Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

This Time it's Dividend

Investigations into the Valuation of Future Dividends

Jacco Christian Kragt

PhD 01/2018-016

This time it's dividend

Kragt, Jac

This Time it's Dividend

Investigations into the Valuation of Future Dividends

JACCO CHRISTIAN KRAGT

January 17, 2018

This Time it's Dividend

Investigations into the Valuation of Future Dividends

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnficus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op woensdag 17 januari 2018 om 16.00 uur door

JACCO CHRISTIAN KRAGT

geboren op 21 februari 1967 te Voorst.

PROMOTIECOMMISSIE:

PROMOTORES:	Prof. dr. F.C.J.M. de Jong
	Prof. dr. J.J.A.G. Driessen
OVERIGE LEDEN:	Prof. dr. P.C. Schotman
	Prof. dr. R.S.J. Koijen
	Dr. R.G.P. Frehen

Acknowledgements

My promotores Frank de Jong and Joost Driessen have been a tremendous inspiration to me. I have thoroughly enjoyed discussing thoughts, results, literature and all else relevant to achieving the goal of writing this thesis. I am grateful.

I would like to thank the other members of the committee to promote: Peter Schotman, Ralph Koijen and Rik Frehen. They have each provided valuable comments for improvement and suggestions for application of the work in this thesis.

A true word of thanks is due to my room mate, Andreas Rapp. We talked through many of our ups and downs and helped each other on the way.

Much of this thesis has been presented and discussed at seminars. I would like to thank Adlai Fisher, Frederico Belo, my colleagues at the Finance department of Tilburg University and conference participants at AFA 2016, SFS Cavalcade 2015, Essec Business School, Netspar 2014, OptionMetrics research 2016 and Eurofidai 2017.

Heleen and my children Annet, Wouter, Ceciel and Dirk Jan are in my heart and mind. Their love and support is beyond description.

Contents

Chapter 1

Summary

The present value identity is omnipresent in academic Finance, yet its validity for stocks in the shape of the dividend discount model has never been proven definitively. This thesis contains attempts the reconcile stock prices with the value of future dividends.

Present value reasoning is easy to follow; like-for-like cash flows must have the same value, whatever the package they come in. And consequently, the value of a financial asset should equal the sum of its parts, which is the value of the cash flows it produces. There is no free lunch, or unaccounted-for loss. In the case of bonds, the present value of a set of cash flows actually does come very close to the price of the bond that pays them. Any meaningful difference in price will be quickly arbitraged down, if not to zero, then to a point where only transaction costs limit the final push to complete equality.

Obvious as this seems, it is perhaps small wonder that the extant literature does not often raise the fact that we can't be too sure whether present value thinking holds up for stocks as well. The present value of a future dividend is defined as the expectation of the dividend discounted at the time value of money and a risk premium. The dividend discount model is an elaboration of the present value identity, relating the present value of all future dividends to the stock price.

It is quite impossible to acquire prices or valuations of all dividends that a stock pays in the future. There is no contractual commitment to pay dividends, there is no defined pay date, no amount promised. In fact, there may be nothing at all as many companies do not pay any dividend even if they make large profits. And without observable valuations of future dividends, the dividend discount model can only be tested by making strong modeling assumptions about dividend expectations and risk premiums!

The goal of this thesis is to investigate the present value of future dividends and how they are related to stocks, using the prices of derivatives. Particularly, I apply futures and swaps on dividends of stock indices to estimate the future development of dividend values and I compare its model value to stock market indices. Next, I derive valuations of dividends from stock options for individual companies. I test the extent to which these implied valuations predict actual dividends and how they interact with stock returns across companies.

Future dividends paid by stock indices

The introduction of dividend derivatives in the early 2000's has removed the complete lack of information about future stock payoffs. A dividend that is to be paid in the future can actually be traded by means of these dividend swaps and futures. Prices of dividends paid in years to come thus are known. Mainstream stock indices such as the Eurostoxx 50 and the S&P 500 have such derivatives traded on their dividends up to some ten years into the future. Equipped with these price data, we can attribute a proportion of 20 to 30% of the value of the index to specific dividend payments. Clearly, that reduces the uncertainty about how the cash flows that are expected from stocks are valued.

But valuing a stock by means of all its future dividends requires prices for the indefinite future. To get to 100% actual market prices for future dividends is, of course, utopia. Yet dividend price data for the next ten years might be enough for answering several research questions. The first of which is: What do the dynamics of these valuations actually look like? Both academic authors and investment banks have established that buying a dividend that is paid in the next two to three years has a higher return and a higher volatility than the stock market itself, and that their Sharpe ratio is better than that of stocks. But an analysis of the shape of the term structure of dividend values has not yet seen the light. Does it have inflection points, how many, at what horizon? Second is how to deal with what happens beyond the first ten years. Surely investors entertain some expectation about the distant future. Is it a variable growth rate, or is it fixed? What is its level?

In the first chapter of this thesis a methodology that is known for approximating bond curves is applied to model dividend valuations as a term structure. The analysis provides reasonable answers. The first is that, starting from a one-year discounted dividend price, a simple model with only two state variables are needed to concisely describe a daily changing term structure of discounted dividend prices; one at a horizon within one year from the first dividend and the second at a business cycle horizon of about five years. The answer to the question about dividends in the more distant to indefinite future is that investors anticipate a long run fixed decline beyond the business cycle of these valuations by 2 to 3% per year. I investigate several major stock indices and the decline sits within this range for each of them. As a risk-adjusted discount rate, negative growth of this magnitude makes sense economically. Dividend values can't be rising indefinitely, only if they become less valuable at some point in the future will stock prices be stable.

The ingredients are now available for constructing a valuation of the stock market itself: a value of current dividends, two state variables and a fixed indefinite growth rate which together construct a growth path for dividend valuations into infinity. A comparison of stock market prices to this modeled stock market valuation based on dividends reveals that levels are reasonably close to each other. The fixed decline in dividend valuations into infinity of 2 to 3% also appears reasonable for valuating the stock index. Variation in this rate is not required to stay close to the market, so investors do not appear to change their expectations of the very distant future. A simple regression of price changes confirms the dynamic relationship. More than half the daily stock market returns can be explained by changes in dividend valuations, of which both the valuation of current dividends and the term structure of future dividend values each account for about 25%-points. No other models of daily stock market returns using prices of financial products are capable of achieving such a high explanatory power. It is high enough to answer the overarching question in this thesis: the dividend discount model defined as a present value identity enhanced by a modeled term structure of dividend valuations is indeed appropriate for matching up dividends to stocks. Using the present value of future dividends, as dividend discounting intends, provides the evidence.

The three elements to the present values of dividends, expectations of dividends, the time value of money and a risk premium can be modeled individually. Bond yields are observable and can stand in for time value, but objective dividends and the risk premium are not observable independently from each other. The term structure of risk premiums is an important topic of academic research, so many authors delve into disentangling the two, sometimes from the values of future dividends.

There are two major benefits to modeling a term structure without disentangling. The first is that it removes a modeling step. Disentanglement involves modeling the dividend risk premium and objective dividend growth first. This introduces errors, which, who knows, may be substantial. The second step then amounts to modeling modeled values and the mounting of errors may cause intractable results. Reconciling such modeled values to the stock market or its returns is thus on a back-foot.

The second advantage is that modeling present values includes discounting at the time value of money. When investors formulate expectations about future dividends, the state of the world is an information set that conditions dividend growth, risk premiums and interest rates alike. For example, a positive shock to the economy may increase dividends and reduce the risk premium, both of which are good for the value of stocks, but it may also increase rates, which is bad for stocks. By considering the three elements in one go, it is the balance of such effects that is modeled, not estimates of three individual effects. I investigate the extent of this potential issue by modeling future values of dividends without discounting as well. The fit of this model is similar to the discounted dividend term structure, but the power of the undiscounted term structure to explain the stock market is much weaker.

Future dividends paid by individual companies

It stands to reason that future dividends may be informative not only for stock indices, but for the stocks of individual companies too. Understanding returns in a cross-section for portfolios of individual stocks based on future dividends requires access to data for the dividend derivatives of individual stocks. However, such data are not good enough to conduct an analysis at the company level. Individual dividend derivative products exist only for a handful of European companies, and they trade infrequently and with low turnover.

An alternative source of data for the valuation of future dividends is found in options. The price of a stock option depends in part on the dividend that the company is expected to pay between the trading date and the expiry date of the option. It is this forward-looking aspect of such implied dividends by which they can serve the purpose.

It is characteristic to option models that a model price depends on the volatility and the future dividends of a stock. Each should be the same when modeling the price of otherwise similar call and put options of the same stock. Practitioners and academic authors tend not to follow this line of thought, however. Usually an assumption is made for the dividend, which often is that it will remain the unchanged in the future from the last dividend paid. The reason to apply a fixed assumption for dividends is that solving the equation of a modeled option price to an observed price can surface only a single unknown variable and that volatility is the variable of interest. Of course, the actual valuation of a future dividend is not necessarily the same as the last paid dividend. A direct consequence of assuming that dividends never change is that the implied volatility calculated differs between a pair of call and put options that is otherwise equal even though the volatility concerns the same stock.

As I am interested not in volatility but in future dividends, the equivalent approach to implying a dividend value from option prices would be to assume that stock volatility will remain the same as it was in the past. I suspect that would carry little support. I therefore propose to make no assumptions about either one and to imply them both from two option prices at the same time. The basic thought is simple: to surface an extra unknown variable an extra equation is required. And there are two in the case of options, by equating the model prices of a pair of otherwise identical call and put options to their observed prices. By doing so simultaneously, the two unknowns dividend and volatility can be calculated.

Many US stocks have options traded on them and the data series for their prices are long.

However, the value of options on US stocks contain a premium for exercising the option before its expiry. Applying put-call parity without regard to this premium disturbs implying the value of future dividends from the option price. But the solving method I propose is well-suited for the benchmark binomial tree model for calculating prices of options that account for the early exercise premium. The binomial tree model is reverse-engineered by guessing values for implied dividends and implied volatility until the model price for the two options are sufficiently close to their observed prices. The methodology turns out to produce values for the two implied unknowns that are sensible. The result is a vast data set of implied dividends, with values for different horizons of all US stocks that have options traded on them at a daily interval over a period of nearly twenty years.

Implied dividends have several typical characteristics. Across companies, the average implied dividend is below actual dividends. The decline increases for longer horizons, but at a decelerating pace. There is a lot of variety in implied dividends among companies. The top 10% at least double at a horizon of six months, and the bottom 10% decline by at least half. Although noise in the data may add to this degree of dispersion, such swings are large enough to be caused by objective expectations rather than risk premiums. Good dispersion is highly relevant for engaging in cross-sectional analysis.

I test implied dividends for their predictive power for dividend changes. In the second chapter I show an example of dividend initiation by Apple Inc. in 2012 that is well picked-up by implied dividends as of about six months prior to the announcement of the initiation. I establish empirically that the data set confirms the predictive power of implied dividends. They prove meaningful to forecast actual future dividend changes, notably dividend increases, even at a horizon of six months. That far into the future I find that dividend cuts are not well predicted, although other authors establish that implied dividends often do forecast cuts within one month.

Implied dividends are highly relevant to stock returns when a company makes an announcement to change the dividends it pays. When companies announce a dividend cut, the event generally causes a stock price to sink by on average 2.6%. But if implied dividends correctly predict a future dividend cut, the stock price response to the announcement largely disappears.

Implied dividends and stock returns

In the third chapter I investigate the relationship between stock returns and implied dividends more generally. Are stock returns affected when investors have high expectations of future dividends? The dividend discount model suggests that stock prices change when future dividends change, but whether future dividends are high or low does not necessarily matter to expected returns. But it does. A measure to determine whether implied dividends are high or low is to compare them to the last dividend that is paid on a stock. Sorting stocks by this measure of dividend growth demonstrates that it is indeed associated with returns. In the year before the sorting month, the top quartile portfolio of dividend growth returns just under 1% per month better than the lowest quartile. The difference is even larger among stocks with a low dividend yield.

In the year *following* sorting by implied dividends, however, returns of stocks with high implied dividends are actually significantly *lower* than average. This turnaround in returns is more pronounced for stocks with a low dividend yield. The data suggest that a stock may have become expensive, as measured by a low dividend yield, in a period of high implied dividends that does not necessarily materialize into increasing actual dividends. Low dividend yield portfolios have a small but consistently lower return than high dividend yield portfolios. This finding is well-established in the literature, and from my results this seems an expression of irrational stock expensiveness associated with high implied dividends.

I look into the impact of implied dividends on stock pricing in the context of a CAPM and the Fama and French five-factor model by extending both models by implied dividends. Portfolios are delineated by sorting on company fundamentals that are known to bear on returns, such as *Book-to-Market*, *Operating Profitability* and *Investment*, and are sorted on implied dividends as well. The usual effects, or anomaly returns, are under pressure in double-sorted portfolios. Implied dividends override the value effect on returns from stocks with a high *Book-to-Market*. Return effects from *Operating Profitability* and *Investment* are visible only when implied dividends are slow or fast respectively. Just like the dividend yield effect, these effects may be the result of a stock's earlier returns related to the level of implied dividend causing it to be priced irrationally.

Interference with company fundamentals also surfaces when implied dividends are introduced as a risk factor. The premium paid by a *Fast minus Slow* factor of implied dividend growth depends strongly on the level of dividend yields in the portfolio. A similar dependence exists for portfolios sorted by *Book-to-Market*, *Operating Profitability* and *Investment*; the implied dividend risk factor matters to the returns of these portfolios when company fundamentals provide a basis. For example, only the returns of stocks of profitable companies respond to higher dividend growth.

Conclusion

A key observation from these investigations is that the valuation of future dividends link up with stock prices and that, as a consequence, we can reasonably conclude that they are formed in conjunction with stock prices. A important question remains: how exactly?

The anticipation of future dividends by means of two state variables may be a point of departure. Is the short-term state more influenced by company-specific risk and the medium-term more by the state of the economy? Which variables drive these states and do they line up with other term structures?

Present value accounting presumes that investors consider the *future* values of company fundamentals. I do not suggest that testing stock returns using *current* values of fundamentals instead is inappropriate. But the results from the cross-sections in this thesis suggest that the relationships found in, for example, the five-factor model may work through a different mechanism than present value reasoning in which they serve as proxies for future values. The question how to gauge expectations of future performance of company fundamentals should be a centerpiece in academic Finance.

The state of the economy is characterized by many macroeconomic variables and, here too, current observations of these variables are often used as proxies for future states. I contend that the relationship from current macro-indicators to stock prices, the equity risk premium in particular, may operate with short-term and medium-term dividends as a channel. Many financial market products contain prices for payoffs at a business cycle horizon, such as credit default swaps, index options and other derivatives. Prices of such products are known to interact with macroeconomic variables and they are a gauge of the state of the economy in their own right. Their linkage to future dividends may serve to better understand the channel of the impact on stock prices stemming from the state of the economy. For such questions, this dissertation is only a prelude.

Chapter 2

The Dividend Term Structure

We estimate a model for the term structure of discounted risk-adjusted dividend growth using prices of dividend swaps and futures in four major stock markets. A two-state model capturing short-term mean reversion within a year and a medium-term component reverting at business-cycle horizon gives an excellent fit of these prices. The model-implied dividend term structure aggregates to a price-dividend ratio. This model-implied ratio, combined with current dividends, captures most of the daily stock index return variation, despite the fast mean reversion to long-run growth. Using forward-looking dividend prices, this result confirms our dividend discount model dynamically. Another result is that investors do not appear to update the valuation of dividends much beyond the business cycle horizon.

CO-AUTHORS: JOOST DRIESSEN AND FRANK DE JONG

2.1 Introduction

"Since the level of the market index must be consistent with the prices of the future dividend flows, the relation between these will serve to reveal the implicit assumptions that the market is making in arriving at its valuation. These assumptions will then be the focus of analysis and debate.", (Brennan, 1998).

Dividends are a key ingredient for valuing stocks. Investors attach a present value to expected dividends and sum them to arrive at the value of a stock. As Campbell and Shiller (1988) have shown, stock prices thus vary because of changes in expected dividends, changes in interest rates, and changes in risk premiums. However, these elements may be horizon-dependent. For interest rates this is obvious as they can be readily observed. But also the expectations of dividends paid in the short run may at least partly be driven by other considerations than those of dividends paid in the distant future. Equally, risk premiums are likely to differ for various maturities, see for example van Binsbergen et al. (2012). Hence, investors will not only change the price of expected dividends from moment to moment, they may also change them for various maturities relative to each other, similar to a term structure of interest rates. In this paper, we focus on this term structure of the prices of expected dividends.

Given that the stock price is simply the sum of the present values of all dividends expected, Michael Brennan called in the late nineties for the development of a market for dividend derivatives. His wish came to life at the beginning of this century, with the introduction of derivatives referring to future dividend payments. These products exchange uncertain future dividends of an underlying stock or stock index in exchange for cash at the time of expiry. As such, they are forward looking in nature as they contain price information about expected dividends corrected for their risk. More precisely, the price of a single dividend future or OTC swap is the expected dividend for a given maturity discounted at the risk premium for this maturity. Finding present values of expected dividends only requires discounting these prices at the risk-free rate.

In this paper, we use data on these new dividend derivatives to study the dividend term structure for four major stock markets. A key starting point of our analysis is that we show that modeling the dynamics of a single variable is sufficient to describe the entire term structure of discounted dividend derivative prices, and to obtain a total value for the stock index. This single variable is equal to dividend growth minus the one period risk-free rate and a variable capturing the risk premium. We call this variable discounted risk-adjusted dividend growth. Hence, we do not need to separately assume processes for interest rates, risk premiums and dividend growth rates, the simplicity of which is a major advantage of our approach.

It is important to stress that this approach is nonstandard. Existing theoretical work usually separately models dividend growth and the preferences that determine discount rates. Empirically, the present value literature uses econometric models for the expectations about dividend growth and/or returns given past returns and dividend data. One of the earliest and best known examples is given by Campbell and Shiller (1988), who use vector autoregressive methods to predict returns based on past dividends, and use this to decompose returns into discount rate news and cash flow news. Many other attempts at decomposition of dividend growth and risk premiums have since followed (see Cochrane (2011) for an overview). Clearly, the ultimate goal of asset pricing is to understand both discount rate and cash flow dynamics, but it has proven to be difficult to reliably separate discount rates from cash flows. We show that we can learn a great deal about how investors value dividends without making restrictive assumptions on preferences and dividend processes.

Inspired by the affine models often used for modeling the term structure of interest rates, we show how to set up a standard affine model for discounted risk-adjusted dividend growth. Specifically, our model resembles the interest rate model of Jegadeesh and Penacchi (1996), who use a two factor model, where the first factor reverts to a second factor, which in turn reverts to a long run constant. This model thus distinguishes a short-term component, a medium-term component and constant asymptotic growth. We cast this model in state space form and apply the Kalman filter Maximum Likelihood approach to estimate it using dividend derivative prices of one to ten years. The resulting discounted risk-adjusted dividend growth term structure describes the maturity curve of dividend present values in full, including an estimate for long-term growth beyond the medium term until infinity.

We use data for four markets of dividend derivatives and contracts that extend out to horizons of up to ten years. Dividend derivative products exist in the shape of futures listed on stock exchanges and as swaps traded over the counter (OTC) between institutions. Minimum criteria for liquidity and transparency restrict the application of daily data in the estimation procedures to listed futures referring to the Eurostoxx 50 and the Nikkei 225 indices. Daily prices of OTC dividend swaps for the FTSE 100 and the S&P 500 indices are available as well, but they are less liquid and their representativeness of daily variation in dividend expectations is questionable. We therefore perform the same tests for these data using a monthly frequency.

Our key findings are as follows. First, we find evidence that our simple two factor affine model describes the term structure of dividends well. It captures the dynamics of measured growth rates and it delivers an estimate for infinite growth that is economically sensible. We do not need many factors, complex specifications for the factor volatilities or drift terms to generate a good fit.

Second, we find that the factors driving this term structure have rather strong mean reversion. The first factor has a half-life of 6 months to one year (for reversion to the second factor) and thus captures short-term movements in expected risk-adjusted dividends. The second factor reverts to a constant at a horizon of business-cycle proportion.

Third, we perform a relative pricing exercise, comparing the calibrated prices of future dividends to the observed value of the total stock index. Dividend derivatives have maturities up to ten years, but using our term structure model the extrapolated growth rates beyond that are summed to arrive at a model based estimate of the price-dividend ratio. Together with a market price for current dividends, a comparison is made to the actual stock market. This can be interpreted as an out of sample test of our dividend discount model, since the model is estimated using dividend derivatives only, and not the stock index value. At an R^2 of over 50%, we find that most of the variation in the stock market is explained by current dividends and our model implied price-dividend ratio. This demonstrates that the stock market can be understood quite well in terms of the market for dividend derivatives.

Fourth, given the good fit to the aggregate stock market, our results show that most of the variation in stock prices is captured by short-term and business-cycle movements in discounted risk-adjusted dividends. As the infinite growth rate is fixed, our results suggest that investors update their day-to-day valuation of dividends beyond the business cycle horizon only to a limited degree. Apparently, depicting long term investor expectations to be fixed is not a major impediment to capturing most of the observed stock market volatility.

Fifth, the fast mean reversion in growth rates also has implications for the term structure of dividend futures volatilities. We show that our model generates a good fit of this term structure. Using results of Binsbergen et al. (2012) we compare our estimates to the volatility term structures implied by theoretical asset pricing models. We find that our empirical estimates of the volatility term structure are broadly in line with the long-run risk model of Bansal and Yaron (2004). In contrast, the habit formation model of Campbell and Cochrane (1999) generates a volatility term structure that strongly differs from our estimates. Binsbergen and Koijen (2017) provide similar findings for the volatility term structure.

This paper adds to a recent literature that uses dividend derivatives in asset pricing. Our work complements Binsbergen et al. (2013). They introduce the concept of equity yields, which is related to our discounted risk-adjusted growth measure. However, Binsbergen et al. (2013) do not estimate a pricing model for the term structure of discounted risk-adjusted dividend growth nor price the stock market using this model. Instead, they focus on an empirical decomposition of dividend prices into dividend growth rates and risk premiums. They conclude that the term structure for risk premiums is pro-cyclical, whereas expected dividend growth is countercyclical. Our work complements their study, as we show that, without separating dividend growth and risk premiums, one can price the entire term structure of dividends and learn about its dynamics in a formal pricing model.

In other related work, various authors (Binsbergen et al., 2012, Cejnek and Randl, 2016, Golez 2014) focus on realized returns of short and long-term horizon dividend derivatives or forward dividend prices derived from stock index futures and options, and find evidence for a downward sloping term structure of risk premiums. Wilkens and Wimschulte (2010) compare dividend derivative prices with dividend prices implied by index options. Suzuki (2014) assumes risk that premiums are proportional to dividend volatility and then models the dividend growth curve implied by derivative prices using a Nelson-Siegel approach.

This paper is organized as follows. The next section deals with the theory of dividend expectations and their fit into the present value model. It lays out the state space model which parameterizes the dividend term structure. Dividend swaps and futures and the treatment to prepare them for empirical tests is next. The empirical results are discussed in the subsequent section. These results are used for a reconciliation to the stock market in section five. Several robustness checks and a comparison of results to structural Macro models follow in section six, before the paper summarizes its conclusions in the closing section.

2.2 Theory

This section starts by proposing the general framework for discounted risk-adjusted dividend growth, represented in terms of a stochastic discount factor. The section continues to lay out the state space model to capturing time- and horizon-varying dividend growth.

2.2.1 The general framework

To apply the present value framework, we define g_{t+1} as the realized dividend growth rate for period t to $t+1$, so that the dividend payable at maturity n is: $D_{t+n} = D_t exp\left(\sum_{i=1}^{n} g_{t+i}\right)$. We then apply the standard asset pricing equation to price this payoff for maturity n, where its current present value $P_{t,n}$ equals the expected product of the pricing kernel and the payoff:

$$P_{t,n} = E_t\left[D_t \exp\left(\sum_{i=1}^{n} m_{t+i}\right) \exp\left(\sum_{i=1}^{n} g_{t+i}\right)\right], \tag{2.1}$$

and where m_{t+1} is the log pricing kernel for period t to $t+1$. The pricing kernel consists of the one-period risk-free rate y_t and an additional term θ_{t+1}:

$$m_{t+1} = -(y_t + \theta_{t+1}), \tag{2.2}$$

where y_t is observed at time t and reflects the risk free return over the period t to $t+1$.[1] We aim to model a combined growth variable for the present value of future dividends and rewrite the pricing formula (2.1) accordingly:

$$P_{t,n} = D_t \left[E_t exp\left(\sum_{i=1}^{n} \pi_{t+i} \right) \right]. \tag{2.3}$$

Equation (2.3) shows that the basic building block of the term structure model is what we denote discounted risk-adjusted dividend growth:

$$\pi_{t+1} = g_{t+1} - y_t - \theta_{t+1}. \tag{2.4}$$

In our data we observe dividend futures or swap prices. The relation of dividend present values to the prices of these dividend derivatives is achieved by discounting the futures prices at the n-period risk-free rate $y_{t,n}$:

$$P_{t,n} = F_{t,n} exp\left(-n y_{t,n}\right), \tag{2.5}$$

which demonstrates that dividend present values are observable directly from market data $F_{t,n}$ and $y_{t,n}$.

If the risk-adjusted growth rate π_t follows a lognormal distribution, equation (2.3) can be rewritten as:

$$lnP_{t,n} - lnD_t = E_t \left(\sum_{i=1}^{n} \pi_{t+i} \right) + \frac{1}{2} Var_t \left(\sum_{i=1}^{n} \pi_{t+i} \right). \tag{2.6}$$

The left-hand-side variable is related to the key modeling variable of Binsbergen et al. (2013). Specifically, they refer to $-(lnP_{t,n} - lnD_t)/n$ as the *equity yield*.

One may ask why we choose to model π_{t+1}, rather than to assume separate models for its elements dividend growth, risk premium and risk-free discount rates. Decomposition of stock prices into dividend growth and risk premiums knows many attempts, seminal among which is the VAR based approach by Campbell and Shiller (1988). Information from dividend

[1] To be precise, y_t is defined as the continuously compounded, one period risk free rate. Using the relation $exp(-y_t) = E_t[exp(m_{t+1})]$, it follows that the conditional expectation of θ_{t+1} must equal half the conditional variance of θ_{t+1} if the pricing kernel follows a lognormal distribution.

derivatives is also used in the VAR model of Binsbergen et al. (2013). We choose to do the exact opposite of decomposition and instead amalgamate the three variables into one; the proposed model variable is the growth rate of present values of expected dividends π_{t+1}. This amalgamation facilitates to focus on the term structure of the discounted growth trajectory alone. Connecting these growth rates via the present value identity to the stock market allows for a judgment call on the relevance of the horizon decomposition without being side tracked by additional assumptions on the constituent variables. In fact, since we aim to value the stock market as the sum of dividend present values, a decomposition is not needed.

Furthermore, the components of π_{t+1} are likely to be correlated. For example, Bekaert & Engstrom (2010) calculate the correlation between 10 year nominal bond yields and dividend yields in the US over a 40 year period at no less than 0.77. Binsbergen et al. (2013) perform a principal components analysis of equity yields based on dividend derivatives prices. They show that the first two principal components of nominal yields explain about 30% of $g - \theta$ movements. Taken together into a single variable π_{t+1}, it should be possible to model it with a limited number of factors due to the high correlation among its components.

2.2.2 The state space model

In order to build a full term structure of discounted risk-adjusted dividend growth, we model it in state space form. We discuss the state equations and the measurement equations.

State equations

The crucial question is how to model the evolution of risk-adjusted growth rates π_{t+1}. The approach that we advocate is a decomposition of π_{t+1} *by horizon*. Its growth rates differ by maturity, the pattern of which is the object of this paper. Our modeling approach to execute the decomposition by horizon closely follows Jegadeesh and Pennacchi (1996), who propose a model for estimating Libor futures with an aim to construct a term structure of interest rates based on three horizons. Their set up is a state space model in which the short-term interest rate is a latent variable. The prices of the Libor futures of different horizons are estimated by an equation consisting of the interest rates growth for the three horizons. Instantaneous growth and medium-term growth are both factors, infinite growth is a constant. This approach falls into the set of affine term structure models. Dai and Singleton (2000) derive the most general versions of affine term structure models, allowing for time-varying volatilities and time-varying risk premiums. We choose a rather restrictive two-state model with constant volatilities, and show that such a simple approach already generates a very good fit of the dividend term structure.

In this paper, we model discounted risk-adjusted dividend growth according to the same horizons. We specify most of the model in discrete time, following the approach in Campbell, Lo and MacKinley (1997). Specifically, we model π_{t+1} as the sum of a time-varying conditional mean p_t and a stochastic shock:

$$\pi_{t+1} = p_t + \nu_{t+1}, \tag{2.7}$$

where ν_{t+1} is normal i.i.d. with zero mean. Using the definition of π_{t+1} in equation (2.4), we can interpret p_t as the one-period ahead expected dividend growth minus the expected log of the pricing kernel

$$p_t = E_t g_{t+1} - y_t - E_t \theta_{t+1}. \tag{2.8}$$

The stochastic shock ν_{t+1} then is composed of the unexpected dividend growth and the stochastic part of the pricing kernel

$$\nu_{t+1} = g_{t+1} - E_t g_{t+1} - \left(\theta_{t+1} - E_t \theta_{t+1}\right). \tag{2.9}$$

The short-term factor p_t follows a mean reverting process to a medium-term factor $\tilde{p}_t$ which itself is mean reverting to a long-term constant $\overline{p}$, where for convenience we first define their processes in continuous time:

$$dp_t = \varphi \left(\tilde{p}_t - p_t\right) dt + \sigma_p dW_p, \tag{2.10}$$

$$d\tilde{p}_t = \psi \left(\overline{p} - \tilde{p}_t\right) dt + \sigma_{\tilde{p}} dW_{\tilde{p}}. \tag{2.11}$$

dW_p and $dW_{\tilde{p}}$ are Wiener processes, with σ_p and $\sigma_{\tilde{p}}$ scaling the instantaneous shocks to the factors. The horizon at which investors adjust their growth expectation from one state to the next is captured by mean reversion parameters φ and ψ. This two-state system results in the state equations for discrete intervals[2]:

$$\begin{aligned} \begin{pmatrix} p_{t+1} \\ \tilde{p}_{t+1} \end{pmatrix} &= \begin{pmatrix} 1-e^{-\varphi} & -\frac{\varphi}{\varphi-\psi}\left(e^{-\psi}-e^{-\varphi}\right) \\ 0 & 1-e^{-\psi} \end{pmatrix} \begin{pmatrix} \overline{p} \\ \overline{p} \end{pmatrix} + \\ & \begin{pmatrix} e^{-\varphi} & \frac{\varphi}{\varphi-\psi}\left(e^{-\psi}-e^{-\varphi}\right) \\ 0 & e^{-\psi} \end{pmatrix} \begin{pmatrix} p_t \\ \tilde{p}_t \end{pmatrix} + \varepsilon_{t+1}. \end{aligned} \tag{2.12}$$

Finally, we model correlation between the innovation in the growth rate ν_{t+1} and the errors ε_{t+1} in these state equations as $\nu_{t+1} = \beta' \varepsilon_{t+1}$, where $\beta = (\beta_p, \beta_{\tilde{p}})'$ is a 2-by-1 vector.

[2]Refer to Appendix B for further details.

One could incorporate an independent shock to the growth rate, but this does not have an important effect on the term structure of dividend prices or the dynamics of these prices. In terms of the mathematical structure, this setup resembles the approach of Campbell, Lo and MacKinley (1997). They derive affine term structure models in discrete time by modeling the log-pricing kernel, $m_{t+1} = -(y_t + \theta_{t+1})$, in a similar way as we model the discounted risk-adjusted growth rate $\pi_{t+1} = g_{t+1} - (y_t + \theta_{t+1})$. The key difference is that our growth variable depends both on the pricing kernel and the dividend growth rate. As discussed above, we only model the aggregate variable π_{t+1} and do not need to make specific assumptions on its components. This is important for the interpretation of the results. For example, when modeling interest rates, Campbell, Lo and MacKinley (1997) show that the β vector captures the risk premiums on long-term bonds. In our setup, the vector β could represent dividend risk premiums, but can also be the result of correlation of current dividend growth and the factors driving future dividend growth. Again, for pricing dividend derivatives there is no need to specify the source of the correlation between shocks to π_{t+1} and the factors.

Measurement equations

Given the dynamics of π_{t+1}, it follows that the average growth rate of dividend present values from time t to its expiry date at time n corresponds to a function of p_t and $\tilde{p}_t$. Specifically, as shown in Appendix B, filling in the dynamics of π_{t+1} in the pricing equation (2.6) and adding i.i.d. measurement error $\eta_{t,n}$ for each derivatives maturity n, the measurement equations for the state space model are:

$$\begin{aligned} lnP_{t,n} - lnD_t \; &= n\overline{p} + \varphi_n \left(p_t - \overline{p}\right) + \frac{\varphi}{\varphi - \psi}\left(\psi_n - \varphi_n\right)\left(\tilde{p}_t - \overline{p}\right) + \\ \frac{1}{2}\sum_{i=1}^{n} &\left(\sigma_p^2(\beta_p + \varphi_i)^2 + \sigma_{\tilde{p}}^2\left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi - \psi}\left(\psi_i - \varphi_i\right)\right)^2\right) + \eta_{t,n}, \end{aligned} \tag{2.13}$$

in which β_p and $\beta_{\tilde{p}}$ are the covariance betas of the errors of the first and second factor and σ_p^2 and $\sigma_{\tilde{p}}^2$ are their variances. We define φ_n and ψ_n as follows:

$$\varphi_n = \frac{(1 - e^{-n\varphi})}{(1 - e^{-\varphi})},$$

$$\psi_n = \frac{\left(1 - e^{-n\psi}\right)}{\left(1 - e^{-\psi}\right)},$$

with $\varphi_0 = 0$ and $\psi_0 = 0$.

2.2.3 The single-state model

We benchmark the ability of the two-state model to fit the dividend term structure by a state space model with a single factor. In essence, the medium-term factor is set to the long-term constant estimate, rendering the same estimation equations as a Vasicek model:

$$dp_t = \varphi\left(\overline{p} - p_t\right)dt + \sigma_p dW_p. \tag{2.14}$$

Its state equation and measurement equations are:

$$p_{t+1} = \overline{p} + \left(p_t - \overline{p}\right)e^{-\varphi} + \varepsilon_{t+1}, \tag{2.15}$$

$$lnP_{t,n} - lnD_t = \; n\overline{p} + \varphi_n\left(p_t - \overline{p}\right) + \frac{1}{2}\sum_{i=1}^{n}\left(\beta + \varphi_i\right)^2\sigma^2 + \eta_{t,n}. \tag{2.16}$$

2.3 Dividend swaps and futures

2.3.1 The market for dividend derivatives

The estimation methodology uses prices of dividend derivatives referring to four major stock markets: Eurostoxx 50 and Nikkei 225 dividend futures and S&P 500 and FTSE 100 OTC dividend swaps. Dividend futures were introduced in 2008 to the European market and in 2010 in Japan. Maturities extend out to ten years with annual intervals. Price data are available on a continuous basis from the relevant stock exchanges. Liquidity of European dividend futures is good with Euro billions of notional outstanding for maturities up to three years and hundreds of millions for the longest maturities (Table A1). The market for Nikkei dividend futures is smaller with maturities up to two years featuring notionals of over half a billion ($-equivalent) and tens to hundreds of millions for longer maturities (Mixon and Onur, 2017). All maturities normally trade on a daily basis and we apply the estimation procedure to daily prices in the case of dividend futures.

Before 2008, dividend derivatives existed as dividend swaps traded over-the-counter (OTC) only. They date back to 2002, well before the onset of listed futures. Maturities extend out to ten years and more for Eurostoxx 50, Nikkei 225, FTSE 100 and the S&P 500. We obtained dividend swap price data from several investment banks for all four stock indices mentioned[3], but there are problems. Before 2005, prices are stale and, throughout the data period prices, not always consistent which each other among suppliers. Moreover, turnover is very low with days passing by without a single trade taking place across all

[3] Deutsche Bank, Goldman Sachs and Credit Suisse.

maturities more often than not[4].

We nonetheless perform the estimations with data sets both of listed dividend futures at daily frequency and OTC dividend swaps at monthly frequency. The main conclusion from these results is that, although OTC price data originate from pricing models, they still explain variation in stock index levels.

Dividend derivatives exchange the value of a dividend index for cash at set expiry dates. The difference between the transaction price and the amount of dividends actually paid is the amount settled between buyer and seller. The transaction price reflects the growth path expected from the current level of dividends and the premium required for the risk of the actual payment differing from what is expected. It is a risk-adjusted price and equals the present value of a dividend once the time value of money is accounted for (see equation (2.5)).

The dividend index measures the amount of dividends paid by the companies constituent to a stock index during a calendar year[5]. At the end of the year, the index equals the fixing at which the dividend derivative is settled. Manley and Mueller-Glissmann (2008) provide an overview of the market for dividend derivatives and its mechanisms.

Listed futures on dividends paid by the companies in the Eurostoxx 50 and the Nikkei 225 indices are the main subject of this paper. Dividend futures are available for other markets as well. They are referenced to dividends of the FTSE 100, Hang Seng and Hang Seng China Enterprises and several other less liquid markets, and since 2016 also for the S&P 500. Only for the Eurostoxx 50 and the Nikkei 225 dividend index futures are traded with a maturity range of up to ten years, the other markets extend out to four years[6]. The purpose of this paper is to estimate a term structure of dividend risk-adjusted growth, for which longer dated maturities are required, over a reasonably long history. We therefore exclude the other markets from the data set and focus on Europe and Japan.

2.3.2 Constant maturity construction

Dividend derivatives usually expire at a fixed date near the end of the calendar year[7] and therefore their time to maturity shortens by one day for each day that passes. For application in the state space model, growth rates of a constant horizon are required. The horizons of the measurement equations regard annual increments, the state equations regard one day increments. To obtain growth rates from prices with constant maturities, we interpolate

[4]Dividend swaps are said not to trade daily, "sometimes not even for months". Turnover figures are not public, but Mixon and Onur (2017) provide further insight.

[5]Derivatives relating to dividends paid by individual companies exist as well.

[6]With exception of the S&P 500 dividend futures, which trade out to ten years too.

[7]The Nikkei 225 dividend index runs until the last trading day in March.

derivatives with adjacent expiry dates. The interpolation is weighted by a scheme which reflects the uneven distribution of dividends through the year. For example, in the spring season 60% of the Eurostoxx 50 dividends of a full index year are paid in a matter of a few weeks (Figure A1).

Derivatives prices which have a constant horizon from any observation date are constructed from observed derivatives prices. Such Constant Maturity (CM) derivative prices $F_{t,n}^{CM}$ take the following shape, attaching the seasonal pattern of the dividend index as weights to the observed derivatives prices w_i, with i standing for the day in the dividend index year, $i = 1$ being the first day of the count of the dividend index[8]:

$$F_{t,n}^{CM} = (1 - w_i) F_{t,n} + w_i F_{t,n+1}. \tag{2.17}$$

The weight w_i of the dividend index reflects the cash dividend amount paid as a proportion of the total amount during a dividend index year. The average of the years 2005 to 2013 is taken. $F_{t,n}$ is the observed price of the derivative which expires n^{th} in line into the future from the observation date onwards, $F_{t,n+1}$ expiring the following year. This weighting scheme reduces the impact of the n^{th} derivative to expire on the constant maturity derivative as time passes by the proportion w_i of dividends that have actually been declared. Its complement $(1 - w_i)$ is the proportion that remains to be declared until the expiry date and is therefore an expectation of undeclared dividends for year n at the observation date. In order to produce a derivative price with constant maturities, this undeclared amount is balanced by the proportion of the price of the derivative expiring the year after. In so doing, the constant maturity price reflects no seasonal pattern, while still accounting for the seasonal shift in impact from the n^{th} derivative to the next. For example, during the dividend season in Spring, the weight is shifted more quickly from the first to the second derivative[9] than in other parts of the year[10].

2.3.3 Dealing with current dividends

At the heart of the present value model are the discounted values of risk-adjusted dividends. These present values $P_{t,n}$ take current dividends D_t as the starting point from which growth is projected forward at growth rate π_{t+i} (equation (2.3)). It is sometimes assumed[11] that

[8] which is the first trading day following the expiry date of a dividend derivatives contract.

[9] First and second derivatives is shorthand for the derivatives that are first and second to expire.

[10] A linear weighting scheme would reflect the adjacent derivative prices unevenly. For example, half way through the dividend index year already 80% of annual dividends is declared and paid. Linear weighting would then overemphasize the information contained in the price of the derivative in equation (2.17) that is the soonest to expire.

[11] E.g. Binsbergen et al. (2013), Cejnek and Randl (2016).

current dividends can be reasonably approximated by realized dividends. For daily data as applied in this paper, however, this assumption causes issues.

The asset underlying dividend derivatives is the amount of cash dividend thrown off by a stock or a stock index during the year in which the derivative expires. The index companies pay dividends throughout the calendar year[12] which implies that taking realized dividends as current dividends at a certain day of the year would require looking back for twelve months. The dividend paying capacity of index companies does not stay constant for a year, hence a twelve month backward-looking dividend measure will not accurately reflect current dividends.

To take a strong example, around the days of the Lehman bankruptcy on the 15^{th} of September 2008, the one year dividend history of Eurostoxx 50 companies amounted to 154. Due to the bankruptcy, investors would have changed their opinion strongly downwards about the dividend that companies would pay if they would have had to pay on these days. Even if dividends reflect the past year of earnings, company management is likely to reduce dividends if their near term outlook changes for the worse by precautionary motive. After Lehman, taking a dividend history of twelve months would then overestimate current dividends as they stood in the fall of 2008. In the weeks following the default, the Eurostoxx 50 dividend future expiring in 2009 dropped from 140 to 100. Therefore, if twelve month realized dividends are used as current dividends, the shortest horizon observation for growth from 2008 to 2009 on the dividend curve would attain a strongly negative figure even though the actual growth expectation, starting from a level that would have been revised downwards, could be flat or even positive.

This problem rules out considering the dividend index itself, or a rolling twelve month estimate of it, as a starting point from which to calculate the growth rate until the first derivative to expire. The first derivative to expire would also not perfectly capture current dividends. The first derivative contains investor expectations about dividends to be paid in the remaining period until the first expiry date and is not a reflection of current dividends on the observation date itself.

To avoid these data difficulties, we propose an alternative base. In lieu of an estimate for current dividends, we use dividend derivatives with one year remaining life to expiry as the base from which to calculate growth rates:

$$P_{t,1}^{CM} = F_{t,1}^{CM} exp\,(-y_{t,1}), \tag{2.18}$$

[12]In fact, the dividend index year usually runs from the first working day following the third Friday in December until and including the third Friday in December of the following year. Dividend derivatives also apply the third Friday of December as the expiry date.

and the first year of growth is deducted from the subsequent growth path accordingly. Discounted risk-adjusted dividend growth rates are then given by:

$$n\pi_{t,n} - \pi_{t,1} = ln\left(F_{t,n}^{CM}\right) - ny_{t,n} - \left(ln\left(F_{t,1}^{CM}\right) - y_{t,1}\right). \tag{2.19}$$

As a consequence of estimating $n\pi_{t,n} - \pi_{t,1}$ as a single variable, we do not account for the first year of discounted dividend growth as part of the dividend term structure. At the same time, the one year dividend present value $P_{t,1}^{CM}$ includes short-term derivatives prices which encompass investor expectations extending from the observation date until a year later. Although growth for the first year is not observed, the one year discounted derivative price is included in the present value identity ensuring that no information is lost when reconciling the model estimates to the stock market due later in this paper.

Subtracting the first period present value gives the following measurement equation for growth rates and replaces equation (2.13):

$$\begin{aligned} lnP_{t,n} - lnP_{t,1} = {} & (n-1)\,\overline{p} + \varphi_n\left(p_t - \overline{p}\right) + \frac{\varphi}{\varphi - \psi}\left(\psi_n - \varphi_n\right)\left(\tilde{p}_t - \overline{p}\right) + \\ & \frac{1}{2}\sum_{i=1}^{n-1}\left(\sigma_p^2(\beta_p + \varphi_i)^2 + \sigma_{\tilde{p}}^2\left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi - \psi}\left(\psi_i - \varphi_i\right)\right)^2\right) + \eta_{t,n}. \end{aligned} \tag{2.20}$$

State equations (2.10) and measurement equations (2.20) together form the system of which the variables are estimated by maximum likelihood. The procedure is recursive by means of a Kalman filter (Jegadeesh and Pennacchi, 1996).[13]

2.4 Empirical results

Tables 2.1 and 2.2 provide the results of the two-state model and a benchmark single-state model for Eurostoxx 50 and Nikkei 225 dividend markets – the two markets for which listed futures data with sufficiently long horizons of 10 years exist. Estimations are performed on daily data[14].

