>
<td>3.66**</td>
<td>0.45</td>
<td>5.38</td>
<td>3.01**</td>
</tr>
<tr>
<td>( r_{t+1}^{\text{reqnom}} )</td>
<td>0.23</td>
<td>2.77</td>
<td>2.02**</td>
<td>0.49</td>
<td>6.01</td>
<td>3.39**</td>
</tr>
<tr>
<td>( r_{t+1}^{\text{realeal}} )</td>
<td>0.41</td>
<td>4.7</td>
<td>3.5**</td>
<td>0.41</td>
<td>4.67</td>
<td>2.6**</td>
</tr>
<tr>
<td>1965</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_{t+1}^{\text{nom}} )</td>
<td>-0.07</td>
<td>-0.91</td>
<td>-0.64</td>
<td>0.19</td>
<td>2.39</td>
<td>1.3*</td>
</tr>
<tr>
<td>( r_{t+1}^{\text{real}} )</td>
<td>0.06</td>
<td>0.77</td>
<td>0.55**</td>
<td>0.13</td>
<td>1.59</td>
<td>0.85*</td>
</tr>
<tr>
<td>( r_{t+1}^{\text{reqnom}} )</td>
<td>-0.11</td>
<td>-1.36</td>
<td>-0.95</td>
<td>0.14</td>
<td>1.68</td>
<td>0.9*</td>
</tr>
<tr>
<td>( r_{t+1}^{\text{realeal}} )</td>
<td>0.04</td>
<td>0.43</td>
<td>0.31*</td>
<td>0.06</td>
<td>0.7</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Motivated by the findings of Lacerda and Santa-Clara (2010), I now consider a moving average dividend growth control variable. Similarly to Lacerda and Santa-Clara (2010), I create a 10 year moving average dividend growth variable. I then lag it to make sure there is no look ahead bias when estimating. I estimate the following two expanding window equations:

\[
y_{t+1} = \beta_0 + \beta_1 d_{t}^{L} (\omega) + \beta_2 \Delta d_{t,MA}^{L} + \epsilon_{t+1} \tag{14}
\]

\[
y_{t+1} = \beta_0 + \beta_1 d_{t}^{L} (\omega) + \beta_2 eqis_{t}^{L} + \beta_3 \Delta d_{t,MA}^{L} + \epsilon_{t+1} \tag{15}
\]

where \( y_t^L \) is either returns (\( r_{t+1} \)) or the equity premium (\( r_{t+1}^{ep} \)) (nominal and real), \( \Delta d_{t,MA}^{L} \)
is the 10 year moving average of dividend growth, and eqis is percent equity issuing. I do not bootstrap the MSE-F statistics, but rather obtain the p-values from the tables in McCracken (2007).

Table 8: Out-Of-Sample Performance- Equation 15

This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are either returns or the equity premium, both nominal and real. Panel A reports results for the CRSP data while Panel B reports results for the SP500 data. All numbers reported are in percentage terms. ∆RMSE is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the McCracken (2007) asymptotic values.

<table>
<thead>
<tr>
<th></th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: CRSP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r_{t+1}^{nom}</td>
<td>0.78</td>
<td>9.21</td>
<td>6.29***</td>
<td>0.79</td>
<td>9.27</td>
<td>5.52***</td>
</tr>
<tr>
<td>r_{t+1}^{real}</td>
<td>0.52</td>
<td>5.95</td>
<td>3.92**</td>
<td>0.42</td>
<td>4.84</td>
<td>2.75**</td>
</tr>
<tr>
<td>r_{t+1}^{epnom}</td>
<td>0.73</td>
<td>8.39</td>
<td>5.68***</td>
<td>0.68</td>
<td>7.82</td>
<td>4.58***</td>
</tr>
<tr>
<td>r_{t+1}^{epreal}</td>
<td>0.55</td>
<td>6.13</td>
<td>4.05***</td>
<td>0.44</td>
<td>4.83</td>
<td>2.74**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: SP500</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r_{t+1}^{nom}</td>
<td>0.91</td>
<td>10.97</td>
<td>7.64***</td>
<td>0.64</td>
<td>7.66</td>
<td>5.14***</td>
</tr>
<tr>
<td>r_{t+1}^{real}</td>
<td>0.62</td>
<td>7.32</td>
<td>4.89***</td>
<td>0.49</td>
<td>5.59</td>
<td>3.67***</td>
</tr>
<tr>
<td>r_{t+1}^{epnom}</td>
<td>0.76</td>
<td>9.28</td>
<td>5.52***</td>
<td>0.42</td>
<td>5.05</td>
<td>2.87***</td>
</tr>
<tr>
<td>r_{t+1}^{epreal}</td>
<td>0.48</td>
<td>5.64</td>
<td>3.23**</td>
<td>0.32</td>
<td>3.62</td>
<td>2.03**</td>
</tr>
</tbody>
</table>

