

MFI Working Paper Series
No. 2010-010

On the Timing and Pricing of Dividends

Jules H. van Binsbergen
Northwestern Kellogg, Stanford GSB and NBER

Michael W. Brandt
Duke University and NBER

Ralph S.J. Koijen
Chicago Booth School of Business and NBER

October 2010



1126 East 59th Street
Chicago, Illinois 60637

T: 773.702.7587
F: 773.795.6891
mfi@uchicago.edu

On the Timing and Pricing of Dividends*

Jules H. van Binsbergen[†]

Northwestern Kellogg
and Stanford GSB
and NBER

Michael W. Brandt[‡]

Duke University
and NBER

Ralph S.J. Koijen[§]

Chicago Booth
and NBER

This version: October 2010

Abstract

We recover prices of dividend strips on the aggregate stock market using data from derivatives markets. The price of a k -year dividend strip is the present value of the dividend paid in k years. The value of the stock market is the sum of all dividend strip prices across maturities. We study the properties of strips and find that expected returns, Sharpe ratios, and volatilities on short-term strips are higher than on the aggregate stock market, while their CAPM betas are well below one. Short-term strip prices are more volatile than their realizations, leading to excess volatility and return predictability.

*First draft: June 2009. We are grateful to Ravi Bansal, Robert Battalio, Alessandro Beber, John Campbell, John Cochrane, George Constantinides, Zhi Da, Peter DeMarzo, Joost Driessen, Darrell Duffie, Pengjie Gao, John Heaton, Xavier Gabaix, Nicolae Garleanu, Lars Hansen, Otto van Hemert, Mark Hendricks, Dirk Jenter, Martin Lettau, Lars Lochstoer, Hanno Lustig, Pascal Maenhout, Toby Moskowitz, Christian Mueller-Glissmann, Stefan Nagel, Stavros Panageas, Dimitris Papanikolaou, Lubos Pastor, Monika Piazzesi, Lasse Pedersen, Francisco Perez-Gonzalez, Anamaria Pieschacon, Sergio Rebelo, Ken Singleton, Rob Stambaugh, Sheridan Titman, Stijn Van Nieuwerburgh, Adrien Verdelhan, Pietro Veronesi, Rob Vishny, Evert Vrugt, Jessica Wachter, Jeff Wurgler, Amir Yaron, Motohiro Yogo and seminar participants APG Investments, AQR Capital, Berkeley Haas, Boston College, Brigham Young University, CEPR Gerzensee meetings, Chicago Booth, EFA 2010 meetings, Maastricht University, NBER Asset Pricing Meetings, Notre Dame, SITE 2010 meetings, Stanford GSB, Tilburg University, UT Austin, and the Empirical Asset Pricing Retreat at the University of Amsterdam for useful comments and discussions. We thank the Center for Research in Security Prices (CRSP) for generous financial support.

[†]j-vanbinsbergen@kellogg.northwestern.edu, (847) 491-3838

[‡]mbrandt@duke.edu, (919) 660-1948

[§]ralph.koijen@chicagobooth.edu, (773) 834-4199. Koijen is also associated with Netspar (Tilburg University)

A central question in economics is how to discount future cash flows to obtain today's value of an asset. For instance, total wealth is the price of a claim to all future consumption (Lucas (1978)). Similarly, the value of the aggregate stock market equals the sum of discounted future dividend payments (Gordon (1962)). The majority of the equity market literature has focused on the dynamics of the value of the aggregate stock market. However, in addition to studying the value of the sum of discounted dividends, exploring the properties of the individual terms in the sum, also called dividend strips, provides us with a lot of information about the way stock prices are formed. Analogously to zero-coupon bonds, which contain information about discount rates at different horizons for fixed income securities, having information on dividend strips informs us about discount rates of risky cash flows at different horizons. Studying dividend strips can therefore improve our understanding of investors' risk preferences and the endowment or technology process in macro-finance models. This paper is the first to empirically measure the prices of dividend strips. Our approach only requires no-arbitrage relations and does not rely on a specific model.

With this approach, we shed new light on the composition of the equity risk premium. The equity premium puzzle, identified by Mehra and Prescott (1985), Hansen and Singleton (1982) and Hansen and Singleton (1983), states that, for plausible values of the risk aversion coefficient, the difference of the expected rate of return on the stock market and the riskless rate of interest is too large, given the observed small variance of the growth rate in per capita consumption. When decomposing the index into dividend strips, a natural question that arises is whether dividends at different horizons contribute equally to the equity risk premium or whether either short or long-term dividends contribute proportionally more than the other. We find that short-term dividends have a higher risk premium than long-term dividends, whereas leading asset pricing models predict the opposite.

More specifically, we decompose the S&P500 index, which is a broad US equity index, into a portfolio of short-term dividend strips, which we call the short-term asset, and a portfolio of long-term dividend strips, which we call the long-term asset. The short-term asset entitles the holder to the realized dividends of the index for a period of up to three years. Our main focus is to compare the properties of the short-term asset to those of the index, both empirically and theoretically.

In the absence of arbitrage opportunities, there exists a stochastic discount factor M_{t+1} that can be used to discount future cash flows. More formally, the value of an equity index

S_t is given by the discounted value of its dividends $(D_{t+i})_{i=1}^{\infty}$:

$$S_t = \sum_{i=1}^{\infty} E_t(M_{t:t+i}D_{t+i}),$$

where $M_{t:t+i} = \prod_{j=1}^i M_{t+j}$ is the product of stochastic discount factors. We can decompose the stock index as:

$$S_t = \underbrace{\sum_{i=1}^T E_t(M_{t:t+i}D_{t+i})}_{\text{price of the short-term asset}} + \underbrace{\sum_{i=T+1}^{\infty} E_t(M_{t:t+i}D_{t+i})}_{\text{price of the long-term asset}},$$

where the short-term asset is the price of all dividends up until time T , and the long-term asset is the price of the remaining dividends. To compute the price of the short-term asset, we use a newly-constructed data set on options and futures on the S&P500 index.

We document five properties of the short-term asset in comparison with the aggregate stock market. First, expected returns, volatilities, and Sharpe ratios on the short-term asset are on average higher. Second, the slope coefficient (or beta) in the Capital Asset Pricing Model (Sharpe (1964) and Lintner (1965)) of the short-term asset is 0.5. Third, the CAPM alpha of short-term asset returns is 9% per year, which suggests that the short-term asset has a substantially higher expected return than predicted by the CAPM. Fourth, the prices of the short-term asset are more volatile than their realizations, pointing to excess volatility. Fifth, the returns on the short-term asset are strongly predictable.

Our results have several important additional implications for empirical and theoretical asset pricing. First, since Shiller (1981) pointed out that stock prices are more volatile than subsequent dividend realizations, the interpretation has been that discount rates fluctuate over time and are persistent. The long duration of equity makes prices very sensitive to small persistent movements in discount rates, thereby giving rise to excess volatility in prices and returns. We show, however, that the same phenomenon arises for the short-term asset. This suggests that a complete explanation of excess volatility must be able to generate excess volatility both for the aggregate stock market and for the short-term asset. The excess variation in prices also suggests that discount rates fluctuate, and we should therefore find that prices, normalized by some measure of dividends, forecast returns on the short-term asset. We show that this is indeed the case, leading to the fifth property. Second, the first four properties we document, combined with the fact that the CAPM alphas are virtually unaffected if we include additional well-known asset pricing factors such as size or value, suggest that short-term assets are potentially important new

test assets that may be useful in cross-sectional asset pricing tests.¹

To provide a theoretical benchmark for our results, we compute dividend strips in several leading asset pricing models. Recent consumption-based asset pricing models have made substantial progress in explaining many asset pricing puzzles across various markets. Even though such models are not often used to study the pricing of dividend strips, they do have theoretical predictions about the values of these securities, which we explore in this paper.² We focus on the external habit formation model of Campbell and Cochrane (1999), the long-run risks model of Bansal and Yaron (2004), and the variable rare disasters model of Gabaix (2009), which builds upon the work of Barro (2006) and Rietz (1988). We find that both the long-run risks model and the external habit formation model predict that expected returns, volatilities, and Sharpe ratios of short-term dividend strips are lower than those of the aggregate market. Further, the risk premium on short-term dividend strips in those models are near zero. In the rare disasters model, the volatilities and Sharpe ratios of short-term dividend strips are lower than the aggregate market. Expected returns on the other hand are equal across all maturities of dividend strips, and therefore also equal to those on the aggregate market. Our results suggest that risk premia on the short-term asset are higher than predicted by leading asset pricing models.

Our paper relates to Lettau and Wachter (2007) and Croce, Lettau, and Ludvigson (2009). Lettau and Wachter (2007) argue that habit formation models as in Campbell and Cochrane (1999), generate higher expected returns for long-term dividend strips as shocks to the discount factor are priced. Firms with long-duration cash flows have a high exposure to such shocks and should therefore have a higher risk premium than firms with short-duration cash flows. If one adheres to the view that value firms have short-duration cash flows and growth firms have long-duration cash flows, this implies that there is a growth premium, not a value premium (see also Santos and Veronesi (2006)). Lettau and Wachter (2007) propose a reduced-form model that generates higher expected returns for short-term dividend strips. They illustrate the correlation structure between (un)expected cash flow shocks and shocks to the price of risk and stochastic discount factor that is sufficient to generate a value premium in their model. Croce, Lettau, and Ludvigson (2009) argue that the long run risk model as proposed by Bansal and Yaron (2004) also generates higher risk premia for long-term dividend strips. However, Croce, Lettau, and Ludvigson (2009) also show that if the agents cannot distinguish between short-term and

¹Lewellen, Nagel, and Shanken (2009) argue that the standard set of test assets has a strong factor structure, and that it would be valuable to have new test assets.

²Notable exceptions are Lettau and Wachter (2007) and Croce, Lettau, and Ludvigson (2009).

long-term shocks, risk premia on short-term dividend strips can be higher.

Studying the properties of short-term assets is not only interesting from an academic perspective. Recently, dividend strips, futures, and swaps have received a lot of attention in the practitioners' literature.³ First, several banks offer dividend swaps on a range of stock indices. With such a contract, the dividend purchaser pays the market-implied dividend level (the fixed leg). The counterparty, with a long position in the equity index, pays the realized dividend level (the variable leg). Secondly, for the S&P500, Standard and Poor's has introduced the S&P500 Dividend Index, which is a running total of dividend points. The index is reset to zero after the close on the third Friday of the last month of every calendar quarter, to coincide with futures and options expirations. It measures the total dividend points of the S&P500 index since the previous reset date and is used by derivative traders to hedge their dividend positions. Third, from 1982 to 1992, investors could invest in derivatives at the American Stock Exchange (AMEX) that split the total return on individual stocks into a price appreciation part and a dividend yield part. Also in the UK, split-capital funds offered financial instruments that separate investment in a fund's price appreciation and its dividend stream in the late 90s. Finally, Wilkens and Wimschulte (2009) discuss the European market of dividend futures that started mid-2008.

1 The market for dividends

There are two ways to trade dividends in financial markets. First, dividend strips can be replicated using options and futures data, which is the approach we follow in this paper. In 1990, the Chicago Board Options Exchange (CBOE) introduced Long-Term Equity Anticipation Securities (LEAPS), which are long-term call and put options. The owner of a call (put) option has the right to purchase (sell) the stock index at maturity at a predetermined price X . LEAPS have maturities up to three years. The maximum maturity of LEAPS for the sample period in our data set is displayed in Figure 1. The set of maturities of these claims is not constant and varies depending on the issuing cycle. On average, there are around six maturities greater than three months available at any particular time, spaced closer together for shorter maturities, and further apart for longer maturities.

To compute dividend strip prices from options data, we only require the absence of arbitrage opportunities. Under this condition, put-call parity for European options holds

³See for example Brennan (1998), Manley and Mueller-Glissmann (2008) and Barclays Capital (2008).

(Stoll (1969)):

$$c_{t,T} + Xe^{-r_{t,T}(T-t)} = p_{t,T} + S_t - \mathcal{P}_{t,T}, \quad (1)$$

where $p_{t,T}$ and $c_{t,T}$ are the prices of a European put and call option at time t , with maturity T , and strike price X . $r_{t,T}$ is the interest rate between time t and T . We use the symbol $\mathcal{P}_{t,T}$ to denote the value of the short-term asset, which we defined in the introduction as:

$$\mathcal{P}_{t,T} \equiv \sum_{i=1}^T E_t(M_{t,t+i}D_{t+i}). \quad (2)$$

We can rewrite (1) to obtain the price of the short-term asset:

$$\mathcal{P}_{t,T} = p_{t,T} - c_{t,T} + S_t - Xe^{-r_{t,T}(T-t)}. \quad (3)$$

This parity relation shows that purchasing the short-term asset is equivalent to buying a put option, writing a call option, buying the stocks in the index, and borrowing cash.

A second way to synthetically create the short-term asset is by using futures contracts. The owner of the futures contract agrees to purchase the stock index for a predetermined price, $F_{t,T}$, at maturity. Absence of arbitrage opportunities implies the cost-of-carry formula for equity futures:

$$\mathcal{P}_{t,T} = S_t - e^{-r_{t,T}(T-t)}F_{t,T}. \quad (4)$$

Hence, buying the short-term asset is the same as buying the stock index and selling a position in a futures contract. In both cases, the key insight we exploit is that payoffs of derivatives contracts are based on the ex-dividend price, which allows us to recover the price of the short-term asset.

In addition to computing the prices of short-term asset using equity derivatives, it is also possible to trade dividends directly via dividend derivatives such as dividend swaps, dividend futures, and options on dividends. Most of these transactions take place in over-the-counter (OTC) markets, but several exchange-traded products have been introduced recently. For instance, the CBOE introduced options on S&P500 index dividends in May 2010. This development follows the introduction of an array of dividend derivatives at the Eurex. In June 2008, the Eurex introduced dividend futures on the Dow Jones EURO STOXX 50 Index,⁴ and in February 2010, futures are available on five different

⁴See http://www.eurexchange.com/download/documents/publications/index_dividend_swaps_1_en.pdf for more information.

indices.⁵ In addition, the Eurex now introduced dividend futures on the constituents of the Dow Jones EURO STOXX 50 in January 2010. As measured by open interest, the size of the market for index dividend futures is already 20% of the size of the market for index futures, illustrating the rapid developments of dividend trading.⁶ The advantage of dividend derivatives is that the maturities available are longer (up to 15 years).

2 Dividend prices in a Lucas economy

To illustrate how dividend prices can be understood in an equilibrium asset pricing model, we compute the k -period short-term asset, which is the sum of the first k dividend strips, in the consumption CAPM of Lucas (1978). In Section 5, we extend these results to more recent consumption-based asset pricing models. Consumption growth, $\Delta c_t = \log C_t - \log C_{t-1}$, is assumed to be i.i.d.:

$$\Delta c_{t+1} = g + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \sigma^2), \quad (5)$$

where g is the constant average growth rate and σ the growth rate volatility. The representative agent has time-separable Constant Relative Risk Aversion (CRRA) utility:

$$E_t \left(\sum_{s=0}^{\infty} \beta^s \frac{C_{t+s}^{1-\gamma}}{1-\gamma} \right), \quad (6)$$

where beta is the subjective discount factor and γ is the coefficient of relative risk aversion, which for CRRA utility is equal to the inverse of the intertemporal elasticity of substitution. The one-period stochastic discount factor is in this case given by:

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}, \quad (7)$$

and the k -period stochastic discount factor equals:

$$M_{t:t+k} = \beta^k \left(\frac{C_{t+k}}{C_t} \right)^{-\gamma}. \quad (8)$$

Assuming that the aggregate dividend equals aggregate consumption, $C_t = D_t$, the price of the k -period short-term asset, which is the sum of the first k dividend strips, is given

⁵A more detailed description can be found at:
http://www.eurexchange.com/download/documents/publications/Eurex.Produkte_2010_en.pdf.

