On the Timing and Pricing of Cash Flows*

Jules H. van Binsbergen[†]
Stanford GSB

Michael W. Brandt[‡]
Fuqua School of Business
and NBER

Ralph S.J. Koijen[§]
Chicago Booth

This version: April 2010

Abstract

We study the pricing of short-term assets that are claims to the dividends of the aggregate stock market for a period of up to three years. To compute these prices, we apply put-call parity to a newly constructed, and importantly better synchronized, data set of liquid, exchange-traded S&P500 index options. We compare the asset pricing properties of the claim to short-term dividends to the pricing of the aggregate stock market, which is the claim to all future dividends. We find that the short-term asset has high expected returns, a beta to the market of 0.5, is excessively volatile, and has returns that are highly predictable. The returns on short-term dividend claims cannot be explained by standard asset pricing models, which makes such claims important candidate test assets. We compare our empirical results to their theoretical equivalents in leading asset pricing models and find that none of them predict the empirical findings we document.

*First draft: June 2009. We are grateful to 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, Mark Hendricks, Dirk Jenter, Martin Lettau, Pascal Maenhout, Toby Moskowitz, Christian Mueller-Glissmann, Stefan Nagel, Anamaria Pieschacon, Sheridan Titman, Otto Van Hemert, Stijn Van Nieuwerburgh, Pietro Veronesi, Rob Vishny, Jessica Wachter, Motohiro Yogo and seminar participants at Berkeley, Chicago Booth, Notre Dame, UT Austin, and the Empirical Asset Pricing Retreat at the University of Amsterdam for useful comments and discussions. We thank CRSP for generous financial support.

[†]jvb2@gsb.stanford.edu, (650)-721-1353

[‡]mbrandt@duke.edu, (919)-660-1948

[§]ralph.koijen@chicagobooth.edu, (773)-834-4199.

The asset pricing literature on stocks mainly focuses on claims to all future dividends.[1] However, analogously to zero-coupon bonds, which contain information about discounting at different horizons for fixed income securities, having information on the prices of dividends at different horizons can improve our understanding of equity prices. To this end, we study the prices of assets which entitle the holder to the realized dividends of the aggregate stock market for a period of up to three years. We refer to these assets as "short-term" assets. To compute the prices of these short-term assets, we apply put-call parity to a newly constructed, and importantly better synchronized, data set of liquid, exchange-traded S&P500 index options. We then compare the asset pricing properties of the claims to short-term dividends to the pricing of the aggregate stock market, which is the claim to all future dividends.

Using this approach, we document seven properties of short-term assets:

1. The price of the first two years of dividends constitutes about 5% of the total price.
2. Expected returns and Sharpe ratios on the short-term asset are higher than on the aggregate market.
3. The return volatility of the short-term asset is higher than on the aggregate market.
4. The CAPM beta with respect to the aggregate stock market is 0.5.
5. The CAPM alpha is 10% per annum.
6. The prices of short-term dividends are more volatile than their realizations, pointing to excess volatility on the short end of the equity curve.
7. The returns on the short-term asset are strongly predictable.

These properties have important implications for empirical and theoretical asset pricing. The first five properties, combined with the fact that the CAPM alphas are virtually unaffected if we include size, value or momentum factors, suggest that the short-term assets are a new set of test assets that might be useful in cross-sectional asset pricing tests.[2] Second, 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 movements in discount rates, thereby giving rise to excess volatility. We show, however, that the same phenomenon arises at the short end of the equity claims curve. This suggests that a complete explanation of excess volatility is able to generate excess volatility both for the aggregate stock market and for short-term dividends. The excess

[1]Four notable exceptions are Da (2009), Lettau and Wachter (2007), Hansen, Heaton, and Li (2008), and Croce, Lettau, and Ludvigson (2009).

[2]Lewellen, 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 a new set of test assets.

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 seventh property.

As we are the first to empirically explore the pricing properties of these short-term assets, it is useful to compare them to a theoretical benchmark. Recent consumption-based asset pricing models have made substantial progress in explaining many asset pricing puzzles across various markets. 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). All of these models allow us to directly compute the implied theoretical price of short-term assets. We find that these models are not able to reproduce the facts that we document, which suggests that our short-term assets offer an interesting and non-trivial new set of moments to match.

Our paper also 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 an upward sloping term structure of expected returns 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 would imply that there is a growth premium, not a value premium. Lettau and Wachter (2007) propose a reduced-form model that generates a downward sloping term structure of expected returns and 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 downward sloping term structure. Croce, Lettau, and Ludvigson (2009) argue that the long run risk model as proposed by Bansal and Yaron (2004) also generates an upward sloping term structure of expected returns. However, if the agent in the model can not distinguish between short-term and long-term shocks, the term structure can be downward sloping.

Studying the properties of the 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] Several banks are offering OTC dividend swaps on a range of stock indices. With such a contract, the dividend purchaser pays the market-implied level that is derived from an equity index derivative multiplied by the overall exposure per index point (the fixed leg). The counterparty, with a long position in the equity index, pays the realized dividend level multiplied by the exposure per index point.

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

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 since the previous reset date and is used by derivative traders to hedge their dividend positions. Further, 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. Wilkens and Wimschulte (2009) discuss the European market of dividend futures that started mid-2008.

The basic idea of the paper is summarized in the following three equations. The price of a stock or equity index S_t is given by the discounted value of its dividends D_t:

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

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

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

where the short-term asset is the price of the first few dividends, and the long-term asset is the price of the remaining dividends. We use $\mathcal{P}_{t,T}$ to denote the value of the short-term asset given by:

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

We synthetically construct the price of the first few dividend payments using tick-level options data obtained from Market Data Express, the data vendor of the CBOE. In principle, any derivative contract can be used to back out the price of dividend strips up until the maturity of the derivative contract. We choose to use option contracts for most of our analysis because there is a liquid market for these contracts even for longer maturities. There is a liquid market for futures contracts for shorter maturities, but the liquidity drops markedly as the maturity of the futures contracts increases. For maturities for which both futures and options contracts are available, the implied prices of the short-term asset are very similar.

We proceed as follows. In Section 1, we discuss various ways to trade the short-term asset, either by creating it synthetically using index derivatives or by trading dividend derivatives directly. Section 2 discusses our dataset. In Section 3 we discuss the empirical results. In Section 4, we compare our findings with several leading asset pricing models. We present several robustness checks in Section 5. In Section 6 we discuss several possible extensions. Section 7 concludes.

1 The market for dividends

In this section, we describe several ways to trade dividends separately from the capital gains component. First, the short-term assets can be created synthetically using S&P500 index options or futures contracts. This is the approach we follow empirically in this paper. In 1990, the Chicago Board Options Exchange (CBOE) introduced Long-Term Equity Anticipation Securities (LEAPS), which are options with a relatively long maturity ranging up to three years. The maximum maturity of LEAPS for the sample period in our dataset is displayed in Figure 1. As a result of the issuance cycle of LEAPS, the maximum available maturity fluctuates over time. We can then use the put-call parity relationship for a European option to compute its value:

$$\mathcal{P}_{t,T} = p_{t,T} - c_{t,T} + S_t - X\mathrm{e}^{-r_{t,T}(T-t)}, \tag{2}$$

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. We use TAQ options data on the S&P500 index to measure the prices of the short-term asset as accurately as possible.

One potential disadvantage of replicating the asset described above, is that a long position in the index is required. As index replication is not costless, we also consider investing in a so-called dividend steepener, given by:

$$\begin{aligned}\mathcal{P}_{t,T_1,T_2} &= \mathcal{P}_{t,T_2} - \mathcal{P}_{t,T_1} \\ &= p_{t,T_2} - p_{t,T_1} - c_{t,T_2} + c_{t,T_1} - X\left(\mathrm{e}^{-r_{t,T_2}(T_2-t)} - \mathrm{e}^{-r_{t,T_1}(T_1-t)}\right).\end{aligned} \tag{3}$$

This asset entitles the holder to the dividends paid out between period T_1 and T_2, $T_1 < T_2$. 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. Furthermore, this strategy does not involve any dividend payments until time T_1, which could be important for tax reasons. Finally, the steepener is interesting 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 as to increase them when the economy rebounds. By choosing T_1 further into the future, one takes a bet on a later recovery.[4]

A second way to synthetically create the short-term asset is by using futures contracts. The cost of carry formula for equity futures implies:

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

By applying the cost of carry formula for equity to two different maturities, the price of the dividend steepener for period T_1 to T_2, $T_1 < T_2$, can be computed as:

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

The steepener now 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 that the prices obtained from either market are close indeed.

In addition to synthetically creating the short-term asset, it is now also possible to trade dividends directly via dividend derivatives such as dividend swaps, dividend futures, and dividend options. Most transactions are OTC, but several exchange traded products have been introduced recently. For instance, the CBOE announced in December 2009 to introduce options on S&P500 index dividends. This development follows the introduction of an array of dividend derivatives at the Eurex. The Eurex introduced on June 30, 2008 dividend futures on the Dow Jones EURO STOXX 50 Index,[5] and in February 2010, futures are available on five different indices.[6] 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 at March 30 2010, illustrating the

[4]See also "Dividend Swaps Offer Way to Pounce on a Rebound," Wall Street Journal, April 2009.

[5]See `http://www.eurexchange.com/download/documents/publications/index_dividend_swaps_1_en.pdf` for more information.

[6]A more detailed description can be found at: `http://www.eurexchange.com/download/documents/publications/Eurex_Produkte_2010_en.pdf`.

rapid developments of dividend trading.[7]

2 Data and dividend strategies

2.1 Data sources

We use a new data set provided by the CBOE on intra-day trades and quotes on S&P500 options between January 1996 and May 2009 with maturities up to 3 years. The data contains all options contracts for which the S&P500 index is the underlying asset. We obtain tick-level data between January 1996 and May 2009 of the index values and futures prices of the S&P500 index from Tick Data Inc. Further, equation (5) requires a continuously-compounded interest rate as input. This interest rate is calculated from a collection of continuously-compounded zero-coupon interest rates at various maturities and provided by OptionMetrics. This zero curve is derived from BBA LIBOR rates and settlement prices of CME Eurodollar futures.[8] For a given option, the appropriate interest rate input corresponds to the zero-coupon rate that has a maturity equal to the options expiration date, and is obtained by linearly interpolating between the two closest zero-coupon rates on the zero curve.[9] Finally, we obtain daily return data with and without distributions from S&P index services. Daily dividends can then be computed by taking the difference between these two returns and multiplying by the lagged value of the index.

For our purposes, it is important that the inputs in the put-call parity formula are recorded as close to each other as possible during the day.[10] Our option dataset combined with index values from Tick Data allows us to match trades within a minute interval. Using closing prices as available in 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)). Instead, we therefore select option quotes for puts and calls 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 10-11am or 1-2pm has no effect on our results, as we demonstrate in Section 5.

