

Robert Goldstein, Frederico Belo and Pierre Collin-Dufresne

**Endogenous Dividend Dynamics
and the Term Structure of
Dividend Strips**

Endogenous Dividend Dynamics and the Term Structure of Dividend Strips

Abstract

Many leading asset pricing models predict that the term structure of expected returns and volatilities on dividend strips are upward sloping. Yet the empirical evidence suggests otherwise. This discrepancy can be reconciled if EBIT dynamics are combined with a dynamic capital structure strategy that generates stationary leverage ratios. This combination endogenously determines dividend dynamics that are cointegrated with EBIT, implying that long-horizon dividend strips are no riskier than long-horizon EBIT strips. This capital structure policy also implies that shareholders have their position 'managed', creating stock volatility that is higher than long-horizon dividend volatility (i.e., excess volatility).

1 Introduction

Many leading asset pricing models (e.g., Campbell and Cochrane (CC, 1999), Bansal and Yaron (BY, 2004)) predict that the term structure of expected returns and volatilities on dividend strips are upward sloping. Yet the empirical evidence suggests otherwise (Binsbergen, Brandt and Kojien (BBK, 2011)). While Boguth, Carlson, Fisher and Simutin (2011) have questioned some of the findings of BBK, both papers seem to agree that the strongly upward sloping term structures predicted by BY and CC are inconsistent with the historical evidence on returns of dividend strips. In this paper, we show that both the CC and BY frameworks are consistent with downward sloping term structures of expected returns and volatilities for dividend strips if dividend dynamics are derived endogenously from capital structure policies that generate stationary leverage ratios. That is, one need not change the pricing kernel dynamics of these models, but rather only their specifications of dividend dynamics.

Compared to unleveraged cash flows such as EBIT or consumption, dividends are a leveraged cash flow. It is thus not surprising that claims to dividends (i.e., equity) are more volatile and have higher average historical returns than claims to EBIT (i.e., debt and equity).

Yet, over long horizons, EBIT and dividends should be cointegrated in that, path-by-path, dividends and EBIT should share the same long run growth rate. We state this claim more as an accounting identity than as a theory. Indeed, economists would be hard-pressed to write down a structural model of cash flows where this cointegration does not hold, especially if the model captures the empirical fact that leverage ratios are stationary. This is even more true if the dividend and EBIT processes we focus on are those of an aggregate index, where poorly performing firms are eliminated from the index long before they default. Note that cointegration implies predictability (see, e.g., Granger and Engle (1987)). This is important, since many researchers have reported that dividends mostly follow a random walk (Cochrane (2007)). Below, we show that variance ratio tests strongly reject this random walk hypothesis: long-horizon dividend volatility is significantly lower than short-horizon volatility.

Interestingly, leading asset pricing models either ignore the leveraged nature of dividends, or its cointegration with unleveraged cash flows, or both. Moreover, even if they do account for leverage, they do so in a reduced-form way by introducing free parameters that are not directly tied down to observed leverage ratios. For example, Campbell and Cochrane (CC, 1999) specify consumption and dividends as iid with the same drift, and therefore disregard leverage. Bansal and Yaron (BY, 2004) capture leverage by assuming that dividends have greater exposure to shocks in expected growth rates than does consumption. However, their

model does not capture cointegration. Abel (1999, 2005) models cash flows to be of the form y^λ , where $\lambda = 0$ for fixed income securities, $\lambda = 1$ for EBIT, and $\lambda > 1$ for dividends. This framework also does not capture cointegration.¹

In this paper, we investigate a framework that captures the leveraged nature of dividends while maintaining cointegration between dividends and unleveraged cash flows. In particular, we investigate an endowment-like economy that combines an exogenous unleveraged cash flow process with a dynamic capital structure strategy that leads to stationary leverage ratios. These two ingredients generate an *endogenously* obtained ‘leveraged’ dividend process that is internally consistent with the EBIT process. Claims to this dividend process (i.e., equity) have higher expected return and higher volatility than claims to EBIT (i.e., equity plus debt). Yet, this framework generates dividend and EBIT processes that are cointegrated.

Compared to frameworks that do not structurally account for the leveraged nature of dividends and its cointegration with unleveraged cash flows, our framework can explain certain “puzzling” properties of asset prices. First, our model generates stock return volatility that is higher than long-horizon dividend volatility (Shiller (1981), LeRoy and Porter (1981)), even if we specify a constant market price of risk.² This result is in contrast to the standard Gordon growth model prediction that long horizon dividend volatility equals stock return volatility, and in stark contrast to the long-run risk model of Bansal and Yaron (2004), which predicts that stock returns are less volatile than long-horizon dividends. Second, our framework generates a term structure of expected returns and volatilities for dividend strips that are decreasing in horizon, consistent with the empirical findings of BBK, and in contrast to the models of BY and CC.

The intuition for why it is important to jointly model EBIT and dividend dynamics in an internally consistent manner stems from the fact that when a firm rebalances its debt levels over time to maintain a stationary leverage process, shareholders are being forced to divest (invest) when leverage is low (high). Thus, even if investors follow a “static” strategy of holding a fixed supply of stock, their position is effectively being ‘managed’ by the capital structure decisions of the firm. Below, we show that these imposed investments/divestments conceal the ‘leveraged nature’ of dividends in that, even though instantaneously dividends are leveraged

¹Models such as Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2005) directly model cointegration between consumption and dividends, but their mechanism is through labor share, and not stationary leverage ratios, which generates our results below. Many recent papers investigate the implication that dividends and consumption are cointegrated. See, for example, Bansal, Dittmar, and Lundblad (2001 and 2005), Hanson, Heaton and Li (2008), Bansal Dittmar and Kiku (2009).

²Below, we will argue that it is more important to focus on long horizon dividend volatility than short horizon volatility, since managers can (and do) choose to smooth dividends in the short run. See Marsh and Merton (1986), and Shiller (1986).

(in the sense that returns on equity are more volatile and have higher expected values than returns on (equity + debt)), over the long run, EBIT and dividends are cointegrated, and therefore have the same long run growth rate and volatility (i.e., same level of risk).

This intuition allows us to explain the two asset pricing puzzles mentioned above: First, we demonstrate that when dividend dynamics are specified to exhibit stationary leverage ratios, it automatically generates “excess volatility” in that long-horizon dividend volatility is lower than stock return volatility, even if market prices of risk are constant. Intuitively, since dividends are cointegrated with EBIT, its long-horizon volatility is shown to be equal to the volatility of (unleveraged) EBIT. In contrast, stock return volatility is pushed up by a “leverage factor” $\left(\frac{1}{1-L}\right)$. So for an average leverage ratio of approximately 40%, the stock price volatility is about 67% higher than the long-run dividend volatility.³

Second, due to the implicit divestments (investments) that the firm imposes in good (bad) times on stockholders via capital structure decisions, long-maturity dividend strips are not as risky as typically imagined – rather, they are about as risky as long-maturity EBIT strips, since dividends and EBIT are cointegrated. However, claims to *all* future dividends (i.e., equity) are riskier than claims to EBIT (ie., equity plus debt). The implication is that dynamic capital structure decisions that generate stationary leverage ratios shift the risk in dividends from long-horizons to short horizons, and thus generate a downward shift in the slope of the term structure of dividend strip returns compared to the slope of EBIT strips. We demonstrate that this impact is very large for both the CC and BY models. Indeed, calibrating a simple parsimonious model assuming separation of investment and capital structure decisions, we obtain downward sloping term structures for dividend strip returns for both the BY and CC models in spite of the fact that their term structures for EBIT strip returns are upward sloping.

Our framework makes two other important predictions, which we confirm empirically. First, since dividends are correctly interpreted as a leveraged cash flow in the short run, but are cointegrated with EBIT in the long run, our model predicts that dividend variance ratios should be a decreasing function of horizon. This prediction differs from both CC, where iid dividend dynamics generates constant variance ratios, and BY, where long run risk generates variance ratios that increase with horizon. Second, due to the forced investments/divestments imposed by a stationary leverage ratio policy, our model predicts that leverage positively forecasts dividend growth. That is, for example, if leverage is low today, management will issue debt to push leverage ratios back to their target value, in turn increasing the level of

³We note that this effect alone does not explain the entire excess volatility puzzle identified by Shiller (1981). Some amount of time variation in the market price of risk is still needed.

dividends paid today. Furthermore, this equity payout funded by the debt issuance reduces the size of the investment owned by equity-holders, and thus, reduces future dividend growth. The opposite interaction occurs if the current leverage is high. Together, these imply that leverage positively forecasts dividend growth.⁴

There is a large related literature on the time variation of corporate cash flows and discount rates. While firms can (and do) choose to smooth dividends in the short run (Marsh and Merton (1986), Chen (2009))⁵, it is more difficult to explain why long-horizon dividend volatility is lower than stock volatility (Shiller (1986)). As such, we argue that the literature should focus on this relationship, and not on the relation between stock returns and short horizon volatility.

Other related papers include Campbell and Shiller (1988), who find that variation in dividend yield is driven mostly by changes in discount rates. However, others have questioned the power of return predictability (Stambaugh (1999), Campbell and Yogo (2006)). Further, Larrain and Yogo (2008) find that discount rates do not need to be so volatile when focusing on the overall cash flows of the firm rather than just dividends. The issue of dividend growth predictability and smoothing has been investigated in Chen (2009) and Chen and Da (2011). Our paper adds to this literature by pointing out long-run variations in dividends are significantly impacted by the capital structure decisions of the firm. Aydemir et al (2006) investigate the effect of leverage in a habit formation model, but their focus is very different from ours.⁶

The rest of the paper is as follows. In Section 2 we provide empirical evidence that dividend variance ratios decrease with horizon. We also show that, consistent with our model, leverage ratios are stationary and that leverage predicts dividend growth. We investigate a model that captures long-run risk similar to BY in Section 3. We then demonstrate the robustness of our findings by applying it to a model of habit formation similar to CC in Section 4. In both cases, even though the term structures of EBIT strip returns are upward sloping, the term structures of dividend strip returns are downward sloping, consistent with the empirical evidence of BBK. We conclude in Section 5. Proofs are found in the Appendix.