⁶The dollar volume averages to \$33.4 Billion in 2009, see:
<http://www.reuters.com/article/idUSLDE60A10020100111>.

by:

$$\mathcal{P}_{t,t+k} = \sum_{s=1}^k \exp \left(s \log \beta + s(1 - \gamma)g + \frac{s}{2}(1 - \gamma)^2 \sigma^2 \right). \quad (9)$$

3 Data and dividend strategies

3.1 Data sources

We measure dividend prices using put-call parity in equation (1), which is a no-arbitrage relationship. To compute dividend prices as accurately as possible, we record each of the components in equation (1) within the same minute of the last trading day of each month. To this end, we use data from four different sources. First, we use a new data set provided by the CBOE containing intra-day trades and quotes on S&P500 index options between January 1996 and October 2009. The data contains information about all option contracts for which the S&P500 index is the underlying asset. Second, we obtain minute-level data between January 1996 and October 2009 of the index values and futures prices of the S&P500 index from Tick Data Inc. Third, the interest rate is calculated from a collection of continuously-compounded zero-coupon interest rates at various maturities and provided by IvyDB (OptionMetrics). This zero curve is derived from LIBOR rates from the British Bankers' Association (BBA) and settlement prices of Chicago Mercantile Exchange (CME) Eurodollar futures.⁷ For a given option, the appropriate interest rate corresponds to the zero-coupon rate that has a maturity equal to the option's expiration date. We obtain these by linearly interpolating between the two closest zero-coupon rates on the zero curve.⁸ Fourth, to compute daily dividends, we obtain daily return data with and without distributions (dividends) from S&P index services. Cash dividends are then computed as the difference between these two returns, multiplied by the lagged value of the index.

3.2 Data selection and matching

As mentioned before, our data allows us to match call and put option prices and index values within a minute interval.⁹ We therefore select option quotes for puts and calls

⁷We use data from Bloomberg to replicate the OptionMetrics yield curves and obtain very similar results.

⁸Alternative interpolation schemes give the same results at the reported precision.

⁹Using closing prices from OptionMetrics for all quantities does not guarantee that the index value and option prices are recorded at the same time and induces substantial noise in our computations, see also Constantinides, Jackwerth, and Perrakis (2009)). For instance, the options exchange closes 15 minutes later than the equity exchange, which leads to much wider bid-ask spreads in options markets

between 10am and 2pm that trade within the same minute, and match these quotes with the tick-level index data, again within the minute. Changing the time interval to either 10am to 11am or 1pm to 2pm has no effect on our results, as we demonstrate in Section 6.

We compute dividend prices at the last trading day of the month. For a given strike price and maturity, we collect all quotes on call option contracts and find a quote on a put option contract, with the same strike price and maturity, that is quoted closest in calendar time. Of the resulting matches, we keep the matches for each strike and maturity that trade closest to each other in time. This typically results in a large set of matches for which the quotes are recorded within the same second of the day, making the matching procedure as precise as possible. For each of these matches, we use the put-call parity relation to calculate the price of the dividend strip. We use mid quotes, which are the average of the bid and the ask quotes. We then take the median across all prices for a given maturity, resulting in the final price we use in our analysis. By taking the median across a large set of dividend prices, we mitigate potential issues related to measurement error or market microstructure noise.

To illustrate the number of matches we find for quotes within the same second, Figure 2 reports the average number of quotes per maturity during the last trading day of the month in a particular year. We focus on option contracts with a maturity between 1 and 2 years. The number of quotes increases substantially over time, presumably as a result of the introduction of electronic trading. However, even in the first year of our sample, we have on average nearly a thousand matches per maturity on a given trading day for options with maturities between 1 and 2 years.

3.3 Dividend strategies

Holding a long position in the short-term asset has the potential disadvantage that a long position in the index is required (see equations (1) and (4)). As index replication is not costless, we also consider investing in a so-called dividend steepener. This asset entitles the holder to the dividends paid out between period T_1 and T_2 , $T_1 < T_2$. The price of the dividend steepener is given by:

$$\begin{aligned} \mathcal{P}_{t,T_1,T_2} &= \mathcal{P}_{t,T_2} - \mathcal{P}_{t,T_1} \\ &= p_{t,T_2} - p_{t,T_1} - c_{t,T_2} + c_{t,T_1} - X \left(e^{-r_{t,T_2}(T_2-t)} - e^{-r_{t,T_1}(T_1-t)} \right). \end{aligned} \quad (10)$$

during this period. OptionMetrics reports the last quote of the trading day, which is likely to fall in this 15-minute interval. We reproduced our results using OptionMetrics data, and find similar results for average returns, but the volatility of prices and returns is substantially higher.

This strategy can be interpreted as buying the first T_2 periods of dividends and selling the first T_1 periods of dividends, which results in a long position in the dividends paid out between periods T_1 and T_2 . This strategy does not involve any dividend payments until time T_1 . Replicating this asset does not require a long position in the index and simply involves buying and writing two calls and two puts, in addition to a cash position. The dividend steepener is also interesting to study as a macro-economic trading strategy, as it can be used to bet on the timing of a recovery of the economy following a recession. During severe recessions, firms slash dividends and increase them when the economy rebounds. By choosing T_1 further into the future, investors bet on a later recovery.¹⁰

By applying the cost-of-carry formula for equity index futures to two different maturities T_1 and T_2 , where $T_1 < T_2$, the price of the dividend steepener can also be computed as:

$$\mathcal{P}_{t,T_1,T_2} = e^{r_t,T_1(T_1-t)} F_{t,T_1} - e^{r_t,T_2(T_2-t)} F_{t,T_2}.$$

In this case, the steepener only involves two futures contracts and does not require any trading of the constituents of the index. By no-arbitrage, the prices implied by equity options and futures need to coincide. Since LEAPS have longer maturities than index futures, we rely on options for most of our analysis. For the maturities for which both futures and options data is available, we show in Section 6.2 that the prices obtained from both markets are close, and our main findings are unaffected by using either options or futures.

Apart from reporting dividend prices, we also implement two simple trading strategies. The first trading strategy goes long in the short-term asset. The monthly return series on this strategy is given by:

$$R_{1,t+1} = \frac{\mathcal{P}_{t+1,T-1} + D_{t+1}}{\mathcal{P}_{t,T}}. \quad (11)$$

The monthly returns series on the second trading strategy, which is the dividend steepener, is given by:

$$R_{2,t+1} = \frac{\mathcal{P}_{t+1,T_1-1,T_2-1}}{\mathcal{P}_{t,T_1,T_2}}, \quad (12)$$

which illustrates that this trading strategy does not return any dividend payments until time T_1 . Further details on the implementation of these strategies can be found in Appendix A.

¹⁰See also “Dividend Swaps Offer Way to Pounce on a Rebound,” Wall Street Journal, April 2009.

4 Main empirical results

In this section, we document the properties of the prices and returns on the short-term asset. First, we study dividend prices in Section 4.1. In Section 4.2, we study the properties of dividend returns. In the remaining subsections, we study excess volatility of dividend strip prices, and the predictability of the return series that we compute.

4.1 Properties of dividend prices

Figure 3 displays the prices of the first 6, 12, 18, and 24 months of dividends during our sample period. To obtain dividend prices at constant maturities, we interpolate over the available maturities. For instance, in January 1996, the price of the dividends paid out between that date and June 1997 is \$20. As expected, the prices monotonically increase with maturity. Violations of this condition would imply the existence of arbitrage opportunities. Further, the dividend prices for all maturities, drop during the two NBER recessions in our sample period, which occur between March and November 2001 and between December 2007 and June 2009. This is to be expected, as during recessions expected growth of dividends drops and discount rates on risky cash flows are likely to increase. This effect is more pronounced for the 24-month price. The 6-month price is less volatile.

As dividend prices are non-stationary over time, it is perhaps more insightful to scale dividend prices by the value of the S&P500 index. In Figure 4, we plot the prices of the first 6, 12, 18, and 24-month dividend prices as a fraction of the index value. The ratios are highly correlated. They drop between 1997 and 2001, and slowly increase afterwards. When comparing Figure 4 with Figure 3, one interesting observation is that during the recession of 2001, both the ratio and the level of dividend prices drop, whereas for the recent recession the level of dividend prices drops, but not by as much as the index level. This leads to an increase in the ratio. One interpretation of this finding is that the most recent recession has a longer-lasting impact than the recession in 2001. The index level is more sensitive to revisions in long-term cash flow (dividend) expectations and discount rates than the short-term asset, see Shiller (1981) and Lettau and Wachter (2007). A more severe recession can therefore lead to a decline in the index value that is proportionally larger than the decline in the price of the short-term asset.

4.2 Properties of dividend returns

We now report the return characteristics of the two investment strategies. Figure 5 and Figure 6 plot the time series of monthly returns on the two trading strategies. Figure 7 and Figure 8 display the histogram of returns. The two trading strategies are highly positively correlated, with a correlation coefficient of 92.1%. Panel A of Table 1 lists the summary statistics alongside the same statistics for the S&P500 index for the full sample period. Both dividend strategies have a high monthly average return equal to 1.16% (annualized 14.8%) for trading strategy 1 and 1.12% (annualized 14.3%) for trading strategy 2 (the steeper). Over the same period, the average return on the S&P500 index was 0.56% (annualized 6.93%). The higher average returns also come with a higher level of volatility than the S&P500 index, with monthly return volatilities of 7.8% for strategy 1 and 9.6% for strategy 2. Over the same period the monthly volatility of the return on the S&P500 index equals 4.7%. Despite the higher volatility, the dividend strategies result in substantially higher Sharpe ratios, defined as the ratio of the average monthly excess returns and the volatility of the excess returns (Sharpe (1966)). The Sharpe ratios of the dividend strategies are about twice as high as the Sharpe ratio of the S&P500 index. Duffee (2010) shows that Sharpe ratios are lower for Treasury bonds with longer maturities. We document a similar property in equity markets; Sharpe ratios are higher for dividend claims with shorter maturities.

We find that the volatility of dividend returns is lower in the second part of our sample. To further analyze the volatility of dividend returns, we estimate a GARCH(1,1) model (Bollerslev (1986)) for each return series and for the returns on the S&P500 index. As the returns on the dividend strategies are predictable (see Section 4.4), we include an AR(1)-term in the mean equation. In Figure 9, we show that the volatility of dividend returns and the index broadly follow the same pattern. The correlation between the volatility of the dividend returns of strategy 1 and the S&P500 index is 0.55. Table 2 reports the estimates of the GARCH(1,1)-specification, illustrating that the parameters of the volatility equations are very similar as well.

To further assess the difference in volatility between the early and the late part of our sample, Table 1 also presents summary statistics for the period before January 2003 (Panel B) and for the period afterwards (Panel C). We are mostly interested in the average return and volatility of the dividend strategies relative to the same statistics of the S&P500 index. Consistently across both sample periods, the average return and the volatility on the dividend strategies is higher than the average return and volatility of the S&P500 index. The volatility of the dividend strategies is high in both sub-periods, even though the volatilities in the more recent sample are closer to the levels of volatility

that we record for the index. The Sharpe ratios of the dividend strategies are comparable across subperiods, and always higher than the ones of the S&P500 index. Overall, the conclusions we draw from the full sample are consistent with our findings in both sub-samples.

The high average returns on short-maturity dividend strips may be due to exposures to systematic risk factors that are priced in financial markets. To verify whether well-known empirical asset pricing models, such as the CAPM and the Fama and French three-factor model (Fama and French (1993)), can explain the average returns on short-maturity dividend strips, we regress excess returns of both strategies on (i) the excess return on the market (mktrf) and (ii) on Fama and French’s three factors (mktrf, hml and smb). The three Fama French factors are (i) the excess returns on the market (mktrf), (ii) the returns on a portfolio that goes long in stocks with a high book-to-market ratio, also called value stocks, and short in stocks with a low book-to-market ratio, also called growth stocks (hml), and (iii) the returns on a portfolio that goes long in small stocks and short in large stocks (smb). See Fama and French (1993) for further details.

Table 3 presents OLS regressions of the returns of the two trading strategies in excess of the one-month short rate on a constant and the market portfolio’s returns in excess of the one-month short rate (the CAPM). We find that both dividend strategies have a CAPM beta (or slope) of around 0.5. Secondly, R^2 values of the regression are rather low. The intercept (also called CAPM alpha) of the regression equals 0.73% for the first dividend strategy and 0.69% for the second strategy (the steeper), which in annualized terms corresponds to 9.1% and 8.6% respectively. Despite these economically significant intercepts, the results are not statistically significant at conventional levels due to the substantial volatility of these two return strategies and the rather short time series that is available for dividend returns. Generally the p-values vary between 10% and 20%, using Newey-West standard errors. When including an AR(1) term in the regression, to account for the negative autocorrelation (predictability) in returns, the standard errors are somewhat smaller.

In Table 4 we repeat the analysis of Table 3, but instead of using excess returns on the aggregate market, we use as the regressor the excess returns on the S&P500 index. The table shows that the results are nearly identical to those of Table 3, with betas of around 0.5 and monthly intercepts (alphas) of around 0.7%.

Table 5 presents regression results for the three-factor model, in which we also include the AR(1) term in the second and fourth column. The market beta, that is the slope coefficient on mktrf, is unaffected by the additional factors and is estimated between 0.48

and 0.60, depending on the strategy and specification. We find positive loadings on the value factor (hml), which seems consistent with duration-based explanations of the value premium. An important element of this theory is that the portfolio of value firms have cash flows that are more front-loaded than the cash flows of the portfolio of growth firms. As such, this theory suggests that the short-term asset loads more on value firms than on growth firms, which corresponds to a positive coefficient on the book-to-market factor. The coefficient on the size portfolio switches sign depending on the specification, and has very low significance.

Perhaps most interestingly, the intercepts (or alphas) are hardly affected by including additional factors; monthly alphas are estimated between 0.53% and 0.68%. These results suggest that the short-term asset has rather high expected returns that cannot be explained easily by standard empirical asset pricing models. As a comparison, the monthly value premium, which is defined as the average return on the hml factor, over our sample equals 0.35% which corresponds to 4.3% annualized. As the alphas cannot be explained by the Fama French model, the high expected returns that we find for our dividend strategies are not (solely) driven by value firms or smaller firms in the S&P500 index.

