Time-disaggregated dividend-price ratio and dividend growth predictability in large equity markets

Panagiotis Asimakopoulos*
University of Piraeus

Stylianos Asimakopoulos†⊛
University of Stirling

Nikolaos Kourogenis‡
University of Piraeus

Emmanuel Tsiritakis§
University of Piraeus

Abstract

We consistently show that in large equity markets, the dividend-price ratio is significantly related with the growth of future dividends. In order to uncover this relationship, we use monthly dividends and a mixed data sampling technique which allows us to cope with within-year seasonality. Our approach avoids the use of overlapping observations, and at the same time reduces the implications of the impact of price volatility on the dividend-price ratio. An empirical analysis using market level data from U.S., U.K., Canada and Japan strongly supports the dividend growth predictability hypothesis, suggesting that time-aggregation of dividends eliminates significant information.

Keywords: dividend growth, dividend-price ratio, predictability, dividend smoothing, mixed data sampling

---

*Department of Banking and Financial Management, University of Piraeus, e-mail: pasimako@unipi.gr

†Division of Economics, Stirling Management School, University of Stirling, e-mail: Stylianos.Asimakopoulos@stir.ac.uk

‡Department of Banking and Financial Management, University of Piraeus, e-mail: nkourog@unipi.gr and nkourogenis@yahoo.com

§Department of Banking and Financial Management, University of Piraeus, e-mail: manolis@unipi.gr

JEL Classification: G12, G14, G15, G17

⊛: Corresponding author: stylianos.asimakopoulos@stir.ac.uk

1 Introduction

A main theoretical implication of the present value approach on firm valuation is that the dividend-price ratio should be significantly related to at least one of future returns and future dividend growth. Since the late '80s, the starting point for testing this hypothesis is mainly the work of Campbell and Shiller (1988a,b). Campbell and Shiller generalized the result of Williams (1938) and Gordon (1962) by obtaining an approximate relationship, which describes the log dividend-price ratio, dy, as the difference between expected discounted future log returns and expected discounted future dividend growth (plus a constant).[1]

More than twenty-five years since the publication of the Campbell-Shiller model, the finance literature seems to have reached a consensus about the existence of a significant linear relationship between dy and future returns. The examination, however, of the relationship between dy and future dividend growth has not yielded uniform conclusions (Ang and Bekaert, 2007, Koijen and van Nieuwerburgh, 2011, Maio and Santa-Clara, 2014).[2] In order to explain the lack of consistent results among different countries, a series of recent studies provided evidence that in large equity markets, dividend growth predictability by dy

[1] From another perspective, several theoretical models that examine the information embedded in dividend announcements predict that changes in dividend policy convey news about future cash flows (Bhattacharya, 1979, John and Williams, 1985, and Miller and Rock, 1985). Specifically, dividend increases (decreases) convey good (bad) news. According to Acharya and Labrecht (2011) executives in companies adjust dividend payments to market expectations at the company level in order to keep their shareholders satisfied and therefore keep their positions. Although asymmetric information theories view the dividend process from a different angle, they nevertheless suggest that we should expect to observe a significant negative relationship between dy and future dividend growth, if good firm prospects are embedded in the stock price while dividends are sticky or smoothed.

[2] Lettau and Ludvigson (2005) concluded that although US dividend growth rates are predictable by an estimated consumption-dividends-labor income ratio, they are not predictable by dy itself. Lettau and Nieuwerburgh (2008) concluded that the simple dividend-price ratio does not predict future dividend growth. They also used an adjusted, locally demeaned, dividend-price ratio, but even then they did not find any evidence of dividend growth predictability. Cochrane (2008) used the absence of evidence supporting a significant negative relationship between dy and subsequent dividend growth, in order to argue that this fact provides strong evidence against the hypothesis that returns are not forecastable. Van Binsbergen and Koijen (2010) found that while US market-wide dividends are predictable, the lagged price-dividend ratio does not contribute to a higher R^2. They finally achieved a higher R^2 by filtering out the information in the dividend-price ratio which is related to expected return variation. Engsted and Pedersen (2010) concluded that "*... real dividend growth in the US is unpredictable in the pre war period but significantly predictable in the 'wrong' direction in the post war period.*" Maio and Santa-Clara (2014) conclude that at the market level dividends are not predictable by dy, while this is not true for portfolios of small and value stocks.

is weak or even absent, arguing that this is a result of dividend smoothing policies applied mainly by large firms (Chen, 2009, for the US in the post WWII period, Rangvid *et al.*, 2014).[3]

All aforementioned results are based on annual dividends, mainly due to the strong seasonality issues that emerge when higher frequency dividends are used. Moreover, the studies which provided evidence on dividend growth predictability in large equity markets (mainly the U.S.), were unable to relate this predictability with dy. In other words, so far the dividend growth predictability relationship which stems from Campbell-Shiller's approximate identity seems not to be supported by the empirical evidence. In this paper, we argue that the use of time-aggregated (annual) dividends in dy washes out significant information concerning dividend growth predictability. We use time-disaggregated dividend-price ratios in order to reveal the link between dy and future dividend growth.[4] We deal with possible seasonality effects by applying the Mixed frequency Data Sampling (MiDaS) approach of Ghysels et al. (2004). MiDaS allows us to use annual data for the dependent variables (dividend growth) and data sampled at a higher than annual frequency for the variables on the right hand side of our regressions.[5] It also allows us to avoid the use of overlapping observations.

Because the current literature relates large market size with the absence of dividend growth predictability by dy, our empirical analysis focuses on four of the world's largest equity markets, namely, S&P 500 (U.S.), FTSE 100 (U.K.), SPTSX 60 (Canada) and Nikkei 225 (Japan). Our findings suggest that for every country in our sample, the time-disaggregated dividend-price ratio, which involves monthly dividends, is significantly related with the future

[3] Chen *et al.* (2012) showed that at the firm level, "... *even if dividends are supposed to be predictable without smoothing, dividend smoothing can bury this predictability in a finite sample.*"

[4] By the term 'time-disaggregated dividend-price ratio' we refer to any dividend-price ratio in which the value of the dividend corresponds to the aggregate dividends paid within a period of less than one year. In that sense, the term 'aggregation' corresponds to aggregation of information in the time domain, and not to summation of higher frequency variables.

[5] In a very recent paper, Golez (2014) uses data from derivatives on the S&P 500 at monthly frequency in order to extract information that predicts future returns and dividend growth for the period January 1994 - June 2011. Specifically, Golez combines the futures pricing (cost-of-carry) and put-call parity formulas under no-arbitrage, and extracts implied dividend yields (IDY) for S&P 500. Then, he defines the implied dividend growth as $idg = \ln(IDY) - dy$, where the log twelve-month trailing sum of dividends, d^{12}, is involved in dy, and shows that idg predicts the growth of d^{12} over a horizon of one, three, six and twelve months.

dividend growth. The results also identify a component of the time-disaggregated dividend-price ratio, which, in all cases, offers predictive power. We also repeat our analysis using quarterly dividends but then the predictability of dividend growth vanishes, implying that the effect of time-aggregation is significant even when higher-than-annual frequencies are used.[6] The existence of dividend growth predictability by the dividend-price ratio, especially for the U.S., is in contrast with what is suggested in recent studies (see Chen, 2009, and Rangvid *et al.*, 2014)[7]. It is worth noting, however, that in a recent paper, Kelly and Pruitt (2013) show that the use of cross-sectionally disaggregated (firm level) information in a latent factor system can significantly improve the predictability of dividend growth at the market level. On the other hand, their analysis concerning future cash flows is based on annual data and uses information from the cross section of book-to-market ratios due to the lack of dividend payments for a substantial fraction of U.S. firms in their sample.

The paper is organized as follows. Section 2 describes the main variables and the set of equations that are used as a starting point in our analysis. This section also outlines MiDaS. Section 3 presents how MiDaS is used in order to obtain the predictive regression for dividend growth. Section 4 presents the results of our empirical analysis. Section 5 concludes the paper.

2 Preliminaries

In this section we introduce the variables used throughout the paper and we briefly present the model of Campbell and Shiller (1988a,b). Then, we outline the Mixed Data Sampling approach.

[6]Monthly (quarterly) dividends of an index correspond to the aggregate dividends paid by all companies in the index during the period of one month (quarter) and not to annual dividends sampled at monthly (quarterly) frequency.

[7]Rangvid *et al.* (2014) suggest that dividend growth predictability via dividend yield for the US is accidental and the result does not extent to the other large equity markets.

2.1 Dividend Predictability

Let P_t and D_t denote the price of a stock (or the value of an index) at time t and the corresponding aggregate dividend that has been paid during the time interval $(t-1, t]$, respectively. Let also $p_t := \ln P_t$ and $d_t := \ln D_t$. The returns are defined by:

$$r_t = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right) .$$

The log dividend-price ratio is given by $dy_t = d_t - p_t$. The literature on dividend predictability has been motivated from the work of Campbell and Shiller (1988a,b). They showed that a good approximation of dy_t is given by:

$$dy_t \simeq c + E_t \sum_{i=1}^{\infty} \rho^{i-1} r_{t+i} - E_t \sum_{i=1}^{\infty} \rho^{i-1} \Delta d_{t+i} , \tag{1}$$

where E_t is the conditional expectation operator at time t. Equation (1) implies that dy_t should predict revisions on future returns and/or dividend growth.