⁴We acknowledge that this argument ignores the issue of investment/divestment of projects. For example, a firm that raises debt may use it to invest in a project rather than pay it out as dividends. This would imply that current dividends are not increased, and that future expected dividends are not decreased. In a robustness section we model such debt-funded investments, and show that it does not change our main results.

⁵Chen (2009) shows that management began smoothing dividends in the post-war era, which explains their lack of predictability since then.

⁶Aydemir et al (2006) investigate how much of the variation in stock volatility can be explained by time variation in leverage.

2 Empirical Support

In this section, we provide empirical support for the three most fundamental features of the model that drive our results. First, we show that dividend variance ratios are decreasing with horizon. That is, long horizon dividends are not as risky as iid models would predict, and much less risky than what ‘long-run risk’ models (which, as we show below, generate dividend variance ratios that increase with horizon) predict. Second, we provide support for the assumption that the aggregate leverage ratio is stationary. Finally, we provide evidence that leverage positively forecasts dividend growth.

2.1 Data

The two main variables required for our empirical work are the dividends on the aggregate stock market, and the aggregate leverage ratio. In this section we explain how these variables are constructed.

We consider three alternative measures of aggregate dividends to help establish the robustness of the findings. We perform the analysis using annual data to avoid the seasonality in dividend payments.⁷ The use of an annual dividend series implies that we need to take a stance on how dividends received within a particular year are reinvested. We consider two alternative reinvestment strategies. In the first strategy, we assume the monthly dividends are reinvested in the aggregate stock market. As in Binsbergen and Koijen (2009), we refer to this dividend series as market-invested dividends. This measure of dividends is by far the most common in the dividend-growth and return-forecasting literature, and thus we focus on this definition for the main part of our analysis.⁸ In the second strategy, we invest the monthly dividends in cash, and obtain a time series of annual dividends which we call cash-invested dividends. As shown by Binsbergen and Koijen (2010) and Chen (2009), the two dividend series have different time series properties in the post-war sample period.

We obtain the data for the two dividend series from Long Chen’s webpage (the data is used in Chen (2009)). We use this dataset because it covers a long sample period from 1873 to 2008, thus covering the pre Center for Research in Security Prices (CRSP) period. Focusing on this long sample allows us to address Merton’s (1987) concern about the lack of research in the pre-CRSP period, as well as to obtain more robust results. To construct the two dividend series, Chen (2009) combines the pre-CRSP data compiled by Schwert (1990) with the data from the

⁷For a similar approach, see also Cochrane (1994), Lettau and Ludvigson (2005), and Binsbergen and Koijen (2010).

⁸A non comprehensive list of studies that use this measure of dividends includes Lettau and Ludvigson (2005), Cochrane (2008), and Lettau and Van Nieuwerburgh (2008).

CRSP (NYSE/Amex/Nasdaq) value-weighted market portfolio at monthly frequency. We refer the reader to Chen (2009) for additional details on the construction of the two dividend series. We transform the nominal dividends into real dividends by deflating the annual dividends by the consumer price index (CPI), which is available from Robert Shiller’s webpage.

In addition to the previous two dividend series, we investigate a third alternative measure of dividends that includes share repurchases. The data for this alternative dividend series is available from Motohiro Yogo’s webpage (the data is used in Gomes, Kogan and Yogo (2009)), and covers a relatively shorter sample period from 1927 to 2007. Examining this alternative definition of dividends is motivated by a growing view that changing corporate finance policy has led many firms, in recent years, to compensate shareholders through repurchase programs rather than through dividends (Fama and French (2001), Grullon and Michaely (2002)). As discussed in Lettau and Ludvigson (2005), still, large firms with high earnings have continued to increase traditional dividend payouts over time (DeAngelo, DeAngelo and Skinner, 2002). The impact on aggregate dividends is therefore unclear. To show that our main findings are not altered by adjusting dividends to account for share repurchase activity, since 1971, we consider a dividend series augmented with equity repurchases using Compustat’s statement of cash flows. We transform the nominal dividends into real dividends by deflating the annual nominal dividends by the CPI.

Finally, to construct the time series of the aggregate leverage ratio, we use data from the Flow of Funds Accounts of the United States (Board of Governors of the Federal Reserve System, 2005). The aggregate leverage ratio is defined as the ratio of total value of liabilities to the sum of the total value of liabilities and the total market value of equity. Liabilities are the sum of accounts payable; bonds, notes, and mortgages payable; and other liabilities. The data is for the nonfarm, nonfinancial corporate sector and is available annually since 1946. Larrain and Yogo (2008) extend the data back to 1927. We use this dataset which is available on Motohiro Yogo’s webpage, but this data ends in 2004. As such, we update the data to 2008 by collecting the updated aggregate total liabilities data from the Flow of Funds Accounts, and by constructing the total market value of equity in the nonfarm, nonfinancial corporate sector by replicating the approach in Larrain and Yogo (2008). We refer the reader to Larrain and Yogo (2008) for further details on the data construction.

2.2 Dividend Variance Ratios

If dividends follow a random-walk, then the variance of dividend growth increases linearly with the observation interval. That is, for example, the variance of two year dividend growth

will equal twice the variance of one year dividend growth, implying that the ratio of the two variances per unit of time equals unity. Following the approach of Lo and MacKinlay (1988), we construct the dividend variance ratio statistic (VR) across horizons from one to twenty years for each of the three alternative dividend series. We then show that dividend variance ratios are decreasing with horizon.

To compute the VR statistic, we directly apply the test formulas from Lo and MacKinlay (1988) (see their Section 1). For completeness, and to help in the calibration of the theoretical model proposed below, we also report the dividend volatilities at each horizon. We define dividend volatility over a given horizon using two different approaches.⁹ First, the more standard approach is to specify dividend volatility over a horizon T as

$$\sigma_{D,1}^T = \sqrt{\left(\frac{1}{T}\right) \text{Var}_0 \left[\log \left(\frac{D(T)}{D(0)} \right) \right]}. \quad (1)$$

We also consider a second definition:

$$\sigma_{D,2}^T = \sqrt{\left(\frac{1}{T}\right) \log \left[\frac{\text{E}_0 [D^2(T)/D^2(0)]}{(\text{E}_0 [D(T)/D(0)])^2} \right]}. \quad (2)$$

Note that for the case of log-normal (i.e., iid random walk) dynamics

$$\frac{dD}{D} = g dt + \sigma dz, \quad (3)$$

implying that

$$D(T) = D(0)e^{(g-\sigma^2/2)T+\sigma z(T)}, \quad (4)$$

both definitions produce the result $\sigma_{D,1}^T = \sigma_{D,2}^T = \sigma$ for all horizons T . The reason we consider the second definition is that it is defined even if dividends are negative (that is, if equity issuances are larger than dividend payments.)

Table 1 reports the VR test results for the three alternative measures of dividends. It reports the per year variance of dividend growth across each horizon T , for the two alternative definitions of dividend variance ($\sigma_{D,1}$ and $\sigma_{D,2}$). In addition, it reports the VR test statistic at each horizon, the corresponding standard errors (s.e.(VR)), and its p-value. The p-value is for the test of the null hypothesis that dividends follow a random walk, in which case the VR test

⁹These formulas are not the ones used in the variance ratio test of Lo and MacKinlay (1988), who instead use unbiased estimators of the variance by appropriately adjusting for the degrees of freedom. As a result, the variances reported here do not exactly match the variances used in the reported VR statistics, but the difference between the two is minimal.

statistic is 1. In specifying the null hypothesis, we consider the most general case in which the shocks to dividends can be heteroskedastic, not necessarily iid.

Table 1 shows that dividends do not follow a random walk. The VR test statistic decreases strongly with horizon for the three alternative dividend measures, implying predictability in dividends. Both definitions of dividend variance show that the variance of dividend growth is much smaller at long horizons than at short horizons. Regardless of the measure of dividends used and of how dividend variance is computed, the difference between short (1-year)- and long (10 or 20 years)-run dividend volatility is always greater than 5.2%. The conclusion that dividend variance decreases with horizon thus seems to be robust to how the monthly dividends are reinvested during the year, and to the inclusion of share repurchases in the measurement of dividends.¹⁰

For the first measure of dividends (market-invested dividends), Table 1 shows that the VR test statistic rejects the hypothesis that dividends follow a random walk at the 10% significance level for the 4- and 15-year horizon, and at the 5% significance level for the 6-, 8- and 10-year horizon. Using the first definition of dividend variance, the volatility is 15.0% for the one year horizon, but only 7.5% for the 20-year horizon, a large difference of 7.6%. Using the second definition ($\sigma_{D,2}$), the difference between short- and long-run dividend volatility is even larger. The volatility is 14.7% for the one year horizon, and only 6% for the 20-year horizon, a difference of 8.9%. For the other two alternative measures of dividends, the statistical rejection of the random walk hypothesis is weaker, but it is clear that the volatility of dividends decreases significantly with the horizon as well. For the second measure of dividends (cash-invested dividends), the difference between short- and long-run dividend volatility is 6.2% using the first definition of dividend variance, and is 7.3% using the second definition of dividend variance. For the last measure of dividends (with equity repurchases), the corresponding differences using the two dividend variance definitions is 8.3% and 9.1%, respectively.

The previous results have implications for the evaluation of leading asset pricing models. To illustrate the implications in a clear manner, Figure 1 graphically demonstrates the

¹⁰Table 1 shows that the variance of dividend growth for the first two measures of dividends (market-invested and cash-invested) is very similar. This result seems in contrast with the descriptive statistics reported in Binsbergen and Koijen (2010) who shows that the volatility of the cash-invested dividend growth is almost half the volatility of the market-invested dividend growth. The difference is the sample period. In Binsbergen and Koijen (2010) the sample period is from 1946 to 2007, whereas we examine a larger sample from 1873 to 2008. When we restrict the analysis to the shorter sample from 1946 to 2007, we confirm the Binsbergen and Koijen (2010) results using Chen's (2009) measure of cash-invested dividends. The larger volatility of cash-invested dividends in the pre-1946 period makes the properties of the two dividend series more similar in the full sample. Chen (2009) reports a similar sub-sample analysis and confirms that the different properties of the two series varies across sub-samples.