[13]The error variance terms are assumed to be the same for all measurement equations (σ_η), except for the first one (which we denote σ_η^1). This is because the definition of the first derivative to expire (set to a constant maturity of one year following the observation date) differs slightly from subsequent derivative prices due to an alternative weighting scheme for finding constant maturity values as explained in Appendix A.

[14]For robustness, we perform the same tests with monthly data (not shown here). None of the parameter estimates and test coefficients change meaningfully relative to the daily data set.

2.4.1 Pricing errors

Before we discuss the parameters of the growth rate model, we first establish that the two-state model fits the data well[15]. To this end, we calculate mean absolute errors for the measurement equations (2.20). Given that they are specified for log prices of dividend futures, these mean absolute errors can be interpreted as relative pricing errors[16]. The first measurement equation produces a mean absolute pricing error of 0.015 (1.5%) and pricing errors of subsequent expiries are between 0.002 and 0.005 (Figures 2.3 and 2.4). The error levels are clearly small, confirming a good fit of the model to the data. A test for serial correlation in the residuals of the first measurement equation and potential impact on the parameters is conducted in the robustness section.

2.4.2 Mean reversion estimates

The mean reversion towards medium-term growth φ attains levels which translate to a half-life of less than a year. The Eurostoxx 50 mean reversion at 1.51 is twice as fast as for the Nikkei 225 ($\varphi = 0.74$), which is due to the global credit crisis in 2008/09 being included in the Eurostoxx 50 data period and not in the Nikkei 225 data period.[17] Mean reversion towards the long run constant ψ is broadly measured in half-lives of 3 to 4 years in both markets, a space of time that comes close to that of a business cycle. All mean reversion parameters are significant at the 1% level. The estimates for φ and ψ are positive, which implies that the growth rate is stationary and thus tends to a long-term constant.

A benchmark for the speed of reversion cannot be provided since there are no other attempts in the literature to fit the dividend term structure. Jegadeesh and Pennacchi (1996) apply the two-state model to interest rates and find the opposite pattern; at a half-life of 4.5 years short mean reversion is slower than medium-term mean reversion at 2.3 years in interest rates. The first factor thus mean reverts much faster in dividends ($\phi = 1.51$ equals 0.5 year half-life) than in bonds, whereas the second states are comparable.

The model imposes the long-run growth rate to be constant, while the speed at which medium-term growth adjusts to it is estimated from the data. The interpretation from these results is that investors change their opinion about growth only as far ahead as the anticipated business cycle. We do not formally link an economic interpretation to the three growth stages, but given the estimates of the mean reversion parameters some intuition can

[15] The short term beta β_p is set to zero, as discussed further below.

[16] These errors are thus not annualized. Transformed to annual growth rates, the errors are even smaller.

[17] Estimating the model for the Eurostoxx 50 data over a partial data period that coincides with the Nikkei 225 data period yields mean reversion parameters that are closer to those found for the Nikkei 225: $\varphi = 0.88$ and $\psi = 0.09$.

be provided. Instantaneous growth can be thought of as the expectation of the immediate future. Shocks to risk aversion and to the volatility of the current business climate are likely to influence investors' valuations of dividends several months ahead, but perhaps not much further. Developments in the business cycle, on the other hand, such as credit conditions, investment growth and monetary policy set the stage for the business cycle influencing dividend expectations over a longer period ahead, measured in several years. Structural factors such as population growth and technological progress determine how investors perceive the long run, extending from the business cycle horizon into the infinite future.

Structural developments should be slow moving, if at all, and are approximated by imposing asymptotic constancy. Thus, at horizons extending well beyond business cycles, investors may have time-varying opinions of economic and financial variables, but they do not change them once taken together. This means that any rise in long-maturity interest rates is exactly offset by a rise in long-term dividend growth or a fall in long-term risk premiums. Mean reversion towards such a constant implies therefore that a horizon exists at which investors never change their opinion about present value growth.

2.4.3 Discounted risk-adjusted dividend growth rates

Given the mean reversion estimates, the instantaneous factor reflects short-term movements in risk-adjusted growth, the medium-term factor reflects an assessment of the business cycle, while $\overline{p}$ depicts a structural level which can be linked closely to the dividend yield. The four panels in Figure 2.5 provide estimates of expected growth rates by recalculating the factors by means of the measurement equations (2.20) into 1-year growth and 1-year forward 4-year growth of discounted risk-adjusted dividends. Forward growth rates imply the level of growth expected after the 1 year growth rate has materialized[18].

1-Year growth is mostly determined by the instantaneous factor. Panel 2.5(a) shows that it is highly volatile for the Eurostoxx 50, with the global credit crisis in 2008/09 showing a decline by nearly half and during the Eurozone sovereign debt crisis in 2011 by a quarter. Outside these periods, it moves between broadly − 10 and + 5 percent. Nikkei 225 1-year growth rates move in the same range until late 2012 (Panel 2.5(b)). The period following the announcement of "Abenomics" in 2012[19] portrays high optimism with 1-year growth rates attaining 10 percent and more.

Given the values found for the mean reversion parameters, the medium-term factor largely

[18] As discussed later, the short term beta is set to zero for these data, but different fixed levels do not materially change growth rates.

[19] Late 2012 the government of Shinzo Abe proclaimed a policy of monetary and fiscal expansion combined with economic reform. The two main consequences for financial markets were a substantial weakening of the Japanese Yen and a rise in the stock market.

determines 1-year forward 4-year growth depicted in Panels 2.5(c) and 2.5(d). In Europe forward growth circles around the long run constant between – 2 and – 6 percent. The sovereign debt crisis in 2011 shows a somewhat more negative rate than the global credit crisis. Investors apparently expected that the serious short-term blow to dividends in 2008/09 would not be corrected or reversed (by positive growth) afterwards. However, the less negative blow in 2011 would be followed by a period more negative than the long run constant (Panel 2.5(b)), implying that investors expected that the European sovereign debt crisis would bear consequences for the business cycle.

2.4.4 The volatility term structure

The volatility of dividend futures prices across maturities provides further insight into the relation between the risk and the maturity of dividends. Specifically, we calculate the annualized variance of changes in the log dividend price for each maturity n, both as observed in the data and as implied by the two-state model:[20]

$$\begin{aligned}\sigma_t^2\left(lnP_{t+1,n} - lnP_{t,n}\right) = \sigma_t^2\left(\Delta lnP_{t,1}\right) + exp\left(\sigma_p^2\right)\left(\sum_{i=1}^{n-1} exp\left(-\varphi i\right)\right)^2 + \\ exp\left(\sigma_{\tilde{p}}^2\right)\left(\frac{\varphi}{\varphi - \psi}\left(\sum_{i=1}^{n-1} exp(-\varphi i) - \sum_{i=1}^{n-1} exp(-\psi i)\right)\right)^2,\end{aligned} \tag{2.21}$$

Figure 2.6 shows that the Eurostoxx 50 and the Nikkei 225 dividend markets portray an increasing but concave volatility curve as maturities increase, and that our model fits this pattern quite accurately. The concavity of the volatility term structure is consistent with the fast mean reversion of the two state variables in our model. The volatility of the long maturity dividend futures prices converges to a value of more than 20%, consistent with typical values for the volatility of the stock market return. We return to this point, and discuss the relation with macro asset pricing models in the subsection on OTC data (6.4), where we calibrate the model to U.S. data.

[20]We do not form a model for the volatility of current dividend changes, instead we choose this volatility to match the 1-year observed volatility exactly.

2.4.5 Long-term growth and dividend yields

The economic interpretation of the long-term discounted risk-adjusted dividend growth constant is briefly recapitulated. The present value identity for stock prices S_t is recalled as:

$$S_t = \sum_{n=1}^{\infty} P_{t,n} = D_t \sum_{n=1}^{\infty} exp(n\pi_{t,n}), \tag{2.22}$$

where $\pi_{t,n}$ is the observed annualized discounted risk-adjusted growth rate of dividends payable at maturity n, $\pi_{t,n} = (lnP_{t,n} - lnD_t)/n$, which is the negative of what Binsbergen et al. (2013) call the equity yield. The dividend-price ratio is found by rearranging identity (2.22) to:

$$\frac{D_t}{S_t} = \frac{1}{\sum_{n=1}^{\infty} exp(n\pi_{t,n})}, \tag{2.23}$$

in which $\pi_{t,n}$ is the discounted risk-adjusted growth rate of dividends payable at maturity n. For the sake of interpretation, if both factors in the two-state model are equal to the mean, $\pi_{t,n}$ is a horizon invariant constant and identity (2.23) simplifies to:

$$\frac{D_t}{S_t} = -\overline{p}^*. \tag{2.24}$$

The dividend yield equals the negative of long-term growth for which the state space approach thus provides an estimate. Combined with the constant convexity term, the estimate for $\overline{p}$ constitutes a measure of long-term growth $\overline{p}^*$. On an annual basis, this measure is given by:

$$\overline{p}^* = \overline{p} + \frac{1}{2}\left(\sigma_p^2(\beta_p + \varphi_{i\to\infty})^2 + \sigma_{\tilde{p}}^2\left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi - \psi}\left(\psi_{i\to\infty} - \varphi_{i\to\infty}\right)\right)^2\right), \tag{2.25}$$

in which the values for $\varphi_{i\to\infty}$ and $\psi_{i\to\infty}$ are set for i approaching infinity. The second factor sigma $\sigma_{\tilde{p}}^2$ is small, so the term $\sigma_p^2(\beta_p + \varphi_{i\to\infty})^2$ delivers most of the impact.

For longer horizons, the convexity term on the right hand side of measurement equations (2.20) approaches constancy. It includes the betas and the factor sigmas. While the sigmas are identified by means of the state equations, the betas are only present in the measurement equations. The long-term growth parameter $\overline{p}$ balances out the betas under the optimization procedure. As a result, the estimates for long-term growth $\overline{p}$ as well as of both covariance betas β_p and $\beta_{\tilde{p}}$ come out unstable with sizable standard errors. The estimation technique of the Kalman filter finds optimal solutions for various combinations of $\overline{p}$ and betas, a fact which indicates multicollinearity.

Finding a relevant and well identified value for long-term growth thus requires fixing the short term beta β_p while solving the model for the other variables. In Figure 2, the results of this exercise is shown. There is a consistent parabolic trade-off between the long-term growth constant and the short term beta, which is as expected in view of the quadratic nature of the convexity term.

The difference between $\overline{p}^*$ and $\overline{p}$ is smallest for values of β_P equal to $-\varphi_{i\to\infty}$. For these values, the inverted parabola in Figure 2 reaches its maximum. The interdependence between β_P and $\overline{p}$ in the output of the estimation procedure is such that different combinations along the parabola render very little influence on the value of $\overline{p}^*$.

In Tables 2.1 and 2.2 the parameters are shown of a single estimation in which β_p is set to zero. Long-term growth obtains reasonable levels, a discussion of which follows below. Standard errors show the estimates are (close to) significant at the 5% level. The mean reversion estimates and the variance estimates do not change materially when the short term beta is fixed. This is as expected since they have only a small impact in the long run, approaching zero impact at the limit. The medium term beta $\beta_{\tilde{p}}$ remains large but insignificant. For the purpose of all of the subsequent discussion, the short term beta is thus set to zero.

Seen in this light, there is an economic rationale in the estimates from the state space model for the long-term growth constant $\overline{p}^*$. The levels found equal – 2.6 percent in Japan and Europe, which appears reasonable relative to dividend yields (2.24). Table 2.3 contains some metrics for comparison. The average dividend yield in Europe was 4.3 percent and in Japan it was 1.9 percent during the short data period. The average 1 year forward 4 year growth rate also deviates less than 1 percent from the average dividend yield, but the average short-term growth rate deviates substantially more. A tentative conclusion is that the business cycle stood close to the long-term average during the data period, but sentiment was more negative in Europe and more positive in Japan[21]. Overall, the estimates for long-term growth seem a fair assessment of the long-term cash run rate of the stock market. It is noteworthy that the estimates are produced without input from the stock market itself.

It is also important to observe that the state space model estimates discounted long-term growth to be negative, since present value theory requires stock valuations to be finite. The flexibility of the model would allow for positive values, but the estimates correctly imply that dividend present values decline at a horizon that is sufficiently long.

[21] In fact, in particular in Japan it turned more positive during the data period.

2.4.6 The Dividend Term Structure

Equipped with model estimates for the growth parameters, a Dividend Term Structure (DTS) can be calibrated. The DTS depicts the present values that investors attach to expected dividends per horizon n expressed as a proportion of the total present value:

$$DTS_n = \frac{\hat{P}_{t,n}}{\sum_1^\infty \hat{P}_{t,n}}. \tag{2.26}$$

The value for $\hat{P}_{t,1}$ is the calibrated discounted price of the derivative expiring one year from t. The values for subsequent expiries $n \geq 2$ are calibrated from the estimated growth parameters. Figure 2.7 shows that the average term structure of the Nikkei 225 starts sloping upwards, and then becomes downward sloping as the horizon increases. The transition is slow given the low mean reversion and the moderate levels of the estimated averages for instantaneous and medium-term growth. The Eurostoxx 50 DTS, by contrast, is strongly negatively sloping at the outset, but adjusts to the long-term growth path rather quickly. The first dividend point on the Eurostoxx 50 DTS is therefore high, which translates into an equally high current dividend yield. The Nikkei 225 first dividend points are lower on average, which fits with the positive slope at the start of their DTS. It is also in line with the fact that the estimate for long-term growth is somewhat higher than the first dividend point.

Its DTS indicates that the fundamental value of the European stock market is more front loaded, or more heavily weighted towards the near future, than that of the Japanese stock market. The surface below the calibrated DTS equals one by definition. The relative present values of dividends of Japan cross over the European values after about thirty years into the future. Relative to the European stock market, the present value of Japanese dividends beyond the cross over makes up for their lower contribution before it.

2.4.7 Other long-term growth estimates

Giglio, Maggiori and Stroebel (2014) compare prices of houses of different contractual ownership to arrive at a very long-term discount rate. Leased housing reverts to the owner of the land after the lease expires, while freehold housing remains with the owner of the house indefinitely. The difference in price between the two for comparable properties equals today's present value put to ownership once the lease has expired. At lease expiries of over one hundred years, this provides an interesting comparison to the estimates for long-term discounted risk-adjusted dividend growth.

The discounts Giglio et al. (2014) find in the data equate to a value for infinite growth of around – 2% for periods of 100 years and more. This level makes sense economically and is also reasonably close to the long-term discounted risk-adjusted dividend growth estimates[22].

2.5 Reconciliation to the stock market

The second part of our research agenda is to analyze the implications of the model for the value of the stock market. Given that we estimate the model using dividend derivative data only, this constitutes an out-of-sample test of the model. Alternatively, if one takes the model assumptions for granted, it can be seen as a relative pricing exercise of the dividend derivative prices versus stock market levels.

2.5.1 The empirical approach

The present value model incorporates expected index dividends which can be extrapolated from the estimated dividend term structure. This provides the following estimate for the stock market:

$$\hat{S}_t = D_t \sum_{n=1}^{\infty} exp(n\widehat{\pi}_{t,n}) = D_t \widehat{PD}_t, \tag{2.27}$$

with the summation of fitted growth rates $\widehat{\pi}_{t,n}$ equal to the estimated dynamic price-dividend ratio $\widehat{PD}_t$, and where the fitted growth rates satisfy[23]:

$$\begin{aligned} n\widehat{\pi}_{t,n} = n\overline{p} + \varphi_n \left(p_t - \overline{p}\right) + \frac{\varphi}{\varphi - \psi} \left(\psi_n - \varphi_n\right) \left(\tilde{p}_t - \overline{p}\right) + \\ \frac{1}{2} \sum_{i=1}^{n} \left(\sigma_p^2 (\beta_p + \varphi_i)^2 + \sigma_{\tilde{p}}^2 \left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi - \psi} \left(\psi_i - \varphi_i\right) \right)^2 \right). \end{aligned} \tag{2.28}$$

It is a well-known and critical problem of the present value model that it depends on a

[22]It is clear that not the level of the rents D, but only the growth of rents (being part of p) matters for establishing the lease discount. We can therefore consider growth in rents with or without maintenance cost, depreciation and taxes assuming they stay constant in proportion to rents over the very long-term considered. Another aspect is the convenience provided to the occupier of a house. Growth comparisons should be made only for sufficiently remote horizons. Since the notion of convenience yield is that there is a benefit to the current user that a future user cannot currently enjoy, nearer horizon comparisons are distorted.

[23]Subtracting the first growth rate $\pi_{t,1}$ from equation (2.22) provides an alternative representation which can be directly applied to the present value identity in equation (2.23):

$$\begin{aligned} n\widehat{\pi}_{t,n} - \widehat{\pi}_{t,1} = n\overline{p} + \varphi_n \left(p_t - \overline{p}\right) + \frac{\varphi}{\varphi - \psi} \left(\psi_n - \varphi_n\right) \left(\tilde{p}_t - \overline{p}\right) + \\ \frac{1}{2} \sum_{i=2}^{n} \left(\sigma_p^2 (\beta_p + \varphi_i)^2 + \sigma_{\tilde{p}}^2 \left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi - \psi} \left(\psi_i - \varphi_i\right) \right)^2 \right). \end{aligned}$$

reasonable estimate for the expected growth and the risk premium of dividends. Historical analysis of dividend growth followed by risk premium decomposition provides such estimates. (Campbell and Shiller, 1988). Binsbergen et al. (2013) execute the decomposition by making use of the price data of dividend derivatives. A key contention in this paper is that for the purpose of the present value model reconciliation, without decomposition of dividend expectations into growth and risk premiums, decomposition of risk-adjusted discounted dividend growth by horizon alone is very informative.

Successful reconciliation of dividend derivative price information to the stock market is uncommon in the literature. For example, Suzuki (2014) builds a Nelson Siegel model of the Eurostoxx 50 dividend growth term structure and makes assumptions about the level for longer dated values. These include a fixed level imposed at 4% for discounted growth after 25 years. Under these conditions, Eurostoxx 50 dividends reconcile well with the stock market dynamically.

In contrast to Suzuki (2014), we do not impose a fixed level as the state space model itself renders an estimate for the long-term growth path of the present value of dividends independent from stock market information, while it captures the shape and the dynamics of the term structure up to the medium-term at the same time. The entirety of the present value term structure is thus described by a handful of variables from two markets[24] in a single estimation procedure. The fit of the reconciliation to the observed stock market acts as a joint check on the validity and the robustness of the two-state model and the present value identity. To that end, equation (2.28) is used to calculate the fitted dividend growth rates and present values as implied by the estimated state space model.

All variables are taken as estimated by the state space model applied to dividend derivative data. Current dividends in (2.27) are approximated by the value of the first constant maturity derivative $F_{t,1}$, which is discounted at the risk-free rate. This is a better approximation for investors' estimate of current dividends than twelve month historical dividends. We thus get for the model implied stock market level:

$$\hat{S}_t = F_{t,1} exp\left(-y_{t,1}\right) \left(1 + \sum_{n=2}^{\infty} exp\left(n\widehat{\pi}_{t,n} - \widehat{\pi}_{t,1}\right)\right) = F_{t,1} exp\left(-y_{t,1}\right) \left(1 + \widehat{PD}_t^1\right), \tag{2.29}$$

in which $n\widehat{\pi}_{t,n} - \widehat{\pi}_{t,1}$ are the fitted values, estimated as a single variable, of the measurement variables in equations (2.20) and $\widehat{PD}_t^1$ represents the estimate for the price-dividend ratio as implied by the sum of exponential growth rates, where growth starts from the present value

[24] The interest rate swap market and the dividend derivative market.

of the dividend derivative expiring one year following the observation date[25].

2.5.2 Stock market level reconciliation

We first discuss the empirical results of the reconciliation with stock market levels[26]. In the European market, the two-state model estimates applied to equation (2.29) cause the stock index to be overestimated at a reasonably constant level distance to the actual stock index for most of the data period (Figure 2.8). There is no clear trend among the factors driving the estimated valuation away or towards the stock index. The historical dividend yield (4.3%) is somewhat higher than the negative of the long-term estimate (–2.6%) and the index is overestimated at some 20 to 30 percent except during the outbreak of the global credit crisis[27]. The level estimate of the stock index is highly sensitive to the long-term growth parameter. For the mean squared errors of this level comparison to be minimized, the estimate for long-term discounted growth would have to be closer to the historical dividend yield, or about 0.7% higher.

Dividend present values underestimate the Nikkei 225 index level at the beginning of the data period, but the gap closes from 2012 onwards. Short-term growth ranges between – 0.20 and + 0.05 percent initially, but at the onset of Abenomics in late 2012, it turns strongly positive (Figure 2.9). At – 2.6 percent, long-term growth is more pronounced than the average Japanese dividend yield (1.9%), which contributes to the underestimation[28].

2.5.3 Dynamic reconciliation

Following the present value model, *stock returns* are a consequence of investors changing their valuation of future dividends. The dynamics of stock indices can be retrieved from the present value model estimate as provided in equation (2.29). The present value of the first dividend amount to be paid over the year to come is the starting point of the growth term structure. The first dividend is observable and the growth path of discounted risk-adjusted

[25]The stock index estimate is approached by numeric summation, which is approximated by:

$$\hat{S}_t \approx F_{t,1}^{CM} exp\left(-y_{t,1}\right) \left[1 + \sum_{n=2}^{\overline{n}} exp\left(n\widehat{\pi}_{t,n} - \widehat{\pi}_{t,1}\right) + \frac{exp\left(\overline{n}\widehat{\pi}_{t,\overline{n}}\right)}{-\widehat{\overline{p}}^{*}}\right]. \tag{2.30}$$

In the estimations $\overline{n}$ is set at 50 years. Unless reduced to single digits, the number of years which $\overline{n}$ is set to is not material to the stock index estimates.

[26]The state space model estimations are produced setting the short term beta to zero.

[27]Some market participants consider Eurostoxx dividend derivatives prices around the turn of 2008/09 as unrepresentative of dividend expectations due to one-sided interests.

[28]A principal component analysis of the dividend growth rates turns out that the first two components explain over 99% of their total variance. Once taken as the regressors for the stock market similar to (30), the two principal components do not outperform our model: at 53%, the R2 is slightly lower.

dividends starting after it is a model implied estimate. The dynamic fit as well as the relative importance to stock returns of the first derivative on the one hand and the growth path on the other requires testing. For this reason the estimated returns of the stock market is split into its drivers. Equation (2.27) is repeated with logs denoted in lower case as a regression equation:

$$s_t = \alpha + \beta_f f_t + \beta_{\Delta y} \Delta y_t + \beta_{\widehat{pd}} \widehat{pd}_t + \varepsilon_t. \tag{2.31}$$

Stock index log returns are regressed by OLS on the log return of the first constant maturity derivative f_t, changes in the 1 year risk-free rate Δy_t and the log returns of the estimated price-dividend ratio $\widehat{pd}_t^1$, which is the sum of the normalized dividend present values of the state space model. The betas[29] of the returns of the first dividend and the price-dividend ratio are predicted to be close to + 1, while the beta of the risk-free rate is expected at – 1. Data are daily.

Eurostoxx 50 and the Nikkei 225 index returns respond well to the prediction of the present value model, shown in Table 2.4. The model is quite capable of explaining variation in stock returns, reaching an R^2 of above 50 percent. Although we cannot benchmark this explanatory power, it appears substantial given that the model does not incorporate any direct information of the stock market. Each of the regressors add considerably to the explanatory power, while the constant is close to zero. Both stock markets appear highly sensitive to changes in the first constant maturity derivative. The daily betas are in the order of 0.85 for the Nikkei 225 to 0.90 for the Eurostoxx 50. The beta of the price-dividend ratio is close to 0.86 and 0.66 respectively. In the case of Japan, most of the explanatory power comes from the price-dividend ratio, for Europe it is more evenly divided between short-term dividends and the price-dividend ratio.

The 1 year zero-coupon interest rate brings the price of the first derivative to its present value. Its relevance seems limited and the expected beta is – 1. It is highly significant in the estimates for the Eurostoxx 50, but reaches values of no more than 0.20. In Japan, the risk-free beta is close to zero.[30]

The interpretation of the assumption that long run discounted risk-adjusted dividend growth is constant is not that investors do not change their opinion about what value to attach to dividend present values far into the future. The value ascribed to dividends expected ten years and, for example, twenty years from today is influenced by the estimate of

[29] Coefficients of $\widehat{PD}_t^1$ are expected below 1 due to the errors in the regressor estimates increasing their variance.

[30] The impact of short-term dividends and the price-dividend ratio is mitigated by negative coefficients found once lags are added to the set of regressors (not shown here). This suggests that either the stock market overreacts to shocks to dividends, which is corrected in the following day, or that dividend prices may partly follow stock prices by at least a one day lag.

present values in the near term and medium-term. But the value of the twenty year dividend does not change relative to that of the ten year dividend regardless of changes in near and medium-term expectations – the relationship between them is (approximately) fixed. Therefore, long-run constancy excludes mean reversion to levels. Dividend levels attained in the past are not a target for investors to project their long-term expectations onto. Only long-term growth is.

2.6 Robustness

2.6.1 The single-state model

The two-state model distinguishes instantaneous from medium-term growth. Its ability to fit the dividend present value term structure is benchmarked by a state space model with a single factor, in which the medium-term factor is set to a long-term constant[31]:

$$dp_t = \varphi\left(\overline{p} - p_t\right)dt + \sigma_p dW_p, \tag{2.32}$$

Figures 2.3 and 2.4 show absolute estimation errors of the single-state model in comparison to the two-state model. While still not substantial, single-state estimation errors are larger by a factor of 2 to 3. The better fit of the two-state model is also indicated by the log likelihood statistics. Per observation the log likelihood contribution is a third higher in the two-state model than it is in the single-state variation (Table 2.1). The estimated parameters for the single-state model are significant and attain reasonable levels, but the two-state model is superior. The Eurostoxx 50 estimate for mean reversion at 1.73 is even higher than the short-term mean reversion in the two-state model, its standard error is larger. This adjustment speed implies a half-value time of instantaneous growth of only 5 months. Long-term growth is slightly lower and its standard error is smaller than in the two-state model. Figures 2.5(a) to 2.5(d) depict the forward discounted risk-adjusted dividend growth rates as delivered by both models. Single-state model growth rate are less volatile, which reflects the quick fading of the instantaneous growth factor.

For the Nikkei 225, the picture is rather different. Mean reversion attains a value in the middle ground of the two parameters in the two-state model. Long-term growth is somewhat lower, but again economically sensible. Standard errors are smaller for both parameters. The log likelihood contribution is again a third higher for the two-state model .The 1 year growth rate largely overlaps with that of the two-state model.

[31] Alternatively one can depict this model as a nested two-state model with medium-term mean reversion parameter ψ constraint to infinity.

Benchmarking against the single-state model indicates that it appears plausible to distinguish between investors gauging the immediate future on the one hand and their considerations about the business cycle on the other. A state space model with two factors caters for such a distinction.

2.6.2 Serial correlation in residuals

The estimated growth rates fit the data well, but the residuals exhibit some serial correlation in the growth rate from the first to the second constant maturity futures and swaps. This estimated growth rate sometimes varies from the data by several percentage points, although mostly during periods of stronger than average negative growth. Growth rates of longer horizons do not share this pattern. In order to check for the importance of this phenomenon on the estimated parameters, we allow for serial correlation in the measurement equation of the first growth rate. Followingle Eraker (2004), we assume a first order autoregressive structure for the first measurement equation ($n = 2$):

$$\eta_{t+1,2} = \ \xi\eta_{t,2} + u_{t+1} \tag{2.33}$$

The fit as measured by the log likelihood contribution improves. Since the short term growth rate captures most of the short term mean reversion, there should be some effect from an extra degree of freedom on the speed of mean reversion of the factors. However, the mean reversion parameters do not change drastically. In the case of Eurostoxx 50 dividend futures, φ decreases from 1.51 to 1.43, and ψ decreases from 0.24 to 0.22. For the other markets the impact is even smaller. So, we conclude that our estimates are robust and the two factor model seems to be correctly specified.

2.6.3 An alternative model

Our modeling approach focuses directly on discounted risk-adjusted dividend growth $\pi_{t+1} = g_{t+1} - y_t - \theta_{t+1}$. We thus incorporate discounting at the risk-free rate when valuing future dividends. An obvious alternative to this approach would be to model $z_{t+1} = g_{t+1} - \theta_{t+1}$ using a term structure model to value dividend derivatives, and subsequently discount it at observed interest rates to calculate present values. This latter step requires the assumption that interest rates and $g_{t+1} - \theta_{t+1}$ are independent. In addition, we assume the expectations hypothesis holds for bonds, so that bond risk premiums equal zero and long-term interest rates equal expected future short rates. Given these assumptions we rewrite equation (2.3) as:

$$
\begin{aligned}
P_{t,n} &= D_t \left[E_t exp \left(\sum_{i=1}^{n} g_{t+i} - y_t - \theta_{t+i} \right) \right] \\
&= D_t \left[E_t exp \left(\sum_{i=1}^{n} -y_{t+i} \right) \right] \left[E_t exp \left(\sum_{i=1}^{n} z_{t+i} \right) \right] \\
&= D_t exp(-n y_{t,n}) \left[E_t exp \left(\sum_{i=1}^{n} z_{t+i} \right) \right].
\end{aligned}
\tag{2.34}
$$

and the derivatives price equals:

$$
F_{t,n} = D_t \left[E_t exp \left(\sum_{i=1}^{n} z_{t+i} \right) \right]. \tag{2.35}
$$

This shows that, to fit the futures price data, only a model for z_{t+1} is needed. To reconcile this model with the stock index level, the independence assumption and expectations hypothesis for bonds are necessary and equation (2.29) can be used to calculate present values of dividends and the stock index value. Using this pricing equation, one can again specify a two-state model, in this case for one period growth z_{t+1}, and estimate it using the Kalman filter in the same way as described for the base model.

As mentioned, this model assumes independence of interest rates and risk-adjusted growth rates. In the real world, however, correlation between the risk-free rate, dividend growth and the dividend risk premium is expected since often the same drivers apply: economic growth, the investment cycle, slack in the labor market and other economic variables will affect all of them. For estimating the term structure model, such correlation is not a problem if z_{t+1} is the subject of state space estimation instead of π_{t+1}, but it will cause misestimation of the implied stock market levels. It is easy to show that this separation of the two correlated variables would produce overestimation of the stock index in equation (2.26) if the actual correlation is positive.

Turning to the results, the long-term estimate for risk-adjusted growth $\overline{z}$ is estimated rather high, at 0.3 percent for the Eurostoxx 50 and – 1.0 percent for the Nikkei 225, which translates to 1.9 percent and 0.1 percent once $\overline{z}$ is corrected for the convexity term (Table 2.5). Standard errors are larger than in the case of the base model. The mean reversion parameters obtained remain reasonable and significant. However, reconciling the dividend market to the stock market based on these estimates overstates the stock market by a large margin and reduces the fit of the dynamic return reconciliation (Table 2.7). Compared to the base model, the coefficient of the estimated price-dividend ratio maintains its presence in the Japanese data with a coefficient of 0.78. In the data period considered, yen interest rates were close to zero and showed less variation than in the European market. The reduction to the explanatory power when modeling growth without discounting is relatively small for the Nikkei 225. In Europe, the picture is quite different due to the steep drop in interest rates in the period of 2008 to 2015. The coefficient of the estimated price-dividend ratio is almost negligibly small for the Eurostoxx 50 market. Therefore, the European data set in particular demonstrates the correlation among the three elements of π_{t+1}, which confirms the advantage of estimating risk-adjusted dividend growth *after* discounting at the risk-free rate.

2.6.4 OTC data

We retain price data of dividend swaps from several investment banks[32] for the dividend futures markets under investigation, and also for the S&P 500 and the FTSE 100. These data extend back to December 2005. Over-the-Counter (OTC) prices for dividend derivatives are not readily observable as are, for example, interest rate swaps, money market derivatives or foreign exchange derivatives which are posted on information systems such as Reuters. Mixon and Onur (2017) provide insight into the OTC market for dividend swaps. They investigated data from a Swap Data Repository to which participants in swap markets must report at transaction-level. It is shown that OTC swaps trade infrequently; even for the S&P 500, which is the largest OTC dividend market, they trade less than daily between dealers and only once every few weeks between a dealer and a non-dealer end-user.

Investment banks update their pricing sheets on a daily basis, but often prices remain stale and extended periods go by without a single trade taking place. The data set of OTC prices for dividend swaps, therefore, is impacted by the model investment banks use for pricing them. We find price differences for same maturity transactions among the pricing sheets of investment banks of on average 3% with a standard deviation of 3%.

Since the OTC market does not trade regularly, it seems likely that fitting the state space model to its price data is akin to mimicking the pricing models used by the investment banks. We nonetheless perform the same set of estimations and reconciliations as above on the OTC price data of dividend swaps referring to the S&P 500 and the FTSE 100 indices. The results shown are restricted to monthly frequencies, as the daily data are stale. The results are shown in Tables 2.6 and 2.8.

The two-state model produces a high estimate for the long run growth constant $\overline{p}^*$ of S&P 500 dividends. Indeed, at – 1.3 percent for this constant, the S&P 500 present value as estimated by the model (equation (2.26)) overestimates its observed values by a factor of

[32] Deutsche Bank, Goldman Sachs and Credit Suisse.

more than 2. Both mean reversion parameters attain reasonable levels, but they attract fairly large standard errors. In the case of the FTSE 100, the two-state model estimate for long run growth equals – 5.3 percent, with a standard error even exceeding that level in absolute terms. At the same time, the second mean reversion parameter comes out low at 0.04, which translates into a half-value time running into decades. At such slow moving mean reversion, the role of the long run constant is essentially taken over by the medium-term factor. The single-state estimate for long run growth is more reasonable at – 3.3 percent.

We then turn to the reconciliation regressions (equation (2.31)). The variation in the modeled price-dividend ratio produced by the estimates does not depend on the long run constant and the dynamic reconciliation to the stock indices demonstrates that it has meaningful explanatory power. Table 2.8 shows that the model produces a coefficient of 0.1 to 0.2 for the price-dividend ratio, with reasonable significance for both the S&P 500 and the FTSE 100. Overall explanatory power is high for the S&P 500 with the adjusted R^2 reaching 0.58. However, most of it stems from the observed first dividend price F_t rather than the modeled price-dividend ratio.[33] The same applies to the FTSE 100, albeit with the adjusted R^2 at a lower level of explanatory power. In both markets variation in the price-dividend ratio accounts for 8%, against 28% and 43% for the Eurostoxx 50 and Nikkei 225.

2.6.5 Comparison to structural Macro models

Our model is not a structural model that separates expected dividends from risk premiums. But although it does not provide means of dividend growth or expected dividend strip returns, we can still calculate the model-implied volatility of returns of dividends across maturities. In a similar way as Binsbergen and Koijen (2017), we compare the volatility term structure of dividend returns for the S&P 500 to those implied by theoretical asset pricing models. Binsbergen et al. (2012) report in their Figure 5 the volatility term structures for three important models: the habit formation model of Campbell and Cochrane (1999), the Bansal-Yaron (2004) long-run risk model, and the rare disasters model of Gabaix (2009). Figure 2.10 graphs the annualized volatility of log dividend price changes implied by our model estimates, for maturities of one to ten years[34]. The figure also plots the volatility curve for the habit-formation model of Campbell and Cochrane (1999) and long-run risk model of Bansal and Yaron (2004), as reported in Figure 5 of Binsbergen et al. (2012).[35] It shows that the volatility curve calibrated using our model is increasing and concave, similar

[33] Similar regressions based on daily estimates and stock index data produce R^2 of less than 5 percent.

[34] We do not estimate the volatility of current dividend changes. We set this value equal to 11.2% in Figure 10, which is the value used by Campbell and Cochrane (1999).

[35] We thank Ralph Koijen and Jules van Binsbergen for providing us with these data.

to the results for the Eurostoxx and Nikkei data. In sharp contrast, the volatility curve of the habit formation model is almost flat in this range and out of line with the data. The long run risk curve approximates the calibrated volatility curve better than the habit formation model, but is somewhat less steep than our calibrated volatility curve. The volatility of the long maturity futures implied by the long run risk model (17 to 18%) is also lower than the volatility fitted by our model, and lower than typical values for the stock market volatility. Notice that Bansal and Yaron (2004) also report that their long-run risk model produces a volatility of the price-dividend ratio that is lower than the value found in the data.[36]

2.7 Conclusion

This paper proposes a method to extract information about the expectations that investors entertain of stock dividends from dividend derivatives. We show that modeling a single variable is sufficient to describe the dynamics and term of structure of dividend values. This variable is equal to the dividend growth minus the risk-free rate and a term capturing the risk premium. We propose a two-factor model for this discounted risk-adjusted growth variable, capturing the dynamics of short-term and medium-term dividend growth. The two factors shape a term structure of dividend growth which fits the data well and they determine the dynamics of the price-dividend ratio. Applied to the Eurostoxx 50 and the Nikkei 225, most of the variation of the stock market can be traced back to the model and short-term dividends together. We conclude that dividend derivatives and stock prices line up well enough to consider the information contained in one market for use of understanding the other. Several inferences from these findings can be drawn.

The distance into the future considered by investors affects the fit of dividend derivatives. At the extreme, a model which assumes a constant discount rate would show poor fit and explanatory power. But even a model where short-term variation in growth expectations is described by a single factor is significantly outperformed by a two-state model. The short-term factor reflects a horizon of under one year and the medium-term factor a horizon of several years. Deploying two states next to each other allows some distinction between sudden occurrences and those at business cycle proportions. Pursuing different explanations for the two states, or in other words, finding different determinants of how investors think of the short and the medium-term, seems an appropriate research avenue.

The state space model imposes the return to a constant mean level of growth in the long run. This assumption is loosely interpreted as that investors do not change their opinion

[36]Bansal and Yaron (2004, Table IV) report a price-dividend ratio volatility of 29% in the data and 18% in their benchmark model.

about the sequence of present values of dividends in the long run, which seems very restrictive intuitively. Nevertheless, small estimation errors, the explanatory power of the reconciliation of the model to stock returns and the near unity of the coefficients of the short-term dividend and the price-dividend ratio in the regressions of stock returns on these determinants, add credence to the imposition that long run growth is no major source of stock market variation. Interest rates are a part of discounted risk-adjusted dividend growth, and they are observable to investors. Under the assumption that they do not change their opinion about discounted risk-adjusted dividend growth $\overline{p}$ in the long run, then our results suggest that most interest rate variation is balanced by risk-adjusted dividend growth expectations $g - \theta$ at these long horizons.

The estimation of the dividend term structure improves when we directly discount risk-adjusted dividend growth for the time value of money. An alternative approach, which does not discount dividends at the risk-free rate and assumes independence between interest rates and risk-adjusted growth, implies estimates of dividend growth that are not economically sensible and which reconcile poorly to the stock market. Hence, jointly modeling interest rates, dividend growth and risk premium is preferred.

We perform the estimations using prices of OTC dividend swaps as well as of listed dividend futures. The prices produced by the OTC market are relevant, but generate less precise results and much lower explanatory power, while dividend futures provide intuitive and highly significant results. Not only do the long run estimates come out poorly, also the added value of the two-state model is not confirmed by OTC prices. It seems that stale prices, large price discrepancies among investment banks and infrequent trading cautions their interpretation when applied in present value analysis. Fortunately, the set of listed data will only expand as time passes.

Robert Shiller contends that realized dividends are not volatile enough to justify the observed volatility of stock markets, if discount rates are assumed to be constant over time and maturities (Shiller, 1981). The approach we take constructs a rationally expected price for stocks in a different way. Rather than a model of future realized dividends, the term structure contains both actual expectations of future dividends and risk-adjusted discount rates. While Shiller finds observed stock return volatility to be five to thirteen times larger than modelled volatility, we find that the model produces about as much stock market volatility as is observed, regardless of the stock market that we consider.

Any model limited to using realized dividends as a proxy for expectations of dividends will fail to pick up the variation in discount rates that the market applies to those future dividends, as well as estimation error of those expectations which may well display substantial volatility of their own. The approach taken by Shiller confirms that dividends turn out much less volatile than the stock market. But it does not confirm that the volatility of the present value of dividends is too high, only because the drivers of such valuations aren't observed.

It would be interesting to study to which fundamental variables the variation of short-term and medium-term growth in discounted risk-adjusted dividends can be ascribed. Armed with such linkages, the ability to understand stock market dynamics will improve. At the same time, Cochrane (2011) is clear in his assertion that: "We do not have to *explain* discount rates – relate expected returns to betas and understand their deep economics – in order to *use* them.". Opportunities are plentiful.

2.8 Appendix A: Dividend derivatives data

2.8.1 The first-to-expire constant maturity derivative

The weighting scheme in equation (2.36) is applied to obtain all constant maturity (CM) derivatives prices, except for the first CM derivative, because the proposed approach carries measurement problems. At time t the expected dividend to be delivered at the expiration of the first derivative $E_t\left(D_1\right)$ is the sum of the dividend index DI_t as it accretes throughout the year and its unknown complement $E_t\left(UD_1\right)$

$$E_t\left(D_1\right) = DI_t + \ E_t\left(UD_1\right). \tag{2.36}$$

For CM derivatives with horizons longer than the first, the weight w_i in (2.17) is the average seasonal pattern in the preceding decade, which may not necessarily resemble that of a particular dividend index year $DI_t/E_t\left(D_1\right)$. The difference between the two is shown in Figure A2; for example in April 2013 the payments of Eurostoxx 50 dividends had already reached 33% of the annual total, while on average in the years 2005 to 2013 it stood at 20%. This advance dropped below ten percent not until a month later. In general, dividend payments in 2012 and 2013 seem to have taken place earlier in the calendar year than usual in the preceding years. Weighting the first derivative by the average of the preceding decade when dividends realize sooner in the year than the average, as was the case in April 2013, overemphasizes the importance of that first derivative to the one year CM derivative. This first CM derivative will then contain backward looking information as well underemphasize the unrealized proportion of the contemporaneous dividend index both to the tune of the difference between the historical average and the realized dividend index. To avoid this issue, the first CM derivative is construed by defining the weight as the proportion of the dividend index that has been realized of the total expected dividend for that year only:

$$F_{t,1}^{CM} = F_{t,1} - DI_t + \frac{DI_t}{F_{t,1}} F_{t,2}. \tag{2.37}$$

For building a first CM derivative with a constant one year horizon as a stochastic variable, we include unknown $E_t(UD_1)$ and exclude known DI_t. The expectation of full year dividends is proxied by the equivalent observation. Later CM derivatives do not weight variables which have already been partly realized, hence the weighting issue of the first CM derivative does not reoccur. For $n \geq 2$, the prices of CM derivatives remain constructed as in the weighting equation (2.36).

2.8.2 Calculating seasonal weights for different dividend index years

Expiry years do not have the same number of trading days every year or across markets. Not only do trading holidays differ, also the expiry date is set to the third Friday in December in every expiry year. This day falls anywhere between the 15^{th} and the 21^{st} of December[37] and the number of trading days fluctuates accordingly.

In order to establish a seasonal pattern for w_i that is correct for the actual number of trading days in each expiry year, realized dividends are normalized and averaged. First, the amount of dividends paid on a given day is expressed as a percentage of the total dividends paid in the matching dividend index year. Next, for each expiry year these percentages are normalized to a set number of trading days. Finally, they are averaged. For calculating the values in the weighting equation, they are rescaled to the actual number of trading days in the dividend index year in question. This approach guarantees that in every expiry year, weight w_i starts at zero and ends the year at 100%, regardless of the number of trading days.