Considering the approximate identity (1), a reasonable starting point for the identification of the main driving force of equity markets is to model the vector $[r_t, \Delta d_t, dy_t]'$ as a VAR(1) process, where the first two columns of the coefficients' matrix are zeros (see Cochrane, 2008 and 2011, and Chen *et al.*, 2012). Specifically, the vector autoregression can be expressed as

$$\Delta d_{t+1} = c_0 + c_1 dy_t + u_{t+1,d} \tag{2}$$

$$r_{t+1} = c_{0,r} + c_{1,r} dy_t + u_{t+1,r} \tag{3}$$

$$dy_{t+1} = c_{0,y} + c_{1,y} dy_t + u_{t+1,y} , \tag{4}$$

where a common approach is to use annual data for every variable in order to avoid seasonality issues. The approximation of Campbell and Shiller (1988a) determines the set of admissible joint null hypotheses on the coefficients of the VAR. For example, under the assumption that $c_{1,y} < 1$, we cannot assume a null hypothesis where $c_{1,r} = c_1 = 0$, which implies that at least one of r_{t+1} and Δd_{t+1} must be significantly related with dy_t (Cochrane, 2008). A consequence of this observation is that if the empirical evidence supports the null hypothesis $\{c_1 = 0\}$, then it also supports the rejection of the hypothesis $\{c_{1,r} = 0\}$. On the other hand, a rejection of the hypothesis $\{c_1 = 0\}$ is not informative about whether $c_{1,r} = 0$.

2.2 The Mixed Frequency Data Sampling Approach (MiDaS)

A useful tool for empirical analyses, when regressor and regressand are sampled at different frequencies, is the Mixed Data Sampling approach (MiDaS), introduced by Ghysels *et al.* (2004). MiDaS has been extensively applied in financial data for assessing volatility predictions and stock returns (i.e. Forsberg and Ghysels, 2006 and Ghysels *et al.,* 2006), as well as in forecasting macroeconomic variables using intra-annual data (i.e. Bai *et al.,* 2009, Kuzin *et al.,* 2011 and Clements and Galvao 2008, 2009) and more recently, in forecasting annual fiscal data using quarterly announcements (Asimakopoulos *et al.*, 2013). To the best of our knowledge, it is the first time that MiDaS is applied on a dividend growth predictability model.

Let us assume that the higher frequency data (monthly in our case study) and the low frequency data (annual data) are denoted by X_t^M and Y_t^A, respectively. The standard MiDaS regression is:

$$Y_{t+1}^A = \beta_0 + \beta_1 B(L^{1/m}; \boldsymbol{\theta}) X_t^M + \varepsilon_{t+1} \tag{5}$$

where $L^{1/m}$ is the higher frequency (monthly) lag operator (here $m = 12$), and $B(L^{1/m}; \boldsymbol{\theta}) = \sum_{j=0}^{K-1} \omega_j(\boldsymbol{\theta}) L^{k/m}$ is a polynomial of $L^{1/m}$ that also depends on a vector of parameters, θ, which determine the curvature of the weighting scheme.

The above expression determines the effect of the higher frequency explanatory variable on the lower frequency dependent variable. Ghysels *et al.* (2007) provide several weighting schemes. They show that the exponential Almon lag polynomial has the most flexible shape and therefore is assumed to be the most general weighting scheme and this is the main reason that we also incorporate that to our analysis. The exponential Almon lag polynomial is fully determined by two parameters θ_1 and θ_2, hence $\boldsymbol{\theta} = (\theta_1, \theta_2)'$. The corresponding weights, $\omega_j(\boldsymbol{\theta})$, are given by

$$\omega_j(\boldsymbol{\theta}) = \frac{\exp\{\theta_1 j + \theta_2 j^2\}}{\sum_{j=1}^{m} \exp\{\theta_1 j + \theta_2 j^2\}} . \tag{6}$$

The advantage of MiDaS when compared to alternative approaches, such as State Space and mixed frequency VAR models that use Kalman filter, is that it is parsimonious and less sensitive to specification errors due to the use of non-linear lag polynomials. In addition, MiDaS does not suffer from the parameter proliferation issue. This is important in our analysis because the time span of the data is not large enough for most of the countries in our sample. Concerning the small sample size issue, Ghysels *et al.* (2006) show that MiDaS performs better than State Space models that make use of Kalman filter, as the time span of data decreases.

Another significant advantage of MiDaS is that the weighting scheme is purely data driven and no prior assumption is necessary. Note that it is common in the literature to simply take the average of the higher frequency variables to transform them into low frequency, but the equal weighting assumption might lead to inefficient and, in some cases, to biased or inconsistent estimators (Andreou et al., 2010). By definition, MiDaS avoids this issue. On the other hand, when compared to a purely high frequency model, MiDaS avoids the seasonality issues that appear in higher frequencies due to the lower frequency sampling of the regressand and the flexibility of the weighting scheme.

3 MiDaS Predictive Regression for Dividend Growth

The literature on dividend growth predictability uses annual observations in order to avoid potential seasonality issues that appear in higher frequency dividend data (see, e.g. Rangvid *et al.*, 2014, among others). This approach, however, ignores potentially important information from higher frequency observations, which vanishes when aggregated over the periods that correspond to the lower frequency. In this paper, we propose a solution to this problem using mixed frequency data.

We consider that the lower frequency observations are annual and that the time variable, t, takes integer values and corresponds to the last day of the corresponding year. Specifically, p_t is the logarithm of the value of an equity index the last day of year t, and d_t is the logarithm of the aggregate dividend paid by all companies in this index during year t. By $t - j/12$, $j \in \{0, 1, \ldots, 11\}$, we denote the months within year t. For example, $P_{t-1/12}$ is the value of the index the last day of November of year t, $p_{t-1/12} = \log(P_{t-1/12})$, while $D^m_{t-1/12}$ is the aggregate dividend paid by all companies in the index during November of year t and $d^m_{t-1/12} = \log(D^m_{t-1/12})$. The corresponding monthly log dividend-price ratio is denoted by $dy^m_{t-1/12} := d^m_{t-1/12} - p_{t-1/12}$. Note that by $dy^m_{t-j/12}$, $j \in \{0, 1, \ldots, 11\}$, we denote the log dividend-price ratio of month $t - j/12$, where the dividends correspond to this month only, and not to the twelve-month period ending at the end of month $t - j/12$.

3.1 A first model with monthly dividend-price ratios

Our first task will be to introduce higher frequency (monthly) variables in (2). We will maintain the low frequency dividend growth on the left hand side of the equation in order to avoid the effects of high frequency seasonalities. Concerning the right hand side, the reasoning behind the choice of a mixed frequency method implies that we have to incorporate a higher frequency variable. Specifically, we will use the monthly dividend-price ratios, $dy^m_{t-j/12}$.

A 'naive' approach would be to replace dy_t in (2) by $\overline{dy^m_t} := \sum_{j=0}^{11} dy^m_{t-j/12}$ yielding the following equation

$$\Delta d_{t+1} = c_0 + c_1 \overline{dy^m_t} + u_{t+1} \ . \tag{7}$$

We observe that $\overline{dy^m_t}/12$ is the average monthly dividend-price ratio during year t. In other words, $\overline{dy_t^m}$ is the annualized average monthly dividend-price ratio for year t. When compared to dy_t, $\overline{dy_t^m}$ combines more synchronous information, because while the annual dividend in dy_t aggregates throughout a whole year, the price in dy_t corresponds to the end of year t. When compared to (2), the right-hand side of (7) is much less sensitive to end-of-year price volatility, while $\overline{dy_t^m}$ may be seen as a smoothed version of the dividend-price ratio.[8] The correlation of $\overline{dy_t^m}$ with dy_t, $corr(\overline{dy_t^m}, dy_t)$, is high. This should be expected because d_t is a transformation of the information contained in every $d_{t-j/12}^m$ within year t, while $p_{t-j/12}$ is strongly persistent. The high value of $corr(\overline{dy_t^m}, dy_t)$ is also supported by our dataset. Specifically, we find that the sample correlation between $\overline{dy_t^m}$ and dy_t is 0.93 for the U.S., 0.87 for the U.K., 0.93 for Canada and 0.85 for Japan.