Maturity (T)	1	2	4	6	8	10	15	20	Diff 1-10	Diff 1-20
Dividend Definition 1: Market-invested dividends										
$\sigma_{D,1}^T$	15.04	14.59	12.62	10.92	9.13	8.07	7.95	7.45	6.97	7.59
$\sigma_{D,2}^T$	14.74	14.06	11.82	9.96	8.21	7.14	6.70	5.99	7.60	8.87
VR	1.00	0.97	0.72	0.54	0.42	0.32	0.31	0.32	–	–
s.e.(VR)	–	0.09	0.17	0.23	0.28	0.31	0.37	0.48	–	–
p-value	–	0.74	0.10	0.04	0.04	0.03	0.06	0.15	–	–
Dividend Definition 2: Cash-invested dividends										
$\sigma_{D,1}^T$	13.27	13.88	12.44	10.63	9.21	8.00	7.91	7.10	5.27	6.17
$\sigma_{D,2}^T$	12.99	13.36	11.66	9.71	8.27	7.07	6.67	5.73	5.92	7.26
VR	1.00	1.10	0.87	0.64	0.51	0.38	0.40	0.35	–	–
s.e.(VR)	–	0.10	0.21	0.28	0.34	0.38	0.45	0.53	–	–
p-value	–	0.30	0.54	0.19	0.15	0.11	0.18	0.22	–	–
Dividend Definition 3: With equity repurchases										
$\sigma_{D,1}^T$	13.71	14.14	13.60	10.43	7.43	7.24	6.96	5.44	6.47	8.27
$\sigma_{D,2}^T$	13.40	13.62	12.73	9.58	6.70	6.36	5.80	4.28	7.04	9.12
VR	1.00	1.08	1.02	0.54	0.32	0.31	0.36	0.21	–	–
s.e.(VR)	–	0.13	0.23	0.30	0.35	0.39	0.49	0.54	–	–
p-value	–	0.53	0.92	0.12	0.05	0.08	0.19	0.14	–	–

Table 1: Dividend variance ratio test demonstrates that dividend volatility drops significantly with horizon in the data. We can reject the hypothesis that dividends follow a random walk. The data for dividend definitions 1 and 2 are annual from 1873 to 2008, and the data for dividend definition 3 are annual from 1927 to 2007.

results in Table 1. The figure focuses on the main definition of dividends (market-invested dividends). The large difference between short- and long-run dividend volatility implies that we can strongly reject the random walk assumption of CC in specifying dividend dynamics, which naturally implies a dividend variance that is constant across the different horizons and hence a VR test statistic that is always equal to one. Moreover, we can reject even more strongly the long run risk dividend dynamics posited in BY (using their calibration), which we also plot in Figure 1. Here, due to long run risk, dividend growth volatility *increases* with horizon, in sharp contrast with the data. These results extend the analysis of Beeler and Campbell (2011), who show that dividend variance ratios in the U.S. aggregate stock market increase with horizon in the long run risk model, but not in the real data, using a sample of annual

dividend data for the 1930 to 2008 period.

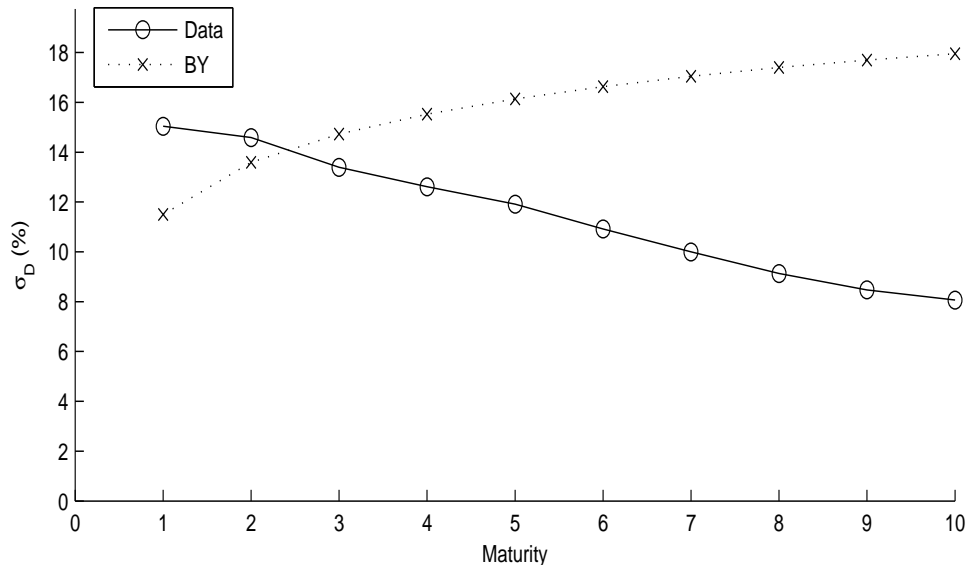


Figure 1: Expected dividend growth volatility as a function of horizon in the data and in Bansal and Yaron (2004). The data are annual from 1873 to 2008.

At a fundamental level, the finding that the dividend variance decreases with horizon must reflect negative serial correlation in the dividend growth series. To show this formally, here we consider a simple econometric approach that is based on a linear regression. Specifically, we investigate if past values of dividend growth help predict future dividend growth by running a regression of the form:

$$d_{t+1} - d_t = a + \sum_{k=1}^K b_k (d_{t+1-k} - d_{t-k}) + \varepsilon_t, \quad (5)$$

where d_t is log dividend at time t , and K is the number of lagged observations of dividend growth included in the regression. We consider $K=1$ and 2 (the main conclusion is robust to including other lags). By construction, this test is designed to capture the existence of serial correlation in dividend growth, which is ruled out by the random walk assumption.

The results reported in Table 2 show that past values of dividend growth help predict future dividends. In particular, in specification 2, the twice-lagged value of dividend growth helps forecasting dividends growth (the slope coefficient of b_2 is significantly different from zero with a p-value of 1%). When the one- and two-year lagged values of dividend growth

are included (specification 2), the chi-squared test rejects the hypothesis that all the slope coefficient are zero with a p-value of 4%. Finally, the slope coefficient on the lagged values of dividend growth are negative. Thus, an unusually high value of dividends growth today, predicts lower dividend growth. It is this negative autocorrelation that drives the decreasing pattern of dividend volatility across maturities.

Spec.		Parameter Estimates			Tests		
		a	b ₁	b ₂	R ²	χ ² (b = 0)	p-val(χ ²)
1	Slope	1.18	-0.04		-0.58	0.19	0.66
	p-val	0.38	0.67				
2	Slope	1.32	-0.05	-0.21	3.15	6.29	0.04
	p-val	0.32	0.61	0.01			

Table 2: Predictability regressions of real dividend growth on lagged values of dividend growth demonstrates that dividend growth is not a random walk. Data are annual from 1872 to 2008.

2.3 Leverage and Future Dividend Growth

As discussed previously, dynamic capital structure policies that generate stationary leverage ratios will cause leverage to positively forecast aggregate dividend growth. In this section we provide empirical support for both the stationarity of leverage ratios and the prediction that leverage forecasts future dividends.

Previous studies show empirical support for the claim that leverage ratios are stationary. As discussed in Collin-Dufresne and Goldstein (2001), at an aggregate (industry) level, leverage ratios have remained within a fairly narrow band even as equity indices have increased ten-fold over the past thirty years. At the firm level, Opler and Titman (1997) provide empirical support for the existence of target leverage ratios within an industry.¹¹ Further, dynamic models of optimal capital structure by Fischer, Heinkel, and Zechner (1989), and Goldstein, Ju and Leland (2001) find that firm value is maximized when a firm acts to keep its leverage ratio within a certain band.

Our empirical measure of aggregate leverage ratio is stationary as well. To demonstrate this, we run a regression of changes in log aggregate leverage ratio on lagged values of the log aggregate leverage ratio. We obtain the following results (Newey-West corrected t-statistics in

¹¹Additional studies providing empirical support for the claim that leverage ratios are stationary at the firm level include Flannery and Rangan (2006) and Fama and French (2002).

parenthesis):

$$\Delta \text{Lev}_{t+1} = \underset{(-2.92)}{-0.129} - \underset{(-2.99)}{0.137} \times \text{Lev}_t + e_{t+1}, \quad R^2 = 5.90\%, \quad \sigma(e_{t+1}) = 0.12.$$

The negative slope coefficient on the lagged value of leverage implies mean reversion in the aggregate leverage ratio.

The time-series properties of the aggregate leverage ratio are an important input for the calibration of the theoretical models that we present below. Here, we briefly report the relevant summary statistics of this process. The mean leverage ratio in our sample is 39.5% (in logs, the mean is -0.957 ± 0.05). The standard deviation of the aggregate leverage ratio is 9.69% (0.242 in logs) and the first order autocorrelation is 87.8% (in logs the autocorrelation is 86.6%)

In the theoretical sections below we will be modeling log-leverage dynamics in continuous time as an AR1 process similar to:

$$d\ell = \kappa(\bar{\ell} - \ell) dt + \sigma_\ell dz. \quad (6)$$

The data above allows us to calibrate the parameters: ($\bar{\ell} = -0.957 \pm 0.05$, $\kappa = 0.147 \pm 0.046$, $\sigma_\ell = 0.12 \pm 0.01$.)

To examine the relationship between leverage and future dividends, we run standard short- and long-horizon predictive regressions (e.g. Fama and French (1989), Lettau and Ludvigson (2002)). Let d_t be the log dividend. The dependent variable in the predictive regression is the T-year cumulated log dividend, in which T is the forecast horizon ranging from one year to 20 years. In examining predictability regression using horizons up to 20 years, we follow Cochrane (2008) who argues that very long horizons can be especially useful for detecting predictability. For completeness, we also include the results from a contemporaneous regression, denoted horizon T=0, to investigate the link between current dividend growth and current leverage.

