Four centuries of return predictability

Benjamin Golez†
University of Notre Dame

Peter Koudijs‡
Stanford University

This version: March 15, 2016

Abstract

Using the Dutch, English and U.S. stock markets, we provide evidence for return predictability across four centuries. Most of the predictability stems from recessions. In downturns, the dividend-to-price ratio increases and predicts higher returns going forward, both over annual and multi-annual horizons. For the period as a whole, dividend growth rates are also predictable, but in the recent U.S. period return predictability dominates. We hypothesize that this is the result of firms retaining more earnings for investment, which extended the duration of the stock market as a whole and increased the importance of (persistent) discount rate shocks.

Key words: Dividend-to-price ratio, return predictability, dividend growth predictability, stock market 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), John Campbell (discussant), Anna Cieslak, Zhi Da, Arvind Krishnamurthy, Hanno Lustig, Stijn van Nieuwerburgh (discussant), Monika Piazzesi, Martin Schneider, Robert Stambaugh (discussant), Jessica Wachter, Michael Weber (discussant), Nicholas Hirschey (discussant) and seminar participants at University of Notre Dame, UC San Diego, Stanford University, European Central Bank, University of Konstanz, USI Lugano, Nova School of Business and Economics, Bocconi University, University of Zurich, 2014 SITE Conference, 2015 Rodney L. White Center for Financial Research Conference, 2015 Western Finance Association Conference, 2015 Red Rock Conference, 2016 American Economic Association Conference, and 2016 European Winter Finance 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 stock price 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).

A number of issues remain. First of all, most studies use relatively recent U.S. data. This raises the question how representative these findings are for financial markets in general (Schwert 1990). A complicating factor is that the dividend-to-price ratio is highly persistent in the recent U.S. period, virtually indistinguishable from a unit root. Combined with a relatively short sample, this biases estimates in favor of finding return predictability (Stambaugh 1999). Moreover, parameter estimates appear instable when adding or removing years from the sample (Goyal and Welch 2003, Schwert 2003, Koijen and Van Nieuwerburgh 2011).[1] 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 annual asset prices and dividends to cover the whole history of modern financial markets starting in 17th century Amsterdam. In particular, we

---

[1] There is also substantial variation in the estimated parameters across countries and in the cross-section of portfolios (Campbell and Shiller 2005; Rangvid, Schmeling, and Schrimpf 2014; Santa-Clara and Maio 2015).

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 separation of ownership and control, an active secondary market for shares, and (generally) limited liability for shareholders.

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. Using the extended time series, we then revisit the question of return predictability, explore the main sources of price movements, and analyze how predictability varies over different horizons and over the business cycle. The long time perspective also provides new insights about modern markets and helps us delve in the driving forces behind the dominant role of discount rate news in the recent U.S. data.

The paper has four key empirical findings. First, across all four centuries, the dividend-to-price ratio is stationary and fluctuates around a long-run average of five percent (minus three in logs, see Figure 1 for details). Only after around 1945, the dividend-to-price ratio decreases and becomes increasingly persistent. Despite this decrease, expected returns have been approximately stable over time, with returns increasingly coming from capital appreciation in lieu of decreasing dividends.

Second, we find strong evidence for return predictability. For the entire 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. Given the long time series and the stationarity of the dividend-to-price ratio, the results are not prone to small sample bias

(Stambaugh 1999). Moreover, the return predictive coefficient is remarkably stable across subperiods (although not always statistically significant). Overall, this suggests that excess volatility is a pervasive characteristic of financial markets.

Third, the dividend growth rate is also predictable in the full period for annual and multiannual horizons. However, there exist important differences between periods. While the dividend-to-price ratio strongly predicts dividend growth rates in the earlier periods, such predictability completely disappears around 1945. This implies that changes in cash-flows were more important for price movements before 1945 than nowadays. The dominance of discount rate news is therefore a relatively recent phenomenon.

Fourth, the predictability of returns is concentrated in recession periods – more so than for dividend growth rates – and the dividend-to-price ratio increases during downturns. This supports the view that expected returns are countercyclical and that business cycles are an important driver of return predictability.

In sum, our analysis reveals that returns have always been predictable, especially in downturns. At the same time, the recent (post 1945) period stands out with an exceptionally low dividend-to-price ratio and a dominant role for discount rate news. In light of relatively stable expected returns across periods, these changes do not seem to be driven by differences in equity valuations. Instead we hypothesize that the changes are closely linked to a recent increase in the fraction of companies' earnings that are being reinvested. As companies reinvest more and push cash flow distributions into the future, the duration of the stock market increases. This suggests that modern companies can be seen as longer-lived high duration assets. As such, they have lower dividend-to-price ratios and, as long as expected returns are more persistent than dividend growth rates, prices are more prone to changes in discount rates. We illustrate this intuition with a simple stylized present value model where the market consists of finitely lived companies.

Consistent with our hypothesis, we show empirically 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.

We investigate a number of alternative explanations for the dominant role of discount rate news. Chen, Da, and Priestly (2012) argue that dividend smoothing biases 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 variation. In addition, we document that dividend smoothing in the earlier part of our sample was comparable to today. We also study the role of alternative ways to distribute cash flows. The evidence suggests that this cannot explain long term patterns. Payout ratios are decreasing even when we account for net repurchases and the results are robust to ending the recent period before repurchases became economically important in 1982.

Our paper is related to several strands in the literature. First, we contribute to the literature on return predictability (Campbell and Shiller 1988; Fama and French 1988; Stambaugh 1999; Boudoukh, Richardson and Whitelaw 2008). Relative to these studies, we use a much longer time series that spans four centuries of dividends and prices. We show that the extended time period is free of statistical problems rendering return predictability tests in the recent U.S. data unreliable, and establish robust evidence for return predictability over different horizons. The longer time series also features many more recession periods and therefore allows us to establish the importance of business cycles variation for time-varying expected returns (Campbell and Cochrane 1999; Henkel, Martin, and Nardari 2011).

Second, we contribute to the literature that emphasizes the role of dividend growth predictability and that analyzes the relative importance of discount rate and cash flow news for stock prices (Campbell and Shiller 1988; Menzly, Santos, and Veronesi 2004; Lettau and Ludvigson 2005; Cochrane 2008; Binsbergen and Koijen 2010; Koijen and Van Nieuwerburgh 2011; Chen, Da, and Priestly 2012; Golez 2014). In comparison, we provide a long run perspective on the sources of prices movements. Chen (2009), Goyal and Welch (2003), and Schwert (2003) show that, in the U.S. data, dividend growth predictability is stronger and (return predictability weaker) before 1945. We find that dividend growth rates have been predictable over most of modern financial markets' course of history, reinforcing the notion that the dominance of discount rate news is a recent phenomenon. Importantly, the paper's long run perspective provides new insights for this phenomenon. Based on the falling payout ratios and the notion that the market consists increasingly of growth oriented stocks (Fama and French 2001; Campbell and Voulteenaho 2004), we argue that the increased duration of the stock market as a whole provides a natural explanation for the importance of (persistent) discount rates in the recent period.

Finally, our paper is related to other studies analyzing price movements in a historical context. Relying on primary sources, Goetzmann, Ibbotson, and Peng (2001) estimate a stock 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 (see also Schwert 1990). In comparison, we obtain actual price and dividend data across all four centuries. Le Bris, Goetzmann, and Pouget (2014) analyze six hundred years of dividend and price data for an individual company, namely the Bazacle Company in France. In comparison, we analyze the aggregate market. Interestingly, both papers find that the dividend-to-price ratio fluctuates around a long run average of five percent.

2. Data

We extend the time series of annual stock prices and dividends back in time until 1629 using the most important financial markets of a specific period. In particular, for the period 1629 through 1809, 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 through 2014, with only a small gap for the years between 1809 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. We are 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 (*forthcoming*) 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. 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 (VOC, from 1629 onwards)[3], the (United) British East India Company (EIC, 1692), the Bank of England (BoE, 1697; at the time a private bank with strong ties to the government), the South Sea Company (SSC, 1711), and the (Second) Dutch West India Company (WIC, 1719). Even though VOC share prices are available starting in 1629, our annual regressions use data starting in 1686. Before that, the VOC only paid out dividends every two or three years and it is impossible to calculate annual dividend growth rates and dividend-to-price ratios. Appendix A provides details on data sources and the underlying securities.

It is important to note that, even though these are only five securities, they constitute the virtual 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.[5] In comparison, for the U.S., stock market capitalization amounted to 39% of GDP in 1913 and 152% in 1999 (Rajan and Zingales 2003).

---

[2] Both were on metallic standards and exchange rate fluctuations were only of minor importance.

[3] 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.

[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. Historical evidence suggests that Dutch investors mostly lived in this area. We use information from Bowen (1989) and Wright (1997) to calculate what fraction of English securities were held by Dutch investors.

[5] 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.

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 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, newly issued companies usually had full shareholder liability (Hickson, Turner, and Ye 2011). 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 frequently traded domestic equities 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.[6] 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.

To facilitate comparison with the existing literature, we rely on U.S. stock market data starting in 1871 using the data from Amit Goyal's website. For the period between 1871 and 1925, these data rely on information from Cowles (1939) that covers between 50 (1871) and 258

---

[6] Adjustment for the survivorship bias affects the level of the risk premium, but it does not materially impact the predictability results. Our results are robust to using the data that are not adjusted for attrition.

(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.[7]

In alternative estimates, we focus on the London market for the entire 19th century (relying on the data from Grossman 2002), only switching to the U.S. after 1900, as one can argue that it was not before the 1900 that the U.S. became the dominant economy in the world with a well-developed capital market in New York. This has the additional advantage that London featured many more securities (~750) than New York in this period. In the Online Appendix (Table OA.4), we show that all our main results are robust to this alternative approach. In the same appendix, we also show that results are similar when using the broad based CRSP index after 1925 that, again, is comprised of a significantly larger number of securities (522 in 1926).

To account for changes in the purchasing power, we additionally obtain data on price levels. For the Netherlands/U.K. period (1629-1809), we rely on information from the International Institute of Social History.[8] Clark (2015) provides us with the data on inflation for the U.K. period (1825-1870/1900). For the U.S. period, we use the inflation index (CPI) from Robert Shiller's webpage.

3. Methodology

We study the present value relations between asset prices and cash flows. 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-

---

[7] In the Cowles data the end-of-year price is the average price in December rather than the price on the last trading day of a year. We verify in the post 1925 data that this does not materially affects our results.

[8] http://www.iisg.nl/hpw/data.php#netherlands.

flows. Other types of distributions (e.g. repurchases) are assumed to be reinvested in the company.[9]

3.1 Present value relations

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

$$R_t = \frac{P_t + D_t}{P_{t-1}}, \tag{1}$$

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:

$$r_{t+1} \simeq dp_t + dg_{t+1} - \rho dp_{t+1}, \tag{2}$$

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:

$$dp_t \simeq r_{t+1} - dg_{t+1} + \rho dp_{t+1}. \tag{3}$$

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

---

[9] See Section 6.3 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.

growth predictability should best be studied jointly (Lettau and Ludvigson 2005; Cochrane 2008; Binsbergen and Koijen 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):

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

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:

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

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:

$$x_{t+1} = \phi x_t + \varepsilon_{t+1}, \tag{6}$$

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:

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

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. The restrictions can be written as:

$$\left(e_1' - e_2' + \rho e_3'\right)\phi = e_3'. \tag{8}$$

In total we have nine moment conditions, nine parameters, and three linear restrictions. The VAR model is therefore overidentified. We estimate the model using two-step GMM, and we test for overidentifying restrictions using a *J*-test.[10] 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):

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

---

[10] In the Online Appendix (Table OA.2), we show that our results are robust to using iterative GMM rather than two-step GMM.

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

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

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.

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

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

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

(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. Next, we illustrate the basic present value relations in the data using scatterplots. Finally, we report the VAR estimates and the decomposition results.

4.1 Summary statistics

Table 1 presents summary statistics for the annual data, both in nominal (Panel A) and real terms (Panel B). In early periods, prices were stable, featuring moderate inflation or even deflation. In sharp contrast, inflation in the post-1945 period was relatively high at 3.82% per year (on average). Although inflation does not affect the ratio of dividends and prices, any variable that predicts inflation may also predict nominal returns and nominal dividend growth rates. Therefore, in our main analysis, we always rely on real data.[11] Figure 1 plots the main variables of interest: annual real returns, annual real dividend growth rates, and the annual dividend-to-price ratio. Dashed lines separate the different time periods: Netherlands/U.K. (1686-1809), 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 (1945-2014).

The dividend-to-price ratio is remarkably stationary in the early periods. It always oscillates around a long-run average of approximately 5% (minus 3 in log terms). Only in the recent period, the dividend-to-price ratio decreases substantially; it is 4.85% on average in 1686-1945 period and 3.40% in the recent 1945-2014 period. At the same time, the dividend-to-price ratio becomes increasingly persistent; the AR(1) coefficient increases from between 0.51 and 0.68 in the first three periods 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.

Nominal returns are between 5.5% in the Netherlands/U.K period and 10.3% in the post-1945 period. The higher nominal returns in the recent period are mostly due to inflation; real returns are very comparable across periods and vary between 5.1% and 7.5%. The risk premium varies between 2.3% in the Netherlands/U.K. period and 6.1% in the recent U.S. sample. There is no distinct pattern in the Sharpe ratios though, mainly because returns are more volatile in the two U.S. periods. Persistence of real returns is relatively low, with the AR(1) ranging between -0.08 and 0.11.

---

[11] In the Online Appendix (Table OA.3), we show that our results are robust to using nominal data.

In the recent period returns consist for a large part of capital appreciation. Whereas in the earlier periods around two thirds of nominal returns stem from dividends (and only one third comes from capital appreciation), this is exactly the opposite in the recent U.S. sample. In terms of real returns, the difference is somewhat less pronounced, but still apparent; half of the real returns stem from capital appreciation in the recent period, and around one quarter on average in the early periods (the only exception is the 1825-1870 U.K. period, where due to deflation, capital appreciation increases to 42% of real returns).