Concerning the applicability of equation (7), we have to bear in mind that its purpose is mainly to show the direction we have to follow in order to introduce higher frequency variables in a model for the predictability of dividend growth without having to resort to other variables than monthly dividend-price ratios. Unfortunately, the use of $\overline{dy_t^m}$ on the right hand side of (7) has two main drawbacks. First, although $\overline{dy_t^m}$ involves monthly dividend-price ratios, it is a variable constructed at the annual frequency as a simple equally-weighted sum. This fact does not support the identification of different sensitivities between Δd_{t+1} and $dy_{t-j/12}^m$ for different values of j. An argument supporting the existence of different sensitivities at the aggregate, country level, is that only a relatively small number of companies of an index pay dividends at monthly frequency, while the value of an index depends on the prices of all its constituents. A second argument, concerns the fact that prices of the late months of the year represent expectations based on a richer information set than the one that corresponds to the early months of the same year. Moreover, it should be noted that $\overline{dy_t^m}$ is highly persistent. This is supported by the high correlation it has with dy_t, and also, by an initial inspection of the statistical properties of $\overline{dy_t^m}$ for the data involved in our analysis. Specifically, for each

[8] In fact, $\overline{dy_t^m}/12$ is a smoothed version of the monthly dividend-price ratio. However, the use of $\overline{dy_t^m}$ instead of $\overline{dy_t^m}/12$ in (7) and its subsequent variations does not affect the signs and p-values in the results of the corresponding regressions. Note that the dividends in $\overline{dy_t^m}$ do not correspond to ovelapping periods.

country in our sample, the values of the ordinary least squares estimator, $\widehat{\rho}$, of the coefficient of an AR(1) specification for $\overline{dy_t^m}$, $\overline{dy_t^m} = \rho \overline{dy_{t-1}^m} + \varepsilon_t$, are: 1.002 for the U.S., 0.994 for the U.K., 0.997 for Canada and 0.99 for Japan. Although based in a very small sample, these estimates provide an indication that $\overline{dy_t^m}$ is strongly persistent.[9] In case that the persistence of $\overline{dy_t^m}$ is of a unit root type, a nonzero c_1 in equation (7) would imply that Δd_{t+1} has a unit root too, contradicting the empirical findings that are reported in the relevant literature so far (see van Binsbergen and Koijen, 2010, among others). From another perspective, since the persistence of Δd_{t+1} is much weaker than that of a (near) unit root process, if $\overline{dy_t^m}$ has a unit root the estimated value of c_1 will converge to 0 as the sample size increases.[10]

In the next subsection we relax equation (7) in order to allow for deviations from the underlying assumptions of (1) and (7). The resulting model nests (7). In other words, we embed equation (7) in a more general framework, allowing for the data to "tell their story."

3.2 The Role of Time and MiDaS

In the previous subsection we identified two main drawbacks in equation (7). These drawbacks concern the selection of the explanatory variable and are (i) the persistence of $\overline{dy_t^m}$, and (ii) the fact that although $\overline{dy_t^m}$ consists of monthly dividend-price ratios, it is still a low frequency variable.

In order to deal with the first drawback, the near unit root behavior of $\overline{dy_t^m}$, we relax equation (7) by also incorporating the changes (first differences) of $\overline{dy_t^m}$ on the right hand side. Specifically, we allow for the sensitivity of Δd_{t+1} to recent information on the dividend-price ratio ($\Delta\overline{dy_t^m}$) to be different from the sensitivity of Δd_{t+1} to $\overline{dy_{t-1}^m}$. This leads to the following equation:

$$\Delta d_{t+1} = c_0 + c_1 \Delta \overline{dy_t^m} + c_2 \overline{dy_{t-1}^m} + u_{t+1} \ . \tag{8}$$

[9] Further examination of the persistence of $\overline{dy_t^m}$ is provided in the empirical section of the paper.

[10] This is an implication of the fact that the percentage of time a unit root process passes outside any fixed bounded interval tends to 1, while the time series on the left hand side of the regression is stable. The observation that a persistent right-hand side could lead to serious over-rejections, is not novel (see, for example, Nelson and Kim, 1993, and Campbell and Yogo, 2006).

We observe that equation (8) nests (7) as the special case where $c_1 = c_2$. In this sense, equation (8) is not incompatible with (1). However, equation (8) allows for the sensitivity of Δd_{t+1} to the strongly persistent term, $\overline{dy_{t-1}^m}$, to be determined separately from $\Delta\overline{dy_t^m}$ which might contain significant information for the future dividend growth. Therefore, if only c_1, and not c_2, appears to be significant we will have identified $\Delta\overline{dy_t^m}$ as the informative component of $\overline{dy_t^m}$.

Equation (8) seems to deal with the persistence issue of $\overline{dy_t^m}$. It is, however, reasonable for us to wonder if there is an explanation in terms of economic behavior or asset pricing that supports a relationship between the first differences of the dividend-price ratio and future dividend growth. We will not provide a general answer to this question, but we will focus in the case where prices are formed under the perspective that dividend smoothing takes place. This is the relevant case in our study, because dividend smoothing is a well documented characteristic of dividend policies in large equity markets after the WWII (see Chen, 2009, among others).

We start by rewriting the first difference of the dividend-price ratio as $\Delta dy_t := d_t - p_t - (d_{t-1} - p_{t-1}) = \Delta d_t - \Delta p_t$. At the end of period t, investors are already informed about the change in the dividend paid during period t. The price, p_t, of the corresponding asset is, however, sampled at the end of period t. Therefore, the information set used for the formation of p_t includes the change of dividends, Δd_t. Note that over the years, many theories considered dividends either as a signaling device to mitigate information asymmetry problems or as an efficient way to resolve agency problems.[11]

In case that an increase in Δd_t is considered as "good news" for future cash flows, p_t increases accordingly. Dividend smoothing implies that expectations on future cash flows are higher than the ones revealed by the increase of Δd_t . Consequently, in order to capture the effect of dividend smoothing on Δd_t investors set the price p_t at an even higher level, which leads to a smaller Δdy_t . On the other hand, future dividend growth has to be higher

[11] Allen and Michaely (2003), Frankfurter and Wood (2003), Baker (2009), and DeAngelo, DeAngelo, and Skinner (2009) provide excellent reviews of these theories and the related empirical facts.

in order to compensate for the smoothed change in the current dividend. In other words, under rational expectations, an increase in the expectations of future cash flows implies that the growth rate of future dividends has to increase with respect to current dividend growth, because the latter was suppressed due to dividend smoothing. Similarly, a specific decrease of Δd_t under dividend smoothing, signals lower expectations of future cash flows than in the case of no dividend smoothing. Consequently, p_t is set at a lower level than in the case of the same Δd_t under no dividend smoothing, leading to a higher Δdy_t. Again, the future change of dividend growth has to be negative in order to compensate for the smoothed change of the current dividend. The previous reasoning leads us to conjecture that when investors know that dividends are smoothed, a negative relationship exists between the change of the dividend-price ratio and future dividend growth.[12]

Equation (8) can be considered as an intermediate step towards the derivation of an even more flexible model that will be able to exploit any differences in the sensitivities of the variables that correspond to high frequency (monthly) information. In order to derive a model with this feature, we focus on the component of (8) that corresponds to recent information, $\Delta\overline{dy_t^m}$. We observe that

$$\Delta\overline{dy_t^m} = \sum_{j=0}^{11} \left[(d^m_{t-j/12} - p^m_{t-j/12}) - (d^m_{t-1-j/12} - p^m_{t-1-j/12}) \right] \tag{9}$$

where each summand of the right hand side of (9) corresponds to the annual growth of a monthly log dividend-price ratio, $gdy^m_{t,j} := (d^m_{t-j/12} - p^m_{t-j/12}) - (d^m_{t-1-j/12} - p^m_{t-1-j/12})$, $0 \le j \le 11$. According to (8), the sensitivity of Δd_{t+1} to each $gdy^m_{t,j}$, $0 \le j \le 11$, is the same. Because each term, $gdy^m_{t,j}$, corresponds to information available at a different point in time, it is reasonable for us to require a more flexible structure than $\Delta\overline{dy_t^m}$, that will allow

[12] It would be tempting to try using this reasoning in order to obtain a similar relationship for the predictability of returns. However, an increase of the smoothed current dividend growth does not imply higher future returns, given that investors have already incorporated this smoothing in the formation of the current price.

for different sensitivities.[13] A straightforward approach would be to estimate the following relaxed version of (7):

$$\Delta d_{t+1} = c_0 + \sum_{j=0}^{11} c_{1,j} gdy_{t,j}^m + c_2 \overline{dy_{t-1}^m} + u_{t+1} \ . \tag{10}$$

Unfortunately, equation (10) has fourteen degrees of freedom! This fact alone would make any estimation result unreliable, given that the availability of monthly dividends at the market level covers at most two and a half decades, which corresponds to less than twenty-five annual observations. We deal with the issue of parameter proliferation by using the approach of MiDaS.