Both CRSP and the SP500 had positive, statistically significant out-of-sample performance when equation (14) was estimated. For brevity, I do not report the results of equation (14) since they are only an improvement on equation (13), but not on equation (12). The out-of-sample results for equation (15) can be found in Table 8. From Table 8 we can see that there are substantial improvements in the OOS R2’s for both CRSP and the SP500 over both Tables 6 and 7 if we add a lagged moving average dividend growth control variable.

To evaluate the out-of-sample performance over time for equation (15), I construct the Welch and Goyal (2007) out-of-sample plots. These can be seen in the Appendix. From the plots we can see that equation (15) has consistent positive performance out-of-sample since the start of the out-of-sample evaluation period. There is a dip in performance during the 1990’s, which is consistent with the findings of Chen (2009) and Ang and Bekaert (2006). Despite this fall in out-of-sample predictability the past 20 years out-of-sample are better than that of equation (12). In conclusion, we can see that the
mixed frequency approach combined with a slow moving dividend growth component and percent equity issuing can result in consistently positive and highly statistically significant out-of-sample performance for both CRSP and the SP500.

4 Long Horizon Empirical Results

In this section, I continue to present empirical results for large equity markets. All empirical results that follow are for long horizons. We first look at the in-sample long horizon results and then the out-of-sample long horizon results.

4.1 Long Horizon Regressions

Within this subsection I explore predictability at longer horizons following the method of Ang (2011), Engsted and Pedersen (2009), Chen (2009), and many others. Specifically, I directly estimate\(^{15}\) the following long horizon regressions of future log returns, log equity premium and log dividend growth:

\[
\sum_{j=1}^{k} r_{t+j}^{L} = a_r^{k} + b_r^{k} dp_t^{L} (\omega) + \epsilon_{r+1}^{L}
\]

\[
\sum_{j=1}^{k} \Delta d_{t+j}^{L} = a_d^{k} + b_d^{k} dp_t^{L} (\omega) + \epsilon_{d+1}^{L}
\]

\[
\sum_{j=1}^{k} r_{ep}^{L} t+j = a_{r, ep}^{k} + b_{r, ep}^{k} dp_t^{L} (\omega) + \epsilon_{r, ep}^{L+1}
\]

where again, \(dp_t^{L} (\omega)\) is our MIDAS regression weighted dividend price ratio. The null hypothesis of no predictability I want to test for each equation is as follows: \(b_r^{k} = 0\) for equation (16), \(b_d^{k} = 0\) for equation (17), and \(b_{r, ep}^{k} = 0\) for equation (18). An acceptance of the null hypothesis implies that the MIDAS regression weighted dividend price ratio has no predictive power at long horizons.

\(^{15}\)Equations (16), (17) and (18) are estimated via OLS using the Almon polynomial specification. Once the weights are estimated via OLS, I re-scale them to obtain the slope coefficient and normalized weights. For this reason, I am able to write the equations with the normalized weights and recovered slope coefficients.
I estimate over-lapping long horizon regressions, which results in an error term ($\epsilon_{t+k}^d$, $\epsilon_{t+k}^{rep}$ or $\epsilon_{t+k}^r$), that is no longer serially uncorrelated, but rather follows a MA($k - 1$) process. With a large enough sample and if the horizon is not too long, then the OLS estimator will be consistent. Hence, I am more so concerned with correct inference as this is what becomes distorted due to the over-lapping observations used to construct the long horizon return or dividend growth.