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 (so-called dividend lumpiness). 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 has flattened out once again in the recent years, probably due to increased dividend smoothing. Accordingly, dividend growth rates become increasingly persistent over time: the AR(1) coefficient is -0.03 in the Netherlands/U.K period and increases to 0.37 in the recent U.S. period. To address the concern that differences in the dividend growth process affect our main results, we also analyze data sampled at lower frequencies. In particular, we sample the data every three years (triennial), where we sum dividends over three years and take the price of the end of that period. This also allows us to extend the time period backwards by 57 years to 1629. The summary statistics for the triennial data are reported in the Online Appendix, Table OA.1.

4.2 Scatterplots

We first illustrate the basic present value relations in scatterplots, similarly to Campbell and Shiller (2005). In particular, for each sample period, we plot annual real returns and dividend growth rates against the lagged dividend-to-price ratio. All variables are demeaned. Figure 2 shows that there is a positive relation between the dividend-to-price ratio and returns. Consistent with the present value framework, a relatively high dividend-to-price ratio predicts higher next period returns in all sub-periods. The slope is steepest in the U.K. period and flattest, although still upward sloping, in the early U.S. period. Consistent with theory, we find the opposite relation for dividend growth rates where a high dividend-to-price ratio is generally associated with lower dividend growth. The only exception is the recent U.S. period where the slope is virtually flat. Combining the data from the different sub-periods suggests that both returns and dividend growth rates are predictable across four centuries of data.

4.3 VAR estimates

To analyze the present value relations more formally, we present VAR estimates and decomposition results, using both annual and triennial real data. We estimate the model for each period separately as well as for the full sample. All the variables are demeaned by the respective sample mean: in the full sample, variables are demeaned at the full sample mean. When appending the data, we are careful to avoid cross-period predictions.[12] We always report the full parameter matrix associated with the VAR estimates, but focus our attention on the parameters associated with the lagged dividend-to-price ratio.

---

[12] That is, we avoid predicting first year return and dividend growth rate in the U.K. period using the last year observations from the Netherlands/U.K. period. Similarly, we avoid predicting first year U.S. data return and dividend growth using the last year of the U.K. data. As a result, a predictive regression in our full sample period has 308 observations (rather than 310).

4.3.1 Annual real data: 1686-2014

Table 2 reports the VAR estimates based on annual data. Consistent with the previous evidence, the dividend-to-price ratio predicts both returns and dividend growth rates in the full period. The estimated parameters on the dividend-to-price ratio have the expected sign and are of the same magnitude; positive at 0.11 in the return regression, and negative at -0.09 in the dividend growth regression. Both parameters are significant at the one percent level. Importantly, these estimates do not seem to suffer from small sample bias. The dividend-to-price ratio in the full sample is stationary, and our sample spans 308 annual observations. This substantially reduces problems associated to predicting returns with highly persistent predictors in small samples. We confirm this in simulations based on Stambaugh (1999), reported in Appendix B. Small sample bias only plays an important role in the recent U.S. period where the dividend-to-price ratio is highly persistent. In other sample periods and in particular in the full sample, the bias is rather small. Moreover, the bootstrapped *p*-value that takes into account small sample bias is below one percent in the full period.[13]

Looking at the different sub-periods separately, the first thing that stands out is the remarkable stability of the predictive coefficient in the return regressions across all four centuries. In particular, the estimated parameter on the dividend-to-price ratio is always positive, between 0.10 and 0.24.[14] It is significant at the one percent level in the Netherlands/U.K. period and in the recent U.S. sample, and at the five percent level in the 19th century U.K. period. Even in the early U.S. period, where the estimated parameter is not statistically significant, a Wald test

---

[13] This smaller bias in the full sample versus the recent U.S. period is the result of a combination of three factors: many more observations (308 vs. 68), lower correlation between innovations in the dividend-to-price ratio and errors in the predictive regression (-0.69 vs. -0.93), and lower persistence of the dividend-to-price ratio (0.83 vs. 0.92).

[14] Note that the return and dividend growth parameters in the full sample are smaller than the average parameters in the sub-periods. This is mostly due to the differences in the mean of the dividend-to-price ratio; in the full period, we demean all the variables using the full period mean. If we demean the variables for each sub-period separately and then append the data, the estimated return parameter is 0.14, and the estimated dividend growth parameter is -0.12. Both parameters remain significant at the one percent level.

implies that the coefficient is not statistically significantly different from the rest of the period. This suggests that returns have always been predictable by the dividend-to-price ratio.[15]

In comparison, the estimated parameter in the dividend growth regression is much less stable. It is negative and relatively large, between -0.21 and -0.28, and overall significant in the first three periods, but turns out to be almost zero (-0.01) 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.[16] Whereas the dividend-to-price ratio predicts itself with the estimated parameter between 0.57 and 0.63 in the first three periods, this coefficient increases to 0.92 in the recent U.S. sample. Again, the difference is highly statistically significant.

This is also reflected in the decomposition results reported in Panel B of Table 2. While in the first three samples, cash-flow news accounts for approximately half of the price variation, all the variation in the dividend-to-price ratio appears to be driven by discount rates in the recent U.S. sample. Focusing on the decomposition of unexpected returns yields similar results. Thus, the dominance of discount rate news appears to be a relatively recent phenomenon. To verify that the recent period is statistically different, we first conduct a Wald test on the VAR model in the full period imposing a break in 1945. The Wald test confirms the presence of a structural break with a p-value of 0.00.[17] Next, we test for the difference in decomposition results more directly by conducting Monte Carlo simulations. We use the distribution of parameters and errors

---

[15] Note that our evidence for return predictability in the early U.S. sample, although insignificant, is somewhat stronger than in Chen (2009). This is because we use real data rather than nominal data and because we use fully specified VAR model with present value restrictions rather than predict returns with the dividend-to-price ratio only.

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

[17] Interestingly, when we endogenously search for one break, we again arrive at 1945 (using a Sup statistic for Wald test).

estimated using the data up to 1945 to simulate 100,000 datasets that match the length of the post-1945 period. We then perform the same decompositions as in the main analysis. Figure 2 shows the resulting distribution of the discount rate component. The post-1945 is far in the right tail with a *p*-values of 0.01 (decomposition based on the dividend-to-price ratio) and 0.00 (decomposition based on unexpected returns).

4.3.2 Triennial real data: 1629-2014

To investigate the impact of changes in the underlying dividend process, with dividends being either "lumpy" or smoothed out over time, we replicate our main results using triennial data. This also enables us to extend the data by 57 years to 1629-2014. Relying on triennial data, however, reduces the number of observations. We therefore pool the data before 1945 and consider three different samples: the 1629-1945 period, the recent U.S. period 1945-2014, and the full period. For each period, we estimate the VAR model on three different non-overlapping triennial samples and report the mean of the estimated parameters across these samples.

Table 3 shows that all the main conclusions from the annual data (Table 2) carry over to triennial data. The dividend-to-price ratio predicts both returns and dividend growth rates in the full period.[18] The same holds for the early period. In the post-1945 period, however, the dividend-to-price ratio predicts returns only. Both a Wald test and simulation results confirm that the importance of discount rate news is significantly higher in the recent period. Thus, the documented patterns appear to be a deep characteristic of the market that seems to go beyond dividend lumpiness (or smoothing).

---

[18] We also verified that returns and dividend growth rates are predictable in the full period for data sampled at even lower, e.g. five-year frequency.

5. Further evidence

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

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.5), we show that all the results from the main analysis (Table 2) are robust to using this reduced form model.

5.1 Longer horizon predictability

In this section we investigate whether the predictability results we document for annual data can be extended to longer horizons. This is important for two reasons. First, the presence of long run predictability would suggest that prices do not simply move in response to transitory shocks, but reflect fundamental changes in the underlying economy. Second, most studies find that return predictability becomes stronger at longer horizons (see for example Fama and French 1988 and Cochrane 2005). The intuition for this result is that, in the recent data, the dividend-to-price ratio is highly persistent and therefore commands highly persistent expected returns. 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 was much less persistent before 1945, our data is perfectly suited to revisit the evidence for longer horizon predictability.

In the previous section, we document that our results are robust to 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:

$$dp_t \simeq \sum_{h=1}^{H} \rho^{h-1} r_{t+h} - \sum_{h=1}^{H} \rho^{h-1} dg_{t+h} + \rho^H dp_{t+H}. \tag{13}$$

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):

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

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 two-step 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 1686. We are careful in appending the data from different sub-periods 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 expected signs. Ruling out bubbles, all the variation in the dividend-to-price ratio is ultimately driven by the return and dividend growth predictability. Consistent with this notion, the predictive coefficients for returns and dividend growth rates increase with the horizon. At the same time, the extent to which the dividend-to-price ratio simply predicts itself falls with the length of the horizon. All coefficients are significant at usual confidence levels. All in all, the results suggest that, over the last four centuries, returns and dividend growth rates are predictable over both annual and multiannual horizons.

5.2 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 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 parameters estimated for the post-1945 period. We evaluate the “out-of-sample” performance by an out-of-sample R-squared:

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

where x is either return or dividend growth rate. For consistency with the previous analysis, all the variables are demeaned; x^{Actual} denote the realized period-specific demeaned returns or dividend growth rates and $x^{Predicted}$ are the predicted returns or dividend growth rates.[19] 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 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. For a comparison, we also calculate the usual in-sample R-squares.

Results reported in Table 5 suggest that we can indeed use the estimated parameters based on the pre-1945 data to predict returns in the post-1945 period with an out-of-sample R-square of 3.8%.[20] This is sizable, considering that the comparable in-sample R-square for the recent period is 6.5%. Similarly, using the post-1945 period estimates, we obtain an out-of-sample R-square of 4.6% for predicting returns in the earlier part of the sample. This is almost the same as the in-sample R-square for that period.

The return predictability results, however, do not carry over to dividend growth predictability. While the in-sample R-square in the early data is 15.0%, 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. We also obtain a very small out-of-sample R-

---

[19] Note that demeaning introduces a slight look-ahead bias. However, it 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.

[20] If we estimate parameters for the pre-1945 by demeaning the dividend-to-price ratio period by period, the out-of-sample R-square for the post-1945 period increases to 5.0%.

square (0.08%) when using estimated parameters from the recent period to predict dividend growth rates 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, and very close to zero 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 20th century.

5.3 Predictability over the business cycle

An extensive literature argues that market risk premia are countercyclical and business cycle variation drives return predictability (Campbell and Cochrane 1999; Menzly, Santos, and Veronesi 2004; Mele 2007; 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 through 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 through 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. By period, we split the sample into 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 two step GMM 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 (1945-2014), and the full period (1700-2014). It is important to note that in the early period 46% of the years were in recessions, whereas only 17% 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. Return predictability is consistently stronger during recession periods. While the estimated parameter is always positive, it is much higher during recessions and always significant at the one percent level, whereas it is insignificant in expansions. The difference in parameters is significant for the full sample and the recent U.S. period. It is (borderline) insignificant for the early period. Dividend growth rates also tend to be somewhat more predictable in recessions than during expansions, but the difference between expansions and recessions is economically less pronounced and never statistically significant.[21] As expected, we find no evidence for dividend growth predictability in the recent period. Finally, the dividend-to-price ratio is always less persistent in recessions than in expansions. This difference is significant in the pre-1945 period as well as in the full sample, but insignificant in the post-1945 period.

The results provide compelling evidence for the idea that return predictability is driven by countercyclical discount rates. In recessions, stock prices fall more than dividends, suggesting

---

[21] 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.

that the equity premium goes up. Most of the return predictability that we find in the data as a whole originates from the subsequent mean-reversion of increased risk premia. The predictability of cash flow growth is also stronger in recession years, suggesting that expected dividend growth rates adjust more in recessions. However, this difference is economically smaller and statistically insignificant.

6. Reconciling the evidence

Despite the many institutional changes that the markets have undergone throughout the last four centuries, the recent U.S. market shares many important features with earlier periods. Average returns are comparable throughout all four centuries, and returns are predictable across all the periods. Furthermore, return predictability is concentrated in recessions and is remarkably stable. At the same time, there are a number of differences. First, the dividend-to-price ratio has decreased considerably after 1945. 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. Starting point is the prominent decline in the dividend-to-price ratio in recent decades. We consider three possible explanations for this phenomenon: (1) higher valuations or a fall in expected returns, (2) the substitution of dividends for repurchases, and (3) higher reinvestment of firm profits. After reviewing the evidence it seems that (3) is most important in understanding the long term trend in the dividend-to-price ratio. We then argue that, if firms are retaining more of their earnings, the duration of the stock market increases. This provides a natural explanation for the dominance of discount rate news after 1945.

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

Figure 1 shows that the dividend-to-price ratio has decreased significantly after 1945. There are at least three possible explanations.

First, the valuation of firms may have changed, with expected returns falling below historical levels. The empirical evidence, however, is not consistent with this hypothesis. Table 1 shows that average returns (adjusted for inflation) are surprisingly stable over time at around 6%. Compared to the pre-1945 period, the U.S. Sharpe ratio has actually increased somewhat in recent decades.

Second, firms can use share repurchases to pay out earnings to investors. This would decrease dividends and increase prices. Though important in recent years, repurchases cannot explain the long term decline in the dividend-to-price ratio. Before 1982, repurchases played only a minor role (Fama and French 2001).[22] Nevertheless, between 1945 and 1982, the S&P 500’s dividend-to-price ratio dropped from an average of 5.3% for the pre-1945 period to an average of 4.2. This decrease has to be related to factors other than repurchases. Furthermore, data for the CRSP index from Larrain and Yogo (2008) show that, adjusting for repurchases, there is still a decline in dividend yields in the recent period. Dividends plus equity repurchases as a fraction of market equity were on average 5.5% between 1927 and 1945, dropping to 4.6% between 1946 and 2004. In comparison, the average CRSP dividend-to-price ratio payout was 5.1% between 1927 and 1945 and 3.6% between 1946 and 2004.

---

[22] In 1982, the SEC changed its rules on manipulative trading, which allowed firms to reduce the tax burden on investors by substituting dividends for repurchases that are effectively taxed at a lower rate (Grullon and Michaely 2002, Boudoukh, Michaely, Richardson, and Roberts 2007).