As described in the previous section, MiDaS imposes a structure on $c_{1,j}$s. Specifically, according to MiDaS, $c_{1,j} = c_1 w_j$, $0 \leq j \leq 11$, where the w_js are determined by the parameters θ_1 and θ_2 of the Almon lag polynomial. The corresponding MiDaS equation becomes

$$\Delta d_{t+1} = c_0 + \sum_{j=0}^{11} c_1 w_j gdy_{t,j}^m + c_2 \overline{dy_{t-1}^m} + u_{t+1} = c_0 + c_1 \Delta w \overline{dy_t^m} + c_2 \overline{dy_{t-1}^m} + u_{t+1} \ , \tag{11}$$

where

$$\Delta w \overline{dy_t^m} = \sum_{j=0}^{11} w_j gdy_{t,j}^m \ . \tag{12}$$

Equation (11) can be considered as a version of (2) which is robustified with respect to end-of-year price volatility, and which allows for separate treatment of information of different lags, annual and monthly. It also nests equations (7) and (8) as special cases. Therefore, (11) is not incompatible with the Campbell-Shiller model, while, at the same time it allows for deviations from the strict structure of equation (7).[14] Moreover, the term with the higher frequency observations, $\Delta w \overline{dy_t^m}$, has low persistence, because it involves annual changes

[13] As already mentioned, only a small number of companies pay dividends every month. This fact causes time-variation to the number of companies that pay dividends each month, and consequently implies that the degree at which monthly dividends affect market expectations on future dividend growth may be also varying; probably following approximately the same pattern every year.

[14] These deviations may be a result of a violation of the underlying assumptions of equation (7).

of monthly dividends at monthly frequency. In other words, although the changes of the monthly dividend-price ratios in $\Delta w\overline{dy_t^m}$ correspond to overlapping periods, the monthly dividends involved in this term correspond to periods of one month. Each of these periods is considered only once in the derivation of $\Delta w\overline{dy_t^m}$.

Note that in MiDaS regressions as (11), the term $\Delta w\overline{dy_t^m}$ can not be considered as a simple weighted average where the weights are known *a priori*, because the high frequency coefficients w_j, $0 \leq j \leq 11$, are estimated along with c_0, c_1, and c_2. In other words, equation (11) is a compact way to describe model (10) under the restriction $c_{1,j} = c_1 w_j$, $0 \leq j \leq 11$, where the w_js are determined by the estimated θ_1 and θ_2.

As far as the number of degrees of freedom of equation (11) is concerned, five coefficients (c_0, c_1 and c_2 along with θ_1 and θ_2) are estimated when equation (11) is considered without any restriction. On the other hand, under the joint hypothesis that Δd_{t+1} does not have a unit root while $\overline{dy_{t-1}^m}$ has a unit root, we can estimate (11) under the restriction $c_2 = 0$. In this case, only four coefficients are estimated.[15]

We finally note that c_1 is not affected by a possible strong correlation between the high frequency components of $\Delta w\overline{dy_t^m}$. A significant c_1 in equation (11) is interpreted as evidence that the annual changes of the monthly dividend-price ratios contain significant information with respect to future dividend growth.

4 Empirical Results on Dividend Growth Predictability

In this section we present the results concerning the predictability of dividend growth using index data from U.S., U.K., Canada and Japan. Specifically, we consider the following indices: S&P 500 (U.S.), FTSE 100 (U.K.), SPTSX 60 (Canada) and Nikkei 225 (Japan). The

[15] As we will see in the next section, the results concerning the significance and the sign of c_1 under the assumption $c_2 = 0$ are almost always in agreement with the corresponding results from the estimation of the unrestricted equation (11).

aggregate monthly dividend paid by all companies in the index is reported by Bloomberg. The available information on monthly dividend starts at 1988 for U.S., and 1994 for U.K., Canada and Japan.[16] The end point of the analysis for every country is the end of 2012. The sample size for each country depends on the availability of the corresponding monthly dividends.[17] An approach to obtain monthly dividends prior to this date is to interpolate quarterly dividends. For the U.S. in particular, monthly dividends via interpolation are provided in Robert Shiller's web page (www.econ.yale.edu/∼shiller/data.htm). Note, however, that this approach does not recover the information in the variation of higher frequency dividends, and cancels the benefit of using a mixed frequency data sampling approach. The importance of monthly information becomes apparent in the following subsections.

4.1 On the time series properties of $\overline{dy_t^m}$

Before we proceed to the main results of this section, we will look for indications of nonstationarity for the time series $\{\overline{dy_t^m}\}$. Table 1 presents the results of a set of unit root tests including an intercept for $\{\overline{dy_t^m}\}$ for each country in our sample. We observe that the results seem to support the null of a unit root in $\{\overline{dy_t^m}\}$ for the U.S., Canada and Japan. In addition, the KPSS LM-statistic has a p-value of less than 10% under the null of stationarity for Canada and Japan. However, the KPSS statistic does not exceed the 10% critical value for the U.S. Concerning U.K., the combined results seem not to support the unit root hypothesis, while the p-value of the KPSS stationarity test is well above 10%. Taking into consideration

[16] For each index, monthly dividends are calculated by aggregating all dividends paid withing the corresponding month. Bloomberg aggregates daily information since 1988 for S&P 500, and since 1994 for FTSE 100, SPTSX 60 and Nikkei 225. Concerning S&P 500, Standard & Poor's began reporting daily dividends in 1988.

[17] Japan is the only country in our dataset, for which the problem of zero monthly dividend occurred. Specifically, during the nineteen years of monthly dividend data for Japan, April dividends were always zero, while nonzero October dividends occurred only twice. In view of this fact, our analysis actually omits these two months for Japan. On the other hand, concerning the same country, we had to deal with an additional number of thirteen zero monthly dividends. This number is relatively small when compared with a dataset of 190 observations (after the exclusion of April and October zero dividends). In order to avoid the issue of applying the logarithmic function to zero, we treat the zero dividends as not available data when MiDaS is used.

the results for all countries, we conclude that the univariate tests do not provide a clear indication about whether the time series $\left\{\overline{dy_t^m}\right\}$ is stationary or not, in general.

[Table 1 about here]

Another approach to the stationarity issue of $\left\{\overline{dy_t^m}\right\}$ is that of performing unit root tests for pooled data. Table 2 reports the results of the tests of Levin, Lin and Chu (2002) for a common unit root in the four series, and of Im, Pesaran and Wu (2003), as well as the Fisher tests based on the Augmented Dickey-Fuller and Phillips-Perron statistics, for individual unit roots. All results seem to support the unit root hypothesis.

[Table 2 about here]

We have to bear in mind that the small number of observations for $\left\{\overline{dy_t^m}\right\}$ may render the tests used in Table 1 and Table 2 unable to reject the null hypothesis of a unit root. Consequently, in the subsequent analysis we have to estimate equation (11) both under no restrictions and under the restriction $c_2 = 0$ (which corresponds to the acceptance of the unit root hypothesis). As we will see, however, the results under the two alternatives are almost always similar, and the acceptance of the theoretical arguments supporting the stationarity of $\left\{\overline{dy_t^m}\right\}$ (hence, focusing on the unrestricted specification) does not have any implications in the main findings of our analysis.[18]

4.2 Dividend growth predictability

The results of our empirical study concerning dividend growth predictability are presented in Tables 3 to 8. Each one of Tables 3 to 7 has two panels (Panel A and Panel B). Panels A concern regressions with the annual dividend growth, Δd_t, being the dependent variable.

[18] Given the strong persistence of $\left\{\overline{dy_t^m}\right\}$, and the mixed results of Tables 1 and 2, we are obliged to examine both specifications ($c_2 \in \mathbb{R}$ and $c_2 = 0$) for reasons of statistical consistency.

Panels B present results for the same regressions, but with the left hand side variable, Δd_t, being replaced by the average annual growth of monthly dividends, given by

$$\overline{\Delta^a d_t^m} := \frac{1}{12} \sum_{j=0}^{11} \left(d_{t-j/12}^m - d_{t-1-j/12}^m \right) \ .$$

Finally, note that none of the regressions in our empirical analysis will use overlapping data.

4.2.1 One-year-ahead predictability of dividend growth

The results on one-year-ahead dividend growth predictability are presented in Table 3. The first block of columns of Table 3 corresponds to equation (2). The second block of columns corresponds to an application of MiDaS without the decomposition of the dividend-price ratio. In other words, although mixed frequency data are used in the second regression, the high frequency data correspond to monthly dividend-price ratios, $dy_{t-j/12}^m$, $0 \leq j \leq 11$. The corresponding equations are given by:

$$\Delta d_{t+1} = c_0' + c_1' w dy_t^m + u_{t+1} \ , \tag{13}$$

where

$$w dy_t^m = \sum_{j=0}^{11} w_j' dy_{t-j/12}^m \ .$$

The third block of columns corresponds to equation (11). As already mentioned, Panel 3.B presents the results of the same regressions with Δd_t replaced by $\overline{\Delta^a d_t^m}$.