2.9 Appendix B: Measurement equations

In this appendix the details of the derivation of the measurement equations are described. We rewrite the state equations (2.10) and (2.11) in vector form and then derive the discrete-time implications of the model. Denote $Q_t = \begin{pmatrix} p_t \\ \tilde{p}_t \end{pmatrix}$ the 2×1 vector of the factors and $\overline{Q} = \begin{pmatrix} \overline{p} \\ \overline{p} \end{pmatrix}$ as the 2×1 vector of the constant infinite growth rate. In a two equation matrix format, the system becomes:

$$dQ_t = \begin{pmatrix} dp_t \\ d\tilde{p}_t \end{pmatrix} = \left[\begin{pmatrix} -\varphi & \varphi \\ 0 & -\psi \end{pmatrix} \begin{pmatrix} p_t \\ \tilde{p}_t \end{pmatrix} + \begin{pmatrix} 0 \\ \psi\overline{p} \end{pmatrix} \right] dt + \begin{bmatrix} \sigma_p & 0 \\ 0 & \sigma_{\tilde{p}} \end{bmatrix} \begin{pmatrix} dW_p \\ dW_{\tilde{p}} \end{pmatrix}. \quad (2.38)$$

This system of differential equations in matrix notation is:

$$dQ_t = C \left[Q_t - \overline{Q} \right] dt + \varSigma dW, \quad (2.39)$$

which has the general solution:

$$Q_{t+1} = \overline{Q} + \varPhi \left(Q_t - \overline{Q} \right) + \varepsilon_{t+1}, \quad (2.40)$$

[37] With exception of the Nikkei 225.

and of which the eigenmatrix solves to:

$$\Phi = \begin{pmatrix} e^{-\varphi} & \frac{\varphi}{\varphi-\psi}(e^{-\varphi} - e^{-\psi}) \\ 0 & e^{-\psi} \end{pmatrix}. \tag{2.41}$$

Substituting the expression for the eigenmatrix into equation (2.40) delivers state equations:

$$\begin{aligned} \begin{pmatrix} p_{t+1} \\ \tilde{p}_{t+1} \end{pmatrix} = \begin{pmatrix} 1 - e^{-\varphi} & -\frac{\varphi}{\varphi-\psi}\left(e^{-\psi} - e^{-\varphi}\right) \\ 0 & 1 - e^{-\psi} \end{pmatrix} \begin{pmatrix} \overline{p} \\ \overline{p} \end{pmatrix} + \\ \begin{pmatrix} e^{-\varphi} & \frac{\varphi}{\varphi-\psi}\left(e^{-\psi} - e^{-\varphi}\right) \\ 0 & e^{-\psi} \end{pmatrix} \begin{pmatrix} p_t \\ \tilde{p}_t \end{pmatrix} + \varepsilon_{t+1}. \end{aligned} \tag{2.42}$$

We model correlation between the innovation in the growth rate ν_{t+1} and the errors ε_{t+1} in these state equations as $\nu_{t+1} = \beta'\epsilon_{t+1}$, where $\beta = (\beta_p, \beta_{\tilde{p}})'$ is a 2-by-1 vector. Next we use this process to write the n-period ahead growth rate as a function of the factors:

$$\pi_{t+n} = \alpha'\left(\overline{Q} + \Phi^{n-1}\left(Q_t - \overline{Q}\right)\right) + \alpha'\sum_{i=1}^{n-1}\Phi^{n-i}\varepsilon_{t+i} + \beta'\varepsilon_{t+n}, \tag{2.43}$$

in which $\alpha^{'} = (1\ 0)$. This can be substituted into the pricing equation:

$$lnP_{t,n} - lnD_t \ = E_t\left(\sum_{i=1}^{n}\pi_{t+i}\right) + \frac{1}{2}Var_t\left(\sum_{i=1}^{n}\pi_{t+i}\right), \tag{2.44}$$

The right hand side can be worked out as follows:

$$lnP_{t,n} - lnD_t \ = \alpha'\left(n\overline{Q} + B_n\left(Q_t - \overline{p}\right)\right) + \frac{1}{2}Var_t\left(\sum_{i=1}^{n}\left(\alpha'\sum_{j=1}^{i-1}\Phi^{n-j}\varepsilon_{t+j} + \beta'\varepsilon_{t+i}\right)\right), \tag{2.45}$$

which in turn implies:

$$lnP_{t,n} - lnD_t = \alpha'\left(n\overline{Q} + B_n\left(Q_t - \overline{p}\right)\right) + \frac{1}{2}\sum_{i=1}^{n}(\beta^{'} + \alpha^{'}B_i)\Sigma(\beta + B_i^{'}\alpha), \tag{2.46}$$

where matrix B_i is an expression constructed from the eigenmatrix:

$$B_i = (I + \Phi + \cdots + \Phi^{i-1}) = (I - \Phi)^{-1}\left(I - \Phi^i\right). \tag{2.47}$$

The equations are written without vector notation. By the definition of Φ, B_n is worked out as:

$$B_n = \begin{pmatrix} \frac{\left(1-e^{-n\varphi}\right)}{\left(1-e^{-\varphi}\right)} & \frac{\varphi}{\varphi-\psi}\left(\frac{\left(1-e^{-n\psi}\right)}{\left(1-e^{-\psi}\right)} - \frac{\left(1-e^{-n\phi}\right)}{\left(1-e^{-\phi}\right)}\right) \\ 0 & \frac{\left(1-e^{-n\psi}\right)}{\left(1-e^{-\psi}\right)} \end{pmatrix} = \begin{pmatrix} \varphi_n & \frac{\varphi}{\varphi-\psi}\left(\psi_n - \varphi_n\right) \\ 0 & \psi_n \end{pmatrix}, \tag{2.48}$$

with shorthand notation:

$$\phi_n = \frac{\left(1-e^{-n\phi}\right)}{\left(1-e^{-\phi}\right)} \tag{2.49}$$

$$\psi_n = \frac{\left(1-e^{-n\psi}\right)}{\left(1-e^{-\psi}\right)}. \tag{2.50}$$

An expression which consists of scalars only is obtained by substituting all elements of the above in the measurement equation:

$$\begin{aligned} lnP_{t,n} - lnD_t = &(1 \;\; 0) \left(n \begin{pmatrix} \overline{p} \\ \overline{p} \end{pmatrix} + \begin{pmatrix} \varphi_n & \frac{\varphi}{\varphi-\psi}\left(\psi_n - \varphi_n\right) \\ 0 & \psi_n \end{pmatrix} \begin{pmatrix} p_t - \overline{p} \\ \tilde{p}_t - \overline{p} \end{pmatrix} \right) + \\ &\frac{1}{2}\sum_{i=1}^{n} \left((\beta_p \; \beta_{\tilde{p}}) + (1 \;\; 0) \begin{pmatrix} \varphi_i & \frac{\varphi}{\varphi-\psi}\left(\psi_i - \varphi_i\right) \\ 0 & \psi_i \end{pmatrix} \right) \begin{pmatrix} \sigma^2 & 0 \\ 0 & \widetilde{\sigma}^2 \end{pmatrix} \times \\ &\left(\begin{pmatrix} \beta_p \\ \beta_{\tilde{p}} \end{pmatrix} + \begin{pmatrix} \varphi_i & 0 \\ \frac{\varphi}{\varphi-\psi}\left(\psi_i - \varphi_i\right) & \psi_i \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) + \eta_{t,n} \end{aligned} \tag{2.51}$$

$$\begin{aligned} = &n\overline{p} + \varphi_n\left(p_t - \overline{p}\right) + \frac{\varphi}{\varphi-\psi}\left(\psi_n - \varphi_n\right)\left(\tilde{p}_t - \overline{p}\right) + \\ &\frac{1}{2}\sum_{i=1}^{n}\left(\sigma_p^2(\beta_p + \varphi_i)^2 + \sigma_{\tilde{p}}^2 \left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi-\psi}\left(\psi_i - \varphi_i\right) \right)^2 \right) + \eta_{t,n}, \end{aligned}$$

which is the same as equation (2.13) in the main text. The right hand term on the right hand side is referred to in the paper as the ”convexity term”. Dividend return variance follows from equation (2.46). Conditional variance is reduced to:

$$\sigma_t^2\left(lnP_{t+1,n} - lnP_{t,n}\right) = \sigma_t^2\left(lnD_{t+1}\right) + \frac{1}{2}\alpha^{'} B_{n-1} \Sigma B_{n-1}^{'} \alpha. \tag{2.52}$$

Substituting for the variables in the two state model yields equation (2.21), from which volatilities in Figures 2.6 and 2.10 are shown by taking square roots.

2.10 Tables

Table 2.1: **Base model of Discounted Risk-adjusted Dividend Growth**: $\pi_t = g_{t+1} - y_t - \theta_{t+1}$

Estimates using listed Dividend Futures of the **Eurostoxx 50 Index**.
Sample period: 4 August 2008 – 16 February 2015.

Maximum Likelihood estimates are based on daily prices of dividend futures and interest rates. Measurement equations capture discounted dividend growth starting one year following the observation date. The estimates include eight measurement equations: from one to eight years, except for its begin until 13^{th} May 2009 in which the number is five due to a lack of data. σ_η measures the standard deviations of the second until the eighth measurement equations, σ_ε of the first. This distinction is made to reflect that the base from which growth rates are determined is calculated by applying an alternative weighting scheme between first and second derivatives to expire. See the Data section 3. Standard errors in parentheses.

	Two state		Single state	
	$dp_t = \varphi\left(\tilde{p}_t - p_t\right)dt + \sigma_p dW_p$ $d\tilde{p}_t = \psi\left(\overline{p} - \tilde{p}_t\right)dt + \sigma_{\tilde{p}} dW_{\tilde{p}}$		$dp_t = \varphi\left(\overline{p} - p_t\right)dt + \sigma_p dW_p$	
$\overline{p}$	-0.0586	-0.0404	-0.2067	-0.0435
	(9.5339)	(0.0197)	(28.7770)	(0.0144)
φ	1.5130	1.5132	1.7297	1.7292
	(0.3160)	(0.3158)	(0.4894)	(0.4894)
ψ	0.2433	0.2434		
	(0.1089)	(0.1088)		
β_p	0.1553	*Set to 0*	0.6246	*Set to 0*
	(67.007)		(69.2935)	
$\beta_{\tilde{p}}$	-2.6695	-2.6693		
	(6.2539)	(6.2523)		
σ_p	0.5701	0.5704	0.7033	0.7033
	(0.7876)	(0.7870)	(1.2245)	(1.2245)
$\sigma_{\tilde{p}}$	0.0437	0.0437		
	(0.0947)	(0.0946)		
σ_ε	0.0219	0.0219	0.0177	0.0177
	(0.0295)	(0.0294)	(0.0071)	(0.0071)
σ_η	0.0063	0.0063	0.0441	0.0441
	(0.0025)	(0.0025)	(0.0806)	(0.0806)
Log Likelihood per contribution	24.57	24.57	18.35	18.35

Table 2.2: **Base model of Discounted Risk-adjusted Dividend Growth**: $\pi_t = g_{t+1} - y_t - \theta_{t+1}$

Estimates using listed Dividend Futures of the **Nikkei 225 Index**.
Sample period: 17 June 2010 – 16 February 2015.

Maximum Likelihood estimates are based on daily prices of dividend futures and interest rates. Measurement equations capture discounted dividend growth starting one year following the observation date. The estimates include eight measurement equations: from one to eight years. σ_η measures the standard deviations of the second until the eighth measurement equations, σ_ε of the first. This distinction is made to reflect that the base from which growth rates are determined is calculated by applying an alternative weighting scheme between first and second derivatives to expire. See the Data section 3. Standard errors in parentheses.

	Two state		Single state	
	$dp_t = \varphi(\tilde{p}_t - p_t)\,dt + \sigma_p dW_p$ $d\tilde{p}_t = \psi(\overline{p} - \tilde{p}_t)\,dt + \sigma_{\tilde{p}} dW_{\tilde{p}}$		$dp_t = \varphi(\overline{p} - p_t)\,dt + \sigma_p dW_p$	
$\overline{p}$	-0.0320	-0.0371	-0.0719	-0.0487
	(1.3833)	(0.0264)	(3.8264)	(0.0304)
φ	0.7381	0.7381	0.2837	0.2837
	(0.2360)	(0.2345)	(0.0306)	(0.0306)
ψ	0.1784	0.1784		
	(0.0539)	(0.0537)		
β_p	-1.5513	*Set to 0*	-7.6306	*Set to 0*
	(92.822)		(211.137)	
$\beta_{\tilde{p}}$	-3.4234	-3.4229		
	(16.191)	(16.1862)		
σ_p	0.1531	0.1531	0.0630	0.0630
	(0.2193)	(0.2193)	(0.1189)	(0.1186)
$\sigma_{\tilde{p}}$	0.0251	0.0251		
	(0.0731)	(0.0730)		
σ_ε	0.0147	0.0147	0.0137	0.0137
	(0.0197)	(0.0197)	(0.0041)	(0.0041)
σ_η	0.0040	0.0040	0.0170	0.0170
	(0.0015)	(0.0015)	(0.0285)	(0.0285)
Log Likelihood per contribution	29.37	29.36	22.04	22.04

Table 2.3: **Key results from the state space model with two factors and sample period averages.**

		Eurostoxx 50	**Nikkei 225**
Sample period		4 August 2008 – 16 Feb 2015	17 June 2010 – 16 Feb 2015
Estimated LT growth (beta = 0)	$\overline{p}$	–4.0%	–3.7%
Estimated convexity term for $i \to \infty$[39]		1.4%	1.1%
Estimated LT growth plus convexity term[39]	$\overline{p}^*$	–2.6%	–2.6%
Average dividend yield	$\frac{D_t}{S_t}$	4.3%	1.9%
Average estimated 1 year growth	p_t	–9.8%	1.1%
Average estimated 1 year forward 4 year growth	$\tilde{p}_t$	–3.7 %	–1.6 %
Average calibrated first dividend point	$\frac{\hat{P}_{t,1}}{\sum_1^\infty \hat{P}_{t,n}}$	3.1%	2.1%

[39] Refer to equation (2.25) in the main text:

$$\overline{p}^* = \overline{p} + \frac{1}{2}\left(\sigma_p^2(\beta_p + \varphi_{i\to\infty})^2 + \sigma_{\tilde{p}}^2\left(\beta_{\tilde{p}} + \frac{\varphi}{\varphi - \psi}\left(\psi_{i\to\infty} - \varphi_{i\to\infty}\right)\right)^2\right)$$

Table 2.4: Reconciliation of the Base Present Value Model (two state) constituent returns to stock market returns: **listed Dividend Futures.**

The modeled present values of dividends are tested for their explanatory power of the dynamics of the stock market. The OLS regression estimates equation (2.31) $s_t = \alpha + \beta_f f_t + \beta_{\Delta y} \Delta y_t + \beta_{\widehat{pd}} \widehat{pd}_t + \varepsilon_t$, in which s_t is stock index log returns, f_t is the log return of the first constant maturity dividend derivative, Δy_t is the change in the one year zero-coupon swap rate and $\widehat{pd}_t$ is the first differenced log of the sum of the normalized present value of dividends as estimated in the two state space model. β is fixed at zero. Daily data for periods as in Tables 1 and 2. Standard errors in parentheses.

	Eurostoxx 50				**Nikkei 225**			
Constant	0.0005	0.0002	0.0006	-0.0001	-0.0002	0.0002	0.0005	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)
f_t	0.8978	1.0009			0.8488	0.6582		
	(0.0337)	(0.0426)			(0.0508)	(0.0719)		
Δy_t	0.1446		0.2022		0.0751		-0.0081	
	(0.0127)		(0.0178)		(0.0619)		(0.0912)	
$\widehat{pd}_t$	0.6587			0.6893	0.8619			0.8156
	(0.0216)			(0.027)	(0.0251)			(0.0278)
Adj. R^2	0.540	0.248	0.071	0.280	0.540	0.068	0.000	0.429

Table 2.5: **Alternative Model** of Undiscounted Risk-adjusted Dividend Growth: $z_t = g_{t+1} - \theta_{t+1}$

Estimates using **listed Dividend Futures.**

Maximum Likelihood estimates are based on daily prices of dividend futures and interest rates. Measurement equations capture discounted dividend growth starting one year following the observation date. The estimates include eight measurement equations: from one to eight years, except for its begin until 13^{th} May 2009 in which the number is five due to a lack of data. σ_η measures the standard deviations of the second until the eighth measurement equations, σ_ε of the first. This distinction is made to reflect that the base from which growth rates are determined is calculated by applying an alternative weighting scheme between first and second derivatives to expire. See the Data section 3. Standard errors in parentheses.

	Two state		Single state	
	$dz_t = \varphi(\tilde{z}_t - z_t)\,dt + \sigma_z dW_p$ $d\tilde{z}_t = \psi(\overline{z} - \tilde{z}_t)\,dt + \sigma_{\tilde{z}} dW_{\tilde{z}}$		$dz_t = \varphi(\overline{z} - z_t)\,dt + \sigma_z dW_z$	
	Eurostoxx 50	**Nikkei 225**	**Eurostoxx 50**	**Nikkei 225**
Sample period	4 August 2008 – 16 Feb 2015	17 June 2010 – 16 Feb 2015	4 August 2008 – 16 Feb 2015	17 June 2010 – 16 Feb 2015
$\overline{z}$	0.0027	-0.0104	-0.0147	-0.0406
	(0.0241)	(0.0416)	(0.0212)	(0.0549)
φ	1.4108	0.6547	1.4142	0.2944
	(0.2552)	(0.1874)	(0.4588)	(0.0350)
ψ	0.1860	0.2006		
	(0.0779)	(0.0591)		
β_z	*Set to 0*	*Set to 0*	*Set to 0*	*Set to 0*
$\beta_{\tilde{z}}$	-3.0428	-3.6131		
	(6.8937)	(35.6244)		
σ_z	0.5125	0.1335	0.4858	0.0794
	(0.6433)	(0.1814)	(1.0167)	(0.1830)
$\sigma_{\tilde{z}}$	0.0362	0.0286		
	(0.0687)	(0.0810)		
σ_ε	0.0233	0.0138	0.0259	0.0144
	(0.0311)	(0.0183)	(0.0136)	(0.0042)
σ_η	0.0060	0.0038	0.0569	0.0183
	(0.0023)	(0.0014)	(0.1331)	(0.0336)
Log Likelihood per contribution	24.78	29.64	15.75	21.58

Table 2.6: **Base model** of Discounted Risk-Adjusted Dividend Growth:$\pi_t = g_{t+1} - y_t - \theta_{t+1}$

Estimates using **OTC Dividend Swaps** (monthly).

Maximum Likelihood estimates are based on daily prices of dividend futures and interest rates. Measurement equations capture discounted dividend growth starting one year following the observation date. The estimates include eight measurement equations: from one to eight years. σ_η measures the standard deviations of the second until the eighth measurement equations, σ_ε of the first. This distinction is made to reflect that the base from which growth rates are determined is calculated by applying an alternative weighting scheme between first and second derivatives to expire. See the Data section 3. Standard errors in parentheses.

	Two state		Single state	
	$dp_t = \varphi\left(\tilde{p}_t - p_t\right)dt + \sigma_p dW_p$ $d\tilde{p}_t = \psi\left(\overline{p} - \tilde{p}_t\right)dt + \sigma_{\tilde{p}} dW_{\tilde{p}}$		$dp_t = \varphi\left(\overline{p} - p_t\right)dt + \sigma_p dW_p$	
	S&P 500	**FTSE 100**	**S&P 500**	**FTSE 100**
Sample period	Dec 2005 – June 2014	Dec 2005 – June 2014	Dec 2005 – June 2014	Dec 2005 June 2014
$\overline{p}$	-0.0188	-0.0841	-0.0186	-0.0430
	(0.0231)	(0.1513)	(0.0108)	(0.0093)
φ	1.0651	1.6347	0.3537	1.7702
	(0.7296)	(0.5865)	(0.0583)	(0.5828)
ψ	0.1809	0.0371		
	(0.1431)	(0.1422)		
β_p	*Set to 0*	*Set to 0*	*Set to 0*	*Set to 0*
$\beta_{\tilde{p}}$	-2.5935	-2.1624		
	(10.4524)	(8.5798)		
σ_p	0.1756	0.5865	0.0584	0.6770
	(0.2975)	(0.9707)	(0.0674)	(1.1518)
$\sigma_{\tilde{p}}$	0.0293			
	(0.0642)			
σ_ε	0.0167	0.0199	0.0138	0.0141
	(0.0072)	(0.0326)	(0.0044)	(0.0056)
σ_η	0.0078	0.0054	0.0261	0.0298
	(0.0021)	(0.0026)	(0.0167)	(0.0687)
Log Likelihood per contribution	22.62	23.79	19.79	19.29

Table 2.7: Reconciliation of the **Alternative Model** of Undiscounted Risk-Adjusted Dividend Growth (two state) constituent returns to stock market returns.

The modeled present values of dividends are tested for their explanatory power of the dynamics of the stock market. The OLS regression estimates equation (2.31) $s_t = \alpha + \beta_f f_t + \beta_{\Delta y}\Delta y_t + \beta_{\widehat{pd}}\widehat{pd}_t + \varepsilon_t$, in which s_t is stock index log returns, f_t is the log return of the first constant maturity dividend derivative, Δy_t is the change in the one year zero swap rate and $\widehat{pd}_t$ is the first differenced log of the sum of the normalized present value of dividends as estimated in the two state space model. β is fixed at zero. Daily data for periods as in Tables 1 and 2. Standard errors in parentheses.

	Eurostoxx 50				**Nikkei 225**			
Constant	0.0006	0.0002	0.0006	0.0000	-0.0002	0.0002	0.0005	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0003)
f_t	0.9555	1.0009			0.8781	0.6582		
	(0.0419)	(0.0426)			(0.0572)	(0.0719)		
Δy_t	0.1693		0.2022		0.1448		-0.0081	
	(0.0163)		(0.0178)		(0.0694)		(0.0912)	
$\widehat{pd}_t$	0.0725			-0.0113	0.7800			0.7121
	(0.0157)			(0.0179)	(0.0293)			(0.0318)
Adj. R^2	0.293	0.248	0.071	0.000	0.424	0.068	0.000	0.305

Table 2.8: Reconciliation of the Base Present Value Model (two state) constituent returns to stock market returns: **OTC Dividend Swaps** (monthly data).

The modeled present values of dividends are tested for their explanatory power of the dynamics of the stock market. The OLS regression estimates equation (2.31) $s_t = \alpha + \beta_f f_t + \beta_{\Delta y} \Delta y_t + \beta_{\widehat{pd}} \widehat{pd}_t + \varepsilon_t$, in which s_t is stock index log returns, f_t is the log return of the first constant maturity dividend derivative, Δy_t is the change in the one year zero swap rate and $\widehat{pd}_t$ is the first differenced log of the sum of the normalized present value of dividends as estimated in the two state space model. β is fixed at zero. Monthly data for periods as in Table 2.6. Standard errors in parentheses.

	S&P 500				**FTSE 100**			
Constant	-0.0019	-0.0016	0.0065	0.0038	0.0005	0.0005	0.0032	0.0016
	(0.0029)	(0.0028)	(0.0042)	(0.0040)	(0.0034)	(0.0035)	(0.0040)	(0.0039)
f_t	1.1677	1.1901			0.5812	0.6273		
	(0.1118)	(0.1073)			(0.1111)	(0.1033)		
Δy_t	-0.0041		0.0514		0.0047		0.0320	
	(0.0158)		(0.0221)		(0.0163)		(0.0173)	
$\widehat{pd}_t$	0.1869			0.2561	0.1190			0.1428
	(0.0600)			(0.0881)	(0.0410)			(0.0459)
Adj. R^2	0.582	0.552	0.051	0.078	0.307	0.269	0.033	0.088

Table A.1. Summary static data of the dividend derivatives per underlying stock indices.

		Eurostoxx 50	**S&P 500**	**FTSE 100**	**Nikkei 225**
Number of companies in the index		50	500	100	225
Currency		Euro	US$	GBP	JPY
Market capitalization US$ per 7^{th} May 2014		US$ 3.3 trillion	US$ 17.2 trillion	US$ 3.1 trillion	US$ 2.7 trillion
Data period	Dividend swaps	N/A	19 December 2005 – 13 June 2014	19 December 2005 – 13 June 2014	N/A
	Dividend futures	4 August 2008 – 16 February 2015	N/A	N/A	17 June 2010 – 16 February 2015
Source of the data	Dividend swaps	N/A	OTC	OTC	N/A
	Dividend futures	Eurex	N/A	N/A	Singapore exchange
Average number of trading days		256	252	253	245
Liquidity[40]		Good	Low	Low	Reasonable
Expiry horizon	Dividend swaps	N/A	10 years	10 years	N/A
	Dividend futures	10 years	N/A	4 years	10 years
Expiry date		3^{rd} Friday of December	3^{rd} Friday of December	3^{rd} Friday of December	Last trading day in March
Data frequency		Daily	Daily	Daily	Daily
Stock index ticker		SX5E	SPX	UKX	NKY
Dividend index ticker		DKESDPE	SPXDIV	F1DIVD	JPN225D

[40] Mixon & Onur (2017)

2.11 Figures

Figure 2.1: **Eurostoxx 50 Dividend Futures**. Price curve of dividend futures and discounted dividend futures on a random day (January 29^{th}, 2014) for purpose of illustration. Discounted dividend futures equal the present value of future dividends. Expiries occur on the third Friday in December of each expiry year.

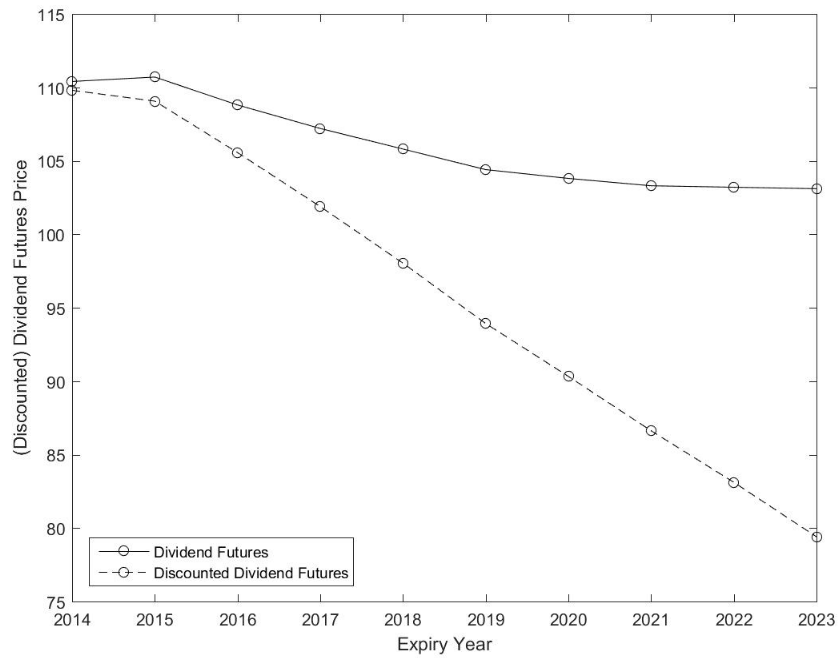
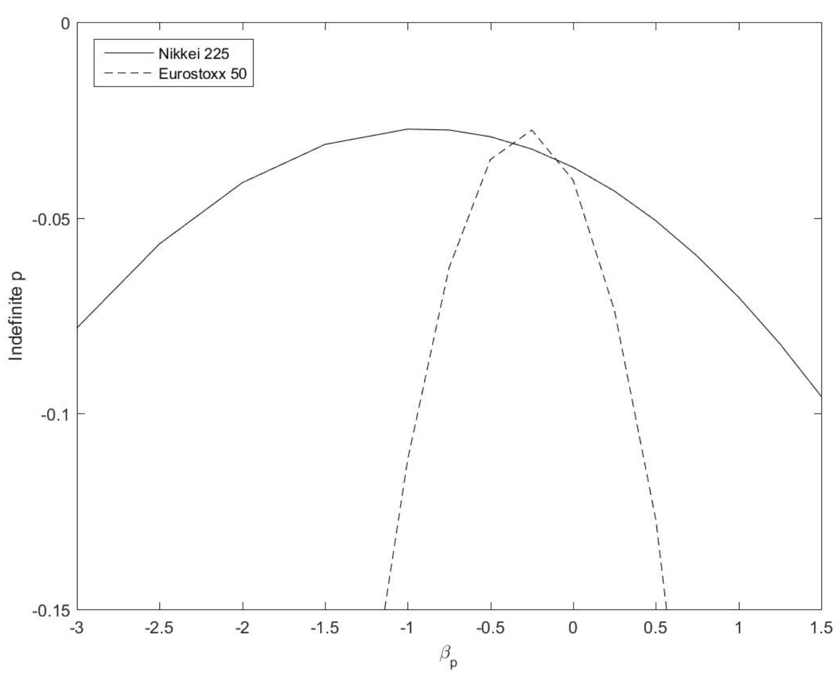


Figure 2.2: **Values of $\overline{p}$ for a given value of** β_p. Values of β_p are fixed to calculate $\overline{p}$. See Tables 1 and 2 for estimation results ($\beta_p = 0$). Parameters other than $\overline{p}$ do not change materially when β_p is varied.



Figures 2.3 and 2.4: **Mean absolute estimation errors**. Figures depict the average of the absolute estimation error of the two-state and the single state base model for the Eurostoxx 50 and Nikkei 225 index. The measurement variables are discounted dividend risk-adjusted growth rates of 1 to 8 years.

Figure 2.3: Eurostoxx 50

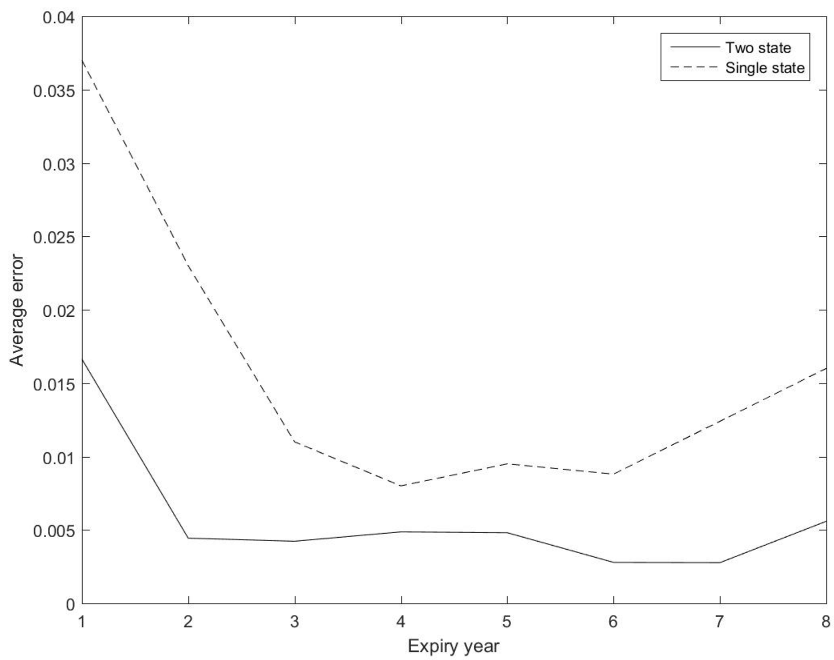


Figure 2.4: Nikkei 225

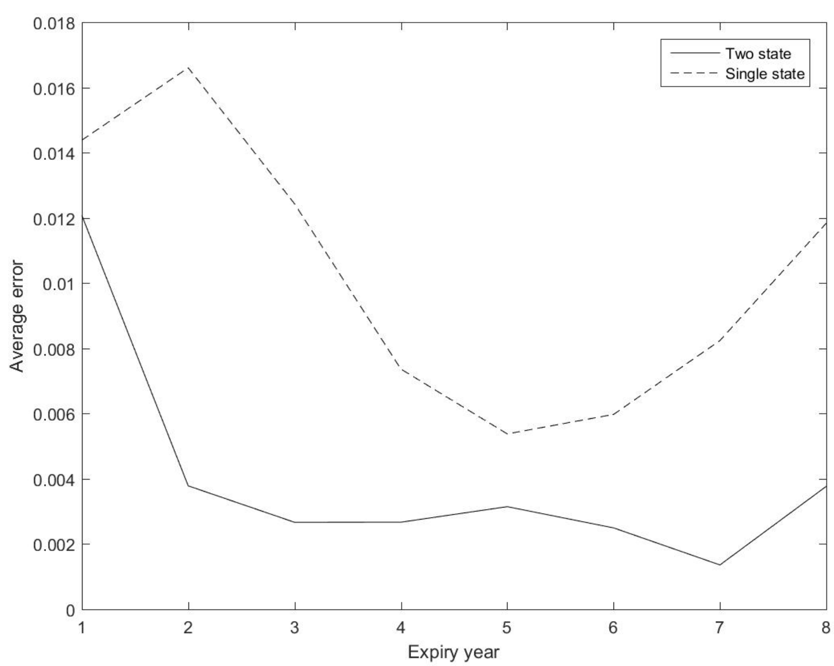
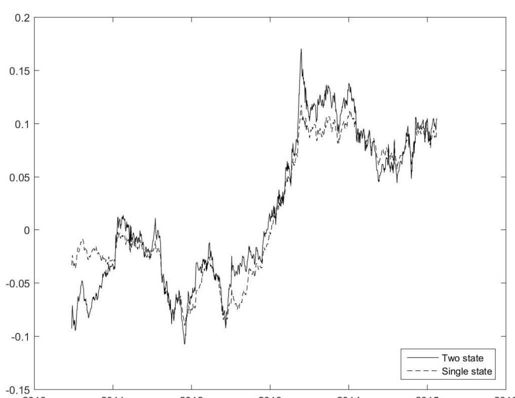


Figure 2.5: **Calibrated risk-adjusted dividend growth rates**. Figures 5(a) and 5(c) contain the 1 year growth rate $\pi_{t,1}$. Figures 5(b) and 5(d) contain average annual growth rates of the 4 years following the first year of growth: $\pi_{t,t+1\rightarrow t+5}$.

(a) Eurostoxx 50: 1 year growth

(b) Eurostoxx 50: 1 year forward 4 year growth

(c) Nikkei 225: 1 year growth



(d) Nikkei 225: 1 year forward 4 year growth

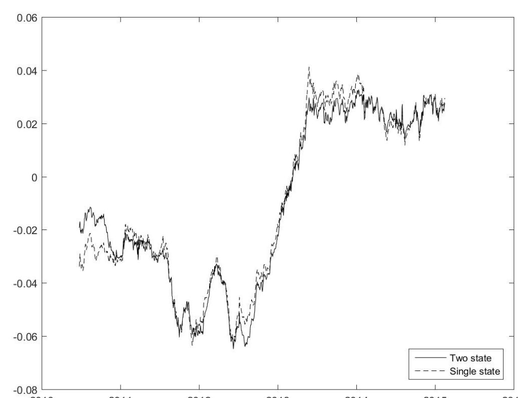


Figure 2.6: **Volatility of Dividend returns**: $\sigma_t\left(lnP_{t+1,n} - lnP_{t,n}\right)$
Volatilities are calculated both by the two state model and as observed in the data.

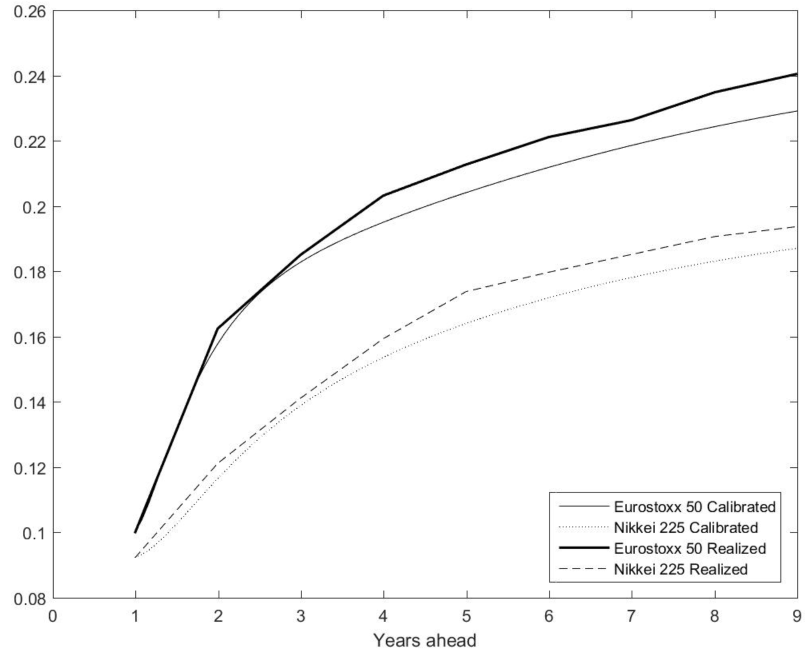
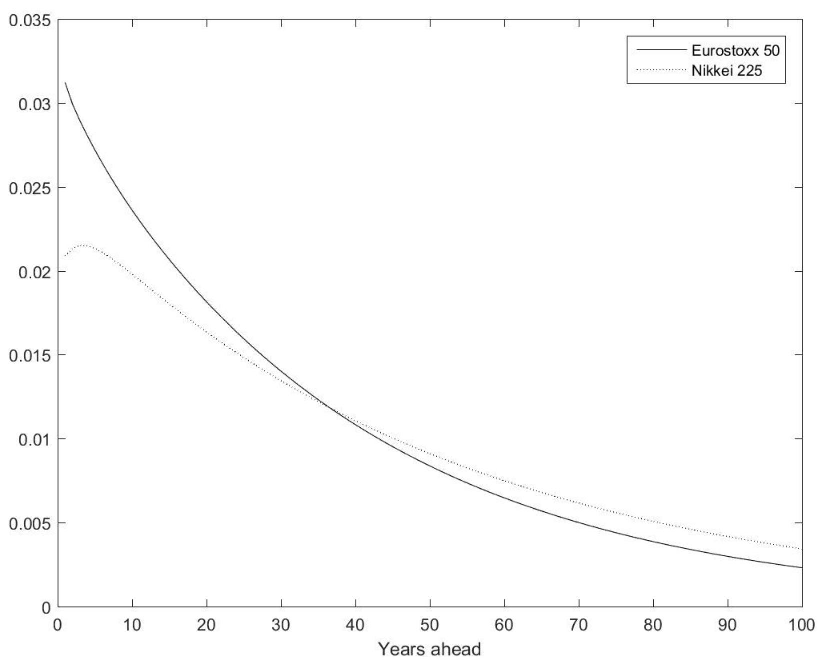


Figure 2.7: **Calibrated average Dividend Term Structure**. The average of calibrated present values of dividends per expiry year $\overline{P}_{t,n}$ is divided by the sum of the averages. This represents the average dividend yield per expiry year in present value terms.



Figures 2.8 & 2.9 portray the **present value model estimates for the level of stock indices** as described in $\hat{S}_t = F_{t,1} exp\left(-y_{t,1}\right)\ \left(1 + \sum_{n=2}^{\infty} exp\left(n\widehat{\pi}_{t,n} - \widehat{\pi}_{t,1}\right)\ \right)$ (equation 2.29), their ranges with an estimate for $\overline{p}$ of $-2\sigma_{\overline{p}}$ to $+2\sigma_{\overline{p}}$ and stock market observations S_t.

Figure 2.8: Eurostoxx 50

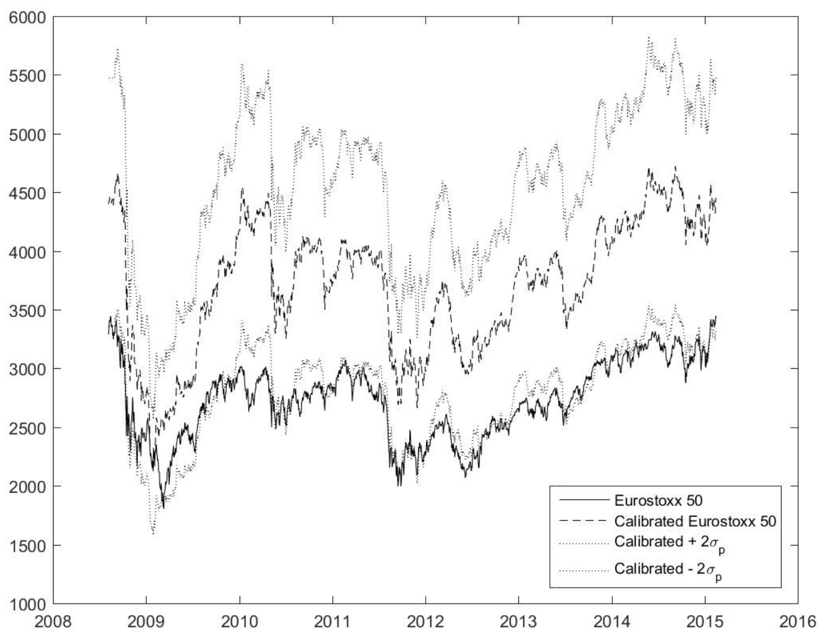


Figure 2.9: : Nikkei 225

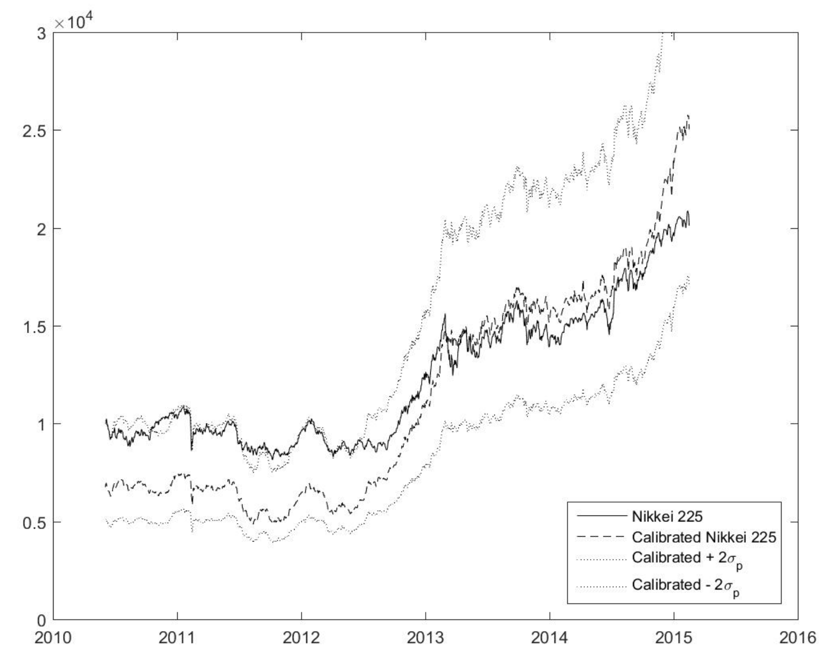


Figure 2.10: **Model implied dividend return volatility**: $\sigma_t\left(lnP_{t+1,n} - lnP_{t,n}\right)$. Our S&P 500 data is contrasted with the data for Long Run Risk and Habit Formation as compiled in Binsbergen et al. (2012).

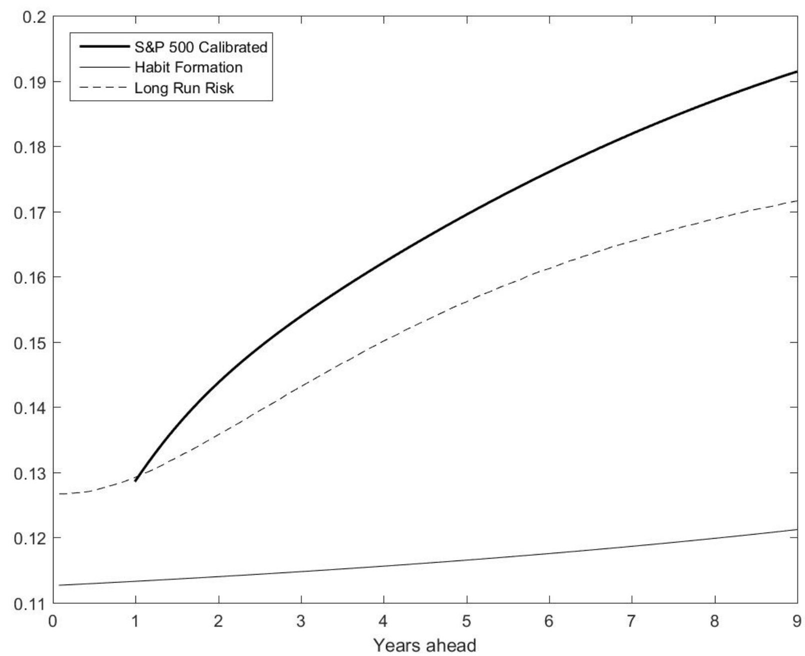


Figure A1: **Proportion of dividend payments** throughout the Eurostoxx 50 dividend index year. The first trading day of a dividend index year is the Monday following the third Friday of December. The chart depicts the average of the years 2005 to 2013.