Specifically, we run a long-horizon forecasting regression of the form:

$$d_{t+T} - d_t = a + b \times \text{Lev}_t + e_{t+T}, \quad (7)$$

in which Lev_t is the current value of the aggregate leverage ratio. For each horizon, we report the estimated slope associated with leverage, the corresponding p-value for the test of the hypothesis that the slope coefficient is zero, and the regression R^2 . In computing the p-value of the slope coefficient, we use standard errors corrected for autocorrelation per Newey and West (1987) with lag equal to three years plus the overlapping period ($\text{lag} = (3+T)$), and a GMM correction for heteroskedasticity. The correction for autocorrelation is especially important here due to the use of overlapping data. Finally, to evaluate the economic (not only

statistical) significance of the predictability results, we follow Cochrane (2008) and report the implied volatility of the *conditional* expected dividend growth (computed as $\sigma[E_t(\Delta D_T)] = \sigma(\hat{a} + \hat{b} \times \text{Lev}_t)$, using the estimated slope coefficients at the appropriate horizon), which we compare against the unconditional mean of dividend growth for each horizon (computed as $E(\Delta D_T)$). The ratio of the previous two measures tells us how large the variation in the conditional mean of dividend growth is relative to its unconditional mean. Naturally, high values of this ratio suggest an economically large variation in the conditional mean of dividend growth, and hence that the predictability is large when evaluated on economic grounds.

Table 3 reports the results of the long-horizon predictability regressions. The results shows that leverage can forecast future dividend growth. Importantly, consistent with the discussion in the introduction section, the slope coefficients are all positive, and thus periods with high leverage tend to be followed by periods with low dividend growth. The slope coefficient is statistically significant at the one- and 20- year horizons, but not at the other horizons. Finally, the regression R^2 increases from 3.19% at the one-year horizon to 14.2% at the 20 year-horizon.

The statistical significance of the predictability regressions reported here is modest, which is perhaps not surprising given the well established fact that dividends are difficult to predict, especially at the aggregate level (see, for example, Cochrane (2008), among others). However, the economic significance of the reported predictability is large. To see this, note that at the one-year horizon, the standard deviation of the fitted values ($\sigma[E_t(\Delta D_T)]$) is 2.7%, whereas the unconditional mean of annual dividend growth is only 0.83%. Thus, the conditional mean of dividend growth varies by about 3 times the value of its unconditional mean, an implied very large variation of the conditional mean judged on economic grounds. At longer horizons, the ratio is smaller, but still large in economic terms (across all horizons, the variation in the conditional mean of dividend growth is always larger than 38% of its unconditional mean).

3 Endogenous Dividend Dynamics in a ‘Long Run Risk’ Model

Here we investigate a model which captures the essential features of the “one-channel long-run risk model” of Bansal and Yaron (BY, 2004). We specify a state price density and aggregate cash-flow processes that correspond to a continuous time version of the exponential affine (approximate) solution presented in BY (2004).¹² BY demonstrate that their model can capture high expected returns, volatility and Sharpe ratios of stocks even with moderate levels of risk aversion. However, rather than exogenously specifying dividend dynamics as BY did, here we

¹²BY show that the affine approximation is very accurate relative to the numerical approximation of the exact model.

	Forecast horizon in years (T)					
	0	1	5	10	15	20
Lev _t	-0.08	0.11	0.19	0.22	0.35	0.43
p-val	0.08	0.03	0.31	0.22	0.07	0.05
R ²	0.61	2.55	2.68	5.00	12.64	14.17
σ [E _t (ΔD _T)]	1.87	2.66	4.71	5.25	8.63	10.75
$\frac{\sigma[E_t(\Delta D_T)]}{E(\Delta D_T)}$	2.24	3.19	0.69	0.39	0.42	0.38

Table 3: Long-horizon predictability regression of future dividend growth on current leverage. The table shows that high values of leverage forecast high future dividend growth. Data are annual from 1927 to 2008.

specify EBIT dynamics that are similar to BY's dividend dynamics, and then combine that with a dynamic capital structure policy in order to determine dividend dynamics endogenously. Specifying EBIT dynamics and leverage dynamics separately is consistent with the standard approach of assuming a separation of investment and capital structure policies.

3.1 EBIT Dynamics

We specify the dynamics for log-EBIT y_t to have a small but persistent shock to its expected growth x_t :

$$dy = \left(g + x - \frac{\sigma_y^2}{2} \right) dt + \sigma_y dz_1 \quad (8)$$

$$dx = -\kappa_x x dt + \sigma_{x_1} dz_1 + \sigma_{x_2} dz_2. \quad (9)$$

Since state vector dynamics are affine, it follows that date- t expectations of the first two moments of date- T EBIT take the exponential affine forms:

$$\mathbb{E}_t [e^{y_T}] = e^{y_t + A_0(T-t) + x_t A_1(T-t)} \quad (10)$$

$$\mathbb{E}_t [e^{2y_T}] = e^{2y_t + A_2(T-t) + x_t A_3(T-t)}, \quad (11)$$

where the deterministic coefficients $\{A_0(\tau), A_1(\tau), A_2(\tau), A_3(\tau)\}$ are given in the Appendix.

The term structure of EBIT expected growth rates over horizon τ is defined as

$$\begin{aligned} g_{y,\tau} &\equiv \left(\frac{1}{\tau} \right) \log (\mathbb{E}_0 [e^{y_\tau - y_0}]) \\ &= \left(\frac{1}{\tau} \right) [A_0(\tau) + x_0 A_1(\tau)]. \end{aligned} \quad (12)$$

Similarly, define the term structure of EBIT volatilities to be

$$\begin{aligned}\sigma_{y,\tau} &\equiv \sqrt{\left(\frac{1}{\tau}\right) \log \left[\frac{\mathbb{E}_0 [e^{2(y_\tau - y_0)}]}{(\mathbb{E}_0 [e^{y_\tau - y_0}])^2} \right]} \\ &= \sqrt{\left(\frac{1}{\tau}\right) [A_2(\tau) - 2A_0(\tau)]}.\end{aligned}\tag{13}$$

Note that $\sigma_{y,\tau}$ is independent of x , since $A_3(\tau) = 2A_1(\tau)$. We plot these term structures at their long run mean ($x_t = 0$) in Figure (2).

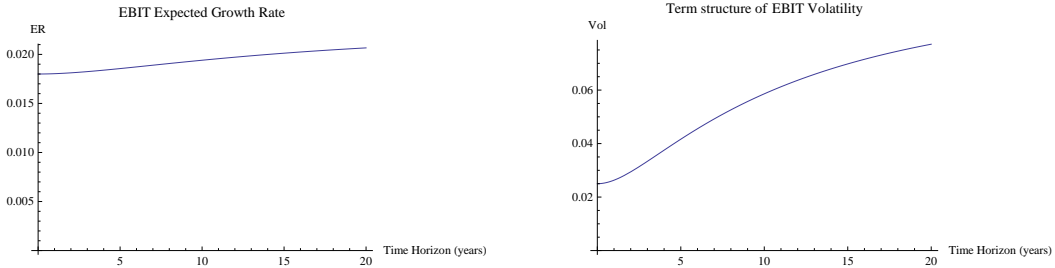


Figure 2: Term structure of EBIT expected growth rate (equation (12)) and volatilities (equation (13)) for the BY economy. Parameters are set as in Table 4.

Note that the term structure of volatilities is upward sloping. Intuitively, this is because over short horizons, the random variable x_T does not differ too much from its current value x_0 , and therefore log-EBIT approximately follows a random walk. Over longer horizons, however, the value of x_T becomes more uncertain (hence the name, “long run risk”), in turn generating an increasing term structure of volatilities.

3.2 EBIT Strips

BY demonstrate that their specified endowment dynamics combined with recursive preferences (Epstein Zin (1989)) generate pricing kernel dynamics that are well-approximated by constant market prices of risk:¹³

$$\frac{d\Lambda}{\Lambda} = -r dt - \theta_1 dz_1 - \theta_2 dz_2.\tag{14}$$

This implies that risk-neutral dynamics are:

$$dy = \left(g^Q + x - \frac{\sigma_y^2}{2} \right) dt + \sigma_y dz_1^Q\tag{15}$$

$$dx = \kappa_x (\bar{x}^Q - x) dt + \sigma_{x1} dz_1^Q + \sigma_{x2} dz_2^Q,\tag{16}$$

¹³For simplicity, we have set the risk free rate to a constant since it has no bearing on the issues at hand.

where we have defined $g^Q \equiv (g - \sigma_y \theta_1)$, $\bar{x}^Q \equiv -\left(\frac{\theta_1 \sigma_{x_1} + \theta_2 \sigma_{x_2}}{\kappa_x}\right)$.

The date- t price $P^T(t, x_t, y_t)$ of the security whose payoff is the date- T EBIT flow e^{y_T} is:

$$P^T(t, x_t, y_t) = e^{-r(T-t)} \mathbb{E}_t^Q [e^{y_T}]. \quad (17)$$

The solution takes the exponential affine form:

$$P^T(t, x_t, y_t) = e^{y_t + F(T-t) + G(T-t)x_t}, \quad (18)$$

where the deterministic functions ($F(\tau)$, $G(\tau)$) are derived in the Appendix.

Expected excess returns on the EBIT strips satisfy

$$\begin{aligned} \frac{1}{dt} \mathbb{E} \left[\frac{dP^T(t, x_t, y_t)}{P^T(t, x_t, y_t)} - r dt \right] &= -\frac{1}{dt} \mathbb{E} \left[\frac{d\Lambda}{\Lambda} \frac{dP^T(t, x_t, y_t)}{P^T(t, x_t, y_t)} \right] \\ &= \theta_1 [\sigma_y + G(T-t)\sigma_{x_1}] + \theta_2 [G(T-t)\sigma_{x_2}]. \end{aligned} \quad (19)$$

EBIT strip volatility is

$$\begin{aligned} \sigma^{P,\tau} &\equiv \sqrt{\frac{1}{dt} \left(\frac{dP^{t+\tau}(t, x_t, y_t)}{P^{t+\tau}(t, x_t, y_t)} \right)^2} \\ &= \sqrt{(\sigma_y + G(\tau)\sigma_{x_1})^2 + (G(\tau)\sigma_{x_2})^2}. \end{aligned} \quad (20)$$

We calibrate this model using the parameter values in Table 4, and plot the resulting term structures in Figure (3). We choose the parameter $\theta_1 = 0.0$ to be small and $\theta_2 = 0.4$ to be large in order to capture the notion in BY that consumption risk per se is low – it is expected consumption growth that agents with Epstein-Zin (1989) preferences are extremely sensitive to.

g	σ_y	κ_x	σ_{x_1}	σ_{x_2}	r	θ_1	θ_2
0.018	0.025	0.15	0.0	0.015	0.025	0.0	0.4

Table 4: Calibrated Parameters for the BY model.