[Table 3 about here]

The results of the first (annual frequency) regression of Panel 3.A can be compared with the corresponding results in Rangvid *et al.* (2014). Only for Canada the dividend-price ratio seems to be a significant component of dividend growth variability, with a p-value equal to

9%. On the other hand, the dividend-price ratio does not significantly affect the future dividend growth for both U.S. and U.K. This result is comparable and towards the same direction with previous studies, which are mostly based on U.S. data (i.e. Chen, 2009). The corresponding results in Panel 3.B indicate no significant relationship between dy and $\overline{\Delta^a d_t^m}$ for all countries in our sample.

When monthly dividend-price ratios, as described in equation (13), are used instead of annual dividend-price ratios, the results are not uniform (second block of columns in Table 3). Specifically, when Δd_t is the dependent variable (Panel 3.A), only for the cases of Canada and Japan, we observe statistically significant dividend-price ratios with a p-value of 4% and 8% respectively. On the other hand, when Δd_t is replaced by $\overline{\Delta^a d_t^m}$, the coefficients that correspond to U.K., Canada and Japan are significant, with p-values between 1% and 2%, while no significant relationship is identified between $\overline{\Delta^a d_t^m}$ and wdy_t^m, concerning S&P500. The results indicate that a simple approach that directly applies MiDaS to higher frequency dividend-price ratios, does not consistently reveal signs of dividend growth predictability.

When equation (11) is estimated, the situation changes radically. The decomposition of the dividend-price ratio and the application of MiDaS to its growth component yields significant results for all countries. Specifically, $\Delta w\overline{dy_t^m}$ is always significant and the sign of c_1 is always negative, in agreement with the theoretically expected sign. Table 3 also reports the results of a test on the hypothesis that $c_1 = c_2$ for all countries. The only country for which this hypothesis cannot be rejected is Canada when Δd_t is the dependent variable, while the hypothesis cannot be rejected for Canada and U.S., when $\overline{\Delta^a d_t^m}$ is the dependent variable. Finally, Table 3 presents the results of the estimation of equation (11) under the restriction $c_2 = 0$. Again, for all countries, the coefficient of the growth of the smoothed dividend-price ratio is statistically significant. In the supplemental material of the paper, we present the figures of the estimated and realized dividend growth through equation (11) for the four markets under consideration.

Figure 1 illustrates the weighting schemes of the four markets as they result from the

MiDaS estimation of equation (11). It reveals that the dividends paid during the late months of the year, play a much more significant role in the prediction of the annual dividend growth than the dividends paid during the early months of the year. This fact provides and indication of the importance of allowing different weights, and, therefore, of the MiDaS approach.

[Figure 1 about here]

Concerning equation (11), it would be reasonable to wonder whether the application of an additional MiDaS scheme to its second term would yield different results. To answer this question, we would have to estimate the following equation:

$$\Delta d_{t+1} = c_0 + c_1 \Delta w\overline{dy_t^m} + c_2 wdy_{t-1}^m + u_{t+1} \ . \tag{14}$$

The results of the second block of table 3 provide a first indication about whether c_2 in equation (14) is statistically significant. Specifically, when wdy_t^m is included as the only regressor of Δd_{t+1} there is no evidence of a significant relationship between these two variables. Therefore, it seems unlikely that the lagged value of wdy_t^m is significantly related to Δd_{t+1}, given that another regressor is also included in equation (14). The results of the estimation of equation (14) verify this conjecture. Specifically, $\Delta w\overline{dy_t^m}$ remains statistically significant while the same does not hold for wdy_{t-1}^m. For economy of space, the corresponding results, are included in the supplemental material of the article.

Summarizing the results of Table 3, we conclude that the decomposition of the (smoothed) dividend-price ratio revealed a component ($\Delta w\overline{dy_t^m}$) that always contains predictive information. On the other hand, the remaining component ($\overline{dy_{t-1}^m}$) of the dividend-price ratio is not always significant. For each country, the sign of the relationship between future dividend growth and the corresponding significant component of the dividend-price ratio is always negative, being in agreement with the theoretically predicted sign. Concerning the U.S., in particular, it is worth noting that our conclusions seem to be in contrast with what has been suggested for the post WWII period in the literature so far (see Chen, 2009, and Rangvid *et al.*, 2014).[19]

4.2.2 The importance of high frequency data

Table 3 provided evidence that the involvement of dividends at a higher than annual frequency, changes the picture concerning the predictability hypothesis of dividend growth. Consequently, it is reasonable to ask whether the selection of monthly frequency is necessary. Table 4 presents the results of a MiDaS estimation of equation (11) with the only difference that quarterly data are used (in other words, we have only four subperiods each year). It is directly observed that when Δd_t is the dependent variable, only for the U.S. a significant relationship between Δd_{t+1} and $\Delta w\overline{dy_t^m}$ is still identified (with a p-value of 6%), while for Japan, the only significant relationship is between Δd_{t+1} and $\overline{dy_{t-1}^m}$. As far as $\overline{\Delta^a d_t^m}$ is concerned (Panel 4.B), the only significant c_1 corresponds to U.K. data. Under the restriction $c_2 = 0$, c_1 is not statistically significant for all countries in panel 4.A, while it remains significant for U.K. in Panel 4.B. The results of Table 4 reveal the importance of the choice of the highest possible (monthly) frequency in order to obtain globally uniform results. They also reveal how quickly time aggregation destroys useful information concerning the relationship between future dividend growth and $\Delta w\overline{dy_t^m}$.

[Table 4 about here]

4.2.3 Predictability of dividend growth at longer horizons

Let us now examine the relationship between the dividend-price ratio and the future dividend growth at longer horizons. Table 5 presents the results of the estimation of equation $z_{t+i} =$

[19] Any predictability of future dividend growth shown in previous studies, did not stem from the dividend-price ratio.

$c_0 + c_1 \Delta w\overline{dy_t^m} + c_2\overline{dy_{t-1}^m} + u_{t+i}$, with i =2, 3 and 4, where $z_t = \Delta d_t$ (panel 5.A) and $z_t = \overline{\Delta^a d_t^m}$ (panel 5.B). We observe that the relationship between Δd_{t+i} and $\Delta w\overline{dy_t^m}$ has always negative sign and is not significant only for Japan when i =4. Under the restriction $c_2 = 0$, the p-value of c_1 is higher than 10% only for the U.K., for i =3, and for Japan, for i =4. When $z_t = \overline{\Delta^a d_t^m}$, however, the results are not so uniform. Specifically, while c_1 is always negative, the corresponding p-values are considerably larger than 10% for S&P500. On the other hand, the p-values of c_1 for the rest three countries are less than 5% most of the times, and do not exceed the levels of 10%, when c_2 is also estimated, and 11%, under the restriction $c_2 = 0$, respectively.

[Table 5 about here]

Tables 3 and 5 provide evidence of a significant relationship between future dividend growth and the monthly dividend-price ratios. Although the number of monthly dividends involved in the estimation of equation (11) is more than 150 for all countries in our sample, it corresponds to only 25 years for U.S. and 19 years for U.K., Canada and Japan. This implies that an attempt to evaluate the forecasting performance of model (11) would be subject to small-sample effects, because both Δd_t and $\overline{\Delta^a d_t^m}$ are sampled at annual frequency. On the other hand, it still remains interesting to see the in-sample performance of a model which is based only on market-level monthly dividend-price ratios. Table 6 presents the in-sample adjusted R^2 of equation $z_{t+i} = c_0 + c_1 \Delta w\overline{dy_t^m} + c_2\overline{dy_{t-1}^m} + u_{t+i}$, with i =1, 2, 3 and 4, where $z_t = \Delta d_t$ (panel 6.A) and $z_t = \overline{\Delta^a d_t^m}$ (panel 6.B).

[Table 6 about here]

Concerning the dividend growth, Δd_t, we observe that the values of the in-sample adjusted R^2 for any horizon of up to four years, are at least 16%, 21%, 22% and 20% for U.S., U.K., Canada and Japan, respectively, when the unrestricted model is estimated. The corresponding values when equation (11) is estimated under the restriction $c_1 = c_2$ are 15%, 21%, 16% and 17%. When the sum of annual changes of the monthly dividend-price ratio, $\overline{\Delta^a d_t^m}$, is the dependent variable, the values of the in-sample adjusted R^2 for any horizon of up to four years, are at least 10%, 28%, 24% and 20% for U.S., U.K., Canada and Japan, respectively, when the unrestricted model is estimated. The corresponding values under the restriction $c_1 = c_2$ are 10%, 24%, 18% and 24%. It is worth pointing out that in the case of FTSE100, the adjusted R^2 of the two and four years ahead MiDaS predictive regressions for $\overline{\Delta^a d_t^m}$ reaches 55%.[20]

4.2.4 The added value of MiDaS

The results of our analysis are based on the decomposition of the smoothed dividend-price ratio, $\overline{dy_t^m}$ and the application of MiDaS on its growth component, $\Delta w\overline{dy_t^m}$. We have also shown that the application of MiDaS alone does not suffice to reveal the link between the information contained in monthly dividend-price ratios and future dividend growth (Table 3). In order to provide further support to our approach, we show that the decomposition of $\overline{dy_t^m}$ alone is also unable to reveal this link. Table 7 presents the results of the estimation of equations $z_{t+1} = c_0 + c_1 \Delta dy_t + c_2 dy_{t-1} + u_{t+1}$ and $z_{t+1} = c_0 + c_1 \Delta\overline{dy_t^m} + c_2 \overline{dy_{t-1}^m} + u_{t+1}$, where $z_t = \Delta d_t$ (Panel 7.A) and $z_t = \overline{\Delta^a d_t^m}$ (Panel 7.B). It can be easily observed that in most of cases the p-values of the estimated coefficients are significantly larger than 10%, while only for Canada, and only for c_2, the corresponding p-values are smaller than 10% in both regressions.