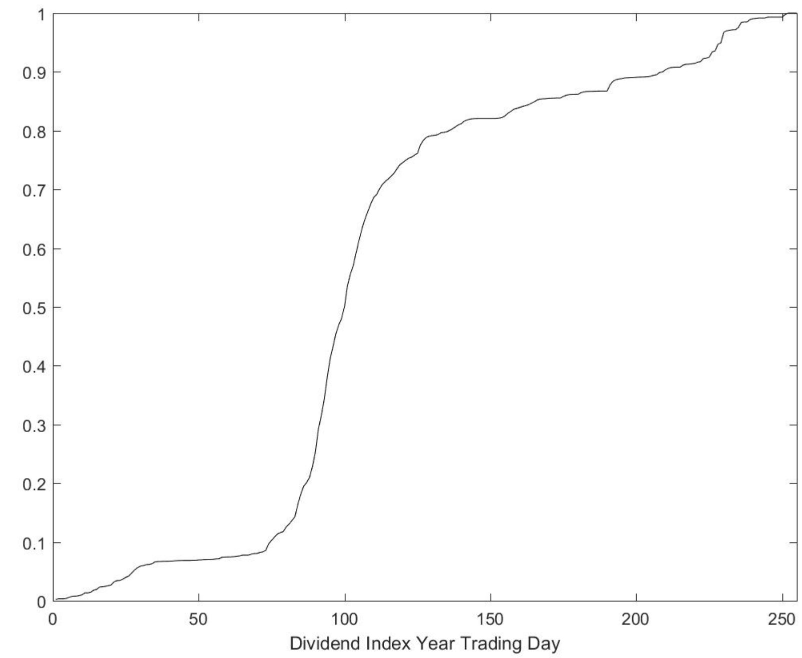
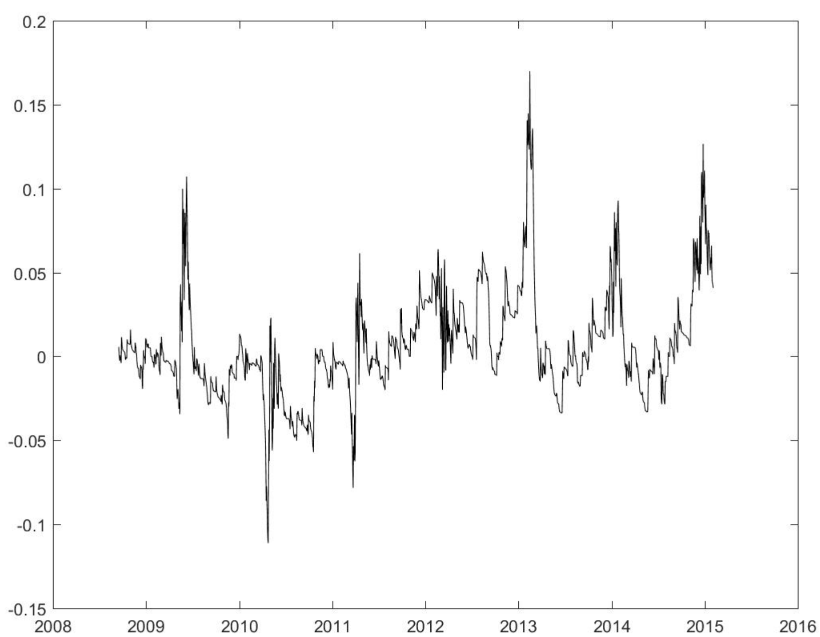


Figure A2: **Difference between the proportion of annual dividends paid out at a given date and their average** over the period 2005 to 2013 (Eurostoxx 50).



Chapter 3

Option Implied Dividends

I determine the valuation of future dividends for US companies as implied by option prices. This is the first paper in which the early exercise premium included in these prices is explicitly accounted for as part of finding these dividend valuations. From these implied dividend data, I build company-specific term structures of dividend growth relative to actual dividends. These term structures show substantial variation in slope over time as well as in the cross-section. Implied dividends predict actual dividends, particularly upward dividend changes. Stock prices do not respond to a dividend cut when it is priced in. An announcement to cut dividends causes a stock's price to drop by 2.6% on average, but if it is correctly predicted by implied dividends the response is negligible.

3.1 Introduction

The main component of a stock's value is its promise to future dividends. The present value identity for stock prices lays out the precise terms of the relationship, yet an adequate empirical connection between stocks and dividends is hard to establish. Shiller (1981) documents that the variation of dividends over time is far exceeded by the volatility of stock prices. This and other authors work with actual dividends as a proxy for future dividends. In view of these weak findings, could it be that actual dividends are not the best available representation for application in the present value of future dividends? I make the case to investigate the market valuation of future dividends for finding information about share prices.

Central to this argument is that the present value identity attributes value from the *discounted expectation* of dividends, and that the valuation of such expectations are more accurately measured than that future dividends are modeled. At the level of stock indices this is argued by several authors, summarized in Binsbergen and Koijen (2016). This literature circles around a particular product in financial markets: dividend derivatives. These products allow market participants to trade the dividends thrown off by the constituent companies of a stock index usually several years into the future. Their pricing thus provides an excellent gauge of future dividend valuations in the aggregate.

Addressing cross sectional asset pricing puzzles by means of dividend valuations requires access to priced dividends for individual companies. Unfortunately, such products are few and far between. Dividend futures for individual European stocks exist in small numbers since 2010, but they are scarcely traded and their liquidity is limited. In the US market, publicly traded dividend futures were introduced for the S&P 500 in 2016, but do not exist for individual companies. In order to find such pricing information, I propose to measure implied dividends for individual companies from the prices of stock options. This paper is the first to display such an approach and to apply its result in asset pricing.

The rationale to investigate option prices for this purpose is that they are a function of the assumptions made by market participants about the volatility of the underlying stock and about the dividends it will throw off. Options are the only instrument in financial markets which contain such forward-looking information about dividends for individual companies in meaningful numbers. They also exist for several expiry dates, which gives insight into a term structure of future dividends.

The case to imply dividend valuations paid by all constituent companies of the S&P 500 index from option prices has been made by van Binsbergen, Brandt and Koijen (2012). These authors apply the put-call parity (PCP) to index options for finding the implied dividends. This methodology presupposes that payoffs are the same for stock plus put option and strike plus call option throughout the life of the options, which is true for options that are not exercisable before their expiry. For index options, the direct application of PCP is an appropriate methodology, since they are indeed such European style. Options of individual US stocks, however, are American style and exercisable before their expiry date. The embedded early exercise premium introduces a deviation in the PCP which, without proper adjustment, would disqualify it for the purpose of finding implied dividends directly from the prices of American options. I propose a methodology to circumvent this issue by modeling option prices and solving for implied volatility and implied dividends at the same time, using a binomial tree model. The core of this procedure is the minimization of the squared difference between the model prices and the market prices of put-call option pairs in a single step.

The procedure provides a database of implied dividends for each firm out to three quarters ahead of the observation date. The data set is as large as the set of companies on which stocks options are traded. A subset of companies have stock options expiring in excess of two years (LEAPS), with an according horizon for implied dividends[1].

The data set bears out that implied dividends vary substantially over time and in the cross-section. Nevertheless they track actual dividends well, despite that these observable data do not play a part in calculating implied dividends. Over time, they are on average valued lower than actual dividends, except in periods immediately following bear markets. In the period of 1996 to 2015, the average actual dividend yield is 56 basis points per quarter. Expressed as an implied quarterly dividend yield, the average term structure shows a decline of a little over 1 basis points per quarter. Implied dividends in the highest cross-sectional quartile are often substantially above actual dividends and in the lowest they are mostly below. For longer horizons this dispersion is smaller than for shorter horizons.

Implied dividends show that market participants anticipate changes in dividends. In fact, 57% of dividend increases are correctly predicted at horizon of six months by implied dividends whereas dividend cuts are not predicted by implied dividends (close to 50 %). For five to twenty days ahead Fodor, Stowe and Stowe (2017) find stronger predictive strength in implied dividends. More generic tests also reveal that upward changes in dividends in particular are predicted by implied dividends. Future dividend raises are better predicted with higher sensitivity to implied dividends for longer horizons than dividend cuts.

These predictions are, however, obfuscated by the nature of dividend changes. Actual

[1] As a check on the robustness of minimization procedure, I include stock loan fees as part of the return to shareholders. Such fees can be regarded as a source of income to be derived from owning a stock in addition to its dividend. By expanding the methodology to find implied dividends, stock loan fees are implied simultaneously but separately from implied dividends. Implied fees correspond to actual fees, while showing a negative slope in their term structure (see the Appendix).

dividends are often left unchanged for an extended period of time by company management pursuing a dividend smoothing policy and the prevalence of such no-changes in dividends is vast. An estimation methodology that takes account of this feature of the data orders them into a separate category. Actual dividend changes are categorized into upward small and large changes, negative changes and zero changes. An ordered Probit estimation then reveals the probabilities of a raise, a cut or a no-change occurring for a given implied dividend being higher or lower than the dividend that was last paid. Although significant, the material effect of implied dividends on a change in such probabilities is not large. The Probit results show that the marginal effect to the probability of predicting an actual increase by implied dividends moving from below to above actual dividends amounts to about 2%.

Just as implied dividends predict changes in actual dividends, the effect of an announcement of a dividend change on stock prices is also affected if implied dividends correctly predict the change. Larkin, Leary and Michaely (2016) note that stock prices typically fall around an announcement to cut dividends by about twice as much as they rise in case of a dividend increase. They argue that market participants are less surprised by a dividend change if a company has cut dividends before and they show that this reduces the stock price response. This mitigated element of surprise plays a part once implied dividends are included in testing the announcement effect as well. Although dividend cuts are not well-predicted by implied dividends, the three-day stock price response to a dividend cut is reduced from –2.6% to about zero if it is correctly predicted by implied dividends. Any response reduction in case of dividend raises, however, is invisible.

This paper is organized as follows. The subsequent section sets out the procedure to imply dividend valuations from US stock options. This involves a discussion of the minimization procedure and the criteria to calculate the data and is followed by a description of the dividends as implied. An example of the data and the inclusion of stock loan fees in the minimization procedure concludes the section. In the subsequent section these data are deployed in several tests for their predictive power, both for actual dividends and the stock price response to announcements to change dividends. The final section contains the conclusions.

3.2 Implying dividends from option prices

The market valuation of dividends is found from the relationship between stock prices and futures. Pricing a future contract $F_{t,n}$ that matures at time n needs to take account of the time value of money r and for the fact that the buyer of a future does not receive the cash distributed by the asset, which is its dividend $P_{t,n}$ paid up to n in the case of a stock S_t:

$$P_{t,n} = S_t - e^{-r_{t,n}} F_{t,n}. \tag{3.1}$$

As futures of individual stocks are not sufficiently available for obtaining an empirically meaningful data set, I use options instead. Put-call parity (PCP) is a model free relationship which enables to find dividends as implied by option prices. The intuition is that the combination of opposite positions in otherwise equal call and put options[2] creates the payoff of a future. This pricing relationship is given by the equality of the difference between the price of a European call option $c_{t,n}$ and put option $p_{t,n}$ to the discounted difference between option strike K and future price $F_{t,n}$

$$c_{t,n} - p_{t,n} = (F_{t,n} - K)e^{-r_{t,n}}. \tag{3.2}$$

Combining (3.1) and (3.2) describes the PCP as I apply it, where the price of the future is replaced by option prices:

$$P_{t,n} = S_t + p_{t,n} - c_{t,n} - Ke^{-r_{t,n}}. \tag{3.3}$$

This depiction is clear cut for European options as the value of dividends can be traced back directly to the variables on the right hand side, which are all observable. But American style options on US stocks allow for exercise earlier than the expiry date resulting in a strictly positive early exercise premium. There is no need to deviate from the principle of the PCP approach, but a correction for the premium that the early exercise presents is required to apply (3.3).

3.2.1 The model

The standard for pricing American options applies a binomial tree, introduced by Cox, Ross and Rubinstein (1979). The binomial CRR model builds up stock prices and option prices by up and down scenarios in constant time steps as the horizon extends into the future until the expiry date of the option. The put and call option prices in the CRR model are

$$p_{t,n} = f(\sigma_{t,n}, P_{t,n}, S_t, n, K, r_{t,n}), \tag{3.4}$$

$$c_{t,n} = f(\sigma_{t,n}, P_{t,n}, S_t, n, K, r_{t,n}), \tag{3.5}$$

[2]Same underlying stock, strike and expiry date.

in which $P_{t,n}$ is the present value at time t of all dividends to be paid between t and n as implied by the call and put option pair expiring at n, $r_{t,n}$ is the risk-free rate and $\sigma_{t,n}$ is the implied volatility of the stock with price S_t.

The build-up of the binomial tree requires to specify whether dividends are paid at discrete points in time, or continuous through time. Discretely timed dividends, as I apply the model, allows for the possibility of a payment shortly before expiry, which would trigger exercise at such a date in the life of the option. The CRR model thus takes into account the pricing consequences for options on the stock of such precipitous exercise.

If no assumptions are made about either implied volatility or implied dividends, then unique solutions cannot be found for each equation (3.4) and (3.5) independently from each other. At the same time, each of the two implied unknowns should be equal for calls and puts. The implied volatility refers to the volatility of the stock price, which is the same stock for the call and the put option. The implied dividend refers to the dividend implied to be paid by the company issuing the stock, which is also the same for both options. It follows that, with two equations and two unknowns, unique values can be found for implied dividend and implied volatility under this presumption for pairs of call and put options which have the same strike, expiry date and underlying stock characteristics.

There are concerns that the implied volatility in particular is not the same for calls and puts. The literature on the pricing of options and potential causes for mispricing in the Black-Scholes-Merton (1973) framework looks for trading demand and supply pressure as a potential cause. Bollen and Whaley (2004) argue that changes in implied volatility of index and stock options are related to net buying pressure measured by trades occurring away from the bid/ask midpoint. This causes the implied volatility function across moneyness of the options not to be flat as expected in the classic framework. Cremers and Weinbaum (2010) state that differences between call and put implied volatilities do not reflect pure arbitrage opportunities, rather these represent deviations from PCP, which can be viewed as proxies for price pressures. The authors conclude that such deviations in option prices can lead stock prices by days. Gârleanu, Pedersen and Poteshman (2009) identify the relative demand for index and stock options by end-users. Option prices are priced away from the Black-Scholes-Merton framework as they cannot be hedged perfectly by intermediaries. The authors measure the expensiveness of options as their implied volatility relative to the expected volatility for the life of the option and establish a positive relation between option expensiveness and end-user demand.

In this paper concerns about short-lived price pressure are mitigated by the aggregation of the data. The implied dividend for an individual firm for a certain trading day is itself the product of several option pairs. Daily data are then averaged into a quarterly data set, which further reduces the impact of one-sided pressures as well as of noise.

There are other concerns with the exact application of the PCP. Kalay, Karakaş and Pant (2014) use deviations to establish the value of the voting premium. The difference of voting share prices from their non-voting synthetic cash flow equivalent using options measures the value of the right to vote as shareholders. They find the voting premium to be positive and increasing with the expiry of the options.

The two pricing equations for put and call options (3.4) and (3.5) each have the same two unknown implied dividend and implied volatility and the same observables. Cremers and Weinbaum (2010) as well as Gârleanu, Pedersen and Poteshman (2009) and Kalay, Karakaş and Pant (2014) employ in their analysis the implied volatilities delivered in the Ivy DB data set from OptionMetrics. As in Bollen and Whaley (2004), these are calculated under the constant dividend yield assumption (OptionMetrics, 2011). Harvey and Whaley (1992) state: "Since firms tend to pay stable quarterly dividends at regular periodic intervals during the calendar year, little uncertainty exists about the dividend parameters for short term index options.". This means that future dividends are proxied by the most recent dividend that actually has been paid. With that unobservable input to the model pinned down, it is possible to imply a value for the volatility of the underlying stock from the price of a single option.

From the market for dividend derivatives, however, it is established empirically that risk neutral dividend expectations are neither flat for any horizon nor constant over time (Binsbergen et al., 2013, and the previous Chapter). The actual expectation about future dividends prevailing in the market will be different from its last known value for two key reasons. The first is that the present value of a dividend expected to be paid at time n depends on several ingredients:

$$P_{t,n} = PV_t(D_{t+n}) = D_t exp(g_{t,n} - \theta_{t,n} - r_{t,n}). \tag{3.6}$$

Expected dividends $E_t(D_{t+n})$ grow at rate $g_{t,n}$ from currently paid dividends D_t and are discounted at the risk premium $\theta_{t,n}$ and the risk free rate $r_{t,n}$ for finding its present value $PV_t(D_{t+n})$. The assumption that the last known dividend will be continued at the same level beyond the observation date thus implicitly equates the objective expected growth rate $g_{t,n}$ to $\theta_{t,n} + r_{t,n}$ for all n.

Furthermore, instead of projecting dividends at a constant rate into the future, market participants are likely to change their expectations of future dividends over time based on their assessment of a company's willingness and ability to pay dividends. Such expectations may change whether the company publicly states to do so or not. It is the variation in the market valuation of dividends, as denoted in equation (3.6), that the following procedure is intended to capture.

I identify a call and put option pair that is the same for all characteristics, among which future dividends and volatility are unobservable. These two unknowns can be found, in principle, exactly given the option pricing equations (3.4) and (3.5). However, model implied prices $\hat{c}_{t,n}$ and $\hat{p}_{t,n}$ depend on building up the CRR binomial trees. They cannot be determined analytically by reverse engineering as this would involve starting from the back-end of the tree. This problem is resolved by equating model and observed prices respectively, while the option pricing equations for call and put are inverted at the same time.

In practice, the procedure starts by guessing values for the two unknown implied variables, calculating model prices and reiterating new guesses until convergence of model and observed prices. As both pricing equations require the same guessed value for the inputs, this is pursued by minimizing the quadratic difference between model prices and the observed end-of-day mid-prices of both a pair of call and put options in a single equation:

$$F = (p_{t,n} - \hat{p}_{t,n})^2 + (c_{t,n} - \hat{c}_{t,n})^2, \tag{3.7}$$

where F is a function of the implied dividend and implied volatility of the modeled option prices. The minimization of F given the observed option prices results in a unique value for both implied dividends between t and n and for implied volatility for horizon n for each option pair in the data set.

The minimization method is based on two main assumptions: Markets are transparent and efficient, and end-of-day close option prices contain all economically relevant information relevant to market participants. Implied volatility and implied dividends are the same for call and put options with otherwise equal characteristics. As these implied values refer to a single stock, this is a statement of fact rather than an assumption.

3.2.2 Data and aggregation

The Ivy DB database provided by OptionMetrics contains all options traded on US stocks starting in 1996 and are included in the data set until August 2015. The minimization is performed under the condition that prices relate to companies with CRSP share codes 10 or 11, which restricts the set to regular companies only[3]. Overall, the data set contains option prices for 5,700 individual companies, which constitues most of the stocks traded on the NYSE, AMEX and NASDAQ. The prices used are daily and are taken as the average of the

[3]This excludes REITS, ETFs and several other underlying types.

bid and ask closing prices[4]. The data set includes a total of 337 million option prices. The daily stock prices S_t and current dividends D_t are also taken from the Ivy DB database. The risk free rate $r_{t,n}$ is proxied by the Eurodollar Libor spot rate, downloaded from Datastream.[5]

To construct a data set that is a large as possible, but from which irrelevant options are eliminated, I apply the following filters:

1. Options of which OptionMetrics does not provide a value for implied volatility are excluded. As OptionMetrics applies the constant dividend assumption, implied volatilities cannot be calculated for options which are far in-the-money or far out-of-the-money. Such options are less frequently traded, which is the main reason to exclude them.

2. Options expiring within 90 days of the observation date are excluded. Once a dividend is announced, there is little uncertainty about the magnitude of the payment being made. As some firms announce dividend payments farther in advance than others, the degree of risk embodied in implied dividends payable in the near future will vary among them. Only a small fraction of firms announce dividends beyond a horizon of 90 days, so restricting the data set to longer option maturities allows for a like-for-like comparison of implied dividends among firms.

Many stocks exist in the data set for just a part of the data period, they sometimes have options traded on them and not all companies pay dividends consistently. Consequently, it is insightful to regard the prevalence of the data per period[6]. Over the 79 quarters in the data period, there are 153,589 quartercompanies with options, or some 64% of all available stocks (see Panel A in Table 3.1). More often than not, a single stock will have options of the same expiry date with several strikes traded on it. Options with different strikes do not have the same implied volatilities, as the volatility surface is not flat but skewed. But as long as the expiry dates of option pairs are the same, the implied value for dividends remain equal to the market's value put on the dividends paid before expiry regardless of the option strike. Option pairs with the same expiry will thus produce the same implied dividends, but varying implied volatilities for varying strikes.

On average, for every observation date and expiry date combination, there are three option pairs with different strikes in the data set. To reduce measurement error, it is useful to find the implied dividends from all option pairs available. I calculate implied dividends

[4]Market closing time of options on individual stocks is at 4:00 pm, which matches the closing time of the stocks. Timing differences between the two markets should not be a major concern.

[5]Gârleanu, Pedersen and Poteshman (2009), Cremers and Weinbaum (2010) also deploy Libor for this purpose.

[6]I work with quarters for reasons that are clarified later.

from available options pairs for a given expiry date and underlying stock and take their median to determine the relevant value for each observation date. Alternative methods to find averages deliver materially the same results.[7]

The minimization is performed for dividends to be paid discretely, as this accounts for the early exercise premium in the CRR model. This requires identifying dates beyond the observation date at which dividends are assumed to be attributed to the stock. The relevant date to include in the pricing model is the ex-dividend date, which is the first business day following the date on which the dividend is attributed to the stock. The settlement date at which the dividend is actually paid into the account of the owner is not relevant. If the owner decides to sell the stock as of the ex-dividend date, but before the payment date, then the dividend will not be paid into the owner's account.

The latest date at which a dividend has actually been paid is known at any observation date and is used as a starting point. Most US companies pay dividends at a quarterly frequency (see Panel B in Table 3.1). The subsequent date at which the next dividend is assumed to be paid is set to the last ex-dividend date plus 91 calendar days, the following quarterly payment goes ex-dividend after 182 calendar days and so on. Such assumptions about the exact dates cannot be avoided even if company statements about ex-dividend dates were analyzed piecemeal. Many stocks have LEAPS[8] traded on them, which are options with expiry dates extending out to over two and a half years (see Panel A in Table 3.1). In general, companies provide guidance about dividend dates of at most one year in the future. For these long dated options, an assumption about the date of going ex-dividend always has to be made.

Options with an expiry date of more than three months beyond the observation date can span more than a single implied dividend payment. The dollar value of each of them

[7]An alternative weighting scheme takes account of the impact of the bid ask spread on the precision of the estimates for the implied dividends. Bid ask spreads tend to be large for options that are traded infrequently, such as when their strikes are far in or out of the money. Large bid ask spreads involve a wide range of consistent mid prices, which increases the potential for measurement error. The larger the bid ask spread, the smaller the weight in the averaging of the dividend implied from among the option pairs

$$\hat{P} = \sum_{i=1}^{n} \frac{w_i \hat{P}_{K_i}}{\hat{P}_{K_i}}$$

in which the weight w_i for the option pair with strike K_i is set by

$$w_i = (0.05 + |BO_c| + |BO_p|)^{-1}$$

with $|BO_c|$ and $|BO_p|$ representing the bid ask spread of the call and put options respectively. A constant 0.05 is added to avoid inversion of zero. Although this approach appears more robust to measurement error, it delivers materially the same results and is not pursued.

[8]Long-term Equity Anticipation Securities.

is assumed to be the same. Relaxing this assumption would increase the number of free variables in the minimization which prevents finding unique solutions in the minimization of an individual option pair with more than one dividend payments in (3.7). At the same time, at any observation date the present value of the first to be paid dividend is implied from the first to expire option during which life a dividend is expected to be paid. From the options with the next quarterly expiry date, both the first and the second future dividends present values are implied simultaneously. By deducting the first quarter implied value, the second dividend's implied value can be isolated, and by iteration the same applies to dividends further into the future.

Companies do not necessarily have to be current dividend payers to have dividends implied from their stock options. Often firms do not pay dividends for a protracted period of time, or never pay dividends at all. Sometimes they are current payers, but interrupt payment for a few years before they return as payers. Several hundred companies in the data set have paid dividends at an interval different than quarterly. And many among them switch payment frequency.

The choice made for aggregation deals with these issues by dividing the data into subsets of observation quarters of companies which pay dividends at a quarterly rate and of observation quarters companies which do not pay dividends. The quarterly frequency is chosen since this allows to establish a data set of implied dividends of which the horizon is constant. In practice, the expiry dates of regular stock options with expiries shorter than nine months are dispersed over a calendar year in a manner that is not the same for all companies. They tend to follow a pattern of expiries in January/April/July/October, February/May/August/November or March/June/September /December. Exceptions exist, particularly in the latter half of the data period. Also, LEAPS always expire in January of up to two years and eight months ahead. If a monthly frequency were pursued, then the horizon would reduce by a month twice for each month that passes and then jump ahead for two months. Aggregation of daily implied dividends to a monthly frequency thus produces a term structure of which the horizon is not the same across companies. Aggregating to buckets of quarterly averages resolves this issue. Appendix I sets out the procedure for this aggregation.

3.2.3 Summary description of the minimization results

The results of the minimizations are pictured in dividend yields, that is, in present values of implied dividends relative to the share price at quarter t. All companies in the data set have regular options traded on their stock, while some companies also have LEAPS traded on their stock. The expiry of regular options extends out to no more than nine months, but the expiry of LEAPS can be as long as two years and eight months. As a result, growth rates beyond the third quarter are based on LEAPS and thus consists of a smaller sample of companies (see Panel B in Table 3.1). All data discussed are value weighted.

The term structure of dividend yields averaged over the data period for stocks without LEAPS (Figure 3.1) starts at 0.60% for currently paid dividends and drops to 0.53% for the average of the first, second and third quarter forward. Average implied dividends of stocks with LEAPS are at about the same level and show a noticeable decrease for quarters beyond the third. The implied dividend yield over the eight quarters following the current quarter is valued at no more than 0.45% to 0.48%, which is a fifth to a quarter less than the current dividend yield. These figures broadly match the levels and slope pattern for dividends based on option prices reported elsewhere (Binsbergen, Brandt and Koijen, 2012).[9]

Figures 3.2 and 3.3 illustrate implied dividend yields over time next to the current dividend yield for value weighted averages across all companies who pay dividends at a quarterly rate. Figure 2 shows the implied dividend yield of one quarter ahead and the sum of one and two quarters ahead, Figure 3 contains the implied dividend yield of one to three and one to eight quarters ahead. Several observations stand out. First of all, option prices appear to capture the expectation of dividends sensibly over time. The implied dividend yields follow the current dividend yield well, despite that they are calculated in the minimization process without reference to actual dividends. Second, implied dividends are below current dividends for longer dated horizons. This means that growth rate $g_{t,n}$ is smaller on average than the risky discount rate $\theta_{t,n} + r_{t,n}$ (see Equation (3.6)). Lastly, the level difference between current dividend yields and implied dividend yields decreases during the time span of the data set. In the latter half, the implied dividends are higher relative to current dividends for both short and long dated options. Lower interest rates can partly explain this structural shift.

The term structure emerges clearly in Figures 3.4 and 3.5, which portray sample average growth rates of implied dividends relative to actual dividends. Over time, there is substantial variation in dividend growth, with an upward trend as time passes. Implied dividend growth is negative throughout for longer horizons, but hovers around zero for horizons of one and two quarters. In 2008 a change has occurred taking dividend growth substantially higher.

[9] These authors calculate dividend prices of the S&P 500 index based on PCP over the period 1996 to 2009. The levels they find are close despite differences in data period, smaller set of firms and S&P-weighting versus value-weighting. The slope of the term structure they report at an average of -5.15% for horizon 6 to 12 months ahead and -2.95% for 12 to 18 months ahead. The average slopes found here are a little less than twice as steep. The authors report a positive slope of 0.50% for 18 to 24 months ahead, whereas I find a decrease of -3.2% (figure 3.1). A slope from actual dividends to 6 months ahead implied dividends can not be calculated from the data they present.

From their peak in 2008, current dividends measured as dividends per share fell by 10%. In part, the sudden increase in dividend growth is therefore a denominator effect.

Returning to the combined set of companies with and without LEAPS, current dividend yields show substantial variation. Breaking them down into deciles indicates a dividend yield of less than 0.2% per quarter for the lowest decile, the highest decile companies pay in excess of 1.1% dividend per quarter (Figure 3.6).

Growth rates of implied dividends vary substantially among individual companies as well. The shortest horizons give the biggest dispersion in growth rates (Figures 3.7 and 3.8). The breakpoint of the highest decile in first quarter growth is nearly 100%, while the breakpoint of the lowest decile is minus 50%. Options of companies in these deciles thus price a near doubling and halving respectively of dividends in the following quarter. Growth rates in consecutive quarters remain widely dispersed. The average growth rate after the third quarter ranges from nearly minus 40% for the lowest decile breakpoint, to over 60% and 30% growth for the highest decile breakpoints of three quarters and two years average dividend yields.

3.2.4 An example of a change in dividend policy and anticipatory pricing in options

The present value relationship assumes that the value of a stock is derived from expected dividends, but such valuation does not necessarily require that any dividend is currently paid. In fact, the expectation that dividends will be paid may lay far ahead in the future. Implied dividends give some insight into the length of that expectation horizon.

Often companies attain substantial market capitalization even though they do not pay dividends. A case in point is Apple Inc. (AAPL US) which attained a market capitalization in 2011 of US$ 354 billion following seventeen years of not paying dividends. Its operating profitability reached US$ 35 billion in fiscal year 2011.

The procedure to imply dividends from option prices works equally well for stocks whether dividends are being paid or not. The implied values found for such stocks are usually close to zero. Only when market participants anticipate dividend payments in the future will implied dividends be significantly above zero.

Throughout the data set, there are 683 instances in which a company started to pay dividends after at least one year of no dividend payments. In these instances, there is some anticipation discernible from implied dividends ahead of time. Of these companies, the dividend valuation implied one quarter ahead amounted to a little over 30% of the newly installed dividend paid one quarter later. A quarter earlier, this had already reached 28%

and three quarters before the payment the valuation reached 17%. For companies with LEAPS traded on their stock, dividend valuation going back even further could be obtained. However, there are only 19 instances of newly started dividend payments in the data set among such companies. In view of this small number, it is instructive to single out an example.

On 19 March 2012 Apple Inc. announced its intent to start paying dividends later in the year[10]. Figure 3.9 shows implied dividends for Apple priced into LEAPS expiring in January 2013. In the six months preceding the announcement, the implied dividend had already risen from zero to about $ 0.50 per share. Market participants adjusted their expectations, anticipating at least some dividend payment until the start of the year 2013, which is clearly reflected in option prices before the announcement was made.

Next to the uncertainty over dividend initiation itself, there is uncertainty over its timing. Figure 3.10 shows the Apple dividends implied from an option series which expires in July 2012, which is before the first payment went ex-dividend in August 2012. It is clear from the implied dividends in this series too that the market anticipated payment of dividends during the first quarter of 2012 before the announcement on 19 March. But when Apple announced that the ex-dividend date would be in August, the implied dividend vanished from this expiry series. Market participants appeared to have realized that payment was to occur after its expiry in July.

Following the announcement, the option implied dividend per share per quarter until January 2013 rose from about $0.50 first to about $1.25 and then to $2.00. The implied dividend remains lower than the announced level of $2.65 because the exact number of discrete dividend payments deviates from the model assumptions at the time of the press release on the 19^{th} of March.[11]

3.2.5 Stock loan fees and robustness

The PCP relationship contains all instruments relevant to an investor wishing to run a portfolio of options. However, market makers in particular use long and short positions in stocks as well to hedge their exposure to stock prices as a result of their option book. In cases where such market participants take short positions in the stock, they need to take

[10] Apple stated: "... the Company plans to initiate a quarterly dividend of $2.65 per share sometime in the fourth quarter of its fiscal 2012, which begins on July 1, 2012."

[11] Apple paid two dividends of $2.65 before the expiry of the January 2013 LEAPS. Since Apple didn't pay dividends up to that point, payments are modeled to occur 91 days, 182 and 273 days following the observation date. There are thus three payment dates over which to divide the present value of two anticipated dividend payments between the announcement on 19 March and 20 April, which is 273 days before the expiry of this LEAPS in January 2013. It is noteworthy, however, that after 20 April the implied dividend still doesn't reach the announced amount of $ 2.65.

into account the cost of doing so. D'Avolio (2002) documents that these stock loan fees amount to 25 basis points per annum on average. In his sample 91% of all stocks are not hard-to-borrow, and these can be borrowed at a running cost of 17 basis points.[12]

If cash income from stocks not only appears as dividends, this should be incorporated in the PCP. It is straightforward to add loan fees f_t to the dividend $P_{t,n}$ paid by the company on its shares. The left-hand side of the PCP as depicted in equation (3.3) is thus expanded with fees as follows,

$$P_{t,n} + S_t(e^{f_{t,n}} - 1) = S_t + p_{t,n} - c_{t,n} - Ke^{-r_{t,n}}, \tag{3.8}$$

which, for small $f_{t,n}$, approximately equals

$$P_{t,n} = S_t e^{-f_{t,n}} + p_{t,n} - c_{t,n} - Ke^{-r_{t,n}}. \tag{3.9}$$

Market data for stock loan fees are available and they can in principle be included in the minimization in (3.7). However, stock loan fees may be adjusted throughout the period that the stocks are on loan. Stock lenders and borrowers may terminate lending at any point in time (Avellaneda and Lipkin, 2009). As a consequence, the actual cost of borrowing a stock is not precisely captured by loan fee data. Moreover, fees in the PCP in equation (3.9) presume a cost for borrowing stocks throughout the life of the option. The typical period for borrowing stocks is measured in days. In the data sample, option expiries extend out to over two years. As a consequence, market data about fees that represent the actual cost anticipated for borrowing stocks until the expiry date of particular options do not exist and including fees in the PCP would require unverifiable assumptions about their future value.

An alternative route to take account of stock loan fees is to imply them in a manner similar to implied dividends. The minimization in (3.7) is performed on two equations, allowing for two unknowns: implied volatility and implied dividend. To find another implied variable is feasible by expanding the number of option pairs in the minimization from one to two. The two pairs are the same except for the strike, which increases the number of unknowns by a third variable: the implied volatility of the second option pair. The number of equations is increased from two to four, thus allowing for a fourth variable to be estimated as well: implied stock loan fees. Similar to dividends, the stock loan fee refers to the share and not the options. Therefore, the stock loan fee is the same for the modeled prices of both option pairs and can be implied as a fourth variable. The fact that they can be identified

[12]These stocks are *General Collateral*, which are all loaned at the same fee since they are readily available and the fee only needs to compensate for the service of stock lending. Stocks that are hard-to-borrow are loaned at a sometimes much higher fee, which makes up for the difference between the average fee of all stocks and the General Collateral fees documented by Avellaneda and Lipkin (2009).

separately from dividends is due to the fact that stock loan fees are deemed continuous, whereas dividends are paid discretely. Appendix II contains a section which explains how the implied stock loan fee is included in the minimization procedure in more detail.

The implied stock loan fees produced by the minimization are shown in Figures 3.11 and 3.12. Throughout the data period, the fees move between 15 and 20 bppa up to two quarters ahead and slightly lower in the quarters beyond. Implied fees for different expiries move in conjunction with each other.

The levels of the implied fees come close to the market data for General Collateral fees of 17 bppa documented by D'Avolio (2002). The fact that implied fees move in a tight range reflects that investor opinions about future fees change little compared to current fees as time passes. This is confirmed by the term structure of implied stock loan fees in Figure 3.13. Its slope is negative but shallow. If average stock loan fees do not change much over time, investors may be motivated to make an assumption to this effect in the pricing of options.

Cross-sectional dispersion is shown in Figure 3.14. The breakpoint for the decile with smallest implied stock lending fees stands at 12 bppa against 22 bppa for the decile with the largest fees. Such a difference is not substantial. Stocks that are hard-to-borrow can require fees for several percentage points per annum (Avellaneda and Lipkin, 2009) and the largest decile would be expected to have a higher breakpoint. However, such fees occur more often for small stocks (Avellaneda and Lipkin, 2009) and these are less likely to be present in the database of stock which have options traded on them. Moreover, in the minds of investors high fees do not necessarily remain high for a prolonged period. Implied fees represent their expectations of stock lending until expiry of the options, which may be led more by the average of fees than by the fee of a particular stock that happens to be in demand.

As implied stock loan fees do not show large dispersion, their effect on the dispersion of implied dividends should be small. Implied dividends which are calculated while taking account of stock loan fees according to (3.9) indeed show very similar patterns and levels as those calculated without regarding fees. Although the minimization method to find implied dividends appears robust to the inclusion of fees, I chose to omit their mention for the remainder of this paper. The interpretation of implied dividends is thus to include stock loan fees throughout this paper.

3.3 The predictive power of implied dividends

An obvious question is whether implied dividends have predictive power for actual dividends. The opinion that investors express about the value of future dividends as implied in option prices can be tested as an explanatory variable for the same dividends being paid later.

In the period 1996 to 2015, nearly 74,000 quarterly dividend payments were made by companies with stocks on which options were traded (Table 3.3). Most of the time companies leave dividend payments unchanged relative to the previous quarter, which confirms that a commonly accepted policy of dividend smoothing is widespread. Nonetheless, there are many instances of a change in dividend policy. Four types of policy changes are identified. In over 11,000 cases firms increase dividend, where in more than 2,000 cases they decreased it. At only 66 occurrences, firms terminating dividends are rare, and the majority of them go out of business not long after. In total 683 times a company started paying dividends, which is defined as a dividend payment following at least four quarters of non-payment.

At a quarterly rate of 0.79%, the dividend yield is on average highest in the last quarter for companies that make a payment before they cease to do so. Presumably a high yield reflects a low share price in which discontinuation of the firm is anticipated. The group of companies decreasing dividends, reduce it from a dividend yield of 0.66% to 0.48% (the median \$ cut is -46%). When they increase payment, the change is more modest from 0.55% to 0.62% (the median \$ increase is 11%). Companies initiating dividend payments are the most frugal relative to their stock price and introduce payment at a yield of 0.44%. The biggest group of companies leaves their dividends unchanged at a dividend yield of 0.57%. This pattern confirms that companies are careful to increase their dividends, but once it happens they act more determined.

The example of implied dividends anticipating Apple's dividend initiation in 2012 suggests that dividends implied by option prices are forward looking not only in theory, but also in practice. An interesting question is whether this relationship holds up across companies and through time. In particular, I investigate whether implied dividends provide information about future dividends over and above lagged dividends. Since dividend changes are often zero and do not show a continuous distribution, an ordered Probit estimation is pursued in addition to a standard OLS. I conclude by considering the stock price response to company announcements to change dividends in order to establish whether dividend changes that are correctly anticipated by implied dividends reduce this response.

3.3.1 Predictive OLS regressions

If market participants were to predict dividend changes correctly, then the relationship between actual and implied dividend changes would be linear, with a coefficient equal to one and zero intercept. The following regression equation serves as a first attempt to test the hypothesis:

$$d_{j,t+n} = \alpha_t + \beta_t p_{j,t,n} + \epsilon_{j,t} \tag{3.10}$$

The actual dividend change $d_{j,t+n}$ is defined as the relative change in the dividend paid n quarters following the observation quarter t for a given (portfolio of) stock j:

$$d_{j,t+n} = \frac{D_{j,t+n} - D_{j,t}}{D_{j,t}}. \tag{3.11}$$

The implied dividend change $p_{j,t,n}$ is the dividend change for n quarters ahead relative to the observation quarter as implied by the data $P_{j,t,n}$:

$$p_{j,t,n} = \frac{P_{j,t,n} - D_{j,t}}{D_{j,t}}. \tag{3.12}$$

Implied dividends, at the same time, are valuations of future dividends and not objective expectations of future dividends (see equation (3.6)). The difference between the two is the dividend risk premium, which is not the subject of investigation here. Under the hypothesis that this risk premium is constant over time, it should surface as a non-zero intercept. However, if the dividend risk premium is time-varying, it may also show up in the coefficient of implied dividends. The appropriate hypothesis for testing the predictive content of implied dividends is that their relationship to actual dividends is close to linear with a coefficient close to one.

The implied dividend data are quite noisy. Figure 3.15 contains a scatter plot of actual and implied dividend changes two quarters ahead of the observation quarter. The plot shows that they are heavily scattered away from a hypothesized regression line of model (3.10), which is at least in part caused by measurement error at the stage of the calculations laid out in the previous section.

The plot also bears out that downward and upward dividend changes do not mirror each other. Under the model, downward changes should be concentrated in the southwest quadrant and upward changes in the northeast quadrant. For upward changes this appears reasonably clearly, but it is only barely visible for downward changes. As far as visible, downward changes align vertically close to zero implied dividend growth, while positive changes align horizontally slightly above zero actual dividend growth. Neither pattern fit well to model (3.10).

To assess these patterns statistically in addition to the base model, upward and downward changes are separated and introduced as two different independent variables to an alternative model, as follows:

$$d_{j,t+n} = \alpha_t + \beta_{up,t} I_{p>0} \times p_{j,t,n} + \beta_{down,t} I_{p<0} \times p_{j,t,n} + \epsilon_{j,t}, \tag{3.13}$$

in which dummies $I_{p>0}$ and $I_{p<0}$ attain the value of 1 for a positive implied dividend change and negative respectively and 0 otherwise.

OLS regressions are run cross-sectionally for each quarter t of the data period of 1996 to 2015. This quarter-by-quarter approach allows for calculating standard errors of quarterly beta's, in a manner similar to Fama and MacBeth (1973). The predictive power of implied dividends is tested for n one to three quarters ahead, resulting in 77 to 79 regressions per n.

The estimated coefficients from the regressions of models (3.10) and (3.13) are displayed in Table 3.2. The averages of the coefficients of model (3.10) bear out that implied dividends do not line up strongly with actual dividends. Only 7 to 10% of a change in implied dividends translates into actual dividend changes. Even though these coefficients are small, they are significant, particularly for predictions two and three quarters ahead.

The alpha's in these regressions are close to but significantly larger than zero. There may be at least two explanations that implied dividend changes are smaller than actual dividend changes. Dividends increase throughout the data period, by on average 2.4% per quarter. Considering that only a fraction of such average change is picked up by implied dividends, the intercept reflects a substantial part of the average increase. A second explanation is that implied dividends are risk-neutral. If model (3.10) is true, while expected values of dividends coincide with their realization later, then the alpha would reflect the risk premium and the time value of money as a result of which a positive value is expected.

The test results of the alternative model (3.13) confirms the visual inspection of the data in Figure 3.15. Applying up and down dummies to implied dividend changes shows that negative predicted changes attract an unexpected negative sign and are not significant. Only positive changes demonstrate statistically relevant predictive power for actual dividends[13]. Their coefficients come out slightly stronger than in the base model, which is to be expected given the filtering of the weak power present in downward changes. Market participants only reflect upwards dividend changes in option implied dividends as dividend cuts remain largely unpredicted by implied dividends.

[13]The weakness of negative predictions is also reflected in an average R^2 of 0.35% for a model that only includes negative predictions (1.10% to 1.60% for model (3.13)).

3.3.2 Predictive ordered Probit

The scatter plot in Figure 3.15 depicts another particularity of the relationship between actual and implied dividends. Dividends are set by company management and are thus discretionary in nature. The plot shows that firms often choose a "round percentage" for a dividend change, for example, minus 50%[14] or plus 50%. One in five dividend payments constitutes a change in the size of the payment, with increases occurring about five times as often as decreases (Table 3.3). At four out of five times, the most prevalent dividend change is therefore no change as also noted by Guttman, Kadan and Kandel (2010) and Baker, Mendel and Wurgler (2016). The last paid actual dividend is therefore highly relevant to the next payment due to the wide-spread policy of dividend smoothing.

While implied dividends may signal a change in future dividends, only when a change in actual dividends occurs will implied dividends be identified as relevant to future dividends. Moreover, any change in dividends tends to be large. The median increase amounts to about 11 percent while the median decrease is –46 percent (Table 3.3). If implied dividends predict a small change in actual dividends, it may often be too small to be effected by company management. Baker, Mendel and Wurgler (2016) note that this obscures a highly non-linear relationship where changes around zero occur much more frequently than larger movements. The high persistence in actual dividends thus presents a challenge to the linearity assumed in models (3.10) and (3.13).