Note also that it is likely that adjusting for repurchases understates the decline in the dividend-to-price ratio. To begin with, not all repurchases should be counted as payout to investors. Share buybacks often arise from employee stock compensation or are used to finance mergers and acquisitions.[23] Fama and French (2001) estimate that only half of all repurchases are actual payouts to investors. In addition, repurchases are the exact opposite of share issuances and it seems natural to consider the two together, especially since firms often issue and repurchase shares in the same fiscal year (Farre-Mensa, Michaely, and Schmalz 2015). Data from Larrian and Yogo (2008) indicate that net equity payout (dividends plus repurchases minus issuances as a fraction of market equity) decreased from 4.0% between 1927 and 1945 to as low as 1.7% between 1946 and 2004.[24] All in all, the evidence suggests that it is unlikely that the long term decline in the dividend-to-price ratio is purely the result of repurchases (see also Campbell and Shiller 2005). In Section 6.3 we study whether increased repurchases after 1982 could still play an important role in explaining the recent dominance of discount rates.

Third, it is possible that firms simply retain more of their earnings to re-invest in the firm rather than distribute dividends to investors. The evidence in Fama and French (2001) suggests that this is an important factor to understand long-term trends. To illustrate this point, Figure 4 documents the dividend-to-earnings ratio for the first part of our sample until 1809 and for the U.S. since 1871 (data for the intermediate period are unavailable).[25] 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

---

[23] In addition, Hong and Wang (2008) provide evidence that companies engage in stock buybacks to provide liquidity in times of distress. Almeida, Fos, and Kronlund (2016) show that companies use repurchases strategically to meet analyst EPS forecasts.

[24] Data from Boudoukh, Michaely, Richardson, and Roberts (2007) yields a similar picture.

[25] 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 U.S. period (1871-2012), the data are from Robert Shiller’s webpage.

towards the end of the 18th century.[26] Companies in the early U.S. 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 U.S. period, the payout ratio decreased from around 0.8 at the end of the 19th century to close to 0.4 nowadays.[27]

Reviewing the empirical evidence suggests that the long term decline in the dividend-to-price ratio is importantly driven by higher rates of reinvestment on part of firms. As dividends are substituted for investment, current dividends correspond to a smaller fraction of total market valuation, and the dividend-to-price ratio falls. As a result, the stock market as a whole has become much more growth oriented in the recent period. Fama and French (2001) point to the fact that starting in the 1960s many more small and growth oriented companies have issued shares. At the same time, even large, profitable firms have decreased investor payout and increased investment. Consistent with these developments, the co-movement of growth stocks with the market has increased in recent decades (Campbell and Vuolteenaho 2004).

There are a number of explanations for this pattern. First of all, for most of the 17th, 18th, and 19th century, income taxes were negligible, and taxes did not affect dividend decisions. The U.S. 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. period 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

---

[26] There is a big spike in the payout ratio between 1782 and 1794. This is due to the fact that the VOC 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 observe a smooth decline in dividends-to-earnings after 1750.

[27] Fama and French (2001) estimate that for the period 1982-1998 this number lies close to 0.6, after accounting for repurchases.

decrease dividend payments and reinvest free cash-flows. 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.[28] Yet another reason for a decrease in dividends may be improved transparency due to more detailed and frequent reporting requirements. This may have decreased the role of dividends as signals. Relatedly, investors' preference for dividends (the "bird in the hand effect") may have decreased, with companies optimally responding by lowering dividend payments.

6.2 Increased duration of the stock market

The reduction in dividend payouts suggests that the nature of companies has changed in the recent period. Compared to the past, companies nowadays reinvest a larger share of their earnings and are thus postponing the distribution of cash flows into the future. Also, they often do not distribute cash flows for several years, sometimes even decades, after they are publicly listed. All this suggests that the stock market nowadays consists of longer-lived higher duration companies. We hypothesize that this transformation is an important driver of the differences that we observe between the pre- and post-1945 periods.

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 and become longer-lived assets, future dividend payments become relatively more important than current dividend payments. This has three implications that are consistent with our empirical observations. First, as we already observed, the dividend-to-price ratio decreases with duration (see also Lettau and Wachter 2007). Second,

---

[28] Note that, as discussed further in Section 6.3, dividends were relatively smooth in the 18th century as well.

for the same reason, the dividend yield decreases and capital appreciation becomes a more important part of total returns. Third, if expected returns are highly persistent, an increase in stock duration may also increase the importance of discount rate news. The intuition is as follows. In a present value model, changes in the current price (or more formally: the dividend-to-price ratio) are a function of shocks to expected returns and dividend growth rates. The more persistent these shocks, the bigger their impact on the current valuation (Binsbergen and Koijen 2010). When duration increases, the persistence of these shocks becomes even more important. In particular, compared to assets with short duration, the price of a longer-lived asset will not be very sensitive to transient shocks to the expected return or dividend growth rate. Persistent shocks, on the other hand, will have a more pronounced impact because they effectively last for a larger number of periods. Importantly, the literature suggests that expected returns are substantially more persistent than expected dividend growth rates (Binsbergen and Koijen 2010; Van Nieuwerburgh and Koijen 2012; Golez 2014). In the recent period, the estimated persistence of expected returns is typically over 0.9, whereas the persistence of expected dividend growth rates is below 0.5. This suggests that an increase in stock duration should lead to a relatively more important role of discount rate news and a relatively less important role of cash flow news. This is exactly what we find in the data. To illustrate these effects, we derive a simple stylized model.

6.2.1 Duration in a present value model

We assume that the stock market consists of a collection of projects or assets. These can be thought of as individual companies. Alternatively, a single company can feature multiple projects. Each asset has a specific cash flow profile. To fix ideas, we think of the early trading companies of the 17th century as a collection of (overlapping) short duration projects. The VOC and EIC sent out expeditions to Asia that lasted around two years. Upon return of the fleet, companies paid out virtually all profits to shareholders. In contrast, we think of current (Tech) companies as long duration assets: current dividends are low (or non-existent) and most of the value of the firm depends on future payouts.

To keep the analysis tractable, we simplify and assume that dividend growth rates are stationary but that each project ends after a finite number of periods. That is, assets pay out dividends for a fixed number of periods $\tau = T - t$.[29] A larger τ implies that future dividends are more important. We assume that expected returns μ_t and expected dividend growth rates g_t are common to all projects and follow an AR(1) process (Binsbergen and Koijen 2010):

$$\begin{aligned} \mu_{t+j} &= \delta_0 + \delta_1(\mu_{t+j-1} - \delta_0) + \varepsilon^{\mu}_{t+j}, \\ g_{t+j} &= \gamma_0 + \gamma_1(g_{t+j-1} - \gamma_0) + \varepsilon^{g}_{t+j}, \end{aligned} \tag{19}$$

where $\mu_t = E_t(r_{t+1}),\ g_t = E_t(\Delta d_{t+1}),\ \Delta d_{t+1} = g_t + \varepsilon^d_{t+1}$, and the error terms follow a multivariate normal distribution. Parameters δ_1 and γ_1 capture the persistence of shocks.

When we log-linearize an asset's returns around the long term trend of the dividend-to-price ratio, the actual dividend-to-price ratio for an individual project *i* can be approximated as:

$$dp_{i,t}(\tau) \simeq C + D\left(\delta_0 - \gamma_0\right) + \left(\mu_t - g_t\right) + D_\mu\left(\mu_t - \delta_0\right) - D_g\left(g_t - \gamma_0\right), \tag{20}$$

where $C = -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k}$, $D = \sum_{j=1}^{\tau-1} \prod_{k=1}^{j} \rho_{t+k}$, $D_\mu = \sum_{j=1}^{\tau-1} \delta_1^j \prod_{k=1}^{j} \rho_{t+k}$, and $D_g = \sum_{j=1}^{\tau-1} \gamma_1^j \prod_{k=1}^{j} \rho_{t+k}$;

$\rho_{t+1} = \dfrac{\exp(-\overline{dp}_{t+1})}{1+\exp(-\overline{dp}_{t+1})}$, and $\kappa_{t+1} = \log(1+\exp(-\overline{dp}_{t+1})) + \rho_{t+1}\overline{dp}_{t+1}$ (details are in Appendix C).

---

[29] Our focus on duration is similar in spirit to Lettau and Wachter (2007), who relate cross-sectional differences in stock duration to the value premium. However, in their approach the duration of the stock market as a whole is effectively kept constant. Moreover, we focus purely on testing present value relations and do not take a stance on which risks are priced in the economy.

Eq. (20) captures the main intuition of our model. In particular, changes in expected returns and expected dividend growth rate affect the dividend-to-price ratio directly as well as through D_μ and D_g. Both terms are increasing in τ and the persistence parameters. As long as $\delta_1 \neq \gamma_1$, an increase in τ changes the relative importance of discount rate and cash flow news. If, as suggested by the existing empirical evidence, expected returns are more persistent than dividend growth rates $\left(\delta_1 > \gamma_1\right)$, the relative contribution of expected returns to price movements increases with τ. Using parameter values from Binsbergen and Koijen (2010), we show in cross-sectional simulations in Appendix B that the contribution of discount rates to price movements is increasing and concave in duration.[30] Consistent with the argument that long-duration assets have a lower dividend-to-price ratio, we also show that the average dividend-to-price ratio is decreasing and convex in duration.

Next, we provide time-series evidence at the aggregate market level. We assume that the market consists of multiple finitely lived assets. All projects are exposed to the same shocks to expected returns and dividend growth rates. They only differ in the number of periods they have left until they die. Every year, one project ends and is replaced by a new one that will last τ periods. To preserve the same proportions of assets in different stages, we assume that the economy consists of τ projects. Because a project's market cap falls with the time it has left, we use value-weights. As before, we apply the parameters from Binsbergen and Koijen (2010) and let the simulations run for 10,000 years. Details are in Appendix C.

We consider three cases. We first assume that the market consists of very long lived assets, that is we set $\tau = 100$. Next, we keep the same parameter values, but we decrease τ to 40

---

[30] Parameters from Binsbergen and Koijen (2010) are based on a present value model applied to the (nominal) S&P 500 index between 1946 and 2007 (Table VI, p. 1463).

and then to 10 years. Table 7 presents the main results. As expected, for very long lived assets, the simulation results match the recent period. Returns come mainly from capital appreciation. The dividend-to-price ratio is low, highly persistent, and mostly predicts returns. The predictive coefficient on the dividend growth rate is close to zero, which implies that price movements are dominated by discount rate changes. When we decrease τ, an increasingly smaller part of returns comes from capital appreciation. The dividend-to-price ratio decreases and becomes less persistent. At the same time, the predictive coefficient on the dividend growth rate increases in magnitude more than the predictive coefficient on returns, which implies that an increasingly large part of price movements is due to cash flow news. While we do not exactly match the economic magnitudes in the data, our simulations confirm all the sign predictions. Thus, an increase in stock market duration due to longer-lived assets can largely explain the differences between the pre- and post-1945 periods.

6.2.2 Further empirical evidence

Motivated by this model, we next attempt to verify empirically 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. One minus this fraction is a proxy for duration dur_t. For simplicity, we assume that dividend growth rates follow a random walk with drift $\overline{g}$. Our duration measure is defined as:

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

As proxies for $\overline{g}$ and $\overline{r}$, we use the average per-period dividend growth rate and market return. Note that our measure for duration can be interpreted as the dividend-to-price ratio adjusted for drift $\overline{g}$ and expected returns $\overline{r}$.[31]

We first calculate duration separately for each of the four periods within our sample. The average values for stock duration are 0.59 for Netherlands/U.K., 0.62 for U.K., 0.58 for the early U.S. period, and 0.72 for the recent U.S. period. As expected, stock duration increased in recent years. The difference in duration between the earlier and the post-1945 period is statistically significant with a *t*-statistic of 2.78.[32]

To better understand the increased importance of discount rate news in the recent period, we next explore the time-varying relationship between duration and discount rates in the combined U.S. period using a rolling windows approach. We set the length of the window to 75 years to match the pre-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 expected returns are estimated within the same window). Figure 5 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).

Duration and the discount rate component are increasing over time and are positively correlated. A regression of discount rate news on a constant and duration yields an adjusted R-squares of 55% (based on Panel A) and 71% (based on Panel B). Thus, duration appears to explain a large part of discount rate news variation. Note also that the increasing duration in the

---

[31] The results are similar if we use the dividend-to-price ratio in place of the *Dur* variable.

[32] 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.

U.S. period is consistent with the reduction in the dividend-to-earnings ratio documented in Figure 4.[33]

6.3 Discussion

We find strong support for our duration hypothesis. At the same time, other explanations may also play a role.

Section 6.1 documents that in 1982 firms have started using share repurchases to pay out earnings to investors. This is unlikely to explain our results. First, the presence of repurchases does not invalidate our predictability and decomposition results. 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.[34] In this framework, repurchasing shares does not automatically increase the importance of discount rates. The intuition is that the expected dividend growth rate per share increases for the shares that remain outstanding. As a result, the dividend-to-price ratio becomes more sensitive to both discount rate and cash flow news.[35] Second, in the Online Appendix (Table OA.6) we replicate our predictability and decomposition results for the years between 1945 and 1982 when repurchases were only of minor importance. The results are similar to the entire post-1945 period. If anything, return predictability appears even stronger and discount rate news more important.

---

[33] If we regress discount rate news on the payout ratio within our rolling window approach, we obtain adjusted R-squares between 74% and 81%.

[34] 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. Moreover, the log-linearized present value model cannot accommodate negative net payouts.

[35] We thank John Campbell and Xavier Gabaix for pointing this out.

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 0.56 in the pre-1945 period and 0.23 after 1945. In other words, compared to earnings, dividends were far less volatile in the recent period. However, dividend smoothing is not unique to the recent period: our data indicates that for the period 1686-1809 the dividend smoothing parameter was 0.22, virtually the same as in the recent U.S. period.[36]

Dividend smoothing and stock market duration are potentially interrelated. As we noted before, when companies retain more earnings, it is easier to smooth dividends. 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 75 years rolling windows that we use to estimate the discount rate news components. Regressing the discount rate news component on the smoothing parameter gives an expected negative parameter estimate. The adjusted R-squared is 8% or 14%, depending on whether we measure the discount rate news component of the dividend-to-price ratio or unexpected returns. As noted above, using our measure for duration yields an adjusted R-square of 55% or 71%. When we include both explanatory variables, the adjusted R-square is 54% or

---

[36] 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). If we exclude the first years with highly volatile dividends and focus on the period 1721-1809, the smoothing parameter is even lower at 0.13.