To construct the prices of dividend strips, we first find all couples of put and call

[7] The daily dollar volume averages to $33.4 Billion in 2009, see: `http://www.reuters.com/article/idUSLDE60A10020100111`.

[8] We also use data from Bloomberg to replicate the OptionMetrics yield curves and using this interest rate instead is inconsequential for the results.

[9] Alternative interpolation schemes give the same results at the reported precision.

[10] This is particularly important for the options and index data and less so for the interest rate data.

contracts that have the same maturity and strike price within at the last trading day of a particular month. 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 provided within the same second of the day. For each of these matches, we use the put-call parity to calculate the price of the dividend strip. 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, for instance as a result of the introduction of electronic trading, but even in the first year of our sample, on average have nearly a thousand matches per maturity on a given trading day for maturities between 1 and 2 years.

2.2 Dividend prices

We first construct for each date t and for all maturities T longer than three months the prices of the dividend claims:

$$\mathcal{P}_{t,T} = p_{t,T} - c_{t,T} + S_t - X\mathrm{e}^{-r_{t,T}(T-t)}, \tag{5}$$

where $p_{t,T}$ and $c_{t,T}$ are the prices of a European put and call option on the S&P500 index at time t, with maturity T and strike price X, and S_t is the value of the index at time t.

The set of maturities T of these claims is not constant and varies depending on the option issuing cycle. On average there are around six maturities greater than three months available at any particular time.

2.3 Returns on dividend strategies

Apart from reporting dividend prices, we also implement two simple trading strategies. The first trading strategy goes long in the dividend claim. We hold this claim for one month, receive the dividends within that month, and subsequently sell the claim at the end of the next month. For example, on January 30^{th} 2009, we go long in a 1.411 year dividend claim, which entitles the holder to the first 1.411 years of dividends on the index. We hold this contract for one month, collect the dividends between January 30^{th} and February

27^{th} and sell the dividend claim on February 27^{th}, which now has a remaining maturity of 1.329 years. The monthly return series on this strategy is given by:

$$R_{1,t+1} = \frac{\mathcal{P}_{t+1,T-1} + D_{t+1}}{\mathcal{P}_{t,T}}. \tag{6}$$

The implementation of the trading strategy above assumes that the investor can costlessly replicate the index. To avoid a long position in the index, we also consider the price at time t of the dividends between T_1 and T_2, given by:

$$\begin{aligned} \mathcal{P}_{t,T_1,T_2} &= \mathcal{P}_{t,T_2} - \mathcal{P}_{t,T_1} \\ &= p_{t,T_2} - p_{t,T_1} - c_{t,T_2} + c_{t,T_1} - X\left(\mathrm{e}^{-r_{t,T_2}(T_2-t)} - \mathrm{e}^{-r_{t,T_1}(T_1-t)}\right), \end{aligned} \tag{7}$$

where $T_2 > T_1 > 1$. We hold this claim for one month and sell it in the next period. The return on this strategy, called strategy 2 hereafter, is given by:

$$R_{2,t+1} = \frac{\mathcal{P}_{t+1,T_1-1,T_2-1}}{\mathcal{P}_{t,T_1,T_2}}. \tag{8}$$

This return strategy 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. Furthermore, this return strategy does not involve any dividend payments until time T_1. Further details on the implementation of these strategies can be found in Appendix A.

3 Main empirical results

In this section, we document the seven facts about the prices and returns on the short-term asset listed in the introduction.

3.1 Properties of dividend prices

Figure 3 displays the prices of the first 0.5, 1, 1.5, and 2 years of dividends during our sample period. As expected, the dividend prices drop during the two recessions in our sample period, as expected growth of dividends drops and discount rates increase. This effect is more pronounced for the 2-year price. The 0.5 year price of is much more stable.

As dividend prices are non-stationary, 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 0.5, 1, 1.5, and 2 year dividends as a fraction of the index value. The ratios are highly correlated.

They drop between 1997 and 2001, and slowly increase afterwards. The relative prices of the dividend steepener can be found by taking the difference of these ratios across maturities. Second, and perhaps more striking, is that the first two years of dividends seems low relative to the total. The price of the first two years of dividends only makes up a small fraction of the total price, only once exceeding 5%, and with an average of 3.36%. At the bottom in 2001, this fraction is lower than 2% of the total value of the S&P500 index. As a point of reference, if one considers a Gordon growth model with a constant discount rate of 10% and a dividend growth rate of 6%, then the first two years of dividends amount to just under 8% of the total index value. We show below that in recent asset pricing models, the ratio is closer to 10% on average. Even in May 2009, when presumably risk premia were high and expected dividend growth rates were low, which would imply that the first dividends should be a large fraction of the index, the first two years of dividends only make up just under 5% of the total index value in our data-set. This result provides the first piece of evidence that short-term dividends seem relatively cheap.

3.2 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 92.2%. Panel A of Table 1 lists the summary statistics alongside the same statistics for the S&P500 and the market for the full sample period. For the market, we use the CRSP value-weighted return of all stocks traded on the AMEX, NYSE, and Nasdaq. Both dividend strategies have a high monthly average return equal to 1.20% (annualized 14.4%) for trading strategy 1 and 1.15% (annualized 13.8%) for trading strategy 2. Over the same period, the average return on the market portfolio was 0.54% (annualized 6.5%) and the return on the S&P500 index was only 0.49% (annualized 5.9%). The higher average returns also come with a higher level of volatility than both the aggregate stock market and the S&P500 index, with monthly return volatilities of 7.9% for strategy 1 and 9.8% for strategy 2. Over the same period the monthly market volatility is 4.9% and the volatility of the return on the S&P500 index equals 4.7%. The summary statistics also indicate that dividend returns tend to have fatter tails than both equity indices. Despite the higher volatility, the dividend strategies result in substantially higher Sharpe ratios. The Sharpe ratios of the dividend strategies are about twice as high as the Sharpe ratios of both the aggregate stock market and 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 lower for dividend claims with longer 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 for each return series and for the returns on the aggregate stock market. In Figure 10, we show that the volatility of dividend returns and the aggregate stock market broadly follow the same pattern. The correlation between the volatility of the dividend returns of strategy 1 and the S&P500 is 0.52. Table 4 reports the estimates of the GARCH(1,1)-specification, illustrating that the parameters of the volatility equations are very similar as well.

The volatility of the two return strategies was substantially higher before 2003 than after. Table 1 therefore 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 equity indices. We find consistently across both sample periods that the average return on the dividend strategies is about twice as high as the average return on the market or 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 much closer to the levels of volatility that we record for the equity indices. The Sharpe ratios of the dividend strategies are comparable across subperiods, and always substantially higher than the ones of either the aggregate stock market or the S&P500 index. Overall, the conclusions we draw from the full sample are consistent with our findings in both sub-samples.

Table 2 presents OLS regressions of the two trading strategies on the market portfolio in excess of the one-month short rate.[11] We find that both dividend strategies have a CAPM beta of 0.49. Secondly, R^2 values of the regression are low and close to 10%. The intercept of the regression equals 0.79% for strategy 1 and 0.73% for strategy 2, which in annualized terms equals 9.48% and 8.76%. 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.

Our return series does exhibit some negative autocorrelation, pointing to mean-reversion in dividend returns. We return to the predictability of dividend returns in

[11] Using S&P500 index returns instead of the market portfolio leads to almost identical regression results.

Section 3.4. Including an autoregressive (AR) term in the regression explains around 8-12% of the variation in returns. Once we include this additional regressors, we find that the intercept is statistically significant at the 10%.[12]

Table 3 presents regression results for the Fama-French three factor model, in which we also include the AR(1)-term in the second and fourth column. The market beta is unaffected by the additional factors and is estimated between 0.5 and 0.6, depending on the strategy and specification. We find positive loadings on the book-to-market factor, which seems consistent with duration-based explanations of the value premium. The loading is statistically significant in case of the dividend steepener. The coefficient on the size portfolio switches sign depending on the specification, and has very low significance. Perhaps most interestingly, the intercepts are hardly affected by including additional factors; monthly alphas are estimated between 0.56% and 0.70%. These results suggest that the short-term asset has rather high expected returns that cannot be explained easily by standard asset pricing models.

The high monthly alphas compensate investors for the risk in the dividend strategy that cannot be explained by the other priced factors. Our results becomes 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 (2006), 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 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 summarized in Table 5. 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. Further, both expected returns and expected dividend growth rates have a persistent component, but expected returns are more persistent than expected dividend growth rates. Interestingly, as the starting date of our sample period increases, both the R^2 value of returns as well as the R^2 value of dividend growth rates strongly increases. Over the data period starting

[12] The reported intercept in the AR specification is comparable to the intercepts of the other specifications and is adjusted for the persistence coefficient.

in 1989, we find an R^2 value for dividend growth rates of 56%.

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. The unpredictable part of dividend growth, which is rather small, then needs to be highly priced.

3.3 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 9 plots the price of the next year of dividends and the realized dividends during the next year. We shift the latter time series such that the price and 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" at the short end of the equity curve. 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 assets.

3.4 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 dividend strategies. Some of this evidence is present already in Table 2, which shows that dividend returns are to a certain extent mean-reverting. We extend this evidence by regressing monthly dividend returns on the lagged value of $\mathcal{P}_{t,T}/D_t$. This is the equivalent of the price-dividend ratio for the short-term asset.[13] The results are presented in Table 6. We find

[13] See, among others, Fama and French (1988), Campbell and Shiller (1988), Cochrane (1991), Cochrane (2006), Lettau and Van Nieuwerburgh (2006), Wachter and Warusawitharana (2009), and Binsbergen and

that $\mathcal{P}_{t,T}/D_t$ forecasts dividend returns with a negative sign, and is highly significant. We use OLS standard errors to determine the statistical significance of the predictive coefficient. To mitigate concerns regarding measurement error in the predictor variable, we smooth $\mathcal{P}_{t,T}/D_t$ over three months and use this predictor variable instead.[14] The results are reported in the second column and are very comparable to the first column. In the third column, we simply add an AR(1)-term, which enters significantly. Nevertheless, the coefficient on $\mathcal{P}_{t,T}/D_t$ is still negative and highly significant.

4 Comparison with asset pricing models

We then compare our findings to the implications for dividend strips of several recent asset pricing models, including the Campbell and Cochrane (1999) external habit formation framework, the Bansal and Yaron (2004) long run risk model, the Barro-Rietz rare disasters framework (Barro (2006)) as presented by Gabaix (2009), and the Lettau and Wachter (2007) model.