Isolating the many zero-change occurrences in actual dividends in the data set requires a simple comparison of the number of instances in which a dividend change is correctly predicted by implied dividends. The north-east quadrant in Figure 3.15 contains correctly predicted increases, the south-west quadrant contains correctly predicted decreases and the other quadrants show incorrect predictions. A positive implied dividend, defined as an implied dividend value above its median, correctly predicts a future increase in dividends in 54 to 57% of the observed instances, depending on the horizon of the dividend implied. A decrease is predicted correctly by implied dividends below their median just about as often as it is predicted incorrectly (Table 3.4). This poor result for dividend cuts is matched by the weak explanatory power of implied decreases in the OLS estimates (Table 3.2). The only other study of this relationship in the literature is by Fodor, Stowe and Stowe (2017). These authors investigate 389 dividend payments made by US companies in 2008 and 2009 and find that a bigger discount of an implied dividend relative to the previous dividend increases the likelihood that the dividend will be cut next time one is scheduled. Their findings are more conclusive and a possible reason is that the horizon at which they appraise the predictive

[14] Nearly one in four decreases is exactly minus 50%. Baker, Mendel and Wurgler (2016) discuss salience and round numbers in dividends and dividend changes in more detail.

power of implied dividends is five days and twenty days, whereas the data in this paper use implied dividends to predict at a horizon of two to three quarters. The example of Apple Inc. in Figure 3.9 demonstrates that six months ahead of their dividend initation announcement, implied dividends did not signal a change in dividends but clearly did in the month prior to the announcement[15].

When implied dividends are larger than actual dividends, there should be at least some predictive effect on the probability of actual dividends increasing in the future. The particular characteristics of the data set induces to order the actual dividends data set into categories. This is followed by Probit estimation which produces the probabilities of the dividends being changed as predicted by positive or negative implied dividends.

Actual dividend changes d are ordered into four categories: a middle no-change category that is identified as actual dividend growth close to zero, a decrease category and two increase categories for small and large increases. These categories are represented formally as follows:

$$y = \begin{cases} 1 & if \quad d \leq -0.01 \\ 2 & if \quad -0.01 < d \leq 0.01 \\ 3 & if \quad 0.01 < d \leq b \\ 4 & if \quad d > b \end{cases}$$

Actual dividend changes d are ordered into a category $y = 1$ for cuts, of which the upper boundary is set to -0.01. Zero changes are categorized separately in $y = 2$.[16] Small increases are allocated to the third category $y = 3$ and large increases are allocated to the fourth category $y = 4$. Carving up increases in y serves to account for the wide-spread policy of dividend smoothing, which leads to increases being both small and plentiful. Carving up the first category into one for small and large cuts is not material to the results due to the small number of small dividend cuts and because cuts are large more often than not. Boundary b between the third and the fourth category will be varied in order to exhibit the sensitivity of the parameter coefficients to it.

Ordinal values y represent latent variable y^* which is a linear function of the model:

$$y^*_{j,t} = \beta_{up,t} I_{p>0} + \epsilon_{j,t}. \tag{3.14}$$

[15]It is clear that the data set used by Fodor, Stowe and Stowe (2017) includes only a fraction of US firms that pay dividends and have options traded on their stock. In the data set used in this paper, the number of quartercompanies for 2008 and 2009 is more than 5.000.

[16]The purpose of the boundaries that define $y = 2$ is solely to isolate the zero-change instances into a separate category. Setting these boundaries to small fractions other than but close to -0.01 and 0.01 is not material to the results.

The up-dummy $I_{p>0}$ attains a value of 1 when implied dividends are larger than actual dividends and 0 otherwise. The probabilities of selecting one of the categories are equal to the probabilities of the latent variable falling into an estimated intercept (α_i) category. These probabilities then are equal to the standard cumulative distribution function Φ value of the estimated parameters:

$$P(y = 1) = P(y^* \leq \alpha_1 | p, \beta_{up}) = \Phi(\alpha_1 - I'_{p>0}\beta_{up})$$

$$P(y = 2) = P(\alpha_1 < y^* \leq \alpha_2 | p, \beta_{up}) = \Phi(\alpha_2 - I'_{p>0}\beta_{up}) - \Phi(\alpha_1 - I'_{p>0}\beta_{up})$$

$$P(y = 3) = P(\alpha_2 < y^* \leq \alpha_3 | p, \beta_{up}) = \Phi(\alpha_3 - I'_{p>0}\beta_{up}) - \Phi(\alpha_2 - I'_{p>0}\beta_{up})$$

$$P(y = 4) = P(\alpha_3 < y^* | p, \beta_{up}) = 1 - \Phi(\alpha_3 - I'_{p>0}\beta_{up})$$

Coefficient β_{up} indicates the impact of a unit change in implied dividends on the probability of one of the categories occurring. The model is estimated using maximum likelihood for each quarter in the data set. The coefficients are then averaged over all quarters and their t-statistics are calculated following the method of Fama & MacBeth (1973), similar to the t-statistics in the OLS regression in the previous section. Regressions are performed for dividends one, two and three quarters ahead of the observation quarter.

Results of the Probit estimation are shown in Table 3.5 for boundary value $b = 0.10$.[17] For one quarter ahead the intercepts are further removed from the distribution mid-point than for two quarters ahead and, in turn, they are further removed than for three quarters ahead (Panel A). A change in implied dividends thus signals a larger likelihood of predicting a change in actual dividends the further ahead such change is implied. A potential reason is that nearby dividend changes may be known as they are announced sometimes several months before the ex-dividend date.

The positive implied dividend coefficient β_{up} has the expected positive sign. The inference is that a change in implied dividends from negative to positive reduces the likelihood of actual dividends falling in *cut* category 1, and increases the likelihood of falling in category 4, in which dividends rise by more than b. The coefficient of positive changes varies between 0.056 to 0.090 over the three quarters horizon, differing statistically from zero based on Fama & MacBeth t-statistics.

Are these coefficients material to the probabilities of actual dividends moving from one category to another? The marginal effects of the coefficients in this Probit estimation are

[17] Figures 3.16 and 3.17 show marginal effects for other values of b.

defined as the change in the likelihood of falling into a particular category in response to implied dividends pointing to an increase instead of a cut. Translating this marginal effect to the discrete character of the dummy-variable is performed by differencing the probability of a certain category of a dividend increase from a dividend cut:

$$\delta P(\alpha_{i-1} < y^* \leq \alpha_i | \beta_{up} I_{p>0}) = [\Phi(\alpha_i + \beta_{up} I_{p>0}) - \Phi(\alpha_{i-1} + \beta_{up} I_{p>0})] - [\Phi(\alpha_i) - \Phi(\alpha_{i-1})]$$

Panel B in Table 3.5 shows the probabilities of a category occurring in case of a negative implied dividend. For example, among observations of negative implied dividends, the probability of dividends actually being cut two quarters later is 0.047. Marginal effects on the likelihood of falling into the $y = 1$ (decreasing) or $y = 2$ (zero change) actual dividend category are negatively affected by implied dividend changes (Panel C in Table 3.5), as is determined by coefficient β_{up}. In the example given, if implied dividends move from negative to positive, the probability of dividends being cut two quarters later drops by 0.008 to 0.039. An implied dividend larger than actual dividend thus reduces the chance of falling into these categories and increases the chance of falling into the one of the two dividend increase categories $y = 3$ and $y = 4$.

The marginal effect of a positive implied dividend for an actual dividend change to fall in one of the positive categories is 2.1% to 2.7%. For dividend increases cut-off at 10%, the chances of falling into the increase $y = 4$ category above 10% is larger than to fall into the $y = 3$ small increase category. On average, a positive implied dividend hints at a larger increase in dividends rather than a small one.

Increases are accounted for in two separate categories to deal with dividend smoothing. This is a policy practiced by company management to change dividends only slowly in response to earnings changes. Based on survey evidence, Lintner (1956) observes that many managements believe that most stockholders prefer a reasonably stable dividend rate and that they put a premium on stability or gradual growth in the rate. Survey evidence in Brav et al. (2005) shows that managers feel strongly that the penalty for reducing dividends is substantially greater than the reward for increasing them. Evidenced by survey data, companies often pursue a policy to smooth dividend payments as long as conditions regarding stable earnings, among others, are fulfilled (Brav et al., 2005). Together with a desire to avoid dividend cuts, such a policy leads to increases in dividends being more frequent but also smaller than dividend cuts. It is worthwhile to find out whether smaller dividend increases, induced by regular policies, are indeed better predicted by implied dividends than larger ones that are more likely to reflect a one-off adjustment to the firm's structurally improved financial position. Cutting off the data set of dividend increases into two separate increase categories $y = 3$ and $y = 4$ serves this purpose.

The sensitivity of the marginal effect of a positive implied dividend to the value of the upper category boundary b is instructive (the results discussed thus far refer to $b = 0.10$ (Table 3.5)). The marginal effect relative to the probability of a dividend falling into the $y = i$ category for decreasing implied dividends provides insight into the importance of moving from negative to positive implied dividends:

$$\frac{\delta P(\alpha_{i-1} < y^* \leq \alpha_i | \beta_{up} I_{p>0})}{P(\alpha_{i-1} < y^* \leq \alpha_i | \beta_{up} I_{p>0})} = \frac{[\Phi(\alpha_i + \beta_{up} I_{p>0}) - \Phi(\alpha_{i-1} + \beta_{up} I_{p>0})] - [\Phi(\alpha_i) - \Phi(\alpha_i - 1)]}{\Phi(\alpha_i) - \Phi(\alpha_{i-1})}.$$

Figures 3.16 and 3.17 shows this marginal effect conditional on boundary d for the two raise categories $y = 3$ and $y = 4$. A boundary slightly above 1% constructs a narrow category, which involves that a positive implied dividend has a limited effect on improving the probability of predicting an increase correctly to that small raise category. Larger boundaries increase the size of the category and thus the probability of a raise falling into it, but the probability of predicting the small raise category increases more than proportionally up to a boundary of approximately 25%[18]. Predicting the large raise category deteriorates already at a small boundary. Consequently, a move from negative to positive implied dividends improves the chances of predicting a raise smaller than 25%. Raises larger than that are predicted better by implied dividends only marginally. The policy of firms to raise dividends at a steady moderate pace rather than in large jumps thus appears in market pricing.

3.3.3 Stock price response to dividend changes

The degree of predictability of future dividends may be picked up in the response of the stock price to a change in dividend policy. If an announcement to change dividends is expected, then it should be reflected in the stock price before the fact. Larkin, Leary and Michaely (2016) document that market reaction, measured as three-day stock return in absolute terms, to dividend reductions is more than double that of increases. They also find that institutional investors, mutual funds in particular, are more likely to hold dividend-smoothing stocks, but retail investors are less likely to do so. Dividend smoothing thus affects the composition of a firm's shareholders but, according to the authors, has little impact on its stock price.

The previous section discusses that negative implied dividends appear to have very little to say about actual dividends, while positive implied dividends do, and with the expected sign. Implied dividends thus indicate that investors do not see dividend cuts coming and may consequently be more surprised by them than in the case of increases. I investigate

[18]The level difference between the proportional marginal effects of differing horizons shown in these figures is due to the differences found in the intercepts (Table 3.5).

whether the difference in stock price responses between cuts and increases as well as the size of the response is connected with the expectations of future dividends that is part of the option implied valuation. I also test the impact to stock prices of dividend changes for being correctly or incorrectly predicted by implied dividends.

To establish the data sample, I identify all dividend changes of companies paying dividends at a quarterly rate in the period of 1996 to 2015 with stocks that have options traded on them as before. Following the approach of Grullon et al. (2002)[19], small changes due to rounding and recording of stock splits, as well as extreme observations, are eliminated by limiting the dividend announcements to those with absolute value of changes in quarterly common dividend per share between 12.5% and 500%. The sample is restricted to distribution events in which the declaration date is a non-missing trading date and there is no more than one dividend announcement made per event. For every dividend change, the three-day Cumulative Abnormal Return ($CAR_{j,t}$) is the sum of daily returns of the stock of the announcing firm i around the announcement ([-1,+1] trading days) minus the CRSP value-weighted market return. The implied dividends $p_{j,t}$ are calculated as the median of the ten trading days preceding the dividend announcement date.

The stock price response is regressed on the actual dividend changes and implied dividend changes in the following model:

$$\begin{aligned} CAR_{j,t} = \alpha_t + \beta_{1,t} I_{d<0} + \beta_{2,t} I_{d>0} \times d_{j,t} + \beta_{3,t} I_{d<0} \times d_{j,t} + \beta_{4,t} I_{p<0} + \\ \beta_{5,t} I_{d>0} \times I_{p>0} + \beta_{6,t} I_{d>0} \times I_{p<0} + \beta_{7,t} I_{d<0} \times I_{p>0} + \beta_{8,t} I_{p<0} \times I_{p<0} + \epsilon_{j,t}, \end{aligned} \tag{3.15}$$

from which parameters are included in the regression in several compositions. Results are presented in Table 3.6, showing only the two-quarter ahead implied dividends for ease of presentation. One and three quarter ahead implied dividends deliver broadly the same results, as usually different implied horizons jointly point to an increase or a cut.

Starting without implied dividends, the pattern found by Larkin, Leary and Michaely (2016) that dividends cuts weigh more on stock returns than dividends raises is repeated. The intercept shows a 0.7% three-day price return in response to dividend increases and a return of -1.6% for dividend cuts ($\alpha + \beta_1$, model (1)). Once the interaction of the size of the change is included, taking into account that the average cut is -0.42, the average response to a cut is about the same ($\alpha + \beta_1 + \beta_3 \times \mu_d$, model (2)). This data sample also shows that the impact of the size of dividend cuts on stock prices β_3 is ten times as large as that of dividend increases β_2[20].

[19] Their approach is followed by Larkin, Leary and Michaely (2016) as well.

[20] Larkin, Leary and Michaely (2016) report a ratio of four between dividend cuts and increases.

Next I test whether implied dividends decrease the surprise impact of a dividend change on stock prices. If the implied dividend points to a dividend cut by means of a down-dummy $I_{p<0}$ then the positive surprise impact is indeed reduced to -0.4% ($\alpha + \beta_1 + \beta_4$, model (3) versus -0.6%, model (2)). Statistically these results are not strong, the coefficient of the down-dummy is not significant and the R^2 does not improve once the dummies are included.

Will stock prices respond differently to a dividend change that is correctly predicted by implied dividends from one that is not? A direct test of this question is pursued in model (4) in Table 3.6, which deploys four interaction terms. The regressors capture the four possible scenarios when a dividend change occurs: a correctly predicted increase or cut and an incorrectly predicted increase or cut. The results point to a very clear distinction between the response to the kind of change. A dividend raise is followed by a positive three-day stock return of $0.6 - 0.7\%$ whether correctly predicted by implied dividends or not. A cut, however, causes a substantial drop in stock prices (-2.6%) if it is not anticipated by implied dividends. If the cut is predicted, the response is only -0.2%. The coefficient for unpredicted cuts is strongly significant and subsumes most of the impact of a dividend cut on stock returns (β_1).

In summary, as a predictor of stock returns around dividend announcement, implied dividends do not necessarily affect the impact from an announcement. But once implied dividends are used to predict dividend changes, it is the incorrectly predicted cuts that move stock prices. Taking this conclusion together with the earlier results leads to the following interpretation on the element of surprise from changing dividends. Testing equation (3.13) shows that increased dividend payments are significantly predicted by implied dividends, whereas dividend cuts are not. It is reasonable to presume that the expectations contained in implied dividends also prevail in the valuation of stocks; if market participants have reason to anticipate that a firm will raise dividends, then they will reflect this in the equilibrium price of its stock as well as in its dividends. Once the firm delivers on a higher dividend, the stock price may be unmoved by this fact alone. Since dividend cuts are not well predicted by implied dividends, a dividend cut usually is a surprise causing a stock's price to drop.

Larkin, Leary and Michaely (2016) find an explanation for the large negative response to dividend cuts in controlling for whether firms have cut at least once before. If they have, then the absolute impact on the stock price is about equal for an increase or a cut. A previous cut raises the awareness to the possibility of another cut in the future, reducing the element of surprise if a subsequent cut occurs. Their argument is thus similar to implied dividends predicting a dividend change: the announcement effect on stock prices reduces if there is reason to be less surprised about it.

3.4 Conclusion

There is a lot of information about the valuation of future dividends of individual firms to be gained from option prices. The extant literature on future dividends focuses on derivatives on stock indices. E.g. Binsbergen, Brandt and Koijen (2012) connect implied dividends to the term structure of risk premiums. But a cross-sectional approach to identifying future dividends has been lacking thus far. This study is the first to present a methodology to do so and applies the data found to their predictive power for dividends and stock returns.

A novel element in the methodology to imply dividends from option prices is that it clears the hurdle of dealing with the early exercise premium that is not present in index options but is an element of value in American style options of US stocks. The Cox, Ross and Rubinstein (1979) binomial tree takes account of this premium. The tree is constructed simultaneously for a pair of put and call options with otherwise equal characteristics by guessing the same values for implied dividends and implied volatilities for the put and call pricing models.

The implied data found do not equal actual dividend data. This should be a warning sign to the assumption often made in the model-pricing of options that future dividends are equal to actual dividends. On average, however, they are consistent with and slightly lower than actual dividend data, showing a modest negative term structure.

At the level of individual companies implied dividends are prone to measurement error. Nonetheless, they prove meaningful to predict actual future dividend changes, notably dividend increases. Dividend cuts are not well predicted. But if implied dividends do correctly predict a future dividend cut, the stock price response to the announcement thereof reported by Larkin, Leary and Michaely (2016) is reduced substantially.

In view of these results, there is good reason to pursue further research on implied dividends. The first is to improve on the methodology itself and to further investigate their relationship with actual dividends and stock prices. The second is to investigate whether implied dividends contain idiosyncratic risk that can be isolated from systemic risk, contradicting traditional CAPM reasoning. An obvious pursuit is to construct a factor based on implied dividends and establish whether portfolios of high implied dividends produce high or low stock returns.

3.5 Appendix

3.5.1 Aggregation of daily implied dividend data into quarterly buckets

The daily dollar value of the implied dividend as found in the minimization of equation (3.7) is divided by the stock price of the day. This daily implied dividend yield is then averaged over the period from the first day following the third Friday in a given month until and including the third Friday of the next month. Constructing months this way guarantees that averages refer to the same option expiry date. Only if an implied dividend falls before an expiry date is it included in the option price. If it is paid after the expiry date it is not, even though it may still occur in the same calendar month. Hence the redefiniton of months. The average is calculated after dropping the two most extreme values in each month.

The next step in the reduction of the data frequency is to bucket the values for each three months into a quarterly dividend yield. This involves taking an average again, but the horizon of the three implied dividends is not constant and it is therefore appropriate to interpret these as buckets instead of averages. Quarterly buckets are only taken if daily solutions for implied values exist for at least 75% of the days of the quarter. Equipped with quarterly dividend yields, a growth rate with a quasi-constant horizon is calculated, starting from a dividend actually paid in a given quarter to implied dividends in the ensuing quarters. A quarterly data set has the added benefit that it brings the frequency of the dividends implied in line with the dividend payment frequency that is the most prevalent among US companies. Lastly, I assume that any distortion in implied dividends due to inventory imbalances among market makers, asymmetric information or fixed cost is of constant mean over the daily implied values averaged to quarterly buckets.

There are two important arguments not to reduce the frequency to less than quarterly. First, the data period from 1996 to 2015 is only 20 years. Time series analysis would suffer from less variation due to fewer data points in case of an annual or semi-annual frequency. Second, most stocks have options maturing eight to nine months into the future, but many among them do not have longer dated LEAPS traded on them. A consistent pattern of implied dividends cannot be offered for annual frequency for these companies, although semi-annual would still be feasible.

3.5.2 Stock loan fees included in the minimization

In order to find implied stock loan fees, two pairs of call and put options are identified that have the same characteristics except for the strike instead of one pair of options. The modeled prices of all four options share the same implied dividend and implied stock loan fee. For each option pair with the same strike volatilities are implied separately from each other. This leads to four unknown variables to be implied: dividend, stock loan fees and the volatility of the first pair and the volatility of the second pair. F is a function of all four:

$$F = (p_{t,n,k} - \hat{p}_{t,n,k})^2 + (c_{t,n,k} - \hat{c}_{t,n,k})^2 + (p_{t,n,k+1} - \hat{p}_{t,n,k+1})^2 + (c_{t,n,k+1} - \hat{c}_{t,n,k+1})^2, \quad (3.16)$$

and is minimized as follows. Each model price is calculated starting by guessing values for the four unknown variables which serve as input to the CRR binomial tree. The four model prices are compared to their market prices in equation (3.16) after which new values are guessed until convergence of model and observed prices.

The procedure is performed for option pairs with strikes that are next to each other. For example, if there are 6 adjacent strikes available of options with the same expiry date, the minimization is performed on option pairs with strikes 1 and 2, 2 and 3, and so on, until strikes 5 and 6. The implied values for dividends and stock loan fees represented as the output of the minimization procedure are the medians of the five values implied.

3.6 Tables

Table 3.1: **Option and LEAPS Distribution and Dividend Payment Frequencies**

Prevalence of options and dividend payment frequencies for US stocks traded from 1996 to August 2015. A quartercompany is a quarter of daily data in which at least 45 trading days of options are available for a given stock.

Panel A: Excluded from the data set are options of which OptionMetrics does not provide implied volatilities and with maturities shorter than three months. LEAPS are options with a maturity at start of at least two years.

Panel B: Payment frequency of dividends per quartercompany. The data set includes options of stocks with quarterly dividend frequencies only.

	Number of Quartercompanies
Panel A: Prevalence of options	
Stocks total	238,964
Stocks without options	85,375
Stocks with options	153,589
Stocks with LEAPS	39,389
Stocks without LEAPS	114,200
Panel B: Dividend frequency of stocks with options	
Any	153,589
Annual	813
Semi-annual	2,284
Quarterly	78,623
Monthly	531
None paid	71,338

Table 3.2: **Predictive regressions of actual dividend changes on implied dividend changes.**

Actual n-quarter dividend change is defined as the growth rate of a dividend paid n quarters ahead of the observation quarter t for a given company i: $d_{i,t+n} = \frac{D_{i,t+n}-D_{i,t}}{D_{i,t}}$. Implied dividend change is defined as the value in quarter t of the dividend implied by option prices for quarter $t+n$ relative to the actual dividend paid in quarter t: $p_{i,t,n} = \frac{P_{i,t,n}-D_{i,t}}{D_{i,t}}$. For each quarter in the data set the regression equation is estimated for one, two and three quarters ahead. The coefficient estimates are averaged over t and their t-stats are calculated by dividing the average coefficient values by their standard deviation, multiplied by the square root of the number of quarters. The t-statistics are reported in parentheses. The estimated equations are:

$$d_{i,t+n} = \alpha_t + \beta_t p_{i,t+n} + \epsilon_{i,t+n}$$

$$d_{i,t+n} = \alpha_t + \beta_{up,t} I_{d>0} \times p_{i,t,n} + \beta_{down,t} I_{d<0} \times p_{i,t,n} + \epsilon_{i,t+n}$$

	Number of quarters ahead of implied dividends					
	1		2		3	
Intercept	0.004	0.002	0.018	0.016	0.037	0.029
	(0.97)	(0.33)	(4.95)	(4.27)	(6.26)	(5.04)
β_t	0.097		0.102		0.168	
	(2.38)		(6.68)		(5.88)	
$\beta_{up,t}$		0.100		0.105		0.184
		(2.37)		(6.55)		(6.04)
$\beta_{down,t}$		-0.015		-0.002		-0.062
		(-0.77)		(-0.08)		(-1.87)
R^2	0.009	0.011	0.011	0.014	0.012	0.016

Table 3.3: **Prevalence of dividend changes.**

Actual n-quarter dividend growth is defined as the growth rate of a dividend paid n quarters ahead of the observation quarter t for a given company i: $d_{i,t+n} = \frac{D_{i,t+n} - D_{i,t}}{D_{i,t}}$. The upper panel describes the number of quarterly dividend changes and their median and average size for five different types of dividend changes from 1996 to 2015 in US listed stocks that have options traded on them. The second panel shows the average dividend yield during the quarter before the quarter in which the dividend payment was made or terminated. The third panel shows the same for the quarter in which the dividend payment was made or terminated. The definitions for dividend changes made by Baker and Wurgler (2004, p 1134) partly coincide. Initiations as defined here equal new and list dividend payers, while Terminations equal new non-payers in Baker and Wurgler (2004).

Panel A	Increases	Decreases	Initiations	Terminations	Unchanged	Total
Number	11,564	2,357	683	66	59,211	73,881
Median change (%)	10.81	-45.71		-100	0	0
Average change (%)	24.16	-42.32		-100	0	2.46
Panel B	*Quarter before dividend change*					
Average Dividend Yield	0.55	0.66		0.79	0.57	0.57
StDev	0.37	0.54		0.60	0.40	0.40
Panel C	*Quarter of dividend change*					
Average Dividend Yield	0.62	0.48	0.44		0.57	0.57
StDev	0.38	0.43	0.39		0.41	0.41

Table 3.4: **Correctly and incorrectly predicted dividend changes.**

The table shows the number of instances in the data set of correctly and incorrectly predicted dividend changes (zero-changes are excluded) and the fraction of correct and incorrect predictions. Correct implied dividend predictions of non-zero changes in actual dividends are defined as follows: if an n-quarter ahead implied dividend at time t is higher than its median and the n-quarter ahead actual dividend of the corresponding firm is increased relative to the actual dividend at time t, then the implied dividend correctly predicts the increase. Correctly predicted dividend cuts are determined similarly but in the downward direction. Incorrectly predicted dividend changes occur when the implied dividend points to an increase while a decrease materializes and vice versa. Definitions of implied and actual dividend changes are as before.

	1 quarter ahead		2 quarters ahead		3 quarters ahead	
	Number	Fraction	Number	Fraction	Number	Fraction
Correct increases	4,807	0.577	8,790	0.573	11,118	0.545
Correct decreases	795	0.497	1,404	0.507	1,850	0.510
Incorrect increases	3,531	0.423	6,546	0.427	9,295	0.455
Incorrect decreases	805	0.503	1,365	0.493	1,779	0.490
Total	9,938		18,105		24,042	

Table 3.5: **Ordered Probit estimations of actual dividend changes on implied dividend changes.**

Ordinal values y represent latent variable y^* which is a linear function of the model: $y^*_{i,t} = \beta_{up,t} I_{d>0} + \epsilon_{i,t}$. The up-dummy $I_{d>0}$ attains a value of 1 for values of $p_{i,t,n}$ larger than zero, which means that the implied dividend is higher than the actual dividend paid in quarter t, and zero otherwise. Actual dividend changes d are ordered into categories as follows:

$$y = \begin{cases} 1 & if \quad d \leq -0.01 \\ 2 & if \quad -0.01 < d \leq 0.01 \\ 3 & if \quad 0.01 < d \leq b \\ 4 & if \quad d > b \end{cases}$$

Probabilities of falling into category $y = i$ are given by:

$P(y = i) = P(\alpha_{i-1} < y^* \leq \alpha_i | \beta_{up} I_{d>0}) = \Phi(\alpha_i + \beta_{up} I_{d>0}) - \Phi(\alpha_{i-1} + \beta_{up} I_{d>0})$

for $y = [1, 4]$, and where Φ is the cumulative normal distribution function. The coefficients indicate the impact of a unit change in implied dividends on the probability of one of the categories occurring.
The increase in the likelihood of falling into category i due to a positive implied dividend is calculated by deducting the likelihood of a cut falling into a category i from the likelihood of an increase falling into that category:
$\delta P(\alpha_{i-1} < y^* \leq \alpha_i | \beta_{up} I_{p>0}) = [\Phi(\alpha_i + \beta_{up} I_{p>0}) - \Phi(\alpha_{i-1} + \beta_{up} I_{p>0})] - [\Phi(\alpha_i) - \Phi(\alpha_{i-1})]$
Panel A describes the coefficients of the Probit estimation, t-statistics are reported in parentheses. Panel B shows the probabilities of a category occurring when implied dividends are negative. Panel C contains the marginal effects of a move in implied dividends from negative to positive on the probability of falling into category $y = i$. Boundary b is set to 0.10.

		Number of quarters ahead of implied dividends		
		1	2	3
Panel A	$\hat{\alpha}_{1,t}$	-1.855	-1.596	-1.442
		(-73.40)	(-55.09)	(-45.91)
	$\hat{\alpha}_{2,t}$	1.077	0.630	0.297
		(48.06)	(24.95)	(10.84)
	$\hat{\alpha}_{3,t}$	1.465	1.095	0.843
		(54.18)	(33.20)	(21.09)
	β_{up}	0.090	0.080	0.056
		(6.48)	(4.98)	(3.29)
Panel B	Probability of category y being chosen for negative implied dividends ($\beta_{up} = 0$)			
$d \leq -0.01$	$y = 1$	0.026	0.047	0.067
$-0.01 < d \leq 0.01$	$y = 2$	0.812	0.662	0.528
$0.01 < d \leq 0.10$	$y = 3$	0.077	0.136	0.189
$d > 0.10$	$y = 4$	0.085	0.155	0.216
Panel C	Marginal effect from a move in implied dividends from negative to positive			
$d \leq -0.01$	$y = 1$	-0.006	-0.008	-0.008
$-0.01 < d \leq 0.01$	$y = 2$	-0.015	-0.018	-0.014
$0.01 < d \leq 0.10$	$y = 3$	0.008	0.009	0.006
$d > 0.10$	$y = 4$	0.013	0.018	0.016

Table 3.6: **Stock price response to dividend announcements.**

The dependent variable $CAR_{i,t}$ is the stock price response to dividend change announcements, defined as the return of a stock on which a dividend change is announced during the three days around the announcement [-1,1] minus the CRSP value-weighted market return. Actual n-quarter dividend change is defined as the growth rate of a dividend paid n quarters ahead of the observation quarter t for a given company i: $d_{i,t+n} = \frac{D_{i,t+n} - D_{i,t}}{D_{i,t}}$. Implied dividend change is defined as the value in quarter t of the dividend implied by option prices for quarter $t+n$ relative to the actual dividend paid in quarter t: $p_{i,t,n} = \frac{P_{i,t,n} - D_{i,t}}{D_{i,t}}$. Dummies $I_{d>0}/I_{d<0}/I_{p>0}/I_{p<0}$ take on the value 1 when actual (D) or implied (P) dividends are increased/cut and are 0 otherwise. The regressions are performed by including either dummies or the interaction of the parameters with dummies. The t-statistics are reported in parentheses.
The estimated equations are:

$$CAR_{i,t} = \alpha + \beta_1 I_{d<0} + \beta_2 I_{d>0} \times d_{i,t} + \beta_3 I_{d<0} \times d_{i,t} + \beta_4 I_{p<0} +$$
$$\beta_5 I_{d>0} \times I_{p>0} + \beta_6 I_{d>0} \times I_{p<0} + \beta_7 I_{d<0} \times I_{p>0} + \beta_8 I_{p<0} \times I_{p<0} + \epsilon_{i,t}$$

	(1)	(2)	(3)	(4)
Intercept	0.007	0.006	0.005	
	(6.82)	(4.62)	(3.52)	
$I_{d<0}$	-0.023	-0.007	-0.006	
	(-11.38)	(-1.35)	(-1.25)	
$I_{d>0} \times d$		0.003	0.003	
		(1.74)	(1.71)	
$I_{d<0} \times d$		0.030	0.030	
		(3.49)	(3.49)	
$I_{p<0}$			0.002	
			(0.90)	
$I_{d>0} \times I_{p>0}$				0.006
				(2.03)
$I_{d>0} \times I_{p<0}$				0.007
				(4.08)
$I_{d<0} \times I_{p>0}$				-0.026
				(-8.81)
$I_{d<0} \times I_{p<0}$				-0.002
				(-0.52)
R^2	0.021	0.024	0.024	0.016

3.7 Figures

Figure 3.1: **Average actual and implied dividend yields.** Data are shown over the 1996-2015 data set for a given number of quarters ahead of the observation quarter for all firms (value-weighted) with and without LEAPS traded on their stocks.

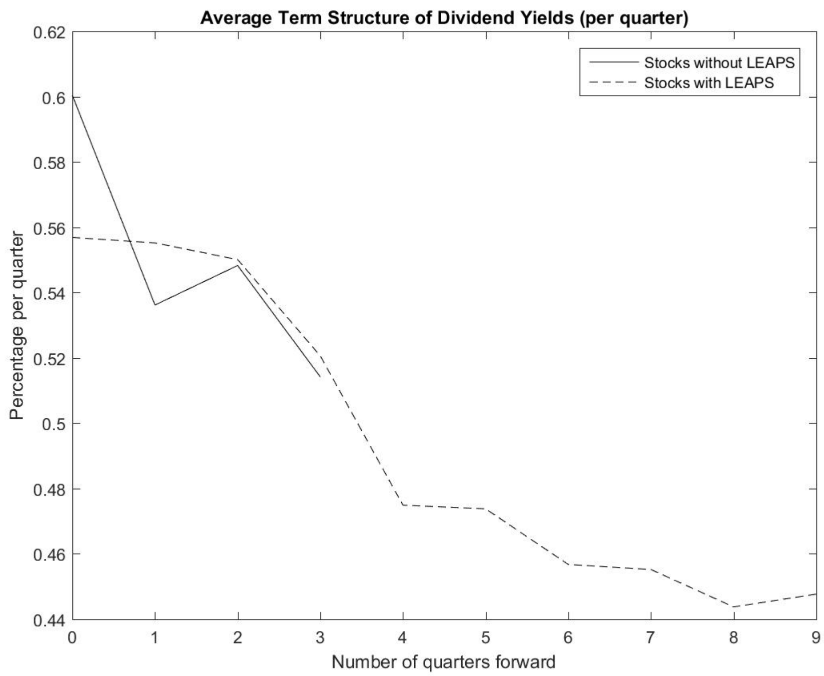


Figure 3.2: **Dividend yields per quarter, current, 1 quarter ahead and 1 to 2 quarters ahead.** Yields are value-weighted over all firms who pay dividends at a quarterly rate.

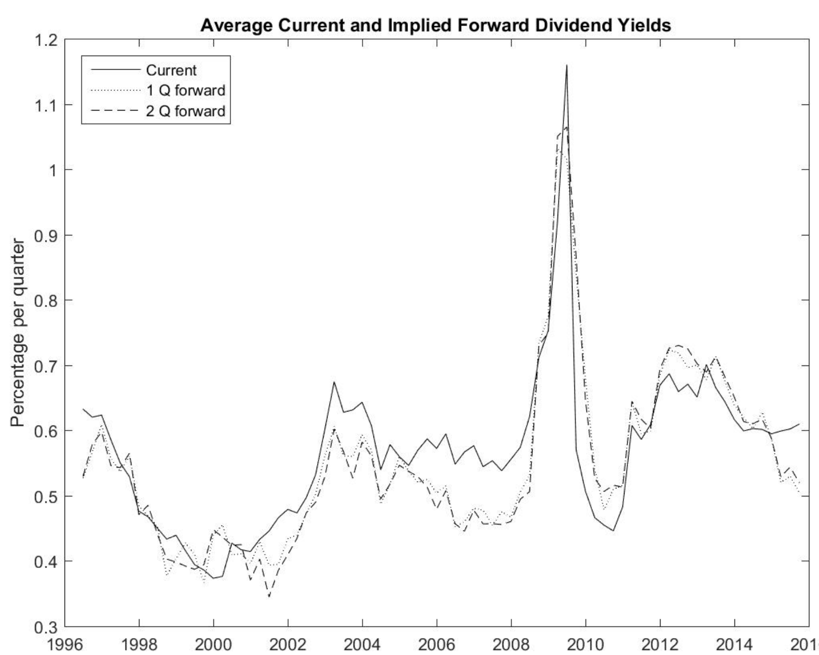


Figure 3.3: **Dividend yields per quarter, current, 1 to 3 quarters ahead and 4 to 8 quarters ahead.** Yields are value-weighted over all firms who pay dividends at a quarterly rate.

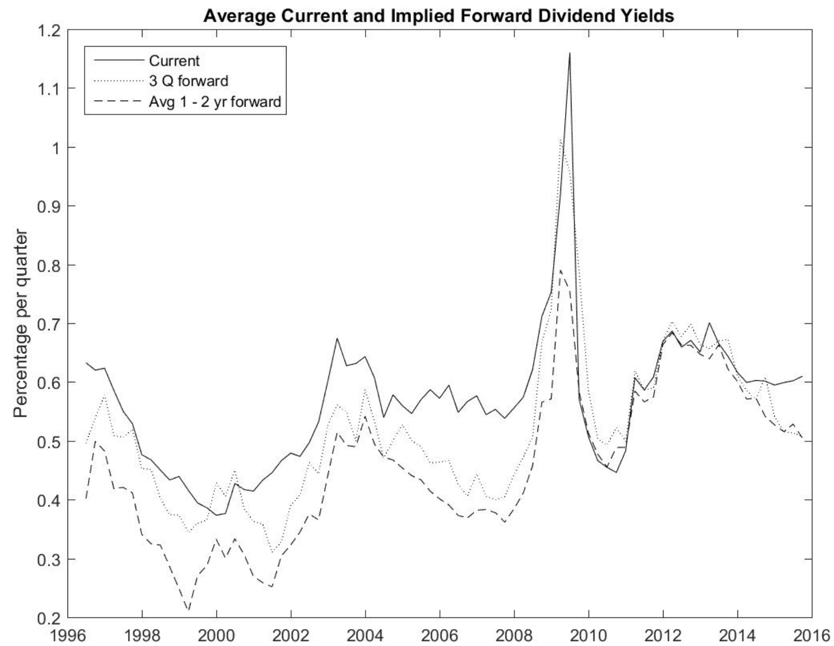


Figure 3.4: **Implied growth rate for dividends 1 quarter ahead and 1 to 2 quarters ahead relative to current dividends.** Growth rates are value-weighted over all firms who pay dividends at a quarterly rate.

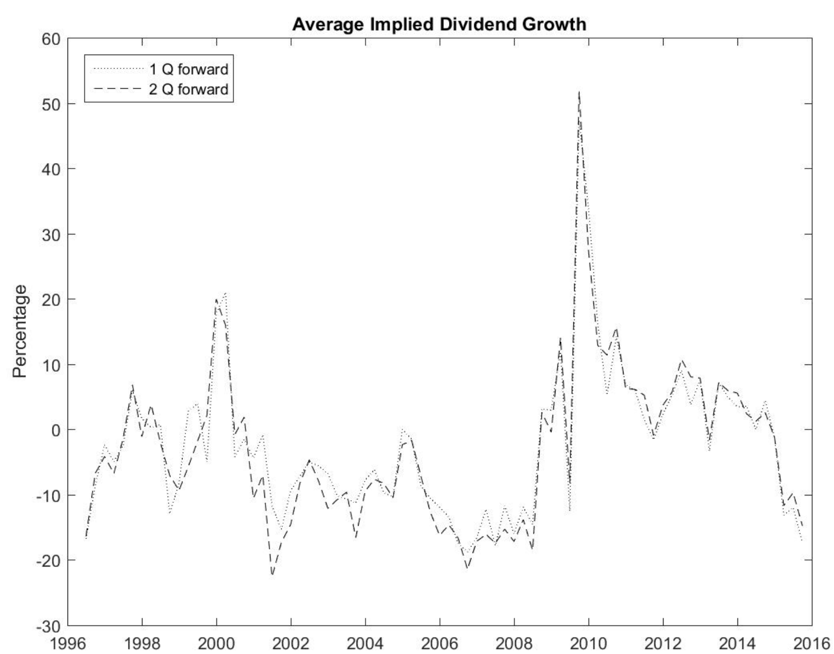


Figure 3.5: **Implied growth rate for dividends 1 to 3 quarters ahead and 4 to 8 quarters ahead relative to current dividends.** Growth rates are value-weighted over all firms who pay dividends at a quarterly rate.

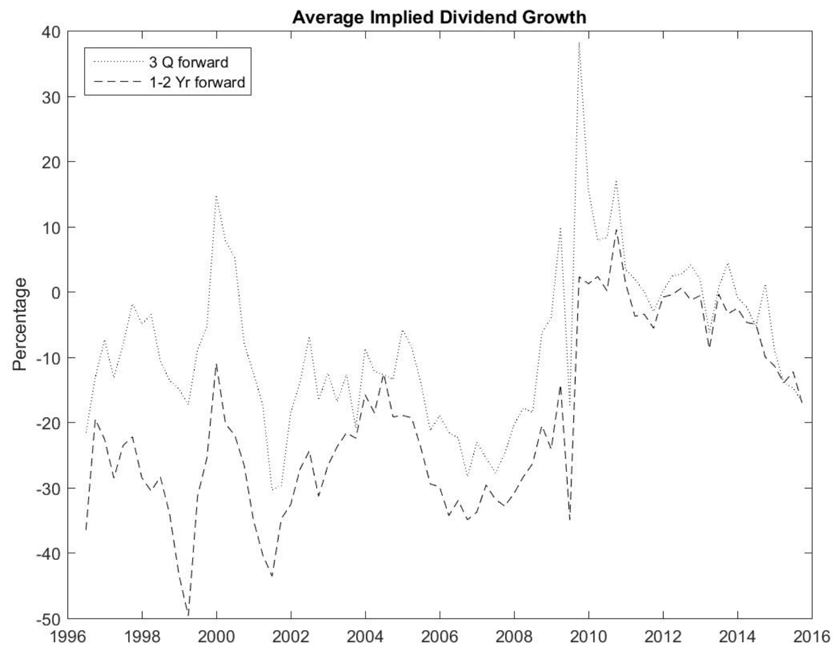


Figure 3.6: **Decile breakpoints of current quarterly dividend yields for all quartercompanies.**

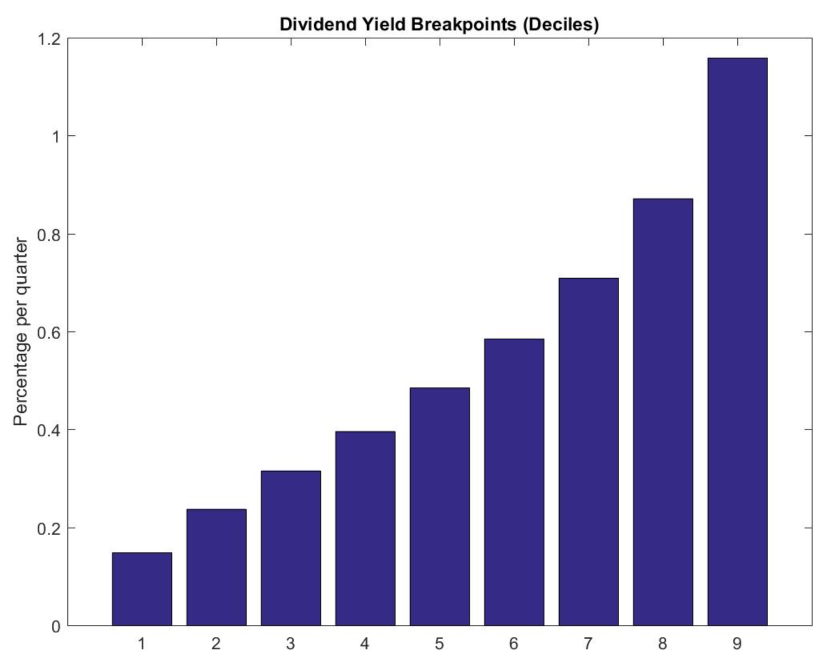


Figure 3.7: **Decile breakpoints of implied dividend growth rates for 1 quarter ahead and 1 to 2 quarters ahead for all quartercompanies.**

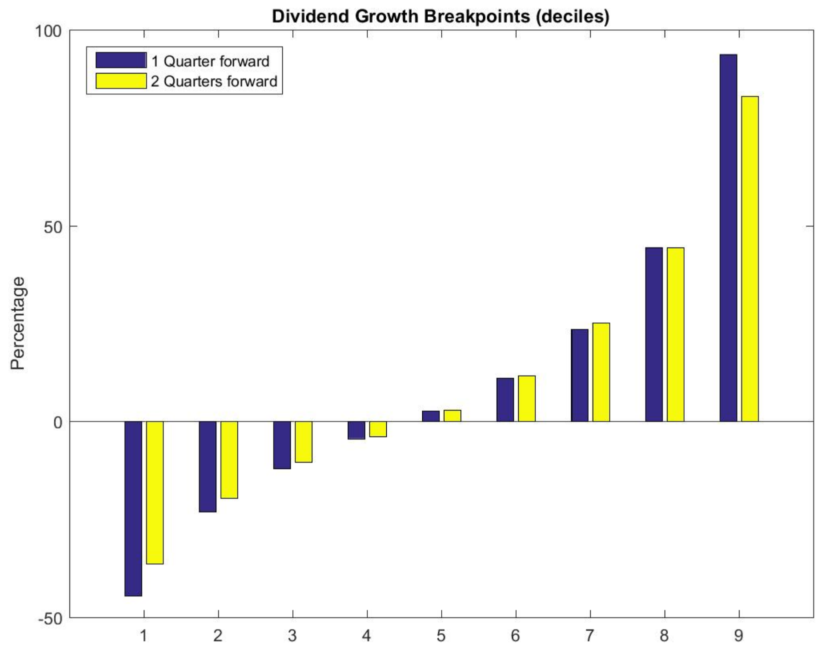


Figure 3.8: **Decile breakpoints of implied dividend growth rates for 1 to 3 quarters ahead and 1 to 8 quarters ahead for all quartercompanies.**

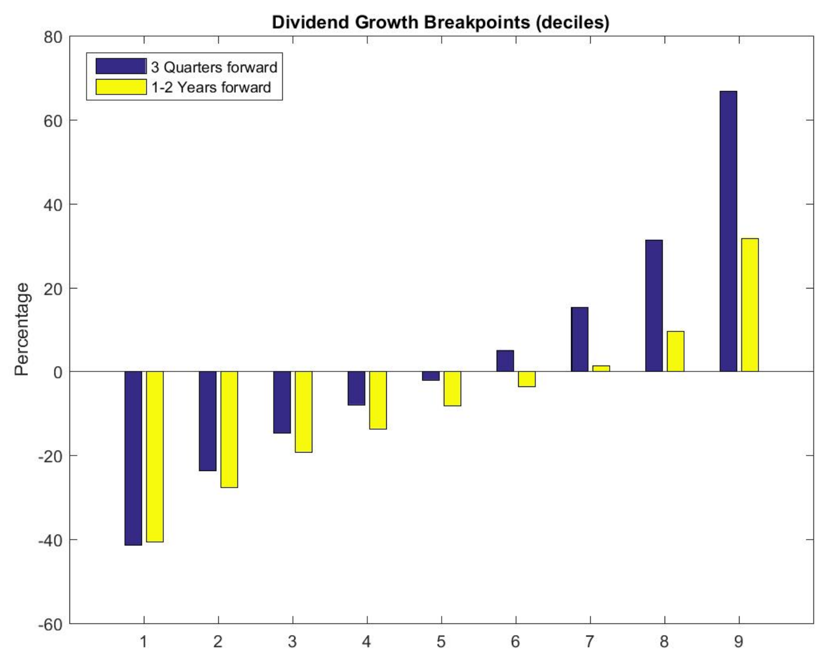


Figure 3.9: **Apple Inc. dividends as implied by the LEAPS expiring in January 2013.** Normalized to a quarterly rate.

