

Four centuries of return predictability

Benjamin Golez†
University of Notre Dame

Peter Koudijs‡
Stanford University

This version: June 26, 2015

Abstract

We analyze aggregate market prices and dividends throughout modern financial history. Focusing on the most important market in each period, we start with the Dutch and English markets in the early 17th century and move to the U.S. market at the end of the 19th century. We find that the dividend-to-price ratio predicts returns throughout all four centuries. “Excess volatility” is thus a pervasive feature of financial markets. The dividend-to-price ratio also predicts dividend growth rates in all but the most recent (post 1945) period. Cash-flows were therefore much more important for price movements in the past, and the dominance of discount rate news is a relatively recent phenomenon. This is consistent with the increased duration of the stock market in the recent period.

Key words: Dividend-to-price ratio, return predictability, dividend growth predictability, duration

JEL classification: G12, G17, N2

†256 Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556, USA, Tel.: +1-574-631-1458, bgolez@nd.edu

‡655 Knight Way, Stanford Graduate School of Business, Stanford, CA 94305, Tel.: +1-650-725-1673, koudijs@stanford.edu

*We thank Jules van Binsbergen (discussant), Anna Cieslak, Zhi Da, Monika Piazzesi, Robert Stambaugh (discussant), Jessica Wachter, and seminar participants at University of Notre Dame, University of California San Diego, Stanford University, European Central Bank, 2014 SITE Conference, 2015 Rodney L. White Center for Financial Research Conference, and 2015 Western Finance Association Conference for comments and suggestions. We thank Gary Shea for generously sharing his data about 18th century British companies, John Turner for his advice on how to use British information from the 19th century, and Manisha Goswami and Cathy Quiambao for research assistance.

1. Introduction

One of the most important questions in asset pricing is whether prices (or rather the dividend-to-price ratio) can predict returns. If so, asset prices would be “excessively volatile;” that is, they move more than is warranted by fundamentals, such as dividends (Shiller 1981; LeRoy and Porter 1981). The empirical evidence suggests that returns are indeed partially predictable (Campbell and Shiller 1988; Fama and French 1988; Cochrane, 2008; Binsbergen and Koijen 2010). This has motivated an important theoretical literature that incorporates time-varying returns in equilibrium models (Campbell and Cochrane 1999; Bansal and Yaron 2004; Albuquerque, Eichenbaum, and Rebelo 2014).

Two issues remain. First, in the recent period, the dividend-to-price ratio is highly persistent, virtually indistinguishable from a unit root. Combined with the relatively short sample, this biases estimates in favor of finding return predictability (Stambaugh 1999). Second, the existing evidence not only suggests that returns are predictable, it also indicates that dividend growth rate predictability is limited (Campbell and Shiller 1988; Campbell 1991; Cochrane 1992; 2008; 2011). This implies that “excess volatility” is extreme: prices seem to move only in response to changing expected returns and not to news about future dividends. The jury is still out on what can explain this feature of the data.

In this paper we extend the time series of asset prices and dividends to cover the whole history of modern financial markets starting in 17th century Amsterdam. In particular, we focus on the dominant stock markets of the time: the Dutch stock market in the 17th and 18th century, the U.K. stock market in the 18th and 19th century, and the U.S. stock market from the end of the 19th century onwards. With this focus, we cover a large fraction of global market capitalization. The included companies are very similar to those of today, with limited liability for shareholders,

separation of ownership and control, and an active secondary market for shares. By extending the time series, we add independent variation to the data. This is more difficult to achieve in the cross-section, where markets often move together, especially in the recent period.

The paper has four key findings. First, over the long run the dividend-to-price ratio is stationary, fluctuating around a long-run average of five percent (minus three in logs, see Figure 1 for details). It is only after around 1945 that the dividend-to-price ratio starts to persistently decrease. This is interesting in its own right, but also means that, for the period as a whole, results are not biased due to a non-stationary predictor. Although the dividend-to-price ratio has decreased, expected returns have been approximately stable over time, with returns increasingly coming from capital appreciation in lieu of decreasing dividends.

Second, we find robust evidence for return predictability. In the full period covering all four centuries, the predictive coefficient on the dividend-to-price ratio is positive and highly significant, for both annual and multiannual horizons. In subperiods, the predictive coefficient is remarkably stable (although not always statistically significant). Thus, excess volatility appears to be a pervasive characteristic of financial markets.

Third, there are important differences between periods in terms of dividend growth predictability. While the dividend-to-price ratio strongly predicts dividend growth rates in the earlier periods, such predictability completely disappears around 1945. In line with this observation, our analysis implies that, before 1945, changes in cash-flows were more important for price movements than changes in discount rates. The dominance of discount rate news is therefore a relatively recent phenomenon.

Fourth, return predictability is concentrated in recession periods - more so than dividend growth rate predictability. This suggests that discount rate news is more important than cash flow news during recessions.

We hypothesize that the growing importance of time-varying expected returns is related to the fact that dividend payments have decreased since 1945 and that returns have started to predominantly consist of capital appreciation. This is consistent with firms reinvesting cash-flows and pushing dividend payments into the future (Fama and French 2001).¹ We hypothesize that this has led to the increased duration of the stock market as a whole. The bond pricing literature suggests that when cash flow distributions (dividend payments) are shifted into the future and duration increases, prices become more sensitive to discount rate shocks.

This hypothesis is supported by two empirical observations. First, we document that in the post-1945 period, the discounted value of the next ten years of (expected) dividends accounts for a significantly smaller portion of current stock valuations than in earlier periods. Using one minus this discounted value as a proxy for stock duration, we find that it is highly correlated with the importance of discount rate news. Second, we document that the market betas of long maturity Treasury bonds have increased relative to the betas of short maturity bonds in the recent U.S. period. Thus, the stock market as a whole has become more correlated with longer-duration assets.

Although these results are supportive of our hypothesis, other explanations put forward in the literature may also play a role. Chen, Da, and Priestly (2012) argue that dividend smoothing may bias results towards discount rate news. Our tests suggest that both dividend smoothing and stock duration matter, although stock duration explains a much larger fraction of discount rate news

¹ Many factors may have contributed to this: taxes, buybacks, dividend smoothing, as well as improved transparency.

variation. In addition, we document that dividend smoothing in the earlier part of our sample was comparable to today.

Our paper relates to a large literature on the importance of dividend growth and discount rates for stock prices. See Kojien and Van Nieuwerburgh (2011) for an overview. Most closely related are Goetzmann, Ibbotson, and Peng (2001), Chen (2009), and Rangvid, Schmeling, and Schrimpf (2014). Using primary sources, Goetzmann et al. estimate an index for the New York stock market between 1815 and 1925. They find little evidence for return predictability, but due to data limitations, they have to approximate dividends for the period before 1870. Chen examines the differences between the pre- and post-1945 U.S. periods and argues that before 1945 returns were not predictable, assigning almost all variation in the dividend-to-price ratio to expected dividend growth rates. Rangvid et al. show that discount rates are less important in countries with relatively small companies and less dividend smoothing.

Our paper is also related to a growing literature that emphasizes the predictability of dividend growth rates (Menzly, Santos, and Veronesi 2004; Lettau and Ludvigson 2005; Cochrane 2008; Binsbergen and Kojien 2010; Golez 2014). Relative to this literature, we provide a long term perspective on the sources of asset price movements that help us understand why discount rate news dominates in the recent period. Finally, our paper is related to le Bris, Goetzmann, and Pouget (2014), who analyze six hundred years of dividend and price data for the Bazacle Company in France.

2. Data

We extend the time series of stock prices and dividends back in time until 1629 using the most important financial markets of a specific period. In particular, for the period between 1629

and 1811, we focus on the equity market in Amsterdam (that included a number of English securities). For 1825-1870, we look at London. For the period after 1870, we rely on U.S. market data. In total, we construct an annual time series from 1629 to 2012, with only a small gap for the years between 1811 and 1825.

While the American data have been extensively studied, the Dutch and English data have received limited attention in the literature and merit closer inspection. This paper is the first to look at return and dividend growth rate predictability in these markets.

During the 17th and 18th century, Amsterdam was the financial capital of the world, and it was closely integrated with the London market (Neal 1990). Although technologically less advanced, the market functioned much like the one today. Harrison (1998) provides evidence that returns in these markets had similar distributions and time series properties as today. Koudijs (2014) shows that the Amsterdam market responded to the arrival of news in an efficient way and that trading costs were very similar to the recent period. The (negative) autocorrelation of returns on a daily level is comparable to today. We take the perspective of an Amsterdam investor, assuming that he held a value-weighted portfolio of Dutch and English securities. We use exchange rate information to convert returns in Pounds Sterling into Dutch Guilder returns.² There is information available for five securities: the Dutch East India Company (from 1629 onwards), the Bank of England (1694), the (United) British East India Company (1693), the South Sea Company (1711), and the (Second) Dutch West India Company (1719). Details on data sources are provided in Appendix A.

The Dutch East India Company (VOC) was the world's first publicly traded corporation; its shares were freely tradable, and shareholders enjoyed limited liability. There was a clear

² Since both England and the Dutch Republic were on metallic standards, exchange rate fluctuations were only of minor importance.

separation between ownership and control. The Company was founded in 1602, and its capital became permanent in 1613 (Gelderblom, De Jong, and Jonker 2013). It held the Dutch monopoly on trade with Asia and operated an extensive trade network there. It started to pay annual dividends in 1685; before that it paid dividends every two to three years.³ The company was nationalized by the government in 1796. One of the contributions of this paper is the collection from the original sources of a complete VOC price series for the years between 1629 and 1719.

The Dutch West India Company (WIC) was founded in 1675 and was involved in slave trade and the administration of (slave) colonies in Africa and Caribbean. It paid out dividends sporadically and was nationalized in 1791. Price information is only available for 1719 onwards. In that year it constituted only one percent of our index, so the omission of the WIC during 1675-1718 probably has little impact on our estimates.

The Bank of England (BoE) was founded in 1694 to help finance the English government debt. It held an effective monopoly over issuing banknotes and provided short-term credit to merchants and banks. It was also an important lender to the British East India Company (EIC). The EIC was created in 1708 through a merger of the Old and New East India Companies (between 1693 and 1707 we use the prices of the Old EIC). It held the English monopoly on trade with Asia. The South Sea Company (SSC) started in 1711 after receiving a monopoly on the trade with South America. These activities never materialized, and the Company was mainly a vehicle to finance the English government debt. It performed a number of debt-for-equity swaps; the final one resulted in the South Sea Bubble in 1720. In that year the company accounts for 60% of our value-weighted index. After the bubble burst, the company was largely liquidated; in 1732, it constituted only 6% of our index. Remaining shares were mainly backed

³ It is impossible to calculate annual dividend growth rates and dividend-to-price ratios for the VOC up to 1684. Since it is the only company in our sample for that period, we can only run our annual regressions for the years after 1684.

by government debt. It matters very little for our results whether or not we keep the company in our index after 1732. The English companies have a complicated history of capital calls, rights issues, and other “capital events.” We use the work by Shea (*in preparation*) to adjust stock prices where necessary.

It is important to note that, even though these are only five securities, they effectively constitute the universe of traded equities in Amsterdam and London. Only during the bubble year of 1720 did new equities enter the market; most of these new companies were liquidated before the end of the year (Frehen, Goetzmann, and Rouwenhorst 2013). The few surviving companies were relatively small and were not widely traded.

The companies in our index were quite large: the total market capitalization to GDP of securities held by Dutch investors ranged from 15% (during the 1630s and again in the early 1800s) to 64% (during the 1720s).⁴ This means that diversification was provided within companies, rather than between them. In comparison, for the U.S., stock market capitalization amounted to 39% of GDP in 1913 and 152% in 1999 (Rajan and Zingales 2003). In addition, there were many investment opportunities available outside the stock market such as shipping, trade and small manufacturing that would have expanded the efficient portfolio frontier.⁵

For the period between 1825 and 1870 we focus on the London market. After the Napoleonic Wars, London became the financial capital of the world, and the United Kingdom was the largest economy in the world. Starting in the 1810s many new equities were issued. Initially, these were mainly canals and insurances companies. Later on, banks and railroad

⁴ To compute these numbers we use the GDP of Holland, the only Dutch region for which reliable figures are available for the 17th and 18th centuries (Van Zanden and Van Leeuwen 2012). Holland was the largest province of the Dutch Republic, comprising the most populous and developed parts of the country, including important cities like Amsterdam and Rotterdam. The historical evidence suggests that most Dutch investors lived in this area. We used information from Bowen (1989) and Wright (1997) to calculate what fraction of English securities were held by Dutch investors.