Within the literature it is common to use Newey and West (1987) t-statistics or Hodrick (1992) t-statistics for inference. Valkanov (2003) and Liu and Maynard (2007) have demonstrated that they may not be appropriate for inference of long horizon regressions. Liu and Maynard (2007) showed that the Newey and West (1987) t-statistic has poor size for various specifications\textsuperscript{16} of the $x_t$ (independent) variable. Ang (2011) also demonstrated that the Newey and West (1987) t-statistics did not have high power for long horizon inference.

Liu and Maynard (2007) show that the Campbell and Dufour (1995) and Campbell and Dufour (1997) sign and signed rank tests can be used to conduct inference on long horizon regressions. In order to use the sign and signed rank methods for long horizon regressions, I first rearrange the long horizon equations defined earlier in the section. First, consider the following:

\[
\sum_{j=1}^{k} y_{t+j} = \beta_0 + \beta(k)x_t + \epsilon_{t+k}
\]  

where in our case $y$ is either returns, the equity premium or dividend growth and $x$ is $dp^L_t(\omega)$. The null hypothesis of no predictability is $\beta(k) = 0$.

Liu and Maynard (2007) suggest rearranging the long horizon regression to follow an approach similar to Jegadeesh (1991) and Cochrane (1991). Now, define the long horizon regression instead as:

\[
y_{t+1} = \alpha_0 + \alpha(k) \sum_{j=1}^{k} x_{t+j-k} + \epsilon_{t+1}
\]  

\[\text{\textsuperscript{16}Specifically, they simulate size and power when } x_t \text{ is generated by an AR(1) process, a historical break model, a markov switching model, a long-memory model and a STOP-BREAK model.}\]
where once again $y$ is either returns, equity premium or dividend growth and $x$ is $dp_L^T(\omega)$.

This rearrangement will avoid the MA($k-1$) process in the error terms. The null hypothesis $\beta(k) = 0$ for equation (19) is the same as the null hypothesis $\alpha(k) = 0$ for equation (20). Because of this, Jegadeesh (1991) and Cochrane (1991) suggest testing the $\alpha(k) = 0$ hypothesis instead of the hypothesis $\beta(k) = 0$. Under this specification, a rejection of the null $\alpha(k) = 0$ also implies a rejection of the null $\beta(k) = 0$.

Following the notation of Liu and Maynard (2007), the original sign ($S_T$) and signed rank ($SR_T$) tests from Campbell and Dufour (1995) and Campbell and Dufour (1997) are as follows:

$$S_T(y_{t+1}, g_t, r, s) = \sum_{t=1+r}^{T-s} u((y_{t+1} - b_0)g_t)$$

$$SR_T(y_{t+1}, g_t, r, s) = \sum_{t=1+r}^{T-s} u((y_{t+1} - b_0)g_t)R_{t+1}^+(b_0)$$

where $u$ is an indicator function indicating whether or not $y_{t+1} - b_0$ is greater than or equal to zero and $R_{t+1}^+(b_0) = \sum_{t=j+1}^{T-s} u(|y_{t+1} - b_0| - |y_j - b_0|)$. In equations (21) and (22), $r$ is the rank and $s$ is the number of observations that were truncated. Define $x_k^{**}$ as the median centered\(^{18}\) value of $\sum_{j=1}^{k} x_{t+j-k}$. Again, following Liu and Maynard (2007), I test $S_T(y_{t+1}, x(\omega)_k^{**}, k-1, 1)$, which follows a binomial distribution, and $SR_T(y_{t+1}, x(\omega)_k^{**}, k-1, 1)$, which is described by the Wilcoxon signed-rank test.

The simulations of Liu and Maynard (2007) demonstrated that the above sign and signed rank test perform very well when applied to long horizon regression inference. They are slightly conservative, but they do not over reject. Given how robust the tests are to long horizon regression inference, I only report p-values from the sign and signed rank tests.

From Tables 9 and 10 we can see that there is long horizon predictability of dividend growth by $dp_L^T(\omega)$ and that it is primarily for real dividend growth. Engsted and Pedersen (2009) also found some dividend growth predictability, but the estimated coef-

\(^{17}\)It is important to note that Liu and Maynard (2007) demonstrate that this rearrangement holds true even when $x_t$ is non-stationary.