71%. Thus, while both explanations play a role, the marginal contribution of duration in explaining the discount rate news variation outweighs the importance of dividend smoothing.

An alternative explanation for the drop in the importance of cash flow news could be related to the fact that after 1945 the market consisted of many more companies than before. In particular, if firm level cash flow news is mostly idiosyncratic and discount rate news is largely systematic, as it appears to be the case in the recent period (Vuolteenaho 2002; Chen, Da, and Zhao 2013), an increase in the number of companies in the index would decrease the role of cash flow news at the market level. This is less likely to apply to the 19th and early 20th century because the (baseline) market indices for these periods include between 50 and 258 individual stocks. Moreover, when we use the London market up to 1900 and rely on the CRSP index (instead of the S&P 90/500) after 1925, the data cover many more companies, but cash flows news remains equally important (details in Online Appendix, Table OA.4). The limited number of companies in the early Dutch/English period, however, might still induce a positive bias in the importance of cash flow news. To address this concern, we repeat our analysis for this period at the individual stock level. We include four stocks with uninterrupted annual dividend payments (VOC: 1686-1782; BOE: 1697-1809; EIC: 1700-1809; SSC: 1721-1809). Interestingly, even though returns are predictable for all four companies at the individual stock level, dividend growth rates are predictable in only three out of four cases. Moreover, we find that the average value for the discount rate versus cash flow news is comparable to the evidence for the market as a whole (67% vs. 33% when the decomposition is based on the dividend-to-price ratio and 19% vs. 81% when the decomposition is based on unexpected returns). This suggests that diversification across stocks in the early period does not reduce the importance of cash flow news, and that the number of companies included in the index is unlikely to drive our results.

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, may play a role as well.[37]

7. 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 robustly predicts future returns, both over annual and multi-year horizons. Our findings suggest that return predictability is a deep characteristic of financial markets, and reflects fundamental changes in the underlying economy, rather than short-lived shocks. Consistent with recent theoretical work, the predictability of returns is importantly driven by the business cycle. The equity premium tends to rise during recessions and this strongly predicts higher returns going forward. Fluctuations in the dividend-to-price ratio have significantly less predictive power during expansions.

Despite the similarities in return predictability across four centuries, there are also important differences between periods. In particular, whereas dividend growth rates are highly predictable before 1945, this relationship breaks down in the second half of the 20th century. This

---

[37] 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 (Koijen and Van Nieuwerburgh 2011).

suggests that the dominance of discount rate news documented in the literature is only a recent phenomenon. We argue that this shift is consistent with increased duration of the stock market in recent years. As firms retain more of their earnings, and reinvest rather than pay out dividends, the duration of the stock market increases. As long as expected returns are more persistent than expected dividend growth rates, this will increase the sensitivity of asset prices to shocks in the discount rate. Linking “excess volatility” to stock market duration has important implications for asset pricing theory. Most importantly, it suggests that the importance of time-varying expected returns depends on firms’ investment and payout decisions and the underlying uncertainty associated with long term investments.

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.

Almeida, H., V. Fos, and M. Kronlund. 2016. The real effects of share repurchases. Journal of Financial Economics, forthcoming.

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. Engstom, 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.

Bris le, 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.

Campbell, J. Y., and R. J. Shiller. 2005. Valuation ratios and the long run stock market outlook: An update, in Richard H. Thaler, Advances in Behavioral Finance II, Princeton: Princeton University Press [CFDP 1295].

Campbell, J. Y., and T. Vuolteenaho. 2004. Bad beta, good beta. The American Economic Review 94:1249–1275.

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 X. Zhao. 2013. What drives stock price movements? Review of Financial Studies 26:841–876.

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.

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.

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.

Farre-Mensa, J., R. Michaely, and M. Schmalz. 2015. Financing payouts. Working paper. Harvard Business School, Cornell University, and University of Michigan.

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. 2003. Predicting the equity premium with dividend ratios. Management Science 49:639–654.

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.

Grossman, R., S. 2002. New indices of British equity prices, 1870-1913. Journal of Economic History 62:121–146.

Grullon G., and Michaely R. 2002. Dividends, share repurchases, and the substitution hypothesis. *Journal of Finance* 57:1649–84.

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.

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.

Lettau, M., and J. A. Wachter. 2007. Why is long-horizon equity less risky? A duration based explanation of the value premium. Journal of Finance 62:55–92.

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.

Mele, A. 2007. Asymmetric stock market volatility and the cyclical behavior of expected returns. Journal of Financial Economics 86:446–478.

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 49:1255-1277.

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.

Santa-Clara, P., and P. Maio. 2015. Dividend yields, dividend growth, and return predictability in the cross-section of stocks. Journal of Financial and Quantitative Analysis 50: 33–60.

Schwert, G. William, 1990. Indexes of U.S. Stock Prices from 1802 to 1987. Journal of Business 63:399–426.

Schwert, G. William, 2003. Anomalies and Market Efficiency. Chapter 15 *Handbook of the Economics of Finance* eds. George Constantinides, Milton Harris, and René Stulz, North-Holland, 937–972.

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.

Vuolteenaho, T. 2002. What drives firm-level stock returns? Journal of Finance 57:233–264.

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.

Table 1: Summary statistics: Annual data

This table reports summary statistics for the annual variables, based on nominal data (Panel A), and adjusted for inflation (Panel B). Column (1) reports the statistics for the 1686-1809 period based on the Netherlands/U.K.. data (because dividends were not paid out every year before 1686, 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. 1686-1809	Neth./U.K. 1629-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2014	Full period
Panel A: Nominal data						
r	0.055	0.058	0.069	0.069	0.103	0.071
$Cap.\ app.\ (ca)$	0.009	0.011	0.026	0.018	0.069	0.027
$ca\ /\ r$	0.163	0.193	0.373	0.253	0.675	0.378
$Std.(r)$	0.093	0.108	0.069	0.190	0.161	0.136
$AR1(r)$	-0.062	-0.085	0.043	0.053	-0.023	0.018
rf	0.032	-	0.034	0.025	0.041	0.032
$Risk\ premium$	0.023	-	0.035	0.044	0.061	0.039
$Sharpe\ ratio$	0.249	-	0.508	0.232	0.382	0.283
dg	0.012	-	0.034	0.013	0.059	0.026
$Std.(dg)$	0.173	-	0.100	0.157	0.066	0.143
$AR1(dg)$	-0.065	-	0.182	0.204	0.413	0.071
dp	-3.086	-	-3.126	-2.964	-3.473	-3.148
$Std.(dp)$	0.241	-	0.167	0.253	0.442	0.343
$AR1(dp)$	0.682	-	0.659	0.512	0.908	0.829
$ADF(dp)$	-3.971**	-	-4.638***	-4.765***	-2.174	-4.753***
Panel B: Real data						
r	0.051	0.056	0.075	0.065	0.065	0.061
$Cap.\ app.\ (ca)$	0.005	0.009	0.031	0.013	0.032	0.017
$ca\ /\ r$	0.107	0.163	0.418	0.196	0.491	0.276
$Std.(r)$	0.114	0.130	0.107	0.186	0.171	0.146
$AR1(r)$	0.001	-0.080	0.114	0.013	0.025	0.014
dg	0.009	-	0.039	0.008	0.022	0.016
$Std.(dg)$	0.184	-	0.105	0.151	0.069	0.146
$AR1(dg)$	-0.030	-	0.165	0.074	0.372	0.037
dp	-3.086	-	-3.126	-2.964	-3.473	-3.148
$Std.(dp)$	0.241	-	0.167	0.253	0.442	0.343
$AR1(dp)$	0.682	-	0.659	0.512	0.908	0.829
$ADF(dp)$	-3.971**	-	-4.638***	-4.765***	-2.174	-4.753***

Table 2: Vector autoregression (VAR) estimates: Annual data (1686-2014)

Panel A reports VAR estimates predicting real returns, real dividend growth rates, and the dividend-to-price ratio with the first order lags of the same variables. The data are annual. The model is estimated by two-step 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, a Chi-square statistics reports a test 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. 1686-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2014	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.188***	0.244**	0.139	0.102***	0.105***
	(0.041)	(0.107)	(0.088)	(0.033)	(0.029)
	[4.654**]	[1.663]	[0.333]	[2.600]	
r_t	-0.002	0.180*	0.118	0.004	0.058
	(0.087)	(0.105)	(0.169)	(0.086)	(0.060)
dg_t	-0.023	-0.003	-0.265*	0.222	-0.011
	(0.046)	(0.119)	(0.145)	(0.175)	(0.058)
Dep. variable: dg_{t+1}					
dp_t	-0.208**	-0.209*	-0.282***	-0.008	-0.088***
	(0.087)	(0.110)	(0.074)	(0.019)	(0.031)
	[2.609]	[1.165]	[7.340***]	[18.140***]	
r_t	0.247**	-0.194	0.270***	0.128*	0.196***
	(0.117)	(0.199)	(0.102)	(0.069)	(0.064)
dg_t	0.008	0.290**	0.046	0.323***	0.007
	(0.105)	(0.122)	(0.089)	(0.077)	(0.069)
Dep. variable: dp_{t+1}					
dp_t	0.631***	0.571***	0.609***	0.917***	0.841***
	(0.113)	(0.104)	(0.100)	(0.045)	(0.050)
	[5.002**]	[5.755**]	[6.155**]	[13.167***]	
r_t	0.260**	-0.391**	0.160	0.128	0.144
	(0.132)	(0.189)	(0.214)	(0.124)	(0.090)
dg_t	0.032	0.307**	0.327*	0.104	0.019
	(0.103)	(0.140)	(0.168)	(0.154)	(0.090)
J-test	1.595	0.054	1.806	0.721	11.152
Panel B: Decomposition results					
Dividend-to-price					
DR	0.538	0.483	0.413	1.059	0.613
CF	0.462	0.517	0.587	-0.059	0.387
Unexpected returns					
DR	0.113	0.255	0.313	0.798	0.375
CF	0.889	0.746	0.693	0.202	0.625

Table 3: Vector autoregression (VAR) estimates: Triennial data (1629-2014)

Panel A reports VAR estimates predicting real returns, real dividend growth rates, and the dividend-to-price ratio with the first order lags of the same variables. The data are triennial. Reported are the means of the estimated parameters and statistics estimated on three consecutive non-overlapping samples. The VAR model is estimated by two-step 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. 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)
		1629-1945	1945-2014	1629-2014
Panel A: VAR estimates				
	Dep. variable: r_{t+1}			
	dp_t	0.222***	0.252**	0.182***
		(0.080)	(0.099)	(0.062)
	r_t	-0.061	0.132	-0.014
		(0.081)	(0.125)	(0.088)
	dg_t	-0.092	-0.285	-0.070
		(0.062)	(0.696)	(0.059)
	Dep. variable: dg_{t+1}			
	dp_t	-0.618***	-0.030	-0.443**
		(0.175)	(0.041)	(0.175)
	r_t	0.159	0.065	0.281**
		(0.142)	(0.068)	(0.116)
	dg_t	-0.039	-0.209	-0.125
		(0.124)	(0.212)	(0.124)
	Dep. variable: dp_{t+1}			
	dp_t	0.182	0.780***	0.424**
		(0.142)	(0.101)	(0.198)
	r_t	0.251*	-0.073	0.332**
		(0.134)	(0.138)	(0.138)
	dg_t	0.061	0.082	-0.061
		(0.105)	(0.575)	(0.124)
	J-test	1.822	0.256	4.619
Panel B: Decomposition results				
	Dividend-to-price			
	DR	0.274	0.958	0.339
	CF	0.726	0.042	0.661
	Unexpected returns			
	DR	0.146	0.767	0.160
	CF	0.859	0.231	0.830

Table 4: Longer horizon predictability

This table reports 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 two-step 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 1686 through 2014.

	(1)	(2)	(3)
Horizon (H)	1 year	3 years	5 years
	1686-2014	1686-2014	1686-2014
Panel A: Overlapping observations			
Dep. variable: $\sum_{h=1}^{H}\rho^{h-1}r_{t+h}$			
dp_t	0.092***	0.226***	0.345***
	(0.029)	(0.060)	(0.082)
Dep. variable: $\sum_{h=1}^{H}\rho^{h-1}dg_{t+h}$			
dp_t	-0.141***	-0.264***	-0.300***
	(0.048)	(0.089)	(0.115)
Dep. variable: dp_{t+H}			
dp_t	0.800***	0.579***	0.439***
	(0.065)	(0.123)	(0.159)
J-test	1.244	2.013	2.009
Panel B: Non-overlapping observations			
Dep. variable: $\sum_{h=1}^{H}\rho^{h-1}r_{t+h}$			
dp_t	0.092***	0.241***	0.357***
	(0.029)	(0.084)	(0.106)
Dep. variable: $\sum_{h=1}^{H}\rho^{h-1}dg_{t+h}$			
dp_t	-0.141***	-0.262**	-0.296**
	(0.048)	(0.108)	(0.144)
Dep. variable: dp_{t+H}			
dp_t	0.800***	0.564***	0.428**
	(0.065)	(0.150)	(0.186)
J-test	1.244	2.101	2.074

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 real returns and annual real 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 1686-1945 (1945-2014) is based on the estimated parameters from the period 1945-2014 (1686-1945). Results are based on the reduced form VAR model where returns and dividend growth rates are predicted using lagged dividend-to-price ratio only.

	In-sample [out-of-sample] R-square			
	1686-1945		1945-2014	
	Returns	Dividend growth	Returns	Dividend growth
Parameters from:				
1686-1945	0.047	0.150	[0.038]	[-2.575]
1945-2014	[0.046]	[0.008]	0.065	0.002

Table 6: Predictability over the business cycle

This table reports summary statistics (Panel A) and predictability results (Panel B) over the business cycle. All the data are in real terms. Estimated parameters for expansions and recessions are estimated jointly using two-step 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 1945-2014, or 1700-2014.