⁵ Although the overall riskiness of a portfolio depends on diversification possibilities, it should not affect return or dividend growth predictability.

companies became the most important issuers of new equity. The period covers the so-called Railroad “Manias” of the 1830s and 1840s. It is important to note that before 1855, the newly issued companies had full shareholder liability. Afterwards, it became possible to issue shares with limited liability, but many banks and insurance companies continued to maintain full liability.

We use the value-weighted stock market index constructed by Acheson, Hickson, Turner, and Ye (2009) (henceforth, AHTY) that includes all regularly traded domestic equities traded in London starting in 1825. The index covers between 125 (1825) and 250 (1870) different securities. Total market capitalization accounted for between 10 and 30 percent of British GDP. During this period, there were many new issues and delistings. AHTY (2009) omits all securities that were traded for less than 12 months (most of these companies failed to raise sufficient capital to start their businesses) and adjust for survivorship bias by reconstructing companies’ value at liquidation or delisting. In addition, there were many capital calls, rights issues, and other capital events. AHTY (2009) omits individual security returns for the months in which these events took place. See AHTY (2009) and Hickson, Turner, and Ye (2011) for more details.

Starting in 1871, we rely on the U.S. stock market. By 1900, U.S. had become the largest economy of the world, with a well-developed capital market in New York. We obtain the data from Amit Goyal’s webpage. For the period between 1871 and 1925, these data rely on information from Cowles (1939) that covers between 50 (1871) and 258 (1925) securities. From 1926, the data are based on the S&P 500 index provided by CRSP. Before 1957, this was actually the S&P 90. As before, our U.S. return index is value-weighted.

3. Methodology

As is standard in the asset pricing literature, we take the perspective of an individual investor who is interested in the per share value of a company. An investor receives dividends as the only source of cash-flows. Other types of distributions (e.g. repurchases) are assumed to be reinvested in the company.⁶

3.1 Present value relations

The holding period return per share of equity consists of the dividend yield and any price appreciation:

$$R_t = \frac{P_t + D_t}{P_{t-1}}, \quad (1)$$

where P_t is the per share price at time t and D_t are the per share dividends accumulated from $t-1$ to t . We take logs and define the dividend-to-price ratio as $dp_t = \log(D_t / P_t)$ and the dividend growth rate as $dg_t = \log(D_t / D_{t-1})$. Using a first-order Taylor expansion around the long-run mean of the dividend-to-price ratio \overline{dp} , Campbell and Shiller (1988) show that log returns can be expressed as:

$$r_{t+1} \approx dp_t + dg_{t+1} - \rho dp_{t+1}, \quad (2)$$

where all variables are demeaned and $\rho = \exp(-\overline{dp}) / 1 + \exp(-\overline{dp})$ is the linearization constant.

Rewriting Eq. (2) in terms of the dividend-to-price ratio we obtain:

$$dp_t \approx r_{t+1} - dg_{t+1} + \rho dp_{t+1}. \quad (3)$$

⁶ See Section 5.1 for a discussion on an alternative approach pursued by Larrain and Yogo (2008), which takes the perspective of a representative investor interested in the value of the whole company.

Eq. (3) shows that a high dividend-to-price ratio is related to (and should therefore predict) high future returns, and/or low future dividend growth rates, and/or a high future dividend-to-price ratio. Because the predictive coefficients are interrelated, return and dividend growth predictability should best be studied jointly (Lettau and Ludvigson 2005; Cochrane 2008; Binsbergen and Kojien 2010; Golez 2014).

Iterating Eq. (3) forward and excluding rational bubbles, the dividend-to-price ratio can also be expressed as an infinite sum of discounted returns and dividend growth rates (since the relationship holds ex-ante and ex-post, an expectations operator can be added to the right-hand side):

$$dp_t \approx E_t \sum_{j=0}^{\infty} \rho^j (r_{t+1+j}) - E_t \sum_{j=0}^{\infty} \rho^j (dg_{t+1+j}). \quad (4)$$

Thus, ultimately, any variation in the dividend-to-price ratio must be related to future changes in expected returns and/or expected dividend growth rates.

Finally, the above present value model also allows study of variation in unexpected returns (Campbell 1991). Subtracting the expectations of Eq. (4) at time $t+1$ from the expectations at time t yields:

$$r_{t+1} - E_t r_{t+1} = -(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j (r_{t+1+j}) + (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j (dg_{t+1+j}). \quad (5)$$

Hence, unexpected return can be high either because the expected future dividend growth rate is high or because future expected returns are low.

3.2 Estimation

We estimate the joint dynamics of returns, dividend growth rates, and the dividend-price ratio through a vector autoregression (VAR) model:

$$x_{t+1} = \phi x_t + \varepsilon_{t+1}, \quad (6)$$

where $x_t = [r_t, dg_t, dp_t]'$ is a column vector of three variables. All variables are demeaned. Denote by $\Sigma = E[\varepsilon_t \varepsilon_t']$ the covariance matrix of residuals, and by $\Gamma = E[x_t x_t']$ the covariance matrix of the variables.

The model is identified by nine moment conditions:

$$E[(x_{t+1} - \phi x_t) \otimes x_t] = 0 \quad (7)$$

The present value relations in Eq. (3) add further restrictions on the estimated parameters. Let I be a three by three identity matrix, and let e_i denote the i th column of the identity matrix. Then the restrictions can be written as:

$$(e_1' - e_2' + \rho e_3') \phi = e_3'. \quad (8)$$

In total we have nine moment conditions, nine parameters and three linear restrictions. The VAR model is therefore overidentified. We estimate the model using iterative GMM, and we test for overidentifying restrictions using a J -test. Heteroscedasticity and autocorrelation consistent statistics are based on Bartlett kernel with optimal bandwidth determined by the Newey-West method. A similar approach is used by Larrain and Yogo (2008), among others.

3.3 Decompositions

Using the VAR model, we infer long-horizon estimates from their short-run analogs. We start by decomposing the variance of the dividend-to-price ratio into the covariances with future returns and dividend growth rates (Cochrane 1992):

$$\text{Var}(dp_t) = \text{Cov}\left(dp_t, \sum_{j=0}^{\infty} \rho^j (r_{t+1+j})\right) + \text{Cov}\left(dp_t, -\sum_{j=0}^{\infty} \rho^j (dg_{t+1+j})\right) \quad (9)$$

In terms of the VAR model, the covariance terms can be written as:

$$\text{Var}(dp_t) = e_3' \Gamma e_3 = e_1' \phi (I - \rho \phi)^{-1} \Gamma e_3 - e_2' \phi (I - \rho \phi)^{-1} \Gamma e_3. \quad (10)$$

The first covariance term can be interpreted as the variation of the dividend-to-price ratio due to discount rates. The second term captures variation due to cash-flows. To determine the relative importance of the two components, we divide the covariance terms by the variance of the dividend-to-price ratio and express them in percentages.

Similarly, we can decompose the variance of unexpected returns from Eq. (5) into a discount rate and a cash-flow component.

$$\text{Var}(r_{t+1} - E_t r_{t+1}) = -\text{Cov} \left[r_{t+1} - E_t r_{t+1}, (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j (r_{t+1+j}) \right] + \text{Cov} \left[r_{t+1} - E_t r_{t+1}, (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j (dg_{t+1+j}) \right]. \quad (11)$$

In the context of the VAR model, the covariance terms can be written as:

$$\text{Var}(r_{t+1} - E_t r_{t+1}) = -e_1' \rho \phi (I - \rho \phi)^{-1} \Sigma e_1 + e_2' (I - \rho \phi)^{-1} \Sigma e_1. \quad (12)$$

(Campbell 1991). Again, to determine the relative importance of each component, we divide the covariance terms by the variance of unexpected returns and express them in percentages.

4. Results

We start by presenting the summary statistics, followed by the VAR estimates, and the decomposition results.

4.1 Summary statistics

Figure 1 plots the main variables of interest: annual returns, dividend growth rates, and the dividend-to-price ratio. Dashed lines separate the different time periods: Netherlands/U.K. (1685-1809), U.K. (1825-1870), and the two U.S. samples. Following Chen (2009), we split the

U.S. sample in the early U.S. period (1871-1945) and the recent U.S. period (1945-2012). Table 1 presents the corresponding summary statistics.

What stands out most vividly is the remarkable stationarity of the dividend-to-price ratio. It always oscillates around the long-run average of approximately 5% (minus 3 in log terms). Only in the recent period (post-1945), the dividend-to-price ratio becomes persistently decreasing. Whereas the persistence of the dividend-to-price ratio in the first three periods is between 0.51 and 0.66, it increases to as much as 0.91 in the post-1945 data. The recent U.S. period is the only period for which we cannot reject a unit root.

Average returns are between 5.10% and 6.94% in the earlier years and increase to 10.00% in the post-1945 period. At the same time, the volatility of returns is somewhat higher in the two U.S. periods. Therefore, we do not observe any clear differences in the Sharpe ratios across the different periods. Persistence of returns is relatively low, with the AR(1) ranging between -0.09 and 0.05.

Capital appreciation is much more important for returns in the recent period. Whereas in the earlier periods around two thirds of returns stem from dividends (and only one third comes from capital appreciation), this is exactly the opposite in the recent U.S. sample.

We also see highly volatile dividend growth rates in the early years of the first period. This is due to the fact that companies did not always pay out dividends each and every year, and this is reflected by the increased volatility of aggregate dividend growth. Volatility is substantially lower after 1720, which can be interpreted as a first indication of dividend smoothing. The volatility of dividend growth rates increased again in London in the 19th century and in the early U.S. period. It flattened out once again in the recent years, probably due to increased dividend smoothing. Accordingly, dividend growth rates become increasingly

persistent over time. While the AR(1) for dividend growth is -0.22 in the Dutch/English period, it increases to 0.39 in the recent U.S. period. To reduce the effect of both lumpy dividends in the initial years (and overly smooth dividends in the recent period), we also report results based on triennial data (in addition to annual data). By using triennial data, we also are able to extend the time period backwards by 56 years to 1629.⁷

4.2 VAR estimates and decomposition results

Table 2 presents the VAR estimates and decomposition results based on annual data. We estimate a VAR for each period as well as for the full sample (by appending periods). The same results based on the triennial data are reported in Table 3. When using triennial data, we estimate our model on three different non-overlapping samples and then report the mean of the estimated parameters across those samples. We always report the full parameter matrix associated with the VAR system, but focus our attention on the parameters associated with the lagged dividend-to-price ratio.

We start by analyzing the results reported in Table 2 that are based on annual data. We first note that, in the full period, the dividend-to-price ratio predicts both returns and dividend growth rates. The estimated parameters on the dividend-to-price ratio are of the expected sign; positive at 0.08 in the return regression, and negative at -0.12 in the dividend growth regression. Both parameters are also highly significant. This stands in sharp contrast to the recent evidence that the dividend-to-price ratio only predicts returns.

Note once more that the dividend-to-price ratio in the full sample is stationary, and our sample spans 310 annual observations. Thus, our inference does not suffer from the general

⁷ See footnote 3 for details.

problems associated with predictability regressions that use highly persistent predictors in small samples (Stambaugh 1999).⁸

We observe substantial variation between periods. Whereas the estimated parameter in the return regression is fairly stable, the estimated parameters in the dividend growth and the dividend-to-price regressions change importantly in the recent period. Effectively, the predictability of the dividend growth rate disappears in the recent period and is substituted with increased persistence of the dividend-to-price ratio.⁹

In particular, in the return regression, the estimated parameter on the dividend-to-price ratio is always positive and relatively stable, between 0.08 and 0.12. It is significant in the Dutch/English period and in the recent U.S. sample. Even in the 19th c. U.K. period and the early U.S. period, where the estimated parameters are statistically insignificant, Wald tests suggest that these two coefficients are not statistically significantly different from the rest of the period. This implies that returns have always been predictable by the dividend-to-price ratio. In comparison, the estimated parameter in the dividend growth regression is much less stable. It is negative, between -0.19 and -0.34, and significant in the first three periods, but turns out to be (exactly) zero in the recent U.S. period. According to a Wald test, this difference is highly statistically significant.

The disappearance of dividend growth predictability is associated with an increased persistence of the dividend-to-price ratio. Whereas the dividend-to-price ratio predicts itself with the estimated parameter between 0.61 and 0.73 in the first three periods, this coefficient increases to 0.90 in the recent U.S. sample.

⁸ We also note that the correlation between residuals in the return regression and innovations in the dividend-to-price ratio is less negative in the full sample at -0.69 as compared to the recent period -0.93.

⁹ Estimated parameters on the dividend-to-price ratio need to satisfy the linear restriction $\phi_{r,dp} - \phi_{dg,dp} + \varphi\phi_{dp,dp} = 1$. This means that a change in one of the parameters needs to be accompanied with a change in at least one other parameter.