As noted in Binsbergen, Brandt and Kojien (2011), this long-run risk model generates an upward sloping term structure of expected returns and volatilities. That is, the return variances on EBIT strips have inherited the upward sloping term structure associated with the variance ratios of the EBIT cash flows.

The enterprise value of the firm is equal to the present value of the claim to all EBIT strips:

$$P(x_t, y_t) = e^{y_t} \int_t^\infty dT e^{F(T-t) + G(T-t)x_t}. \quad (21)$$

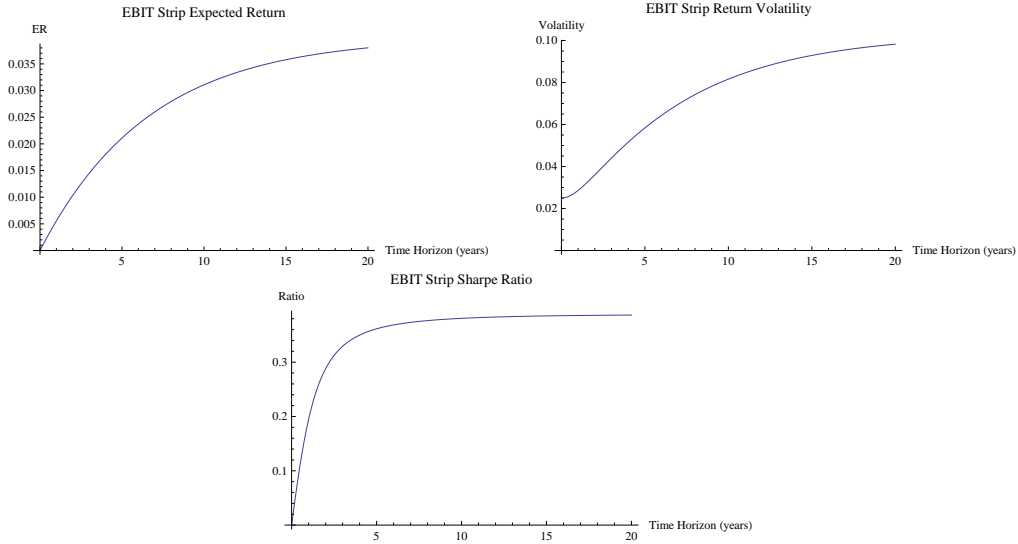


Figure 3: Term structure of EBIT Strip Expected Return (equation (19)) and volatilities (equation (20)) for the BY economy. Parameters are set as in Table 4.

As noted by Bansal and Yaron (2004), this can be well-approximated by a log-linear approximation:

$$P(x_t, y_t) \approx e^{y_t + F + G x_t}, \tag{22}$$

where the coefficients (F, G) are given in Table 5 below.¹⁴ Figure 4 plots the exact and approximate solution, and shows the accuracy of the log-linear approximation.

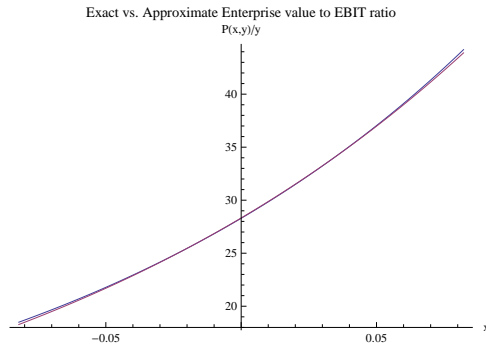


Figure 4: Exact and approximate solutions to the enterprise value as given in respectively equations (21) and (22) for the parameters given in Table 4. The x-axis covers -4 to $+4$ standard deviations of the unconditional distribution of x .

¹⁴We use the approach of Chen, Collin-Dufresne and Goldstein (2008) to derive the log-linear approximation.

In this “one-channel” model, the expected return and volatility of the claim to EBIT are constant under the log-linear approximation:

$$(\mu^P - r)_{BY} \approx \theta_1 (\sigma_y + G\sigma_{x1}) + \theta_2 (G\sigma_{x2}). \quad (23)$$

$$\sigma_{BY}^P \approx \sqrt{(\sigma_y + G\sigma_{x1})^2 + (G\sigma_{x2})^2}. \quad (24)$$

Their values, given the calibrated parameters, are given in the Table 5.

F	G	$(\mu^P - r)_{BY}$	σ_{BY}^P	Sh_{BY}^P
3.342	5.203	0.031	0.082	0.38

Table 5: Enterprise value expected return, volatility and sharpe ratio for the BY model with parameters set in equation 4.

3.3 Dividend Dynamics

Assume that at all dates- t , the firm issues riskless debt that matures at date- $(t + dt)$ with present value equal to

$$\begin{aligned} B(\ell_t, x_t, y_t) &= e^{\ell_t + y_t + F + Gx_t} \\ &\approx e^{\ell_t} P(x_t, y_t). \end{aligned} \quad (25)$$

We interpret $e^{\ell_t} \approx \frac{B(\ell_t, x_t, y_t)}{P(x_t, y_t)}$ as the leverage of the firm. Since it is riskless, the firm must pay $e^{r dt} B(\ell_t, x_t, y_t)$ at date $(t + dt)$. It does so by issuing at this time debt with face value $B(\ell_{t+dt}, x_{t+dt}, y_{t+dt})$, with all residual cash flows paid out as dividends. As such, dividends $dD(t + dt)$ paid out at date- $(t + dt)$ are

$$\begin{aligned} dD(t + dt) &= B(\ell_{t+dt}, x_{t+dt}, y_{t+dt}) - e^{r dt} B(\ell_t, x_t, y_t) + e^{y_{t+dt}} dt. \\ &= dB(\ell_t, x_t, y_t) - rB(\ell_t, x_t, y_t) dt + e^{y_t} dt. \end{aligned} \quad (26)$$

We choose the dynamics of log-leverage so that i) it is mean-reverting, and ii) dividend payments are locally deterministic, that is, $dD(t + dt) = \mathcal{D} dt$ and not $dD(t + dt) = \mathcal{D} dt + (\cdot) dz$. In particular, we choose¹⁵

$$\begin{aligned} d\ell_t &= \kappa_\ell \left(\bar{\ell}^P + \alpha x - \ell \right) dt - (G\sigma_{x1} + \sigma_y) dz_1 - G\sigma_{x2} dz_2 \\ &= \kappa_\ell \left(\bar{\ell}^Q + \alpha x - \ell \right) dt - (G\sigma_{x1} + \sigma_y) dz_1^Q - G\sigma_{x2} dz_2^Q, \end{aligned} \quad (27)$$

¹⁵Note that for D to be locally deterministic it is sufficient that B (and therefore $\log B$) be locally deterministic. Since $d \log B_t = d\ell_t + dy_t + Gdx_t$, it is clear that our choice below achieves this objective.

where

$$\bar{\ell}^Q = \bar{\ell}^P + \left(\frac{1}{\kappa_\ell}\right) [\theta_1 (G\sigma_{x_1} + \sigma_y) + \theta_2 G\sigma_{x_2}]. \quad (28)$$

Since the combination $(d\ell + dy + Gdx)$ is locally deterministic, dividends paid out over the interval $(t, t + dt)$ are equal to $\mathcal{D}(t) dt$, where

$$\begin{aligned} \mathcal{D}(t) &= e^{y_t} \left[1 + e^{\ell_t + F + Gx_t} \left(\kappa_\ell (\bar{\ell} + \alpha x - \ell) + g + x - \frac{\sigma_y^2}{2} - G\kappa_x x - r \right) \right] \\ &= e^{y_t} \left[1 + e^{\ell_t + F + Gx_t} \left(\kappa_\ell (\bar{\ell}^Q + \alpha x - \ell) + g^Q + x - \frac{\sigma_y^2}{2} + G\kappa_x (\bar{x}^Q - x) - r \right) \right]. \end{aligned} \quad (29)$$

Note that the terms inside the square bracket follow a stationary process. Hence, dividends are cointegrated with EBIT e^{y_t} .¹⁶

We calibrate the leverage ratio parameters as in Table 6. Note that our parameters ($e^{\bar{\ell}} = 0.35$, $\kappa_\ell = 0.11$) are well within a one standard deviation estimate of the empirical results ($e^{\bar{\ell}} = 0.39$, $\kappa_\ell = 0.14$), although admittedly our implied leverage volatility $\sigma_\ell = 0.06$ is quite a bit lower than the empirical observation $\sigma_\ell = 0.12$. We will improve upon this somewhat when we consider investment in the robustness section below.

$e^{\bar{\ell}}$	α	κ_ℓ
0.35	2.0	0.11

Table 6: Parameters for the log-leverage process in the BY Model. (Note that the long run mean of the log-leverage process ℓ_t is $\bar{\ell} + \alpha\bar{x} = 0.35$.)

Consistent with the analysis in the empirical section, we define the term structures of i) expected growth rates and ii) standard deviations of dividends over horizon T as:

$$g_{D,T} \equiv \left(\frac{1}{T}\right) \log \left(\mathbb{E}_0 \left[\frac{\mathcal{D}_T}{\mathcal{D}_0} \right] \right) \quad (30)$$

$$\sigma_{D,T} \equiv \sqrt{\left(\frac{1}{T}\right) \log \left[\frac{\mathbb{E}_0 [\mathcal{D}_T^2]}{(\mathbb{E}_0 [\mathcal{D}_T])^2} \right]}. \quad (31)$$

We plot these term structures for $x_t = 0$ in Figure (5).

¹⁶Our model does not restrict leverage to be less than unity, or dividends to be positive. In the appendix, we discuss an extension of the model, which avoids this issue. Unfortunately, we do not obtain closed-form solutions for that model. The numerical results show, however, that for the quantities of interest (dividend variance ratios, strip expected returns and volatilities), the results of our simpler model presented above are similar.