In Table 6 we repeat the analysis of Table 5, but instead of using the three Fama and French factors (that are based on all firms in the CRSP database), we compute the three factors (labeled sp500rf, hml-sp500 and smb-sp500) using firms in the S&P500 index only. To construct the three S&P500 factors, we follow the same construction procedure as Fama and French.¹¹ The results are very comparable to the results in Table 5, with the exception that the link between our return strategies and the value factor (hml-sp500) is somewhat stronger and statistically significant at conventional significance levels. That said, the value spread over the sample period that we consider (1996:1 through 2009:10) for hml-sp500 is a mere 0.16% per month (1.9% annualized). As argued above, this suggests that the high expected returns that we find for our dividend strategies are not (solely) driven by value firms (or small firms) in the S&P500 index.

The high monthly alphas compensate investors for the risk in the dividend strategies that cannot be explained by other priced factors. Our results become even more striking, however, if we account for the fact that dividend growth rates are, to some extent, predictable, see for instance Lettau and Ludvigson (2005), Ang and Bekaert (2007), Chen, Da, and Priestley (2009), and Binsbergen and Koijen (2010). To illustrate the degree of dividend growth predictability in the S&P500 during various sample periods, we follow the

¹¹We refer to <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data.Library/f-f-factors.html> for the construction of the Fama and French factors.

approach developed in Binsbergen and Koijen (2010) to obtain an estimate of expected dividend growth rates. They combine standard filtering techniques with a present-value model as in Campbell and Shiller (1988) to forecast future returns and dividend growth rates. The approach is summarized in Appendix B.

The estimation results are presented in Table 7. We provide parameter estimates for three data periods, the post-war period, starting in 1946, the period for which monthly data on the index is available, starting in 1970, and the period for which daily data is available, starting in 1989. Consistent with Binsbergen and Koijen (2010), we find that both expected returns and expected dividend growth rates are predictable, with R^2 -values for returns varying between 8.5% and 14.3%, and R^2 -values for dividend growth rates varying between 26.6% and 47.3%, depending on the sample period. Further, both expected returns and expected dividend growth rates have a persistent component, but expected returns are more persistent than expected dividend growth rates.

This high level of dividend growth predictability combined with the high volatility of the returns on the short-term dividend claim seems rather puzzling. The volatility of annual dividend growth rates is only 7%, but a substantial part of the variance can be explained by simply predictor variables. As such, it would seem that to correctly price claims on the S&P500 index, we need a model that generates a downward sloping term structure of expected returns and volatilities, and which generates, or allows for, a non-trivial degree of dividend growth predictability. Investors thus seem to require a large compensation for the risk associated with the unpredictable part of dividend growth. This illustrates how studying dividend strips can improve our understanding of investors' risk preferences and the endowment or technology process in macro-economic models.

4.3 Excess volatility of short-term dividend claims

Shiller (1981) points out that prices are more volatile than subsequent dividends, which is commonly known as “excess volatile.” One explanation has been that discount rates fluctuate over time and are persistent. The long duration of equity makes prices very sensitive to small movements in discount rates, thereby giving rise to excess volatility.

Since we study short-term claims, we can directly compare prices to subsequent realizations. Figure 10 plots the price of the next year of dividends and the realized dividends during the next year.¹² We shift the latter time series such that the price and

¹²At the end of 2004, Microsoft paid a one-time large dividend. Even though this dividend payment substantially increased the dividend yield on Microsoft stock, Microsoft's weight in the S&P500 index was (and is) less than 2%. As a consequence this dividend does not substantially affect the aggregate dividend series.

subsequent realization are plotted at the same date to simplify the comparison. This illustrates that the high volatility of dividend returns is mostly coming from variation in dividend prices as opposed to their realizations. This points to “excess volatility” of the short-term asset. An explanation of the excess volatility puzzle therefore ideally accounts for both the excess volatility of the equity index as well as that of the short-term asset. If dividend growth is i.i.d., persistent and slow-moving discount rates that produce sufficient excess volatility of the index value, will induce much less “excess volatility” for the short-term asset.

4.4 Predictability of dividend returns

The previous section shows that prices are more variable than subsequent realizations. This suggests that discount rates fluctuate over time, which in turn implies that we need to be able to uncover a predictable component in the returns on the dividend strategies (Shiller (1981)). Some of this evidence is already present in Table 3, which shows that dividend returns are to a certain extent mean-reverting. We extend this evidence by regressing monthly dividend returns from trading strategy 1 on the lagged log price-dividend ratio of the short term asset.¹³ We compute this price-dividend ratio, denoted by $\mathcal{P}D_t$, by taking the 1.5 year dividend strip price at time t , and dividing it by the sum of the past 12 realized dividends:

$$\mathcal{P}D_t = \frac{\mathcal{P}_{t,t+18}}{\sum_{s=t-11}^t D_s} \quad (13)$$

The results are presented in the second column of Table 8. We find that $\mathcal{P}D_t$ forecasts dividend returns with a negative sign, and is highly significant. This suggests that when the price of the short-term asset is high relative to the past 12 months of realized dividends, the expected return on dividend strategy 1 is low. We use both OLS standard errors (between parentheses) and Newey-West standard errors (between brackets) to determine the statistical significance of the predictive coefficient. For both sets of standard errors, the results are significant at conventional significance levels. To mitigate concerns regarding measurement error in the predictor variable, we perform two additional regressions. First, we use as the regressor $\ln(\mathcal{P}D_{t-2})$, that is the log price-dividend ratio from the end of the previous quarter, instead of the previous month. Second, we take an average over the

¹³See, among others, Fama and French (1988), Campbell and Shiller (1988), Cochrane (1991), Cochrane (2008), Lettau and Van Nieuwerburgh (2008), Wachter and Warusawitharana (2009), and Binsbergen and Koijen (2010) for the predictability of returns by the price-dividend ratio or the dividend yield for the aggregate market.

past three price-dividend ratios and use this predictor variable instead.¹⁴ More formally, the smoothed price-dividend ratio, $\overline{\mathcal{P}D}_t$, is given by:

$$\overline{\ln(\mathcal{P}D_t)} = \frac{\mathcal{P}D_t + \mathcal{P}D_{t-1} + \mathcal{P}D_{t-2}}{3} \quad (14)$$

The results are reported in the third and fourth column of Table 8 and are comparable to the first column: the price-dividend ratio enters with a negative sign and is significant at conventional levels.

To further illustrate the strength of these predictability results, we present in columns six through eight the same regressions, but now for the S&P500 index. We regress monthly returns on the index (including distributions) on the lagged log price-dividend ratio, computed as the ratio of the index level at time t and the sum of the past twelve realized dividends. Also in this case, the sign on the price-dividend ratio is negative, implying that a high price-dividend ratio is indicative of low expected returns. However, over this sample period, both the R-squared and the statistical significance for the index are substantially lower than for the dividend strategy.

5 Comparison with asset pricing models

To provide a theoretical benchmark for our results, we compute dividend strips in several leading asset pricing models in this section. Recent consumption-based asset pricing models have made substantial progress in explaining many asset pricing puzzles across various markets. Even though such models are not often used to study the pricing of dividend strips, they do have theoretical predictions about their values. We consider the Campbell and Cochrane (1999) external habit formation model, the Bansal and Yaron (2004) long-run risk model, the Barro-Rietz rare disasters framework (Barro (2006)) as explored by Gabaix (2009) and Wachter (2010). We focus on the calibration of Gabaix (2009) in this case.

The habit model and the long-run risk model imply that the risk premium and volatility on long-term dividend claims are higher. The risk premium on the short-term asset is virtually zero and lower than on the aggregate stock market, which is contrary to what we measure in the data. In the rare disasters model, expected returns are constant across maturities, but the volatilities are higher for long-term dividend claims than for short-term claims. To generate these results, we use the original calibrations that are

¹⁴See Cochrane and Piazzesi (2005) for a similar treatment of measurement error in the forecasting variable of, in their case, bond returns.

successful in matching facts about the aggregate stock market. It is important to keep in mind though that such models have a relatively simple shock structure and have not been calibrated to match prices of dividend strips. It may be possible to consider alternative calibrations or model extensions that do match the features of the dividend strip prices we report.

We also consider the model of Lettau and Wachter (2007) who exogenously specify the joint dynamics of cash flows and the stochastic discount factor to match the value premium. In their model, expected returns and volatilities of the short-term asset are higher than on the aggregate stock market, and the CAPM beta of the short-term asset is below one, resulting in a substantial CAPM alpha. These features of their model are in line with our empirical findings.

For all models, we describe the intuition and main results below. In Appendix C - F, we summarize the key equations necessary to compute the returns on dividend strips within the models.

5.1 External habit formation model

In the external habit formation model of Campbell and Cochrane (1999), the consumption dynamics are the same as in the Lucas model of Section 2. As in the Lucas model, dividend growth is assumed to be i.i.d., but shocks to dividend growth rates have a correlation of 20% with shocks to consumption growth rates. Furthermore, the agent is assumed to have external habit formation preferences. The habit level is assumed to be a slow-moving and heteroscedastic process. The heteroscedasticity of the habit process, the sensitivity function, is chosen so that the real interest rate in the model is constant.¹⁵ Further details can be found in Appendix C.

We use the same calibrated monthly parameters as in Campbell and Cochrane (1999). We simulate from the model and compute for each dividend strip with a maturity of n months the average annualized excess return (risk premium), the annualized volatility, and the Sharpe ratio. The results are plotted in Figure 12 for the first 480 months (40 years). The graph shows that the term structure of expected returns and volatilities is upward sloping and the Sharpe ratio is upward sloping as well. The early dividend strips have a low annual average excess return equal to 1%.

The intuition behind these results can be summarized as follows. A positive dividend shock is likely to go together with a positive consumption shock due to the positive

¹⁵Wachter (2006) considers an extension to also match the term structure of interest rates.

correlation between consumption and dividend growth. A positive consumption shock moves current consumption away from the habit level, which in turn lowers the effective risk aversion of the representative agent. The lower degree of risk aversion implies that risk premia fall and future dividends are discounted at a lower rate. As a result, prices of dividend strips increase. This effect is more pronounced for dividend strips with longer maturities as they are more sensitive to discount rates. Since dividend prices are likely to increase in case of a positive consumption shock, they earn a positive risk premium. This effect is more pronounced for long-maturity dividend strips, explaining the upward-sloping curves for risk premia and volatilities. We find that the effect on risk premia is quantitatively stronger, which implies that Sharpe ratios also increase with maturity.

5.2 Long-run risks model

We next consider a long run risks model. We use the model and monthly calibration by Bansal and Yaron (2004).¹⁶ This model departs from the Lucas model in Section 2 in two important ways. First, the CRRA preferences are generalized to Epstein and Zin (1989) preferences to separate the coefficient of relative risk aversion from the elasticity of inter-temporal substitution. Second, the dynamics of consumption and dividend growth are modified in two ways. Both growth rates have a small predictable component that is highly persistent. This implies that even though consumption risk may seem rather small over short horizons, it gradually builds up over longer horizons. In addition to the predictable component, Bansal and Yaron (2004) introduce stochastic volatility in the dynamics of consumption and dividend growth. Further details on the model can be found in Appendix D.

We compute dividend strips in the same manner as described in the previous subsection, and we compute the average annualized excess return, volatility, and Sharpe ratio. The results are plotted in Figure 13. Interestingly, the results are very similar to the habit formation model. The terms structure of expected returns and volatilities is upward sloping and the Sharpe ratio is upward sloping as well.

The intuition behind these results can be summarized as follows. Good states of the economy are states in which the predictable component of growth rates is high and where the stochastic volatility is low. Prices of dividends, however, increase in case of higher growth rates, and fall in case of higher uncertainty. In the model, higher stochastic volatility increases discount rates, which leads to a contemporaneous decline in dividend prices. Both effects imply that dividend strips earn a positive risk premium. Long-

¹⁶We obtain comparable results by using the model and calibration by Bansal and Shaliastovich (2009).

maturity dividend strips are more sensitive to fluctuations in the predictable component of growth rates and the stochastic volatility process, which explains the upward-sloping curves for risk premia and volatilities.

5.3 Variable rare disasters model

We then consider the variable rare disasters model by Gabaix (2009). In this model, the representative agent has CRRA preferences as in the Lucas model in Section 2.¹⁷ The consumption and dividend growth process are generalized to allow for rare disasters. In case of a disaster, both consumption and dividends decline by large amounts. The probability of a rare disasters is assumed to fluctuate over time, which induces time variation in risk premia. However, since shocks to the probability of a rare disaster are independent of shocks to consumption growth, discount rate shocks do not affect risk premia, but they do affect the volatility of dividend strips. See Appendix E for further details.

The results for the variable rare disasters model are summarized in Figure 14. In this model, the term structure of expected returns is flat. The reason is that strips of all maturities are exposed to the same risk in case of a disaster. Further, the return volatility is increasing with maturity. The reason is that longer maturity strips have a higher volatility because they are more sensitive to the time variation in disaster probabilities. As a result, the Sharpe ratio is downward sloping.

5.4 Lettau and Wachter (2007) model

We finally consider the model by Lettau and Wachter (2007), which is designed to generate a downward-sloping term structure of expected returns. We use their quarterly calibration and compute dividend strips using the essentially affine structure of the model.¹⁸ For more details on the calibration and the computation of dividend strips within their model, we refer to Lettau and Wachter (2007) and Appendix F.

As before, we report for each dividend strip n the average annualized excess return (risk premium), the annualized volatility, and the Sharpe ratio. The results are plotted in Figure 15. The term structure for the risk premium is downward sloping and the term structure of volatilities is initially upward sloping up until 8 years, and downward sloping thereafter. The Sharpe ratio is downward-sloping as well.

¹⁷See Wachter (2010) for an extension to Epstein and Zin (1989) preferences.

¹⁸We apply a similar method to compute the dividend strips in the long run risk model as described in appendix B.

The model of Lettau and Wachter (2007) specifies an exogenous stochastic discount factor. Dividend growth is assumed to have a predictable component. In this model, unexpected dividend growth is priced, and the price of risk fluctuates over time. Shocks to the price of risk are assumed to be independent of the other shocks in the model. An important feature of the model is that shocks to expected and unexpected dividend growth are negatively correlated. This implies that long-maturity dividend claims are on a per-period basis less risky than short-horizon claims as, for instance, a negative dividend shock today is partially offset by higher expected growth rates going forward.

The model implies a downward-sloping term structure of risk premia and Sharpe ratios. It also results in CAPM alphas of short-maturity dividend claims that are about 10% per annum and in CAPM betas that are below one. These aspects of the model are consistent with the properties of dividend strips that we measure directly in the data. Lettau and Wachter (2010) show to extend the model to also fit important properties of the term structure of interest rates. However, as also pointed out by Lettau and Wachter (2010), the model is not a full-fledged equilibrium model and an important next step is to think of the micro foundations that can give rise to this specification of the technology and the stochastic discount factor.

6 Robustness

In this section, we perform several robustness checks of our empirical results.