All in all, the recent U.S. period appears very different. Return predictability is somewhat higher. Most importantly, there is a complete lack of dividend growth predictability, and a highly persistent dividend-to-price ratio. This also reflects itself in the decomposition results reported in Panel B of Table 2. While in the first three samples, cash-flows are much more important, in recent years, all the variation in the dividend-to-price ratio appears to be driven by discount rates. We obtain qualitatively similar results if we focus on the decomposition of unexpected returns. In both cases, cash-flows account for approximately two thirds of the variation in the first three samples and do not matter at all in the recent period. In comparison, discount rates account for only one third of variation in the early periods and completely dominate in the most recent period.¹⁰

Given the stark differences between results in different periods, we also conduct a Wald test on the full period imposing a break in 1945. The Wald test confirms the presence of a structural break with a p-value of 0.003. When we endogenously search for one break, we arrive at 1961 (using a Sup statistic). Finally, we conduct Monte Carlo simulations. We use the distribution of parameters and errors estimated using the data up to 1945 to simulate 10,000 datasets that match the length of the post-1945 period. We then perform the same decompositions as in the main analysis. The probability density functions for simulated values of discount rate news components along with the actual values from the data are reported in Figure 2. We confirm that the results in the post-1945 period are statistically different with a p-value of 0.002 when the decomposition is based on the dividend-to-price ratio and a p-value of 0.000 when the decomposition is based on unexpected returns.¹¹

¹⁰ In Online Appendix B, we show that the results are robust to the adjustment for inflation.

¹¹ To speed up simulations, we use one step GMM estimation and set the Newey-West bandwidth to 3.

All the results are qualitatively similar when we use triennial (rather than annual) data, as reported in Table 3. The dividend-to-price ratio predicts both returns and dividend growth rates in the full period. The observation that price movements are mostly driven by cash-flow news in the early years and are dominated by the discount rate news in the recent period remains robust and strong. Both a Wald test and simulation results confirm that the importance of discount rate news is significantly higher in the recent period. We also verified that the results hold for data sampled at even lower, e.g. five-year frequency. Thus, the documented pattern appears to be a deep characteristic of the market that seems to go beyond dividend lumpiness (or smoothing).

4.3 Further evidence

In this section, we provide further evidence for predictability results. First, we analyze how predictability changes with the length of the return and dividend growth horizon. Next, we explore stability of predictability results in a cross-period out-of-sample exercise. Finally, we analyze predictability over the business cycle.¹²

We explore these issues within a reduced-form model, where returns, dividend growth rates, and the dividend-to-price ratio are predicted by the lagged dividend-to-price ratio only. This departure from the full VAR model enables us to conduct the analysis in a parsimonious way while still preserving the present value restrictions. Exact model specifications are reported below. In the Online Appendix (Table OA 1), we show that all the results from the main analysis (Section 4.2) are robust to using this reduced form model.

¹² In Online Appendix B, we also show that the main results are robust to adjusting for inflation.

4.3.1 Longer horizon predictability

Because the dividend-to-price ratio is highly persistent, it commands highly persistent expected returns. As a result, the predictable component of returns generally increases with the return horizon, as documented by Fama and French (1988) and Cochrane (2005), among others. The high persistence of the dividend-to-price ratio, however, also means that estimators are almost perfectly correlated across horizons, casting doubt on the benefits of using longer horizon regressions in small samples where the number of independent observations is limited (Boudoukh, Richardson, and Whitelaw 2008). Relatedly, Ang and Bekaert (2007) argue that the statistical evidence for return predictability is weaker for longer horizons. Because our data spans a much longer time period, and the dividend-price ratio in the historical data is much less persistent, our data is perfectly suited to revisit the issue of longer horizon predictability in more detail.

In the previous section, we document that our results are robust using different sample frequencies (one, three, or five years). The drawback of this approach is that it ignores higher frequency time-series information. Therefore, in this section, we follow the more standard approach, where multi-year returns (and dividend growth rates) are regressed on the lagged annual dividend-to-price ratio.

In particular, for a multi-period horizon H , the present value constraint in Eq. (3) can be rewritten as:

$$dp_t \approx \sum_{h=1}^H \rho^{h-1} r_{t+h} - \sum_{h=1}^H \rho^{h-1} dg_{t+h} + \rho^H dp_{t+H}. \quad (13)$$

This motivates a predictive system, where the discounted sum of returns, the discounted sum of dividend growth rates, and the future value of the annual dividend-to-price ratio are regressed on the lagged annual dividend-to-price ratio (as before, all variables are demeaned):

$$\begin{aligned}
\sum_{h=1}^H \rho^{h-1} r_{t+h} &= \beta^r dp_t + \varepsilon_{t+h}^r, \\
\sum_{h=1}^H \rho^{h-1} dg_{t+h} &= \beta^{dg} dp_t + \varepsilon_{t+h}^{dg}, \\
dp_{t+H} &= \beta^{dp} dp_t + \varepsilon_{t+h}^{dp}.
\end{aligned} \tag{14}$$

According to Eq. (13), the estimated coefficients must satisfy the constraint:

$$\beta^r - \beta^{dg} + \rho^H \beta^{dp} = 1. \tag{15}$$

As before, we estimate this predictive system of equations subject to the linear constraint by iterative GMM and report heteroscedasticity and autocorrelation consistent statistics. As in Boudoukh, Richardson, and Whitelaw (2008), we consider horizons up to five years. All results are based on the full period starting in 1685. We are careful in appending the data to avoid cross-period predictions.

Table 4 reports results for one, three, and five year horizons; Panel A reports results based on overlapping observations; in Panel B, we estimate the model on H different non-overlapping samples and report the mean estimates across the different samples. All the estimated parameters have the theoretically correct signs and are highly statistically significant. Consistent with the notion that all the variation in the dividend-to-price ratio is ultimately driven by the return and dividend growth predictability, we note that the estimated parameters on returns and dividend growth rate generally increase with the horizon, and the estimated parameter in the dividend-to-price regression decreases with the horizon. In the case of overlapping observations, the statistical significance of parameters increases marginally with horizon. In the case of non-overlapping observations, the statistical significance across horizons is stable, with the only exception being the five year horizon, where the estimated parameter for dividend growth rates is now significant at the five percent level, rather than the one percent

level. All in all, results suggest that, over the last four centuries, returns and dividend growth rates are predictable over both annual and multiannual horizons.

4.3.2 Cross-period out-of-sample predictability

In the main analysis, we document that returns are predictable throughout all four centuries, but there is a break in the dividend growth predictability around 1945. To provide additional support for these results, we next consider an out-of-sample analysis, where we use parameters estimated in one period to predict returns and dividend growth rates in other periods. In particular, we use parameters estimated on the data up to 1945 to make predictions for returns and dividend growth rates in the post-1945 period. Then we reverse the exercise and make predictions for returns and dividend growth rates in the earlier part of the sample using the estimated parameters from the post-1945 period. We evaluate the “out-of-sample” performance by an out-of-sample R-squared:

$$R^{OOS} = 1 - \frac{\sum_{t=1}^N (x_{t+1}^{Actual} - x_{t+1}^{Predicted})^2}{\sum_{t=1}^N (x_{t+1}^{Actual})^2}, \quad (16)$$

where x is either return or dividend growth rate. For consistency with the previous analysis, all the variables are demeaned; x^{Actual} denotes the realized period-specific demeaned returns or dividend growth rates and $x^{Predicted}$ are the predicted returns or dividend growth rates.¹³ Positive R^{OOS} implies that the predicted values have useful information about future returns or dividend growth rates, while negative R^{OOS} implies the opposite. Note that R^{OOS} is bounded by one on the

¹³ Note that demeaning introduces a slight look-ahead bias. Demeaning, however, applies to all the variables and hence does not mechanically bias results towards finding evidence for out-of-sample predictability. Note also that our approach differs from Goyal and Welch (2008), who calculate out-of-sample R-squares using a rolling (and expanding) window approach.

positive side, and is unbounded on the negative side. We base our predictions on the reduced-form model, where returns and dividend growth rates are predicted by the dividend-to-price ratio.¹⁴ For a comparison, we also calculate in-sample R-square. In-sample R-square is calculated in the same way as the out-of-sample R-square, with the only difference being that we use parameters estimated over the same period.

Results reported in Table 5 suggest that we can indeed use the estimated parameters based on the pre-1945 data samples to predict returns in the post-1945 period with an out-of-sample R-square of 11.0%. This is impressive, considering that the comparable in-sample R-square for the recent period is only marginally higher, 11.5%. Similarly, we note that, using the post-1945 period estimates, we obtain an out-of-sample R square of 3.1% for predicting returns in the earlier part of the samples. Again, this is very close to the in-sample R-square of 3.4% for that period.

The return predictability results, however, do not carry over to dividend growth predictability. While in-sample R-square in the early data is 20.5%, we cannot use the estimates from this period to predict dividend growth rates in the post-1945 period, as denoted by a large and negative out-of-sample R-square. Similarly, we cannot use the estimated parameters from the recent period to predict dividend growth rate in the early part of the samples. This is not surprising since the estimated parameter on the dividend-to-price ratio is large and negative in the pre-1945 sample, but marginally positive in the recent period.

All in all, we find that returns are predictable across periods, whereas dividend growth rates are not. This confirms the stability of return predictability over the sample as a whole and a break in the dividend growth predictability in the middle of the previous century.

¹⁴ Results, based on the full VAR model from the main analysis, where returns and dividend growth rates are predicted using lagged dividend-to-price ratio, lagged returns, and lagged dividend growth rates, are very similar.

4.3.3 *Predictability over the business cycle*

An extensive literature argues that market risk premia are countercyclical and business cycle variation affects return predictability (Campbell and Cochrane 1999; Menzly, Santos, and Veronesi 2004; Bekaert, Engstrom, and Xing 2009). In line with theory, Henkel, Martin, and Nardari (2011) document that, in the recent U.S. sample (1953-2007), the dividend-to-price ratio is less persistent and predicts returns much better in recession than in expansions. In this section, we explore whether this result holds in the earlier years of our sample as well.

To identify recessions, we use data from a number of different sources. For the US period, we follow Henkel, Martin, and Nardari (2011) and rely on the standard NBER chronology of expansions and contractions. We classify a year as a recession if at least six months in that year are characterized as a downturn. For the period before 1870, we focus on recessions in the UK. We rely on peak and trough dates from Ashton (1959), Gayer, Rostow, and Schwartz (1953), and Rostow (1972). These authors apply the same dating methodology as the NBER (originally developed by Burns and Mitchell 1946). We let recessions start in the year an economic trough occurs, and we let it end the year before an economic peak is reached. Note that these dates are only available from 1700 onwards.

To identify the effect of business cycle variation, we proceed as follows. For each period, we first separate expansionary and recessionary years. In particular, if the dividend-to-price ratio in year t coincides with a recession in year t , we call it a recessionary year, or else it qualifies as an expansionary year. As before, we use the dividend-to-price ratio at time t to forecast next year's returns and dividend growth rates, regardless whether the associated next year returns and dividend growth rates are in recessions or expansions. For each sample separately, we then

calculate the linearization constant and demean all variables. Finally, we estimate all the parameters for expansions *EXP* and recessions *REC* jointly in the following system:

$$\begin{aligned}
\begin{pmatrix} r_{t+1}^{EXP} \\ r_{t+1}^{REC} \end{pmatrix} &= \begin{pmatrix} dp_t^{EXP} & 0 \\ 0 & dp_t^{REC} \end{pmatrix} \begin{pmatrix} \beta_r^{EXP} \\ \beta_r^{REC} \end{pmatrix} + \begin{pmatrix} \varepsilon_{r,t+1}^{EXP} \\ \varepsilon_{r,t+1}^{REC} \end{pmatrix}, \\
\begin{pmatrix} dg_{t+1}^{EXP} \\ dg_{t+1}^{REC} \end{pmatrix} &= \begin{pmatrix} dp_t^{EXP} & 0 \\ 0 & dp_t^{REC} \end{pmatrix} \begin{pmatrix} \beta_{dg}^{EXP} \\ \beta_{dg}^{REC} \end{pmatrix} + \begin{pmatrix} \varepsilon_{dg,t+1}^{EXP} \\ \varepsilon_{dg,t+1}^{REC} \end{pmatrix}, \\
\begin{pmatrix} dp_{t+1}^{EXP} \\ dp_{t+1}^{REC} \end{pmatrix} &= \begin{pmatrix} dp_t^{EXP} & 0 \\ 0 & dp_t^{REC} \end{pmatrix} \begin{pmatrix} \beta_{dp}^{EXP} \\ \beta_{dp}^{REC} \end{pmatrix} + \begin{pmatrix} \varepsilon_{dp,t+1}^{EXP} \\ \varepsilon_{dp,t+1}^{REC} \end{pmatrix},
\end{aligned} \tag{17}$$

subject to the constraints imposed by the present value model:

$$\begin{aligned}
\beta_r^{EXP} - \beta_{dg}^{EXP} + \rho^{EXP} \beta_{dp}^{EXP} &= 1, \\
\beta_r^{REC} - \beta_{dg}^{REC} + \rho^{REC} \beta_{dp}^{REC} &= 1.
\end{aligned} \tag{18}$$

In total, we have six moment conditions and two linear constraints. As before, we estimate the system of equations using the iterative generalize method of moments and report heteroscedasticity and autocorrelation corrected standard errors. We use a Wald statistic to test whether estimated parameters for recessions are different from those in expansions.