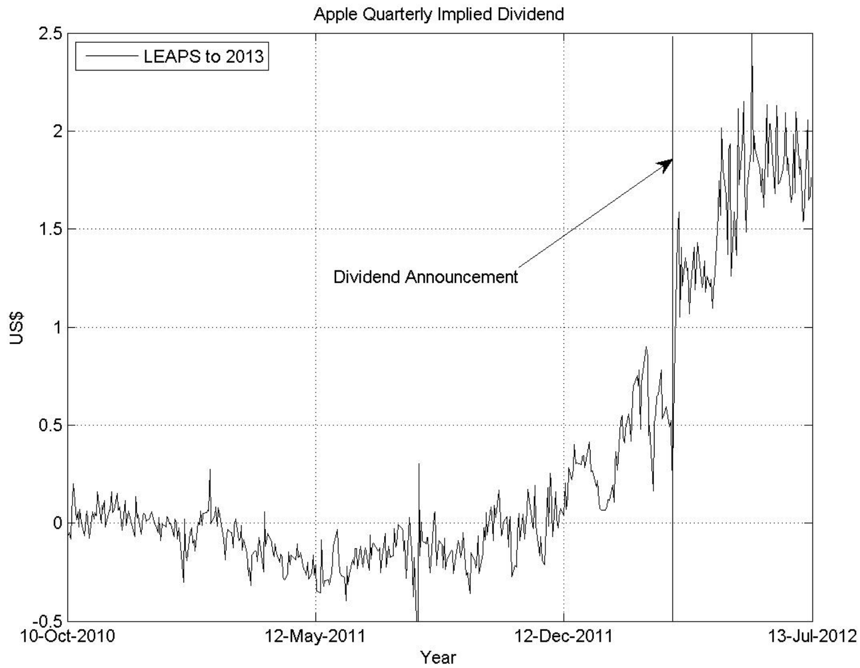


Figure 3.10: **Apple Inc. dividends as implied by the stock option expiring in July 2012.** Normalized to a quarterly rate.

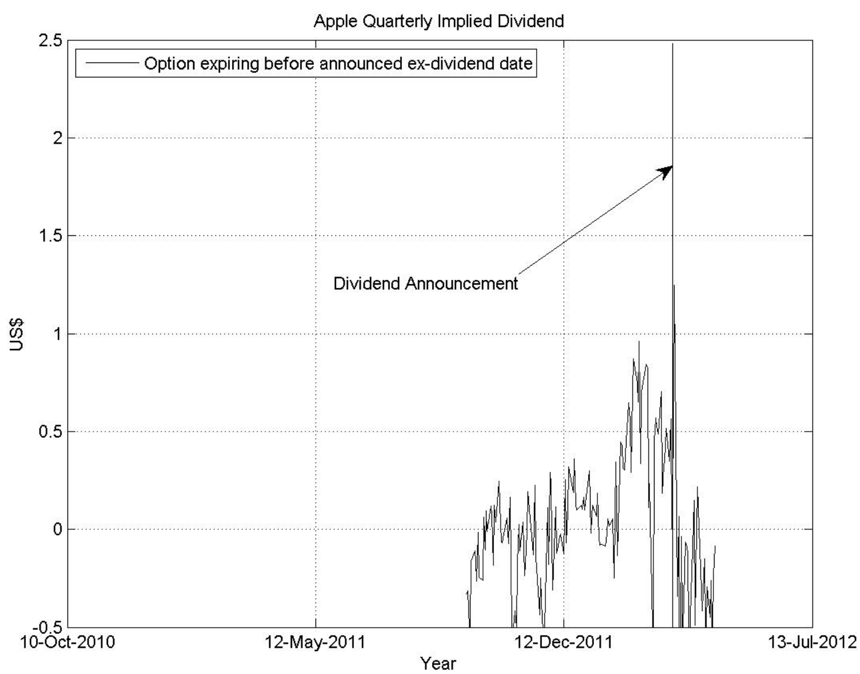


Figure 3.11: **Implied stock loan fees 1 quarter ahead and 1 to 2 quarters ahead.** Normalized to bppa. Fees are value-weighted over all firms who pay dividends at a quarterly rate.

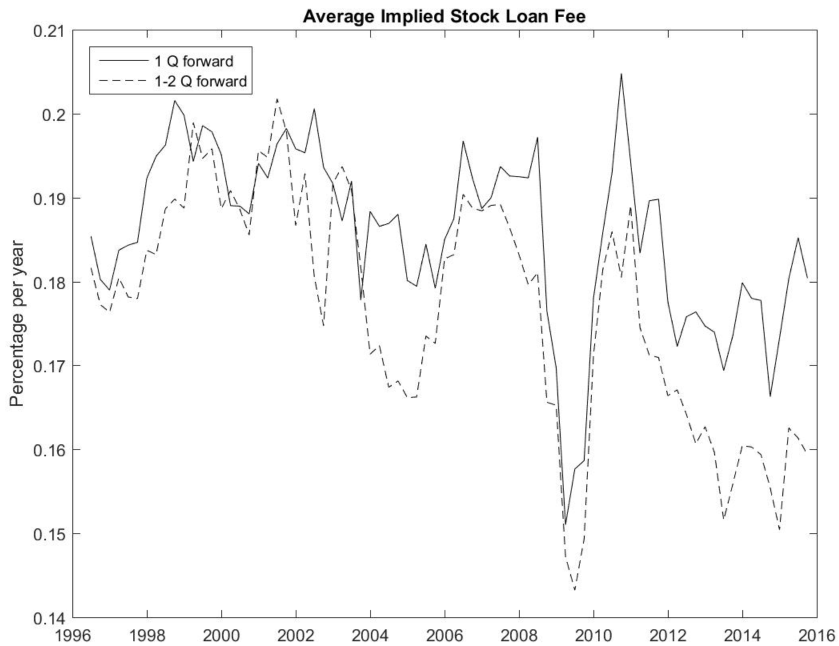


Figure 3.12: **Implied stock loan fees 1 to 3 quarters ahead and 4 to 8 quarters ahea.** Normalized to bppa. Fees are value-weighted over all firms who pay dividends at a quarterly rate.

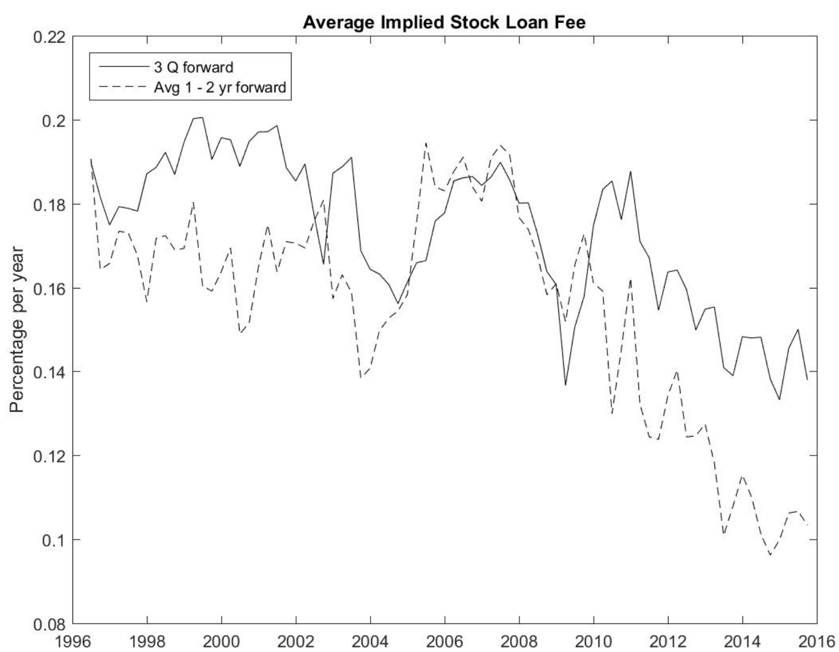


Figure 3.13: **Average implied stock loan fees.** Data are averaged over the 1996-2015 data set for a given number of quarters ahead of the observation quarter for all firms (value-weighted) with and without LEAPS traded on their stocks.

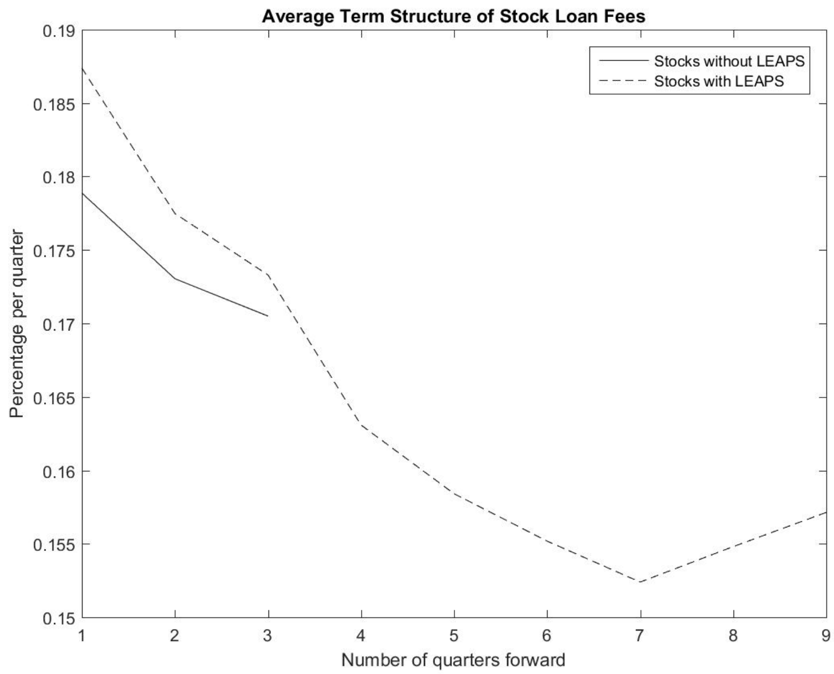


Figure 3.14: **Decile breakpoints of implied stock loan fees.** Data are shown for 1 quarter ahead, 1 to 2 quarters ahead and for 1 to 3 quarters ahead for all quartercompanies.

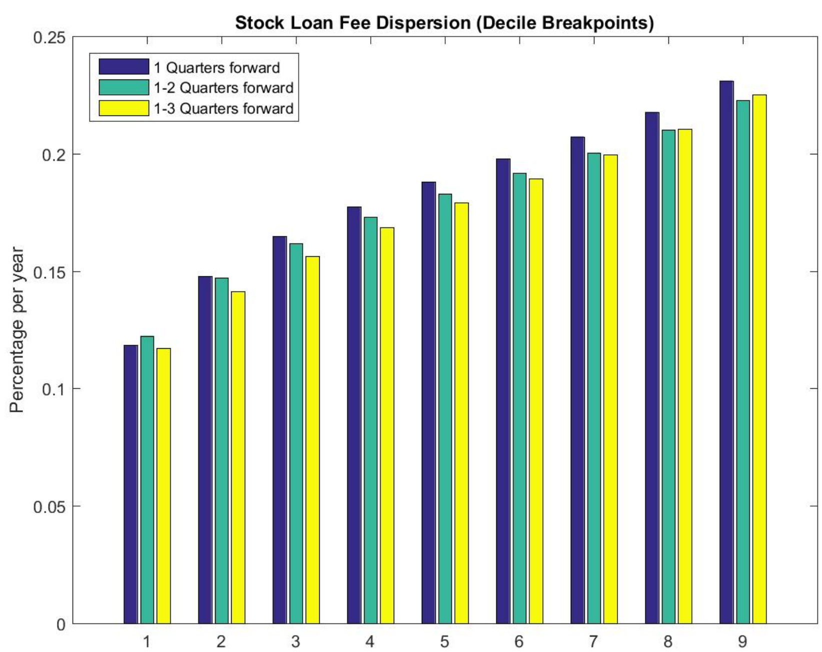


Figure 3.15: **Contingency chart of actual dividends and implied dividends.** Implied dividends refer to 2 quarters ahead. A dot indicates an actual change in dividends against a change as implied by the data $P_{i,t+n}$.

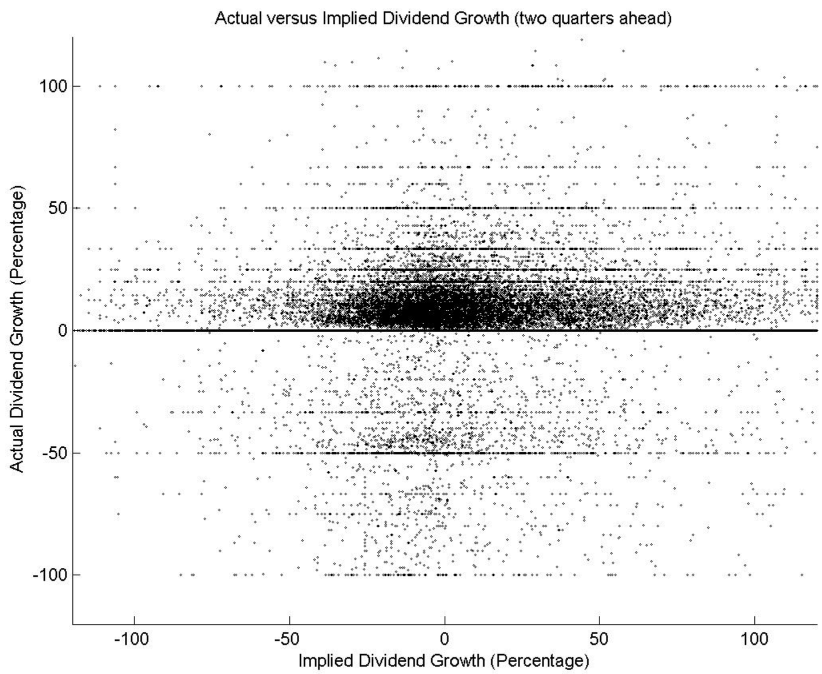


Figure 3.16: **Marginal effects from a change in implied dividends.** Increases in the likelihood of the small dividend raise category $y = 3$, which is defined as $0.01 < d \leq b$, occurring due to a change in implied dividends moving from negative to positive, relative to the probability of a dividend falling into the small dividend increase category. Marginal effects for 1, 2 and 3 quarters ahead.

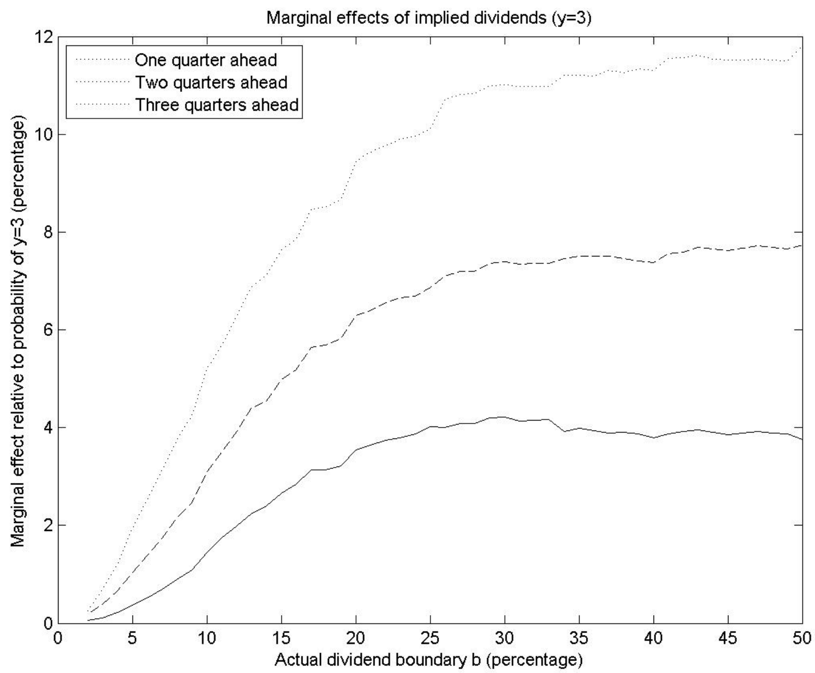
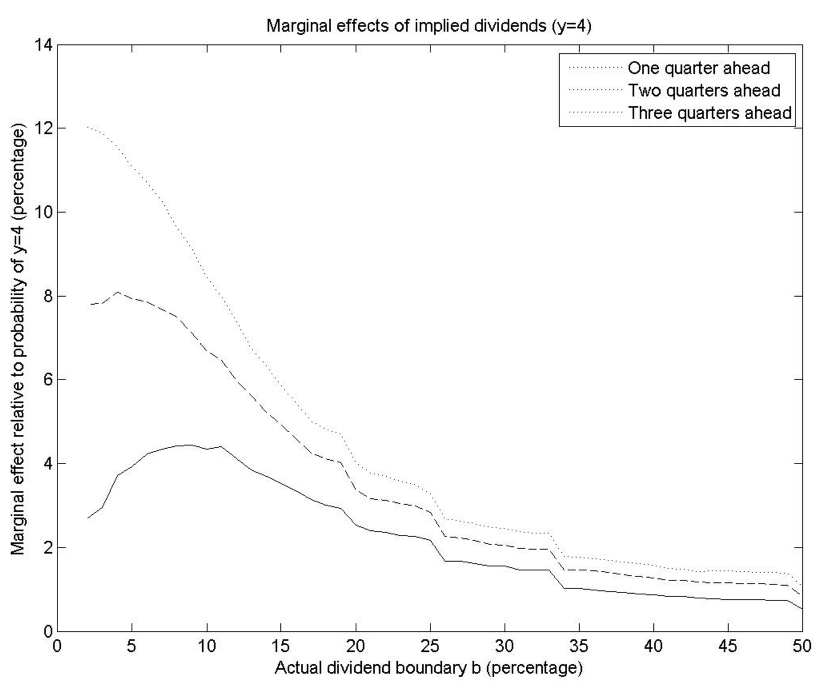


Figure 3.17: **Marginal effects from a change in implied dividends.** Increases in the likelihood of the large dividend raise category $y = 4$, which is defined as $d > b$, occurring due to a change in implied dividends moving from negative to positive, relative to the probability of a dividend falling into the large dividend increase category. Marginal effects for 1, 2 and 3 quarters ahead.



Chapter 4

The Valuation of Future Dividends in Cross-Sectional Models of Stock Returns

The valuation of future dividends as implied in option prices adds to the understanding of cross-sectional stock returns. I consider portfolios sorted by the difference between the valuation of future dividends and actual dividends. If it is fast, such implied dividend growth coincides with high stock returns but is followed by low stock returns, particularly if dividend yields are low. Explaining returns of portfolios sorted by implied dividend growth and accounting variables challenges the value effect and casts light on a possible origin of the profitability effect and the investment effect: a stock's return history associated with implied dividends.

4.1 Introduction

Dividends are a key ingredient to the valuation of stocks and contain predictive power for stock returns. Canonical research shows that returns are predicted at the index level by actual dividends, particularly for longer returns horizons (Campbell and Shiller, 1988, Fama and French, 1988, Cochrane, 2011). More recent research draws attention to the valuations of future dividends, implied from option prices, as a predictor for index returns (Bilson et al., 2015, Golez, 2014)[1]. Actual dividends are also informative about the cross-sectional returns among portfolios of stocks, but authors report a premium of only 0.10% to 0.20% per month for portfolios of stocks paying a high dividend relative to their price over those paying a small one (Fama and French, 1993, Conover et al. 2016). Compared to other well known factors such as *Book-to-Market* this is not substantial. In fact, variation in dividend yields appears to be related to other factors to a point that it is omitted in return predicting models by several authors (for example Fama and French, 2015, Asness et al., 2014)

In view of this research and the perhaps unexpectedly weak performance of dividends in this regard, I investigate the area that has been missed out thus far: do the valuations of future dividends implied from option prices predict returns cross-sectionally? To visualize where this question stands in the literature, I draw the following quadrant:

Return prediction	Actual dividends	Valuation of future dividends, as implied by option prices
Index	Campbell and Shiller (1988) Fama and French (1988) Cochrane (2011)	Golez (2014) Bilson et al. (2015)
Cross-section	Fama and French (1993) Conover et al. (2016) Maio and Santa-Clara (2015)	This paper

In the previous Chapter, I exploit the price data of options traded on US stocks to find implied dividends for individual companies. A Cox, Ross and Rubinstein (1973) binomial tree prices a pair of otherwise equal put and call options and is solved simultaneously for implied dividend and implied volatility, both of which refer to the same stock. The approach results in a data set of term structures of implied dividends for all firms that have options traded on their stocks from 1996 onward. Implied dividends have forecasting power for actual

[1]Other authors focus on implied dividends of the S&P 500 index too, notably Binsbergen, Brandt and Koijen (2012), but they apply these data to other purposes than return predictions.

dividends being paid later. Implied dividends also reflect that, if market participants are wrong about the implied values by not anticipating an actual dividend cut later, the element of surprise bears strongly on the stock price.

In this Chapter, this implied information about future dividend valuation is tested for its relevance to stock returns more generally. The analysis is presented in two parts. First, I investigate the relation between implied dividends and stock returns contemporaneously and sequentially. Second, implied dividends are used to sort portfolios of stocks and to construct a factor which is put to work as an add-on to a CAPM and a five-factor model.

Portfolios of stocks with fast dividend growth, defined as a high valuation of future dividends relative to current dividends, show high returns. In the 12 months *prior* to sorting, the fastest quartile portfolio returns are about 0.90% per month above those of the slowest quartile portfolio. If a stock moves up or down in portfolio quartile rankings, that explains a return of 0.72% in the same quarter.

Future returns, however, paint the opposite picture. In fact, in the year following sorting, portfolios of fast dividend growth stocks underperform those with slow dividend growth by 0.30% per month. Low returns following an expected increase in dividend may seem counter-intuitive, but if the earlier coincidence of high implied dividends and high returns is a stock price overreaction, then low returns afterwards is merely a reversal in the stock price.

The dividend yield premium (documented by among others Fama and French (1993) and Conover et al., (2016)) may be a consequence of the reversal phenomenon. Fast dividend growth causes a stock to rise, which in turn reduces the dividend yield if this rise is larger than the actual dividend increase. An anticipated increase in dividends, when dividend growth is positive, will not necessarily materialize in an actual dividend increase. Low returns following fast dividend growth and a low dividend yield may thus be one and the same: a consequence of a high stock price rather than of a low dividend.

The effect of a change in dividend growth on returns appears to depend on the level of dividend growth. When it is fast or slow, a change in dividend growth down or up matters significantly more to returns than when dividend growth falls into one of two middle quartiles. This is also true for the reversal pattern after portfolios are sorted. For example, when dividend growth is fast, stock prices respond stronger to a decrease in dividend growth because there is a larger potential for a reduction in dividend growth.

Next, I document that future dividends matter to expected returns as measured in a CAPM setting enhanced by dividend yield and dividend growth used to build factor portfolios. The regression intercept of a standard CAPM model decreases once the two factors are added. Moreover, there are two types of interactions between the two dividend factors. Portfolios of stocks with a low dividend yield respond stronger to the returns of the dividend growth factor than high dividend yield portfolios. And portfolios of stocks with slow dividend growth respond stronger to the returns of the dividend yield factor than fast dividend growth portfolios. The interaction indicates that the extent of return effects from dividend growth depends on the actual dividend itself.

Such dependency results exist also when portfolios are controlled for the Fama and French variables. In the latest version of their 1993 model, Fama and French (2015), further referred to as FF, introduce *Operating Profitability* (*OP*) and *Investment* (*Inv*) as variables along which to construct portfolios of companies. In line with the present value accounting identity, they show that portfolios consisting of companies with high *OP* and low *Inv* outperform subsequent to sorting, even if controlled for by *Book-to-Market* (*B/M*) and *Size*.

The authors use current observable values of these variables in their tests to proxy for their expected future values, which are required in present value thinking. Total dividends paid is the difference between *OP* and *Inv* and implied future dividends can thus step into the gap between the theoretically intended expected values and empirically applied actual values of these variables. The implied valuations of future dividends can be found within the horizon of the expiry dates of options, but they remain unobservable for longer horizons. Implied dividends thus cannot close, but can still reduce the gap between the actual data as applied and the expected values intended.

Average returns of portfolio sorted by one of the FF variables show that several of the effects found by FF are in trouble when dividend growth comes into play as a second sorting variable. The value effect is lost entirely on such portfolios, while the relevance of profitability and investment to stock returns depends on the speed of dividend growth. The profitability effect only survives when dividend growth is slow. Earlier results show that returns are poor in the run-up to slow dividend growth, and a stock may have become cheap. A combination of good profitability and slow dividend growth induced cheapness underpins good future returns. Portfolios of high investment companies show lower returns than average (FF). This effect only survives when dividend growth is fast. A stock's relative expensiveness due to fast dividend growth in the past may be a main cause of the investment effect.

I further calculate the alpha's and the loadings of portfolios double-sorted by FF variables and dividend growth on FF factors. Adding a factor formed from the returns of portfolios of fast dividend growth minus slow dividend growth consistently reduces the unexplained abnormal returns of portfolios relative to applying only one or more of the FF factors.

The returns from this *Fast minus Slow* dividend growth factor behave quite different from other FF factors. It correlates positively with the *Market* and the *Size* factor, and negatively with factors based on accounting variables *B/M*, *OP* and *Inv*. The signs of dividend growth factor correlations are the exact opposite of those of the accounting variables. Regressions on other factors confirm that the dividend growth factor is not well explained by them, which justifies its inclusion in a five-factor model.

The dividend growth factor plays a part in the transmission of these company accounting variables into returns. Portfolio returns are more susceptible to dividend growth factor returns when *OP* or *Inv* are high. These sensitivities seem to depend on the sustainability of future dividends as they, in turn, depend on profits and investment.

This paper has three main sections. The subsequent section describes the contemporaneous and forecasting relationship between implied dividends and stock returns. The implied dividend data are then deployed to portfolios sorted by dividends in a CAPM setting enlarged by dividend growth and dividend yield as factors. The closing section investigates the returns of portfolios of companies sorted by the growth in implied dividends in the framework of the Fama and French (2015) five-factor model extended by a dividend growth factor.

4.2 Portfolio returns conditional on dividends

The central item of investigation is the relationship between the valuation of future dividends and expected returns in a cross-section of stocks. Following the introduction to dividend growth as a concept, this section discusses the average portfolio returns in relation to dividend growth, as well as to the combination of dividend growth and dividend yield and closes with cross-sectional regressions of returns on dividend growth. This discussion serves as a prelude to the following two sections, in which dividends are introduced as factors in other models.

The relevance of future dividends to stock prices is evident in the standard dividend discount model (DDM):

$$S_t = \sum_{n=1}^{\infty} PV_t(D_{t+n}), \tag{4.1}$$

which simply equates the value of a stock S_t to the sum of the present value PV_t of all dividends D_{t+n} that it is expected to pay at $t+n$. To assess the level of a future dividend's valuation, I construct dividend growth rate $DG_{t,n}$ which relates its present value to the actual dividend D_t, paid at the moment of observation t:[2]

$$DG_{t,n} = \frac{PV_t(D_{t+n})}{D_t} \tag{4.2}$$

[2]Note that the infinite sum of dividend growth rates equals the price-dividend ratio:

$$\sum_{n=1}^{\infty} DG_{t,n} = \frac{\sum_{n=1}^{\infty} PV_t(D_{t+n})}{D_t} = \frac{S_t}{D_t}.$$

As is standard, the economic interpretation of dividend growth is the rate $g_{t,n}$ at which dividends are objectively expected to grow, discounted by the time value of money $y_{t-1,n}$ and risk-adjusted by risk premium $\theta_{t,n}$, all over the same horizon until $t = n$:

$$DG_{t,n} = \frac{1 + g_{t,n}}{1 + y_{t-1,n} + \theta_{t,n}}. \tag{4.3}$$

In the framework of the DDM, cash returns on a stock R_{t+1} follow from the dividend yield and from price changes, and the latter is linked to actual dividends and the growth rate of future dividends:

$$R_{t+1} = D_{t+1} \sum_{n=1}^{\infty} DG_{t+1,n} - D_t \sum_{n=1}^{\infty} DG_{t,n} + D_{t+1}. \tag{4.4}$$

Consequently, whether dividends are valued to grow fast or slowly does not necessarily influence returns. If growth rates and dividends remain unchanged from one period to the next, cash returns equal the dividend paid and the stock's price remains unchanged. Only if the dividend growth rate or the actual dividend changes can a stock's price change.

4.2.1 Portfolio returns before and after monthly single-sorting

A first step to investigate the claim that dividend growth affects returns is to visualize stock returns as a function of the valuation of future dividends. I use the implied dividends discussed in the second Chapter of this thesis for this purpose. From July 1997 to June 2015, portfolios of stocks are sorted each month by implied dividend growth into quartiles. The horizon for implied growth is 5 to 7 months following the sorting month[3]. Returns are then calculated for the months surrounding the sorting month, from 12 months prior to sorting to 12 months following sorting. Portfolio returns are value-weighted. The data set contains only stocks for which dividend growth rates can be found. This means that stocks on which no options are traded and stocks of companies that pay no dividends are excluded. In the early years of the data period the returns of about 500 companies are available, by the end of the data period this has increased to over 1,000 companies.

In the first DG quartile, future dividends are valued consistently well below actual dividends at -26% and in the fourth quartile they are consistently valued above actual dividends at $+41\%$[4]. Such growth rates are large enough that objective growth $g_{t,n}$ in equation (4.3) should play a part. Very slow DG cannot reasonably be expected to be caused by high discounting values $r_{t,n} + \theta_{t,n}$ alone in the case of first quartile DG and very fast DG cannot

[3]Other horizons do not materially change the estimation outcomes.

[4]Both on a average in the data set and at a horizon of 6 months. See Chapter 2 for further reference.

reasonably be expected to be caused by low discounting values for the high DG quartile. Broadly speaking, dividends are expected to be cut in the first quartile and to be raised in the fourth quartile at an objective measure.

The main message emerging from this exercise is the return reversal at the point of sorting. Returns before sorting are low for slow DG stocks and high for fast DG stocks, which fits in with the DDM as anticipated in equation (4.4). Following sorting, however, the pattern is reversed (Figure 4.1, (a) and (d) respectively).

In the sorting month itself, stocks of which dividends are expected to be cut show average returns, but in the run-up to that point returns are poor. The average monthly return in the 12 preceding months is 0.66%, against 1.10% in the 12 months following sorting. A mirrored pattern is clear in the returns of the fast DG portfolio, showing an average of 1.57% before sorting and 0.80% after sorting. Second and third quartile DG portfolios return closer to average both before and after sorting.

The data shown in Figure 4.1 suggests that the relationship (4.4) between returns and dividends occurs throughout the preceding year, and in particular in the second to fourth month before sorting[5]. In the year-long run-up to sorting, high/low DG stocks outperform/underperform the market average by a cumulative 6.33%/−4.53%. This excellent/dismal performance of high/low DG stocks before sorting suggests that such stocks may have become expensive/cheap by some measure by the time of sorting. This raises the question whether the reversal pattern is a correction of irrational overshooting in the price of the stock, of which a discussion follows below.

4.2.2 Portfolio returns before and after monthly double-sorting

The dividend yield DY can be regarded as a measure of the expensiveness of a stock. Portfolios double-sorted on dividend growth and dividend yield shed some light on their interaction as relevant to returns and thus on reversal patterns. Figure 4.2 shows value weighted monthly returns of the first and fourth quartiles of DG/DY sorted portfolios. The low DY and fast DG portfolio returns in the months preceding sorting are substantial (1.81%) and they fall to below average following sorting (Panel (b)). Such return reversal does not appear for a high DY/fast DG portfolio (Panel (d)). If anything, this portfolio has somewhat higher returns following sorting.

The high DY/slow DG portfolio in Panel (c) shows a similar but opposite reversal pat-

[5]The size of a dividend payment is often announced in the month preceding it. If an announcement is made to change dividends, then the stock has normally undergone the returns associated with the change in dividends in the month immediately before sorting. During second to fourth month before sorting, first quartile DG stocks return as little as 0.35% per month, whereas fourth quartile DG stocks return 1.69%.

tern. Before sorting it performs poorly at only 0.20% per month and it reverses to a slightly above average return in the period following sorting. Again this contrasts starkly to slow DG portfolios that start from a low dividend yield (Panel (a)). Their returns are above 1.50% in months 12 to 6 preceding sorting, then drop to less than 0.50% in the remaining months up to sorting to bounce back to slightly above average in the period following sorting.

The return patterns found in the first and fourth quartile portfolios sorted by DG only in the previous subsection are sharpened when they are sorted by DY as well. If dividends are expected to fall/rise, this negatively/positively influences returns contemporaneously, but much more so if dividend yield is high/low. The level of the dividend yield thus clearly matters to returns associated with fast or slow dividend growth.

A possible interpretation is that the potential for gains or losses in dividends influences returns. For example, a portfolio of high DY has more return from dividend to lose than a low DY portfolio. When market expectation is for dividends to fall, then returns are lower in the run-up to sorting of stocks with high dividend yields that have more to lose than of stocks with low dividend yields.

But this mechanism may be blurred: high DG stocks that are expensive, as measured by a low dividend yield, show strong returns in the run-up to sorting which cause them to stay expensive and remain or end up in the low DY quartile[6]. In such instances, actual dividends may have risen less, if at all, than stock prices gained. If that happens, their subsequent returns are low. The opposite is true for high DG stocks that are cheap as measured by a high dividend yield, and mutatis mutandis for low DG portfolios. Changes in dividend growth may cause stock prices to overshoot due to such prior returns. The tendency towards fair pricing following such a move is the reversal found.

Portfolios sorted by dividend yield alone produce future excess returns. (Fama and French 1993, Conover et al., 2016). Along the line of reasoning above, it may be not the dividend itself that affects subsequent returns, but irrational prior returns associated with dividend growth being at least partly reversed. If a stock rises in conjunction with fast dividend growth, the dividend yield falls if this rise is larger than the actual dividend increase. A low return following a low dividend yield might thus be nothing other than a reflection of a stock's expensiveness. This is not to rule out a rational explanation for the phenomenon, although that requires large moves in risk premiums.

[6] In the months preceding sorting these stocks did not necessarily fall in the low DY quartile.

4.2.3 Cross-sectional regressions

When investors change the valuation of expected dividends, the stock price as a present value of future dividends should change too if markets are efficient (4.4). As an introduction to the factor analysis in subsequent sections, I investigate empirically the connection between dividend growth and returns as well as its predictive power for future returns in a cross-sectional setting.

Stocks are sorted each quarter into 4 portfolios by the growth rate from actual dividends to implied dividends at a horizon of 6 months, as before. I refrain from using dividend growth rates per stock as regressors because the implied dividend data are noisy. The indicator for dividend growth change $I_{\Delta DG_{j,t}}$ is defined as a change in dividend growth quartile of stock j. Its value equals 1 if $DG_{j,t}$ increases in its quartile ranking from quarter $t-1$ to quarter t, it is -1 if the ranking decreases and is zero otherwise. The equation to test is as follows:

$$R_{j,t+i} = \alpha_t + \beta_t I_{\Delta DG_{j,t}} + \epsilon_{j,t+i}, \tag{4.5}$$

where $R_{j,t+i}$ are quarterly returns of stock j in quarter $t+i$ in excess of the data set average. This regression is run for each quarter in the data set from 1996 to 2015. The β_t and its t-statistic are calculated following the Fama-Macbeth method.

The contemporaneous return ($i = 0$) of an individual stock explained by a change in $DG_{j,t}$ quartile amounts to 0.72% during a quarter (Table 4.1, model (5a)) and this coefficient is highly significant. This result confirms that stock prices respond to implied dividends in line with the dividend discount model (4.4).[7]

That can not be said for the power of a stock return prediction by the same variable. Dividend growth has a small impact on stock returns two quarters ahead ($i = 2$) with a coefficient not significantly different from zero (Table 4.1, model (5b)).

I continue by differentiating the sensitivities of returns to up and down dividend growth. Work with this data set in Chapter 2 suggests that downward and upward changes in DG produce different effects on stock returns, notably as a response to a dividend announcement. When companies announce a dividend cut, on average stock prices fall by more than 2%. However, if a cut is anticipated as measured by implied dividends, the stock does not respond by much at all. It is useful to investigate whether returns respond symmetrically to dividend changes or not.

To distinguish between up and down moves, I introduce dummies which capture upward, downward and zero changes in the DG quartile: $I_{\Delta DG_{j,t}>0} = 1$ if the DG quartile indicator

[7]At 1.3% the R^2 seems low, but this is quite reasonable given the large dispersion in individual stock returns and the measurement error in the regressor.

increases from quarters $t-1$ to t, $I_{\Delta DG_{j,t}<0} = 1$ if it decreases and $I_{\Delta DG_{j,t}=0} = 0$ if remains unchanged[8]. The following model reflects this distinction:

$$R_{j,t+i} = \beta_{1,t} I_{\Delta DG_{j,t}>0} + \beta_{2,t} I_{\Delta DG_{j,t}<0} + \beta_{3,t} I_{\Delta DG_{j,t}=0} + \epsilon_{j,t}. \tag{4.6}$$

The regression results of this model indicate a similar relationship to stock returns regardless of whether dividends are implied to increase or to decrease. Both sensitivities are 0.72%, highly significant and very close to the coefficient in model (4.5a) that does not make the distinction (Table 4.1). The quarters in which the dividends growth quartile does not change constitute about 60% of the data set. The dummy for these instances attracts a coefficient that is close to zero, which is as expected.

This symmetric connection between future dividend valuations and stock prices can be reconciled with the result in Chapter 2 that stocks hardly respond to the announcement of a change in dividends if they are predicted by implied dividends but react strongly to an unanticipated cut and much less so to a raise. If a dividend raise/cut is correctly anticipated and the stock price has turned upwards/downwards correspondingly, the shock of the dividend raise/cut once announced has by that time abated since the raise/cut is already reflected in the stock price[9].

Lastly, I combine levels and changes in implied dividends as explanatory variables of stock returns to test whether the impact of a change in dividend growth depends on the level of dividend growth. The following model serves this purpose:

$$\begin{aligned} R_{j,t+i} = \alpha_t + \beta_{1,t} I_{\Delta DG_{j,t}} \times I_{DG=1} + \beta_{2,t} I_{\Delta DG_{j,t}} \times I_{DG=2} + \\ \beta_{3,t} I_{\Delta DG_{j,t}} \times I_{DG=3} + \beta_{4,t} I_{\Delta DG_{j,t}} \times I_{DG=4} + \epsilon_{j,t}. \end{aligned} \tag{4.7}$$

$I_{\Delta DG_{j,t}}$ is the quarterly change in the DG of stock j from $t-1$ to t, as before. $I_{DG=q}$ takes on the value 1 if DG falls in quartile q.

The results in Table 4.1 (model (7a)) show the differentiation in the sensitivities to the different levels of dividend growth as measured by quartiles for contemporaneous stock returns ($i = 0$). If it is either low or high (quartiles 1 or 4), the price response to a change in dividend growth is just under 1.00. If dividend growth is more muted (quartiles 2 and 3), the response is on average 0.40 smaller.

The interpretation of these results is as follows. Dividend growth in quartile 1 tends to be negative and in quartile 4 it tends to be positive. If dividend growth moves up

[8] Increases and decreases in dividend growth each account for 20% of a total of 46,564 dividend growth quarters in the data set.

[9] Practitioners capture this phenomenon by the phrase: "Buy the rumor, sell the fact."

($\Delta DG_{j,t} > 0$) from a decreasing path ($DG_{j,t} = 1$), at least its decrease lessens, and if it moves down ($\Delta DG_{j,t} < 0$) from an increasing path ($DG_{j,t} = 4$), at least its increase lessens[10]. If either happens, stock prices react more strongly than when dividend growth is closer to unchanged in quartiles 2 and 3. Consequently, stock prices respond more sharply to a change in dividend growth when there is high dividend growth to be lost or low dividend growth to be increased[11].

To connect with the earlier findings about reversal, the same model is tested for its predictive power ($i = 2$). This regression shows the opposite relationship suggested by reversal, but it is weaker (Table 4.1, model (7b)). The sensitivity of stock returns in quartiles 1 and 4 have the expected negative sign and returns are reversed by about 0.15% per quartile indicator over two quarters. In comparison to the contemporaneous regression (model (4.7a), $i = 0$), this sensitivity is a relevant proportion of the difference between their coefficients (0.94 and 0.97) and those of quartiles 2 and 3 (0.48 and 0.62). A meaningful degree of reversal thus appears to exist, although the coefficients of the first and fourth quartile regressors are not significant[12].

An et al. (2014) report that portfolios sorted by implied volatilities produce excess returns. They find that stocks with call/put options that have experienced increases in implied volatilities tend to have high/low future returns, which is attributed to informed traders. I suggest a different explanation for the phenomenon. These authors apply implied volatility data provided by OptionMetrics, which are calculated under the assumption that future dividends as inputs to the option pricing model are fixed and equal to actual dividends. As shown in the previous chapter and by other authors (for example Binsbergen et al., 2013), implied dividends often differ drastically from actual dividends. All other things equal, a price increase in a call option may be caused by an increase in implied volatility when implied dividends are fixed, or by a decrease in implied dividends when implied volatility is fixed. The results in this section are that future returns are high when implied dividends are low, which is therefore close to the results of An et al. (2014), albeit labeled differently[13]. But since they fix implied dividends, I argue with their interpretation that the return relationship to option prices is a consequence of a preference among informed investors to trade in option markets first, which then leads stock returns[14]. I contend that no inefficiency between markets is required for a predictive capacity of options: my results show that the relationship may

[10]Note that $\Delta DG_{j,t}$ cannot be positive in quartile 4 and cannot be negative in quartile 1.

[11]A similar response is found when dividend yield replaces dividend growth as an interaction term in each of the regressors.

[12]Regressions including lagged returns as explanatory variables do not find cross-sectional relevance (not shown here).

[13]The authors strictly apply changes in implied volatility as a return predictor, not levels.

[14]Cremers and Weinbaum (2010) make a similar claim.

run past implied dividends. Stock options contain information about dividends which drive returns, and are not necessarily a channel for returns themselves.