6.1 Alternative selection criteria

In constructing the prices of dividend strips, we take the median across all matches of put and call contracts with the same maturity and strike price, for a given maturity and for which the prices are quoted within the same second. We select the time frame from 10am to 2pm. We now consider five alternative procedures to construct dividend prices. In all cases, we report the summary statistics of dividend returns for strategy 1, and the CAPM alpha and beta.¹⁹ For Alternative 1, we first minimize the time difference between contracts with the same maturity and strike price, we then select the moneyness that is closest to one (at-the-money) for a given maturity, and if multiple matches are found, we take the median across the matches for that particular maturity. For Alternative 2, we first minimize the time difference between contracts with the same maturity and strike price, we then select the smallest bid-ask spread for a given maturity, and if multiple matches

¹⁹The results for dividend steepener are highly comparable and are not reported for brevity.

are found, we take the median across the matches for that particular maturity. In case of Alternative 3, we use the same matching procedure as in the benchmark case, but narrow the time frame to 10am to 11am, and in case of Alternative 4, we consider the time frame from 1pm to 2pm. We exclude the lunch period for the latter two alternative matching procedures, which might be a period of lower liquidity. Finally, in case of Alternative 5, we consider all matches between put and call contracts for a given maturity and strike price, but instead of minimizing over the time difference first, we take the median right away. The advantage is that we take the median across a larger set of contracts, but the time difference between the quotes might not be zero, which introduces noise. In practice, there are so many quotes that the difference time stamps of quotes is in most cases small.

The results are presented in Table 9, in which A_i corresponds to Alternative i . Even though the numbers change slightly across different matching procedures, which is not unexpected, none of our main results are overturned for any of the alternatives. The dividend strategy earns high average returns, has a relatively high volatility, has a modest CAPM beta, and, as a consequence, a substantial CAPM alpha. It seems challenging to construct an argument based on microstructure issues that explains all empirical facts of dividend strategies, and is robust to all matching procedures we consider.

6.2 Dividend prices implied by futures contracts

As an alternative robustness check, we consider a different market to synthetically construct dividend prices. Instead of relying on options markets, we use we data on index futures. As discussed above, index futures do not have as long maturities as index options, but we have access to maturities up to one year. Figure 11 displays the dividend prices for a 6-month and 1-year contract implied by either futures data or options data. To make both series stationary, we scale the price series by the level of the S&P500 index. The relative price series clearly have the same level and are highly correlated; the full-sample correlation equals 94% for the 6-month contract and 91% for the 1-year contract. As such, explanations of our findings must also be able to explain the same phenomenon in futures markets. As argued before, explanations for all facts solely based on market microstructure features are therefore, in our view, less convincing as index futures markets are among the most liquid asset markets available.

6.3 Sensitivity to interest rates

To explore the sensitivity to the LIBOR rates that we use, we perform the following sensitivity analysis. We recompute the dividend prices for both strategies, changing the

interest rate by δ , where we let δ vary between -50 and +50 basis points. This leads to the following dividend prices:

$$\mathcal{P}_{t,T} = p_{t,T} - c_{t,T} + S_t - X e^{-(r_{t,T} + \delta)(T-t)}, \quad (15)$$

In case of the dividend steepener, we compute:

$$\begin{aligned} \mathcal{P}_{t,T_1,T_2} &= \mathcal{P}_{t,T_2} - \mathcal{P}_{t,T_1} \\ &= p_{t,T_2} - p_{t,T_1} - c_{t,T_2} + c_{t,T_1} - X \left(e^{-(r_{t,T_2} + \delta)(T_2-t)} - e^{-(r_{t,T_1} + \delta)(T_1-t)} \right). \end{aligned} \quad (16)$$

For each value of δ we recompute the dividend returns and compute the time-series average of $R_{1,t+1}$ and $R_{2,t+1}$. This allows us to assess what increase in interest rates we would need to drive the average return on both strategies to zero. The results are summarized in Table 10. We find that the average return for strategy 1 ($E[R_{1,t+1}]$) is zero when $\delta = +45bp$. Regardless of whether one considers this number to be small or large, more importantly, this increase in the interest rate does not drive the average return on the dividend steepener to zero. The reason is that δ enters twice in equation (16), so the two terms offset each other. As such, within the range of values that we consider, there is no value for δ for which the average return on the dividend steepener can be driven to zero. Further, the baseline case ($\delta = 0$) appears to provide a lower bound on the expected return on the dividend steepener and leads to the smallest difference in expected returns between the two return strategies.

7 Further applications

In this section, we briefly two other applications that can be explored using the dividend strips we compute in this paper.

7.1 Stochastic discount factor decompositions

Building on Bansal and Lehman (1997), Hansen, Heaton, and Li (2008) and Hansen and Scheinkman (2009) show how to decompose the pricing kernel into a permanent and temporary component. These decompositions are useful for various reasons. Alvarez and Jermann (2005) for instance show that the ratio of the variance of the permanent component to the overall variance is equal to one minus the ratio of the long-term bond risk premium to the maximum risk premium across all securities. This insight can be used to identify pricing factors and to generate additional restrictions for general equilibrium

asset pricing models.²⁰ In addition, these decompositions are useful to understand how future dividend prices respond to a shock to a macro-economic state variable today, see Borovicka, Hansen, Hendricks, and Scheinkman (2009). Borovicka, Hansen, Hendricks, and Scheinkman (2009) largely use these results to point out differences across asset pricing models, but there is no empirical counterpart yet to which these models can be compared. The methods we develop in this paper might be useful to advance our understanding of the decomposition of the stochastic discount factor.

7.2 Market-implied expected returns and expected growth rates

Binsbergen and Koijen (2010) show how to use filtering methods to estimate expected returns and expected growth rates. Filtering methods are required as the price-dividend ratio is an affine function of expected returns and expected growth rates (see also Section 5), which are both latent. However, if we use exactly the same model to price dividend strips, it follows immediately that all dividend strips are affine in the same two state variables, but with different loadings. Assuming that the model is correctly specified, this implies, reminiscent to the term structure literature, that we can invert any two dividend strips to recover market-implied expected returns and growth rates.

8 Conclusion

We use data from derivatives markets to recover the prices of dividend strips on the aggregate stock market. The price of a k -year dividend strip is the present value of the dividend paid in k years. The value of the stock market is the sum of all dividend strip prices across maturities. We study the asset pricing properties of strips and find that expected returns, Sharpe ratios, and volatilities on short-term strips are higher than on the aggregate stock market, while their CAPM betas are well below one. Prices of short-term strips are more volatile than their realizations, leading to excess volatility and return predictability.

We shed new light on the composition of the equity risk premium. When decomposing the index into dividend strips, a natural question that arises is whether dividends at different horizons contribute equally to the equity risk premium or whether either short or long-term dividends contribute proportionally more than the other. We find that short-term dividends have a higher risk premium than long-term dividends, whereas leading asset pricing models predict the opposite.

²⁰See for instance Koijen, Lustig, and Van Nieuwerburgh (2009) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010).

References

- ALVAREZ, F., AND U. JERMANN (2005): “Using Asset Prices to Measure the Measure the Persistence of the Marginal Utility of Wealth.,” *Econometrica*, (4), 1977–2016.
- ANG, A., AND G. BEKAERT (2007): “Stock Return Predictability: Is it There?,” *Review of Financial Studies*, 20(3), 651–707.
- BANSAL, R., AND B. N. LEHMAN (1997): “Growth Optimal Portfolio Restrictions on Asset Pricing Models,” *Macroeconomic Dynamics*, 1, 333–354.
- BANSAL, R., AND I. SHALIASTOVICH (2009): “Risk and Return in Bond, Currency and Equity Markets,” Unpublished paper, Duke University.
- BANSAL, R., AND A. YARON (2004): “Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles,” *Journal of Finance*, 59(4), 1481–1509.
- BARCLAYS CAPITAL, R. (2008): “Dividend Swap Indices: Access to equity income streams made easy,” Research Report, Barclays Capital.
- BARRO, R. (2006): “Rare Disasters and Asset Markets in the Twentieth Century,” *Quarterly Journal of Economics*, August.
- BINSBERGEN, J. H. V., AND R. S. KOIJEN (2010): “Predictive Regressions: A Present-Value Approach,” *Journal of Finance*, 65, 1439–1471.
- BOLLERSLEV, T. (1986): “Generalized Autoregressive Conditional Heterskedasticity,” *Journal of Econometrics*, 31, 307–327.
- BOROVICKA, J., L. P. HANSEN, M. HENDRICKS, AND J. A. SCHEINKMAN (2009): “Risk Price Dynamics,” Working Paper.
- BRENNAN, M. J. (1998): “Stripping the S&P 500 Index,” *Financial Analyst Journal*, 54(1), 12–22.
- CAMPBELL, J. Y., AND J. H. COCHRANE (1999): “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, 107, 205–251.
- CAMPBELL, J. Y., AND R. J. SHILLER (1988): “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors,” *Review of Financial Studies*, 1, 195–227.
- CHEN, L., Z. DA, AND R. PRIESTLEY (2009): “Dividend Smoothing and Predictability,” Working Paper.
- COCHRANE, J. H. (1991): “Explaining the Variance of Price-Dividend Ratios,” *Review of Financial Studies*, 5(2), 243–280.
- (2008): “The Dog That Did Not Bark: A Defense of Return Predictability,” *Review of Financial Studies*, 21(4), 1533–1575.
- COCHRANE, J. H., AND M. PIAZZESI (2005): “Bond Risk Premia,” *American Economic Review*, 95, 138–160.

- CONSTANTINIDES, G., J. JACKWERTH, AND S. PERRAKIS (2009): “Mispricing of S&P 500 Index Options,” *Review of Financial Studies*, 22, 1247–1277.
- CROCE, M., M. LETTAU, AND S. C. LUDVIGSON (2009): “Investor Information, Long-Run Risk, and the Duration of Risky Cash-Flows,” Working paper, UC Berkeley.
- DUFFEE, G. (2010): “Sharpe ratios in term structure models,” Working Paper.
- EPSTEIN, L., AND S. ZIN (1989): “Substitution Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 57, 937–968.
- FAMA, E. F., AND K. R. FRENCH (1988): “Dividend Yields and Expected Stock Returns,” *Journal of Financial Economics*, 22, 3–27.
- (1993): “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, 33, 3–56.
- GABAIX, X. (2009): “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance,” Working paper, NYU.
- GORDON, M. (1962): *The Investment, Financing, and Valuation of the Corporation*. Irwin, Homewood, IL.
- HANSEN, L. P., J. C. HEATON, AND N. LI (2008): “Consumption Strikes Back? Measuring Long-Run Risk,” *Journal of Political Economy*, 116 (2), 260–302.
- HANSEN, L. P., AND J. A. SCHEINKMAN (2009): “Long-Term Risk: An Operator Approach,” *Econometrica*, 77, 177 – 234.
- HANSEN, L. P., AND K. SINGLETON (1982): “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models,” *Econometrica*, 50(5), 1269–86.
- (1983): “Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns,” *Journal of Political Economy*, 91, 249–65.
- KOIJEN, R., H. LUSTIG, AND S. VAN NIEUWERBURGH (2009): “The Cross-Section and Time-Series of Stock and Bond Returns,” Working Paper.
- KOIJEN, R., H. LUSTIG, S. VAN NIEUWERBURGH, AND A. VERDELHAN (2010): “Long-Run Risk, the Wealth-Consumption Ratio, and the Temporal Pricing of Risk,” Forthcoming, *American Economic Review, Papers and Proceedings*.
- LETTAU, M., AND S. C. LUDVIGSON (2005): “Expected Returns and Expected Dividend Growth,” *Journal of Financial Economics*, 76, 583–626.
- LETTAU, M., AND S. VAN NIEUWERBURGH (2008): “Reconciling the Return Predictability Evidence,” *Review of Financial Studies*, 21(4), 1607–1652.
- LETTAU, M., AND J. A. WACHTER (2007): “Why is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium,” *Journal of Finance*, 62, 55–92.
- (2010): “The Term Structures of Equity and Interest Rates,” *Journal of Financial Economics*, forthcoming.

- LEWELLEN, J., S. NAGEL, AND J. SHANKEN (2009): “A Skeptical Appraisal of Asset Pricing Tests,” Forthcoming, *Journal of Financial Economics*.
- LINTNER, J. (1965): “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *Review of Economics and Statistics*, 47, 13–37.
- LUCAS, R. (1978): “Asset Prices in an Exchange Economy,” *Econometrica*, 46, 1429–1446.
- MANLEY, R., AND C. MUELLER-GLISSMANN (2008): “The Market for Dividends and Related Investment Strategies,” *Financial Analyst Journal*, 64(3), 17–29.
- MEHRA, R., AND E. C. PRESCOTT (1985): “The Equity Premium Puzzle,” *Journal of Monetary Economics*, 15, 145–161.
- RIETZ, T. A. (1988): “The Equity Risk Premium: A Solution?,” *Journal of Monetary Economics*, 22(1), 117–131.
- SANTOS, T., AND P. VERONESI (2006): “Habit Formation, the Cross Section of Stock Returns and the Cash Flow Risk Puzzle,” *Journal of Financial Economics*, forthcoming.
- SHARPE, W. (1964): “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance*, 19, 425–444.
- SHARPE, W. F. (1966): “Mutual Fund Performance,” *Journal of Business*, 39(1), 119–138.
- SHILLER, R. J. (1981): “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?,” *American Economic Review*, 71, 421–436.
- STOLL, H. R. (1969): “The Relationship between Put and Call Option Prices,” *Journal of Finance*, 24, 801–824.
- WACHTER, J. (2005): “Solving Models with External Habit,” *Finance Research Letters*, 2, 210–226.
- WACHTER, J., AND M. WARUSAWITHARANA (2009): “Predictable Returns and Asset Allocation: Should a Skeptical Investor Time the Market?,” *Journal of Econometrics*, 148, 162–178.
- WACHTER, J. A. (2006): “A Consumption-based Model of the Term Structure of Interest Rates,” *Journal of Financial Economics*, 79, 365–399.
- (2010): “Can Time-varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?,” Working Paper.
- WILKENS, S., AND J. WIMSCHULTE (2009): “The Pricing of Dividend Futures in the European Market: A First Empirical Analysis,” Unpublished paper, University of Regensburg.

Panel A: Full sample 1996:2 - 2009:10			
	$R_{1,t}$	$R_{2,t}$	S&P500
Mean	0.0116	0.0112	0.0056
Median	0.0079	0.0148	0.0106
Std. Dev.	0.0780	0.0965	0.0469
Sharpe ratio	0.1124	0.0872	0.0586
Observations	165	165	165
Panel B: First half sample 1996:2 - 2002:12			
Mean	0.0159	0.0139	0.0065
Median	0.0117	0.0231	0.0093
Std. Dev.	0.0986	0.1212	0.0514
Sharpe ratio	0.1242	0.0843	0.0456
Observations	83	83	83
Panel C: Second half sample 2003:1 - 2009:10			
Mean	0.0072	0.0086	0.0046
Median	0.0058	0.0086	0.0118
Std. Dev.	0.0494	0.0630	0.0422
Sharpe ratio	0.1060	0.1044	0.0615
Observations	82	82	82

Table 1: Summary Statistics

The table presents descriptive statistics of the monthly returns on the two trading strategies described in the main text. As the volatility in the second half of the sample is lower than in the first half of the sample, we also present descriptive statistics for two subsamples: 1996:2-2002:12 and 2003:1-2009:10.