	(1)	(2)	(3)	(4)	(5)	(6)
	1700-1945		1945-2014		1700-2014	
	Recessions	Expansions	Recessions	Expansions	Recessions	Expansions
Panel A: Summary statistics						
r	0.039	0.081	-0.050	0.090	0.030	0.084
$Std.(r)$	0.154	0.124	0.218	0.150	0.162	0.133
t-stat	[-2.246] **		[-2.119] *		[-2.999] ***	
dg	0.001	0.022	-0.010	0.029	0.000	0.024
$Std.(dg)$	0.138	0.124	0.110	0.056	0.135	0.107
t-stat	[-1.191]		[-1.206]		[-1.638]	
dp	-3.023	-3.078	-3.270	-3.515	-3.048	-3.215
$Std.(dp)$	0.247	0.208	0.463	0.429	0.284	0.358
t-stat	[1.791] *		[1.689]		[4.467]***	
No. of years	105	124	12	57	117	181
Panel B: VAR estimates						
Dep. variable: r_{t+1}						
dp_t	0.203***	0.073	0.238***	0.050	0.151***	0.022
	(0.062)	(0.051)	(0.075)	(0.044)	(0.041)	(0.032)
	[2.579]		[4.740] **		[6.248] **	
Dep. variable: dg_{t+1}						
dp_t	-0.248***	-0.195***	0.079	-0.017	-0.151**	-0.084**
	(0.076)	(0.063)	(0.057)	(0.028)	(0.075)	(0.036)
	[0.290]		[2.219]		[0.631]	
Dep. variable: dp_{t+1}						
dp_t	0.574***	0.767***	0.870***	0.961***	0.730***	0.931***
	(0.093)	(0.060)	(0.109)	(0.051)	(0.084)	(0.050)
	[2.995] *		[0.592]		[4.226] **	
J-test	11.178***		0.010		0.710	

Table 7: The effect of stock market duration: Simulations

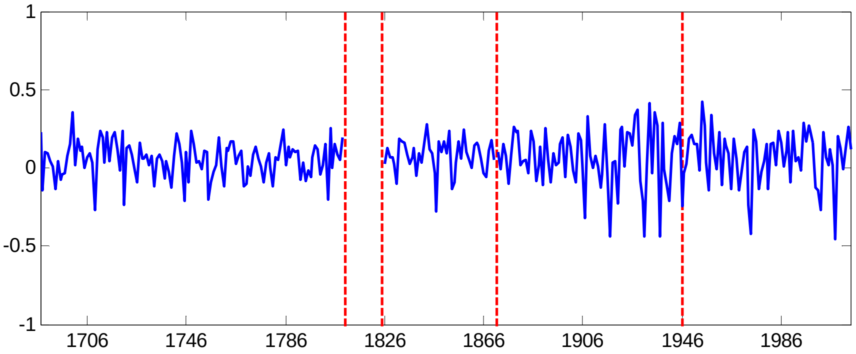
This table reports summary statistics (Panel A), predictability results (Panel B), and decomposition results (Panel C) for simulated data based on our model for a market with finitely lived companies. Each company lives for τ years. The rest of parameters are taken from Binsbergen and Koijen (2010) (Table VI, p. 1463). We let simulations run for 10,000 years. The details are provided in Appendix C.

	(1)	(2)	(3)
	$\tau = 100$	$\tau = 40$	$\tau = 10$
Panel A: Summary statistics			
r	0.091	0.090	0.091
$Cap.\ app.$	0.048	0.019	-0.130
dp	-3.152	-2.627	-1.403
Panel B: VAR results			
Dep. variable: r_{t+1}			
dp_t	0.123	0.148	0.214
Dep. variable: dg_{t+1}			
dp_t	-0.000	-0.010	-0.156
Dep. variable: dp_{t+1}			
dp_t	0.914	0.902	0.785
Panel C: Decomposition of the dividend-to price			
DR	0.999	0.935	0.578
CF	0.001	0.065	0.422

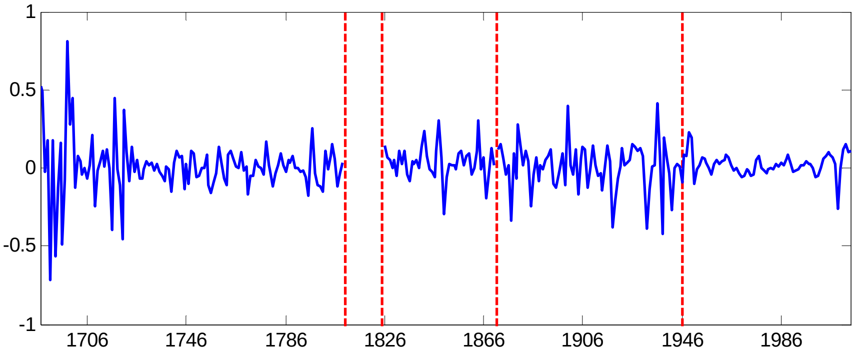
Figure 1: Returns, dividend growth rates, and the dividend-to-price ratio

This figure plots real returns, real dividend growth rates, and the dividend-to-price ratio. The data are annual. Vertical dashed lines denote the time periods: Netherlands and U.K. (1686-1809), U.K. (1825-1870), U.S. early (1871-1945), and U.S. recent (1945-2014).

Panel A: Log annual real returns



Panel B: Log annual real dividend growth rates



Panel C: Log dividend-to-price ratio

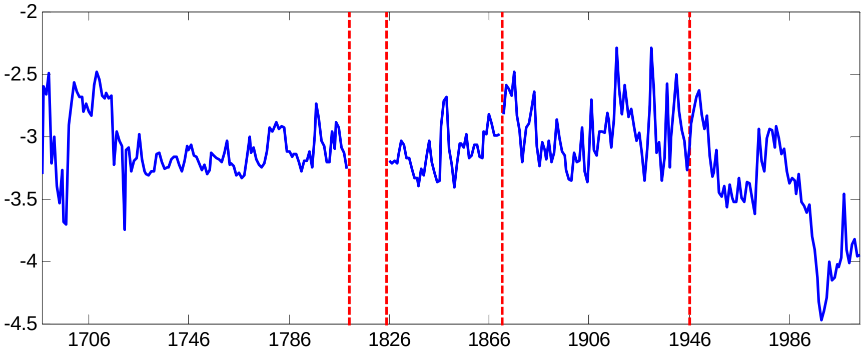
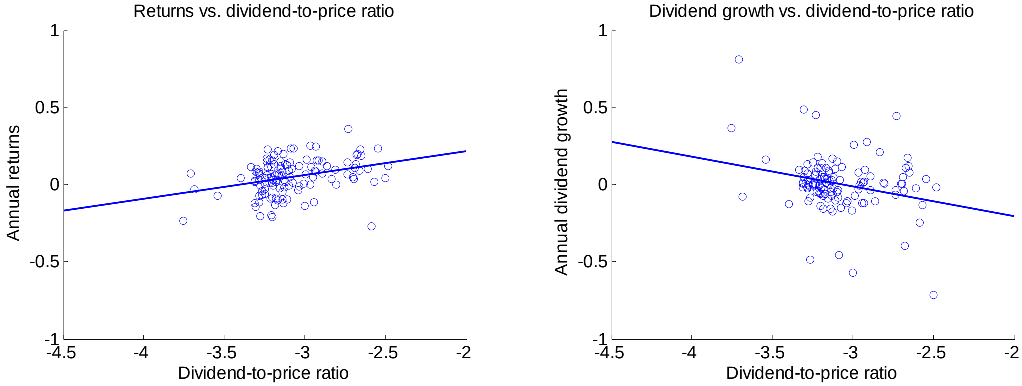


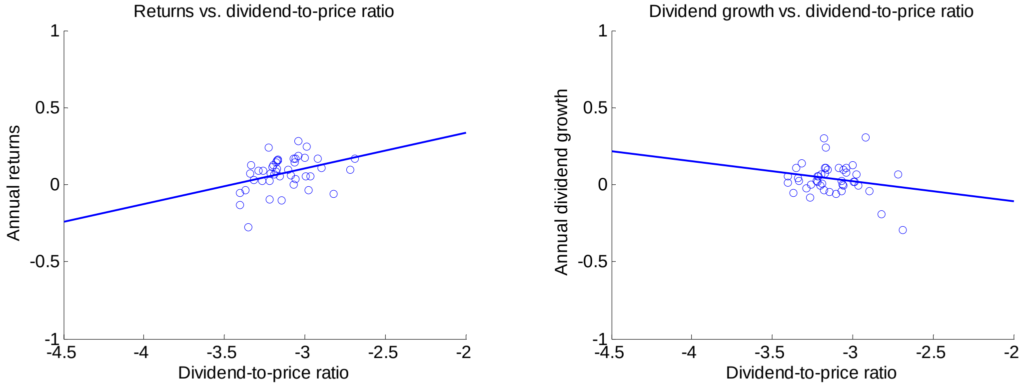
Figure 2: Scatterplots

This figure presents scatterplots of annual real returns (left) and annual real dividend growth rates (right) against lagged values of the dividend-to-price ratio. All variables are demeaned. Panels A through D present scatterplots for each period separately: the Netherlands/U.K. period (1686-1809), the U.K. period (1825-1870), the early U.S. period (1871-1945), and the recent U.S. period (1945-2014). Panel D presents scatterplots for the full period (1686-2014).

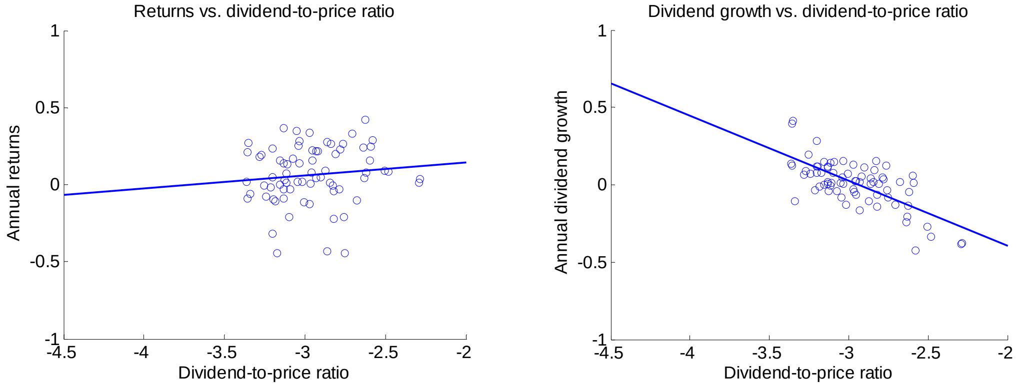
Panel A: Netherlands and U.K.: 1686-1809



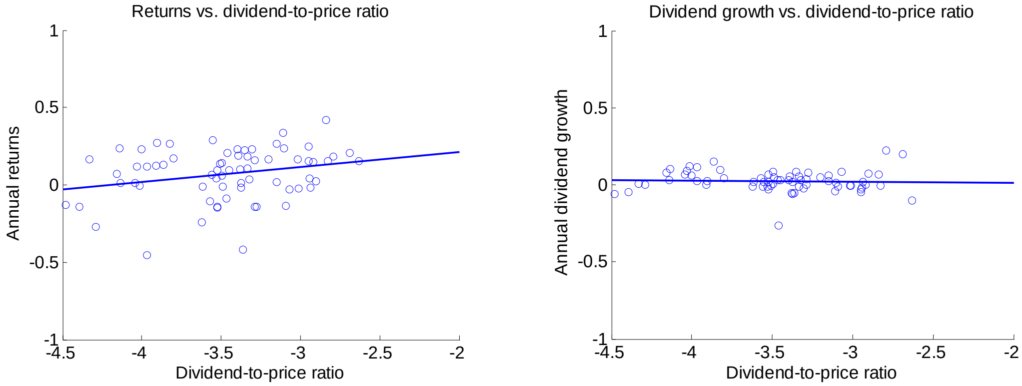
Panel B: U.K.: 1825-1870



Panel C: U.S. early: 1871-1945



Panel D: U.S. recent: 1945-2014



Panel E: Full period: 1686-2014

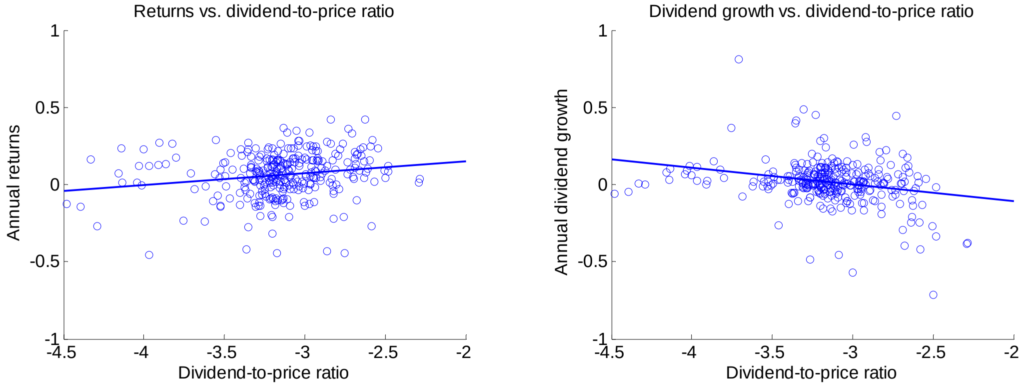
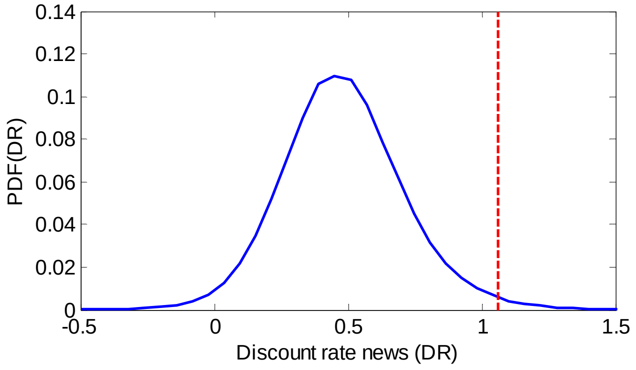


Figure 3: Monte Carlo simulations: Probability density functions (PDF)

This figure plots simulated probability density functions for the discount rate news component in the dividend-to-price ratio (Panel A) and in 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

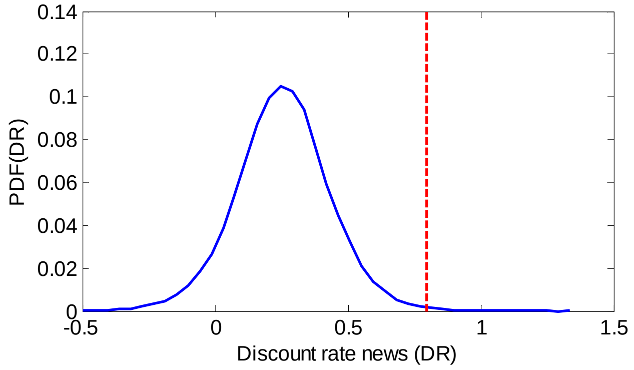


Figure 4: Dividends-to-earnings ratio

This figure plots 20 year rolling averages of the dividends-to-earnings ratio for the Netherlands/U.K. (1686-1809) period and the combined U.S. period (1871-2014).