[Table 7 about here]

[20] Recall that MiDaS regression estimates only five parameters for the unrestricted model. Three of them (c_0, $c_1$and c_2) correspond to the low (annual) frequency part of the model, while the two coefficients of the exponential Almon lag polynomial, θ_1 and θ_2, determine the weights of the variable which is sampled at the higher (monthly) frequency.

4.2.5 Short term predictions

Let d_t^{si} be the log dividends paid within the i-th semester, $i \in \{1, 2\}$, and d_t^{qi} be the log dividends paid within the i-th quarter, $i \in \{1, 2, 3, 4\}$, of year t. We examine whether a significant relationship exists between the monthly dividend-price ratios of year t and the growth of semi-annual and quarterly dividends, $d_{t+1}^{s1} - d_t^{s2}$ and $d_{t+1}^{q1} - d_t^{q4}$, respectively. To this end, we estimate the following MiDaS regressions:

$$d_{t+1}^{s1} - d_t^{s2} = c_0 + c_1 \Delta w\overline{dy_t^m} + c_2\overline{dy_{t-1}^m} + u_{t+1} \tag{15}$$

and

$$d_{t+1}^{q1} - d_t^{q4} = c_0 + c_1 \Delta w\overline{dy_t^m} + c_2\overline{dy_{t-1}^m} + u_{t+1}\ . \tag{16}$$

We have to note that although a seasonal component exists between the dividends paid in sequential semesters or quarters (and their corresponding changes), equations (15) and (16) avoid the implications of this type of seasonality because they avoid the interchange between the two semesters or the four quarters, respectively. Specifically, equation (15) (equation (16)) focuses only on the change of the dividend between the last semester (quarter) of year t and the first semester (quarter) of year $t + 1$. However, because of the existing seasonality, it seems reasonable to wonder about the robustness of the model, which is derived under a set of assumptions that does not include any seasonality. Table 8 presents the results from the estimation of equations (15) and (16). It also reports the corresponding results when the restriction $c_2 = 0$ is imposed. The results of Table 8 seem to support the robustness of the model.

[Table 8 about here]

Specifically, we observe that in all cases the sign of c_1 is negative. Concerning the predictability of the six-month dividend growth, $d_{t+1}^{s1} - d_t^{s2}$, the estimated c_1 has p-values smaller than 5%, as far as U.S., U.K. and Japan are concerned, while the corresponding p-values for

Canada are approximately 10%. Concerning the predictability of the three-month dividend growth, $d_{t+1}^{q1} - d_t^{q4}$, all p-values of the estimated c_1 are smaller than 10%. In particular, when the unrestricted model is estimated, the corresponding p-values for U.K., Canada and Japan are smaller than 5%.

The conclusions of our empirical analysis can be summarized as follows: The application of MiDaS weights to the annual growth of the smoothed dividend-price ratio, revealed a negative and significant relationship between this component of the dividend-price ratio and the future dividend growth. This relationship remains negative and significant at longer horizons for all countries in our sample. When a slightly lower frequency (quarterly) on dividend data is used, the relationship in general vanishes. It is worth noting that only the combination of the decomposition of the smoothed dividend-price ratio with MiDaS is able to reveal a significant relationship for all markets under consideration.

5 Conclusions

In this paper we provided evidence that a significant relationship exists between the dividend-price ratio and the future dividend growth in large equity markets. In order to uncover this relationship we used higher frequency (monthly) data. The analysis focused on the main equity indices of U.S., U.K., Canada and Japan. Our motivation stemmed from the fact that the relevant literature was unable to identify a significant relationship between the dividend-price ratio and the future dividend growth.

Using a mixed data sampling approach (MiDaS) in order to deal with high frequency seasonality issues, and smoothing out the effects of price volatility on the dividend-price ratio, we found that for every country in our sample the smoothed dividend-price ratio contains significant information on the growth of future dividends. We identified a component of the smoothed dividend-price ratio (namely, its annual growth) that is always significantly related with the future dividend growth. The coefficient of this relationship is negative for all countries in our sample, as theoretically expected. We also provided evidence that the predictability of dividend growth emerges only when both MiDaS and the decomposition of the smoothed dividend-price ratio are applied. The weights of the estimated weighting schemes reveal that recent monthly dividends are more significant than the ones paid during the first months of the year.

When we applied exactly the same approach using data of a relatively lower frequency, we did not identify any significant relationship between the dividend-price ratio and future dividend growth for most of the countries in our sample. This result supports the view that when time-aggregated dividends are used significant information is ignored. The effect of time aggregation is quite direct, since it appears even when quarterly dividends are used.

References

[1] Acharya V, Lambrecht B. 2011. A theory of income smoothing when insiders know more than outsiders. *NBER Working Paper* 17696.

[2] Allen F, Michaely R. 2003. *Payout Policy*. In North-Holland Handbook of Economics (eds Constantinides G, Harris M, and Stulz R). North-Holland.

[3] Andreou E, Ghysels E, Kourtellos A. 2010. Regression models with mixed sampling frequencies. *Journal of Econometrics* **158**: 246-261.

[4] Ang A, Bekaert G. 2007. Stock return predictability: Is it there? *Review of Financial Studies* **20**: 651-707.

[5] Asimakopoulos S, Paredes J, Warmedinger T. 2013. Forecasting fiscal policy variables using mixed frequency data. *ECB Working Paper* 1550.

[6] Bai J, Ghysels E, Wright J. 2009. State space models and midas regressions. Working paper, NY Fed, UNC and John Hopkins.

[7] Baker H.K. 2009. *Dividends and Dividend Policy.* (ed. Kolb) in Series in Finance. Wiley.

[8] Bhattacharya S. 1979. Imperfect information, dividend policy, and "the bird in the hand" fallacy . *The Bell Journal of Economics* **10**: 259–70.

[9] van Binsbergen J H, Koijen R S J. 2010. Predictive Regressions: A Present-Value Approach. *Journal of Finance* **65**: 1439-1471.

[10] Campbell J, Shiller R. 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* **1**: 195-228.

[11] Campbell J, Shiller R. 1988b. Stock prices, earnings and expected dividends. *Journal of Finance* **43**: 195–227

[12] Campbell J, Thompson S B. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* **21:** 1509–1531.

[13] Campbell J, Yogo M. 2006. Efficient Tests of Stock Return Predictability. *Journal of Financial Economics* **81**: 27-36.

[14] Chen L. 2009. On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics* **92**: 128-151.

[15] Chen L, Zhao X. 2009. Return Decomposition. *Review of Financial Studies* **22:** 5213–5249.

[16] Chen L, Da Z, Priestley R. 2012. Dividend smoothing and predictability. *Management Science* **58***:*1-20.

[17] Clements P M, Galvao A B. 2008. Macroeconomics forecasting with mixed frequency data: Forecasting US output growth. *Journal of Business and Economic Statistics* **26:** 546-554.

[18] Clements P M, Galvao A B. 2009. Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics* **24**: 1187-1206.

[19] Cochrane J H. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial studies* **21**: 1533-1575.

[20] Cochrane J H. 2011. Discount rates. *Journal of Finance* **66**: 1047-1108.

[21] DeAngelo H, DeAngelo L, Skinner D.J. 2009. Corporate Payout Policy. *Foundations and Trends in Finance* **3**: 95-287.

[22] Engsted T, Pedersen T Q. 2010. The dividend-price ratio does predict dividend growth: International evidence. *Journal of Empirical Finance* **17:** 585-605.

[23] Fama E, French K. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* **22**: 3–25.

[24] Forsberg L, Ghysels E. 2006. Why do absolute returns predict volatility so well? *Journal of Financial Econometrics* **6**: 31-67.

[25] Frankfurter G, Wood B.G. 2003. *Dividend policy: Theory and practice*. Academic Press, San Diego.

[26] Ghysels E, Santa-Clara P, Valkanov R. 2004. The MIDAS Touch: Mixed Data Sampling Regression Models. CIRANO Working Papers 2004s-20, CIRANO.

[27] Ghysels E, Santa-Clara P, Valkanov R. 2006. Predicting volatility: Getting the most of return data sampled at different frequencies. *Journal of Econometrics* **131**: 59-95.