Table 6 reports results for three periods: the early period (1700-1945), the recent U.S. period (1945-2012), and the full period (1700-2012). It is important to note that in the early period 46% of the years were in recessions, whereas only 18% of the years are classified as recessions in the recent period. Panel A reports summary statistics of all variables in years with and without recessions. As expected, realized dividend growth rates and returns are lower and more volatile in recessions than in expansions. At the same time, the dividend-to-price ratio is somewhat higher in recessions, consistent with an increased risk premium in crisis periods. This holds for all three sample periods, although the differences are not always statistically significant.

The return predictability results are in line with Henkel, Martin, and Nardari (2011) across both periods. We first note that the dividend-to-price ratio is always less persistent in recessions than in expansions. This difference is significant at the ten percent level in the pre-1945 period as well as in the full period, but insignificant in the post-1945 period. Next we observe that return predictability is consistently stronger during recession periods. In the earlier period, returns are predictable during both expansions and recessions, but the estimated parameter is much higher in recession periods; the difference is statistically significant at the 10% level. A similar picture emerges in the recent US period; if anything, the difference in return predictability between expansions and recessions increases after 1945. In the full period, returns are only significantly predictable for recessions. The difference with expansions is not statistically significant however.¹⁵

Dividend growth rates also tend to be more predictable during recessions than during expansions, as indicated by the more negative coefficients. However, the difference between expansions and recessions is economically less pronounced, especially for 1700-1945. Differences between coefficients are never statistically significant. In the recent US period, dividend growth is not predictable during expansions and is marginally predictable (but with the wrong sign) during recessions.

All in all, our results confirm that business cycle variation has an important impact on the variation in the dividend-to-price ratio. The 1700-1945 results indicate that especially returns become more predictable during recessions. This suggests that recessions have a more important impact on discount rates than on dividend growth rates.

¹⁵ The regression results for the full period should be interpreted with caution because the earlier part of the sample features relatively more recession years than the recent period. This means that the recession (expansion) estimates for the full period oversample (undersample) pre-1945 observations.

5. Reconciling the evidence

Our analysis reveals that the recent U.S. market is both very similar and very different from the earlier period. Average returns are very comparable throughout all four centuries, and returns are predictable across all the periods. Furthermore, return predictability is concentrated in recessions and is remarkably stable, despite the many institutional changes that the markets have undergone throughout the last four centuries.

At the same time, the recent U.S. period is substantially different from earlier years. Three observations stand out. First, the dividend-to-price ratio has decreased considerably. Second, investors are now receiving most of their returns through capital appreciation. Third, whereas cash-flows seem to be more important for price movements in the earlier period, discount rates appear to be the sole determinant of price movements in the recent years.

In this section, we delve deeper in the driving forces of these differences. We first discuss potential reasons for the decrease in dividends in the recent period. Then, we hypothesize that the documented differences between the recent U.S. period and the earlier years are driven by the increased duration of the stock market. Finally, we discuss the empirical support for our hypothesis and relate it to the alternative explanations put forward in the literature.

5.1 Why did the dividend-to-price ratio decrease in the recent period?

The fall in the dividend-to-price ratio seems to be largely the result of the fact that firms pay out less of their dividends and retain a larger share of their earnings. Figure 3 documents the dividend-to-earnings ratio for the first part of our sample until 1809 and for the US since 1871

(data for the intermediate period are unavailable).¹⁶ In the early years of modern financial markets, companies were paying out most of their earnings to investors with the payout ratio fluctuating around one for most of the 17th and 18th century, and decreasing to about 0.9 towards the end of the 18th century.¹⁷ Companies in the early US period started to retain some of their earnings, but a dramatic reduction in aggregate dividends only happened after the middle of the 20th century. Throughout the US period, the payout ratio decreased from around 0.8 at the end of the 19th century to close to 0.4 nowadays.

Fama and French (2001) document that the propensity to pay dividends decreased for both existing and new companies. This decrease in aggregate dividends is therefore not simply due to a shift in the sector decomposition, but rather due to changes in dividend policy. One reason for the changing dividend policy may be related to taxes. For most of the 17th, 18th, and 19th century, income taxes were negligible, and taxes did not play a role in dividend decisions. The US introduced an income tax in 1913. Initially, dividend income and capital gains were taxed at the same rate, making dividend policy insensitive to taxes. From 1922 on, however, capital gains were taxed at a lower rate. This continued for most of the recent U.S. until the Bush tax cuts in 2003, when taxes on capital gains and dividends were largely equalized (Auten 2005). The extra tax burden on dividends in the intervening years may have incentivized companies to decrease dividend payments and reinvest free cash-flows.

Furthermore, high taxes on dividends, combined with the increased practice of stock option compensation, may have incentivized companies to engage in stock buybacks while

¹⁶ For the period before 1810, earnings data are only available for the VOC, EIC and BoE and the payout ratio is calculated based on these three companies alone. On average, these companies make up 89% of the total market capitalization between 1685 and 1809. For the US period (1871-2012), the data are from Rober Shiller's webpage.

¹⁷ There is a big spike in the payout ratio between 1782 and 1794. This is due to the fact that the Dutch East India Company consistently had negative earnings for those years. After the Fourth Anglo Dutch War, the Company's position in the trade with Asia deteriorated significantly, leading to nationalization in 1795. Omitting the VOC after 1781, we note a smooth decline in dividends-to-earnings after 1750.

simultaneously decreasing dividends. Starting in 1980s, when the difference in taxes between dividends and capital gains was especially pronounced, we observe a sharp increase in stock repurchases (Boudoukh, Michaely, Richardson, and Roberts 2007). Although stock repurchases do not occur on a regular basis, and they are often done for reasons beyond the distribution of cash flows, they are in many ways a substitute for dividends.¹⁸ It is important to note, however, that the presence of repurchases does not invalidate our predictability and decomposition results.¹⁹

A different question is whether repurchases (and issuances) can account for the decrease in the dividend-to-price ratio. Using data for CRSP index from Larrain and Yogo (2008), we note that dividend payout including repurchases (defined as dividends plus equity repurchases as a fraction of market equity) has also decreased in the recent period. On average it was 5.46% between 1927 and 1945, dropping to 4.16% between 1946 and 2004. In comparison, the average CRSP dividend payout was 5.13% between 1927 and 1945 and 3.59% between 1946 and 2004. Thus, repurchases do not seem to be able to fully account for the decrease in dividends in the post-1945 period. Moreover, if, following Larrain and Yogo (2008), we also account for issuances, the decrease is even more pronounced; the equity payout (dividends plus repurchases minus issuances) decreased from 3.98% on average between 1927 and 1945 to an average of 1.66% between 1946 and 2004.

¹⁸ For example, Hong and Wang (2008) provide evidence consistent with companies engaging in stock buybacks to provide liquidity in times of distress.

¹⁹ In our empirical analysis, we follow the standard approach of Campbell and Shiller (1988) and rely on the individual investor, who is interested in the per share value of a company and receives dividends as the only source of cash-flows. Repurchases are assumed to be reinvested in the company. An alternative approach would be to take the perspective of a representative investor who is interested in the value of the whole company. Under this alternative one needs to account for both repurchases and issuances of equity (and debt). Larrain and Yogo (2008) show that this reinforces the importance of cash-flows. Unfortunately, the historical data on repurchases and issuances are scarce, which renders this approach unfeasible in our setting.

Another reason for the decrease in dividend payments is related to the observation that companies have become increasingly reluctant to change dividend payments. As companies smooth dividends, they may set dividends to a lower level in order to be able to meet dividend expectations each and every period.²⁰ Yet another reason for a decrease in dividends may be improved transparency due to requirements of the stock exchanges for more detailed and frequent reporting. This may have decreased the role of dividends as signals. Relatedly, investors' preference for dividends ("bird in the hand effect") may have decreased, with companies optimally responding by lowering dividend payments.

5.2 Increased duration of the stock market

The reduction in dividend payouts suggests that companies reinvest a larger share of their earnings and postpone distribution of cash flows (dividend payments) into the future. We hypothesize that this has led to the increased duration of the stock market as a whole, which may in turn explain the dominance of the discount rate news in the recent period.

Borrowing the intuition from bond pricing, we think of stock duration in terms of the timing of dividend payments that represent the actual stream of cash-flows to an individual investor. When companies reinvest more, current dividends decrease and future cash flow distributions increase. As a result, the dividend-to-price ratio decreases, capital appreciation becomes more important for total returns, and stock duration increases. Following the intuition from bond pricing, the increased stock duration can raise the importance of discount rates. There is, however, an important difference between fixed income and the stock market. In the case of bonds, coupons are fixed, and bond duration therefore maps directly into interest rate sensitivity. In the stock market, however, dividends change over time. This means that increased duration

²⁰ Note that, as discussed further in Section 5.3, dividends were relatively smooth in the 18th century as well.

could affect not only the sensitivity of prices to discount rate news, but also the sensitivity of prices to cash flow news.

To verify that the increased importance of discount rates in the recent period is related to stock duration, we calculate a simple proxy for duration and explore how it correlates with the importance of discount rate news. In particular, for each year, we calculate the fraction of the price that is accounted for by the present value of the next ten years of expected dividends. We take one minus this fraction as a proxy for duration, dur_t . For simplicity, we assume that dividend growth rates follow a random walk with drift μ . Our duration measure is defined as:

$$dur_t = 1 - \frac{D_t}{P_t} \sum_{n=1}^{10} \left(\frac{1+\mu}{1+r} \right)^n. \quad (19)$$

As a proxy for μ , we use the average per-period dividend growth rate. For the discount rate r , we use the average per-period market return. Note that our measure for duration can be interpreted as the dividend-to-price ratio adjusted for the drift μ and the discount rate r .

We first calculate duration separately for each of the four periods within our sample. The average values for stock duration are 0.63 for Netherlands/U.K., 0.62 for U.K., 0.58 for the early U.S. period, and 0.71 for the recent U.S. period. Thus, as expected, stock duration increased in the recent years. The difference in duration between the earlier and the post-1945 period is statistically significant with a t -statistic of 2.16.²¹ We also find that duration is highly correlated with each period's discount rate news component reported in Table 2. Depending on whether we look at the decomposition of the dividend-to-price ratio or unexpected returns, we find a correlation of 0.87 or 0.93 (p-values of 0.13 and 0.08).

²¹ We test for the difference in means using unpaired two sample t -test with unequal variances. To account for overlapping observations, we report the mean t -statistic estimated on ten consecutive non-overlapping samples.

To better account for time-variation and to analyze the within-period relationship between duration and discount rates, we next use a rolling windows approach. We set the length of the window to 68 years to match the post-1945 period. In each window, we estimate the importance of discount rate news and our measure for duration (duration is the average duration within each window; the corresponding values for drift and discount rates are estimated within the same window). Because the 19th century U.K. period is too short, we rely only on the Dutch/English period (1685-1809) and the combined U.S. periods (1871-2012). Figure 4 plots the time series of duration along with either the discount rate component of the dividend-to-price ratio (Panel A) or the discount rate component of unexpected returns (Panel B).

In both periods, we see clearly that duration and the discount rate component are increasing and are positively correlated. In a pooled regression of discount rate news on a constant and duration, we get R-squares of 28% (based on Panel A) or 59% (based on Panel B). The relationship is preserved in both subsamples. The respective R-squares are 52% and 49% in the Netherlands/U.K. period and as high as 74% and 81% in the U.S. period. Thus, duration and discount rate news are strongly correlated across periods as well as within periods. Note also that the increasing duration in the Netherlands/U.K. period and especially in the U.S. period is consistent with the reduction in the dividend-to-earnings ratio documented in Figure 3.