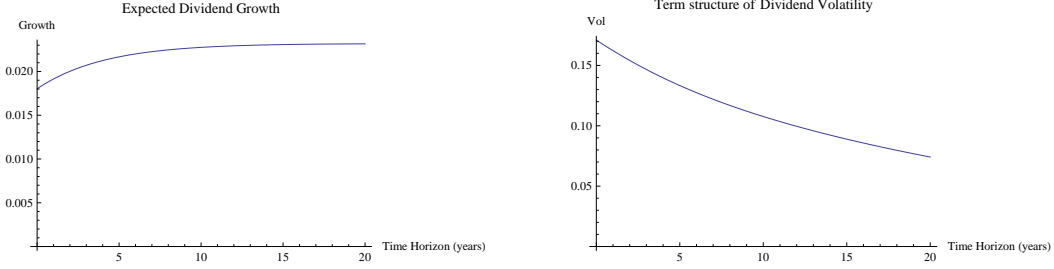


Figure 5: Term structure of expected dividend growth rate (equation (30)) and dividend volatility ($\sigma_{D,T}$ in equation (31)) for the BY economy. Parameters are set as in Table 4.

Note that, in contrast to the upward sloping variance ratios of EBIT, the dividend variance ratios are downward sloping, consistent with our reported empirical results. Admittedly, the levels are a bit high, suggesting that our leverage dynamics, and in turn, our endogenous dividend dynamics) which we chose for tractability and simplicity may be too simple – managers seem to choose a dividend policy that produces smoother dividends than what we are capturing. Of course, the advantage of this simplistic model is that the intuition for what is driving our results is very clear.

The results in this section present the term structure of instantaneous dividends, and estimated at the long-run mean of the state vector. This allows us to make use of the closed-form solutions for all the prices and moments of returns. Instead, in the empirical section we consider dividends aggregated over one full year. In the Appendix, we show that the effect of aggregating dividends over one year for the various statistics presented here in the paper is not economically significant.

3.4 Dividend Strips

Here we provide a closed-form expression for the price of dividend strips, defined as:

$$V^T(t) = \mathbf{E}_t^Q \left[e^{-r(T-t)} \mathcal{D}(T) \right]. \quad (32)$$

As noted previously, at date- T , the firm will issue risk-free debt of value $e^{\ell_T + y_T + F + Gx_T}$, and will retire debt of value $e^{r dt} e^{\ell_{T-dt} + y_{T-dt} + F + Gx_{T-dt}}$. The date- t value of these claims are:

$$\begin{aligned} W^T(t, \ell_t, x_t, y_t) &= \mathbf{E}^Q \left[e^{-r(T-t)} e^{\ell_T + y_T + F + Gx_T} \right] \\ U^T(t, \ell_t, x_t, y_t) &= \mathbf{E}^Q \left[e^{-r(T-t)} e^{r dt} e^{\ell_{T-dt} + y_{T-dt} + F + Gx_{T-dt}} \right] \\ &= W^{T-dt}(t, \ell_t, x_t, y_t). \end{aligned} \quad (33)$$

It therefore follows that the date- t present value of a claim to date- T dividends is:¹⁷

$$\begin{aligned} V^T(t, \ell_t, x_t, y_t) dt &= W^T(t, \ell_t, x_t, y_t) - W^{T-dt}(t, \ell_t, x_t, y_t) + dt \mathbb{E}^Q \left[e^{-r(T-t)} e^{y_T} \right] \\ &= \left[\frac{\partial}{\partial T} W^T(t, \ell_t, x_t, y_t) + P^T(t, x_t, y_t) \right] dt. \end{aligned} \quad (34)$$

From its definition, $e^{-rt}W^T(t, \ell_t, x_t, y_t)$ is a Q-martingale, implying that

$$\begin{aligned} 0 &= -rW + W_t + W_{\ell} \kappa_{\ell} \left(\bar{\ell}^Q + \alpha x - \ell \right) + W_x \kappa_x (\bar{x}^Q - x) + W_y \left(g^Q + x - \frac{\sigma_y^2}{2} \right) + \frac{\sigma_{\ell}^2}{2} W_{\ell\ell} \\ &\quad + \frac{\sigma_x^2}{2} W_{xx} + \frac{\sigma_y^2}{2} W_{yy} - W_{\ell x} [\sigma_{x1} (G\sigma_{x1} + \sigma_y) + G\sigma_{x2}^2] - W_{\ell y} \sigma_y (G\sigma_{x1} + \sigma_y) + W_{xy} \sigma_{x1} \sigma_y, \end{aligned} \quad (35)$$

where we have defined

$$\begin{aligned} \sigma_{\ell}^2 &\equiv (G\sigma_{x1} + \sigma_y)^2 + (G\sigma_{x2})^2 \\ \sigma_x^2 &\equiv \sigma_{x1}^2 + \sigma_{x2}^2. \end{aligned} \quad (36)$$

Since the dynamics of the state vector are affine, it is known (e.g., Duffie and Kan (1996)) that the solution takes an exponential-affine form:

$$W^T(t, \ell_t, x_t, y_t) = e^{y_t + H(T-t) + I(T-t)\ell_t + J(T-t)x_t}, \quad (37)$$

where the deterministic coefficients are determined in the Appendix.

Expected excess returns on the dividend strips satisfy

$$\begin{aligned} \frac{1}{dt} \mathbb{E} \left[\frac{dV^T(t, \ell_t, x_t, y_t)}{V^T(t, \ell_t, x_t, y_t)} - r dt \right] &= -\frac{1}{dt} \mathbb{E} \left[\left(\frac{d\Lambda}{\Lambda} \right) \left(\frac{dV^T(t, \ell_t, x_t, y_t)}{V^T(t, \ell_t, x_t, y_t)} \right) \right] \\ &= \theta_1 \Omega_1(\tau, \ell_t, x_t, y_t) + \theta_2 \Omega_2(\tau, \ell_t, x_t, y_t), \end{aligned} \quad (38)$$

where the terms $\left(\Omega_1(\tau, \ell_t, x_t, y_t), \Omega_2(\tau, \ell_t, x_t, y_t) \right)$ are given in the Appendix. Dividend strip volatility is

$$\begin{aligned} \sigma^{P, \tau} &\equiv \sqrt{\frac{1}{dt} \left(\frac{dV^T(t, x_t, y_t)}{V^T(t, x_t, y_t)} \right)^2} \\ &= \sqrt{\Omega_1^2(\tau, \ell_t, x_t, y_t) + \Omega_2^2(\tau, \ell_t, x_t, y_t)}. \end{aligned} \quad (39)$$

¹⁷There are at least two alternative approaches to derive the solution to the dividend strip in our context. First, one can directly estimate $V^T(t, x_t, y_t) = \mathbb{E}_t^Q \left[e^{-r(T-t)} \mathcal{D}(T) \right]$, where $\mathcal{D}(t)$ is defined in equation (29). Second, the solution can also be computed using equation (34). Given the log-linear approximation used above, the two closed-form solutions will not agree exactly. However, we have verified that the difference between the two is very small (at our parameter values, the difference is less than 10^{-14}). A third approach is to compute the present value of future dividends as the difference between the spot stock price $V(t) = P(x_t, y_t)(1 - e^{\ell_t})$ and the futures price $F^T(t) = \mathbb{E}_t^Q[V(T)]$ using the cash-and-carry formula: $V(t) = e^{-r(T-t)} F^T(t) + \int_t^T V^s(s) ds$. In turn, differentiating with respect to T gives yet another expression for Dividend strip: $V^T(t) = r e^{-r(T-t)} F^T(t) - e^{-r(T-t)} \frac{\partial F^T(t)}{\partial T}$. All of these approaches would give exactly the same values if the log-linear approximation to the enterprise value is not used. In practice, for our parameter choices, the differences between these three approaches are negligible.

We report the dividend strip return and volatility term structures in Figure (6) below.

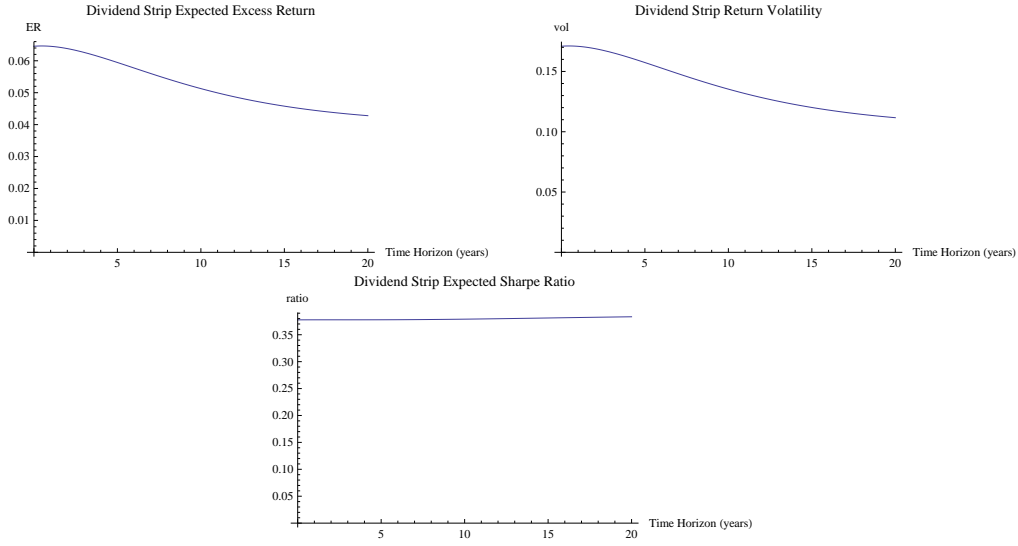


Figure 6: Term structures of dividend strip expected returns (equation (38)) and volatilities (equation (39)) for the BY economy. Parameters are given in Tables 4 and 6.

3.5 Equity Returns

The value of equity equals the claim to all dividends:

$$\begin{aligned}
 V(\ell_t, x_t, y_t) &= \int_t^\infty dT V^T(t, \ell_t, x_t, y_t) \\
 &= \int_t^\infty dT \left[\frac{\partial}{\partial T} W^T(t, \ell_t, x_t, y_t) + P^T(t, x_t, y_t) \right] \\
 &\approx P(x_t, y_t) - B(\ell_t, x_t, y_t) \\
 &\approx P(x_t, y_t) (1 - e^{\ell_t}).
 \end{aligned} \tag{40}$$

This equation is intuitive – it states that equity equals enterprise value minus debt outstanding. Indeed, this equation just follows from Modigliani-Miller’s capital structure irrelevance theorem ($V(t) = P(t) - B(t)$).