Apart from the Lettau and Wachter (2007) model, which is designed to generate a downward sloping term structure of expected returns, all the models above generate either an upward sloping term structure of expected returns (habit formation and long run risk) or a flat term structure of expected returns (rare disasters). In terms of volatilities, the term structure is upward sloping for all the models we consider, with the exception of Lettau and Wachter (2007) who have a slightly upward sloping term structure for the first 8 years and downward sloping thereafter. The Sharpe ratio is upward sloping in the habit formation and long run risk models and downward sloping in the rare disaster and Lettau and Wachter (2007) model.

Because the expected return is so low for both the habit formation model and the long-run risk model, the fraction of the net present value of the first two years of dividends makes up a very large fraction of the total index value, with an average above 8%, and virtually never less than 6%. Recall that the net present value of dividends of the first 2 years that we uncover in the data never exceeds 6% over the 1996-2009 period and averages around 4%.

Koijen (2010) for the predictability of returns by the dividend yield.

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

4.1 External habit formation

In this section we study the term structure of expected stock returns, the term structure of the volatility of stock returns, and the term structure of the Sharpe ratio in a Campbell and Cochrane (1999) habit formation model. In this case, the stochastic discount factor is given by:

$$M_{t+1} = \delta G^{-\gamma} e^{-\gamma(s_{t+1}-s_t+v_{t+1})} \tag{9}$$

where G represents 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:

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1}, \tag{10}$$

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:

$$\Delta d_{t+1} = g + w_{t+1} \tag{11}$$

We use the same calibrated monthly parameters as in Campbell and Cochrane (1999) and set the correlation between the shocks v_{t+1} and w_{t+1} equal to 0.2. 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:

$$D_t^{(1)} = E_t\left(M_{t+1}D_{t+1}\right) = D_t E_t\left(M_{t+1}\frac{D_{t+1}}{D_t}\right). \tag{12}$$

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

$$D_t^{(n)} = E_t\left(M_{t+1}D_{t+1}^{n-1}\right) \tag{13}$$

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

$$R_{n,t+1} = \frac{D_{t+1}^{(n-1)}}{D_t^{(n)}} \tag{14}$$

We simulate from the model and compute for each dividend strip n the average annualized excess return (risk premium), $E(R_{n,t+1}) - R_f$, the annualized volatility $\sigma(R_{n,t+1})$ and the Sharpe ratio, which is the ratio of the previous two quantities. The results are plotted in Figure 12 for the first 480 dividend strips (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 average excess return equal to 1%.

4.2 Long-run risks

We then consider a long run risk model. We use the model and monthly calibration by Bansal and Shaliastovich (2009) which is designed to match return moments across stock, bond and foreign exchange markets. However, highly comparable results are achieved by using the model and calibration by either Bansal and Yaron (2004) or Bansal, Kiku, and Yaron (2006). The log stochastic discount factor in this model is given by:

$$\begin{aligned} m_{t+1} &= \mu_s + s_x x_t + s_{gs}\left(\sigma_{gt}^2 - \sigma_g^2\right) + s_{xs}\left(\sigma_{xt}^2 - \sigma_x^2\right) \\ &\quad -\lambda_\eta \sigma_{gt}\eta_{t+1} - \lambda_e \sigma_{xt} e_{t+1} - \lambda_{gw}\sigma_{gw} w_{g,t+1} - \lambda_{xw}\sigma_{xw} w_{x,t+1}. \end{aligned}$$

The processes for consumption growth Δc_{t+1}, dividend growth Δd_{t+1} are given by:

$$\begin{aligned} \Delta c_{t+1} &= x_t + \mu_g + \sigma_{gt}\eta_{t+1}, \\ \Delta d_{t+1} &= \mu_d + \phi_x x_t + \varphi_d \sigma_{gt}\eta_{d,t+1} \end{aligned}$$

The three state variables in the model are x_t, which is the slowly time-varying mean of consumption and dividend growth (the long run risk component), σ_{xt}^2, which is the stochastic variance of the long-run risk component, and σ_{gt}^2, which is the stochastic variance of the short-term risk component.

$$\begin{aligned} x_{t+1} &= \rho x_t + \sigma_{xt} e_{t+1}, \\ \sigma_{g,t+1}^2 &= \sigma_g^2 + \nu_g\left(\sigma_{gt}^2 - \sigma_g^2\right) + \sigma_{gw} w_{g,t+1}, \\ \sigma_{x,t+1}^2 &= \sigma_x^2 + \nu_x\left(\sigma_{xt}^2 - \sigma_x^2\right) + \sigma_{xw} w_{x,t+1}. \end{aligned}$$

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. More details on how to compute the dividend strips are provided in Appendix B. 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.

4.3 Variable rare disasters

We then consider the variable rare disasters model by Gabaix (2009). In this case, the stochastic discount factor is given by:

$$\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 & \text{if there is no disaster at time t+1} \\ B_{t+1}^{-\gamma} & \text{if there is a disaster at time t+1} \end{cases} \tag{15}$$

where B_{t+1} measures the drop in consumption in case a disaster hits, and δ is the sum of the subjective discount factor and the aggregate growth rate. The dividend process for stock i takes the form:

$$\frac{D_{i,t+1}}{D_{it}} = e^{g_{iD}} \left(1 + \varepsilon_{i,t+1}^{D}\right) \times \begin{cases} 1 & \text{if there is no disaster at time t+1} \\ F_{i,t+1} & \text{if there is a distater at time t+1} \end{cases} \tag{16}$$

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:

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

where the superscript D signifies conditioning on the disaster event and p_t is the probability of a disaster. Instead of modeling each component of H_{it}, Gabaix (2009) assumes that $\hat{H}_{it} \equiv H_{it} - H_{i*}$, follows a near-AR(1) process given by:

$$\hat{H}_{i,t+1} = \frac{1 + H_{i*}}{1 + H_{it}} e^{-\phi_H} \hat{H}_{it} + \varepsilon_{i,t+1}^{H} \tag{17}$$

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. Further details are provided in Appendix D.

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 their duration is higher. As a result, the Sharpe ratio is downward sloping.

4.4 Lettau and Wachter (2007)

We finally consider the model by Lettau and Wachter (2007), which is designed to generate a downward sloping term structure of expected returns. In their framework, the stochastic discount factor, which is specified exogenously, is given by:

$$M_{t+1} = exp(-r_f - \frac{1}{2}x_t^2 + x_t\varepsilon_{d,t+1}) \tag{18}$$

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

$$x_{t+1} = (1-\phi_x)\bar{x} + \phi_x x_t + \sigma_x \varepsilon_{t+1} \tag{19}$$

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

$$\Delta d_{t+1} = g + z_t + \sigma_d \varepsilon_{t+1} \tag{20}$$

where

$$z_{t+1} = \rho_z z_t + \sigma_z \varepsilon_{t+1} \tag{21}$$

We use their quarterly calibration and compute dividend strips using the essentially affine structure of the setup.[15] For more details on the calibration and the computation of dividend strips within their model, we refer to Lettau and Wachter (2007).

As before, we report for each dividend strip n the average annualized excess return (risk premium), $E(R_{n,t+1}) - R_f$, the annualized volatility $\sigma(R_{n,t+1})$ and the Sharpe ratio. The results are plotted in Figure 14. 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.

5 Robustness

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

5.1 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 six alternative procedures to construct dividend prices. In all cases, we report the summary statistics of dividend returns for strategy 1, and

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

the CAPM alpha and beta.[16] 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 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 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. 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. Finally, in case of Alternative 6, we again match all call and put contracts based on maturity and the strike price. However, instead of minimizing the time difference first, we first minimize over the bid-ask spread, and for the set of matches with the same spread for a given maturity, we take the minimum time difference. If multiple matches exist for a particular maturity, we take the median across the matches that have the smallest bid-ask spread and time difference.

The results are presented in Table 7, in which Ai corresponds to Alternative i. Even though the numbers change slightly across different matching procedures, which is not unexpected, none of our main results is overturned for any of the cases. 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 seven empirical facts of dividend strategies, and is robust to all seven matching procedures we consider.

5.2 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

[16]The results for dividend steepener are highly comparable and are not reported for brevity.

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. Explanations for all facts solely based on market microstructure are therefore, in our view, less convincing as index futures markets are among the most liquid asset markets available.

6 Further applications

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

6.1 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.[17] 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 this 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.

[17] See for instance Koijen, Lustig, and Van Nieuwerburgh (2009) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010).

6.2 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.[18]

7 Conclusion

We study the pricing of short-term assets whose payoff equals the dividends of the aggregate stock market during a period of up to three years. To compute these prices, we apply the put-call parity to a new data set of liquid, exchange-traded S&P500 options. We compare the asset pricing properties of the claim to short-term dividends to the pricing of the aggregate stock market, which is the claim to all future dividends. Using this approach, we find that the short-term asset has a high expected returns, a beta to the market of 0.5, is excessively volatile, and has returns that are highly predictable. The returns on short-term dividend claims cannot be explained by standard asset pricing models, which makes such claims important candidate test assets. We compare our empirical results to their theoretical equivalents in leading asset pricing models and find that none of them predict the empirical findings we document.

[18] Additional notes are available upon request.

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 (2006): "Stock Return Predictability: Is it There?," *Review of Financial Studies,* forthcoming.

Bansal, R., D. Kiku, and A. Yaron (2006): "Long Run Risks: Estimation and Inference," Unpublished paper, Duke University and the Wharton School.

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," Forthcoming, Journal of Finance.

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.

——— (2006): "The Dog That Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies,* forthcoming.

Cochrane, J. H., and M. Piazzesi (2005): "Bond Risk Premia," *The American Economic Review,* forthcoming.

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.

Da, Z. (2009): "Cash Flow, Consumption Risk, and the Cross-section of Stock Returns," *Journal of Finance*, 64, 923–956.

Duffee, G. (2010): "Sharpe ratios in term structure models," Working Paper.

Fama, E. F., and K. R. French (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3–27.

Gabaix, X. (2009): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," Working paper, NYU.

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.

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 (2006): "Reconciling the Return Predictability Evidence," Forthcoming Review of Financial Studies.

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.

Lewellen, J., S. Nagel, and J. Shanken (2009): "A Skeptical Appraisal of Asset Pricing Tests," Forthcoming, Journal of Financial Economics.

Manley, R., and C. Mueller-Glissmann (2008): "The Market for Dividends and Related Investment Strategies," *Financial Analyst Journal*, 64(3), 17–29.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71, 421–436.

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.

Wilkens, S., and J. Wimschulte (2009): "The Pricing of Dividend Futures in the European Market: A First Empirical Analysis," Unpublished paper, University of Regensburg.