Finally, if duration of the U.S. stock market has increased over time, we would expect that the stock market's comovement with (other) long duration assets has increased. Following Cornell (1999) we verify this by calculating market betas for medium and long term Treasuries. We use data from Ibbotson SBBI (2013) for the period between 1926 and 2012. To correct for potential bias coming from an increase in the variance of bond and stock returns after 1945, we scale all returns by their pre-1946 standard deviations (Dungey et al. 2005; Forbes and Rigobon

2002). Estimates are reported in Table 7. As expected, we find that betas increased over time, especially for long term treasuries. While betas for medium and long term bonds are 0.12 in the 1926-1945 period, they are 0.30 and 0.48 after 1945 (adding up the level and interaction coefficients). Thus, consistent with our duration hypothesis, the increase in beta is larger for long- than for medium- term bonds (4.00 vs 2.59 times). However, this difference is not statistically significant.

If we exclude the years after 2007, which are heavily influenced by the impact of QE on long term bond yields, the increase in betas is more pronounced. In this case, betas increase to 0.44 and 0.73, and the increase in beta is significant for long-term bonds. Again, the increase in beta is larger for long- than for medium- term bonds (6.12 vs 3.84 times). For the sample up to 2007, this difference is highly statistically significant.

5.3 Discussion

We find strong support for our duration hypothesis. At the same time, however, we note that the importance of discount rate news has increased in the post-war period by much more than our measure of duration (Figure 4). Also, the increase in the correlation with longer maturity assets is not nearly as strong as the increase in the amount of discount rate news. Thus, the increase in stock duration may not account for the whole increase in the discount rate news. Other explanations put forward in the literature may also play a role.

Chen, Da, and Priestly (2012) show that dividend smoothing can mask the predictability of dividend growth and overemphasize the importance of discount rate news. Dividend smoothing measures the relation between volatility of dividends and earnings. For the U.S. period, using data from Robert Shiller's webpage, we find that, consistent with Chen, Da, and

Priestly (2012), U.S. companies more actively smoothed their dividends after 1945. The ratio of the standard deviations of earnings and dividend log-growth rates (the so-called dividend smoothing parameter) is 0.55 in the pre-1945 period and 0.22 after 1945. In other words, compared to earnings, dividends were far less volatile in the recent period. However, based on our earnings data for the period 1685 – 1809, we note that the dividend smoothing parameter was 0.22 as well.²² This indicates that dividend smoothing is not unique to the recent period.

To probe further, we run a horse race between dividend smoothing and an increase in stock duration for the U.S. period. Using a rolling window approach from the previous section, we measure dividend smoothing over the same 68 years rolling windows that we use to estimate the discount rate news components. Figure 5 plots the discount rate news component along with our measure of duration and the smoothing parameter. The dividend smoothing parameter varies over time much more than our stock duration measure, and decreases substantially at approximately the same time when the discount rates news increase, confirming that dividend smoothing may be an important factor driving discount rate news.

To analyze the association between dividend smoothing and discount rate news more formally, we next run a regression of discount rate news component on the smoothing parameter. The estimated parameter is negative, as expected, and the adjusted R-squared is 16% or 25%, depending on whether we measure the discount rate news component of the dividend-to-price ratio or unexpected returns. This compares to the adjusted R-square of 74% and 81% for our measure of duration. When we include both explanatory variables, the adjusted R-square is 76% and 86%. Thus, while both explanations play a role, the marginal contribution of duration in

²² We omit years with negative earnings (12% of the years) in the calculation of the smoothing parameter, inducing a positive bias on the smoothing parameter. For the U.S. period, earnings are always positive.

explaining the discount rate news variation largely outweighs the importance of dividend smoothing.

All in all, the increased importance of discount rate news is likely driven by many different factors. We find strong support for our hypothesis that increased stock duration is responsible for much of the increase in the discount rate news. Other explanations, especially dividend smoothing, however, may also play a role.²³

6. Conclusions

We analyze return predictability and excess volatility in the most important financial markets of the last four centuries. In particular, we analyze the Dutch and English stock markets in the 17th and 18th century, the U.K. stock market in the 18th and 19th century, and the U.S. stock market from the end of the 19th century onwards.

We find that the dividend-to-price ratio is stationary across all periods and predicts both returns and dividend growth rates. There are, however, important differences between the different periods. Whereas returns appear to be always predictable, dividend growth predictability disappeared after 1945. This suggests that cash-flow news used to be much more important for price movements, and the dominance of the discount rate news is a rather recent phenomenon. We argue that this is consistent with increased duration of the stock market in recent years.

Our findings have important implications for the theoretical asset pricing literature. They imply that the role of fundamentals and the relative importance of “excess volatility” depend on

²³ The existing literature also analyzes other statistical properties of the dividend-to-price ratio. For example, Lettau and Van Nieuwerburgh (2008) suggest adjusting the dividend-to-price ratio in the post-war period for one (or two) structural breaks. The dominance of discount rate news, however, still prevails (Kojien and Van Nieuwerburgh 2011).

the timing of dividends (stock market duration). Hence, time-varying expected returns are best analyzed alongside firms' investment and payout decisions.

References

Acheson, G., C. R. Hickson, J. D. Turner and Q. Ye. 2009. Rule Britannia! British stock market returns, 1825-1870. *Journal of Economic History* 69:1107–1137.

Albuquerque, R., M. Eichenbaum, and S. Rebelo. 2014. Valuation risk and asset pricing. Working paper. Boston University and Northwestern University.

Ang, A., and G. Bekaert. 2007. Stock return predictability: Is it there? *Review of Financial Studies* 20:651–707.

Ashton, T.S. 1959, *Economic Fluctuations in England*, Oxford: Clarendon Press.

Auten, Gerald, 2005, Capital Gains Taxation, in: Joseph J. Cordes, Robert D. Ebel, and Jane G. Gravelle (Eds.), *NTA Encyclopedia of Taxation and Tax Policy*, Washington DC: Urban Institute Press.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481–1509.

Bekaert, G., E. Engstrom, and Y. Xing. 2009. Risk, uncertainty, and asset prices. *Journal of Financial Economics* 91:59–82.

Binsbergen, J. H. van, and R. S. J. Koijen. 2010. Predictive regressions: A present-value approach. *Journal of Finance* 65:1439–71.

Boudoukh, J., R. Michaely, M. Richardson, and M. R. Roberts. 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance* 62:877–915.

Boudoukh, J., M. Richardson, and R. F. Whitelaw. 2008. The myth of long-horizon predictability. *Review of Financial Studies* 21:1577–1605.

Bowen, H.V. 1989. Investment and empire in the later 18th century: East India Company stockholding, 1756-1791. *Economic History Review* 42:186-206

Bowen. H.V. 2006. *The business of empire: the East India Company and imperial Britain, 1756-1833*. Cambridge: Cambridge University Press.

Brisle, D., W. N. Goetzmann, and S. Pouget. 2014. Testing asset pricing theory on six hundred years of stock returns: Prices and dividends for the Bazacle company from 1372 to 1946. Working paper. Yale School of Management.

Burns, Arthur F. and Wesley C. Mitchell, 1946, *Measuring Business Cycles*, NBER Studies in Business Cycles, 1946.

Campbell, J. Y. 1991. A variance decomposition for stock returns. *Economic Journal* 101:157–179.

Campbell, J. Y., and J. H. Cochrane. 1999. By force of habit: A consumption-based explanation of stock market behavior. *Journal of Political Economy* 107:205–51.

Campbell, J. Y., and R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195–227.

Chaudhuri, K.N., 1978. *The Trading World of Asia and the English East India Company, 1660-1760*. Cambridge: Cambridge University Press.

Chen, L. 2009. On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics* 92:128–51.

Chen, L., Z. Da, and R. Priestly. 2012. Dividend smoothing and predictability. *Management Science* 58:1834–1853.

Clark, Gregory, 2015, *What Were the British Earnings and Prices Then? (New Series), MeasuringWorth*.

City Archives Amsterdam, 30452, Archief van S. Hart: (gedeeltelijke) toegang op de notariële archieven.

City Archives Amsterdam, 5077, Archief van de Wisselbank, 1609-1820.

Clapham, John, 1945. *The Bank of England, A History*, Cambridge: Cambridge University Press.

Cochrane, J. H. 1992. Explaining the variance of price-dividend ratios. *Review of Financial Studies* 5:243–80.

—. 2005. *Asset Pricing*. Princeton, NJ: Princeton University Press.

—. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21:1533–75.

—. 2011. Presidential address: Discount rates. *Journal of Finance* 66:1047–1108.

Cornell, B., 1999. Risk, duration, and capital budgeting: New evidence on some old questions. *Journal of Business* 72:183–200.

Cowles, A. 1939. *Common Stock Indices, 1871-1937*, Cowles commission for research in economics, Monograph no. 3., Principia Press: Bloomington, Ind.

De Korte, J.P. 1984. *De jaarlijkse financiële verantwoording in de VOC, Verenigde Oostindische Compagnie*, Den Haag: M. Nijhoff.

Dungey, M., R. Fry, B. Gonzalez-Hermosillo, and V.L. Martin. 2005. Empirical modelling of contagion: a review of methodologies. *Quantitative Finance* 5:9–24.

Dutch National Archives, 1.04.02, Archief van de Verenigde Oost-Indische Compagnie (VOC), 1602-1795 (1811).

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

Fama, E. F., and K. R. French. 2001. Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics* 60:3–43.

Forbes, K.J., and R. Rigobon. 2002. No contagion, only interdependence: measuring stock market movements. *Journal of Finance* 57:2223–2261.

Frehen, R. G. P., W. N. Goetzmann, and K. G. Rouwenhorst. 2013. New evidence on the first financial bubble. *Journal of Financial Economics* 108:585–607.

Gayer, A.D., Rostow, W.W. and Schwartz, A.J., 1953, *The Growth and Fluctuation of the British Economy, 1790-1850*, Volume 1, Oxford: Oxford University Press.

Gelderblom, O., A. De Jong and J. Jonker. 2013. The Formative Years of the Modern Corporation: The Dutch East India Company VOC, 1602–1623. *Journal of Economic History* 73:1050–1076.

Goetzmann, W. N., R. G. Ibbotson and L. Peng. 2001. A new historical database for the NYSE 1815 to 1925: Performance and predictability. *Journal of Financial Markets* 4:1–32.

Golez, B. 2014. Expected returns and dividend growth rates implied by derivative markets. *Review of Financial Studies* 27:790–822.

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

Harrison, P. 1998. Similarities in the distribution of stock market price changes between the eighteenth and twentieth Centuries. *The Journal of Business* 71: 55–79.

Henkel, S. J., J. S. Martin, and F. Nardari. 2011. Time-varying short horizon predictability. *Journal of Financial Economics* 99: 560–580.

Hickson, C. R., J. D. Turner and Q. Ye. 2011. The rate of return on equity across industrial sectors on the British stock market, 1825-1870. *Economic History Review* 64:1218–1241.

Hong, H., J. Wang, J. Yu. 2008. Firms as buyers of last resort. *Journal of Financial Economics* 88:119–145.

Ibbotson SBBI 2013 Classic Yearbook, Market results for stocks, bonds, bills, and inflation, 1926-2012. Morningstar Inc.

Klerk de Reus, G.C., 1894, *Geschichtlicher Überblick der administrativen, rechtlichen und finanziellen Entwicklung der Niederländischen-Ostindischen Compagnie*, The Hague: M. Nijhoff.

Koijen, R. S. J., and S. Van Nieuwerburgh. 2011. Predictability of returns and cash flows. *Annual Review of Financial Economics* 3:467–91.

Koudijs, P. 2014. The boats that did not sail: Asset price volatility in a natural experiment. *Journal of Finance*, forthcoming.

Larrain, B., and M. Yogo. 2008. Does firm value move too much to be justified by subsequent changes in cash flow? *Journal of Financial Economics* 87:200–226.

LeRoy, S. F., and R. D. Porter. 1981. The present value relation: Test based on implied variance bounds. *Econometrica* 49:555–574.

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

Lettau, M., and S. Van Nieuwerburgh. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* 21:1607–52.

Luzac, Elias, 1780, *Hollands Rijkdom*, Leiden: Luzac en van Damme.

Menzly, L., T. Santos, and P. Veronesi. 2004. Understanding predictability. *Journal of Political Economy* 112:1–47.

NBER, 2010, *US Business Cycle Expansions and Contractions*.

Neal, L. 1990. *The rise of financial capitalism*, Cambridge: Cambridge UP.

Petram, Lodewijk, 2011, *The World's First Stock Exchange: How the Amsterdam market for Dutch East India Company shares became a modern securities market, 1602-1700*, Dissertation.

Posthumus, N.W., 1946, *Inquiry into the History of Prices in Holland, Volume 1*, Leiden: E.J. Brill.