[28] Ghysels E, Sinko A, Valkanov R. 2007. MIDAS Regressions: Further Results and New Directions. *Econometric Reviews* **26**: 53-90.

[29] Golez B. 2014. Expected Returns and Dividend Growth Rates Implied by Derivative Markets *Review of Financial Studies*, forthcoming. doi:10.1093/rfs/hht131.

[30] Gordon M J. 1962. The Investment, Financing and Valuation of a Corporation, Irwin, Homewood, Ill.

[31] Goyal A, Welch I. 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Review of Financial Studies* **21**: 1455–1508.

[32] Im K.S, Pesaran M.H., Shin Y.2003. Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics* **115**:53-74.

[33] John K, Williams J. 1985. Dividends, dilution, and taxes: A signaling equilibrium. *Journal of Finance* **40**: 1053–1070.

[34] Kelly, B, Pruitt, S. 2013. Market Expectations in the Cross-Section of Present Values. *Journal of Finance* **68**: 1721–1756.

[35] Koijen R, van Nieuwerburgh S. 2011. Predictability of Returns and Cash Flows. *Annual Review of Financial Economics* **3**: 467–491.

[36] Kuzin V, Marcellino M, Schumacher C. 2011. MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting* **27**: 529-542.

[37] Lettau M, Ludvigson S C. 2001. Consumption, aggregate wealth and expected stock returns. *Journal of Finance* **56**: 815–849.

[38] Lettau M, Ludvigson S C. 2005. Expected returns and expected dividend growth. *Journal of Financial Economics* **76**: 583–626.

[39] Lettau M, van Nieuwerburgh S. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* **21**: 1607–1652.

[40] Levin A, Lin C.F, Chu C.J. 2002. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics* **108**:1–24.

[41] Lewellen J. 2004. Predicting returns with financial ratios. *Journal of Financial Economics* **74**: 209–235.

[42] Menzly L, Santos T, Veronesi P. 2004. Understanding Predictability. *Journal of Political Economy* **112:** 1-47.

[43] Maio P, Santa-Clara P. 2014. Dividend Yields, Dividend Growth, and Return Predictability in the Cross-Section of Stocks. *Journal of Financial and Quantitative Analysis*, forthcoming.

[44] Miller M H, Rock K. 1985. Dividend policy under asymmetric information. *Journal of Finance* **40**: 1031–1051.

[45] Nelson C R, Kim M J. 1993. Predictable Stock Returns: The Role of Small Sample Bias. *Journal of Finance* **48**: 641–661.

[46] Pastor L, Sinha M, Swaminathan B. 2008. Estimating the intertemporal risk-return tradeoff using the implied cost of capital. *Journal of Finance* **63**: 2859–2897.

[47] Rangvid J, Schmeling M, Schrimpf A. 2014. Dividend predictability around the world. *Journal of Financial and Quantitative Analysis*, forthcoming.

[48] Williams, J. B. 1938. *The Theory of Investment Value.* Harvard University Press, Cambridge, Mass.1938.

Appendix

Table 1: Values of univariate unit root test statistics for $\overline{dy_t^m}$

Test	Null hypothesis	U.S.	U.K.	Canada	Japan
Augmented Dickey-Fuller	unit root	-1.68	-2.89	-0.79	-1.38
t-statistic		(0.43)	(0.07)	(0.80)	(0.57)
Elliott-Rothenberg-Stock DF-GLS	unit root	-1.45	-2.46	-0.89	-0.91
t-statistic		(>0.10)	(<0.05)	(>0.10)	(>0.10)
Phillips-Perron	unit root	-1.54	-2.80	-0.79	-1.33
adj. t-statistic		(0.50)	(0.08)	(0.80)	(0.59)
Kwiatkowski-Phillips-Schmidt-Shin	stationarity	0.30	0.25	0.38	0.57
LM-statistic		(>0.10)	(>0.10)	(<0.10)	(<0.05)
Elliott-Rothenberg-Stock	unit root	7.19	3.62	9.85	33.61
P-statistic		(>0.10)	($\simeq$0.10)	(>0.10)	(>0.10)
Ng-Perron	unit root	-1.655	-1.66	-0.86	-0.05
MZ_t		($\simeq$0.10)	($\simeq$0.10)	(>0.10)	(>>0.10)

Table 1 presents the results of univariate unit root tests for the series $\{\overline{dy_t^m}\}$ with an intercept (p-values in parentheses).

Table 2: Values of pool unit root test statistics for $\overline{dy_t^m}$

Test	Null hypothesis	
Levin, Lin & Chu t*	common	-0.99
	unit root	(0.16)
Im, Pesaran and Shin	individual	-0.36
W-stat	unit root	(0.36)
ADF-Fisher	individual	8.66
Chi-square	unit root	(0.37)
PP-Fisher	individual	7.98
Chi-square	unit root	(0.44)

Table 2 presents the results of pool unit root tests for the series $\{\overline{dy_t^m}\}$ (U.S., U.K., Canada and Japan) with individual effects (p-values in parentheses).

Table 3: Regressions for the predictability of dividend growth

	Annual dy Equation (2)		MiDaS Monthly dy Equation (13)		MiDaS - Smoothed dy (equation (11))				F-test H_0: $c_1 = c_2$	With restriction $c_2 = 0$	
Panel A (dep. var. Δd_{t+1})	$c_{1,d}$	p-value	c'_1	p-value	c_1	p-value	c_2	p-value	p-value	c_1	p-value
U.S.	-0.04	0.55	-0.05	0.18	-0.16	0.02	-0.05	0.44	0.04	-0.17	0.04
U.K.	-0.24	0.23	0.07	0.35	-0.15	0.01	-0.36	0.56	0.02	-0.22	0.00
Canada	-0.13	0.09	-0.14	0.04	-0.16	0.09	-0.09	0.22	0.19	-0.21	0.04
Japan	-0.07	0.15	-0.11	0.08	-0.18	0.03	-0.07	0.46	0.07	-0.22	0.05
Panel B (dep. var. $\overline{\Delta^a d^m_{t+1}}$)	$c_{1,d}$	p-value	c'_1	p-value	c_1	p-value	c_2	p-value	p-value	c_1	p-value
U.S.	-0.03	0.69	-0.04	0.50	-0.17	0.03	-0.06	0.50	0.28	-0.18	0.03
U.K.	-0.33	0.27	-0.20	0.01	-0.26	0.00	-0.54	0.00	0.01	-0.35	0.01
Canada	-0.15	0.13	-0.20	0.02	-0.19	0.09	-0.21	0.02	0.44	-0.33	0.05
Japan	-0.13	0.22	-0.11	0.01	-0.20	0.06	-0.00	0.99	0.04	-0.20	0.05

Table 3 presents the results on one-year-ahead dividend growth predictability. The first block of columns of Table 3 corresponds to equation (2). The second block of columns corresponds to an application of MiDaS without the decomposition of the dividend-price ratio. Specifically, MiDaS with monthly dy corresponds to equations $\Delta d_{t+1} = c'_0 + c'_1 wdy^m_t + u_{t+1}$ (panel 3.A), and $\overline{\Delta^a d^m_{t+1}} = c'_0 + c'_1 wdy^m_t + u_{t+1}$ (panel 3.B), where $wdy^m_t = \sum_{j=0}^{11} w'_j dy^m_{t-j/12}$. The third block of columns corresponds to equation (11). p-values correspond to Newey-West t-statistics. The F-test tests the hypothesis $c_1 = c_2$.

Table 4: MiDaS with quarterly dividends

Panel A (dep. var. Δd_{t+1})					With restriction $c_2 = 0$	
	c_1	p-value	c_2	p-value	c_1	p-value
U.S.	-0.07	0.06	0.01	0.90	-0.12	0.24
U.K.	0.12	0.23	-0.18	0.29	-0.13	0.16
Canada	-0.07	0.22	-0.17	0.17	-0.07	0.80
Japan	0.25	0.25	-0.21	0.03	-0.57	0.66
Panel B (dep. var. $\overline{\Delta^a d^m_{t+1}}$)					**With restriction $c_2 = 0$**	
	c_1	p-value	c_2	p-value	c_1	p-value
U.S.	-0.10	0.31	0.00	0.35	-0.10	0.32
U.K.	-0.26	0.02	-0.02	0.18	-0.18	0.02
Canada	-0.16	0.38	-0.01	0.12	-0.13	0.25
Japan	-0.59	0.63	-0.37	0.99	-0.60	0.63

Table 4 presents the results of a MiDaS estimation of equation (11) with the use of quarterly data (four subperiods each year). p-values correspond to Newey-West t-statistics.