A Details dividend returns

The two trading strategies described in Section 2.3 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, and continue until May 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 Koijen (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:

$$r_{t+1} \equiv \log\left(\frac{S_{t+1} + D_{t+1}}{S_t}\right), \tag{22}$$

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:

$$\mu_{t+1} = \delta_0 + \delta_1 (\mu_t - \delta_0) + \varepsilon^{\mu}_{t+1}, \tag{23}$$
$$g_{t+1} = \gamma_0 + \gamma_1 (g_t - \gamma_0) + \varepsilon^{g}_{t+1}, \tag{24}$$

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^{μ}_{t+1} and ε^{g}_{t+1} is specified below. The realized dividend growth rate is equal to the expected dividend growth rate plus an orthogonal shock:

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

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 $\overline{pd} = E [pd_t]$, $\kappa = \log \left(1 + \exp \left(\overline{pd}\right)\right) - \rho\overline{pd}$ and $\rho = \frac{\exp\left(\overline{pd}\right)}{1+\exp\left(\overline{pd}\right)}$, as in Campbell and Shiller (1988). If we iterate this equation, and using the AR(1) assumptions (23)-(24), 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^{g}_{t+1}), shocks to expected returns (ε^{μ}_{t+1}), and realized dividend growth shocks (ε^{d}_{t+1}), have mean zero, covariance matrix

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

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:

$$R^2_{Ret} = 1 - \frac{\hat{var}\left(r_{t+1} - \mu_t^F\right)}{\hat{var}\left(r_t\right)}, \tag{25}$$

$$R^2_{Div} = 1 - \frac{\hat{var}\left(\Delta d_{t+1} - g_t^F\right)}{\hat{var}\left(\Delta d_{t+1}\right)}, \tag{26}$$

where $\hat{var}$ 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 long-run risks model

In this appendix we derive dividend strips in a long run risk model as calibrated by Bansal and Shaliastovich (2009). In their model, the log stochastic discount factor is given by:

$$\begin{aligned} m_{t+1} &= \mu_s + s_x x_t + s_{gs}\left(\sigma_{gt}^2 - \sigma_g^2\right) + s_{xs}\left(\sigma_{xt}^2 - \sigma_x^2\right) \\ &\quad -\lambda_\eta \sigma_{gt}\eta_{t+1} - \lambda_e \sigma_{xt} e_{t+1} - \lambda_{gw}\sigma_{gw} w_{g,t+1} - \lambda_{xw}\sigma_{xw} w_{x,t+1}. \end{aligned}$$

The short rate follows from:

$$\begin{aligned} E_t\left(\exp\left(m_{t+1}\right)\right) &= \exp\left(E_t\left(m_{t+1}\right) + \frac{1}{2} var_t\left(m_{t+1}\right)\right) \\ &= \exp\left(\begin{array}{c} \mu_s + s_x x_t + s_{gs}\left(\sigma_{gt}^2 - \sigma_g^2\right) + s_{xs}\left(\sigma_{xt}^2 - \sigma_x^2\right) + \\ \frac{1}{2}\left[\lambda_\eta^2 \sigma_{gt}^2 + \lambda_e^2 \sigma_{xt}^2 + \lambda_{gw}^2 \sigma_{gw}^2 + \lambda_{xw}^2 \sigma_{xw}^2\right] \end{array}\right) \\ &= \exp\left(w_0 + w_1 x_t + w_2 \sigma_{gt}^2 + w_3 \sigma_{xt}^2\right) \\ &= \exp\left(-y_t^{(1)}\right), \end{aligned}$$

with:

$$\begin{aligned} w_0 &= \mu_s - s_{gs}\sigma_g^2 - s_{xs}\sigma_x^2 + \frac{1}{2}\left[\lambda_{gw}^2 \sigma_{gw}^2 + \lambda_{xw}^2 \sigma_{xw}^2\right], \\ w_1 &= s_x, \\ w_2 &= s_{gs} + \frac{1}{2}\lambda_\eta^2, \\ w_3 &= s_{xs} + \frac{1}{2}\lambda_e^2. \end{aligned}$$

The state variables satisfy the following dynamics:

$$
\begin{aligned}
\Delta c_{t+1} &= x_t + \mu_g + \sigma_{gt}\eta_{t+1}, \\
x_{t+1} &= \rho x_t + \sigma_{xt} e_{t+1}, \\
\sigma^2_{g,t+1} &= \sigma^2_g + \nu_g \left(\sigma^2_{gt} - \sigma^2_g\right) + \sigma_{gw} w_{g,t+1}, \\
\sigma^2_{x,t+1} &= \sigma^2_x + \nu_x \left(\sigma^2_{xt} - \sigma^2_x\right) + \sigma_{xw} w_{x,t+1}, \\
\Delta d_{t+1} &= \mu_d + \phi_x x_t + \varphi_d \sigma_{gt} \eta_{d,t+1}.
\end{aligned}
$$

All shocks are independent, apart from: $\tau_{gd} = corr\left(\eta_{t+1}, \eta_{d,t+1}\right) = cov\left(\eta_{t+1}, \eta_{d,t+1}\right)$. All unknown coefficients are defined in Koijen, Lustig, VanNieuwerburgh, and Verdelhan (2009). The 1-period dividend strip follows from:

$$
\begin{aligned}
D^{(1)}_t &= D_t E_t \left(M_{t+1} \frac{D_{t+1}}{D_t} \right) \\
&= D_t E_t \left(\exp \left(\begin{array}{c} \mu_s + s_x x_t + s_{gs}\left(\sigma^2_{gt} - \sigma^2_g\right) + s_{xs}\left(\sigma^2_{xt} - \sigma^2_x\right) \\ -\lambda_\eta \sigma_{gt}\eta_{t+1} - \lambda_e \sigma_{xt} e_{t+1} - \lambda_{gw}\sigma_{gw} w_{g,t+1} - \lambda_{xw}\sigma_{xw} w_{x,t+1} \\ +\mu_d + \phi_x x_t + \varphi_d \sigma_{gt}\eta_{d,t+1} \end{array} \right) \right) \\
&= \exp\left(\mu_s + s_x x_t + s_{gs}\left(\sigma^2_{gt} - \sigma^2_g\right) + s_{xs}\left(\sigma^2_{xt} - \sigma^2_x\right) + \mu_d + \phi_x x_t\right) \times \\
&\quad \exp\left(\frac{1}{2}\left[\lambda^2_\eta \sigma^2_{gt} + \lambda^2_e \sigma^2_{xt} + \lambda^2_{gw}\sigma^2_{gw} + \lambda^2_{xw}\sigma^2_{xw}\right] - \lambda_\eta \varphi_d \tau_{gd} \sigma^2_{gt}\right) \\
&= D_t \exp\left(H^{(1)}_0 + H^{(1)}_1 x_t + H^{(1)}_2 \sigma^2_{gt} + H^{(1)}_3 \sigma^2_{xt}\right).
\end{aligned}
$$

This leads to:

$$
\begin{aligned}
H^{(1)}_0 &= \mu_s - s_{gs}\sigma^2_g - s_{xs}\sigma^2_x + \mu_d + \frac{1}{2}\left[\lambda^2_{gw}\sigma^2_{gw} + \lambda^2_{xw}\sigma^2_{xw}\right], \\
H^{(1)}_1 &= s_x + \phi_x, \\
H^{(1)}_2 &= s_{gs} + \frac{1}{2}\lambda^2_\eta - \lambda_\eta \varphi_d \tau_{gd}, \\
H^{(1)}_3 &= s_{xs} + \frac{1}{2}\lambda^2_e.
\end{aligned}
$$

As before, let $D_t^{(n)}$ denote the price of a dividend at time t that is paid out in n period. The following relationship then holds:

$$
\begin{aligned}
D_t^{(n)} &= E_t\left(D_{t+1}^{(n-1)}M_{t+1}\right) \\
&= D_tE_t\left(\exp\left(H_0^{(n-1)}+H_1^{(n-1)}x_{t+1}+H_2^{(n-1)}\sigma_{g,t+1}^2+H_3^{(n)}\sigma_{x,t+1}^2\right)M_{t+1}\frac{D_{t+1}}{D_t}\right) \\
&= D_tE_t\left(\exp\left(\begin{array}{c} H_0^{(n-1)}+H_1^{(n-1)}x_{t+1}+H_2^{(n-1)}\sigma_{g,t+1}^2+H_3^{(n-1)}\sigma_{x,t+1}^2 \\ +\mu_s+s_xx_t+s_{gs}\left(\sigma_{gt}^2-\sigma_g^2\right)+s_{xs}\left(\sigma_{xt}^2-\sigma_x^2\right) \\ -\lambda_\eta\sigma_{gt}\eta_{t+1}-\lambda_e\sigma_{xt}e_{t+1}-\lambda_{gw}\sigma_{gw}w_{g,t+1}-\lambda_{xw}\sigma_{xw}w_{x,t+1} \\ +\mu_d+\phi_xx_t+\varphi_d\sigma_{gt}\eta_{d,t+1} \end{array}\right)\right)
\end{aligned}
$$

which can be rewritten as:

$$
\begin{aligned}
D_t^{(n)} &= D_tE_t\left(\exp\left(\begin{array}{c} H_0^{(n-1)}+H_1^{(n-1)}\rho x_t+H_2^{(n-1)}\left(\sigma_g^2+\nu_g\left(\sigma_{gt}^2-\sigma_g^2\right)\right) \\ +H_3^{(n-1)}\left(\sigma_x^2+\nu_x\left(\sigma_{xt}^2-\sigma_x^2\right)\right)+\mu_s+s_xx_t+s_{gs}\left(\sigma_{gt}^2-\sigma_g^2\right)+s_{xs}\left(\sigma_{xt}^2-\sigma_x^2\right) \\ -\lambda_\eta\sigma_{gt}\eta_{t+1}+\left(H_1^{(n-1)}-\lambda_e\right)\sigma_{xt}e_{t+1}+\left(H_2^{(n-1)}-\lambda_{gw}\right)\sigma_{gw}w_{g,t+1} \\ +\left(H_3^{(n-1)}-\lambda_{xw}\right)\sigma_{xw}w_{x,t+1}+\mu_d+\phi_xx_t+\varphi_d\sigma_{gt}\eta_{d,t+1} \end{array}\right)\right) \\
&= D_t\exp\left(\begin{array}{c} H_0^{(n-1)}+H_1^{(n-1)}\rho x_t+H_2^{(n-1)}\left(\sigma_g^2+\nu_g\left(\sigma_{gt}^2-\sigma_g^2\right)\right)+H_3^{(n-1)}\left(\sigma_x^2+\nu_x\left(\sigma_{xt}^2-\sigma_x^2\right)\right) \\ +\mu_s+s_xx_t+s_{gs}\left(\sigma_{gt}^2-\sigma_g^2\right)+s_{xs}\left(\sigma_{xt}^2-\sigma_x^2\right)+\mu_d+\phi_xx_t-\lambda_\eta\varphi_d\tau_{gd}\sigma_{gt}^2 \\ +\frac{1}{2}\left[\lambda_\eta^2\sigma_{gt}^2+\left(H_1^{(n-1)}-\lambda_e\right)^2\sigma_{xt}^2+\left(H_2^{(n-1)}-\lambda_{gw}\right)^2\sigma_{gw}^2+\left(H_3^{(n-1)}-\lambda_{xw}\right)^2\sigma_{xw}^2\right] \end{array}\right) \\
&= D_t\exp\left(H_0^{(n)}+H_1^{(n)}x_t+H_2^{(n)}\sigma_{gt}^2+H_3^{(n)}\sigma_{xt}^2\right).
\end{aligned}
$$