Quinn, Stephen, and William Roberds, 2014, 'How Amsterdam got fiat money', *Journal of Monetary Economics* 66: 1-12.

Rajan, R., G., and L. Zingales. 2003. The great reversals: the politics of financial development in the twentieth century. *Journal of Financial Economics* 69:5–50.

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

Report from the Committee of Secrecy on the Bank of England Charter. 1832. London: House of Commons.

Rostow, W.W., 1972, *Cycles in the British Economy: 1790-1914*, in Aldcroft, D.H. and Fearon, P. (eds.), *British Economic Fluctuations 1790-1913*, Basingstoke: Macmillan, 74-96.

Scott, William Robert, 1912, *The Constitution and Finance of English, Scottish and Irish Joint-stock Companies to 1720*, Cambridge: Cambridge University Press.

Shiller, R. J. 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71:421–436.

Stambaugh, R. F. 1999. Predictive regressions. *Journal of Financial Economics* 54:375–421.

Shea, G. Handbook of 18th and early 19th century British corporate finance. In preparation.

Thorold Rogers, James. E., 1902, *A History of Agriculture and Prices in England*, Volumes 6 and 7-2, Oxford: Clarendon Press.

Van Dillen, J.G., 1931, *Effectenkoersen aan de Amsterdamsche Beurs 1723-1794*, *Economisch Historisch Jaarboek* 17: 1-34.

Van Zanden, J. L., and B. Van Leeuwen 2012. Persistent but not consistent: The growth of national income in Holland 1347–1807. *Explorations in Economic History* 49:119-130

Wright, J.F. 1997. The contribution of overseas savings to the funded national debt of Great Britain, 1750-1815. *Economic History Review* 50:657-674

Appendix A: Data sources: Prices, dividends and earnings, 1629-1809

The data for the period 1629-1809 cover two Dutch securities and three English securities. For most of the period, all securities were cross-listed on the Amsterdam and London stock exchanges. We convert all price and dividend data to Dutch guilders. Below we describe in detail the wide variety of sources used to reconstruct end-of-year equity prices, annual dividend payments and earnings.

A.1 Amsterdam 1629-1809

There were two widely traded Dutch equities in the 17th and 18th centuries: the Dutch East India (VOC) and West Indies Company (WIC). The VOC was founded in 1602 and the WIC in 1623. Initially, Amsterdam stock prices were not published in the newspapers of the time, and we need to rely on other sources to reconstruct prices.

For the WIC no continuous end-of-year price series is available until 1719. Since the WIC was a relatively small company (in 1719 it only makes up 1.1% of our overall index), its omission for the earlier period has little effect on our estimates.

For the VOC, we can reconstruct a continuous series of end-of-year prices going back to 1629. To do so, we combine notary records and the VOC's dividend ledgers. The Amsterdam Notary records of the 17th and 18th centuries often contain information about share transactions. Sometimes a buyer or seller wanted to notarize a transaction; sometimes a conflict about a transaction arose, and the notary document details the disputed transaction. Van Dillen (1931) and Petram (2011) provide two (largely) independent sets of share prices extracted from these Notary records. In addition, the Amsterdam City Archives provide an (incomplete) index to the notary records that also contains price observations (City Archives Amsterdam 30452).

The price series that can be reconstructed from the Notary records has a significant number of gaps. To fill these, we use the VOC's dividend ledgers. In the 17th and 18th centuries there were no share certificates yet. Share ownership simply constituted an entry in the Company's dividend books. Whenever the owner of a share changed, there would be a mutation in the dividend ledgers, so that the Company could keep track of to whom they owed dividends. Unfortunately, these share mutations do not list the price. To get this information, we compare the mutations in the dividends books of the VOC chamber in Amsterdam (Dutch National

Archives, 1.04.02) with account transfers in the Amsterdam Bank of Exchange (City Archives Amsterdam, 5077). During the 17th and 18th centuries, all important economic agents had the equivalent of a current account at this bank (Quinn and Roberds 2014). Starting in 1653, many of the Bank's ledgers still exist, and we can reconstruct individuals' bank transfers, including those associated with the transfers of shares. By comparing the size of the share transaction (listed in the VOC's ledgers) and the amount of the bank transfer, we can infer the share price.

In constructing the annual price series, we picked the last available share price of each year. The price series for the VOC goes back to its inception (1602), but a continuous series of end-of-year prices is only available from 1629 onwards.

Starting in 1719 we can rely on the newspapers of the time to get prices for both the VOC and the WIC. For 1719-1722 (following Frehen, Goetzmann and Rouwenhorst 2013), we use information from the *Leydse Courant*. Starting in 1723, we rely on Van Dillen (1931) who lists price information from the *Amsterdamsche Courant*. This coverage continues until 1791 and 1795, the respective years that the WIC and VOC were nationalized by the Dutch government.

Dividends are available for the entire period and come from two sources. For the VOC we rely on the work of Klerk de Reus (1894) who provides information on the exact dates dividends were payable to investors; for the WIC we use Luzac (1780, the year the WIC stopped paying dividends). We are only able to reconstruct earnings for the VOC. Data come from De Korte (1984) and start in 1651.

A.2 London 1691-1809

Starting in 1698, Neal (1990) provides detailed price data for the three English securities in our sample: the British East India Company (EIC), Bank of England (BoE), and South Sea Company (SSC). These prices originate from the Course of the Exchange. For the earlier years, we rely on the work by Thorold Rogers (1902), who reports prices from a series of English newspapers. Prices start in 1691 for the EIC, 1694 for the BoE, and 1711 for the SSC. For the EIC we take prices for the Old EIC until its merger with the New EIC in 1708, from which moment on we use prices for the newly formed United EIC. The EIC was operating before 1691, but there are no frequent price observations available for this period (see Scott 1912, p. 178-9).

Information on dividends was kindly provided by Gary Shea (*in preparation*). The English companies have a complicated history of capital calls, rights issues, and other “capital events.” We closely follow Shea’s work in adjusting stock prices where necessary. We were only able to reconstruct earnings data for the EIC and BoE. Data for the EIC come from Chaudhuri (1978) for 1710 – 1745 and Bowen (2006) for 1757 – 1809. For the BoE we obtain data from Clapham (1945) for 1721 – 1797 and *Report on the Bank Charter* (1832) for 1798 – 1809.

We use exchange rate information listed in Posthumus (1946) to convert Pound Sterling into Dutch guilders.

Table 1: Summary statistics

This table reports summary statistics for the main variables, starting with the 1685-1809 period based on the Netherlands and the U.K. data in column (1) (because dividends were not paid out every year before 1685, column (1a) reports separately the main return statistics for 1629-1809); column (2) reports the same statistics for the U.K. period 1825-1879; column (3) and (4) collect the statistics for the U.S. data before and after 1945; column (5) reports the statistics based on the full sample. Capital appreciation is denoted by *ca*. Sharpe ratio is calculated assuming zero variation in the risk-free rate. The Augmented Dickey-Fuller (ADF) tests for the presence of a unit root. Statistical significance of the ADF at the one, five, and ten percent is denoted by three, two, and one asterisks.

	(1)	(1a)	(2)	(3)	(4)	(5)
	Neth./U.K. 1685-1809	Neth./U.K. 1629-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2012	Full period
<i>r</i>	0.051	0.056	0.069	0.069	0.100	0.069
<i>Cap. app. (ca)</i>	0.008	0.011	0.026	0.018	0.066	0.026
<i>ca / r</i>	0.164	0.194	0.373	0.253	0.661	0.372
<i>AR1(r)</i>	-0.076	-0.090	0.043	0.053	-0.031	0.015
<i>rf</i>	0.032	-	0.033	0.026	0.042	0.033
<i>Risk premium</i>	0.019	-	0.036	0.043	0.058	0.036
<i>Std.(r)</i>	0.089	-	0.069	0.190	0.162	0.135
<i>Sharpe ratio</i>	0.213	-	0.529	0.228	0.356	0.263
<i>dg</i>	0.001	-	0.034	0.013	0.058	0.021
<i>Std.(dg)</i>	0.201	-	0.100	0.157	0.066	0.157
<i>AR1(dg)</i>	-0.223	-	0.182	0.204	0.389	-0.056
<i>dp</i>	-3.163	-	-3.126	-2.964	-3.458	-3.174
<i>Std.(dp)</i>	0.242	-	0.167	0.253	0.440	0.335
<i>AR1(dp)</i>	0.654	-	0.659	0.512	0.910	0.816
<i>ADF(dp)</i>	-4.096***	-	-4.638***	-4.765***	-2.098	-4.709***

Table 2: Vector autoregression (VAR) estimates: Annual data

Panel A reports VAR estimates of returns, dividend growth rates, and the dividend-to-price ratio. The data is annual. The model is estimated by iterative generalized method of moments subject to the present value model constraints. Heteroskedasticity and autocorrelation corrected standard errors based on Bartlett kernel are reported in parentheses below the estimated parameters. The Newey and West method is used for the selection of the optimal bandwidth. In brackets we report a Chi-square statistic on the difference of a coefficient between a specific sub-period and the rest of the sample. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. Panel B reports decomposition results based on the VAR estimates from Panel A.

	(1)	(2)	(3)	(4)	(5)
	Neth./U.K. 1685-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2012	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.112*** (0.038) [3.311*]	0.075 (0.046) [0.512]	0.117 (0.104) [1.433]	0.123*** (0.036) [0.677]	0.081*** (0.029)
r_t	-0.055 (0.078)	0.071 (0.128)	0.144 (0.208)	-0.038 (0.102)	0.051 (0.079)
dg_t	0.030 (0.077)	-0.099 (0.069)	-0.270** (0.117)	-0.095 (0.221)	-0.044 (0.056)
Dep. variable: dg_{t+1}					
dp_t	-0.187** (0.086) [5.374**]	-0.342*** (0.094) [2.126]	-0.225*** (0.059) [0.213]	0.000 (0.018) [13.653***]	-0.123*** (0.030)
r_t	0.019 (0.108)	-0.122 (0.261)	0.420*** (0.075)	0.119 (0.081)	0.279*** (0.056)
dg_t	-0.035 (0.113)	0.255*** (0.091)	0.135* (0.075)	0.415*** (0.084)	0.088 (0.099)
Dep. variable: dp_{t+1}					
dp_t	0.730*** (0.098) [10.79***]	0.609*** (0.128) [2.229]	0.692*** (0.102) [2.089]	0.904*** (0.051) [9.564***]	0.830*** (0.043)
r_t	0.077 (0.152)	-0.202 (0.358)	0.290 (0.227)	0.161 (0.158)	0.238** (0.096)
dg_t	-0.067 (0.123)	0.370*** (0.116)	0.425*** (0.150)	0.526** (0.248)	0.137 (0.120)
J-test	2.674	0.399	1.806	1.391	8.702**
Panel B: Decomposition results					
Dividend-to-price					
DR	0.364	0.204	0.395	1.272	0.438
CF	0.636	0.796	0.605	-0.272	0.562
Unexpected returns					
DR	0.288	0.210	0.243	1.002	0.266
CF	0.713	0.791	0.762	-0.003	0.735

Table 3: Vector autoregression (VAR) estimates: Triennial data

Panel A reports VAR estimates of returns, dividend growth rates, and the dividend-to-price ratio. The data is triennial. Reported are the means of the estimated parameters and statistics estimated on three consecutive non-overlapping samples. The VAR model is estimated by iterative generalized method of moments subject to the present value model constraints. Heteroskedasticity and autocorrelation corrected standard errors based on Bartlett kernel are reported in parentheses below the estimated parameters. Bandwidth is set to three. In brackets we report a Chi-square statistic on the difference of a coefficient between a specific sub-period and the rest of the sample. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. Panel B reports decomposition results based on the VAR estimates from Panel A.