4.3 The CAPM extended by dividends

Dividends play a minor part in the literature using factors for understanding cross sectional stock returns. Well known factors, such as those discussed by Fama and French (2015, further referred to as FF), consist of company accounting fundamentals known to influence stock prices. It seems obvious to construct a factor from dividends as well, if only as a transmission of such fundamentals into cash. Indeed, average stock returns of firms with high dividend yields are somewhat higher than those with low dividend yields (Fama and French, 1993, Keim, 1988). High dividend stocks do even better in down markets (Fuller and Goldstein, 2011). But since the cash payoff from dividends is a substantial part in the evaluation of share prices, the question is whether the risk premium associated with dividends shouldn't be more influential to stock returns. I will pursue this question first by investigating a Sharpe-Lintner CAPM model expanded with actual and implied dividends. In section 4 a similar exercise is performed for the Fama and French five factor model, to which I add implied dividends as a factor.

4.3.1 The present value of stocks derived from dividend yields and dividend growth

A rational motivation for dividends as priced risk factors starts with the present value relationship in (4.1). In this identity, share price S_t at time t equals the sum of the present values of dividends $PV_t(D_{t+n})$. Part of the present values are observable from implied dividends, which are those up to and including $t+n$. Dissecting the identity into observable and non-observable present values then renders:

$$S_t = \sum_{i=1}^{n} PV_t(D_{t+i}) + \sum_{i=n+1}^{\infty} PV_t(D_{t+i}). \tag{4.8}$$

Observable dividends are depicted as their implied growth rate from D_t to $PV_t(D_{t+n})$ while unobservable dividends beyond n are approached by their expected value $E_t(D_{t+n})$ discounted at r:

$$S_t = D_t \sum_{i=1}^{n} \left(\frac{PV_t(D_{t+i})}{D_t} \right)^i + \sum_{i=n+1}^{\infty} \frac{E_t(D_{t+i})}{(1+r)^i}. \tag{4.9}$$

This depiction of the present value relationship motivates to consider the growth in the implied valuation of dividends $PV_t(D_{t+n})/D_t$ and dividend yields D_t/S_t as systematic risk factors for stock returns in addition to the market factor. The CAPM attributes returns to non-diversifiable exposure to the market and this decomposition does not preclude the validity of CAPM. As to the expected effects from risk factors, Fama and French (1993, 2015) make the case for what happens to returns if all variables are fixed except a particular variable and returns. Applied to (4.9) this ceteris paribus reasoning provides the following return relationships. If everything in (4.9) is fixed except dividend yields and the expected stock return r, then higher dividend yields imply higher expected returns. The dividend growth path can change shape without necessarily affecting returns at all. Faster dividend growth upto $i = n$ then merely implies slower dividend growth beyond $i = n$. However, a change in short term dividend growth is an anticipation of a higher dividend level in the near future that may leave the growth path beyond $i = n$ unchanged. In that case faster dividend growth upto $i = n$ implies higher expected returns. The test results discussed in the second section indicate a clear relationship between returns and short term dividend growth, but in the opposite direction as returns are below average following fast dividend growth. Tests in the context of a CAPM confirm this conclusion and provides some interpretation.

4.3.2 Dividend yield and dividend growth added to the CAPM

An additional motivation to extend the CAPM by dividend growth and dividend yield as factors stems from their interaction with stock returns. The earlier results show that portfolios sorted by dividend growth vary in returns. An irrational interpretation for the interaction is that dividend growth stands in for overshooting returns which are subsequently reversed. Dividend yield also provides a basis for adding a factor to the CAPM, as the impact of dividend growth on returns depends on dividend yields.

The analysis pursued here closely follows the procedure in FF. They analyze the average returns of portfolios and test the sensitivity of portfolio returns to risk factors. The model tested is the Sharpe-Lintner CAPM, expanded by the two dividend risk factors:

$$R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + p_j PMF_t + f_j FMS_t + \epsilon_{j,t}, \tag{4.10}$$

where $R_{j,t}$ is the return on (a portfolio of) stocks j for period t, $R_{F,t}$ is the risk-free return, $R_{M,t}$ is the return on the market portfolio and $\epsilon_{j,t}$ is a residual with zero mean. The two dividend related factors are *Prodigal minus Frugal* (*PMF*) and *Fast minus Slow* (*FMS*). *PMF* is the return of a portfolio of high dividend yield stocks minus the return of a portfolio of low dividend yield stocks. *FMS* is the return of a portfolio of companies with fast growing dividends minus the return of a portfolio of companies with slow growing dividends. The definition of growth in dividends is given in (4.2), making use of the implied dividends found in Chapter 2. Further discussion of these factors follows in a subsection below.

The empirical tests investigate the explanatory power of the model in two parts. The first is a check on the average returns of quartiles of portfolios, sorted along the dimensions above. The second part of the analysis focuses on the time series characteristics of the factors.

4.3.3 Average portfolio returns

Portfolios for calculating returns are constructed by sorting the companies into quartiles for both dividend growth (DG) and dividend yield (DY) in the second quarter of a calendar year[15]. Dividend growth is defined as the dividends implied for the two quarters following the sorting quarter relative to the actual dividend in the sorting quarter[16]. Returns are then calculated for the subsequent year starting at the third quarter, which is the quarter following the sorting quarter, up to and including the second quarter of the next calendar year. As before, the data set contains only stocks for which dividend growth rates can be found, which excludes stocks without options and companies that do not pay dividends.

Table 4.2 demonstrates the average monthly value weighted returns of these portfolios sorted on implied dividend growth and dividend yield in excess of the one-month Treasury bill rate. First of all, the higher return for higher dividend yield companies documented in the literature (Fama and French, 1993) clearly emerges. The difference in average return between the highest and the lowest DY quartile is 0.25% per month, frugal companies returning less than companies in the other DY portfolios. At the same time, fast DG firms return less than slow DG firms. This difference is of a similar magnitude at – 0.30% per month.

Portfolios double-sorted on DG and DY may provide insight into the interaction between these dividend variables, but patterns of excess returns are not consistent within DY portfolios and DG portfolios. Among frugal companies, the effect of higher dividend growth is substantial, but in the other groups the effect is not clear. Likewise is the pattern of increasing returns as DY rises not consistent within DG groups. The picture emerging from Table 4.2 is that the sensitivity of average stock returns to dividend growth highly depends on the level of dividends. When the dividend yield is high, there is less potential for it to rise than when it is low. Fast DG portfolios anticipate a dividend increase and their returns

[15] The second quarter of the calendar year is used as the moment to sort portfolios in both this extended CAPM and the FF five factor model discussed in the following section. FF consistently use the second quarter as the moment of sorting and I benchmark the two extended models tested in this paper against theirs.

[16] The data set allows for sorting on the first, the first two and the first three quarters following the sorting quarter. Although the results for returns differ, the general conclusions remain the same. Using LEAPS for this purpose would extend the horizon out to two years, but reduce the size of the data set more drastically.

are high before sorting and show a low return reversal afterwards.

This interpretation matches the data while the ceteris paribus reasoning followed in the subsection 3.1 does not. If all variables are fixed except for DG, DY and expected stock returns, then if both DG and DY are small, expected returns should be small as well. I find the opposite, with the slow and frugal quartile returning the best of all 16 portfolio. This result suggests that such reasoning is not appropriate for the purpose of relating expected returns to other elements of present value identity (4.9). More specifically, the data show that short term dividend growth and the stock price move together. Assuming that the stock price remains fixed within (4.9) while dividend growth is varied will dilute a conclusion for expected returns. In fact, if it is correct to characterize the stock price response to an upward change in dividend growth as irrational overshooting, then future expected returns will drop, which agrees with the data. It appears that ceteris paribus is not a reasonable imposition for anticipating conclusions about expected returns from the present value identity.

4.3.4 Factor definitions

The next step is to construct factors for dividend yield and dividend growth and let them perform as regressors of portfolio returns. Next to market return $R_M - R_F$, the two factors capturing the defining inputs to the share price in equation (4.10) are *Fast minus Slow (FMS)* for dividend growth and *Prodigal minus Frugal (PMF)* for dividend yield. Similar to constructing portfolios, for calculating the returns of *FMS* I apply dividend growth for two quarters ahead of the sorting quarter.[17] The DG sorting variable is thus the growth rate from the dividend paid in the second quarter to the option implied dividend for the fourth quarter[18], calculated for each company. The DY sorting variable is the dividend paid in the second quarter of the year, divided by the average share price in the second quarter.

Factor portfolio returns are constructed at each second quarter of the calendar year by forming three groups of firms demarcated at the 30^{th} and the 70^{th} percentile of the sorting variable. DG factor *FMS* is then defined as the returns of the fast DG group minus the returns of the slow DG group. Similarly, DY factor *PMF* is defined as the returns of the prodigal DY group minus the returns of the frugal DY group. Returns are expressed at monthly rates. Annual rebalancing of portfolios in the second quarter is less frequent than feasible since quarterly data are available, but it is chosen to stay close to the methodology

[17]The choice to use dividends two or three quarters ahead does not bear materially on the results. For limiting space, the results of three quarters ahead dividend growth is omitted here.

[18]For recollection, quarterly periods for the purpose of sorting are defined as the first trading day of the third Friday of the last month of the calendar quarter up to and including the third Friday of the last month of the next calendar quarter. This definition shifts from a calendar quarter by on average just under two weeks. The shift serves to correspond to the expiry schedule of the options used to find implied dividends.

applied by FF.

4.3.5 Factor summary statistics

Table 4.3 contains the statistics of factor returns. The *FMS* portfolio returns are negative on average, which matches the earlier finding that fast *DG* firms return measurably less than slow growing ones. Slow dividend growth companies return 0.40% per month more than fast *DG* companies, with a standard deviation of 2.81. At a t-stat close to 2 this mean return is just about significantly different from zero at the 5% confidence level.

The average return of *PMF* is 0.22%, which is somewhat higher than 0.12% found over the FF data period of 1963 to 2013. There are several reasons for this difference apart from the length of the data period. FF calculate the returns for the year starting at the third calendar quarter for portfolios of dividend yields which are sorted at the end of the previous calendar year, instead of in the second quarter. FF returns thus refer to a period that is further removed from the moment of sorting than is applied here. Moreover, my data set contains fewer companies (see Table 1 in the previous Chapter). The standard deviation of *PMF* is 3.90, which is too volatile for its returns to be different from zero at reasonable significance levels.

The correlation between the two dividend factors is strongly negative and significant at -0.42. They also show opposite correlation with the market factor, *FMS* correlates positively with market excess returns, but the *PMF* factor correlates negatively (Table 4.3). The p-values of these correlations are less than 0.01 (not shown here). Regression of one factor on the others provides further insight in their interaction (Table 4.4). Returns of the *PMF* factor show substantial and significant negative slopes on both market excess returns and *FMS* returns.

The dividend growth factor does not load on market excess returns once PMF is included in the regression. The correlation matrix in Panel B of Table 4.3 reveals some correlation between *FMS* and the market, but in regressions (Table 4.4) this turns out to be superseded by the strong relation between *PMF* and market returns. This remarkable difference among *FMS* and *PMF* in exposure to the market corroborates the impression that portfolios sorted by dividend growth and sorted by dividend yield are different in nature, which furthers their relevance as part of an extended CAPM.

4.3.6 Portfolio regression intercepts

Portfolios are constructed by double-sorting along *DG* and *DY* and forming quartiles out of the data set at each second quarter of the calendar year, similar to the construction of the factors. The 4×4 portfolio returns are regressed on the three factors as described in the dividend-extended CAPM (equation 4.10).

We are interested in the ability of the model to explain average returns of the 4×4 portfolios. For this, the alphas of the portfolio regression need to be statistically zero individually and jointly. The top Panel in Table 4.6 describes the alphas of all 4×4 sorted portfolios. About half of them attain an absolute value of 0.20 or more. Nonetheless, none of them is significant at the 10% level. The average absolute alphas of these portfolio regressions is 0.18 (Table 4.5) against 0.29 in the base CAPM regressions. The inclusion of each dividend factor individually materially improves the performance of the CAPM too judged by their alphas.

In order to establish the validity of the model in explaining excess return, the intercepts in (4.15) should be jointly zero across portfolios. Table 4.5 shows the GRS-statistic of Gibbons, Ross and Shanken (1989) that tests the zero-intercept hypothesis for each set of 16 portfolios and factors. The joint-zero hypothesis is tested by means of the GRS-statistic and its p-value. The GRS-statistic improves a lot from the inclusion of either dividend growth or dividend yield as regressors to test model (4.10). Its p-value triples, primarily due to the addition of *FMS*. Overall, the model does a good job in explaining portfolios sorted by both dividend variables.

The average value of the intercept also decreases relative to the difference of the average return on portfolio j and the average of all portfolio returns when *FMS* and *PMF* are added as factors[19]. This is true both when they are considered as absolute values and as squared values, corrected for sampling error. Although both measures decrease a lot, they remain high. The reason is that the deviation in average returns of individual portfolios is small at 0.12%. With absolute alphas at 0.18%, their ratio is about 1.5 and the ratio of their squares is barely less than one. But, as FF point out as well, I am primarily interested in whether these statistics point to an improvement relative to the CAPM, which they do.

[19]FF assess the extent to which returns are left unexplained by the competing models. They divide the average (of 16 portfolios) absolute estimated intercept $A(a_j)$ by the average absolute deviation $A(\bar{r_j})$ of the time-series average return of each portfolio R_j from their cross-sectional average $A(\bar{R}_j)$. This measure shows the proportion of excess portfolio returns that is left unexplained by a given model. Measurement error causes these estimates to be exaggerated, which can be adjusted for by focusing on squared intercepts and squared errors. Since α_j is a constant, the expected value of the square of its estimate is the squared value of the true intercept plus the sampling variance of the estimate, $E(a_j^2) = \alpha_j^2 + E(e_j^2)$. The estimate $\hat{\alpha_j}^2$ of the square of the true intercept, α_j^2, is the difference of the intercept and its standard error. Similarly, defining the estimate of realized deviation of returns of portfolio j as $\bar{r_j} = \bar{R}_j - A(\bar{R}_j)$, the estimate of $\bar{r_j}^2$ is the difference between its square and the square of its standard error. The ratio of averages $A(\hat{\alpha_j}^2)/A(\hat{\bar{r}}_j^2)$ then reflects the proportion of the variance of *LHS* returns that is not explained by the model.

4.3.7 Portfolio regression slopes

Table 4.6 further contains the slopes of $R_M - R_F$ and the two dividend factors. The slope to the *Market* is measured with small errors and they are usually close to one. Nevertheless, for some portfolios they are significantly different from one and they often do not show a consistent pattern.

Slopes for *FMS* and *PMF* have the expected increasing pattern for rising *DG* and *DY* respectively. There is, however, a notable difference between the steepness of this pattern among portfolios. When dividend yields are low, the difference in loadings between fast and slow *DG* portfolios on *FMS* (0.48 and -0.41 respectively) is about twice as large as it is among the average of the other higher *DY* portfolios. On the *PMF* factor, this difference is also larger for low *DY* portfolios. The interpretation is that firms with low dividend yields are more susceptible to the risk premiums associated with the dividend factors, which is consistent with the reversal phenomenon. Consider for example the *PMF* coefficients of the two portfolio groups with the highest dividend yield. Whether these firms portray slow or fast dividend growth hardly matters to their exposure to the dividend yield risk premium. This lack of distinction agrees with intuition as it indicates that, once dividends are high and further increases may be a stretch, the sensitivity to the dividend yield risk premium does not change with dividend growth projections. But when they are low, there is a lot of room to move from low to high dividend yield due to an adjustment towards fast dividend growth and for the share price to attain positive abnormal returns in conjunction.

4.4 The five-factor model extended by dividend growth

The results of the double-sorting on dividend yield and dividend growth provides guidance on the workings of these variables for returns. Furthermore, dividends are a channel transmitting company fundamentals into tangible returns on stocks. It is therefore useful to investigate the interaction between them in the light of returns of stock portfolio. This section discusses dividend growth in the context of accounting variables in the manner proposed by FF. It follows their analysis with dividend growth added to their model as a factor.

4.4.1 The present value of stocks by company accounting fundamentals

Companies generally pay dividends out of operating profitability, while the non-distributed remainder of profits is added to the capital base of the company. Stimulated by the results of Novy-Marx (2013) in particular, FF identify profitability and investment within the DDM

(4.1), using:

$$S_t = \sum_{n=1}^{\infty} PV_t(D_{t+n}) = \sum_{n=1}^{\infty} \frac{E(D_{t+n})}{(1+r)^n}, \tag{4.11}$$

to pursue a route at the level of company cash flows. Total expected dividend pay-out can be replaced by the difference between profits and investment, analogous to Modigliani and Miller (1961). The market capitalization M_t (the number of shares outstanding times their price S_t) of a company then equals its discounted future balance of profit and capital growth:

$$M_t = \sum_{n=1}^{\infty} \frac{E(Y_{t+n} - dB_{t+n})}{(1+r)^n}. \tag{4.12}$$

The difference between total earnings Y_{t+n} and the change in total book equity dB_{t+n} both for period t to $t+n$ fulfills the role of payout to shareholders in (4.1). FF contend that dividing (4.12) by book equity B_t

$$\frac{M_t}{B_t} = \frac{\sum_{n=1}^{\infty} E(Y_{t+n} - dB_{t+n})/(1+r)^n}{B_t}, \tag{4.13}$$

provides a positive relationship between both the book-to-market equity ratio and total expected future earnings and expected stock return r, as well as a negative relationship between the total expected future change in book equity and expected stock return, under ceteris paribus reasoning.

As market expectations for future values of Y_{t+n} and dB_{t+n} are not observable, Fama and French use their current values in the five-factor-model as proxies for testing (4.13). The approximation puts drawing conclusions from ceteris paribus reasoning at some risk. For example, if a relationship between current total earnings Y_t and expected returns is found that fits such reasoning, then that result hinges on the assumption that the proxy is reasonably accurate. There is no way of knowing that proxying for expectations actually does work, so the relationship, although empirically valid, may follow from another mechanism.

Despite that expected future values are missing elements in the transition from equation (4.13) into an estimable model, future dividend valuation is an observable variable, at least up to the maturity of the options from which future dividends can be implied. I therefore propose to consider the growth in the implied valuation of dividends as an explanatory factor for stock returns. Deployed next to the FF factors, equation (4.13) is matched closely by adding dividend growth as defined in (4.3) and rewriting it to:

$$\frac{M_t}{B_t} = \frac{Y_t - dB_t}{B_t} \times \sum_{n=1}^{\infty} \left(\frac{1 + g_{t,n}}{(1 + y_{t,n-1})(1 + \theta_{t,n})} \right)^n = \frac{Y_t - dB_t}{B_t} \times \sum_{n=1}^{\infty} \left(\frac{PV_t(D_{t+n})}{D_t} \right)^n, \quad (4.14)$$

in which $(1 + y_{t,n-1})(1 + \theta_{t,n})$ equals $(1 + r)$ in (4.13). The first term on the right hand side equals dividends and the second terms equals the divided growth rate.

It is clear from (4.14) that dividend growth steps into the void between current values for earnings and investment and the expectations about their future values, even if it is measured only for a limited horizon. It can do so as a factor in its own right, but also as a term interacting with current payoffs captured by the FF accounting factors.

4.4.2 Dividend growth added to the five-factor model

Motivated by the role of dividend growth shown in (4.14), I test the five-factor model deployed in FF, enlarged by a dividend growth factor:

$$\begin{aligned} R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + s_j SMB_t + h_j HML_t + \\ r_j RMW_t + c_j CMA_t + f_j FMS_t + \epsilon_{j,t}. \end{aligned} \quad (4.15)$$

$R_{j,t}$, $R_{F,t}$, $R_{M,t}$ and $\epsilon_{j,t}$ reflect the same variables as those in the dividend-extended CAPM discussed in the previous section. The five anomaly factors are the difference in returns between diversified portfolios of companies with, for *SMB*, small and big market capitalization, for *HML*, low and high book-to-market, for *RMW*, robust and weak profitability, for *CMA*, conservative and aggressive investment and, for *FMS*, fast and slow growth in dividend valuations, the latter of which is the focus of this paper. The analysis pursued here closely follows the approach and methodology in FF, as before. I omit the dividend yield factor *PMF* as it is strongly correlated to *HML*, of which some discussion follows further below. A key test is whether coefficients $b_j, s_j, h_j.r_j, c_j$ and f_j capture all variation in the expected returns of portfolio j such that the intercept a_j is zero for all portfolios.

Dividend growth takes the role as a gauge of future dividends, forward-looking as it is. At the same time, *DG* has a substantial effect on stock returns and FF and Novy-Marx (2013) show that current profitability and investment do so as well. The research question pursued is whether the proxy seen in *OP* and *Inv* for future stock behavior holds up next to a variable which actually is forward looking by construction.

The tests hereafter investigate the explanatory power of the model in two segments. The first empirical test is a check of the average returns of quartiles of portfolios, sorted along the dimensions of *DG* on the one hand and each of the three FF accounting[20] anomaly variables on the other. I investigate to what extent dividend growth returns can be explained by these variables and whether their own effects on portfolio returns prevail when dividend growth comes into the equation. The second part of the analysis focuses on the time series regressions of portfolios on the factors. I analyze how dividend growth affects regression intercepts and document the slopes of the coefficients, looking in particular at the sensitivity of portfolio returns to the dividend growth risk factor depending on their profitability and investment groups.

4.4.3 Average portfolio returns

Table 4.7 contains the average returns in excess of the one-month U.S. Treasury bill rate of the value weighted double-sorted quartile portfolios over the period July 1996 to June 2015. The details of all sorting are consistent with FF. The breakpoints for the quartiles for book-to-market (*B/M*), (operating) profitability (*OP*) and investment (*Inv*) are sourced from the website of Ken French[21]. Quartiles are formed at the end of June for the three accounting factors and in the second quarter for dividend growth. Returns refer to the year immediately following sorting, as before. The quartile breakpoints use only NYSE stocks, but the sample includes all NYSE, AMEX and NASDAQ stocks in both CRSP and Compustat with share codes 10 or 11.

An ideal disentanglement of variables in (4.14) would sort portfolios jointly for all variables at the same time. As FF point out, this would produce many poorly diversified portfolios that have low power in tests of asset pricing models. They compromise on sorts of *Size* and pairs of the other three variables. In view of the small number of small sized companies in the dividend growth data set, no sort along the *Size* dimension is made here[22]. Instead, the focus will be on pairwise independent sorts of *DG* with each of the other three accounting anomaly variables *B/M*, *OP* and *Inv*.

DG-B/M portfolios in Panel A show that the value effect is in trouble among portfolios sorted by growth in dividends. Returns of the highest *B/M* portfolios average less than 0.10% more than the lowest, and a consistent value effect does not exist for any of the *DG* portfolios. For portfolios formed on *DG-OP* sorts, the picture is not much better. The slowest dividend

[20]The term accounting refers to the variables that appear in equation (4.14), which excludes *Size*.

[21]Not shown in this paper is how factor effects hold up in this data period (1996-2015) relative to the longer period that FF (2015) deploy (1963-2013). As an overall judgment call, the anomaly factors documented by FF hold up in the period 1996-2015, albeit not as well as in their longer data period.

[22]Sorting portfolios on *Size* reveals that the number of companies in the data set under investigation falling in the smallest *Size* group represents only 18% of the universe in the FF data set. This "tiny" group is actually over-represented in numbers as FF do not calculate breakpoints of their universe but of NYSE. The largest *Size* group in this data set overlaps that of FF by about 80%.

growth portfolio increases returns for higher profitability, but faster growing portfolios do not. Portfolios sorted on *DG-Inv* picture the opposite. Here fast growing portfolios show the expected pattern of declining returns for high *Inv* portfolios, whereas there is no clear picture for portfolios with slow dividend growth.

The effect on portfolio returns from profitability and investment found by FF are clearly related to the valuation of future dividends as measured by dividend growth. For average portfolio returns, the investment effect is largely gone when DG is slow, and when it is fast, the profitability effect is largely gone. When DG is slow, stocks are cheap and future returns are high. Returns are particularly high if a firm's high profitability coincides with slow dividend growth (Panel B). In such circumstances a stock is cheap and its past profits have been good, which underpins returns. But the investment effect is found among stocks with fast DG, when stocks are expensive and future returns are low. Returns are particularly low if a firm's high investment coincides with fast dividend growth. In such circumstances a stock is expensive and its pay-out is small, which hurts returns. Both the profitability and the investment effect may thus be primarily transmitted to returns via their relationship to a stock's pricing level. Such transmission may well underlie the success of the ceteris paribus reasoning pursued by FF.

How do the accounting variables and the dividend growth effect interact? Expressed as lower returns subsequent to higher dividend growth, the effect is maintained when portfolios are constructed based on OP and *Inv*. Like the combination with DY, however, when combined with B/M the DG effect is largely lost, except for low B/M portfolios[23].

The reason for the weak performance of B/M may be found in its similarity to DY. Both B/M and DY relate rather stationary values for B_t and D_t to the volatile market capitalization of a company. When the stock price is high, both B/M and DY are low and subsequent returns are low. This simple mechanism may be influential for the value effect and the dividend yield effect to be similar in nature. Moreover, the prevalence of market prices in both parameters may limit the effect from DG on average returns since, as we have seen in the single sorted portfolios, high dividend growth coincides with high returns and is followed by low returns. Sorting portfolios by dividend growth thus introduces pricing that is correlated to stock prices into these portfolios as well. The data bear out that sorting portfolios twice on market pricing, once directly by B/M or DY and once indirectly by DG, causes the expected effects to mostly disappear from view[24].

[23] Notably, FF consider HML redundant next to RMW and CMA, which is investigated further below.
[24] Recall that OP and *Inv* are both scaled by book equity.

4.4.4 Factor definitions

FF pursue several definitions for portfolio sorts and conclude that the choice between them seems inconsequential to the results, and that they are led back to the following method to calculate factors[25]. Portfolios are sorted in June of each year for *Size* at the NYSE median market cap and the breakpoints for *B/M*, *OP*, *Inv* and *DG* are the 30^{th} and 70^{th} percentiles of their respective values for NYSE stocks, except for *DG* where percentiles are taken for all stocks in the data set and sorting occurs per the second quarter of the calendar year.

The *Size* factor is calculated by first taking the average of the returns of the three small portfolios for *B/M* and deducting the returns of their counterparts of the big portfolios. The same is done for *OP* and *Inv*. Next the resulting three returns are averaged to produce the returns of size factor *SMB*.

Value factor *HML* is the average of the return of the two small and big high *B/M* portfolios minus the average of the return of the two small and big low *B/M* portfolios. Profitability and investment factors *RMW* (robust minus weak profitability) and *CMA* (conservative minus aggressive investment) are constructed in the same way. The data used are from Ken French's website[26].

Dividend growth factor *FMS* (fast minus slow dividend growth) is constructed by deducting the average returns of the portfolio with slow dividend growth from the portfolio with fast dividend growth during the two quarters following the quarter of sorting[27]. As before, returns refers to the subsequent 12 months, a period chosen to stay close to the FF methodology.

4.4.5 Factor summary statistics

The summary statistics of factor returns are shown in Panel A of Table 4.3. The period over which these values are calculated refers to 1996-2015 and they deviate from the values of the same variables shown in FF (data period of 1963-2013). The mean return of *RMW* is somewhat higher, but other FF factor returns are smaller and their standard deviations are larger for the more recent data period. The general picture for these values, however, is the same for both periods.

Panel B of Table 4.3 contains the correlation matrix of the six factors. In particular the *RMW* factor stands out, as its correlations with the other factors are considerably stronger

[25] FF label these as 2×3 factors.

[26] Note that, where factors *B/M*, *RMW* and *Inv* are calculated for the universe including stocks on which no options are traded and can be sourced from Ken French's website, the factor *FMS* and all *DG* double-sorted portfolio returns, stem from the smaller universe of stocks that have traded options.

[27] Recall that the lack of small companies in the data set for dividend growth precludes a distinction along the *Size* dimension.

than in the longer FF data period. Its negative correlations with the *Market* factor (−0.46) and the *SMB* factor (−0.53) are stronger (FF: −0.21 and −0.36 respectively), while its correlation with *HML* is 0.52 (FF: 0.08) and the correlation with *CMA* is 0.26, switching signs (FF: −0.10). All other correlations change by less than 0.10. Arguably, the interaction of *RMW* with the other variables is not stable over time and depends on the data period.

The *FMS* factor correlates positively with the *Market* and the *SMB* factor and negatively with the factors composed of portfolios sorted by company accounting variables. This pattern is the opposite of the accounting factors, which correlate positively among themselves and negatively with *Market* and *Size*. As the average return of the *FMS* factor is negative, this is perhaps not surprising. All *FMS* correlations are more than three standard errors from zero, except for correlation with *SMB*. Correlations of the dividend yield factor *PMF* shows the opposite pattern, much like *HML*.

In Table 4.8 the results of multivariate regressions of individual factor returns on the other factors are shown. *Market* factor $R_M - R_F$ has negative coefficients for both *RMW* and *CMA*, the *SMB* factor is strongly negatively correlated to *RMW*, but not to any other, and *RMW* and *CMA* are negatively dependent[28][29].

Turning to the two dividend factors, their dependence on the FF factors differs markedly from each other. *PMF* has a substantial negative coefficient for $R_M - R_F$ and a large but positive one for *HML* as well. A negative coefficient to the *Market* factor implies that the risk premium for high dividend companies goes down in positive markets. Such a pattern fits the view that high dividend companies are defensive (Asness, Frazzini and Pedersen, 2014). The positive relationship between *PMF* and *HML* is discussed earlier. The other three FF factors play much less of a part, to the point of no significance at 5% for any of their coefficients. The intercept is equal to the time-series average at 0.22%. But once *FMS* is added as a factor to the *PMF* regression, it drops to 0.10%. The *FMS* coefficient is strong and significant, which matches the pattern found in the extended CAPM (4.10). In the case of *PMF*, the inclusion of *FMS* as a regressor increases the R^2 by more than 8%, while it barely moves the needle for the other factors. The reason is that *FMS* returns are explained to some degree by the other factors. The intercept of the FMS factor is not significantly

[28] Regression coefficients largely show the same pattern as found by FF for the period of 1963-2013, but the *HML* factor stands out. FF find that it is explained by the other factors to the point where its intercept is close to zero. The authors conclude that *HML* is redundant in their five-factor model for describing average returns since its variation is captured by the exposure to primarily *RMW* and *CMA*. In the data period of 1996 to 2015, however, its regression intercept is −0.34 and significant at 5%. It would be interesting to investigate whether this reduction in the value premium is related to the initiation of value investing following the publication of Fama and French (1993) in which it was first described.

[29] The regression results shown in Table 4.8 include *FMS* and *PMF* as regressors, without them, the intercept for *HML* is −0.32.)

different from zero, but at –0.22%, it equals about half of its time-series mean of –0.40% (Table 4.3, Panel A). The coefficients of the FF factors are relatively small, with a t-statistic around 2. The $R_M - R_F$ and *SMB* coefficients are close to zero. When *FMS* is added in the RHS, the intercepts of all FF factor regressions change by less than 0.04%. *FMS* average return is not strongly captured by the factors of the five-factor-model, while it correlates less strongly with them than these factors correlate among themselves. The conclusion of these results is supportive for viewing dividend growth as a force in its own right in the estimation of model (4.15).

4.4.6 Portfolio regression intercepts

The excess returns of the portfolios regressed on the factors are the next subject of investigation. The five-factor model enhanced by dividend growth is tested next to guises from which one ore more factors are omitted. Note that the *Market* and *SMB* factors are included as regressors in all cases.

The average absolute alpha's decrease for all models when *FMS* is added as a factor (Table 4.9). Alpha's improve most for $DG - Inv$ portfolios (5 basis points) and typically by 2 to 3 basis points for portfolios sorted by one of the other three variables. However, *FMS* never reduces the alpha's by more than the other factors do, when applied on their own (next to *Market* and *SMB*).

The GRS-statistic and its p-value do not reject most model specifications. Portfolios double-sorted by *B/M* or *Inv* show similarly good values, but those of *DG-OP* portfolios fare less well. The original five-factor model without *FMS* does well for *DG-DY* portfolios. When factors are removed or added, the difference in these values among portfolios remains similar. This conclusion is not in line with the results from FF's tests of *Size* sorted portfolios. They find that statistics sometimes deteriorate when a factor is added. The more consistent result found here suggests that portfolio returns sorted along *DG* are more easily explained than along *Size*.

The GRS values found in the elaborate CAPM enhanced by dividends *PMF* (4.10) in Table 4.5 attain values in the same range as those calculated for the five-factor model without *FMS* (4.15). If both models are about as strong, then it makes sense that the inclusion of *FMS* on the right hand side of the five-factor model further improves the regressions results. The p-values of the GRS-statistic improve substantially in whichever combination of portfolios and regressors *FMS* is added. For example, in *DG-OP* and *DG-Inv* sorted portfolios, adding *FMS* often doubles the p-values of the GRS-statistic[30].

[30]Overall, both $A|a_j|/A|\bar{r}_i|$ and $A(\hat{\alpha_j}^2)/A(\tilde{\bar{r}}_i^2)$ measures are higher than in FF's larger data period, leaving

Abnormal returns of *DG-DY* portfolios are explained similarly well by the five-factor model and the dividend-extended *FMS*, whether *FMS* is included as a factor or not. But for all other portfolio specifications returns unexplained by the fully specified five-factor are reduced once dividend growth comes into play as a factor. These improvements confirm that dividend growth addresses risk characteristics in portfolio returns which are not contained in the five-factor model.

Tables 4.10 to 4.13 contain the coefficients of the regressions of *DG* and *FF* variable double-sorted portfolio returns. The intercepts vary quite a lot, with several portfolios in difficulty due to values above 0.30% and below −0.30%, levels at which they attain significance at 5% confidence. Perhaps surprisingly, these larger values do not appear in the same *DG* quartiles. Nevertheless, with 18 out of 64 portfolio regression intercepts more than one standard error away from zero, proportionately this is a better result than the intercepts in the FF regressions of returns of portfolios sorted by *Size* and another factor in which 46 out of 75 intercepts are more than one standard error away from zero (not shown here). The betas to the *Market* factor are close to one, although high *DG* portfolios often attract values near and below 0.90, which is significantly different from one due to the small standard errors of the estimated coefficients.

4.4.7 Portfolio regression slopes

The loadings on the risk factors further cast light on the transmission mechanism through which dividend growth matters to portfolio returns. Before focusing on the dividend growth factor, it is worth noticing that the FF factors associated with the tested portfolios sorted by *DG* show the expected slopes. *HML* coefficients rise strongly as *B/M* rises in *DG-B/M* sorted portfolios and so do *RMW* coefficients in *DG-OP* portfolios. The slopes for *CMA* fall for higher *Inv* portfolios, which is also in line with expectations.

On average, *FMS* slopes do not change much for rising *B/M*, *OP* and *Inv*, but their values are quite different within these groups. Moving from low to high *DG* within *DG-B/M* portfolios, the increase falls as *B/M* increases. The lowest *B/M* quartile *FMS* slopes differ nearly 1.00, while for the largest *B/M* quartile this difference is only 0.10. Returns of *DG-DY* portfolios regressed on the extended CAPM also show such patterns for *FMS* slopes,

a weaker impression of these models for the 1996-2015 data period (Table 4.9). The absolute levels of intercepts and the variance of their estimates are again often larger than those of return deviations, in which cases the dispersion of the intercepts is the larger of the two. For example in *DG-B/M* portfolios the average return deviation is less than 0.10%, while among *Size-B/M* sorted portfolios this value equals almost 0.20% (FF, 2015). It is a harder job to explain these smaller deviations, which is reflected in these statistics. Nevertheless, in most cases the addition of *FMS* as factor on the *RHS* of the regressions improves these measures.

which confirms the overlapping characteristics of DY and B/M for returns.

In *DG-OP* sorted portfolios the *FMS* slopes of the two high *OP* category portfolios both increase by more than 1.00 from slow to fast *DG* portfolios and these slopes for the low portfolios differ much less. Although more moderate, this pattern shows again in *DG-Inv* portfolios. Here the highest *Inv* portfolios attract a difference between fast and slow growing portfolios for *FMS* of nearly 1.00, while this is 0.38 for low *Inv* portfolios.

Susceptibility to *FMS* returns increase for high *OP* and *Inv* portfolios and decrease for high B/M and DY portfolios. When *OP* is high it stands to reason that portfolios sorted by this variable and dividend growth are more susceptible to *FMS* when dividend growth is fast. High profitability gives more scope for dividend increases to materialize than low profitability, as companies are then often loss-making, so when *DG* is fast such stocks respond more strongly to the *FMS* premium. If companies have low profits, but nonetheless pay dividends, there is no consistent *FMS* effect. Many companies in these portfolios have negative *OP* at least temporarily. Given that they do pay dividends, they deplete their capital. It makes sense that sustainable higher dividends increasing the sensitivity to *FMS* emerges only among companies with profitability of some magnitude. Earlier findings in this paper are that fast dividend growth is generally followed by low returns and high profitability increases returns more when dividend growth is slow (Table 4.7). These findings concern average portfolio returns and do not stand in the way of this interpretation of *FMS* slopes, which represent sensitivity to the dividend growth risk premium.

In the case of *DG-Inv* double sorted portfolios, a similar argument for the ability to increase dividends and an associated decreasing effect on returns can be made for a longer horizon. When companies invest, it takes some time for revenues to come to fruition. While the *FMS* coefficient moves up from low to high *DG* for high *Inv* groups, the pattern is also found in low *Inv* but to a lesser degree at about half of the increase in higher *Inv* portfolios. Future profitability due to larger investment gives more scope for a sustainable *FMS* impact on returns. But probably low investment companies are sufficiently profitable on average for dividend growth to matter still to show an effect in both *DG-OP* and *DG-Inv* portfolios.

Portfolios sorted by B/M and DY have similar return characteristics and *FMS* loadings. Both groups show high sensitivity to *FMS* when they are low and *DG* is slow, with returns decreasing for faster *DG*. This sensitivity is lost for high B/M and DY. The returns from factors based on these variables strongly correlate. The difference in sensitivity to *FMS* among *DG* sorted portfolios decrease as B/M and DY rise.

Both B/M and DY are fractions with the price of the stock as their denominator. Strong susceptibility to *FMS* factor returns of a portfolio of highly priced stocks measured as low B/M or DY implies that an adjustment in future dividend valuations weighs heavily on such stocks. Such a response is commonplace among investors; a stock's high price is supported by an anticipated future rise in dividends. Once that anticipation fades, stocks that depend for their expensiveness on this dividend outlook are punished more harshly than stocks that are cheaper relative to book value or dividends. Such stock pricing aligns with the below average future returns of high *DG* and low *B/M* or *DY* portfolios (Table 4.2, Panel A).

Lastly, *SMB* coefficients are significantly different from zero in 19 out of 48 portfolios, without a clear tendency. However, for rising *DG*, two patterns can be found. Among *DG-B/M* portfolios, *SMB* coefficients increase for higher *B/M*, while they fall for higher *Inv* in the case of *DG-Inv* sorted portfolios. Double-sorting with *DG* does not interfere with this finding, high *B/M* and low *OP* remain positively associated with the *Size* effect.

4.5 Conclusion

Market participants adjust the prices of options to reflect their valuations of future dividends. This paper investigates how stock prices adjust in conjunction with changes to these valuations and whether their future returns are impacted.

A high valuation of future dividends explains higher returns during the time when the valuation is made and lower returns afterwards. The results in this paper suggest that this sequence constitutes a reversal to overshooting stock prices. A portfolio sorted by fast dividend growth stocks outperforms a slow growing portfolio by 0.90% per month in the year preceding sorting, but underperforms it by 0.30% per month in the subsequent year.

Portfolios sorted on well known accounting variables do not show a consistent return pattern when sorted by dividend growth as well. The profitability effect only prevails when dividend growth is slow, while the investment effect depends on fast dividend growth. The value effect appears to be lost altogether on dividend growth portfolios.

The returns of a factor based on portfolios of dividend growth are not very dependent on factors returns based on accounting variables. Nonetheless, high profitability and investment increase the susceptibility of portfolio stock returns to dividend growth, a phenomenon which may serve as the transmission mechanism for company accounting fundamentals into returns.

Dividend growth correlates with stock prices and changes to it will thus influence the expensiveness of a stock. The evidence presented here suggests that relationships between stock returns and variables such as dividend yield, *Book-to-Market*, profitability and investment may be caused by a stock's pricing relative to its fundamentals.

4.6 Tables

Table 4.1: **Cross-sectional regressions of excess returns on dividend growth.**

Each quarter in the period 1996-2015 portfolios are formed by sorting stocks on dividend growth (DG). Excess returns $R_{j,t}$ are defined as their returns above the quarter average. These excess returns are regressed on DG in three equations, in which $I_{\Delta DG_{j,t}}$ is a change in the DG quartile ranking equal to 1 if the change from quarter $t-1$ to t is upward and -1 if it is downward, $I_{\Delta DG_{j,t}>0}$ equals 1 if $\Delta DG_{j,t}$ is positive, $I_{\Delta DG_{j,t}<0}$ equals 1 if $\Delta DG_{j,t}$ is negative and $I_{\Delta DG_{j,t}=0}$ equals 1 if ΔDG is unchanged. Regressions are performed each quarter. The coefficients reported are calculated as the average of the quarterly coefficients, the t-statistics are averaged multiplied by the square root of the number of regressions performed (78 for $n = 0$) (Fama and MacBeth, 1973). The R^2 reported are the average of the quarterly R^2.

$$R_{j,t+i} = \alpha_t + \beta_t I_{\Delta DG_{j,t}} + \epsilon_{j,t} \tag{5}$$

$$R_{j,t+i} = \beta_{1,t} I_{\Delta DG_{j,t}>0} + \beta_{2,t} I_{\Delta DG_{j,t}<0} + \beta_{3,t} I_{\Delta DG_{j,t}=0} + \epsilon_{j,t} \tag{6}$$

$$\begin{aligned} R_{j,t+i} = \alpha_t &+ \beta_{1,t} I_{\Delta DG_{j,t}} \times I_{DG=1} + \beta_{2,t} I_{\Delta DG_{j,t}} \times I_{DG=2} + \\ &\beta_{3,t} I_{\Delta DG_{j,t}} \times I_{DG=3} + \beta_{4,t} I_{\Delta DG_{j,t}} \times I_{DG=4} + \epsilon_{j,t} \end{aligned} \tag{7}$$

	(5a)	(5b)	(6a)	(6b)	(7a)	(7b)
	$i = 0$	$i = 2$	$i = 0$	$i = 2$	$i = 0$	$i = 2$
Intercept	-0.001	-0.001			0.001	-0.013
	(-0.33)	(-0.32)			(0.01)	(-0.10)
$I_{\Delta DG_{j,t}}$	0.718	-0.067				
	(13.08)	(-1.39)				
$I_{\Delta DG_{j,t}>0}$			0.716	-0.052		
			(11.45)	(-0.97)		
$I_{\Delta DG_{j,t}<0}$			-0.724	0.083		
			(-12.13)	(1.43)		
$I_{\Delta DG_{j,t}=0}$			0.004	-0.018		
			(0.19)	(-0.68)		
$\Delta DG_{j,t} \times (DG_{j,t} = 1)$					0.935	-0.148
					(9.07)	(-1.59)
$\Delta DG_{j,t} \times (DG_{j,t} = 2)$					0.483	-0.068
					(6.90)	(-0.86)
$\Delta DG_{j,t} \times (DG_{j,t} = 3)$					0.618	0.061
					(7.8)	(0.84)
$\Delta DG_{j,t} \times (DG_{j,t} = 4)$					0.975	-0.162
					(8.25)	(-1.43)
R^2	0.013	0.003	0.015	0.005	0.020	0.009

Table 4.2: **Value-weight portfolios formed on dividend growth and dividend yield.**

Panel A: Average monthly percent returns in excess of the one-month Treasury bill rate for portfolios formed on dividend growth (DG) and on dividend yield (DY). At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to four DG groups and four DY groups independently at their quartile breakpoints of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. DG Average and DY Average are the returns of single-sorted portfolios.

Panel B: The average number of firms per DG-DY portfolio throughout the period 1996-2015.