\(^{18}\)Campbell and Dufour (1997) tried a number of centering methods and showed that this type improves test power.
Table 9: Long Horizon Predictability- CRSP Dividend Growth and Returns

This table provides empirical results for long horizon predictability for both dividend growth and returns (nominal and real). Below, $K$ denotes the horizon in years, Coef. is the estimated coefficient, $S_T$ is as defined in equation 21, and $SR_T$ is as defined in equation 22. As usual, *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th>$K$</th>
<th>Coef. $S_T$</th>
<th>$SR_T$</th>
<th>$R^2$</th>
<th>Coef. $S_T$</th>
<th>$SR_T$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dividend Growth</td>
<td></td>
<td></td>
<td>Returns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nominal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.14</td>
<td>0.37</td>
<td>0.16</td>
<td>0.04</td>
<td>0.45</td>
<td>0.37</td>
</tr>
<tr>
<td>10</td>
<td>-0.18</td>
<td>0.09***</td>
<td>0.04**</td>
<td>0.08</td>
<td>0.62</td>
<td>0.22</td>
</tr>
<tr>
<td>15</td>
<td>-0.35</td>
<td>0.45</td>
<td>0.38</td>
<td>0.15</td>
<td>0.84</td>
<td>0.15</td>
</tr>
<tr>
<td>20</td>
<td>-0.38</td>
<td>0.09***</td>
<td>0.25</td>
<td>0.19</td>
<td>0.81</td>
<td>0.32</td>
</tr>
<tr>
<td>real</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.15</td>
<td>0.01***</td>
<td>0.01***</td>
<td>0.1</td>
<td>0.43</td>
<td>0.29</td>
</tr>
<tr>
<td>10</td>
<td>-0.18</td>
<td>0.01***</td>
<td>0.01***</td>
<td>0.12</td>
<td>0.62</td>
<td>0.33</td>
</tr>
<tr>
<td>15</td>
<td>-0.3</td>
<td>0.01***</td>
<td>0.01***</td>
<td>0.18</td>
<td>0.89</td>
<td>0.36</td>
</tr>
<tr>
<td>20</td>
<td>-0.11</td>
<td>0.03**</td>
<td>0.02**</td>
<td>0.09</td>
<td>1.07</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Coefficients sometimes had the theoretically incorrect sign. Also, they were unable to detect a significant relationship for dividend growth in the CRSP data set. By using MIDAS regressions and leveraging higher frequency data, I am able to estimate all coefficients with the theoretically correct signs.

The lack of statistical significance (for the most part) in Tables 9 and 10 between long horizon returns and $dp_L(t, \omega)$ is not entirely surprising. Chen (2009) was unable to detect long horizon significance when he constructed his dividend price ratio with dividends unreinvested. Ang (2011) was also unable to detect a significant relationship between the dividend price ratio and returns, unless the 1990’s were omitted from estimation.

Results for equation (18) are reported in Table 11. We can see from the top panel of the table that the dividend price ratio for both CRSP and the SP500 does not have any statistically significant predictability at horizons equal to five, ten, fifteen or twenty for nominal equity premium. Though, there is some evidence of predictability at horizons fifteen and twenty for real equity premium for the CRSP data. For the SP500, there is evidence of predictability at a horizon of fifteen years.

The tables reported here underestimate the long horizon predictability found for both CRSP and the SP500 as only four horizons are shown. Figure 9 plots the p-values from the sign and signed rank tests from a horizon of two years to a horizon of twenty years.
Figure 9: Long Horizon $S_T$ and $SR_T$ p-values

This figure plots the p-values of the long horizon predictive coefficients for both CRSP and the SP500 for the signed test and signed rank test. Each plot has the p-values using real and nominal returns/dividend growth/equity premium. The sample is 1928 to 2017.

((a)) CRSP $\Delta d$

((b)) CRSP $r$

((c)) CRSP $r^{ep}$

((d)) SP500 $r^{ep}$

((e)) Sp500 $\Delta d$

((f)) SP500 $r$
Table 10: Long Horizon Predictability - SP500 Dividend Growth and Returns