This implies:

$$
\begin{aligned}
H_0^{(n)} &= H_0^{(n-1)}+H_2^{(n-1)}\left(1-\nu_g\right)\sigma_g^2+H_3^{(n-1)}\left(1-\nu_x\right)\sigma_x^2 \\
&\quad +\mu_s-s_{gs}\sigma_g^2-s_{xs}\sigma_x^2+\mu_d \\
&\quad +\frac{1}{2}\left[\left(H_2^{(n-1)}-\lambda_{gw}\right)^2\sigma_{gw}^2+\left(H_3^{(n-1)}-\lambda_{xw}\right)^2\sigma_{xw}^2\right], \\
H_1^{(n)} &= H_1^{(n-1)}\rho+s_x+\phi_x, \\
H_2^{(n)} &= H_2^{(n-1)}\nu_g+s_{gs}-\lambda_\eta\varphi_d\tau_{gd}+\frac{1}{2}\lambda_\eta^2, \\
H_3^{(n)} &= H_3^{(n-1)}\nu_x+s_{xs}+\frac{1}{2}\left(H_1^{(n-1)}-\lambda_e\right)^2.
\end{aligned}
$$

D 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:

$$E_0\left[\sum_{t=0}^{\infty} e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma}\right] \tag{27}$$

At each period consumption growth is given by:

$$\frac{C_{t+1}}{C_t} = e^g \times \begin{cases} 1 & \text{if there is no disaster at time t+1} \\ B_{t+1} & \text{if there is a disaster at time t+1} \end{cases} \tag{28}$$

The pricing kernel is then given by:

$$\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 & \text{if there is no disaster at time t+1} \\ B_{t+1}^{-\gamma} & \text{if there is a distater at time t+1} \end{cases} \tag{29}$$

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

$$\frac{D_{i,t+1}}{D_{it}} = e^{g_{iD}}\left(1+\varepsilon_{i,t+1}^D\right) \times \begin{cases} 1 & \text{if there is no disaster at time t+1} \\ F_{i,t+1} & \text{if there is a distater at time t+1} \end{cases} \tag{30}$$

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:

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

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:

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

The expected return is the same across maturities, because strips of all maturities are exposed to the same risk in a disaster.[19]

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.

[19] We thank Xavier Gabaix for providing us with this derivation.

Panel A: Full sample 1996:2 - 2009:5				
	$R_{1,t}$	$R_{2,t}$	Market	S&P500
Mean	0.0120	0.0115	0.0054	0.0049
Median	0.0097	0.0148	0.0130	0.0104
Std. Dev.	0.0791	0.0979	0.0492	0.0472
Skewness	0.2199	0.3532	-0.8117	-0.6791
Kurtosis	10.3412	10.8974	4.1703	3.9013
Sharpe ratio	0.1155	0.0876	0.0514	0.0426
Observations	160	160	160	160
Panel B: First half sample 1996:2 - 2002:12				
Mean	0.0159	0.0139	0.0060	0.0065
Median	0.0117	0.0231	0.0136	0.0093
Std. Dev.	0.0986	0.1212	0.0528	0.0514
Skewness	0.1070	0.2931	-0.6014	-0.4598
Kurtosis	8.1457	8.7394	2.8930	2.7453
Sharpe ratio	0.1242	0.0843	0.0564	0.0456
Observations	83	83	83	83
Panel C: Second half sample 2003:1 - 2009:5				
Mean	0.0079	0.0089	0.0047	0.0031
Median	0.0077	0.0067	0.0129	0.0112
Std. Dev.	0.0508	0.0646	0.0453	0.0424
Skewness	0.2183	0.2014	-1.1925	-1.1584
Kurtosis	5.7627	6.3560	6.6587	6.3129
Sharpe ratio	0.1139	0.1050	0.0587	0.0248
Observations	77	77	77	77

Table 1: Descriptive 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 2003:1-2009:5.

Dep. Var.	$R_{1,t+1} - R_{f,t}$		$R_{2,t+1} - R_{f,t}$	
c	0.0079	0.0077	0.0073	0.0069
	(0.0060)	(0.0044)	(0.0075)	(0.0053)
mktrf	0.4879	0.5199	0.4912	0.5250
	(0.1224)	(0.1174)	(0.1539)	(0.1454)
AR(1)	-	-0.2935	-	-0.3396
	-	(0.0731)	-	(0.0738)
R^2	0.0914	0.1764	0.0606	0.1809

Table 2: 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. Standard errors in parentheses.

Dep. Var.	$R_{1,t+1} - R_{f,t}$		$R_{2,t+1} - R_{f,t}$	
c	0.0069	0.0070	0.0057	0.0056
	(0.0061)	(0.0045)	(0.0075)	(0.0053)
mktrf	0.5176	0.5426	0.5987	0.6248
	(0.1303)	(0.1248)	(0.1623)	(0.1533)
hml	0.1954	0.1765	0.4305	0.3988
	(0.1833)	(0.1757)	(0.2283)	(0.2159)
smb	0.0959	0.1074	-0.0158	0.043
	(0.1679)	(0.1607)	(0.2092)	(0.1979)
AR(1)	-	-0.2917	-	-0.3306
	-	(0.0734)	-	(0.0737)
	0.0982	0.1611	0.0854	0.1919

Table 3: 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. Standard errors in parentheses.

Dep. Var.	$R_{1,t+1}$	$R_{2,t+1}$	$R_{SP500,t+1}$
Mean equation			
c	0.0073	0.0083	0.0069
	(0.0046)	(0.0041)	(0.0038)
AR(1)	-0.2905	-0.3954	0.0790
	(0.1059)	(0.1094)	(0.0996)
Variance equation			
c	1.3 x 10^{-4}	2.4 x 10^{-4}	6.8 x 10^{-5}
	(7.2 x 10^{-5})	(1.5 x 10^{-4})	(5.2 x 10^{-5})
squared residual	0.1324	0.1985	0.1937
	(0.0421)	(0.0621)	(0.0761)
GARCH(1)	0.8728	0.8174	0.8071
	(0.0275)	(0.0518)	(0.0610)

Table 4: 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.

Sample period	1946-2007	1970-2007	1989-2007
δ_0	0.0899	0.0782	0.0885
δ_1	0.9267	0.9349	0.8734
γ_0	0.0616	0.0564	0.0716
γ_1	0.4856	0.6966	0.7901
σ_μ	0.0179	0.0166	0.0318
σ_g	0.0465	0.0328	0.0324
σ_D	0.0035	0.0038	0.0025
$\sigma_{g\mu}$	0.4939	0.4981	0.8153
$\sigma_{\mu D}$	0.8581	0.8631	-0.5745
R^2_{Ret}	0.0977	0.0904	0.2451
R^2_{Div}	0.2423	0.4086	0.5670

Table 5: Maximum-likelihood estimates
We present the estimation results of the present-value model. The model is estimated by unconditional maximum likelihood using data over three different sample periods, 1946-2007, 1970-2007, 1989-2007 on cash-invested dividend growth rates and the corresponding price-dividend ratio.

Dep. Var.	$R_{1,t+1}$		
C	0.1904	0.1539	0.1379
	(0.0452)	(0.0496)	(0.0373)
$\mathcal{P}_{t,T}/D_t$	-0.1820	-	-0.1285
	(0.0456)		(0.0377)
$\mathcal{P}_{t,T}/D_t$ smoothed	-	-0.1445	-
		(0.0500)	
AR(1)	-	-	-0.2616
			(0.0840)
R^2	0.0983	0.0547	0.1516

Table 6: Return predictability
The table presents regressions of the monthly return series on trading strategy 1, $R_{1,t+1}$, on the ratio of the one-year dividend strip at time t, denoted by $\mathcal{P}_{t,T}$, and the aggregated dividend paid out over the previous twelve months, where dividends are reinvested in the risk free rate. We also regress returns on a smoothed version of $\mathcal{P}_{t,T}/D_t$, where the smoothed ratio is computed by taking a rolling average over the past three values of $\mathcal{P}_{t,T}/D_t$.

	Benchmark	A1	A2	A3	A4	A5	A6
Mean	0.0120	0.0124	0.0149	0.0122	0.0120	0.0121	0.0141
Median	0.0097	0.0139	0.0056	0.0102	0.0091	0.0091	0.0045
St. Dev.	0.0791	0.1041	0.1332	0.0799	0.0792	0.0790	0.1291
Skewness	0.2199	0.6967	1.0808	0.2856	0.9186	0.1663	0.8522
Kurtosis	10.3412	12.3240	9.6743	8.0964	13.9763	9.8261	8.9516
Sharpe ratio	0.1155	0.0913	0.0901	0.1164	0.1149	0.1165	0.0868
CAPM Alpha	0.0079	0.0083	0.0100	0.0081	0.008	0.0079	0.0094
CAPM Beta	0.4879	0.4528	0.8020	0.4844	0.4236	0.4907	0.7228

Table 7: Alternative selection criteria
The table presents the summary statistics of dividend strategy 1 for six alternative selection criteria (A1 to A6), which are described in the main text. The table also reports the CAPM alpha and beta.

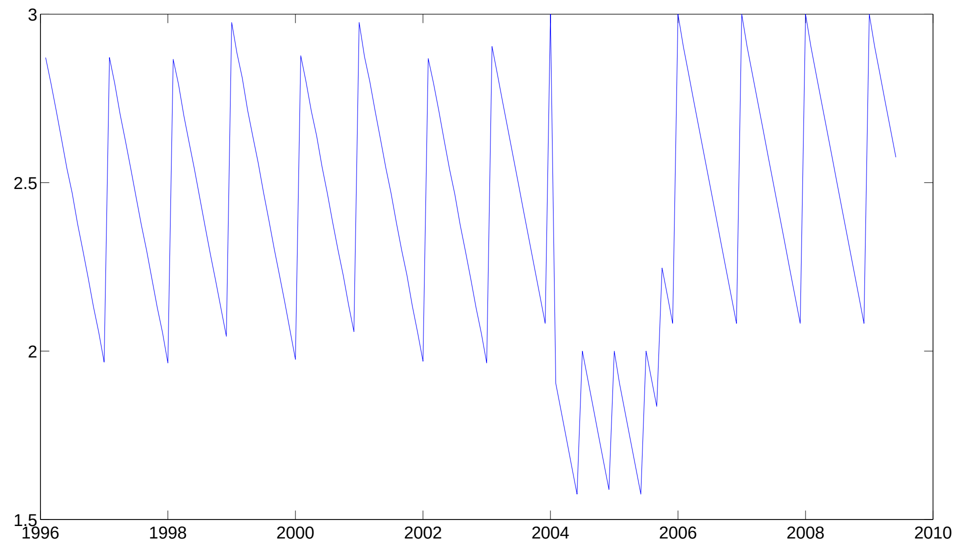

Figure 1: Maximum maturity of LEAPS
The graph displays the maximum maturity of LEAPS contracts that is available at each point of the sample. The sample period is January 1996 up to May 2009.