Dep. Var.	$R_{1,t+1}$	$R_{2,t+1}$	$R_{SP500,t+1}$
Mean equation			
c	0.0074 (0.0047)	0.0085 (0.0042)	0.0070 (0.0038)
AR(1)	-0.2682 (0.0973)	-0.3668 (0.1056)	0.0898 (0.0958)
Variance equation			
c	1.5×10^{-4} (7.4×10^{-5})	2.4×10^{-4} (1.4×10^{-4})	7.2×10^{-5} (6.6×10^{-5})
squared residual	0.1138 (0.0347)	0.1724 (0.0536)	0.1859 (0.0722)
GARCH(1)	0.8773 (0.0263)	0.8287 (0.0536)	0.8056 (0.0639)

Table 2: Estimates of the GARCH(1,1) Model

The top panel provides the estimates of the mean equation; the bottom panel displays the estimates of the variance model. The first two columns report the results for the dividend return strategies, and the third column provides the results for the S&P500.

Dep. Var.	$R_{1,t+1} - R_{f,t}$			$R_{2,t+1} - R_{f,t}$		
c	0.0088 (0.0050)	0.0073 (0.0051)	0.0073 (0.0046)	0.0084 (0.0056)	0.0069 (0.0054)	0.0068 (0.0049)
mktrf	-	0.4721 (0.1612)	0.5006 (0.1517)	-	0.4847 (0.1850)	0.5127 (0.1768)
AR(1)	-	-	-0.2889 (0.1088)	-	-	-0.2904 (0.1164)
R^2	-	0.0877	0.1709	-	0.0604	0.1153

Table 3: Monthly Returns on the Two Trading Strategies and the Market Portfolio.

The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the market portfolio. Newey-West standard errors in parentheses. When an AR(1) term is included, the intercept is adjusted by one minus the AR(1) coefficient, such that the intercept is comparable to the regressions without AR(1) term.

Dep. Var.	$R_{1,t+1} - R_{f,t}$		$R_{2,t+1} - R_{f,t}$	
c	0.0075 (0.0051)	0.0075 (0.0047)	0.0071 (0.0054)	0.0070 (0.0050)
sp500rf	0.4488 (0.1667)	0.4766 (0.1575)	0.4863 (0.1859)	0.5142 (0.1761)
AR(1)	- -	-0.2857 (0.1110)	- -	-0.2880 (0.1185)
R^2	0.0725	0.1542	0.0556	0.1100

Table 4: Monthly Returns on the Two Trading Strategies and the S&P500 Index.

The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the excess returns on the S&P500 index. Newey-West standard errors in parentheses. When an AR(1) term is included, the intercept is adjusted by one minus the AR(1) coefficient, such that the intercept is comparable to the regressions without AR(1) term.

Dep. Var.	$R_{1,t+1} - R_{f,t}$		$R_{2,t+1} - R_{f,t}$	
c	0.0065 (0.0045)	0.0068 (0.0042)	0.0053 (0.0048)	0.0057 (0.0044)
mktrf	0.4880 (0.1485)	0.5086 (0.1396)	0.5712 (0.1610)	0.5915 (0.1529)
hml	0.1393 (0.1900)	0.1136 (0.1813)	0.3744 (0.2202)	0.3530 (0.2322)
smb	0.0751 (0.1522)	0.0862 (0.1517)	-0.0279 (0.1735)	-0.018 (0.1745)
AR(1)	- -	-0.2876 (0.1098)	- -	-0.2806 (0.1144)
R^2	0.0915	0.1739	0.0811	0.1334

Table 5: Monthly Returns on the Two Trading Strategies and the Three Factor Model.

The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the Fama French three factor model. Newey-West standard errors in parentheses. When an AR(1) term is included, the intercept is adjusted with the AR(1) coefficient, so that the intercept is comparable to the regressions without AR(1) term.

Dep. Var.	$R_{1,t+1} - R_{f,t}$	$R_{2,t+1} - R_{f,t}$
c	0.0061 (0.0043)	0.0066 (0.0048)
sp500rf	0.3972 (0.1809)	0.4137 (0.1884)
hml-sp500	0.1526 (0.1651)	0.5668 (0.1838)
smb-sp500	0.3043 (0.2780)	-0.0528 (0.3180)
R^2	0.1000	0.1011

Table 6: Monthly Returns on the Two Trading Strategies and the Three S&P500 Factors. The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the Fama French three factor model, where the three factors are constructed using S&P500 firms only. Newey-West standard errors in parentheses.

Sample period	1946-2009	1970-2009	1989-2009
δ_0	0.0757	0.0488	0.0448
δ_1	0.9174	0.9315	0.7866
γ_0	0.0489	0.0302	0.0281
γ_1	0.5306	0.7263	0.6496
σ_μ	0.0204	0.0191	0.0432
σ_g	0.0569	0.0510	0.0644
σ_D	0.0056	0.0087	0.0098
$\sigma_{g\mu}$	0.4515	0.5313	0.4831
$\sigma_{\mu D}$	0.8877	0.8464	0.8739
R^2_{Ret}	0.1091	0.0852	0.1435
R^2_{Div}	0.2666	0.4743	0.4509

Table 7: Maximum-Likelihood Estimates

We present the estimation results of the latent variables present-value model as proposed by Binsbergen and Koijen (2010), using S&P500 index prices and dividends. The model is estimated by conditional maximum likelihood using data over three different sample periods, 1946-2009, 1970-2009, 1989-2009 on cash-invested dividend growth rates and the corresponding price-dividend ratio.

Dep. Var.	$R_{1,t+1}$				$R_{SP500,t+1}$		
Constant	0.0752 (0.0144) [0.0218]	0.0455 (0.0153) [0.0203]	0.0585 (0.0157) [0.0221]	Constant	0.0635 (0.0622) [0.1052]	0.0793 (0.0634) [0.0924]	0.0773 (0.0638) [0.1006]
$\ln(\mathcal{P}D_t)$	-0.1683 (0.0351) [0.0491]	-	-	$\ln(PD_t)$	-0.0142 (0.0152) [0.0253]	-	-
$\ln(PD_{t-2})$	-	-0.0903 (0.0370) [0.0459]	-	$\ln(PD_{t-2})$	-	-0.0181 (0.0155) [0.0221]	-
$\overline{\ln(PD_t)}$	-	-	-0.1249 (0.0385) [0.0505]	$\overline{\ln(PD_t)}$	-	-	-0.0176 (0.0156) [0.0241]
R^2	0.1235	0.0356	0.0614	R^2	0.0053	0.0084	0.0078

Table 8: Return predictability

Column 2 shows the regression results of the monthly returns of trading strategy 1, $R_{1,t+1}$, on the price-dividend ratio of the short-term asset at time t (a lag of 1 month), computed as the ratio of the 1.5-year dividend strip price at time t , denoted by $\mathcal{P}_{t,t+18}$, and the sum of dividends paid out over the past twelve months. In column 3, we redo the analysis of column 2, but now lag the price-dividend ratio by two more months (one quarter in total). In Column 4, we use as the regressor the smoothed price-dividend ratio of the short-term asset computed as the equal-weighted average over periods t , $t - 1$ and $t - 2$. OLS standard errors are in parentheses, and Newey-West standard errors are in brackets. In columns 6 through 8, we repeat the analysis of columns 2 through 4 for the S&P500 index. We take monthly returns on the S&P500 index ($R_{SP500,t+1}$) and regress those on various lags of the price-dividend ratio of the S&P500 index, computed as the index value at time t , dividend by the sum of dividends paid out over the past twelve months.

	Benchmark	A1	A2	A3	A4	A5
Mean	0.0116	0.0122	0.0142	0.0114	0.0106	0.0116
Median	0.0079	0.0137	0.0058	0.0101	0.0073	0.0070
St. Dev.	0.0780	0.1026	0.1313	0.0755	0.0735	0.0778
Sharpe ratio	0.1124	0.0910	0.0866	0.1136	0.1063	0.1130
CAPM Alpha	0.0073	0.0079	0.0089	0.0070	0.0064	0.0073
CAPM Beta	0.4721	0.4464	0.7736	0.4962	0.4427	0.4757

Table 9: Alternative selection criteria

The table presents the summary statistics of dividend strategy 1 for five alternative selection criteria (A1 to A5), which are described in the main text. The table also reports the CAPM alpha and beta.

δ	-50bp	-40bp	-30bp	-20bp	-10bp	0bp	10bp	20bp	30bp	40bp	50bp
$E[R_{1,t+1}]$	0.053	0.037	0.0278	0.021	0.016	0.012	0.008	0.005	0.003	0.001	-0.001
$E[R_{2,t+1}]$	0.017	0.021	0.0152	0.013	0.011	0.011	0.012	0.012	0.013	0.015	0.016

Table 10: Sensitivity to interest rates

The table presents the sensitivity analysis to interest rates. We add a constant δ to our LIBOR interest rates and recompute the sample averages of the returns for values of δ varying between -50bp and +50bp. For $\delta = 0$, our baseline results are obtained (in bold).

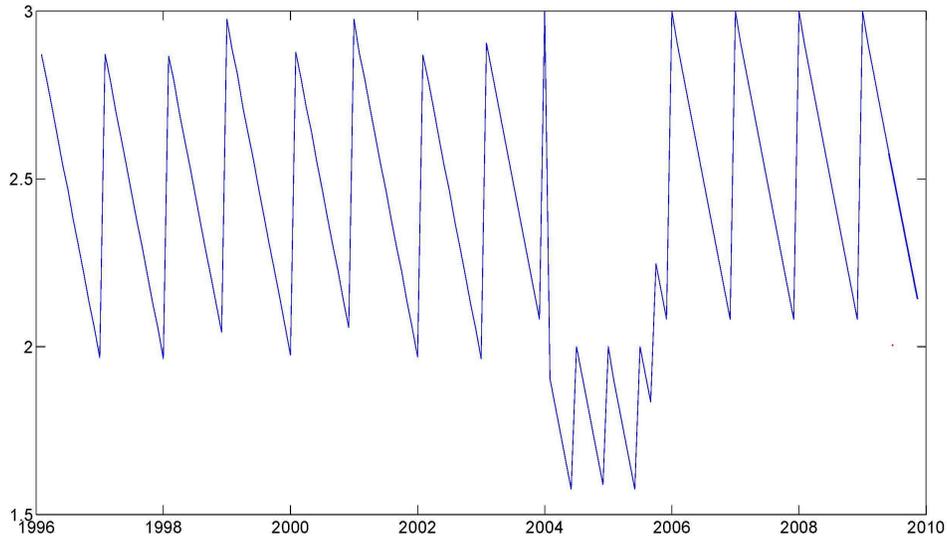


Figure 1: Maximum Maturity of LEAPS

The graph displays the maximum maturity of LEAPS contracts in years that is available at each point of the sample. The sample period is January 1996 up to October 2009.

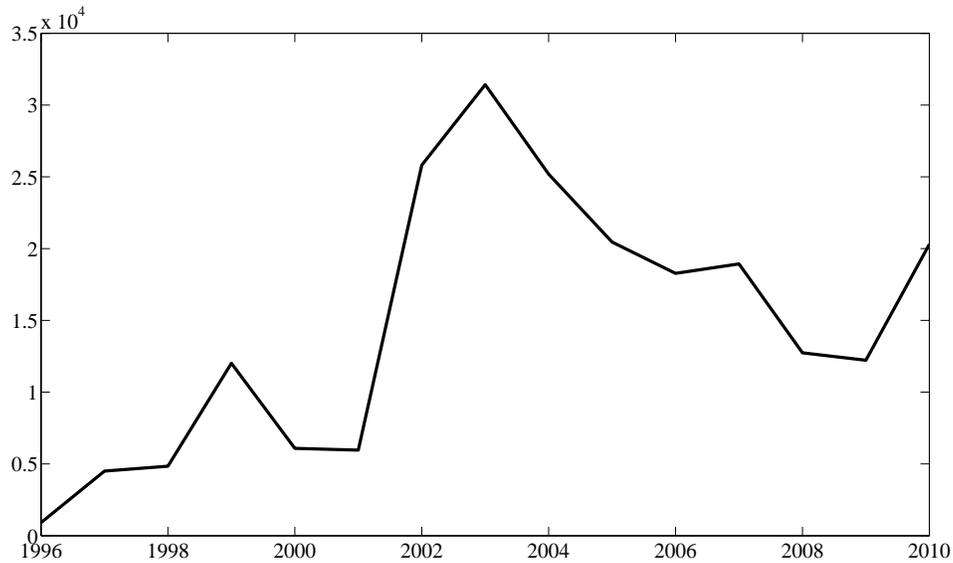


Figure 2: Average Number of Matches

The graph shows the average number of matches of put and call contracts with strike prices and maturities that coincide, and for which the quotes are provided in the same second during the last trading day of the month. We focus on contracts with a maturity between 1 and 2 years, and average the number of matches within a year. The sample period is January 1996 up to October 2009.

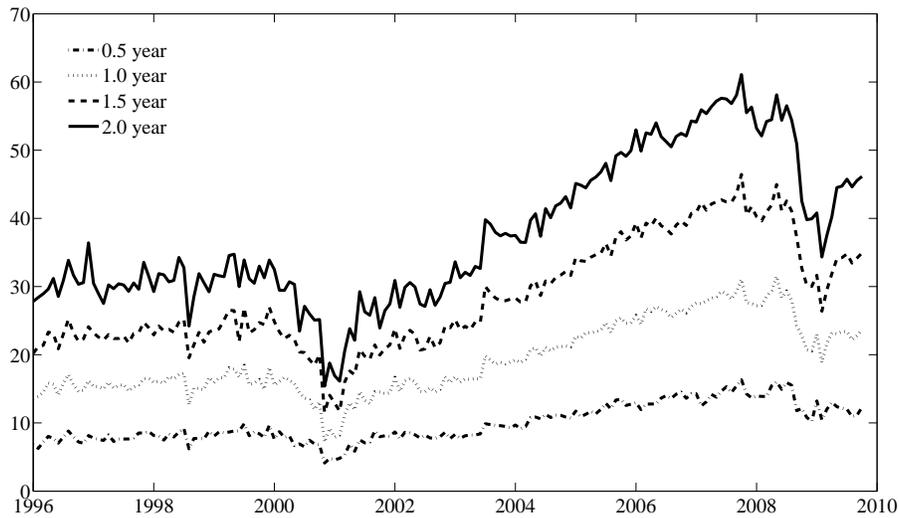


Figure 3: Price Dynamics of the Short-Term Assets (Cumulative)

The graph shows the prices of the first 0.5, 1, 1.5 and 2 years of dividends. The sample period is January 1996 up to October 2009.

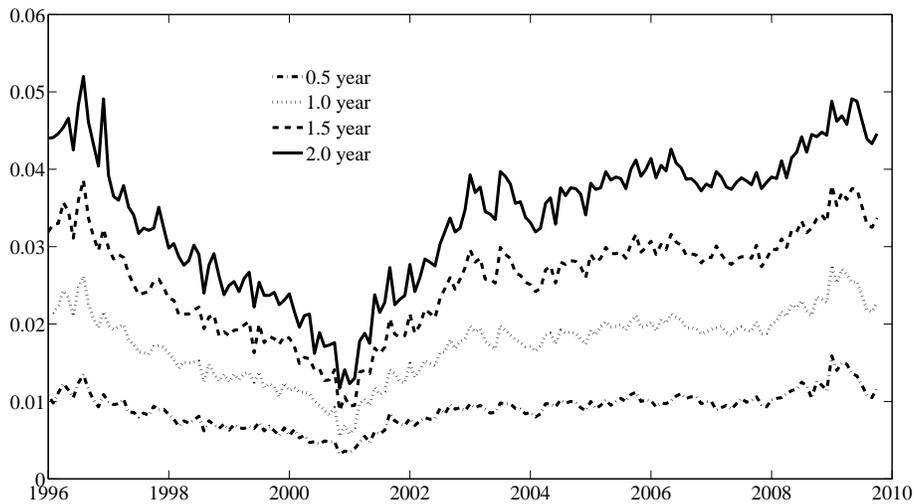


Figure 4: Present Value of Dividends as a Fraction of the Index Value (Cumulative)

The graph shows the net present value of the first 0.5, 1, 1.5 and 2 years of dividends as a fraction of the index value as computed. The sample period is January 1996 up to October 2009.