Excess return on equity is equal to excess return on EBIT (equation (23)) scaled by the leverage factor:

$$\mu^V - r = \left(\frac{1}{1 - e^{\ell_t}} \right) [\theta_1 (\sigma_y + G\sigma_{x1}) + \theta_2 (G\sigma_{x2})]. \tag{41}$$

This equation captures “dividend irrelevance” in that future capital structure decisions do not impact equity returns (or equity value) today. Similarly, equity return volatility is scaled up

by the same factor

$$\sigma^V = \left(\frac{1}{1 - e^{\ell t}} \right) \sqrt{(\sigma_y + G\sigma_{x1})^2 + (G\sigma_{x2})^2}. \quad (42)$$

For the calibration choice we focus on we find:

$(\mu^V - r)_{BY}$	σ_{BY}^V	Sh_{BY}^V
0.048	0.126	0.38

Table 7: Stock return expected return and volatility for the BY model with parameters set in Table 4.

3.6 Discussion

We have shown that if we start from an economy similar to that of BY, but endogenously derive dividend dynamics from an assumption about stationary (mean-reverting) leverage ratios, then, consistent with the empirical findings of BBK, short-maturity dividend strip returns have higher expected excess returns and higher volatilities than stock returns. Indeed, we find that the term structure of dividend strip return variances is downward sloping. This is in contrast to the term structure of EBIT (i.e., total firm cash-flows) strip return volatilities, which are upward sloping. The downward shift in the slope of the term structure of dividend strip returns compared to the slope of EBIT strips is due to the implicit divestments (investments) that the firm imposes in good (bad) times on stockholders via capital structure decisions, which generates stationary leverage ratios. As such, long-maturity dividend strips are not as risky as typically imagined – rather, they are about as risky as long-maturity EBIT strips, since dividends and EBIT are cointegrated. However, claims to *all* future dividends (i.e., equity) are riskier than claims to EBIT (i.e., equity plus debt). The implication is that dynamic capital structure decisions that generate stationary leverage ratios shift the risk in dividends from long-horizons to short horizons. We note that this also results in higher betas (as measured by the covariance of short term dividends with the market return), and therefore higher expected returns, for short-horizon dividend claims compared to long-horizon claims. The term structure of Sharpe ratios on dividend claims thus tends to be flat, whereas it was markedly upward sloping for the claims to EBIT in the BY Model (compare Figures 6 and 3). These results bring the pricing implications of the BY Model for the endogenous dividend claim much more

in line with the empirical results of BBK.¹⁸

Interestingly, the model with endogenous dividend dynamics also generates ‘excess volatility.’ Indeed, the claim to dividends has a volatility and excess return that is higher (by a factor of $(1/(1 - e^l) \approx 1.6)$) than the claim to EBIT, which reflects long-run EBIT risk. Since EBIT and dividends are long-term co-integrated, they will display the same long-run volatility. Thus the excess volatility generated by the model is of that same order of magnitude (stock price volatility is about 50% to 60% more volatile than a claim to long-horizon dividends estimated using a constant discount factor as in Shiller (1986)). This occurs despite the fact that we have a model with a constant market price of risk! As is well understood, (Campbell and Shiller (1987) and Cochrane (1991, 2007)) ‘excess volatility’ can be traced back to time variation in discount rates or predictability in dividends. In our framework, we have both, in that leverage predicts both future dividends (see equation (29)) and expected excess returns on equity (equation (41)). Expected excess returns on stocks are time-varying despite the fact that risk-premia are constant, simply due to the time variation in leverage.

4 Endogenous Dividend Dynamics in an ‘External Habit Formation’ Model

Here we derive the endogenous dividend dynamics in a model where aggregate cash-flows and market price dynamics are similar to those assumed in the external habit formation model of Campbell and Cochrane (CC, 1999). Unlike the BY model, which has a constant market price of risk and predictability in cash-flows, the framework we consider now has no predictability in aggregate cash-flows, but generates predictability in excess returns via time variation in risk-premia. Specifically, we assume that cash flows follow an iid process, and that shocks to the market price of risk are negatively correlated with shocks to the cash flows. Because much of the theory is very similar to the BY framework, we present here only the main results, and place a full description into an on-line appendix.

¹⁸BBK actually find decreasing sharpe ratios for the dividend claims, as, in their sample, short term dividend claims have betas less than 1 and exhibit high predictability and high alphas. However, the statistical significance of these point estimates is debatable, as has been discussed by Boguth et al (2011).

4.1 EBIT Dynamics

Instead of exogenously specifying dividend dynamics as in CC, here we specify the dynamics for log-EBIT y_t to be iid

$$dy = \left(g - \frac{\sigma_y^2}{2} \right) dt + \sigma_y dz. \quad (43)$$

The term structure of EBIT expected growth rate over horizon τ is defined as

$$\begin{aligned} g_{y,\tau} &\equiv \left(\frac{1}{\tau} \right) \log \left(\mathbf{E}_0 [e^{y_\tau - y_0}] \right) \\ &= g \quad \forall \tau. \end{aligned} \quad (44)$$

Similarly, the term structure of EBIT volatility is defined as:

$$\begin{aligned} \sigma_{y,\tau}^2 &\equiv \left(\frac{1}{\tau} \right) \log \left[\frac{\mathbf{E}_0 [e^{2(y_\tau - y_0)}]}{(\mathbf{E}_0 [e^{y_\tau - y_0}])^2} \right] \\ &= \sigma_y^2. \end{aligned} \quad (45)$$

Note that the term structure of volatilities is flat.

4.2 EBIT Strips

CC provide a framework that generates a pricing kernel with a constant risk free rate and a countercyclical market price of risk. We approximate their model with the following dynamics:

$$\frac{d\Lambda}{\Lambda} = -r dt - \theta_t dz, \quad (46)$$

where innovations in the market price of risk are driven by the same Brownian motion that drives EBIT innovations:

$$d\theta = \kappa (\bar{\theta} - \theta_t) dt - \nu dz. \quad (47)$$

Thus, risk-neutral dynamics for the state variables follow

$$dy = \left(g - \frac{\sigma_y^2}{2} - \sigma_y \theta_t \right) dt + \sigma_y dz^Q \quad (48)$$

$$d\theta = \kappa^Q (\bar{\theta}^Q - \theta_t) dt - \nu dz^Q, \quad (49)$$

where we have defined $\kappa^Q \equiv (\kappa - \nu)$, and $\kappa^Q \bar{\theta}^Q \equiv \kappa \bar{\theta}$.

The date- t price $P^T(t, \theta_t, y_t)$ of the security whose payoff is the date- T EBIT flow e^{y_T} is:

$$P^T(t, \theta_t, y_t) = e^{-r(T-t)} \mathbf{E}_t^Q [e^{y_T}]. \quad (50)$$

g	σ_y	$\bar{\theta}$	κ_θ	ν	r
0.018	0.03	0.35	0.2	0.1	0.025

Table 8: Calibrated Parameters for the CC model.

The solution takes the exponential affine form:

$$P^T(t, \theta_t, y_t) = e^{y_t + F(T-t) - G(T-t)\theta_t}, \quad (51)$$

where the deterministic functions ($F(\tau)$, $G(\tau)$) are derived in the Appendix.

Expected excess returns on the EBIT strips satisfy ($\tau = (T - t)$):

$$\begin{aligned} \frac{1}{dt} \mathbb{E} \left[\frac{dP^T(t, \theta_t, y_t)}{P^T(t, \theta_t, y_t)} - r dt \right] &= -\frac{1}{dt} \mathbb{E} \left[\frac{d\Lambda}{\Lambda} \frac{dP^T(t, \theta_t, y_t)}{P^T(t, \theta_t, y_t)} \right] \\ &= \theta_t [\sigma_y + \nu G(\tau)]. \end{aligned} \quad (52)$$

EBIT strip volatility is

$$\sqrt{\frac{1}{dt} \left(\frac{dP^T(t, \theta_t, y_t)}{P^T(t, \theta_t, y_t)} \right)^2} = [\sigma_y + \nu G(\tau)]. \quad (53)$$

We calibrate this model using the parameter values in the following table. We report the results in Figure (7). As noted in Binsbergen, Brandt and Koijen (2011), the CC model generates an upward sloping term structure of expected returns and volatilities. The term structure of Sharpe ratios is constant, and therefore not reported.

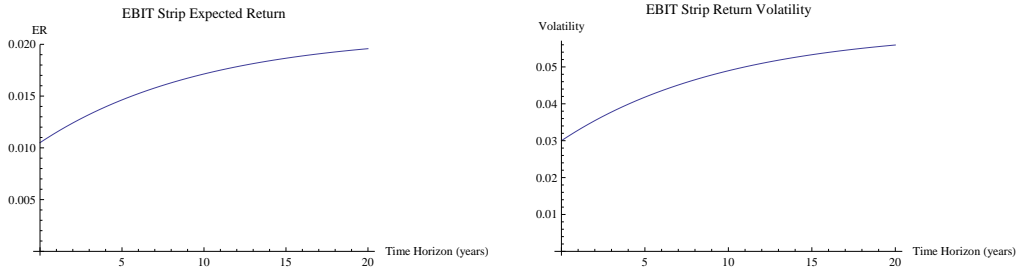


Figure 7: Term structure of EBIT Strip Expected Return (equation (52)) and volatility (equation (53)) for the CC economy. Parameters are set as in Table 8.

The enterprise value of the firm is equal to the present value of the claim to all EBIT strips:

$$P(\theta_t, y_t) = e^{y_t} \int_t^\infty dT e^{F(T-t) - G(T-t)\theta_t}. \quad (54)$$

This can be well-approximated by a log-linear approximation:

$$P(\theta_t, y_t) \approx e^{y_t + F - G\theta_t}, \quad (55)$$

where the coefficients (F, G) are given in Table 9 below.¹⁹ Figure 8 plots the exact and approximate solution, and shows the accuracy of the log-linear approximation.