Table 5: Longer horizon MiDas regressions for dividend growth predictability

Panel A (dep. var. Δd_{t+i})							With restriction $c_2 = 0$		
i	2		3		4		2	3	4
	c_1	c_2	c_1	c_2	c_1	c_2		c_1	
U.S.	-0.18	-0.07	-0.19	-0.09	-0.24	0.01	-0.18	-0.17	-0.22
	(0.01)	(0.37)	(0.01)	(0.36)	(0.02)	(0.88)	(0.03)	(0.01)	(0.02)
U.K.	-0.08	-0.09	-0.07	-0.11	-0.64	0.23	-0.10	-0.09	-0.15
	(0.03)	(0.01)	(0.05)	(0.27)	(0.01)	(0.01)	(0.10)	(0.12)	(0.00)
Canada	-0.18	-0.08	-0.23	-0.10	-0.29	0.17	-0.23	-0.25	-0.31
	(0.07)	(0.30)	(0.01)	(0.17)	(0.02)	(0.13)	(0.02)	(0.07)	(0.02)
Japan	-0.33	-0.05	-0.36	-0.07	-0.36	-0.22	-0.41	-0.37	-0.35
	(0.05)	(0.65)	(0.04)	(0.62)	(0.26)	(0.05)	(0.08)	(0.09)	(0.13)
Panel B (dep. var. $\overline{\Delta^a d^m_{t+i}}$)							With restriction $c_2 = 0$		
i	2		3		4		2	3	4
	c_1	c_2	c_1	c_2	c_1	c_2		c_1	
U.S.	-0.21	-0.07	-0.16	-0.09	-0.16	0.02	-0.18	-0.18	-0.15
	(0.16)	(0.46)	(0.31)	(0.44)	(0.27)	(0.81)	(0.16)	(0.35)	(0.28)
U.K.	-0.44	-0.67	-0.20	-0.04	-0.62	0.72	-0.17	-0.21	-0.28
	(0.07)	(0.63)	(0.01)	(0.66)	(0.00)	(0.00)	(0.00)	(0.04)	(0.00)
Canada	-0.27	-0.25	-0.35	-0.13	-0.72	0.19	-0.35	-0.39	-0.61
	(0.01)	(0.01)	(0.01)	(0.11)	(0.00)	(0.12)	(0.02)	(0.01)	(0.00)
Japan	-0.14	-0.02	-0.24	-0.02	-0.21	0.22	-0.14	-0.24	-0.15
	(0.01)	(0.82)	(0.00)	(0.73)	(0.10)	(0.18)	(0.06)	(0.01)	(0.11)

Table 5 presents the results of the estimation of equations $\Delta d_{t+i} = c_0 + c_1 \Delta w\overline{dy^m_t} + c_2 \overline{dy^m_{t-1}} + u_{t+i}$ (panel 5.A), and $\overline{\Delta^a d^m_{t+1}} = c_0 + c_1 \Delta w\overline{dy^m_t} + c_2 \overline{dy^m_{t-1}} + u_{t+i}$ (panel 5.B), with i =2, 3 and 4. p-values in parentheses correspond to Newey-West t-statistics.

Table 6: In-sample adjusted R^2 (%) of MiDaS regressions for dividend growth predictability

Panel A (dep. var. Δd_{t+i})					With restriction $c_2 = 0$			
i	1	2	3	4	1	2	3	4
U.S.	16	22	24	22	15	20	21	19
U.K.	28	21	25	27	21	22	25	24
Canada	24	22	29	30	16	19	22	23
Japan	20	23	25	20	17	21	21	18

Panel B (dep. var. $\overline{\Delta^a d^m_{t+i}}$)					With restriction $c_2 = 0$			
i	1	2	3	4	1	2	3	4
U.S.	16	18	17	10	12	12	12	10
U.K.	28	55	38	55	24	39	38	39
Canada	24	35	28	32	18	19	21	37
Japan	20	24	27	32	25	24	27	26

Table 6 presents the values of the in-sample adjusted R^2 of equations $\Delta d_{t+i} = c_0 + c_1 \Delta w\overline{dy^m_t} + c_2 \overline{dy^m_{t-1}} + u_{t+i}$ (panel 6.A), and $\overline{\Delta^a d^m_{t+i}} = c_0 + c_1 \Delta w\overline{dy^m_t} + c_2 \overline{dy^m_{t-1}} + u_{t+i}$ (panel 6.B), with i =1, 2, 3 and 4.

Table 7: Regressions for the predictability of dividend growth without MiDaS

	Estimated equation: $z_{t+1}= c_0+c_1\Delta dy_t+c_2 dy_{t-1}+u_{t+1}$							Estimated equation: $z_{t+1}= c_0+c_1\Delta\overline{dy_t^m}+c_2\overline{dy_{t-1}^m}+u_{t+1}$						
Panel A					F-test	With restriction						F-test	With restriction	
$z_{t+1}= \Delta d_{t+1}$					H_0:$c_1 = c_2$	$c_2= 0$						H_0:$c_1= c_2$	$c_2= 0$	
	c_1	p-value	c_2	p-value	p-value	c_1	p-value	c_1	p-value	c_2	p-value	p-value	c_1	p-value
U.S.	-0.24	0.10	-0.04	0.49	0.09	-0.22	0.11	-0.01	0.73	0.01	0.78	0.34	-0.01	0.70
U.K.	-0.12	0.51	-0.12	0.45	0.50	-0.06	0.68	-0.01	0.65	-0.08	0.46	0.25	-0.00	0.75
Canada	-0.09	0.49	-0.13	0.09	0.39	-0.03	0.82	-0.01	0.37	-0.11	0.07	0.04	-0.01	0.38
Japan	-0.12	0.51	-0.08	0.45	0.41	-0.06	0.68	-0.01	0.67	-0.04	0.59	0.35	-0.01	0.73
Panel B					F-test	With restriction						F-test	With restriction	
$z_{t+1}=\overline{\Delta^a d_{t+1}^m}$					H_0:$c_1 = c_2$	$c_2= 0$						H_0:$c_1= c_2$	$c_2= 0$	
	c_1	p-value	c_2	p-value	p-value	c_1	p-value	c_1	p-value	c_2	p-value	p-value	c_1	p-value
U.S.	-0.26	0.17	-0.04	0.54	0.13	-0.23	0.17	-0.01	0.68	0.01	0.83	0.35	-0.01	0.66
U.K.	-0.12	0.67	-0.17	0.50	0.45	-0.03	0.88	-0.00	0.90	-0.11	0.52	0.26	0.00	0.99
Canada	-0.06	0.76	-0.17	0.10	0.29	0.02	0.94	-0.00	0.74	-0.16	0.05	0.02	-0.00	0.80
Japan	0.03	0.83	0.02	0.77	0.47	0.02	0.87	-0.00	0.80	0.01	0.89	0.41	-0.00	0.76

Table 7 presents the results of the estimation of equations $z_{t+1} = c_0 + c_1\Delta dy_t + c_2 dy_{t-1} + u_{t+1}$ and $z_{t+1} = c_0 + c_1\Delta\overline{dy_t^m} + c_2\overline{dy_{t-1}^m} + u_{t+1}$. p-values correspond to Newey-West t-statistics. F-test tests the hypothesis $c_1 = c_2$.

Table 8: Short term MiDaS predictions of dividend growth

Panel A (dep. var. $d^{s1}_{t+1} - d^{s2}_t$)						With restriction $c_2 = 0$		
	c_1	p-value	c_2	p-value	adj. R^2 (%)	c_1	p-value	adj. R^2 (%)
U.S.	-0.22	0.029	-0.01	0.861	15	-0.21	0.025	15
U.K.	-0.34	0.004	0.38	0.008	18	-0.30	0.002	8
Canada	-0.09	0.103	-0.05	0.136	13	-0.09	0.095	17
Japan	-0.72	0.000	0.21	0.151	20	-0.50	0.000	17
Panel B (dep. var. $d^{q1}_{t+1} - d^{q4}_t$)						**With restriction $c_2 = 0$**		
	c_1	p-value	c_2	p-value	adj. R^2 (%)	c_1	p-value	adj. R^2 (%)
U.S.	-0.18	0.091	-0.03	0.265	8	-0.17	0.082	8
U.K.	-0.33	0.036	0.54	0.023	17	-0.46	0.073	5
Canada	-0.20	0.003	-0.03	0.322	11	-0.19	0.006	11
Japan	-1.57	0.021	0.97	0.002	27	-1.27	0.093	2

Table 8 presents the results of the estimation of equations $d^{s1}_{t+1} - d^{s2}_t = c_0 + c_1 \Delta w \overline{dy^m_t} + c_2 \overline{dy^m_{t-1}} + u_{t+1}$ (panel A) and $d^{q1}_{t+1} - d^{q4}_t = c_0 + c_1 \Delta w \overline{dy^m_t} + c_2 \overline{dy^m_{t-1}} + u_{t+1}$ (panel B), where d^{si}_t and d^{qi}_t are the log dividends paid within i-th semester and i-th quarter of year t, respectively, for each index. p-values correspond to Newey-West t-statistics.

Figure 1: Estimated MiDaS weights of the annual changes of monthly dividend-price ratios for the predictive regression of dividend growth (equation 11)