This table provides empirical results for long horizon predictability for both dividend growth and returns (nominal and real). Below, \( K \) denotes the horizon in years, Coef. is the estimated coefficient, \( S_T \) is as defined in equation 21, and \( SR_T \) is as defined in equation 22. As usual, *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th>( K )</th>
<th>Coef.</th>
<th>( S_T )</th>
<th>( SR_T )</th>
<th>( R^2 )</th>
<th>Coef.</th>
<th>( S_T )</th>
<th>( SR_T )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-0.14</td>
<td>0.09*</td>
<td>0.03**</td>
<td>0.04</td>
<td>0.45</td>
<td>0.16</td>
<td>0.12</td>
<td>0.45</td>
</tr>
<tr>
<td>10</td>
<td>-0.14</td>
<td>0.13</td>
<td>0.5</td>
<td>0.04</td>
<td>0.62</td>
<td>0.29</td>
<td>0.39</td>
<td>0.32</td>
</tr>
<tr>
<td>15</td>
<td>-0.3</td>
<td>0.06*</td>
<td>0.13</td>
<td>0.11</td>
<td>0.85</td>
<td>0.24</td>
<td>0.14</td>
<td>0.37</td>
</tr>
<tr>
<td>20</td>
<td>-0.3</td>
<td>0.04**</td>
<td>0.19</td>
<td>0.1</td>
<td>0.86</td>
<td>0.45</td>
<td>0.23</td>
<td>0.23</td>
</tr>
</tbody>
</table>

(K=2 to K=20). We can see that some real return predictability does exist for the SP500, and some nominal return predictability exists for CRSP at the 10% level. We can also see that there are some pockets of nominal dividend growth predictability in both CRSP and the SP500.

Figure 9 also shows us that the predictability of real dividend growth remains statistically significant through most of the horizons. This holds true for both CRSP and the SP500. Specifically, the CRSP real dividend growth predictability is statistically significant, at the 10% level, at 100% of the horizons tested above when judged by both the sign and signed rank tests. For CRSP nominal dividend growth is statistically significant at the 10% level at about 50% of the horizons when judged by both tests. Though we see only one or two instances of real and nominal market returns being statistically significant, the real equity premium has a number of statistically significant horizons at the 10% level.

For the SP500, there is significance for real dividend growth predictability at the 10% level about 89% of the time, when judged by both tests. The SP500 nominal dividend growth predictability has a number of horizons with statistical significance at the 10% level. Similarly to the CRSP results, there are only one or two horizons of real and nominal returns being statistically significant. The real equity premium is significant at...
Table 11: Long Horizon Equity Premium Predictability

This table provides empirical results for long horizon equity premium (nominal and real) predictability for both CRSP and SP500. Below, $K$ denotes the horizon in years, Coef. is the estimated coefficient, $S_T$ is as defined in equation 21, and $SR_T$ is as defined in equation 22. As usual, *, **, *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th>$K$</th>
<th>Coef.</th>
<th>$S_T$</th>
<th>$SR_T$</th>
<th>$R^2$</th>
<th>Coef.</th>
<th>$S_T$</th>
<th>$SR_T$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRSP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nominal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.38</td>
<td>0.46</td>
<td>0.32</td>
<td>0.22</td>
<td>0.39</td>
<td>0.33</td>
<td>0.32</td>
<td>0.19</td>
</tr>
<tr>
<td>10</td>
<td>0.49</td>
<td>0.41</td>
<td>0.25</td>
<td>0.43</td>
<td>0.52</td>
<td>0.5</td>
<td>0.46</td>
<td>0.42</td>
</tr>
<tr>
<td>15</td>
<td>0.75</td>
<td>0.28</td>
<td>0.19</td>
<td>0.51</td>
<td>0.79</td>
<td>0.12</td>
<td>0.17</td>
<td>0.49</td>
</tr>
<tr>
<td>20</td>
<td>1.1</td>
<td>0.5</td>
<td>0.3</td>
<td>0.53</td>
<td>1.21</td>
<td>0.45</td>
<td>0.46</td>
<td>0.47</td>
</tr>
<tr>
<td>SP500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|     |       |       |        |       |       |       |        |       |
| CRSP |       |       |        |       |       |       |        |       |
| real |       |       |        |       |       |       |        |       |
| 5  | 0.36  | 0.33  | 0.2    | 0.21  | 0.37  | 0.33  | 0.38   | 0.18  |
| 10 | 0.49  | 0.18  | 0.15   | 0.35  | 0.51  | 0.25  | 0.31   | 0.37  |
| 15 | 0.8   | 0.08*** | 0.05** | 0.42  | 0.82  | 0.01*** | 0.01*** | 0.44  |
| 20 | 1.37  | 0.12  | 0.07*** | 0.51  | 1.47  | 0.31  | 0.3    | 0.49  |

the 10% level at about 32% of the horizons tested above.