DY	Low	2	3	High	DY Average
Panel A: Average excess returns					
Slow DG	0.93	0.79	0.92	0.56	0.84
2	0.47	0.72	0.68	0.70	0.57
3	0.63	0.52	0.72	0.75	0.65
Fast DG	0.27	0.74	0.65	0.66	0.54
DG Average	0.43	0.66	0.75	0.68	
Panel B: Average number of companies					
Slow DG	40	42	45	43	
2	22	46	60	52	
3	41	54	46	38	
Fast DG	87	38	24	27	

Table 4.3: **Summary statistics of monthly factor returns in the period 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015). The data are sourced from Ken French's website.
At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30^{th} and 70^{th} percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of fast DG group minus the returns of the slow DG group and *PMF* is the returns of prodigal DY group minus the returns of the frugal DY group.

Panel A: Average monthly returns (Mean), the standard deviations of monthly returns (St dev.) and the t-statistics for the average returns.

Panel B: the correlations of the factors among themselves.

	$R_M - R_F$	SMB	HML	RMW	CMA	FMS	PMF
Panel A: Mean returns, standard deviations and t-statistics							
Mean	0.58	0.25	0.22	0.34	0.30	-0.40	0.22
St dev	4.62	3.34	3.35	2.90	2.24	2.81	3.90
t-stat	1.89	1.13	0.98	1.75	2.00	-2.13	0.86
Panel B: Correlations between factors							
$R_M - R_F$	1	0.22	-0.23	-0.46	-0.36	0.20	-0.46
SMB	0.22	1	-0.20	-0.53	-0.03	0.10	-0.18
HML	-0.23	-0.20	1	0.52	0.65	-0.16	0.55
RMW	-0.46	-0.53	0.52	1	0.26	-0.22	0.40
CMA	-0.36	-0.03	0.65	0.26	1	-0.23	0.51
FMS	0.20	0.10	-0.16	-0.22	-0.23	1	-0.42
PMF	-0.46	-0.18	0.55	0.40	0.51	-0.42	1

Table 4.4: **Regressions of factor returns used in the dividend-extended CAPM (4.10) on each other for the period 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30$^{\text{th}}$ and 70$^{\text{th}}$ percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of fast DG group minus the returns of the slow DG group and *PMF* is the returns of prodigal DY group minus the returns of the frugal DY group. *Int* is the regression intercept.

	Int	$R_M - R_F$	PMF	FMS	R^2
$R_M - R_F$					
Coef	0.70		-0.54	0.01	0.21
t-stat	(2.54)		(-6.97)	(0.09)	
	0.70		-0.54		0.21
	(2.55)		(-7.74)		
	0.71			0.33	0.04
	(2.32)			(3.04)	
PMF					
Coef	0.22	-0.33		-0.48	0.33
t-stat	(1.03)	(-6.97)		(8.53)	
	-0.01			-0.59	0.18
	(-0.04)			(-6.95)	
	0.45	-0.39			0.21
	(1.92)	(-7.74)			
FMS					
Coef	-0.33	0.00	-0.30		0.18
t-stat	(-1.91)	(0.09)	(-2.00)		
	-0.33		-0.30		0.18
	(-1.93)		(-6.95)		
	-0.47	0.12			0.04
	(-2.51)	(3.04)			

Table 4.5: **Summary statistics for tests of CAPM extended by dividend growth and dividend yield.**

Test portfolios are 4 × 4 sorted by dividend growth and dividend yield. The *GRS*-statistic tests whether the expected values of all 16 intercept estimates are zero, the average absolute value of the intercepts, A$|a_j|$, A$|a_j|/|\bar{r_j}|$, the average absolute value of the intercept a_j over the average absolute value of $\bar{r}_j$, which is the average return on portfolio i minus the average of the portfolio returns, and $A(\hat{a}_i^2)/A(\hat{\mu}_j^2)$, which is $A(\hat{a}_j^2)/A(\hat{\mu}_j^2)$ the average squared intercept over the average squared value of $\bar{r}_j$, corrected for sampling error in the numerator and denominator.

The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + d_j PMF_t + f_j FMS_t + e_{j,t}$

	GRS	pGRS	A$|a_j|$	$\frac{A|a_j|}{A|\bar{r_j}|}$	$\frac{A(\hat{a_j}^2)}{A(\hat{\mu_j}^2)}$
16 DG-DY portfolios					
$R_M - R_F$	1.30	0.20	0.29	2.48	2.80
$R_M - R_F$ FMS	0.90	0.57	0.22	1.86	1.61
$R_M - R_F$ PMF	1.07	0.39	0.19	1.63	1.49
$R_M - R_F$ PMF FMS	0.83	0.65	0.18	1.52	0.97

Table 4.6: **Regressions for 16 value-weighted DG-DY portfolios: 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30$^{\text{th}}$ and 70$^{\text{th}}$ percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor FMS is the returns of fast DG group minus the returns of the slow DG group and PMF is the returns of prodigal DY group minus the returns of the frugal DY group.

The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + p_jPMF_t + f_jFMS_t + e_{j,t}$

DY	Low	2	3	High	Low	2	3	High
		a				t(a)		
DG Slow	0.32	0.21	0.31	-0.22	1.39	1.01	1.40	-1.09
2	-0.02	0.15	0.13	0.00	-0.09	0.71	0.74	0.03
3	0.26	-0.02	0.23	0.27	1.00	-0.10	1.36	1.67
DG Fast	-0.08	0.24	0.22	0.19	-0.62	1.11	1.01	0.98
		b				t(b)		
DG Slow	0.96	0.82	0.87	1.02	17.37	16.36	16.48	21.15
2	0.93	0.87	0.84	1.04	16.73	16.79	19.95	27.08
3	0.91	0.88	0.85	0.79	14.87	18.92	20.74	20.64
DG Fast	1.03	0.93	0.89	0.90	34.13	17.82	16.81	19.42
		f				t(f)		
DG Slow	-0.41	-0.31	-0.18	-0.24	-4.64	-3.84	-2.07	-3.08
2	-0.09	-0.08	-0.05	0.17	-1.05	-0.94	-0.76	2.76
3	-0.09	-0.05	0.17	0.15	-0.96	-0.69	2.63	2.49
DG Fast	0.48	0.13	0.07	0.28	9.88	1.57	0.87	3.73
		p				t(p)		
DG Slow	-0.54	0.05	0.25	0.61	-7.65	0.79	3.71	9.93
2	-0.27	0.17	0.20	0.85	-3.72	2.64	3.77	17.36
3	-0.06	0.11	0.35	0.64	-0.75	1.93	6.66	12.96
DG Fast	-0.12	0.23	0.31	0.71	-3.22	3.48	4.64	12.09

Table 4.7: **Average monthly percent returns in excess of the one-month Treasury bill rate for portfolios formed on dividend growth and accounting variables.**

Portfolios are 4 × 4 sorted by dividend growth (DG) and *Book-to-Market* (B/M), DG and *Operating Profitability* (OP), DG and *Investment* (Inv) and DG and dividend yield (DY): 1996-2015. At the end of each June, stocks are allocated to four B/M groups using NYSE market cap breakpoints, to four OP groups using accounting data for the fiscal year ending in year $t-1$ (revenues minus cost of goods sold, minus SG&A and interest expenses all divided by book equity) and to four Inv groups using the change in total assets from the fiscal year ending in $t-1$, divided by $t-2$ total assets. At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to four DG groups and four DY groups independently at the quartile breakpoints of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Returns are value-weighted.

	Low	2	3	High
Panel A: 16 DG-B/M portfolios				
Slow DG	0.97	0.61	0.79	0.60
2	0.57	0.83	0.68	0.89
3	0.65	0.71	0.66	0.71
Fast DG	0.40	0.62	0.69	0.51
Panel B: 16 DG-OP portfolios				
Slow DG	0.71	0.52	0.70	1.32
2	0.70	0.68	0.85	0.51
3	0.71	0.74	0.59	0.76
Fast DG	0.40	0.60	0.40	0.43
Panel C: 16 DG-Inv portfolios				
Slow DG	0.87	0.97	0.89	0.77
2	0.73	0.65	0.55	0.65
3	0.77	0.74	0.66	0.47
Fast DG	0.56	0.45	0.58	0.29
Panel D: 16 DG-DY portfolios				
Slow DG	0.93	0.79	0.92	0.56
2	0.47	0.72	0.68	0.70
3	0.63	0.52	0.72	0.75
Fast DG	0.27	0.74	0.65	0.66

Table 4.8: **Regressions of factor returns in the dividend extended five-factor model (4.15) regressed on each other for the period 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015). The data are sourced from Ken French's website. At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30$^{\text{th}}$ and 70$^{\text{th}}$ percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of fast DG group minus the returns of the slow DG group and *PMF* is the returns of prodigal DY group minus the returns of the frugal DY group. *Int* is the regression intercept.

	Int	$R_M - R_F$	*SMB*	*HML*	*RMW*	*CMA*	*PMF*	*FMS*	R^2
$R_M - R_F$									
Coef	1.01		-0.03	0.58	-0.75	-0.70	-0.40		0.40
t-stat	(4.06)		(-0.30)	(5.20)	(-6.49)	(-4.69)	(-5.11)		
	1.05		-0.01	0.45	-0.82	-0.88		0.07	0.32
	(3.97)		(-0.12)	(3.87)	(-6.72)	(-5.63)		(0.71)	
SMB									
Coef	0.44	-0.02		0.06	-0.68	0.14	-0.04		0.30
t-stat	(2.21)	(-0.30)		(0.70)	(-7.75)	(1.17)	(-0.60)		
	0.43	-0.01		0.05	-0.68	0.13		0.00	0.30
	(2.16)	(-0.12)		(0.54)	(-7.68)	(1.08)		(-0.03)	
HML									
Coef	-0.34	0.19	0.03		0.49	0.76	0.21		0.63
t-stat	(-2.34)	(5.20)	(0.70)		(7.74)	(10.41)	(4.79)		
	-0.31	0.14	0.03		0.55	0.91		0.05	0.59
	(-2.01)	(3.87)	(0.54)		(8.29)	(12.78)		(0.92)	
RMW									
Coef	0.52	-0.21	-0.32	0.43		-0.25	0.00		0.56
t-stat	(3.92)	(-6.49)	(-7.75)	(7.74)		(-3.10)	(-0.07)		
	0.49	-0.21	-0.31	0.43		-0.28		-0.09	0.57
	(3.67)	(-6.72)	(-7.68)	(8.29)		(-3.39)		(-1.84)	
CMA									
Coef	0.31	-0.13	0.04	0.43	-0.16		0.07		0.52
t-stat	(2.81)	(-4.69)	(1.17)	(10.41)	(-3.10)		(1.99)		
	0.29	-0.14	0.04	0.47	-0.18			-0.09	0.53
	(2.67)	(-5.63)	(1.08)	(12.78)	(-3.39)			(-2.46)	
PMF									
Coef	0.22	-0.27	-0.04	0.44	-0.01	0.25			0.43
t-stat	(1.04)	(-5.11)	(-0.60)	(4.79)	(-0.07)	(1.99)			
	0.10	-0.25	-0.04	0.47	-0.08	0.14		-0.40	0.51
	(0.52)	(-5.19)	(-0.65)	(5.47)	(-0.78)	(1.16)		(-5.76)	
FMS									
Coef	-0.22	-0.05	-0.02	0.22	-0.18	-0.20	-0.33		0.21
t-stat	(-1.20)	(-1.16)	(-0.26)	(2.71)	(-2.00)	(-1.84)	(-5.76)		
	-0.29	0.03	0.00	0.08	-0.17	-0.29			0.09
	(-1.50)	(0.71)	(-0.03)	(0.92)	(-1.84)	(-2.46)			

Table 4.9: **Summary statistics for tests of the Fama and French five-factor model extended by dividend growth.**

Test portfolios are 4 × 4 formed on dividend growth (*DG*) and *Book-to-Market*, *DG* and *Operating Profitability*, *DG* and *Investment* and *DG* and dividend yield. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015). The data are sourced from Ken French's website. All regressions include $R_M - R_F$ and *SMB* and the factors indicated. The first line of each Panel reflects regressions on $R_M - R_F$ and *SMB* only.
The *GRS*-statistic tests whether the expected values of all 16 intercept estimates are zero, the average absolute value of the intercepts, A$|a_j|$, A$|a_j|/|\bar{r_j}|$, the average absolute value of the intercept a_j over the average absolute value of $\bar{r}_j$, which is the average return on portfolio j minus the average of the portfolio returns, and $A(\hat{a}_j^2)/A(\hat{\mu}_j^2)$, which is $A(\hat{a}_j^2)/A(\hat{\mu}_j^2)$ the average squared intercept over the average squared value of $\bar{r}_j$, corrected for sampling error in the numerator and denominator.

The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + s_jSMB_t + h_jHML_t + r_jRMW_t + c_jCMA_t + f_jFMS_t + e_{j,t}$.

	GRS	pGRS	A$\vert a_i\vert$	$\frac{A\vert a_i\vert}{A\vert\bar{r_i}\vert}$	$\frac{A(\hat{a_i}^2)}{A(\hat{\mu_i}^2)}$
Panel A: 16 DG-B/M portfolios					
	1.11	0.35	0.27	2.80	3.17
HML	0.99	0.47	0.12	1.24	2.13
HML RMW	0.98	0.48	0.13	1.38	1.68
HML CMA	0.87	0.60	0.11	1.09	1.96
RMW CMA	1.21	0.26	0.24	2.49	2.95
HML RMW CMA	1.08	0.38	0.15	1.58	1.66
FMS	0.73	0.76	0.22	2.30	2.47
HML FMS	0.67	0.82	0.12	1.24	1.32
HML RMW FMS	0.73	0.76	0.11	1.16	0.96
HML CMA FMS	0.64	0.85	0.11	1.14	1.27
RMW CMA FMS	1.04	0.42	0.20	2.11	2.46
HML RMW CMA FMS	0.93	0.53	0.12	1.26	1.27
Panel B: 16 DG-OP portfolios					
	1.94	0.02	0.25	1.78	2.74
HML	1.87	0.03	0.20	1.38	2.40
RMW	1.43	0.13	0.22	1.51	1.63
HML RMW	1.44	0.13	0.21	1.47	1.61
HML CMA	1.68	0.05	0.17	1.21	2.18
RMW CMA	1.44	0.13	0.24	1.71	2.04
HML RMW CMA	1.18	0.28	0.19	1.36	1.40
FMS	1.51	0.10	0.23	1.58	2.14
HML FMS	1.49	0.10	0.18	1.29	1.76
RMW FMS	1.17	0.30	0.18	1.24	1.31
HML RMW FMS	1.18	0.29	0.17	1.19	1.33
HML CMA FMS	1.41	0.14	0.17	1.16	1.63
RMW CMA FMS	1.26	0.23	0.21	1.48	1.40
HML RMW CMA FMS	1.03	0.42	0.17	1.19	1.10
Panel C: 16 DG-Inv portfolios					
	1.25	0.23	0.28	1.91	2.50
HML	1.12	0.34	0.19	1.34	1.37
CMA	0.86	0.61	0.15	1.03	0.90
HML RMW	0.88	0.59	0.17	1.19	1.44
HML CMA	0.95	0.51	0.15	1.06	0.92
RMW CMA	0.77	0.72	0.19	1.31	1.33
HML RMW CMA	0.69	0.80	0.15	1.07	0.95
FMS	0.84	0.63	0.21	1.47	1.42
HML FMS	0.77	0.72	0.13	0.90	0.76
CMA FMS	0.61	0.88	0.09	0.64	0.86
HML RMW FMS	0.64	0.85	0.13	0.88	0.50
HML CMA FMS	0.69	0.80	0.10	0.67	0.86
RMW CMA FMS	0.62	0.87	0.14	0.94	0.77
HML RMW CMA FMS	0.56	0.91	0.12	0.83	0.51
Panel D: 16 DG-DY portfolios					
	1.42	0.13	0.32	2.66	3.46
HML	1.26	0.22	0.19	1.64	1.62
CMA	1.09	0.37	0.17	1.45	1.40
HML RMW	1.06	0.40	0.14	1.22	1.39
HML CMA	1.12	0.34	0.18	1.55	1.43
RMW CMA	1.14	0.32	0.17	1.43	1.85
HML RMW CMA	0.99	0.47	0.16	1.31	1.51
FMS	1.01	0.45	0.24	2.04	2.16
HML FMS	0.90	0.57	0.17	1.42	0.99
CMA FMS	0.82	0.66	0.15	1.26	1.07
HML RMW FMS	0.80	0.68	0.12	1.03	1.15
HML CMA FMS	0.85	0.62	0.16	1.36	1.03
RMW CMA FMS	0.97	0.49	0.16	1.32	1.78
HML RMW CMA FMS	0.84	0.63	0.14	1.22	1.40

Table 4.10: **Regressions for 16 value-weighted *DG-B/M* portfolios: 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015). The data are sourced from Ken French's website.
At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30$^{\text{th}}$ and 70$^{\text{th}}$ percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of the fast DG group minus the returns of the slow DG group and *PMF* is the returns of prodigal DY group minus the returns of the frugal DY group.

The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + s_j SMB_t + h_j HML_t + r_j RMW_t + c_j CMA_t + f_j FMS_t + e_{j,t}$.

B/M	Low	2	3	High	Low	2	3	High
	a				t(a)			
DG Slow	0.08	-0.31	-0.23	-0.15	0.60	-1.70	-1.28	-0.60
2	-0.22	0.19	-0.19	0.23	-1.80	1.10	-1.06	1.22
3	-0.04	0.03	-0.07	-0.01	-0.26	0.17	-0.38	-0.07
DG Fast	-0.09	-0.02	-0.04	-0.12	-0.75	-0.09	-0.20	-0.47
	b				t(b)			
DG Slow	1.03	1.00	1.10	0.92	33.22	22.06	24.36	14.41
2	0.99	0.98	1.04	0.81	32.84	23.10	24.05	17.68
3	0.86	1.05	1.10	0.96	25.12	25.65	24.33	19.22
DG Fast	1.10	1.03	0.92	1.02	36.22	20.05	17.33	16.66
	s				t(s)			
DG Slow	-0.10	0.07	0.09	0.16	-2.34	1.19	1.55	1.86
2	-0.26	0.01	-0.14	0.05	-6.38	0.24	-2.42	0.85
3	-0.17	-0.20	-0.01	0.07	-3.75	-3.69	-0.17	1.10
DG Fast	-0.09	0.01	0.19	0.18	-2.29	0.14	2.67	2.16
	h				t(h)			
DG Slow	-0.09	0.22	0.66	0.80	-1.66	2.69	8.26	7.11
2	-0.06	0.57	0.65	1.12	-1.19	7.58	8.56	13.77
3	-0.06	0.31	0.60	0.82	-1.02	4.34	7.49	9.28
DG Fast	-0.10	0.38	0.55	0.78	-1.95	4.24	5.84	7.25
	r				t(r)			
DG Slow	0.27	0.41	0.13	-0.14	4.37	4.58	1.44	-1.09
2	0.45	0.09	0.20	-0.21	7.56	1.07	2.38	-2.28
3	0.28	0.16	0.32	-0.01	4.14	1.94	3.62	-0.13
DG Fast	0.34	0.32	0.20	0.19	5.60	3.16	1.88	1.58
	c				t(c)			
DG Slow	0.12	0.24	0.20	0.15	1.62	2.14	1.77	0.96
2	0.39	-0.06	0.00	-0.19	5.24	-0.58	-0.01	-1.69
3	0.46	0.34	0.04	0.09	5.51	3.45	0.35	0.78
DG Fast	0.17	0.09	0.07	0.01	2.25	0.75	0.51	0.04
	f				t(f)			
DG Slow	-0.49	-0.25	-0.30	-0.01	-11.15	-3.82	-4.71	-0.12
2	-0.10	0.01	-0.15	-0.16	-2.26	0.24	-2.40	-2.52
3	-0.05	0.16	0.20	0.05	-0.98	2.71	3.10	0.69
DG Fast	0.50	0.27	0.07	0.09	11.65	3.76	0.88	1.06

Table 4.11: **Regressions for 16 value-weighted *DG-OP* portfolios: 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015. The data are sourced from Ken French's website.
At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30$^{\text{th}}$ and 70$^{\text{th}}$ percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of the fast DG group minus the returns of the slow DG group and *PMF* is the returns of prodigal DY group minus the returns of the frugal DY group.

The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + s_j SMB_t + h_j HML_t + r_j RMW_t + c_j CMA_t + f_j FMS_t + e_{j,t}$.

OP	Low	2	3	High	Low	2	3	High
		a				t(a)		
DG Slow	0.25	-0.34	-0.17	0.36	0.62	-1.30	-0.59	1.95
2	-0.30	0.10	0.04	-0.32	-0.95	0.37	0.20	-2.02
3	0.02	0.19	-0.02	-0.04	0.06	0.73	-0.12	-0.24
DG Fast	-0.16	-0.32	0.24	-0.02	-0.43	-1.19	0.95	-0.10
		b				t(b)		
DG Slow	0.96	1.07	1.04	0.87	9.83	16.42	14.42	18.97
2	1.18	0.99	1.06	0.91	15.29	14.70	20.92	23.33
3	1.03	0.94	0.97	0.91	13.88	14.57	20.04	20.48
DG Fast	1.13	1.06	1.10	1.01	12.60	16.14	17.96	22.76
		s				t(s)		
DG Slow	0.17	0.13	0.06	-0.02	1.27	1.48	0.64	-0.33
2	0.36	-0.12	0.01	-0.19	3.46	-1.30	0.22	-3.66
3	0.13	-0.08	-0.06	-0.16	1.24	-0.87	-0.91	-2.70
DG Fast	0.25	0.30	-0.09	0.06	2.02	3.32	-1.06	1.03
		h				t(h)		
DG Slow	0.25	0.34	0.18	-0.21	1.45	2.99	1.46	-2.63
2	0.50	0.12	0.41	-0.04	3.64	1.04	4.63	-0.63
3	0.47	0.05	0.28	-0.03	3.60	0.47	3.24	-0.42
DG Fast	0.65	0.31	0.17	-0.07	4.12	2.68	1.61	-0.86
		r				t(r)		
DG Slow	-0.03	0.36	0.09	0.47	-0.13	2.80	0.61	5.17
2	0.32	-0.05	0.47	0.51	2.06	-0.39	4.71	6.58
3	0.08	-0.18	0.15	0.45	0.54	-1.44	1.53	5.09
DG Fast	-0.19	0.55	0.05	0.53	-1.09	4.24	0.44	5.98
		c				t(c)		
DG Slow	-0.20	0.22	-0.36	0.43	-0.81	1.36	-2.04	3.83
2	0.07	0.12	-0.08	0.40	0.37	0.72	-0.64	4.22
3	-0.01	0.54	0.23	0.44	-0.06	3.41	1.98	4.07
DG Fast	-0.10	0.15	-0.43	0.14	-0.46	0.95	-2.88	1.31
		f				t(f)		
DG Slow	0.30	-0.12	-0.71	-0.49	2.18	-1.32	-6.96	-7.56
2	-0.02	-0.05	0.01	-0.17	-0.19	-0.56	0.08	-3.03
3	0.06	0.00	0.21	-0.10	0.60	0.01	3.01	-1.65
DG Fast	0.34	0.02	0.76	0.64	2.69	0.18	8.77	10.19

Table 4.12: **Regressions for 16 value-weighted *DG-Inv* portfolios: 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015. The data are sourced from Ken French's website.
At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three DG groups and three DY groups independently with breakpoints at the 30^{th} and 70^{th} percentile of that quarter. DG is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. DY is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of the fast DG group minus the returns of the slow DG group and *PMF* is the returns of prodigal DY group minus the returns of the frugal DY group.

The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + s_j SMB_t + h_j HML_t + r_j RMW_t + c_j CMA_t + f_j FMS_t + e_{j,t}$.

Inv	Low	2	3	High	Low	2	3	High
		a				t(a)		
DG Slow	0.14	-0.12	0.16	-0.21	0.68	-0.55	1.05	-0.96
2	-0.11	-0.15	-0.23	0.12	-0.62	-1.06	-1.33	0.58
3	-0.01	0.03	-0.04	-0.17	-0.04	0.18	-0.22	-0.83
DG Fast	-0.18	-0.13	-0.02	-0.04	-0.93	-0.76	-0.10	-0.22
		b				t(b)		
DG Slow	0.95	1.01	1.03	1.08	18.37	19.38	28.05	20.19
2	1.04	1.04	1.00	0.93	23.52	30.03	23.34	18.40
3	1.03	0.91	0.92	1.05	20.03	25.08	22.23	21.33
DG Fast	0.99	1.01	1.00	1.14	20.34	24.32	26.06	23.31
		s				t(s)		
DG Slow	-0.06	-0.03	-0.11	0.16	-0.81	-0.49	-2.30	2.15
2	-0.02	-0.18	-0.27	-0.31	-0.27	-3.84	-4.69	-4.45
3	-0.18	-0.08	-0.17	-0.17	-2.65	-1.64	-3.05	-2.59
DG Fast	0.05	0.08	-0.09	-0.09	0.76	1.44	-1.80	-1.38
		h				t(h)		
DG Slow	0.24	-0.22	0.20	0.14	2.65	-2.39	3.16	1.48
2	0.03	0.14	0.32	0.35	0.44	2.34	4.28	3.93
3	0.13	0.19	0.24	0.40	1.44	2.93	3.27	4.56
DG Fast	0.19	0.20	0.14	0.26	2.26	2.79	2.00	2.98
		r				t(r)		
DG Slow	0.11	0.18	0.16	0.45	1.10	1.74	2.22	4.25
2	0.49	0.27	0.28	0.21	5.63	3.89	3.28	2.08
3	0.05	0.28	0.27	0.26	0.45	3.83	3.30	2.64
DG Fast	0.16	0.31	0.24	0.13	1.63	3.83	3.22	1.33
		c				t(c)		
DG Slow	0.15	1.04	-0.24	-0.16	1.19	8.14	-2.64	-1.24
2	0.64	0.27	0.22	-0.22	5.93	3.24	2.15	-1.75
3	0.71	0.39	0.34	-0.04	5.62	4.40	3.39	-0.35
DG Fast	0.62	0.11	0.16	-0.31	5.17	1.10	1.75	-2.60
		f				t(f)		
DG Slow	-0.23	-0.43	-0.34	-0.39	-3.18	-5.86	-6.62	-5.19
2	0.11	-0.10	-0.13	0.03	1.71	-2.09	-2.09	0.45
3	0.02	0.00	0.05	0.16	0.22	0.08	0.79	2.22
DG Fast	0.15	0.39	0.20	0.59	2.20	6.69	3.76	8.53

Table 4.13: **Regressions for 16 value-weighted *DG-DY* portfolios: 1996-2015.**

$R_M - R_F$ is the value-weighted return on the market portfolio of all stocks in the sample of Fama and French (2015) minus the one-month Treasury bill rate. Factors *SMB* for *Size*, *HML* for *Book-to-Market*, *RMW* for *Operating Profitability* and *CMA* for *Investment* are all defined as described in Fama and French (2015. The data are sourced from Ken French's website.
At the second quarter of each calendar year in the period 1996-2015, stocks are allocated to three *DG* groups and three *DY* groups independently with breakpoints at the 30^{th} and 70^{th} percentile of that quarter. *DG* is defined as the growth rate of the average of the implied dividend in the third and the fourth quarter of the calendar year relative to the dividend paid in the second quarter. *DY* is the dividend paid in the second quarter relative to the average daily share price in the second quarter. Factor *FMS* is the returns of the fast *DG* group minus the returns of the slow *DG* group and *PMF* is the returns of prodigal *DY* group minus the returns of the frugal *DY* group.

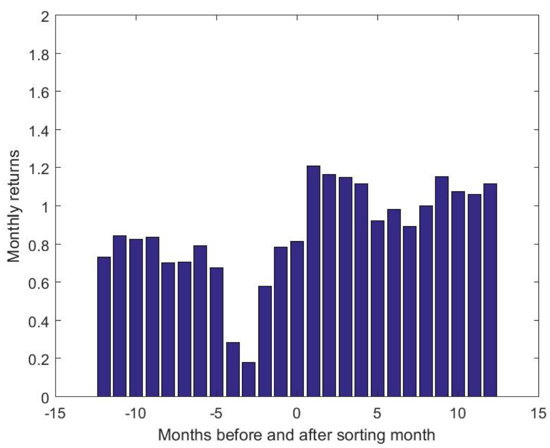
The regression equation is: $R_{j,t} - R_{F,t} = a_j + b_j(R_{M,t} - R_{F,t}) + s_j SMB_t + h_j HML_t + r_j RMW_t + c_j CMA_t + f_j FMS_t + e_{j,t}$.

DY	Low	2	3	High	Low	2	3	High
	a				t(a)			
DG Slow	0.22	-0.15	0.01	-0.42	0.82	-0.77	0.03	-2.02
2	-0.19	-0.25	-0.04	0.07	-0.82	-1.43	-0.32	0.32
3	0.08	-0.30	0.05	0.18	0.33	-1.78	0.29	1.12
DG Fast	-0.18	-0.08	-0.08	0.14	-1.39	-0.42	-0.39	0.70
	b				t(b)			
DG Slow	1.10	0.98	1.02	1.00	16.74	20.84	20.55	19.57
2	1.14	1.07	0.97	0.86	19.49	24.85	28.68	16.74
3	1.05	1.03	0.91	0.74	16.80	25.20	23.19	18.68
DG Fast	1.16	1.02	0.94	0.78	35.91	21.43	19.17	15.51
	s				t(s)			
DG Slow	0.08	0.03	-0.12	0.01	0.85	0.47	-1.86	0.07
2	-0.29	-0.07	-0.24	-0.10	-3.67	-1.19	-5.29	-1.42
3	-0.12	-0.12	-0.11	-0.12	-1.38	-2.26	-2.03	-2.15
DG Fast	-0.06	0.14	0.20	0.02	-1.44	2.13	2.95	0.36
	h				t(h)			
DG Slow	0.10	0.09	-0.01	0.40	0.87	1.09	-0.06	4.46
2	0.05	0.03	0.14	0.65	0.53	0.45	2.35	7.14
3	0.16	0.14	0.20	0.49	1.43	1.99	2.86	6.97
DG Fast	0.10	0.38	0.37	0.72	1.74	4.57	4.25	8.12
	r				t(r)			
DG Slow	0.09	0.43	0.47	0.27	0.72	4.60	4.76	2.71
2	0.13	0.63	0.37	0.03	1.15	7.39	5.56	0.27
3	0.08	0.33	0.32	0.12	0.64	4.03	4.07	1.52
DG Fast	0.19	0.31	0.35	0.15	2.93	3.27	3.59	1.52
	c				t(c)			
DG Slow	-0.42	0.20	0.56	0.38	-2.61	1.77	4.59	3.01
2	0.07	0.42	0.24	-0.04	0.48	3.99	2.87	-0.34
3	0.38	0.33	0.22	0.24	2.47	3.35	2.35	2.47
DG Fast	0.00	0.18	0.21	-0.22	0.05	1.56	1.72	-1.75
	f				t(f)			
DG Slow	-0.24	-0.37	-0.12	-0.31	-2.55	-5.51	-1.68	-4.32
2	-0.06	-0.06	0.00	-0.10	-0.72	-1.00	-0.06	-1.40
3	0.06	-0.05	0.06	-0.02	0.66	-0.91	1.06	-0.35
DG Fast	0.54	0.01	0.00	-0.03	11.89	0.09	-0.04	-0.46

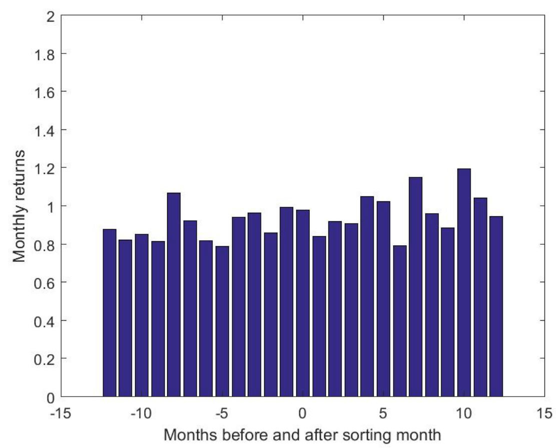
4.7 Figures

Figure 4.1: **Monthly value-weighted returns of portfolios of stock sorted at month = 0 by dividend growth implied 6 months ahead of the sorting month.**

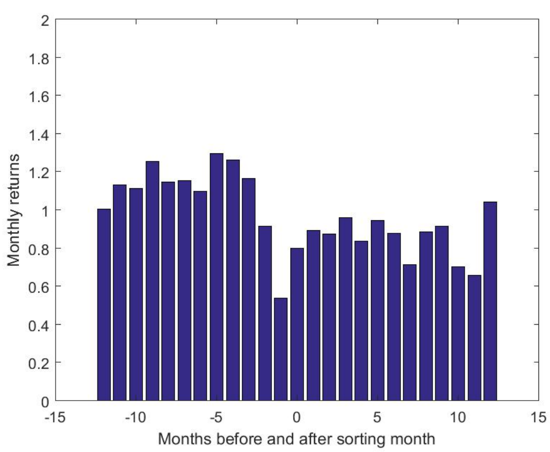
Returns refer to the 12 months before and after the sorting month. Stocks are sorted each month. Period: July 1997 – June 2015.



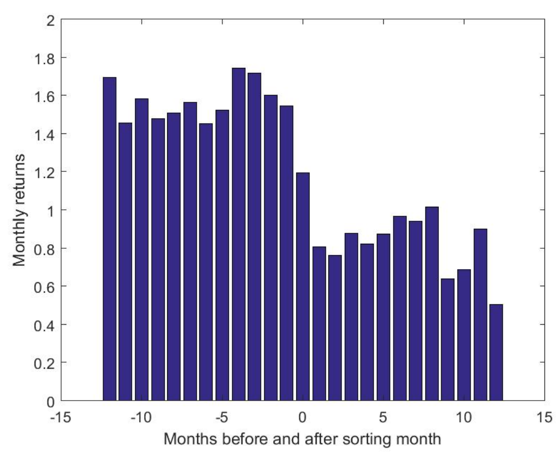
(a) 1^{st} quartile dividend growth



(b) 2^{nd} quartile dividend growth



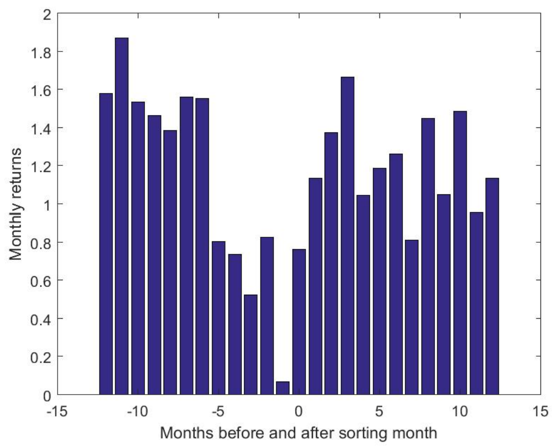
(c) 3^{rd} quartile dividend growth



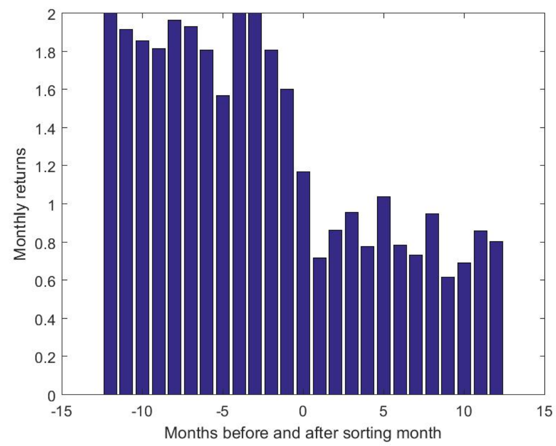
(d) 4^{th} quartile dividend growth

Figure 4.2: **Monthly value-weighted returns of portfolios of stock double-sorted at month = 0 by dividend yield and by dividend growth implied 6 months ahead of the sorting month.**

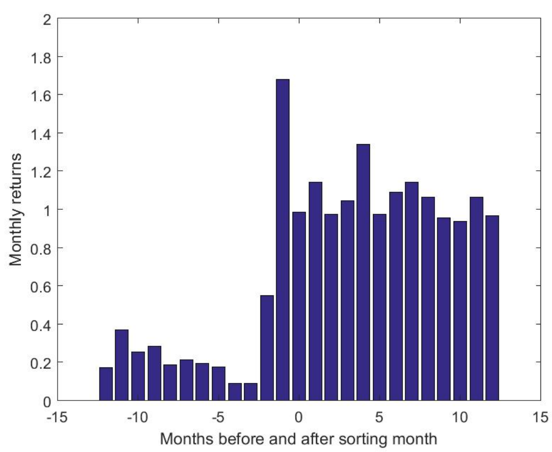
Low dividend yield and slow dividend growth refer to the first quartile and high dividend yield and fast dividend growth to the fourth quartile. Returns refer to the 12 months before and after the sorting month. Stocks are sorted each month. Period: July 1997 – June 2015.



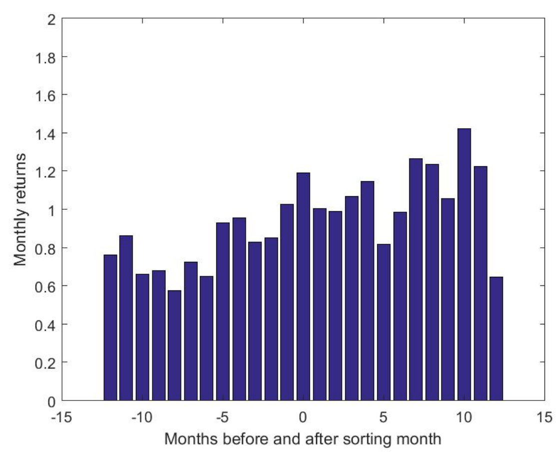
(a) Low dividend yield and slow dividend growth



(b) Low dividend yield and fast dividend growth



(c) High dividend yield and slow dividend growth



(d) High dividend yield and fast dividend growth

Bibliography

An, Byeong-Je, Andrew Ang, Turan Bali and Nusret Cakici, 2014, The joint cross section of stocks and options, *Journal of Finance* 69, 2279-2337.

Asness, Cliff, Andrea Frazzini and Lasse Pedersen, 2014, Quality minus junk, *AQR Working Paper.*

Avellaneda, Marco and Mike Lipkin, 2009, A dynamic model for hard-to-borrow stocks, *Working paper.*

Baker, Malcolm and Jeffrey Wurgler, 2004, A catering theory of dividends, *Journal of Finance*, 59, 1125-65.

Baker, Malcolm, Brock Mendel and Jeffrey Wurgler, 2016, Dividends as reference points: a behavorial signaling approach, *The Review of Financial Studies*, 29, 697-738.

Bansal, Ravi and Amir Yaron, 2004, Risks for the long run: a potential resolution of asset pricing puzzles, *Journal of Finance*, 56, 1481-1509.

Bekaert, Geert and Eric Engstrom, 2010, Inflation and the stock market: Understanding the "Fed Model", *Journal of Monetary Economics*, 57, 278-294.

Bilson, John, Sang Kang and Hong Luo, 2015, The term structure of implied dividend yields and expected returns, *Economics Letters*, 128, 9-13.

Binsbergen, Jules van and Ralph Koijen, 2017, The term structure of returns: facts and theory, *Journal of Financial Economics*, 124, 1-21.

Binsbergen, Jules van, Wouter Hueskes, Ralph Koijen and Evert Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503-519.

Binsbergen, Jules van, Michael Brandt and Ralph Koijen, 2012, On the timing and pricing of dividends, *American Economic Review*, 102, 1596-1618.

Bollen, Nicholas and Robert Whaley, 2004, Does net buying pressure affect the shape of implied volatility functions?, *Journal of Finance*, 59, 711-753.

Brav Alon, John Graham, Campbell Harvey and Roni Michaely, 2005, Payout policy in the 21st century. *Journal of Financial Economics*, 77, 483-527.

Brennan, Michael, 1998, Stripping the S&P 500 Index, *Financial Analysts Journal*, 54, 12-22.

Campbell, John and Robert Shiller, 1988, The dividend price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1, 195-228.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: a consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy*, 107, 205-251.

Campbell, John Y., Andrew W. Lo and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press.

Campbell, John Y. and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *The Review of Financial Studies*, 1, 195-228.

Cejnek, Georg and Otto Randl, 2016, Risk and return of short-duration equity investments, *Journal of Empirical Finance*, 36, 181-198.

Cochrane, John H., 2011, Presidential Address: Discount Rates, *The Journal of Finance*, 66, 4, 1047-88.

Conover, Mitchell, Gerald Jensen and Marc Simpson, 2016, What difference do dividends make?, *Financial Analysts Journal*, 6, 28-40.

Cox, John, Stephen Ross and Mark Rubinstein, 1979, Option pricing: a simplified approach, *Journal of Financial Economics*, 7, 229-263.

Cremers, Martijn and David Weinbaum, 2010, Deviations from put-call parity and stock return predictability, *Journal of Financial and Quantitative Analysis*, 45, 335-367.

Dai, Qiang and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *The Journal of Finance* 55, 5, 1943-1978.

Fama, Eugene and James MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy*, 81, 607-636.

Fama, Eugene and Kenneth French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics*, 22, 3-25.

Fama, Eugene and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, 33, 3-56.

Fama, Eugene and Kenneth French, 2015, A five-factor asset pricing model, *Journal of Financial Economics*, 116, 1-22.

Fodor, Andy, David Stowe and John Stowe, 2017, Option implied dividends predict dividend cuts: evidence from the financial crisis, *Journal of Business Finance & Accounting*, 44, 755-779.

Fuller, Kathleen and Michael Goldstein, 2011, Do dividends matter more in declining markets?, *Journal of Corporate Finance*, 17, 457-473.

Gârleanu, Nicolae, Lasse Pedersen and Allen Poteshman, 2009, Demand-based option pricing, *Review of Financial Studies*, 22, 4259-99.

Giglio, Stefano, Matteo Maggiori and Johannes Stroebel, 2015, Very long-run discount rates, *Quarterly Journal of Economics*, 130, 1-53.

Giglio, Stefano and Bryan Kelly, 2017, Excess volatility: Beyond discount rates, *Quarterly Journal of Economics*, *forthcoming.*

Gibbons, Michael, Stephen Ross and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica*, 57, 1121-1152.

Golez, Benjamin, 2014, Expected returns and dividend growth rates implied in derivative markets, *The Review of Financial Studies*, 27, 790-822.

Grullon G., R. Michaely and R. Swaminathan, 2002, Are dividend changes a sign of firm maturity?, *Journal of Business*, 75, 387-424.

Jegadeesh, Narasimhan and George G. Pennacchi, 1996, The behavior of interest rates implied by the term structure of eurodollar futures, *Journal of Money, Credit and Banking*, 28, 426-46.

Harvey, Campbell and Robert Whaley, 1992, Market volatility prediction and the efficiency of the S&P 100 index option market, *Journal of Financial Economics*, 31, 43-73.

Kalay, Avner, Oğuzhan Karakaş and Shagun Pant, 2014, The market value of corporate votes; Theory and evidence from option prices, *Journal of Finance*, 69, 1235-71.

Keim, Donald, 1988, Stock market regularities: A synthesis of the evidence and explanations, in: E. Dimson. ed., *Stock market anomalies* (Cambridge University Press. Cambridge).

Larkin, Yelena, Mark Leary and Roni Michaely, 2016, Do investors value dividend-smoothing stocks differently?, *Management Science*, 1, 1-23.

Lintner, John, 1956, Distribution of incomes of corporations among dividends, retained earnings, and taxes, *American Economic Review*, 46, 97-113.

Maio, Paulo and Pedro Santa-Clara, 2015, Dividend yields, dividend growth and return predictability in the cross section of stocks, *Journal of Finance*, 50, 33-60.

Manly, Richard and Christian Mueller-Glissmann, 2008, The market for dividends and related investment strategies, *Financial Analysts Journal*, 64, 17-29.

Miller, Merton and Franco Modigliani, 1961, Dividend policy, growth and the valuation of shares, *Journal of Business*, 34, 411-433. Mixon, Scott and Esen Onur, 2017, Dividend swaps and dividend futures: State of play, *Journal of Alternative Investments*, 19, 27-39. OptionMetrics, 2011, *File and Data Reference Manual*, version 3.0.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics*, 108, 1-28.

Shiller, Robert, 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, *The American Economic Review*, 71, 421-36.

Suzuki, Masataka, 2014, Measuring the fundamental value of a stock index through dividend futures prices, working paper.

Wilkens, Sascha and Jens Wimschulte, 2010, The pricing of dividend futures in the european market: a first empirical analysis, *Journal of Derivatives & Hedge Funds* 16, 2, 136-143.