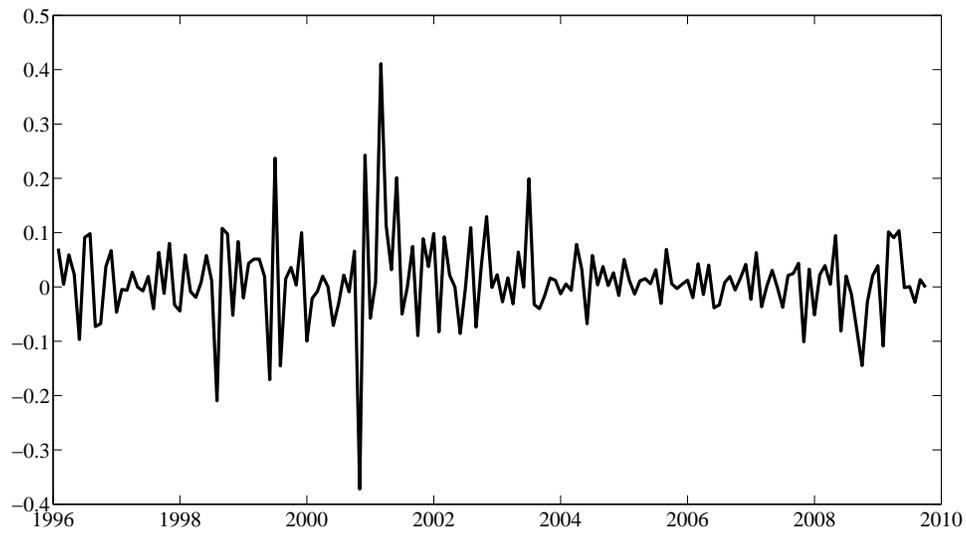


Figure 5: Monthly Returns on Strategy 1 as Defined in (11): 1996:2-2009:10: line graph.

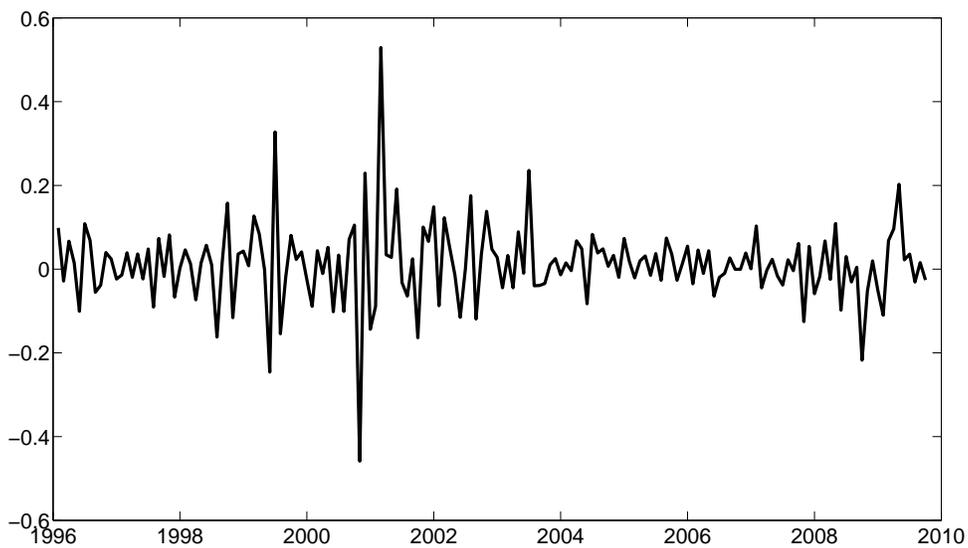


Figure 6: Monthly Returns on Strategy 2 as Defined in (12): 1996:2-2009:10: line graph.

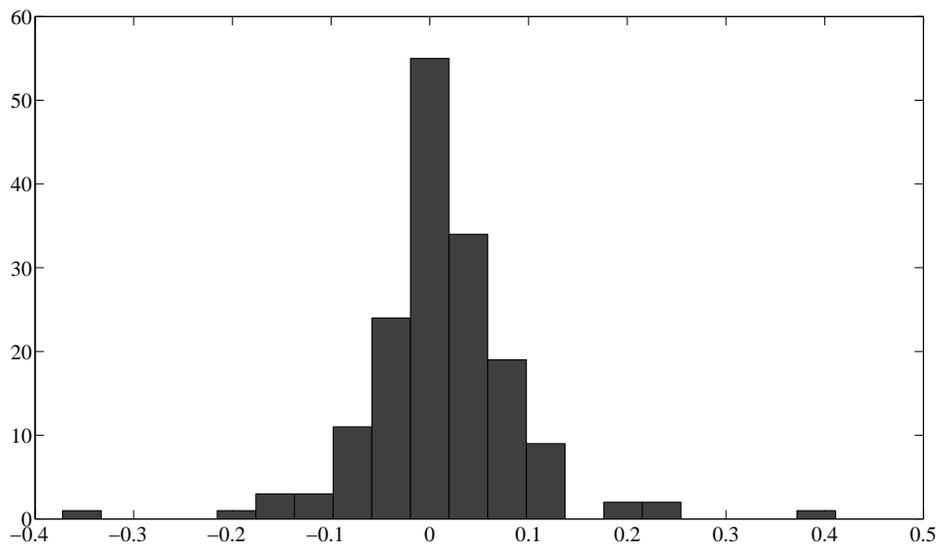


Figure 7: Monthly Returns on Strategy 1 as Defined in (11): 1996:2-2009:10: histogram.

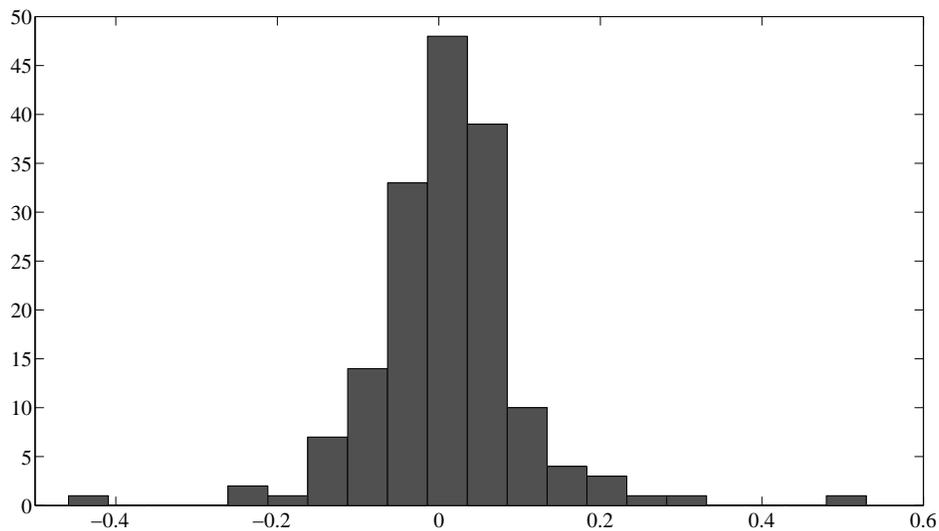


Figure 8: Monthly Returns on Strategy 2 as Defined in (12): 1996:2-2009:10: histogram.

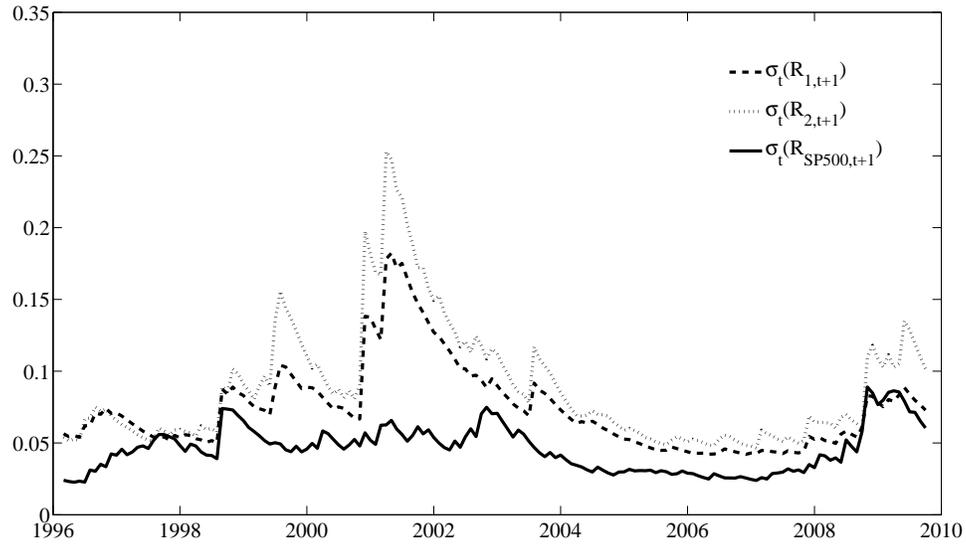


Figure 9: Volatility of Dividend Returns and Returns on the S&P500 Based on a GARCH(1,1) Model.

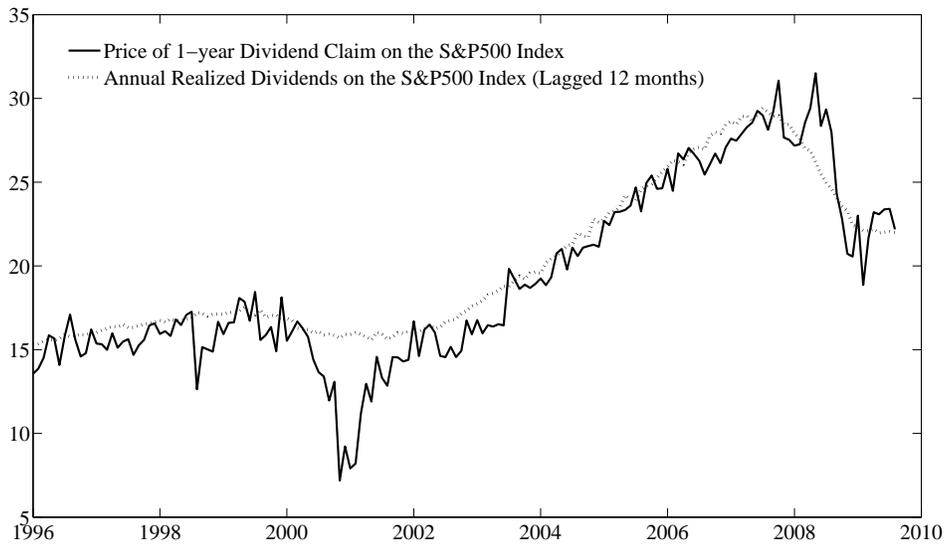


Figure 10: Prices and Realizations of Dividend Claims: 1996:2-2009:08.

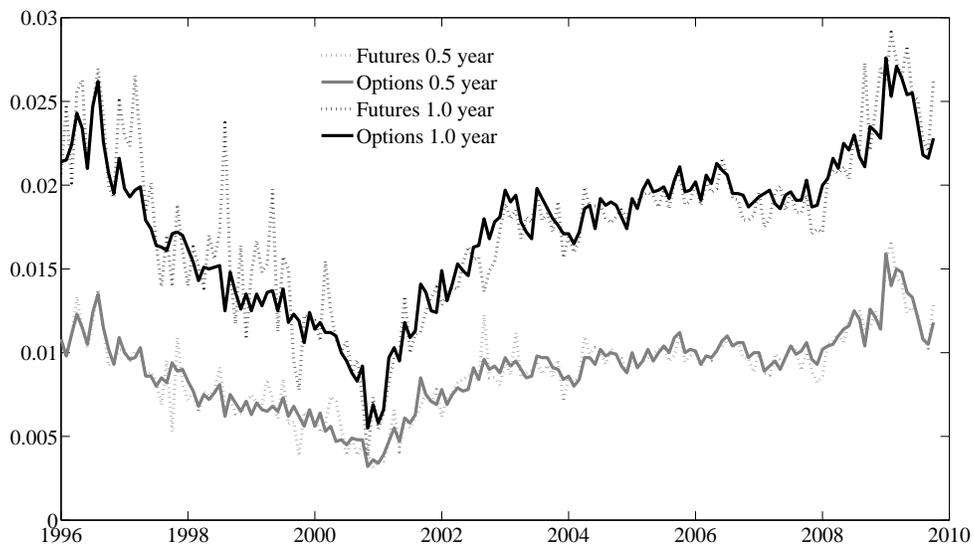


Figure 11: Short-Term Asset Prices Implied by Futures and Options

The graph shows the price of the short-term assets implied by futures and option markets as a fraction of the index. The maturity of the short-term asset equals either 0.5 year or 1 year.

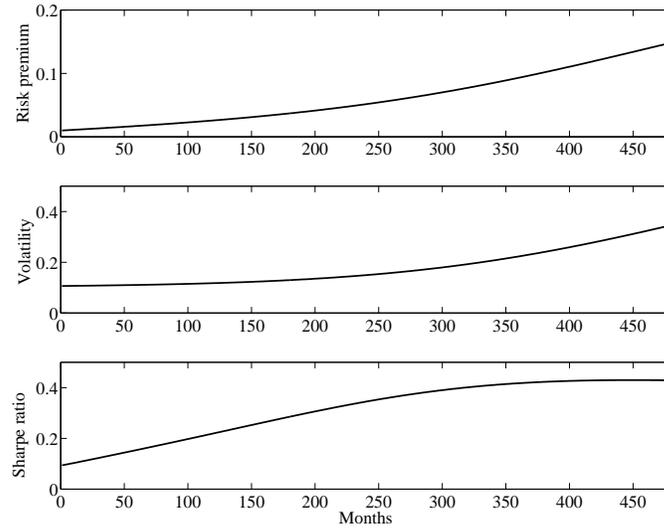


Figure 12: Term Structure of the Risk Premium, Volatility, and Sharpe Ratio for Habits
 The graph shows the term structures of the risk premium, the volatility, and the Sharpe ratio for the Campbell and Cochrane (1999) habit formation model. The graph plots the first 480 months of dividend strips, which corresponds to 40 years.