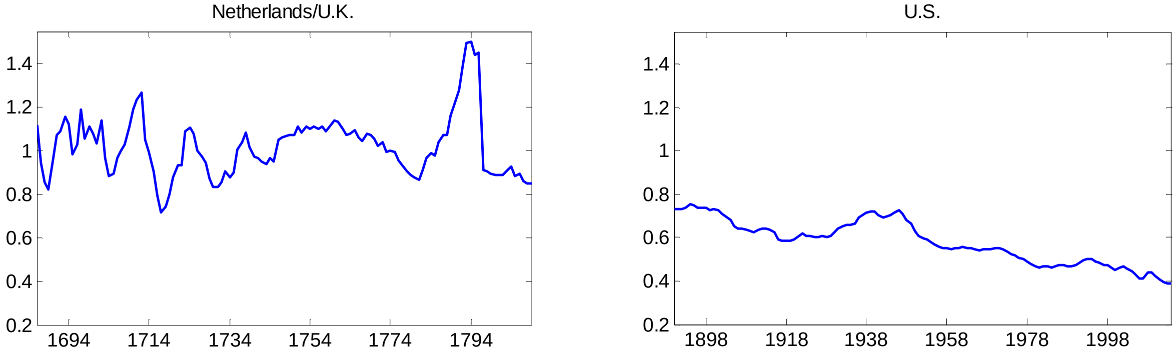
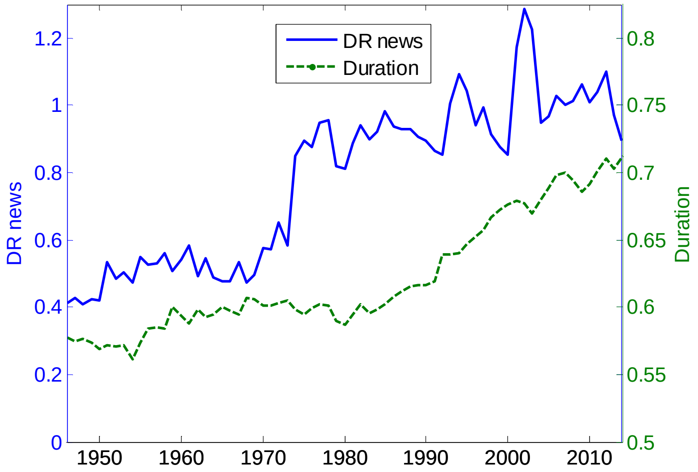


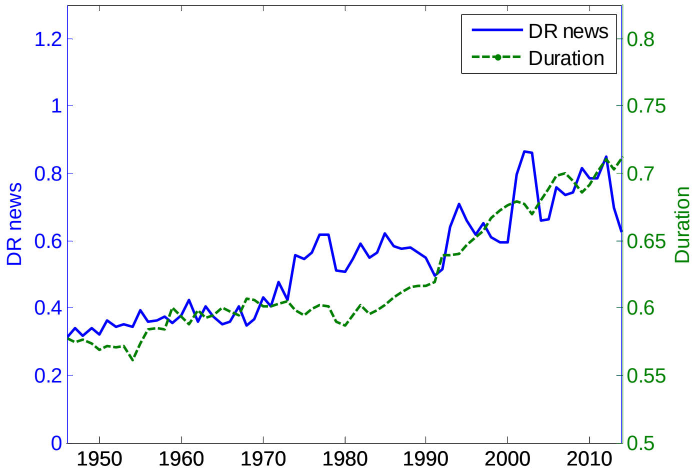
Figure 5: Duration and discount rates: Rolling windows (U.S. period: 1871-2014)

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 annual U.S. data from 1871 through 2014 using rolling windows of 75 years (to match the length of the pre-1945 period). The first rolling window ends in 1945.

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



Panel B: Duration and discount rate component of unexpected returns



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 the world's first publicly traded corporation; its shares were freely tradable, and shareholders enjoyed limited liability. There was a clear 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. The company was nationalized by the government in 1796. The 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.

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

We have information available for three English securities: the Bank of England (BoE), the British East India Company (EIC) and The South Sea Company (SSC). The 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 EIC. This company 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 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 by government debt. It matters very little for our results whether or not we keep the company in our index after 1732.

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 (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.

Appendix B: Small sample bias

In this Appendix, we analyze small sample bias in our return predictability regressions. Following Stambaugh (1999), we depart from the VAR model and consider the system of OLS predictive regressions:

$$\begin{aligned} r_{t+1} &= \beta^{r,OLS} dp_t + \varepsilon^{r}_{t+1}, \\ dp_{t+1} &= \beta^{dp,OLS} dp_t + \varepsilon^{dp}_{t+1}, \end{aligned} \tag{B 1}$$

where all variables are demeaned and errors are multivariate normal. Stambaugh (1999) shows that $\beta^{r,OLS}$ is upward biased when errors are negatively correlated. This leads to over-rejection of the null hypothesis of no return predictability. The bias is largest in small samples and for highly persistent dividend-to-price ratios.

We simulate the system of equations in (B 1) under the absence of return predictability ($\beta^{r,OLS} = 0$) using empirical estimates of the variance-covariance matrix of errors and $\beta^{dp,OLS}$. We always match the length of the simulated data with the sample period. The mean of $\beta^{r,OLS}$ in 100,000 simulated samples represents the magnitude of the bias. The bootstrapped *p*-value is the proportion of estimated $\beta^{r,OLS}$ coefficients in the simulated data that are larger (in absolute value) than the empirically estimated $\beta^{r,OLS}$.

Table B.1 reports results for each sample period separately as well as for the full period. Interestingly, the small sample bias is only important in the post-1945 U.S. sample. In particular, the OLS parameter estimate becomes insignificant when we account for the bias. In all other sample periods, the *p*-values from simple OLS and the bootstrap procedure are largely aligned. In the Netherlands/U.K. and the U.K. period parameters are significant according to both p-values, whereas they are insignificant in the early U.S. period. Most importantly, in the full period, the estimated parameter is highly significant with both *p*-values below one percent. Thus, by extending the period backwards, we alleviate the problem of small sample bias that renders return predictability unreliable in the recent U.S. period. This is driven by many more observations (308 vs. 68), lower correlation of errors (-0.66 vs. -0.91), and less persistence of the dividend-to-price ratio (0.84 vs. 0.92). [38]

---

[38] Small sample bias could also affect dividend growth predictability, if innovations in the dividend-to-price ratio are positively correlated to errors in the dividend growth predictive regression. We repeat simulations for the dividend growth regressions for the full period and find that such bias does not apply as the both OLS and bootstrapped *p*-values are below one percent.

Table B.1: Small sample bias

This table reports results investigating Stambaugh (1999) bias in the OLS real return predictability regressions. All parameters are based on the system of equations in (B 1). The bias is represented by the mean of the $\beta^{r,OLS}$ coefficient estimated in 100,000 samples that are simulated under no return predictability. The bootstrapped *p*-value is the proportion of estimated $\beta^{r,OLS}$ coefficients in the simulated data that are larger (in absolute value) than the empirically estimated $\beta^{r,OLS}$.

	(1)	(2)	(3)	(4)	(5)
	Neth./U.K. 1686-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2014	Full period
$No.$	122	44	73	68	308
$Corr(\varepsilon^{r}, \varepsilon^{dp})$	-0.275	-0.577	-0.851	-0.912	-0.655
$\beta^{dp,OLS}$	0.685	0.671	0.522	0.924	0.844
$\beta^{r,OLS}$	0.155	0.228	0.085	0.097	0.079
$Bias$	0.004	0.032	0.025	0.053	0.006
$OLS\ p-value\ (\beta^{r,OLS})$	0.001	0.013	0.326	0.031	0.001
$Bootstrapped\ p-value\ (\beta^{r,OLS})$	0.001	0.038	0.336	0.219	0.004

Appendix C: Present value model: The effect of duration

In this Appendix we discuss the present value model with finitely lived assets that we introduced in the main text. We first provide analytical solutions and simulation results for an individual asset (project). Next, we analyze the aggregate market that consists of a collection of finitely lived assets.

C.1 Individual assets

We assume that an asset pays out its last dividend in period $T = t + \tau$. For $j \leq \tau$, expected returns μ_t and expected growth rates g_t follow an AR(1) process (Binsbergen and Koijen 2010):

$$\begin{aligned} \mu_{t+j} &= \delta_0 + \delta_1(\mu_{t+j-1} - \delta_0) + \varepsilon^{\mu}_{t+j}, \\ g_{t+j} &= \gamma_0 + \gamma_1(g_{t+j-1} - \gamma_0) + \varepsilon^{g}_{t+j}, \end{aligned} \tag{C.1}$$

where

$$\begin{aligned} &\mu_t = E_t(r_{t+1}) \text{ and } r_{t+1} = \log\left[\left(P_{t+1} + D_{t+1}\right)/P_t\right], \\ &g_t = E_t(\Delta d_{t+1}) \text{ and } \Delta d_{t+1} = \log\left[D_{t+1}/D_t\right], \\ &\Delta d_{t+1} = g_t + \varepsilon^{d}_{t+1}. \end{aligned} \tag{C.2}$$

The error terms follow a multivariate normal distribution. The log dividend-to-price ratio is $dp_t = \log\left[D_t / P_t\right]$.

Following Campbell and Shiller (1988), we log-linearize returns to arrive at closed-form solutions. For infinitely lived assets, it is natural to log-linearize returns around the long run average of the dividend-to-price ratio. In the case of finitely lived assets, however, the dividend-to-price ratio is predictably decreasing in the number of periods. We thus instead log-linearize around the long run trend of the dividend-to-price ratio. Assuming that expected growth rates and expected returns are at their long run average, the price at time *t* is given by:

$$P_t = D_t\left[\frac{G}{R} + \left(\frac{G}{R}\right)^2 + \ldots + \left(\frac{G}{R}\right)^{\tau}\right], \tag{C.3}$$

where $G = \exp(\gamma_0)$ and $R = \exp(\delta_0)$. Thus, the mean dividend-to-price ratio at time *t* for a given τ is:

$$\overline{dp}_t = \log\left(\frac{1-\frac{G}{R}}{\frac{G}{R}-\left(\frac{G}{R}\right)^{\tau+1}}\right), \tag{C.4}$$

which is decreasing in τ . For $t \le T-1$,we log-linearize returns around $\overline{dp}_{t+1}$:

$$\begin{aligned} r_{t+1} &= \log(1+\exp(-\overline{dp}_{t+1})) + \frac{\exp(-\overline{dp}_{t+1})}{1+\exp(-\overline{dp}_{t+1})}(-dp_{t+1}+\overline{dp}_{t+1}) + \Delta d_{t+1} + dp_t, \\ dp_t &= -\kappa_{t+1} + \rho_{t+1} dp_{t+1} + r_{t+1} - \Delta d_{t+1}, \end{aligned} \tag{C.5}$$

where $\rho_{t+1} = \frac{\exp(-\overline{dp}_{t+1})}{1+\exp(-\overline{dp}_{t+1})}$ and $\kappa_{t+1} = \log(1+\exp(-\overline{dp}_{t+1})) + \rho_{t+1}\overline{dp}_{t+1}$.