	(1)	(2)	(3)	(4)	(5)
	Neth./U.K. 1629-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2012	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.204** (0.081) [1.890]	0.125 (0.176) [0.532]	0.405 (0.250) [1.075]	0.343*** (0.088) [0.521]	0.162*** (0.057)
r_t	0.012 (0.118)	-0.215 (0.229)	0.006 (0.208)	0.074 (0.134)	0.034 (0.123)
dg_t	-0.066 (0.055)	-0.030 (0.106)	-0.408** (0.192)	-0.361 (0.532)	-0.034 (0.056)
Dep. variable: dg_{t+1}					
dp_t	-0.592** (0.294) [1.084]	-0.774*** (0.213) [0.308]	-0.338** (0.152) [1.281]	0.050 (0.039) [12.862***]	-0.374** (0.147)
r_t	0.157 (0.411)	-0.100 (0.209)	0.372** (0.152)	-0.026 (0.082)	0.347*** (0.115)
dg_t	-0.045 (0.166)	0.054 (0.097)	-0.226** (0.095)	-0.202 (0.202)	-0.164 (0.111)
Dep. variable: dp_{t+1}					
dp_t	0.231 (0.267) [4.721]	0.113 (0.376) [0.919]	0.298 (0.245) [0.275]	0.768*** (0.093) [9.384*]	0.521*** (0.171)
r_t	0.163 (0.400)	0.129 (0.427)	0.423** (0.208)	-0.109 (0.180)	0.352** (0.167)
dg_t	0.024 (0.156)	0.094 (0.181)	0.210 (0.181)	0.172 (0.469)	-0.146 (0.117)
J-test	2.520	0.351	0.703	0.574	2.529
Panel B: Decomposition results					
Dividend-to-price					
DR	0.278	0.216	0.579	1.090	0.351
CF	0.722	0.784	0.421	-0.090	0.649
Unexpected returns					
DR	-0.064	0.283	0.438	1.073	0.120
CF	1.046	0.725	0.571	-0.082	0.880

Table 4: Longer horizon predictability

This table reports return and dividend growth predictability results over different horizons. Panel A is based on overlapping observations, where the discounted sum of multi-year returns, multi-year dividend growth rates, and the future value of the dividend-to-price ratio are regressed on the lagged dividend-to-price ratio. Panel B reports the same results based on non-overlapping observations; we report the mean estimates across H different non-overlapping samples. The estimation is by iterative generalized method of moments subject to the present value model constraints. Heteroskedasticity and autocorrelation corrected standard errors based on a Bartlett kernel are reported in parentheses below the estimated parameters. Bandwidth is chosen optimally according to Newey-West method. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. The period is from 1685 through 2012.

	(1)	(2)	(3)
Horizon (H)	1 year	3 years	5 years
	1685-2012	1685-2012	1685-2012
Panel A: Overlapping observations			
Dep. variable: $\sum_{h=1}^H \rho^{h-1} r_{t+h}$			
dp_t	0.081*** (0.030)	0.223*** (0.065)	0.371*** (0.091)
Dep. variable: $\sum_{h=1}^H \rho^{h-1} dg_{t+h}$			
dp_t	-0.190*** (0.058)	-0.321*** (0.094)	-0.274*** (0.072)
Dep. variable: dp_{t+H}			
dp_t	0.759*** (0.075)	0.515*** (0.140)	0.436*** (0.140)
J-test	1.273	2.249	2.533
Panel B: Non-overlapping observations			
Dep. variable: $\sum_{h=1}^H \rho^{h-1} r_{t+h}$			
dp_t	0.081*** (0.030)	0.236*** (0.084)	0.377*** (0.121)
Dep. variable: $\sum_{h=1}^H \rho^{h-1} dg_{t+h}$			
dp_t	-0.190*** (0.058)	-0.328*** (0.114)	-0.269** (0.107)
Dep. variable: dp_{t+H}			
dp_t	0.759*** (0.075)	0.493*** (0.173)	0.434** (0.189)
J-test	1.273	1.828	1.935

Table 5: Cross-period out-of-sample predictions

This table reports in-sample R-square and out-of-sample R-square (in brackets) for predicting annual returns and dividend growth rates. Both R-squares are calculated according to Eq. (16). The in-sample R-square is estimated over the same period as the estimated parameters. The out-of-sample R-square is calculated using estimated parameters from a different period; that is the out-of-sample R-square for the period 1685-1945 (1946-2012) is based on the estimated parameters from the period 1946-2012 (1685-1945). Results are based on the reduced form-model where returns and dividend growth rates are predicted using lagged dividend-to-price ratio only.

	In-sample [out-of-sample] R-square			
	1685-1945		1945-2012	
	Returns	Dividend growth	Returns	Dividend growth
Parameters from:				
1685-1945	0.034	0.205	[0.110]	[-4.142]
1945-2012	[0.031]	[-0.022]	0.115	0.011

Table 6: Predictability over the business cycle

This table reports summary statistics (Panel A) and predictability results (Panel B) over the business cycle. Estimated parameters for expansions and recessions are estimated jointly using iterative generalized method of moments subject to the present value model constraints (Eq (17), (18)). Heteroskedasticity and autocorrelation corrected standard errors based on a Bartlett kernel are reported in parentheses below the estimated parameters. In brackets is the Wald test for equality of coefficients. Bandwidth is chosen optimally according to Newey-West method. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. The period is either 1700-1945, or 1945-2012, or 1700-2012.

	(1)	(2)	(3)	(4)	(5)	(6)
	1700-1945		1945-2012		1700-2012	
	Recessions	Expansions	Recessions	Expansions	Recessions	Expansions
Panel A: Summary statistics						
r	0.027	0.087	-0.007	0.123	0.023	0.098
$Std.(r)$	0.140	0.111	0.215	0.140	0.148	0.121
t -stat	[-3.576] ***		[-2.006] *		[-4.555] ***	
dg	-0.005	0.025	0.033	0.063	-0.001	0.037
$Std.(dg)$	0.143	0.117	0.103	0.054	0.140	0.103
t -stat	[-1.730]		[-0.992]		[-2.525] **	
dp	-3.066	-3.126	-3.270	-3.499	-3.087	-3.240
$Std.(dp)$	0.265	0.218	0.463	0.428	0.295	0.344
t -stat	[1.855]		[1.573]		[4.106] ***	
No. of years	105	124	12	55	117	179
Panel B: VAR estimates						
Dep. variable: r_{t+1}						
dp_t	0.265*** (0.082) [3.375]*	0.102*** (0.032)	0.258*** (0.069) [4.699]**	0.081* (0.044)	0.117** (0.052) [1.104]	0.050 (0.037)
Dep. variable: dg_{t+1}						
dp_t	-0.174* (0.089) [0.123]	-0.138*** (0.046)	0.099* (0.055) [2.223]	0.003 (0.027)	-0.190** (0.076) [1.800]	-0.079** (0.031)
Dep. variable: dp_{t+1}						
dp_t	0.587*** (0.098) [3.369]*	0.795*** (0.056)	0.870*** (0.109) [0.436]	0.951*** (0.054)	0.723*** (0.085) [3.129]*	0.906*** (0.059)
J-test	10.507**		0.084		0.948	

Table 7: Treasury betas

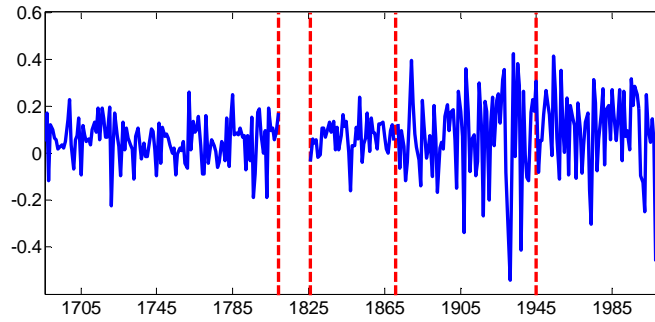
This table reports GMM estimates of the market betas of medium- and long-term Treasury bonds, where we take the S&P 500 as the market and T-bills as the risk-free rate. Monthly data is from Ibbotson and Associates (2013). To account for the change in the variance of bond and equity returns over time (Forbes and Rigobon 2002), we scale bond and stock returns by their standard deviation calculated over the pre-1946 period (Dungey et al. 2005). Heteroskedasticity and autocorrelation corrected standard errors based on Bartlett kernel are reported in parentheses below the estimated parameters. Bandwidth is set to 10. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. We report a Chi-square test on the difference in interaction terms between medium and long term bonds.

	1926-2012		1926-2007	
	Medium-term	Long-term	Medium-term	Long-term
$(R_M - r_f) / \sigma_{1926-1945}$	0.115 (0.080)	0.119 (0.082)	0.115 (0.080)	0.119 (0.082)
$Post - 1945 * (R_M - r_f) / \sigma_{1926-1945}$	0.183 (0.226)	0.357 (0.266)	0.327 (0.242)	0.609** (0.272)
$Post - 1945$	-0.027 (0.070)	-0.025 (0.074)	-0.063 (0.073)	-0.068 (0.074)
N	1044	1044	984	984
$\Delta\beta^{medium} = \Delta\beta^{long}$	1.410		6.770***	
(p-value)	(0.236)		(0.009)	

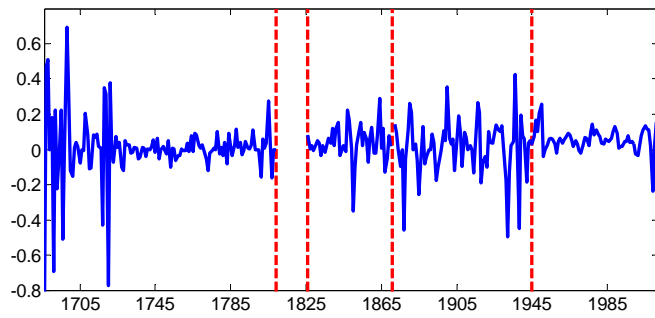
Figure 1: Returns, dividend growth rates, and the dividend-to-price ratio

This figure plots returns, dividend growth rates, and the dividend-to-price ratio. The data is annual. Vertical dashed lines denote the time periods: Netherlands and U.K. (1685-1809), U.K. (1825-1879), U.S. pre-war (1871-1945), and U.S. post-war (1945-2012).

Panel A: Annual returns



Panel B: Annual dividend growth rates



Panel C: Dividend-to-price ratio

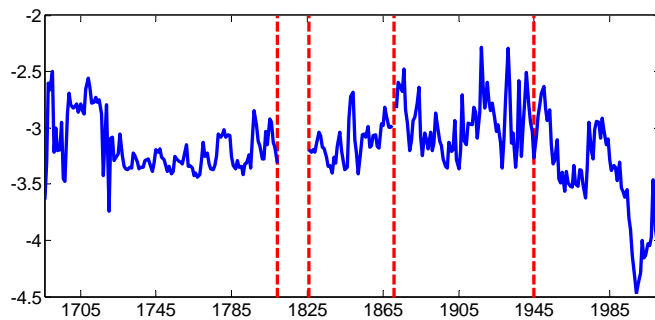
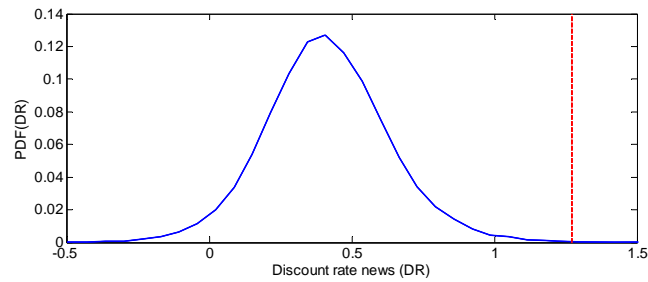


Figure 2: Monte Carlo simulations: Probability density functions (PDF)

This figure plots simulated probability density functions for the discount rate news component in the dividend-to-price ratio (Panel A) and in the unexpected returns (Panel B). We simulate 100,000 data paths based on the distribution of parameters and errors estimated on the pre-1945 data. The simulated data match the length of the post-1945 period. To speed up simulations, we use one step GMM estimation and set the Newey-West bandwidth to 3. Vertical dashed lines denote the post-1945 point estimates.

Panel A: PDF: Discount rate component of the dividend-to-price ratio



Panel B: PDF: Discount rate component of unexpected returns

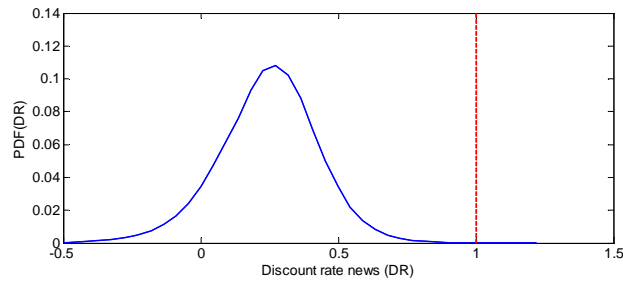


Figure 3: Dividends-to-earnings ratio

This figure plots 20 year rolling averages of the dividends-to-earnings ratio for the Netherlands/U.K. (1685-1809) period and the combined U.S. period (1871-2012).