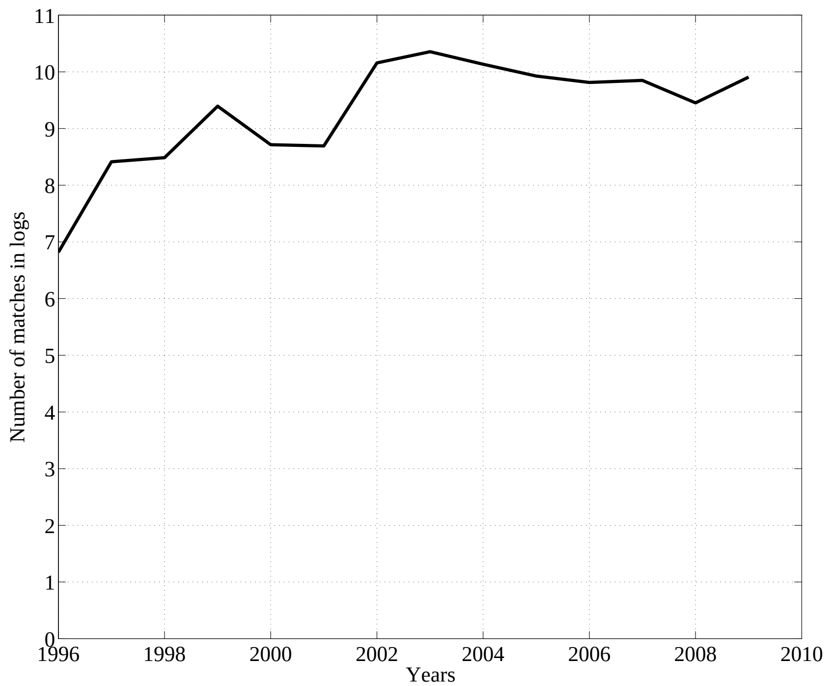

Figure 2: Average number of matches
The graph shows the average number of matches of put and call contracts with strike prices and maturities that coincide, and for which the quotes are provided in the same second during the last trading day of the month. We focus on contracts with a maturity between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

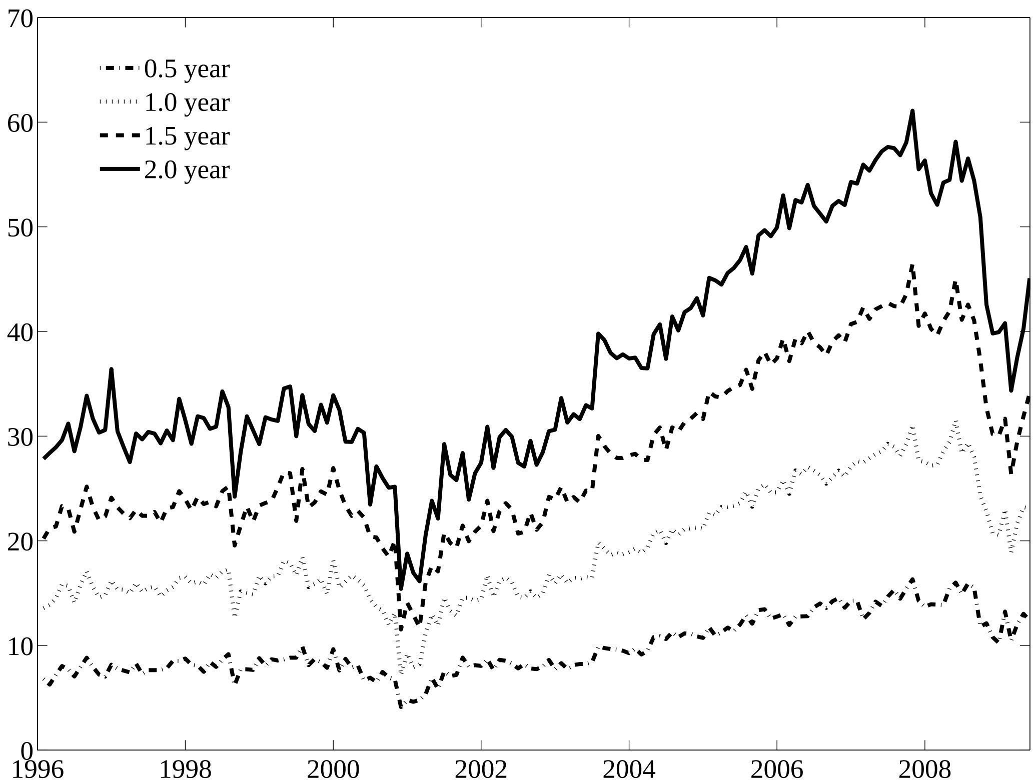

Figure 3: Price dynamics of the short-term assets (Cumulative)
The graph shows the prices of the first 0.5, 1, 1.5 and 2 years of dividends. The sample period is January 1996 up to May 2009.

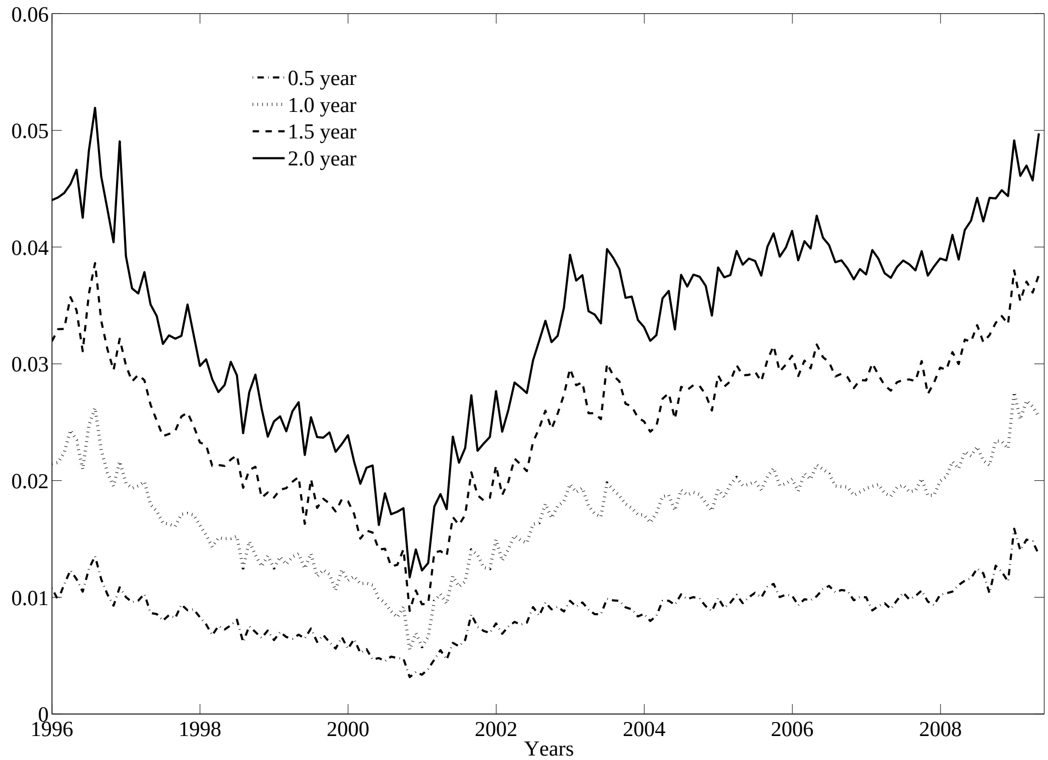

Figure 4: Present value of dividends as a fraction of the index value (Cumulative)
The graph shows the net present value of the first 0.5, 1, 1.5 and 2 years of dividends as a fraction of the index value as computed. The sample period is January 1996 up to May 2009.

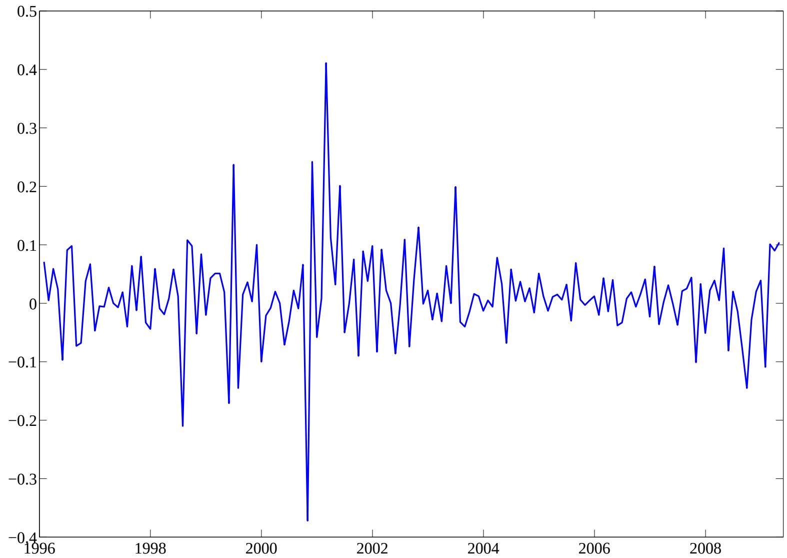

Figure 5: Monthly returns on trading strategy 1: 1996:2-2009:5: line graph.