For $t = T$ returns are simply given by:

$$\begin{aligned} r_T &= \log(D_T / P_{T-1}) = \Delta d_T + dp_{T-1}, \\ dp_{T-1} &= r_T - \Delta d_T. \end{aligned} \tag{C.6}$$

Starting from the log-linearized dp_t, we can iterate forward:

$$\begin{aligned} dp_t &= -\kappa_{t+1} + \rho_{t+1}\left[-\kappa_{t+2} + \rho_{t+2} dp_{t+2} + r_{t+2} - \Delta d_{t+2}\right] + r_{t+1} - \Delta d_{t+1}, \\ &= -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + r_{t+1} + \sum_{j=2}^{\tau-1} r_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} - \Delta d_{t+1} - \sum_{j=2}^{\tau-1} \Delta d_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + dp_{\tau-1} \prod_{k=1}^{\tau-1} \rho_{t+k}. \end{aligned} \tag{C.7}$$

Plugging in for dp_{T-1} and grouping terms we arrive at:

$$dp_t = -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + \left(r_{t+1} - \Delta d_{t+1}\right) + \sum_{j=2}^{\tau} \left(r_{t+j} - \Delta d_{t+j}\right) \prod_{k=1}^{j-1} \rho_{t+k}. \tag{C.8}$$

From here we take expectations conditional on time *t* and plug in for the AR(1) specification:

$$
\begin{aligned}
dp_t &= -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + E_t\left(r_{t+1} - \Delta d_{t+1}\right) + \sum_{j=2}^{\tau} E_t\left(r_{t+j} - \Delta d_{t+j}\right) \prod_{k=1}^{j-1} \rho_{t+k} \\
&= -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + \left(\mu_t - g_t\right) + \sum_{j=2}^{\tau} E_t\left(\mu_{t+j-1} - g_{t+j-1}\right) \prod_{k=1}^{j-1} \rho_{t+k} \\
&= -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + \left(\mu_t - g_t\right) + \sum_{j=1}^{\tau-1} E_t\left(\mu_{t+j} - g_{t+j}\right) \prod_{k=1}^{j} \rho_{t+k} \\
&= -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + \left(\mu_t - g_t\right) + \sum_{j=1}^{\tau-1} \left(\delta_0 + \delta_1^j (\mu_t - \delta_0) - \gamma_0 - \gamma_1^j (g_t - \gamma_0)\right) \prod_{k=1}^{j} \rho_{t+k} \qquad \text{(B 9)} \\
&= -\kappa_{t+1} - \sum_{j=2}^{\tau-1} \kappa_{t+j} \prod_{k=1}^{j-1} \rho_{t+k} + \left(\mu_t - g_t\right) + \left(\delta_0 - \gamma_0\right) \sum_{j=1}^{\tau-1} \prod_{k=1}^{j} \rho_{t+k} \\
&\quad + \left(\mu_t - \delta_0\right) \sum_{j=1}^{\tau-1} \delta_1^j \prod_{k=1}^{j} \rho_{t+k} - \left(g_t - \gamma_0\right) \sum_{j=1}^{\tau-1} \gamma_1^j \prod_{k=1}^{j} \rho_{t+k} \\
&= C + D\left(\delta_0 - \gamma_0\right) + \left(\mu_t - g_t\right) + D_\mu \left(\mu_t - \delta_0\right) - D_g \left(g_t - \gamma_0\right).
\end{aligned}
$$

We are interested how the dividend-to-price ratio varies with τ. Above we already noted that the dividend-to-price ratio decreases with the number of periods τ. Thus, longer duration assets have a lower dividend-to-price ratio. From the equation above, we also see the first insights on how changes in duration affect the relative importance of discount rate and cash flow news. In particular, expected returns and expected dividend growth rates affect the dividend-to-price ratio directly and through D_μ and D_g. These terms are increasing in τ and the persistence of either expected returns δ_1 or expected dividend growth rates γ_1. Thus, a higher τ implies greater sensitivity to both expected returns and expected dividend growth rates. But the relative increase in sensitivities is determined by the persistence parameters. In the special case, when $\delta_1 = \gamma_1$, an increase in τ has no effect on the relative importance of discount rate and cash flow news. When $\delta_1 \neq \gamma_1$, however, an increase in τ disturbs the relative importance of discount rate and cash flow news; $\delta_1 > \gamma_1$ implies that the relative importance of discount rate news increases; equivalently, $\delta_1 < \gamma_1$ suggests an increase in cash flow news. We show analytically, for a simplified version with constant ρ, that discount rate news indeed increases if and only if $\delta_1 > \gamma_1$. Using realistic parameters from Binsbergen and Koijen (2010), we also assess the effect of τ in simulations.

C.1.1 Analytical solutions

To derive closed form solutions, we set ρ_{t+1} equal to its unconditional mean. This is a reasonable approximation as long as *t* is not too close to *T*. In this case, the dividend-to-price ratio becomes:

$$
\begin{aligned}
dp_t &\simeq -\kappa_{t+1}\sum_{j=0}^{\tau}\bar{\rho}^{j} + \sum_{j=0}^{\tau}\bar{\rho}^{j}(\delta_0-\gamma_0) + (\mu_t-\delta_0)\sum_{j=0}^{\tau}\left(\bar{\rho}\gamma_1\right)^{j} - (g_t-\gamma_0)\sum_{j=0}^{\tau}\left(\bar{\rho}\gamma_1\right)^{j} \\
&= -\kappa_{t+1}\left(\frac{1-\left(\bar{\rho}\right)^{\tau+1}}{1-\bar{\rho}}\right) + (\delta_0-\gamma_0)\left(\frac{1-\left(\bar{\rho}\right)^{\tau+1}}{1-\bar{\rho}}\right) \\
&\quad + (\mu_t-\delta_0)\left(\frac{1-\left(\bar{\rho}\delta_1\right)^{\tau+1}}{1-\bar{\rho}\delta_1}\right) - (g_t-\gamma_0)\left(\frac{1-\left(\bar{\rho}\gamma_1\right)^{\tau+1}}{1-\bar{\rho}\gamma_1}\right).
\end{aligned}
\tag{C.10}
$$

We first take derivatives of the dividend-to-price ratio with respect to expected returns and expected growth rates:

$$
\begin{aligned}
\frac{\partial dp_t}{\partial \mu_t} &\simeq \frac{1-\left(\rho\delta_1\right)^{\tau+1}}{1-\rho\delta_1}, \\
\frac{\partial dp_t}{\partial g_t} &\simeq -\frac{1-\left(\rho\gamma_1\right)^{\tau+1}}{1-\rho\gamma_1}.
\end{aligned}
\tag{C.11}
$$

The key question is how these derivatives change as we increase τ.

$$
\begin{aligned}
\frac{\partial\left[\partial dp_t / \partial \mu_t\right] dp_t}{\partial \tau} &= -\frac{\log\left(\rho\delta_1\right)}{1-\rho\delta_1}\left(\rho\delta_1\right)^{\tau+1}, \\
\frac{\partial\left[\partial dp_t / \partial g_t\right] dp_t}{\partial \tau} &= \frac{\log\left(\rho\gamma_1\right)}{1-\rho\gamma_1}\left(\rho\gamma_1\right)^{\tau+1}.
\end{aligned}
\tag{C.12}
$$

Note that $\rho\delta_1 < 1$ and $\rho\gamma_1 < 1$, and therefore $\log\left(\rho\delta_1\right) < 0$ and $\log\left(\rho\gamma_1\right) < 0$. Thus, as we increase τ, $\partial dp_t / \partial \mu_t > 0$ will increase and $\partial dp_t / \partial g_t < 0$ will decrease. In other words, the dividend-to-price ratio becomes more sensitive to changes in dividend growth rates and expected returns as τ increases.

The next question is when is the change in response to an increase to τ larger for $\partial dp_t / \partial \mu_t > 0$ than for $\partial dp_t / \partial g_t < 0$. To evaluate this, we compare the following elasticities:

$$\frac{\partial\left[\partial dp_t / \partial \mu_t\right]}{\partial \tau}\frac{\tau}{\partial dp_t / \partial \mu_t} = -\tau \log\left(\rho\delta_1\right)\frac{\left(\rho\delta_1\right)^{\tau+1}}{1-\left(\rho\delta_1\right)^{\tau+1}},$$
$$\frac{\partial\left[\partial dp_t / \partial g_t\right]}{\partial \tau}\frac{\tau}{\partial dp_t / \partial g_t} = -\tau \log\left(\rho\gamma_1\right)\frac{\left(\rho\gamma_1\right)^{\tau+1}}{1-\left(\rho\gamma_1\right)^{\tau+1}}, \tag{C.13}$$

The elasticities are identical, except for the persistence parameters. Consistent with intuition, the elasticities are increasing in persistence.[39] It then follows that if, and only if, $\delta_1 > \gamma_1$, the dividend-to-price ratio will be relatively more sensitive to discount rates with an increase in τ.

C.1.2 Cross-sectional simulations

To support the analytical solutions and provide the first insight into the magnitude of the effect of τ on the relative importance of discount rate news, we conduct cross-sectional simulations. For a given τ, we start from the unconditional mean of the dividend-to-price ratio, where $\mu_t = \delta_0$ and $g_t = \gamma_0$. Then we impose 10,000 pairs of shocks to expected growth rates and discount rates. For each pair of shocks, we recalculate the dividend-to-price ratio. Finally, we run a cross-sectional regression of the resulting dividend-to-price ratios on shocks to expected returns and separately a regression of the dividend-to-price ratio on shocks to growth rates. We repeat the exercise for each τ between 5 and 50.

We first consider the case where expected returns and expected dividend growth rates have the same persistence $\delta_1 = \gamma_1 = 0.6$, and symmetric, uncorrelated shocks to expected returns and growth rates:

$$\begin{bmatrix}\varepsilon_t^{\mu} \\ \varepsilon_t^{g}\end{bmatrix} \sim N\left[\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}0.01 & 0 \\ 0 & 0.01\end{pmatrix}\right]$$

We set the mean expected returns and growth rates to $\delta_0 = 0.090$ and $\gamma_0 = 0.062$. Figure 1 plots the mean dividend-to-price ratio and a proxy for the relative sensitivity to discount rate news, defined as $\beta_r / \left(\beta_r - \beta_{dg}\right)$, where β_r is the estimated coefficient in the regression of the

---

[39] To see this mathematically, note that in response to an increase in x, for $0 < x < 1$, $-\log\left(x\right)$ decreases at a slower rate than $\frac{x^{\tau+1}}{1-x^{\tau+1}}$ increases.

dividend-to-price ratio on shocks to expected returns and β_{dg} is the estimated coefficient in the regression of the dividend-to-price ratio on shocks to expected growth rates. As expected, the dividend-to-price ratio is decreasing in τ. The relative sensitivity to discount rate news, in this special case of equal persistence of expected returns and growth rates, is constant and always 0.5, in line with our analytical solutions.

Next we use the empirically estimated parameters from Binsbergen and Koijen (2010) (Table VI, p.1463). These parameters are based on the present value model applied to S&P 500 nominal data between 1946 and 2007. Expected returns are more persistent than expected growth rates, $\delta_1 = 0.927$ and $\gamma_1 = 0.485$ and shocks to expected returns and growth rates are correlated:

$$\begin{bmatrix} \varepsilon_t^{\mu} \\ \varepsilon_t^{g} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.013^2 & 0.494\times 0.013\times 0.046 \\ 0.494\times 0.013\times 0.046 & 0.046^2 \end{pmatrix}\right]$$

As depicted in Figure 2, like before, the mean dividend-to-price ratio is convex and decreasing in τ. Unlike before, however, the relative importance of discount rate news is concave and increasing in τ. Thus, using a set of realistic parameters with expected returns more persistent than expected growth rates, our simulations confirm that longer duration assets are relatively more affected by the discount rate news.

C.2 Aggregate market

Next, we provide simulation results at the aggregate market level. We assume that the market consists of τ assets, where each assets lives for τ periods. Every year one asset dies and it is replaced by a new one. The average life of an asset in the market is therefore $\tau / 2$. All assets that are in the market at time *t* are subject to the same expected return and expected dividend growth rate. Thus, the only difference between the assets at a given time *t* stems from differences in τ.

We start by simulating the process for expected returns, expected dividend growth rates, and the realized dividends, using parameters from Binsbergen and Koijen (2010), Table VI, p.1463.

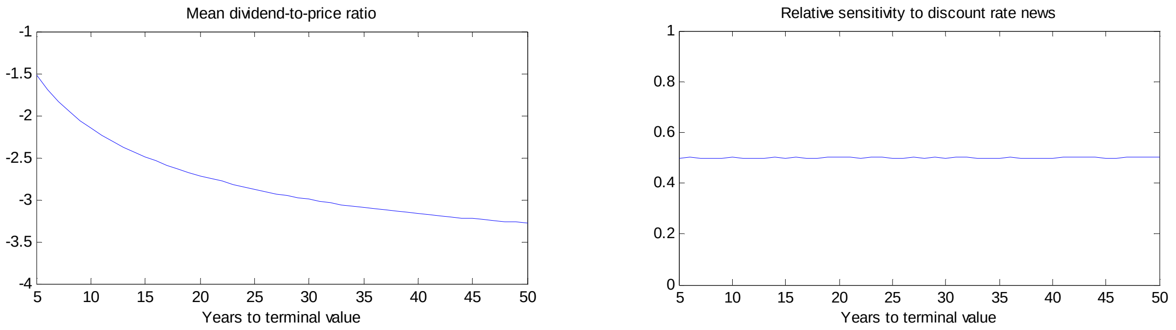
$$\begin{bmatrix} \varepsilon_t^{\mu} \\ \varepsilon_t^{g} \\ \varepsilon_t^{dg} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.013^2 & 0.494\times 0.013\times 0.046 & 0.858\times 0.013\times 0.004 \\ 0.494\times 0.013\times 0.046 & 0.046^2 & 0 \\ 0.858\times 0.013\times 0.004 & 0 & 0.004^2 \end{pmatrix}\right]$$

We then calculate the dividend-to-price ratio for each asset at each time *t*. Assuming an initial value for dividends and using the dividend-to-price ratio data, we calculate the time-series of dividends and prices. From here, we calculate returns and returns without the dividends (capital appreciation). Applying value-weighting, we obtain aggregate returns with and without the dividends. Following Cochrane (2008), p. 1541, footnote 5, we then calculate all the aggregate variables of interest. We consider three cases: $\tau = 100, 40, 10$. In each case, we let simulations run for 10,000 years. Results are reported in Table 7 and discussed in the main text.

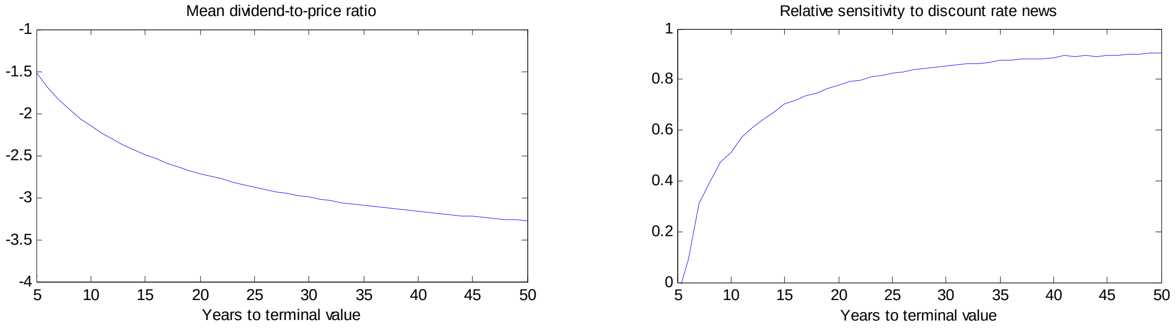
Figure C.1: Cross-sectional simulations

This figure plots the mean dividend-to-price ratio and the relative sensitivity to discount rate news for τ between 5 and 50 years. Panel A is based on the case of equally persistent, symmetric and uncorrelated shocks to expected returns μ and expected dividend growth rates g. Panel B is based on the realistic parameter values from Binsbergen and Koijen (2010), Table VI. The simulation details are provided in the text.