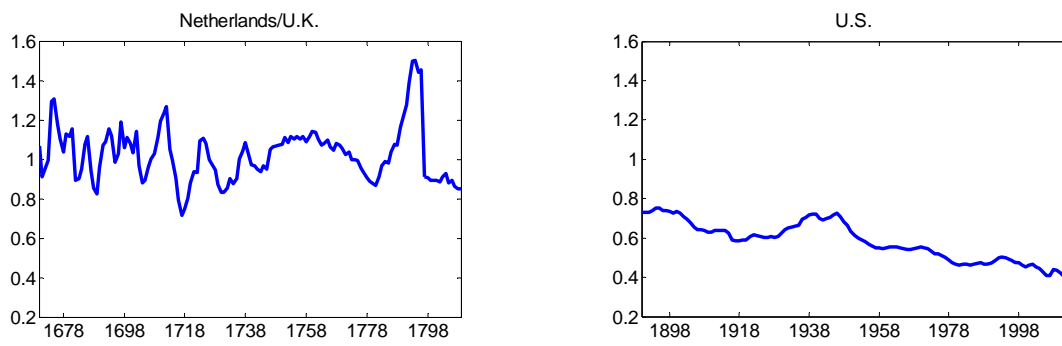
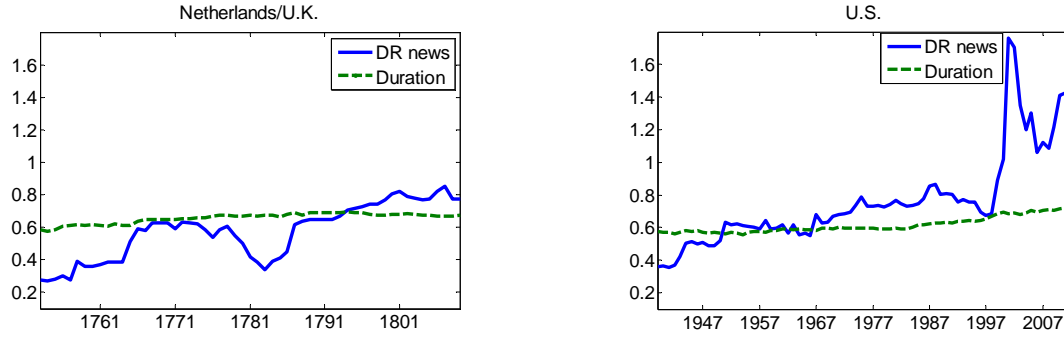


Figure 4: Duration and discount rates: Rolling windows

This figure plots our measure of duration along with either the discount rate component of the dividend-to-price ratio (Panel A) or the discount rate component of unexpected returns (Panel B). All measures are estimated on rolling windows of 68 years (to match the length of the post-1945 period). Estimations are done separately for the Netherlands/U.K. (1685-1809) period and the combined U.S. period (1871-2012). The first rolling window ends in 1753 for the Netherlands/U.K. and in 1939 for the U.S. period.

Panel A: Duration and discount rate component of the dividend-to-price ratio



Panel B: Duration and discount rate component of unexpected returns

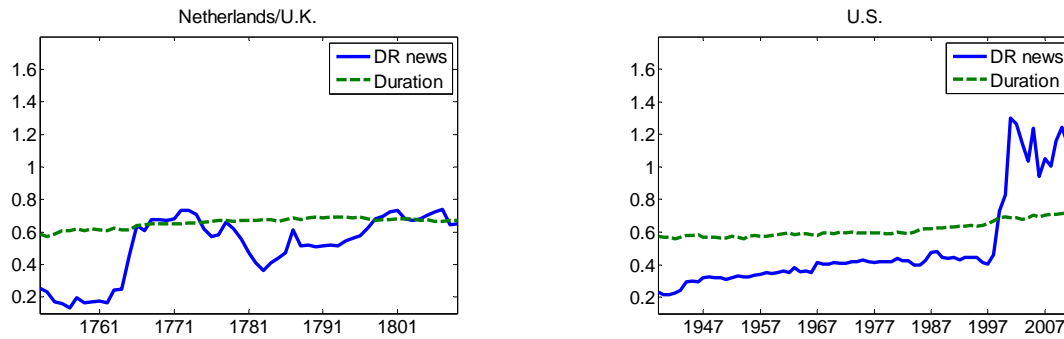
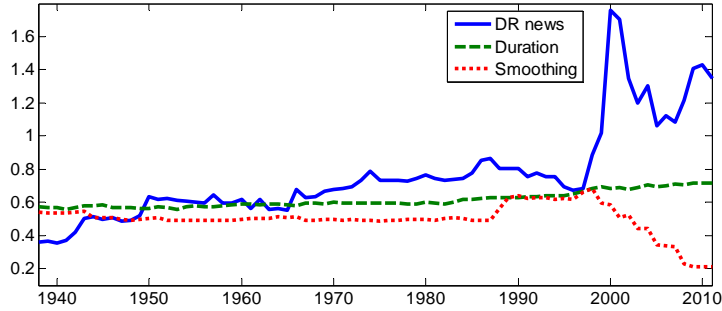


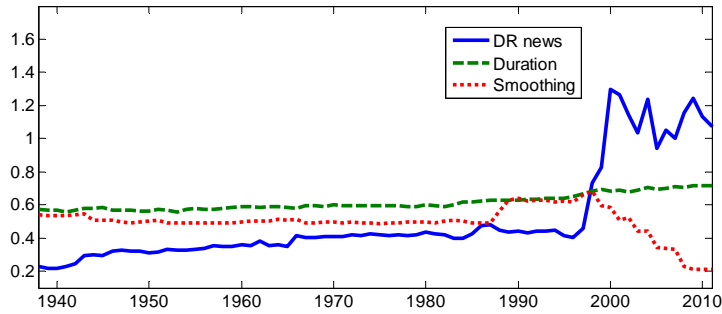
Figure 5: Duration, dividend smoothing parameter, and discount rates: Rolling windows (U.S. period: 1871-2012)

This figure plots our measure of duration along with the dividend smoothing parameter and either the discount rate component of the dividend-to-price ratio (Panel A) or the discount rate component of unexpected returns (Panel B). Dividend smoothing parameter is a ratio of dividend growth volatility and the earnings growth volatility. All measures are estimated on the annual U.S. data from 1871 through 20012 using rolling windows of 68 years (to match the length of the post-1945 period). The first rolling window ends in 1939.

Panel A: Duration, dividend smoothing parameter, and discount rate component of the dividend-to-price ratio



Panel B: Duration, dividend smoothing parameter, and discount rate component of unexpected returns



Online Appendix A: Reduced-form model

We repeat all our main results from Table 2 using a reduced form model. In particular, rather than using the full VAR model, we predict returns, dividend growth rates, and the dividend-to-price ratio by the lagged values of the dividend-to-price ratio only (as before, all variables are demeaned):

$$\begin{aligned}r_{t+1} &= \beta^r dp_t + \varepsilon_{t+1}^r, \\ dg_{t+1} &= \beta^{dg} dp_t + \varepsilon_{t+1}^{dg}, \\ dp_{t+1} &= \beta^{dp} dp_t + \varepsilon_{t+1}^{dp},\end{aligned}\tag{OA 1}$$

subject to the constraint imposed by the present value model:

$$\beta^r - \beta^{dg} + \rho\beta^{dp} = 1.\tag{OA 2}$$

We have three moment conditions and one linear restriction. As in the main analysis, we estimate the system of equation by iterative GMM and report heteroscedasticity and autocorrelation consistent statistics.

Based on the estimated parameters, we infer long-horizon estimates from their short-run analogs. Dividing Eq. (OA 2) by $1 - \rho\beta^{dp}$ and rearranging, we obtain:

$$\frac{\beta^r}{1 - \rho\beta^{dp}} - \frac{\beta^{dg}}{1 - \rho\beta^{dp}} = 1\tag{OA 3}$$

The first term can be interpreted as the variation of the dividend-to-price ratio due to discount rates. Negative the second term captures variation due to cash-flows (see also Cochrane 2008).

Results are reported in Table OA 1 and are very similar to the results from the main analysis reported in Table 2.

Table OA.1: Reduced-form model: Annual data

Panel A reports the regression estimates of returns, dividend growth rates, and the dividend-to-price ratio on the lagged dividend-to-price ratio. The data are annual. The model is estimated by iterative generalized method of moments subject to the present value model constraint. Heteroskedasticity and autocorrelation corrected standard errors based on Bartlett kernel are reported in parentheses below the estimated parameters. The Newey and West method is used for the selection of the optimal bandwidth. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. Panel B reports decomposition results based on the parameter estimates from Panel A.

	(1)	(2)	(3)	(4)	(5)
	Neth./U.K. 1685-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2012	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.120*** (0.034)	0.063** (0.026)	0.055 (0.073)	0.123*** (0.044)	0.081*** (0.030)
Dep. variable: dg_{t+1}					
dp_t	-0.230*** (0.088)	-0.294*** (0.046)	-0.450*** (0.068)	0.015 (0.025)	-0.190*** (0.058)
Dep. variable: dp_{t+1}					
dp_t	0.677*** (0.100)	0.671*** (0.045)	0.520*** (0.091)	0.920*** (0.048)	0.759*** (0.075)
J-test	2.290	0.003	0.201	0.104	1.273
Panel B: Decomposition results					
Dividend-to-price					
DR	0.342	0.175	0.109	1.139	0.300
CF	0.658	0.825	0.891	-0.139	0.700

Online Appendix B: Inflation adjustment

In this Appendix, we repeat our main results from Table 2 using inflation adjusted prices and dividends. The data for the earlier period come from the website of the International Institute of Social History (<http://www.iisg.nl/hpw/data.php#netherlands>) (The Netherlands between 1629 and 1809) and Clark (2015) (England between 1825 and 1870). For the U.S. period, we use the inflation index (CPI) from Robert Shiller's webpage.

Note that the sub-periods differ substantially in terms of inflation. In the early periods, inflation was relatively modest. Prices increased by 45% in the Netherlands/U.K. period 1685-1809. In the U.K. period 1825-1870 prices decreased by 22.5%. In the early part of the U.S. sample 1871-1945, prices increased again by 44%. In sharp contrast to the early samples, inflation in the post-1945 period was substantially higher; the CPI increased 12.6 times, which corresponds to 3.91% per year, on average.

These large differences in inflation, however, cannot account for the increase in the discount rate news in the recent period. As reported in Table OA.2, the results for inflation adjusted data are very comparable to the main results in Table 2. The only notable difference is that returns become somewhat more predictable and dividend growth rates somewhat less predictable in the early periods. The contrast between the recent U.S. period and the early periods, however, remains.

Table OA 2: Vector autoregression (VAR) estimates: Annual inflation adjusted data

Panel A reports VAR estimates of returns, dividend growth rates, and the dividend-to-price ratio. The data is annual and adjusted for inflation. The model is estimated by iterative generalized method of moments subject to the present value model constraints. Heteroskedasticity and autocorrelation corrected standard errors based on Bartlett kernel are reported in parentheses below the estimated parameters. The Newey and West method is used for the selection of the optimal bandwidth. Statistical significance at the one, five, and ten percent is denoted by three, two, and one asterisks. Panel B reports decomposition results based on the VAR estimates from Panel A.

	(1)	(2)	(3)	(4)	(5)
	Neth./U.K. 1685-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2012	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.136*** (0.042)	0.244** (0.107)	0.138 (0.088)	0.118*** (0.037)	0.108*** (0.029)
r_t	-0.002 (0.079)	0.180* (0.105)	0.115 (0.169)	-0.027 (0.094)	0.032 (0.060)
dg_t	0.003 (0.085)	-0.003 (0.119)	-0.261* (0.145)	0.099 (0.194)	0.021 (0.072)
Dep. variable: dg_{t+1}					
dp_t	-0.145 (0.090)	-0.209* (0.110)	-0.282*** (0.074)	-0.007 (0.020)	-0.068** (0.030)
r_t	0.158 (0.107)	-0.194 (0.199)	0.272** (0.102)	0.124* (0.069)	0.175** (0.066)
dg_t	-0.050 (0.092)	0.291** (0.122)	0.047 (0.089)	0.307*** (0.079)	0.027 (0.090)
Dep. variable: dp_{t+1}					
dp_t	0.749*** (0.108)	0.571*** (0.104)	0.610*** (0.100)	0.902*** (0.050)	0.858*** (0.047)
r_t	0.167 (0.119)	-0.391** (0.189)	0.165 (0.214)	0.155 (0.134)	0.149 (0.092)
dg_t	-0.055 (0.100)	0.307** (0.140)	0.323* (0.168)	0.215 (0.185)	0.007 (0.099)
J-test	2.360	0.054	1.804	1.105	6.752*
Panel B: Decomposition results					
Dividend-to-price					
DR	0.524	0.482	0.411	1.084	0.672
CF	0.476	0.518	0.589	-0.084	0.328
Unexpected returns					
DR	0.241	0.254	0.312	0.839	0.447
CF	0.761	0.746	0.694	0.161	0.555