4.2 Long Horizon Out-of-Sample Evaluation

Now that we have seen that there does exist statistically significant predictability at long horizons, I explore the out-of-sample performance here. Again, following Welch and Goyal (2007), I estimate expanding out-of-sample long horizon regressions 20 years after the sample begins. I do this for horizons equal to five, ten, fifteen and twenty years. out-of-sample performance is again measured by the $\Delta RMSE$, the OOS $R^2$ and the MSE-F statistic as defined in section 3.5. The MSE-F p-values reported below are all derived from the bootstrap procedure referenced in section 3.5.

I evaluate performance of (real and nominal) dividend growth, returns, and the equity premium at various horizons. Nominal dividend growth and nominal returns for both CRSP and the SP500 have poor long horizon out-of-sample performance so they are not reported here. Table 12 reports the out-of-sample performance for real dividend growth, for both CRSP and the SP500. We can see from the table that both have statistically significant positive out-of-sample performance for horizons of ten, fifteen and twenty years.

Likewise, Table 13 reports the out-of-sample performance for real returns. We see
Table 12: Out-Of-Sample Performance- Real Dividend Growth
This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are long horizon real dividend growth. All numbers reported are in percentage terms. $\Delta RMSE$ is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in earlier sections.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>$\Delta RMSE$</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>$\Delta RMSE$</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRSP</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>SP500</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.24</td>
<td>-2.68</td>
<td>-1.67</td>
<td>-0.25</td>
<td>-2.62</td>
<td>-1.63</td>
</tr>
<tr>
<td>10</td>
<td>2.32</td>
<td>21.29</td>
<td>14.61</td>
<td>1.96</td>
<td>18.22</td>
<td>12.03</td>
</tr>
<tr>
<td>15</td>
<td>2.33</td>
<td>17.31</td>
<td>9.21</td>
<td>1.9</td>
<td>14.57</td>
<td>7.5</td>
</tr>
<tr>
<td>20</td>
<td>0.88</td>
<td>6.65</td>
<td>2.42</td>
<td>1</td>
<td>7.22</td>
<td>2.65</td>
</tr>
</tbody>
</table>

Table 13: Out-Of-Sample Performance- Real Returns
This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are long horizon real returns. All numbers reported are in percentage terms. $\Delta RMSE$ is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in earlier sections.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>$\Delta RMSE$</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>$\Delta RMSE$</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRSP</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>SP500</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-1.62</td>
<td>-9.59</td>
<td>-5.6</td>
<td>-1.84</td>
<td>-10.13</td>
<td>-6.16</td>
</tr>
<tr>
<td>10</td>
<td>4.21</td>
<td>16.05</td>
<td>10.32</td>
<td>4.26</td>
<td>14.99</td>
<td>9.52</td>
</tr>
<tr>
<td>15</td>
<td>10.32</td>
<td>31.67</td>
<td>22.71</td>
<td>10.81</td>
<td>30.72</td>
<td>19.51</td>
</tr>
<tr>
<td>20</td>
<td>9.45</td>
<td>28.55</td>
<td>13.59</td>
<td>8.96</td>
<td>25.49</td>
<td>11.63</td>
</tr>
</tbody>
</table>

results similar to that of Table 12, statistically significant positive out-of-sample performance for horizons of ten, fifteen and twenty years for both the SP500 and CRSP. For the SP500, it appears that a horizon of five years always results in poor out-of-sample performance for all dependant variables while horizons of ten, fifteen and twenty years result in positive out-of-sample predictability.

The same general pattern appears to hold true for the CRSP data as well. We can see from Table 14 that once again, at horizon equal to five years we have negative out-of-sample performance. However, Table 15 tells a slightly different story for the CRSP data. When it comes to predicting the real equity premium, the CRSP data has positive out-of-sample performance for all four horizons considered, with three being highly statistically significant.