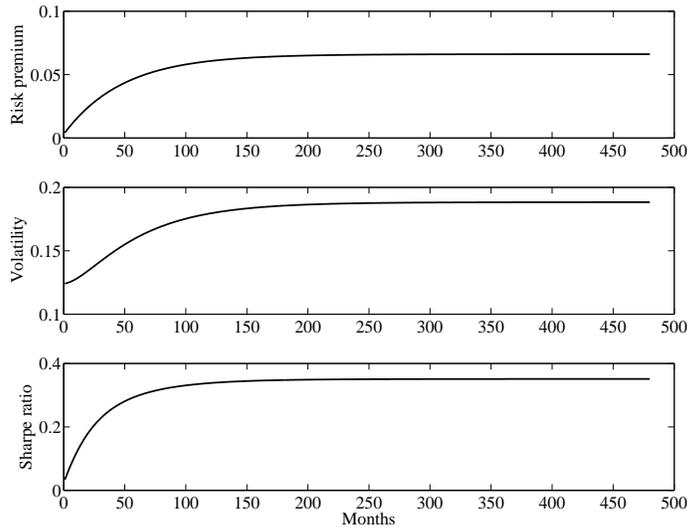


Figure 13: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for the Long Run Risk Model

The graph shows the term structures of the risk premium, the volatility, and the Sharpe ratio for the long run risk model of Bansal and Yaron (2004). The graph plots the first 480 months of dividend strips, which corresponds to 40 years.

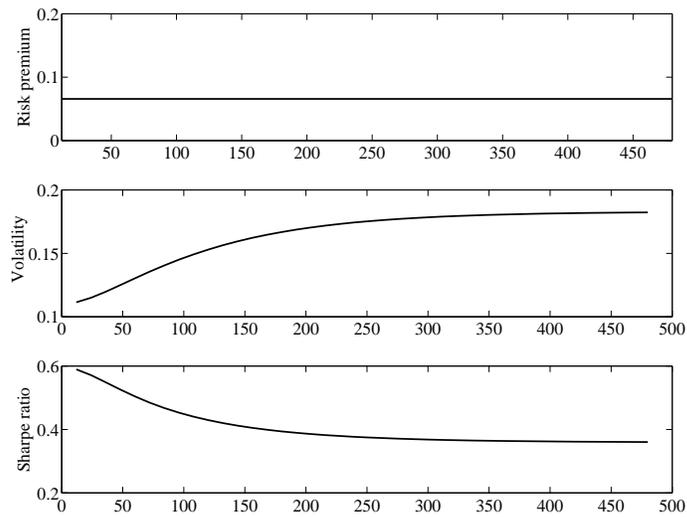


Figure 14: Term Structure of the Risk Premium, Volatility, and Sharpe Ratio for the Variable Rare Disasters Model

The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the variable rare disasters model of Gabaix (2009). The graph plots the first 480 months of dividend strips, which corresponds to 40 years.

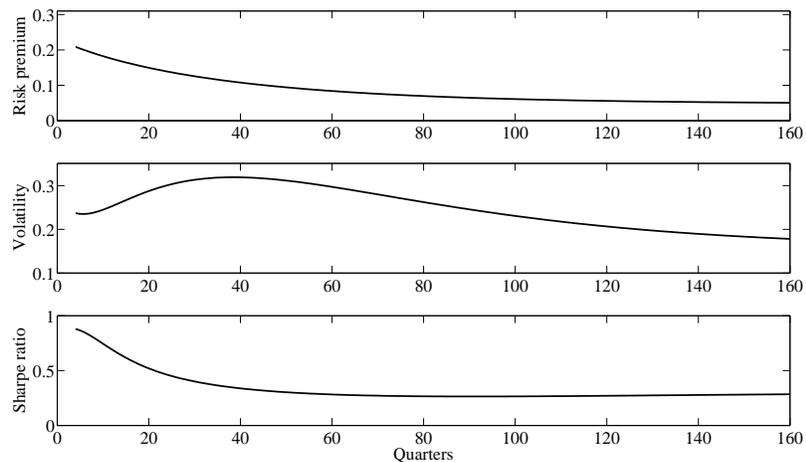


Figure 15: Term Structure of the Risk Premium, Volatility, and Sharpe Ratio for the Lettau Wachter (2007) Model

The graph shows the term structures of the risk premium, the volatility, and the Sharpe ratio for the Lettau Wachter (2007) model. The graph plots the first 160 quarters of dividend strips, which corresponds to 40 years.

A Details dividend returns

The two trading strategies described in Section 3.3 can be implemented for different maturities T . The specific maturities we follow for trading strategy 1 vary between 1.9 years and 1.3 years. To be precise, for trading strategy 1, we go long in the 1.874 year dividend claim on January 31st 1996, collect the dividend during February and sell the claim on February 29th 1996 to compute the return. The claim then has a remaining maturity of 1.797 years. We buy back the claim (or alternatively, we never sold it), go long in the 1.797 year claim, collect the dividend, and sell it on March 29th 1996. We follow this strategy until July 31st 1996 at which time the remaining maturity is 1.381 years. On this date a new 1.881 year contract is available so we restart the investment cycle at this time. We continue this procedure until October of 2009, which is the end of our sample.

For trading strategy 2, we follow the same maturities, apart from the fact that we go long in the 1.874 year dividend claim and short in the 0.874 dividend claim on January 31st 1996. On July 31st 1996 the remaining maturities are 1.381 years and 0.381 years at which point we restart the investment cycle in the 1.881 year contract and the 0.881 year contract available at that time.

B Forecasting returns and dividend growth rates

We follow Binsbergen and Kojien (2010) and use filtering techniques to predict future dividend growth rates and returns. Let r_{t+1} denote the total log return on the index:

$$r_{t+1} \equiv \log \left(\frac{S_{t+1} + D_{t+1}}{S_t} \right), \quad (17)$$

where let PD_t denote the price-dividend ratio:

$$PD_t \equiv \frac{S_t}{D_t},$$

and let Δd_{t+1} denote the aggregate log dividend growth rate:

$$\Delta d_{t+1} \equiv \log \left(\frac{D_{t+1}}{D_t} \right).$$

We model both expected returns (μ_t) and expected dividend growth rates (g_t) as an AR(1) process:

$$\mu_{t+1} = \delta_0 + \delta_1 (\mu_t - \delta_0) + \varepsilon_{t+1}^\mu, \quad (18)$$

$$g_{t+1} = \gamma_0 + \gamma_1 (g_t - \gamma_0) + \varepsilon_{t+1}^g, \quad (19)$$

where $\mu_t \equiv E_t[r_{t+1}]$ and $g_t \equiv E_t[\Delta d_{t+1}]$. The distribution of the shocks ε_{t+1}^μ and ε_{t+1}^g is specified below. The realized dividend growth rate is equal to the expected dividend growth rate plus an orthogonal shock:

$$\Delta d_{t+1} = g_t + \varepsilon_{t+1}^D.$$

Defining $pd_t \equiv \log(PD_t)$, we can write the log-linearized return as:

$$r_{t+1} \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t,$$

with $\bar{pd} = E[pd_t]$, $\kappa = \log(1 + \exp(\bar{pd})) - \rho \bar{pd}$ and $\rho = \frac{\exp(\bar{pd})}{1 + \exp(\bar{pd})}$, as in Campbell and Shiller (1988). If we iterate this equation, and using the AR(1) assumptions (18)-(19), it follows that:

$$pd_t = A - B_1 (\mu_t - \delta_0) + B_2 (g_t - \gamma_0),$$

with $A = \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho}$, $B_1 = \frac{1}{1-\rho\delta_1}$, and $B_2 = \frac{1}{1-\rho\gamma_1}$. The log price-dividend ratio is linear in the expected return μ_t and the expected dividend growth rate g_t . The loading of the price-dividend ratio on expected returns and expected dividend growth rates depends on the relative persistence of these variables (δ_1 versus γ_1). The three shocks in the model, which are shocks to expected dividend growth rates (ε_{t+1}^g), shocks to expected returns (ε_{t+1}^μ), and realized dividend growth shocks (ε_{t+1}^d), have mean zero, covariance matrix

$$\Sigma \equiv \text{var} \left(\begin{bmatrix} \varepsilon_{t+1}^g \\ \varepsilon_{t+1}^\mu \\ \varepsilon_{t+1}^d \end{bmatrix} \right) = \begin{bmatrix} \sigma_g^2 & \sigma_{g\mu} & \sigma_{gD} \\ \sigma_{g\mu} & \sigma_\mu^2 & \sigma_{\mu D} \\ \sigma_{gD} & \sigma_{\mu D} & \sigma_D^2 \end{bmatrix},$$

and are independent and identically distributed over time. Further, in the maximum likelihood estimation procedure, we assume that the shocks are jointly normally distributed.

We subsequently perform unconditional maximum likelihood estimation to obtain estimates for all the parameters and obtain filtered series for μ_t and g_t . The R^2 values

are computed as:

$$R_{Ret}^2 = 1 - \frac{\hat{v}ar(r_{t+1} - \mu_t^F)}{\hat{v}ar(r_t)}, \quad (20)$$

$$R_{Div}^2 = 1 - \frac{\hat{v}ar(\Delta d_{t+1} - g_t^F)}{\hat{v}ar(\Delta d_{t+1})}, \quad (21)$$

where $\hat{v}ar$ is the sample variance, μ_t^F is the filtered series for expected returns (μ_t) and g_t^F is the filtered series for expected dividend growth rates (g_t).

C Dividend strips in the external habit formation model

We first summarize some of the key equations of the Campbell and Cochrane (1999) habit formation model. The stochastic discount factor is given by:

$$M_{t+1} = \delta G^{-\gamma} e^{-\gamma(s_{t+1} - s_t + v_{t+1})}, \quad (22)$$

where G represents consumption growth, γ is the curvature parameter, v_{t+1} is unexpected consumption growth, and s_t is the log consumption surplus ratio whose dynamics are given by:

$$s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1}, \quad (23)$$

where $\lambda(s_t)$ is the sensitivity function which is chosen such that the risk free rate is constant, see Campbell and Cochrane (1999) for further details. Dividend growth in the model is given by:

$$\Delta d_{t+1} = g + w_{t+1}. \quad (24)$$

We solve the model using the solution method described in Wachter (2005). Let $D_t^{(n)}$ denote the price of a dividend at time t that is paid n periods in the future. Let D_{t+1} denote the realized dividend in period $t+1$. The price of the first dividend strip is simply given by:

$$D_t^{(1)} = E_t(M_{t+1}D_{t+1}) = D_t E_t\left(M_{t+1} \frac{D_{t+1}}{D_t}\right). \quad (25)$$

The following recursion then allows us to compute the remaining dividend strips:

$$D_t^{(n)} = E_t(M_{t+1}D_{t+1}^{(n-1)}). \quad (26)$$

The return on the n^{th} dividend strip is given by:

$$R_{n,t+1} = \frac{D_{t+1}^{(n-1)}}{D_t^{(n)}}. \quad (27)$$

D Dividend strips in the long-run risks model

The technology processes are given by:

$$\begin{aligned} x_{t+1} &= \rho_x x_t + \varepsilon_{x,t+1}, \\ \Delta c_{t+1} &= \mu_c + x_t + \varepsilon_{c,t+1}, \\ \Delta d_{t+1} &= \mu_d + \phi x_t + \varepsilon_{d,t+1}, \\ \sigma_{t+1}^2 &= \mu_\sigma + \rho_\sigma (\sigma_t^2 - \mu_\sigma) + \varepsilon_{\sigma,t+1}, \end{aligned}$$

and we define $\varepsilon_{t+1} \equiv (\varepsilon_{c,t+1}, \varepsilon_{x,t+1}, \varepsilon_{\sigma,t+1}, \varepsilon_{d,t+1})'$. We assume:

$$\varepsilon_{t+1} \mid \mathcal{F}_t \sim N(0, \Sigma_t),$$

where:

$$\Sigma_t = \Sigma_0 + \Sigma_1 \sigma_t^2.$$

For the return on total wealth, we have:

$$\begin{aligned} R_{t+1}^c &= \frac{W_{t+1}}{W_t - C_t} \\ &= \frac{\exp(wc_{t+1})}{\exp(wc_t) - 1} \exp(\Delta c_{t+1}), \end{aligned}$$

and thus:

$$\begin{aligned} r_{t+1}^c &= wc_{t+1} + \Delta c_{t+1} - \ln(\exp(wc_t) - 1) \\ &\simeq wc_{t+1} + \Delta c_{t+1} - \ln(\exp(E(wc_t)) - 1) - \frac{\exp(E(wc_t))}{\exp(E(wc_t)) - 1} (wc_t - E(wc_t)) \\ &= \kappa_0^c + \Delta c_{t+1} + wc_{t+1} - \kappa_1^c wc_t, \end{aligned}$$

implying:

$$\kappa_0^c = -\ln(\exp(E(wc_t)) - 1) + \kappa_1^c E(wc_t), \quad (28)$$

$$\kappa_1^c = \frac{\exp(E(wc_t))}{\exp(E(wc_t)) - 1} > 1. \quad (29)$$

The stochastic discount takes the form:

$$m_{t+1} = c_0^m + c_1^m \Delta c_{t+1} + c_2^m (w c_{t+1} - \kappa_1^c w c_t),$$

where:

$$\begin{aligned} w c_t &= A_0^c + A_1^c x_t + A_2^c \sigma_t^2, \\ c_0^m &= -\kappa_0^c - \frac{\gamma - 1}{1 - 1/\psi} (\ln \delta + \kappa_0^c), \\ c_1^m &= -\gamma, \\ c_2^m &= -\frac{\gamma - 1/\psi}{1 - 1/\psi}. \end{aligned}$$

To compute $(\kappa_0^c, \kappa_1^c, A_0^c, A_1^c, A_2^c)$, we start from the Euler condition:

$$E_t (\exp (m_{t+1} + r_{t+1}^c)) = 1,$$

where $r_{t+1}^c = \ln (W_{t+1} / (W_t - C_t))$, which can be rewritten as:

$$E_t (m_{t+1}) + \frac{1}{2} V_t (m_{t+1}) + E_t (r_{t+1}^c) + \frac{1}{2} V_t (r_{t+1}^c) + Cov_t (m_{t+1}, r_{t+1}^c) = 0.$$

The five terms in this equation can be computed explicitly:

$$\begin{aligned} E_t (m_{t+1}) &= E_t [c_0^m + c_1^m \Delta c_{t+1} + c_2^m (A_0^c + A_1^c x_{t+1} + A_2^c \sigma_{t+1}^2 - \kappa_1^c (A_0^c + A_1^c x_t + A_2^c \sigma_t^2))] \\ &= c_0^m + c_1^m \mu_c + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + c_2^m A_0^c (1 - \kappa_1^c) \\ &\quad + c_1^m x_t + c_2^m A_1^c (\rho_x - \kappa_1^c) x_t \\ &\quad + c_2^m A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2 \end{aligned}$$

$$\begin{aligned} m_{t+1} - E_t (m_{t+1}) &= c_1^m \varepsilon_{c,t+1} + c_2^m A_1^c \varepsilon_{x,t+1} + c_2^m A_2^c \varepsilon_{\sigma,t+1} \\ &\equiv \sigma'_m \varepsilon_{t+1}, \end{aligned}$$

$$V_t (m_{t+1}) = \sigma'_m \Sigma_t \sigma_m,$$

$$E_t (r_{t+1}^c) = \kappa_0^c + \mu_c + A_2^c (1 - \rho_\sigma) \mu_\sigma + A_0^c - \kappa_1^c A_0^c + x_t + A_1^c (\rho_x - \kappa_1^c) x_t + A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2,$$

$$\begin{aligned} r_{t+1}^c - E_t (r_{t+1}^c) &= \varepsilon_{c,t+1} + A_1^c \varepsilon_{x,t+1} + A_2^c \varepsilon_{\sigma,t+1}, \\ &\equiv \sigma'_{rc} \varepsilon_{t+1}, \end{aligned}$$

$$V_t (r_{t+1}^c) = \sigma'_{rc} \Sigma_t \sigma_{rc},$$

$$Cov_t (m_{t+1}, r_{t+1}^c) = \sigma'_m \Sigma_t \sigma_{rc}.$$

This results in:

$$\begin{aligned}
& c_0^m + c_1^m \mu_c + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + A_2^c (1 - \rho_\sigma) \mu_\sigma + c_2^m A_0^c (1 - \kappa_1^c) + c_1^m x_t + c_2^m A_1^c (\rho_x - \kappa_1^c) x_t + \\
& c_2^m A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2 + \frac{1}{2} \sigma'_m \Sigma_t \sigma_m + \kappa_0^c + \mu_c + x_t + A_0^c - \kappa_1^c A_0^c \\
& A_1^c (\rho_x - \kappa_1^c) x_t + A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2 + \frac{1}{2} \sigma'_{rc} \Sigma_t \sigma_{rc} + \sigma'_m \Sigma_t \sigma_{rc} = 0.
\end{aligned}$$