Panel A: Equally persistent, symmetric, and uncorrelated shocks to μ and g.



Panel B: Binsbergen and Koijen (2010) parameters



Four centuries of return predictability

Benjamin Golez
University of Notre Dame

Peter Koudijs
Stanford University

Online Appendix

This Online Appendix reports supporting evidence for the results reported in the main paper. Table OA.1 reports the summary statistics for the triennial real data. The rest of the OA analyzes the sensitivity of our main results based on annual real data reported in Table 2. In particular, we replicate the main results using iterative GMM, nominal data, a variation where we extend the U.K. period until 1900 and we use the broad-based CRSP index rather than the S&P 500 from 1925 onward, a reduced version VAR model, and a variation where we stop in 1982.

Iterative GMM (Table OA.2)

In the main analysis, we estimate the VAR model by two-step generalized method of moments with optimal bandwidth determined by the Newey and West method. Now, we establish that our main results from Table 2 are largely the same if we use iterative GMM instead. As reported in Table OA.2, the only notable difference is that the dividend growth is now predictable in the full sample at the five percent statistical significance (rather than one percent). The contrast between the recent U.S. period and the early periods remains pronounced.

Nominal data (Table OA.3)

In the main analysis, we estimate our VAR model using real data. Now, we repeat the analysis using nominal prices and dividends. As reported in Table OA.3, results are very comparable. The only notable difference is that returns are less predictable and the dividend growth is more predictable in the U.K period. The contrast between the recent U.S. period and the early periods remains and appears to be even more pronounced.

Extended U.K. period and CRSP index (Table OA.4)

In the main analysis, we switch from the U.K. data to the U.S. data in 1871 (for better comparability with recent studies). One can argue, however, that it was not until the beginning of the 20th century that that the U.S. became the world's largest economy. Also, the Cowles (1939) data, which we use between 1871 and 1925, includes only 50 companies in 1871 (258 in 1925). Many more companies were traded in the U.K. at the end of the 19th century. Moreover, after 1925 we use the S&P 500 in the main analysis, which was effectively the S&P 90 till March 1957. Again, one may wonder if results change by using an index with a broader coverage.

Here, we present results where we switch from the U.K. to the U.S. market in 1900. The U.K. data for the period 1870-1900 comes from Grossman (2002). These data includes 520 companies in 1870 and around 1,000 companies in 1900. We also use the CRSP value-weighted index from 1925 onward rather than the S&P 500. The CRSP index includes 533 companies in January 1926 (rather than 90) and 7,004 in December 2014 (rather than 500).

Results are reported in Table OA.4 and are surprisingly similar to the main results reported in Table 2. There is some more evidence for return predictability in 1900-1945 (in comparison to 1870-1945) as the estimated return parameter is now significant at the ten percent level.[40] At the same time, in the U.K. period, returns are somewhat less predictable and dividend growth rates are somewhat more predictable (both are significant at the five percent level). Finally, in the recent U.S. period, returns are now predictable at the five percent level (rather than one percent level). The contrast between the recent U.S. period and the early periods remains pronounced.

Reduced-form model (Table OA.5)

In the main analysis, we use a first order VAR model with returns, dividend growth rates, and the dividend-to-price ratio. Now, we repeat our main results using a reduced form model. In particular, 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):

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

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 two-step GMM subject to the above restriction 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.5 and are very similar to the results from the main analysis reported in Table 2. The most notable difference is that dividend growth parameter is

---

[40] In untabulated results, we note that returns are predictable at the ten percent level in the 1900-1945 period even if we use the S&P 500 (rather than the CRSP index) from 1925 onward.

now insignificant in the U.K. period (rather than significant at the ten percent level). We also note that dividend growth is more predictable (at the one percent significance level) in the Netherlands/U.K. period and returns are somewhat less predictable in the recent U.S. period (at the five percent significance level). The contrast between the recent U.S. period and the early periods, however, remains.

Recent U.S. period ending in 1982 (Table OA.6)

In the main analysis, the recent U.S. period runs from 1945 through 2014. To mitigate the effect of repurchases, we replicate our main results ending the recent U.S. period before repurchases became economically important in 1982. Table OA.6 reports results for the U.S. period (1945-1982) and the full period (1686-1982) based on the full VAR model and the reduced form model. Results are very similar to results reported in Table 2 and Table OA.5. If anything, the contrast between the 1945-1982 period and the full period is even more pronounced.

Table OA.1: Summary statistics: Triennial data

This table reports summary statistics for the triennial real data used in the estimations in Table 3. Column (1) reports the statistics for the 1629-1945 period that combines the Netherlands/U.K. data (1629-1809), the U.K. data (1825-1870), and the early U.S. period (1871-1945); column (2) reports the same statistics for the recent U.S. period (1945-2014); column (3) reports the statistics based on the full sample. Capital appreciation is denoted by *ca*.

	(1)	(2)	(3)
	1629-1945	1945-2014	1629-2014
r	0.171	0.200	0.176
$Cap.\ app.\ (ca)$	0.035	0.110	0.049
ca / r	0.207	0.551	0.278
$Std.(r)$	0.225	0.275	0.234
$AR1(r)$	-0.141	0.091	-0.085
dg	0.036	0.055	0.040
$Std.(dg)$	0.396	0.111	0.361
$AR1(dg)$	-0.253	-0.083	-0.254
dp	-1.978	-2.442	-2.063
$Std.(dp)$	0.333	0.430	0.394
$AR1(dp)$	0.145	0.711	0.437

Table OA.2: Vector autoregression (VAR) estimates: Iterative GMM

Panel A reports VAR estimates of real returns, real dividend growth rates, and the dividend-to-price ratio. The data are 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. 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. 1686-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2014	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.195***	0.244**	0.138	0.103***	0.113***
	(0.043)	(0.107)	(0.088)	(0.033)	(0.033)
r_t	0.002	0.180*	0.115	-0.001	0.054
	(0.087)	(0.105)	(0.169)	(0.087)	(0.055)
dg_t	-0.019	-0.003	-0.261*	0.222	0.071
	(0.047)	(0.119)	(0.145)	(0.176)	(0.070)
Dep. variable: dg_{t+1}					
dp_t	-0.216**	-0.209*	-0.282***	-0.009	-0.076**
	0.086	0.110	0.074	0.020	0.032
r_t	0.247**	-0.194	0.272***	0.129*	0.116
	(0.117)	(0.199)	(0.102)	(0.069)	(0.077)
dg_t	0.002	0.291**	0.047	0.323***	-0.010
	(0.109)	(0.122)	(0.089)	(0.077)	(0.063)
Dep. variable: dp_{t+1}					
dp_t	0.615***	0.571***	0.610***	0.916***	0.846***
	(0.115)	(0.104)	(0.100)	(0.045)	(0.054)
r_t	0.257**	-0.391**	0.165	0.134	0.064
	(0.130)	(0.189)	(0.214)	(0.124)	(0.091)
dg_t	0.022	0.307**	0.323*	0.104	-0.085
	(0.109)	(0.140)	(0.168)	(0.155)	(0.096)
J-test	1.430	0.054	1.804	0.715	9.153**
Panel B: Decomposition results					
Dividend-to-price					
DR	0.539	0.482	0.411	1.053	0.639
CF	0.461	0.518	0.589	-0.053	0.361
Unexpected returns					
DR	0.114	0.254	0.312	0.792	0.440
CF	0.888	0.746	0.694	0.208	0.561

Table OA.3: Vector autoregression (VAR) estimates: Nominal data

Panel A reports VAR estimates of nominal returns, nominal dividend growth rates, and the dividend-to-price ratio. The data are annual. The model is estimated by two-step 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. 1686-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2014	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.164***	0.081*	0.117	0.106***	0.077***
	(0.045)	(0.043)	(0.104)	(0.036)	(0.028)
r_t	-0.057	0.072	0.148	-0.015	0.079
	(0.074)	(0.125)	(0.208)	(0.092)	(0.077)
dg_t	-0.036	-0.102	-0.271**	0.039	-0.032
	(0.053)	(0.067)	(0.117)	(0.229)	(0.053)
Dep. variable: dg_{t+1}					
dp_t	-0.237**	-0.350***	-0.225***	-0.001	-0.134***
	(0.094)	(0.087)	(0.059)	(0.018)	(0.038)
r_t	0.239*	-0.149	0.419***	0.123	0.289***
	(0.137)	(0.237)	(0.075)	(0.080)	(0.058)
dg_t	0.002	0.265***	0.134*	0.421***	0.084
	(0.150)	(0.082)	(0.075)	(0.084)	(0.076)
Dep. variable: dp_{t+1}					
dp_t	0.627****	0.594***	0.691***	0.920***	0.823***
	(0.102)	(0.116)	(0.102)	(0.047)	(0.052)
r_t	0.310*	-0.231	0.285	0.141	0.219**
	(0.162)	(0.333)	(0.227)	(0.144)	(0.100)
dg_t	0.040	0.382***	0.426***	0.394	0.122
	(0.139)	(0.103)	(0.150)	(0.241)	(0.099)
J-test	1.286	0.457	1.807	0.745	12.452***
Panel B: Decomposition results					
Dividend-to-price					
DR	0.443	0.209	0.396	1.240	0.405
CF	0.557	0.791	0.604	-0.240	0.595
Unexpected returns					
DR	0.224	0.221	0.242	0.969	0.210
CF	0.778	0.780	0.763	0.030	0.789

Table OA 4: Extended U.K. period and CRSP index

Panel A reports VAR estimates of real returns, real dividend growth rates, and the dividend-to-price ratio. The data are annual. The model is estimated by two-step 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. 1686-1809	U.K. 1825-1900	U.S. 1900-1945	U.S. 1945-2014	Full period
Panel A: VAR estimates					
Dep. variable: r_{t+1}					
dp_t	0.188***	0.121**	0.288*	0.100**	0.093***
	(0.041)	(0.048)	(0.165)	(0.046)	(0.029)
r_t	-0.002	0.142	0.288	0.006	0.044
	(0.087)	(0.090)	(0.242)	(0.133)	(0.060)
dg_t	-0.023	-0.014	-0.291	0.120	0.003
	(0.046)	(0.044)	(0.221)	(0.250)	(0.050)
Dep. variable: dg_{t+1}					
dp_t	-0.208**	-0.194**	-0.287***	-0.005	-0.093***
	(0.087)	(0.094)	(0.073)	(0.023)	(0.034)
r_t	0.247**	-0.132	0.304***	0.109	0.166***
	(0.117)	(0.257)	(0.080)	(0.062)	(0.073)
dg_t	0.008	-0.040	0.107	0.159	-0.045
	(0.105)	(0.101)	(0.074)	(0.117)	(0.063)
Dep. variable: dp_{t+1}					
dp_t	0.631***	0.712***	0.447***	0.922***	0.847***
	(0.113)	(0.086)	(0.168)	(0.045)	(0.052)
r_t	0.260**	-0.285	0.016	0.106	0.127
	(0.132)	(0.234)	(0.282)	(0.159)	(0.089)
dg_t	0.032	-0.027	0.419*	0.040	-0.051
	(0.103)	(0.098)	(0.234)	(0.283)	(0.085)
J-test	1.595	1.719	0.948	0.518	8.178
Panel B: Decomposition results					
Dividend-to-price					
DR	0.538	0.405	0.509	1.062	0.558
CF	0.462	0.595	0.491	-0.062	0.442
Unexpected returns					
DR	0.113	0.189	0.372	0.833	0.358
CF	0.889	0.813	0.633	0.168	0.636

Table OA.5: Reduced-form model

Panel A reports the regression estimates of real returns, real 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 two-step 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. 1686-1809	U.K. 1825-1870	U.S. 1871-1945	U.S. 1945-2014	Full period
Panel A: VAR estimates						
	Dep. variable: r_{t+1}					
	dp_t	0.168***	0.230**	0.083	0.096**	0.092***
		(0.036)	(0.105)	(0.059)	(0.045)	(0.029)
	Dep. variable: dg_{t+1}					
	dp_t	-0.208***	-0.126	-0.420***	-0.007	-0.141***
		(0.065)	(0.091)	(0.064)	(0.027)	(0.048)
	Dep. variable: dp_{t+1}					
	dp_t	0.652***	0.672***	0.523***	0.925***	0.800***
		(0.089)	(0.086)	(0.090)	(0.046)	(0.065)
	J-test	0.656	0.002	0.181	0.032	1.244
Panel B: Decomposition results						
	Dividend-to-price					
	DR	0.446	0.646	0.164	0.934	0.394
	CF	0.554	0.354	0.836	0.066	0.606

Table OA.6: Recent period ending in 1982

Panel A reports predictive estimates; columns (1) and (2) report VAR estimates of real returns, real dividend growth rates, and the dividend-to-price ratio; columns (3) and (4) report similar estimates from a reduced form model. The data are annual and stop in 1982. The models are estimated by two-step 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 estimates from Panel A.

	(1)	(2)	(3)	(4)
	VAR model		Reduced form model	
	U.S. 1945-1982	Full period	U.S. 1945-1982	Full period
Panel A: VAR estimates				
Dep. variable: r_{t+1}				
dp_t	0.274***	0.179***	0.255***	0.163***
	(0.051)	(0.031)	(0.075)	(0.034)
r_t	0.074	0.069		
	(0.060)	(0.061)		
dg_t	-0.051	-0.049		
	(0.190)	(0.057)		
Dep. variable: dg_{t+1}				
dp_t	0.021	-0.142***	0.032	-0.189***
	(0.020)	(0.045)	(0.054)	(0.046)
r_t	0.082	0.181**		
	(0.051)	(0.072)		
dg_t	0.296***	0.014		
	(0.081)	(0.059)		
Dep. variable: dp_{t+1}				
dp_t	0.778***	0.710***	0.809***	0.677***
	(0.061)	(0.055)	(0.085)	(0.051)
r_t	0.008	0.117		
	(0.071)	(0.076)		
dg_t	0.362	0.066		
	(0.222)	(0.083)		
J-test	1.009	4.983	0.140	1.777
Panel B: Decomposition results				
Dividend-to-price				
DR	1.284	0.621	1.145	0.463
CF	-0.284	0.379	-0.145	0.537
Unexpected returns				
DR	1.029	0.357		
CF	-0.029	0.647		