We can see from the results within this section that the out-of-sample $R^2$ tends to increase monotonically with respect to the horizons$^{19}$ under consideration. Table 12 is

$^{19}$Note, we are only considering four different horizons, so this may not necessarily hold at a more granular level where horizon is increased by one year.
Table 14: Out-Of-Sample Performance- Real Equity Premium

This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are long horizon real equity premiums. All numbers reported are in percentage terms. ∆RMSE is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in earlier sections.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRSP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.2</td>
<td>1.03</td>
<td>0.67</td>
<td>-0.18</td>
<td>-0.87</td>
<td>-0.55</td>
</tr>
<tr>
<td>10</td>
<td>8.65</td>
<td>26.46</td>
<td>19.43***</td>
<td>9.27</td>
<td>26.97</td>
<td>19.94***</td>
</tr>
<tr>
<td>15</td>
<td>15.31</td>
<td>34.37</td>
<td>23.04***</td>
<td>16.42</td>
<td>35.47</td>
<td>24.18***</td>
</tr>
<tr>
<td>20</td>
<td>21.46</td>
<td>39.95</td>
<td>22.62***</td>
<td>19.53</td>
<td>35.95</td>
<td>19.09***</td>
</tr>
<tr>
<td>SP500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

an exception, where it appears that the performance improves up through a horizon of ten years and then slightly declines. This is evident in the decreasing MSE-F statistic and OOS R2 for horizons of fifteen and twenty years. For Tables 13 and 15 the OOS R2 increases up through a horizon of fifteen years with a slight decrease in a horizon of twenty years. Table 14 has an increasing OOS R2 through all horizons under consideration for both the SP500 and CRSP.

Table 15: Out-Of-Sample Performance- Nominal Equity Premium

This table provides the out-of-sample results for the mixed frequency estimation. The dependent variables are long horizon nominal equity premiums. All numbers reported are in percentage terms. ∆RMSE is the RMSE difference between the unconditional and conditional forecast for the same sample/forecast period. A positive number signifies superior OOS conditional forecast. The OOS statistics are calculated as reported in section 3.5. Significance levels are based off of the bootstrap procedure described in earlier sections.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
<th>∆RMSE</th>
<th>OOS R2</th>
<th>MSE-F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRSP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.76</td>
<td>4.6</td>
<td>3.23***</td>
<td>-0.09</td>
<td>-0.49</td>
<td>-0.31</td>
</tr>
<tr>
<td>10</td>
<td>9.3</td>
<td>35.2</td>
<td>29.33***</td>
<td>9.14</td>
<td>32.65</td>
<td>26.17***</td>
</tr>
<tr>
<td>15</td>
<td>16.63</td>
<td>46.46</td>
<td>38.18***</td>
<td>16.69</td>
<td>44.71</td>
<td>35.58***</td>
</tr>
<tr>
<td>20</td>
<td>17.9</td>
<td>43.43</td>
<td>26.1***</td>
<td>14.52</td>
<td>35.13</td>
<td>18.41***</td>
</tr>
<tr>
<td>SP500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall, the findings within this section are consistent with those from section 4.1. In section 4.1, we had statistical significance at multiple horizons for real dividend growth and the real equity premium. Given this, it is not surprising to find both with good out-of-sample performance.\(^\text{20}\)

\(^\text{20}\)That is not to say that a highly statistically significant in-sample predictor will always have positive out-of-sample performance or vice verse.
5 Conclusion

The empirical results from this paper indicate that through the application of MIDAS regressions, the dividend price ratio is a predictable component of dividend growth, market returns or the equity premium. Asimakopoulos et al. (2017) found that the dividend price ratio growth is the predictable component of the dividend price ratio when MIDAS regressions were used. The findings of this paper are not necessarily in contrast to theirs.

This paper demonstrates that it matters how we aggregate the dividend price ratio when exploiting high frequency data. By using MIDAS regressions, I show that in many cases we can uncover predictability. In particular, we saw that weighting the dividends in each month differently uncovered statistical significance across large equity markets.

Not only could we uncover statistical significance across large equity markets, but we also saw that there are gains in utilizing high frequency data compared to annual data. The gains for nominal and real dividend growth in-sample were relatively consistent over time. Out-of-sample we saw that the MIDAS regression weighted dividend price ratio resulted in predictability gains when forecasting began in 1965. We also saw that the inclusion of additional predictors with the mixed frequency dividend price ratio resulted in consistent, statistically significant out-of-sample predictability of returns and equity premium.