By matching the coefficients on the constant, x_t , and σ_t^2 , we find the solutions for A_0^c , A_1^c , and A_2^c :

$$\begin{aligned}
0 &= c_0^m + c_1^m \mu_c + (1 + c_2^m) A_2^c (1 - \rho_\sigma) \mu_\sigma + c_2^m A_0^c (1 - \kappa_1^c) + \frac{1}{2} \sigma'_m \Sigma_0 \sigma_m \\
&\quad + \kappa_0^c + \mu_c + A_0^c + \frac{1}{2} \sigma'_{rc} \Sigma_0 \sigma_{rc} + \sigma'_m \Sigma_0 \sigma_{rc} + (1 - \kappa_1^c) A_0^c, \\
0 &= c_1^m + c_2^m A_1^c (\rho_x - \kappa_1^c) + 1 + A_1^c (\rho_x - \kappa_1^c), \\
0 &= c_2^m A_2^c (\rho_\sigma - \kappa_1^c) + \frac{1}{2} \sigma'_m \Sigma_1 \sigma_m + A_2^c (\rho_\sigma - \kappa_1^c) + \frac{1}{2} \sigma'_{rc} \Sigma_1 \sigma_{rc} + \sigma'_m \Sigma_1 \sigma_{rc}.
\end{aligned}$$

We solve this system numerically for (A_0^c, A_1^c, A_2^c) , where we impose:

$$E(wc_t) = A_0^c + A_1^c \mu_\sigma,$$

in (28) and (29).

The price of dividend strips can be computed recursively and are exponentially-affine in the state variables:

$$pd_t^n = A_0^{d(n)} + A_1^{d(n)} x_t + A_2^{d(n)} \sigma_t^2.$$

For a one-period strip, we have:

$$PD_t^1 = E_t(\exp(m_{t+1} + \Delta d_{t+1})),$$

where:

$$\begin{aligned}
E_t(\Delta d_{t+1}) &= \mu_d + \phi x_t, \\
V_t(\Delta d_{t+1}) &= e'_4 \Sigma_t e_4, \\
Cov_t(\Delta d_{t+1}, m_{t+1}) &= e'_4 \Sigma_t \sigma_m,
\end{aligned}$$

with e_4 denotes the fourth unit vector. We then have:

$$E_t(m_{t+1}) + \frac{1}{2} V_t(m_{t+1}) + E_t(\Delta d_{t+1}) + \frac{1}{2} V_t(\Delta d_{t+1}) + Cov_t(\Delta d_{t+1}, m_{t+1}) = A_0^{d(1)} + A_1^{d(1)} x_t + A_2^{d(1)} \sigma_t^2,$$

leading to:

$$c_0^m + c_1^m \mu_c + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + c_2^m A_0^c (1 - \kappa_1^c) + c_1^m x_t + c_2^m A_1^c (\rho_x - \kappa_1^c) x_t + c_2^m A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2 + \frac{1}{2} \sigma'_m \Sigma_t \sigma_m + \mu_d + \phi x_t + \frac{1}{2} e'_4 \Sigma_t e_4 + e'_4 \Sigma_t \sigma_m = A_0^{d(n)} + A_1^{d(1)} x_t + A_2^{d(1)} \sigma_t^2,$$

and thus:

$$\begin{aligned} A_0^{d(1)} &= c_0^m + c_1^m \mu_c + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + \frac{1}{2} \sigma'_m \Sigma_0 \sigma_m \\ &\quad + \mu_d + \frac{1}{2} e'_4 \Sigma_0 e_4 + e'_4 \Sigma_0 \sigma_m + c_2^m A_0^c (1 - \kappa_1^c), \\ A_1^{d(1)} &= c_1^m + c_2^m A_1^c (\rho_x - \kappa_1^c) + \phi, \\ A_2^{d(1)} &= c_2^m A_2^c (\rho_\sigma - \kappa_1^c) + \frac{1}{2} \sigma'_m \Sigma_1 \sigma_m + \frac{1}{2} e'_4 \Sigma_1 e_4 + e'_4 \Sigma_1 \sigma_m. \end{aligned}$$

The general recursion follows from:

$$\begin{aligned} PD_t^n &= E \left(M_{t+1} PD_{t+1}^{n-1} \frac{D_{t+1}}{D_t} \right) \\ &= E_t \left(\exp \left(m_{t+1} + \Delta d_{t+1} + pd_{t+1}^{n-1} \right) \right). \end{aligned}$$

We first compute the moments of $\Delta d_{t+1} + pd_{t+1}^{n-1}$:

$$\begin{aligned} E_t \left(\Delta d_{t+1} + pd_{t+1}^{n-1} \right) &= \mu_d + A_2^{d(n-1)} (1 - \rho_\sigma) \mu_\sigma + \phi x_t \\ &\quad + A_0^{d(n-1)} + A_1^{d(n-1)} \rho_x x_t + A_2^{d(n-1)} \rho_\sigma \sigma_t^2, \\ \Delta d_{t+1} + pd_{t+1}^{n-1} - E_t \left(\Delta d_{t+1} + pd_{t+1}^{n-1} \right) &= \varepsilon_{d,t+1} + A_1^{d(n-1)} \varepsilon_{x,t+1} + A_2^{d(n-1)} \varepsilon_{\sigma,t+1} \\ &\equiv \sigma_{pd}^{n'} \varepsilon_{t+1}, \\ V_t \left(\Delta d_{t+1} + pd_{t+1}^{n-1} \right) &= \sigma_{pd}^{n'} \Sigma_t \sigma_{pd}^n, \\ Cov_t \left(\sigma_{pd}^{n'} \varepsilon_{t+1}, \sigma'_m \varepsilon_{t+1} \right) &= \sigma_{pd}^{n'} \Sigma_t \sigma_m. \end{aligned}$$

This implies:

$$\begin{aligned} &c_0^m + c_1^m \mu_c + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + c_2^m A_0^c (1 - \kappa_1^c) + \\ &c_1^m x_t + c_2^m A_1^c (\rho_x - \kappa_1^c) x_t + c_2^m A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2 + \frac{1}{2} \sigma'_m \Sigma_t \sigma_m + \\ &\mu_d + A_2^{d(n-1)} (1 - \rho_\sigma) \mu_\sigma + \phi x_t + A_0^{d(n-1)} + A_1^{d(n-1)} \rho_x x_t + A_2^{d(n-1)} \rho_\sigma \sigma_t^2 + \frac{1}{2} \sigma_{pd}^{n'} \Sigma_t \sigma_{pd}^n + \sigma_{pd}^{n'} \Sigma_t \sigma_m \\ = &A_0^{d(n)} + A_1^{d(n)} x_t + A_2^{d(n)} \sigma_t^2, \end{aligned}$$

implying for the coefficients:

$$\begin{aligned}
A_0^{d(n)} &= c_0^m + c_1^m \mu_c + c_2^m A_0^c (1 - \kappa_1^c) + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + \frac{1}{2} \sigma_m' \Sigma_0 \sigma_m + \mu_d \\
&\quad + A_2^{d(n-1)} (1 - \rho_\sigma) \mu_\sigma + A_0^{d(n-1)} + \frac{1}{2} \sigma_{pd}' \Sigma_0 \sigma_{pd}^n + \sigma_{pd}' \Sigma_0 \sigma_m, \\
A_1^{d(n)} &= c_1^m + c_2^m A_1^c (\rho_x - \kappa_1^c) + \phi + A_1^{d(n-1)} \rho_x, \\
A_2^{d(n)} &= c_2^m A_2^c (\rho_\sigma - \kappa_1^c) + \frac{1}{2} \sigma_m' \Sigma_1 \sigma_m + A_2^{d(n-1)} \rho_\sigma + \frac{1}{2} \sigma_{pd}' \Sigma_1 \sigma_{pd}^n + \sigma_{pd}' \Sigma_1 \sigma_m.
\end{aligned}$$

Finally, the one-period risk-free rate is given by:

$$\begin{aligned}
-r_t &= E_t(m_{t+1}) + \frac{1}{2} V_t(m_{t+1}) \\
&= c_0^m + c_1^m \mu_c + c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma + c_2^m A_0^c (1 - \kappa_1^c) \\
&\quad + c_1^m x_t + c_2^m A_1^c (\rho_x - \kappa_1^c) x_t + c_2^m A_2^c (\rho_\sigma - \kappa_1^c) \sigma_t^2 + \frac{1}{2} \sigma_m' \Sigma_t \sigma_m \\
&= -r_0 - r_x x_t - r_\sigma \sigma_t^2,
\end{aligned}$$

where

$$\begin{aligned}
r_0 &= -c_0^m - c_1^m \mu_c - c_2^m A_2^c (1 - \rho_\sigma) \mu_\sigma - c_2^m A_0^c (1 - \kappa_1^c) - \frac{1}{2} \sigma_m' \Sigma_0 \sigma_m, \\
r_x &= -c_1^m - c_2^m A_1^c (\rho_x - \kappa_1^c), \\
r_\sigma &= -c_2^m A_2^c (\rho_\sigma - \kappa_1^c) - \frac{1}{2} \sigma_m' \Sigma_1 \sigma_m.
\end{aligned}$$

In the model of Bansal and Yaron (2004) it is assumed that:

$$\begin{aligned}
\Sigma_t &= \Sigma_0 + \Sigma_1 \sigma_t^2 \\
&= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_w^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \varphi_e^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi_d^2 \end{bmatrix} \sigma_t^2.
\end{aligned}$$

E Dividend strips in the rare disasters model

The setup of the Barro-Rietz rare disasters model as presented by Gabaix (2009) is as follows. Let there be a representative agent with utility given by:²¹

$$E_0 \left[\sum_{t=0}^{\infty} e^{-\rho t} \frac{C_t^{1-\gamma}}{1-\gamma} \right] \tag{30}$$

²¹We thank Xavier Gabaix for providing us with this derivation.

At each period consumption growth is given by:

$$\frac{C_{t+1}}{C_t} = e^g \times \begin{cases} 1 & \text{if there is no disaster at time } t+1 \\ B_{t+1} & \text{if there is a disaster at time } t+1 \end{cases} \quad (31)$$

The pricing kernel is then given by:

$$\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 & \text{if there is no disaster at time } t+1 \\ B_{t+1}^{-\gamma} & \text{if there is a disaster at time } t+1 \end{cases} \quad (32)$$

where $\delta = \rho + g$. The dividend process for stock i takes the form:

$$\frac{D_{i,t+1}}{D_{it}} = e^{g_{iD}} (1 + \varepsilon_{i,t+1}^D) \times \begin{cases} 1 & \text{if there is no disaster at time } t+1 \\ F_{i,t+1} & \text{if there is a disaster at time } t+1 \end{cases} \quad (33)$$

where $\varepsilon_{i,t+1}^D > -1$ is an independent shock with mean 0 and variance σ_D^2 , and $F_{i,t+1} > 0$ is the recovery rate in case a disaster happens. The resilience of asset i is defined as:

$$H_{it} = p_t E_t^D [B_{t+1}^{-\gamma} F_{i,t+1} - 1]$$

where the superscript D signifies conditioning on the disaster event. Define $\hat{H}_{it} = H_{it} - H_{i*}$, which follows a near-AR(1) process given by:

$$\hat{H}_{i,t+1} = \frac{1 + H_{i*}}{1 + H_{it}} e^{-\phi_H} \hat{H}_{it} + \varepsilon_{i,t+1}^H$$

where $\varepsilon_{i,t+1}^H$ has a conditional mean of 0 and a variance of σ_H^2 , and $\varepsilon_{i,t+1}^H$ and $\varepsilon_{i,t+1}^D$ are uncorrelated with the disaster event. Under the assumptions above, the stock price is given by:

$$P_{it} = \frac{D_{it}}{1 - e^{-\delta_i}} \left(1 + \frac{e^{-\delta_i - h_{i*}} \hat{H}_{it}}{1 - e^{\delta_i - \phi_H}} \right)$$

where

$$\begin{aligned} \delta_i &= \delta - g_{iD} - h_{i*} \\ h_{i*} &= \ln H_{i*} \end{aligned}$$

Gabaix (2009) shows that the price at time t of a dividend paid in n periods is given by:

$$D_{it}^{(n)} = D_{it} e^{-\delta_i T} \left(1 + \frac{1 - e^{\phi_H n}}{\phi_H} \hat{H}_{it} \right)$$

and that the expected return on the strip, conditioning on no disaster is given by:

$$E_t [\ln R_{n,t+1}] = E_t \left[\ln \frac{D_{t+1}^{(n-1)}}{D_t^{(n)}} \right] \approx \delta - H_{it}$$

The expected return is the same across maturities, because strips of all maturities are exposed to the same risk in a disaster.

The volatility of the linearized return is given by:

$$\sigma_{n,t} = \sqrt{\sigma_D^2 + \left(\frac{1 - e^{-\phi_H n}}{\phi_H} \right)^2 \sigma_H^2}$$

which is increasing with maturity, due to the fact that higher duration cash flows are more exposed to discount rate shocks than short duration cash flows. Given that the expected return is constant across maturities and the volatility is increasing with maturity, the Sharpe ratio is decreasing with maturity.

F Dividend strips in the Lettau and Wachter model

In the model of Lettau and Wachter (2007), the stochastic discount factor is assumed to be of the form:

$$M_{t+1} = \exp(-r_f - \frac{1}{2}x_t^2 + x_t \varepsilon_{d,t+1}) \quad (34)$$

where x_t drives the price of risk and follows an AR(1) process:

$$x_{t+1} = (1 - \phi_x)\bar{x} + \phi_x x_t + \sigma_x \varepsilon_{t+1} \quad (35)$$

where ε_{t+1} is a 3x1 vector of shocks and σ_x is 1x3 vector. Dividend growth is predictable and given by:

$$\Delta d_{t+1} = g + z_t + \sigma_d \varepsilon_{t+1}, \quad (36)$$

where

$$z_{t+1} = \rho_z z_t + \sigma_z \varepsilon_{t+1}. \quad (37)$$

Lettau and Wachter (2007) derive the prices of dividend strips in their model.