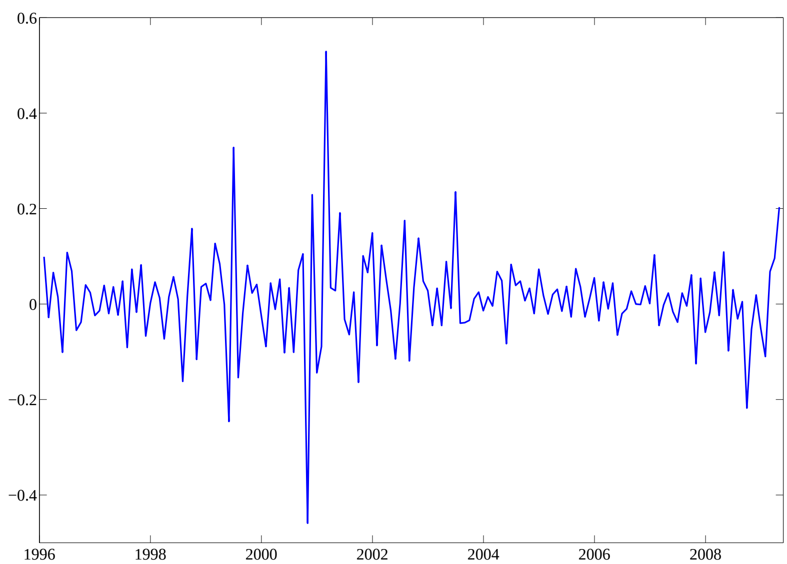

Figure 6: Monthly returns on trading strategy 2: 1996:2-2009:5: line graph.

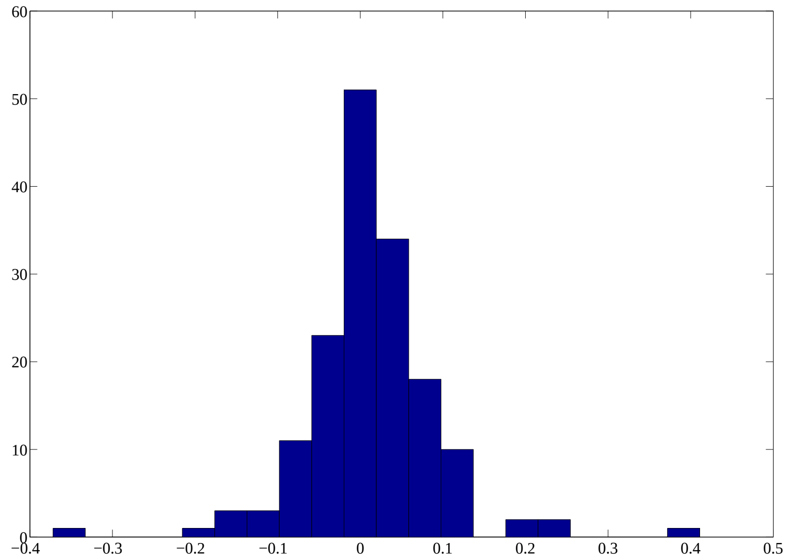

Figure 7: Monthly returns on trading strategy 1: 1996:2-2009:5: histogram.

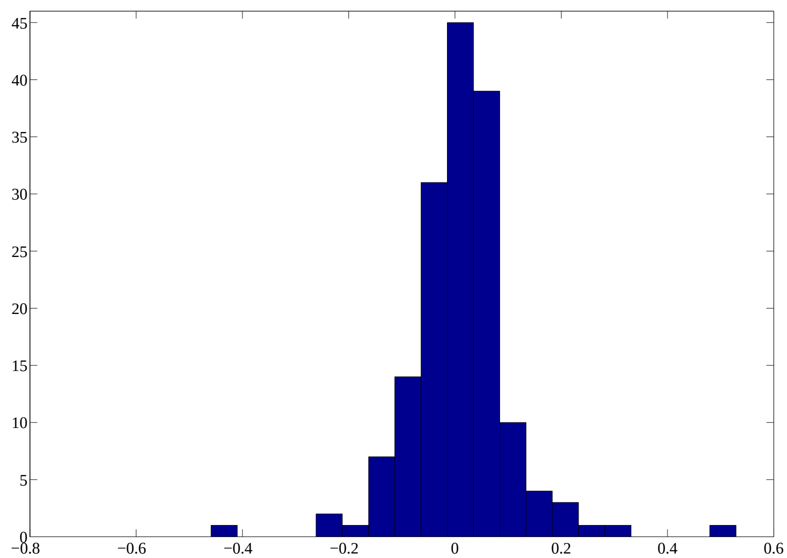

Figure 8: Monthly returns on trading strategy 2: 1996:2-2009:5: histogram.

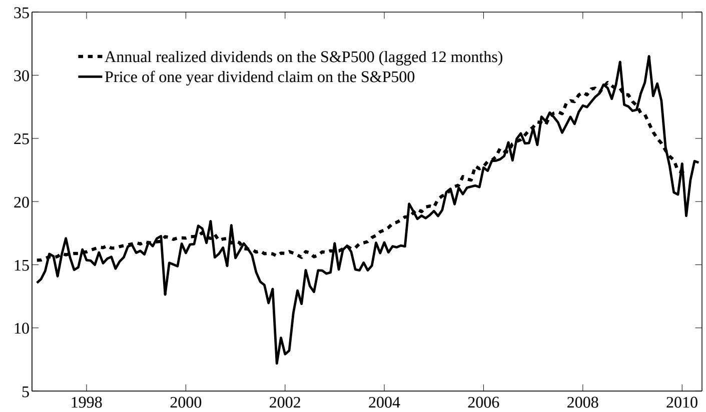

Figure 9: Prices and realizations of dividend claims: 1996:2-2009:5.

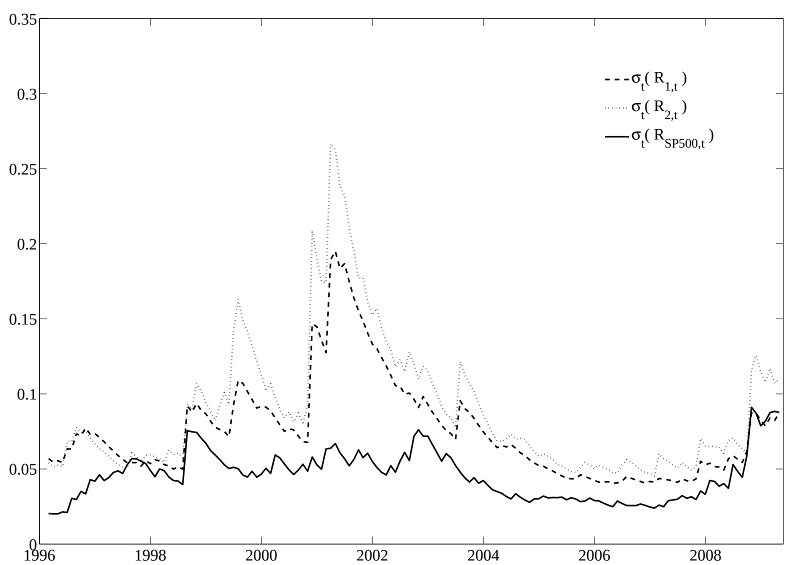

Figure 10: Volatility of dividend returns and returns on the S&P500 based on a GARCH(1,1) model.

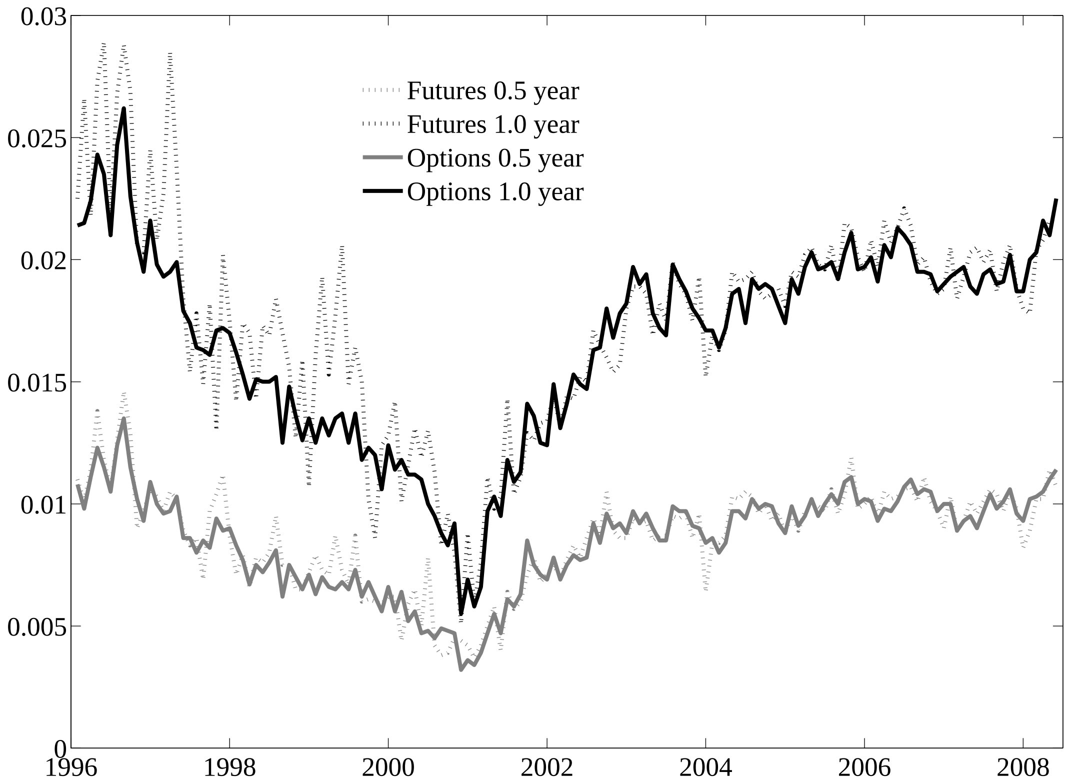

Figure 11: Short-term asset prices implied by futures and options
The graph shows the price of the short-term assets implied by futures and option markets. The maturity of the short-term asset equals either 0.5 year or 1 year.