Engsted and Pedersen (2009) found strong predictability of long horizon real dividend growth, but with the theoretically incorrect sign. Chen (2009) does not find dividend growth predictability at any horizon greater than one when the sample begins in 1926. Similarly, Ang (2011) only finds dividend growth predictability at the one year horizon. With a mixed frequency approach, I found strong predictability of long horizon real dividend growth with the theoretically correct sign. I also showed that there is some predictability with nominal dividend growth in US equity markets. Chen (2009) found a statistically significant relationship between dividend growth and the equity premium, but only for two horizons. Here, I find significance at multiple horizons. In addition, I

\footnote{Note, Ang (2011) does not test a horizon greater than 5. However, we find significance between horizon 1 and 5.}
find strong long horizon out-of-sample performance for both the real and nominal equity premium, real dividend growth and real market returns.
6 Appendix

6.1 F-Test Over Time

Figure 10: Expanding Window Dividend Growth Regressions (CRSP Data)
This figure plots the expanding window estimated p-value for the Wald Test. I test the null hypothesis that all 12 estimated monthly weights are equal to zero. The dependant variable is real and nominal dividend growth. The sample is 1928 to 2017.

Figure 11: Expanding Window Dividend Growth Regressions (SP500 Data)
This figure plots the expanding window estimated p-value for the Wald Test. I test the null hypothesis that all 12 estimated monthly weights are equal to zero. The dependant variable is real and nominal dividend growth. The sample is 1928 to 2017.
Figure 12: Expanding Window Returns Regressions (CRSP Data)
This figure plots the expanding window estimated p-value for the Wald Test. I test the null hypothesis that all 12 estimated monthly weights are equal to zero. The dependent variable is real and nominal returns. The sample is 1928 to 2017.

Figure 13: Expanding Window Returns Regressions (SP500 Data)
This figure plots the expanding window estimated p-value for the Wald Test. I test the null hypothesis that all 12 estimated monthly weights are equal to zero. The dependent variable is real and nominal returns. The sample is 1928 to 2017.
6.2 Out-of-Sample Performance Plots

Figure 14: Russell 2000 Index Out-of-Sample Performance
This figure depicts the Welch and Goyal (2007) annual OOS performance for both returns and the dividend growth ratio. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.

Figure 15: Canada SP/TSX Composite Index Out-of-Sample Performance
This figure depicts the Welch and Goyal (2007) annual OOS performance for both returns and the dividend growth ratio. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.
Figure 16: FTSE All Shares Index Out-of-Sample Performance
This figure depicts the Welch and Goyal (2007) annual OOS performance for both returns and the dividend growth ratio. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.

Figure 17: FTSE Euro First 300 Index Out-of-Sample Performance
This figure depicts the Welch and Goyal (2007) annual OOS performance for both returns and the dividend growth ratio. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.
6.3 Out-of-Sample Performance Plots for equations 12, 13, and 15

Figure 18: CRSP Index Out-of-Sample Performance equation 12
This figure depicts the Welch and Goyal (2007) annual OOS performance for both real and nominal equity premium. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.

Figure 19: CRSP Index Out-of-Sample Performance equation 13
This figure depicts the Welch and Goyal (2007) annual OOS performance for both real and nominal equity premium. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.
Figure 20: SP500 Index Out-of-Sample Performance equation 12
This figure depicts the Welch and Goyal (2007) annual OOS performance for both real and nominal equity premium. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.

Figure 21: SP500 Index Out-of-Sample Performance equation 13
This figure depicts the Welch and Goyal (2007) annual OOS performance for both real and nominal equity premium. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.
Figure 22: CRSP Index Out-of-Sample Performance equation 15
This figure depicts the Welch and Goyal (2007) annual OOS performance for both real and nominal equity premium. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.

Figure 23: SP500 Index Out-of-Sample Performance equation 15
This figure depicts the Welch and Goyal (2007) annual OOS performance for both real and nominal equity premium. As per Welch and Goyal (2007), the y-axis is the cumulative squared prediction errors of the null model minus the cumulative squared prediction error of the alternative model. An increase in the lines indicates better performance of the mixed frequency model, while a decrease in a line indicates better performance of the null.
References