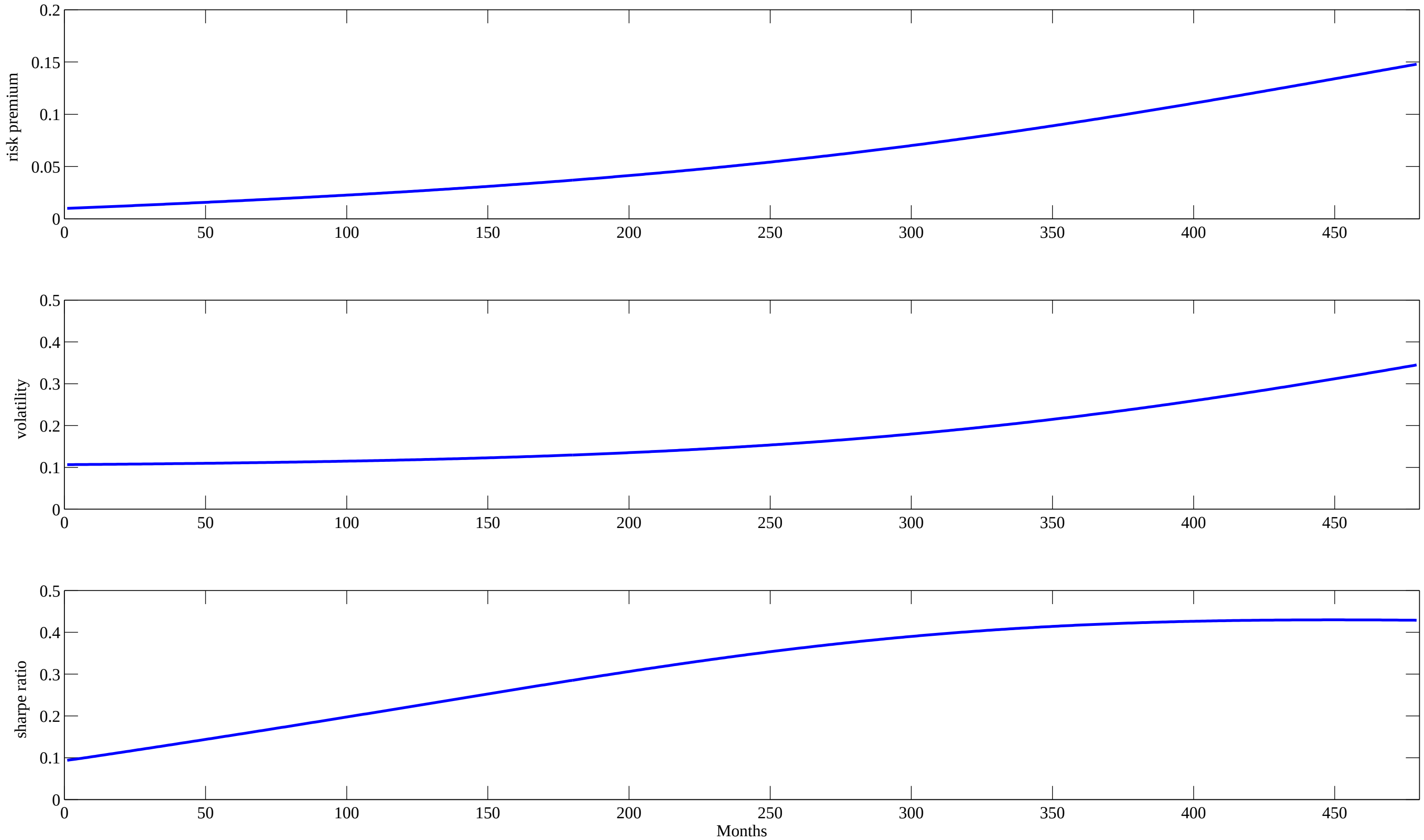

Figure 12: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for External Habits
The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the Campbell Cochrane (1999) habit formation model. The graph plots the first 480 months of dividend strips,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

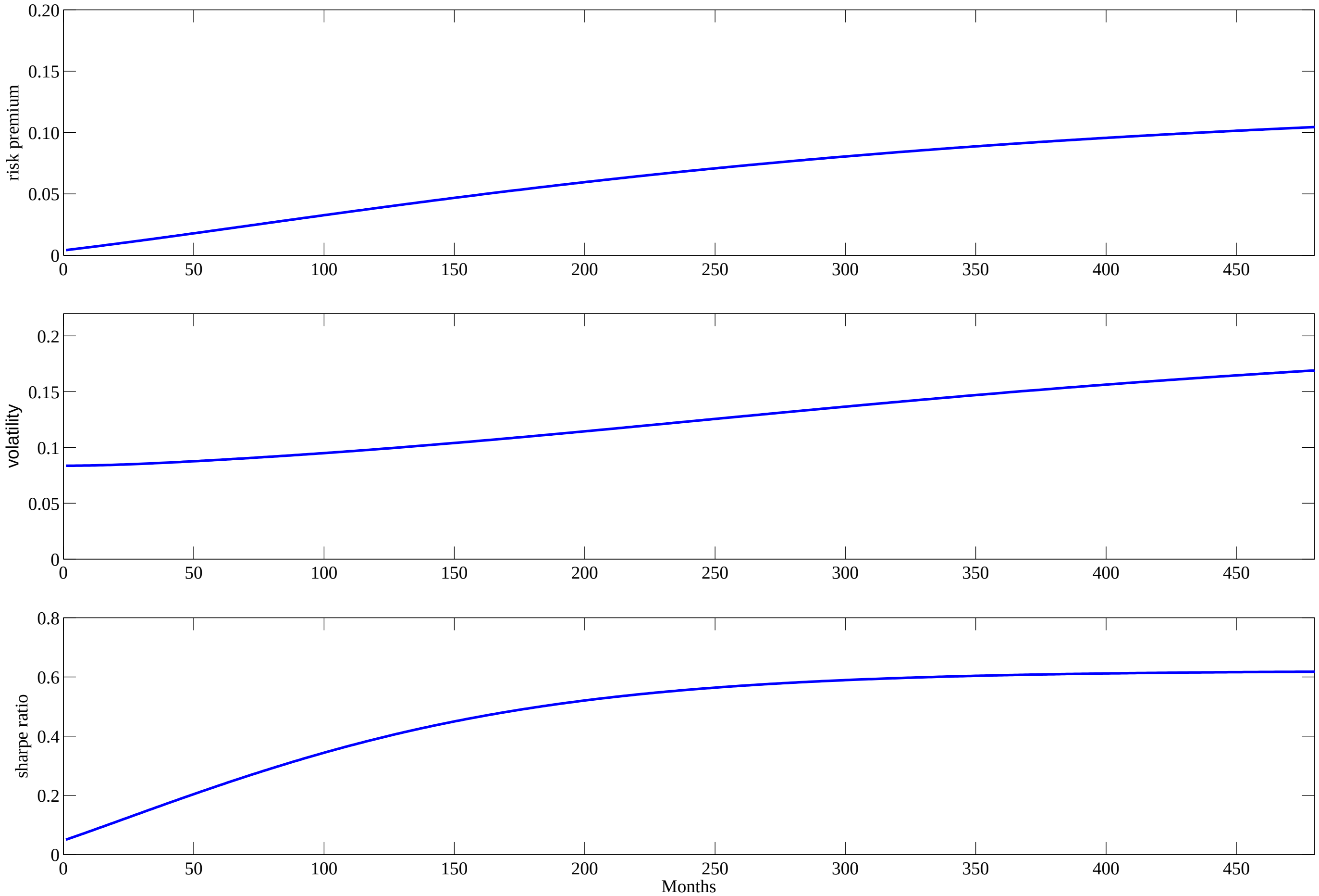

Figure 13: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for the Long Run Risk Model
The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the long run risk model as calibrated by Bansal and Shaliastovich (2009). The graph plots the first 480 months of dividend strips, which corresponds to 40 years.

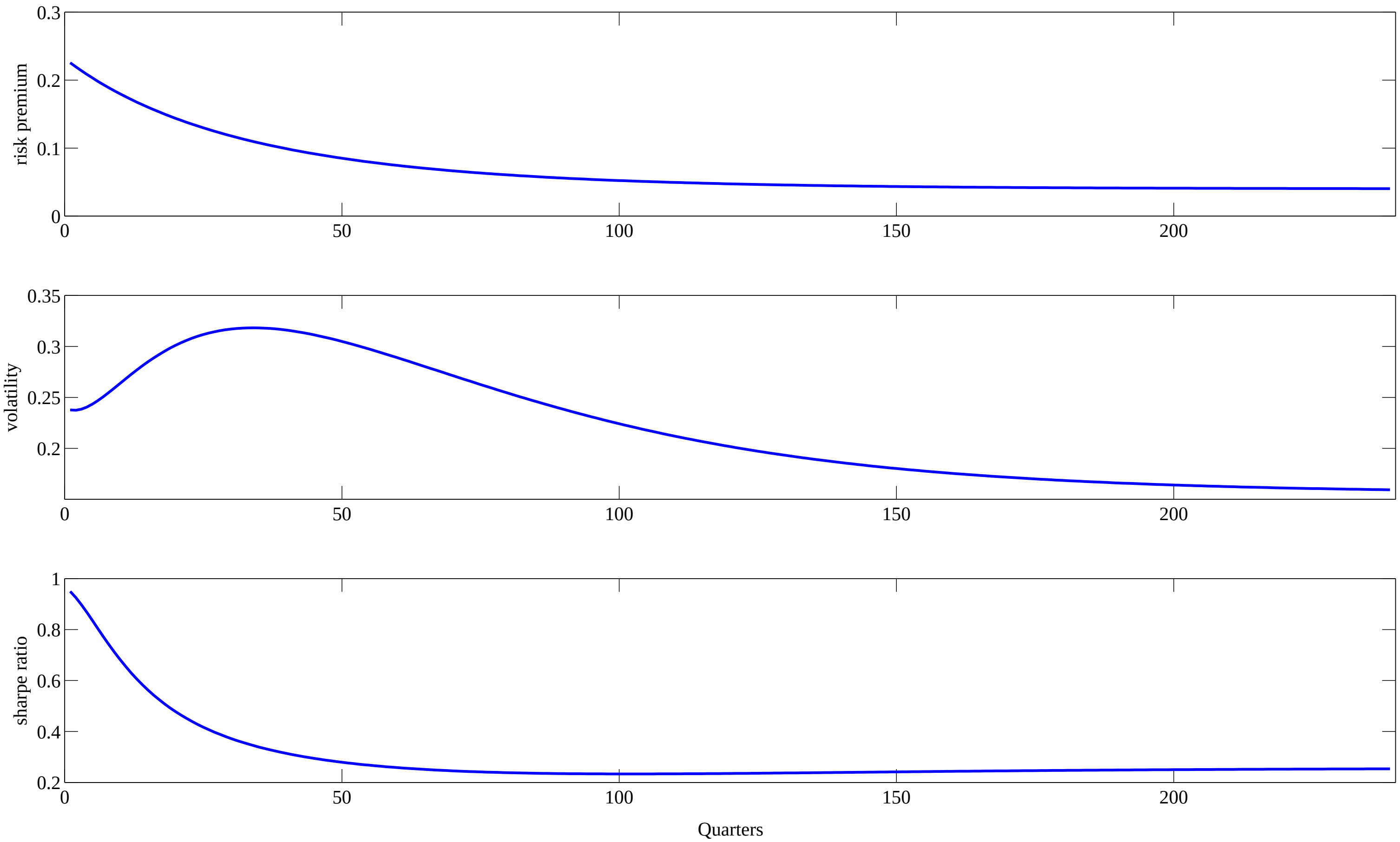

Figure 14: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for the Lettau Wachter (2007) Model
The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the Lettau Wachter (2007) model. The graph plots the first 120 quarters of dividend strips, which corresponds to 40 years.