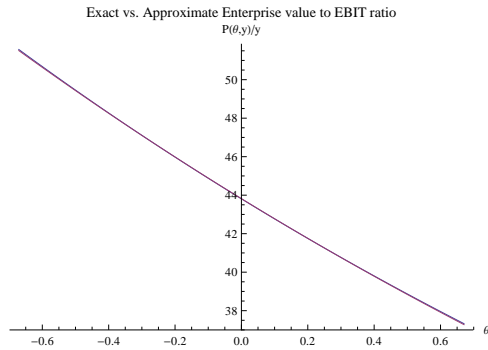


Figure 8: Exact and approximate solutions to the enterprise value as given in respectively equations (21) and (22) for the parameters given in Table 8. The x-axis covers -4 to $+4$ standard deviations of the unconditional distribution of θ .

Under this log-linear approximation, expected return and volatility of the claim to EBIT are:

$$(\mu^P - r)_{CC} \approx \theta_t (\sigma_y + \nu G) \quad (56)$$

$$\sigma_{CC}^P \approx (\sigma_y + \nu G). \quad (57)$$

Their values given the calibrated parameters are given in Table 9:

F	G	$(\mu^P - r)_{CC}$	σ_{CC}^P	Sh_{CC}^P
3.779	0.238	0.019	0.054	0.35

Table 9: Enterprise value expected return, volatility and sharpe ratio for the CC model with parameters set in equation 4.

4.3 Dividend Dynamics

Assume that at all dates- t , the firm issues riskless debt that matures at date- $(t + dt)$ with present value equal to

$$\begin{aligned} B(\ell_t, \theta_t, y_t) &= e^{\ell_t + y_t + F - G\theta_t} \\ &\approx e^{\ell_t} P(\theta_t, y_t). \end{aligned} \quad (58)$$

¹⁹The approach to derive the log-linear approximation is explained in the appendix of Chen, Collin-Dufresne and Goldstein (2008).

As in the previous model, we interpret $e^{\ell_t} \approx \frac{B(\ell_t, \theta_t, y_t)}{P(\theta_t, y_t)}$ as the leverage of the firm. Using an argument analogous to the previous section, we specify leverage dynamics as:

$$\begin{aligned} d\ell_t &= \kappa_\ell (\bar{\ell} + \alpha\theta - \ell) dt - (\nu G + \sigma_y) dz \\ &= \kappa_\ell (\bar{\ell} + \alpha^Q\theta - \ell) dt - (\nu G + \sigma_y) dz^Q, \end{aligned} \quad (59)$$

where

$$\alpha^Q = \alpha + \frac{\nu G + \sigma_y}{\kappa_\ell}. \quad (60)$$

Since the combination $(d\ell + dy - G d\theta)$ is locally deterministic, dividends paid out over the interval $(t, t + dt)$ are equal to $\mathcal{D}(t) dt$, where

$$\begin{aligned} \mathcal{D}(t) &= e^{y_t} \left\{ 1 + e^{\ell_t + F - G\theta_t} \left[\kappa_\ell (\bar{\ell} + \alpha\theta - \ell) + g - r - \frac{\sigma_y^2}{2} - G\kappa (\bar{\theta} - \theta) \right] \right\} \\ &= e^{y_t} \left\{ 1 + e^{\ell_t + F - G\theta_t} \left[\kappa_\ell (\bar{\ell} + \alpha^Q\theta - \ell) + g - r - \frac{\sigma_y^2}{2} - \sigma_y\theta - G\kappa^Q (\bar{\theta}^Q - \theta) \right] \right\} \end{aligned} \quad (61)$$

Note that the terms inside the square bracket follow a stationary process. Hence, dividends are cointegrated with EBIT e^{y_t} .

We define the term structures of expected growth rates and volatilities for dividends over horizon T as in equations (30)-(31). We plot these term structures for $\theta_0 = \bar{\theta}$ in Figure (9).

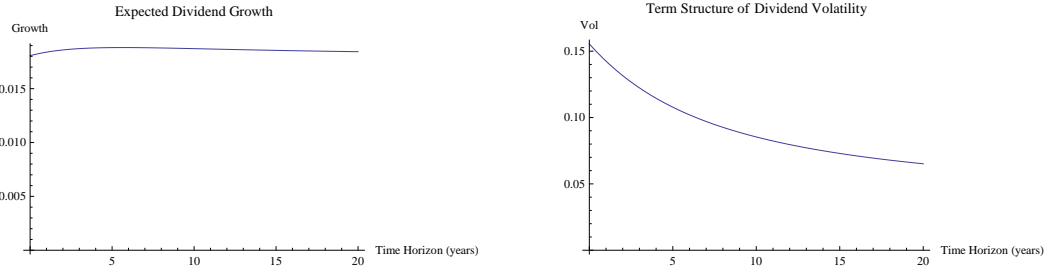


Figure 9: Term structure of expected dividend growth rate (equation (30)) and dividend volatility ($\sigma_{D,T}$ in equation (31)) for the CC economy. Parameters are set as in Table 8.

Note that the term structure of volatilities is downward sloping, consistent with our reported empirical evidence.

4.4 Dividend Strips

As in the previous section, we provide a closed-form expression for the price of dividend strips defined as:

$$V^T(t) = \mathbb{E}_t^Q [e^{-r(T-t)} \mathcal{D}(T)]$$

in terms of the claim:

$$W^T(t, \ell_t, \theta_t, y_t) = \mathbb{E}_t^Q \left[e^{-r(T-t)} e^{\ell_T + y_T + F - G\theta_T} \right],$$

which admits a closed-form exponential affine solution, which we derive in the appendix:

$$W^T(t, \ell_t, \theta_t, y_t) = e^{y_t + H(T-t) + I(T-t)\ell_t - J(T-t)\theta_t}. \quad (62)$$

The date- t present value of a claim to date- T dividends is, as before,

$$V^T(t, \ell_t, \theta_t, y_t) dt = \left[\frac{\partial}{\partial T} W^T(t, \ell_t, \theta_t, y_t) + P^T(t, \theta_t, y_t) \right] dt. \quad (63)$$

Expected excess returns on the dividend strips satisfy

$$\begin{aligned} \frac{1}{dt} \mathbb{E} \left[\frac{dV^T(t, \ell_t, \theta_t, y_t)}{V^T(t, \ell_t, \theta_t, y_t)} - r dt \right] &= -\frac{1}{dt} \mathbb{E} \left[\left(\frac{d\Lambda}{\Lambda} \right) \left(\frac{dV^T(t, \ell_t, \theta_t, y_t)}{V^T(t, \ell_t, \theta_t, y_t)} \right) \right] \\ &= \theta \Omega(\tau, \ell_t, \theta_t, y_t) \end{aligned} \quad (64)$$

where $\Omega(\tau, \ell_t, \theta_t, y_t)$ is given in the Appendix. Dividend strip volatility is

$$\begin{aligned} \sigma^{P, \tau} &\equiv \sqrt{\frac{1}{dt} \left(\frac{dV^T(t, \theta_t, y_t)}{V^T(t, \theta_t, y_t)} \right)^2} \\ &= \Omega(\tau, \ell_t, \theta_t, y_t). \end{aligned} \quad (65)$$

We calibrate the leverage ratio parameters as in Table 10.

$e^{\bar{\ell}}$	α	κ_{ℓ}
.47	-0.5	0.11

Table 10: Parameters for the log-leverage process in the CC Model. (Note that the long run mean of the log-leverage process ℓ_t is $\bar{\ell} + \alpha\bar{\theta} = 0.4$.)

We report the dividend strip return and volatility term structures in Figure (10) below. The instantaneous sharpe ratio is constant across maturities (equal to θ), and hence not reported.

4.5 Equity Returns

The value of equity equals the claim to all dividends:

$$V(\ell_t, \theta_t, y_t) \approx P(\theta_t, y_t) (1 - e^{\ell_t}). \quad (66)$$

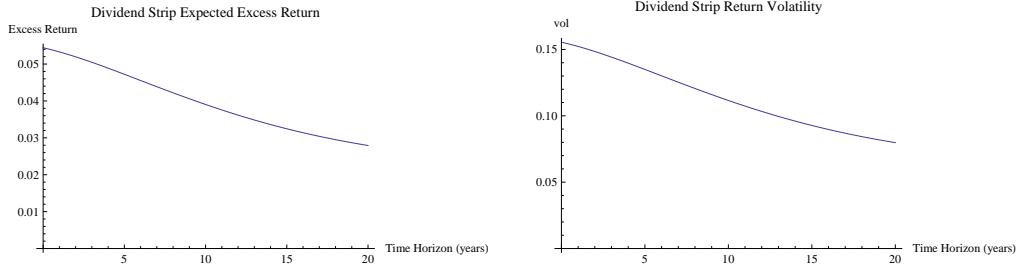


Figure 10: Term structure of Dividend Strip expected returns (equation (64)) and volatilities (equation (65)) for the CC economy. Parameters are set as in Table 8 and 10.

Excess return on equity is equal to excess return on EBIT (equation (56)) scaled by the leverage factor:

$$\mu^V - r = \left(\frac{1}{1 - e^{\ell t}} \right) (\sigma_y + G\nu) \theta_t. \quad (67)$$

This equation captures “dividend irrelevance” in that future capital structure decisions do not impact equity returns (or equity value) today. Similarly, equity return volatility is scaled up by the same factor

$$\sigma^V = \left(\frac{1}{1 - e^{\ell t}} \right) (\sigma_y + G\nu). \quad (68)$$

Results of our calibration are in Table 11.

$(\mu^V - r)_{CC}$	σ_{CC}^V	Sh_{CC}^V
0.036	0.103	.35

Table 11: Stock return expected return and volatility for the CC model with parameters set in Table 8.

4.6 Discussion

Even though in this section we focus on a different asset pricing framework than in the previous section, with time-varying expected return and no cash-flow predictability, we find similar patterns when looking at the term structure of dividend strip return volatilities. With endogenous dividend dynamics derived from a similar mean-reverting process for aggregate leverage, we find that the short term dividend claims are riskier than long-term claims. As a result they display higher volatility and expected returns than long-term claims. (In contrast to the BY

model, here by construction, Sharpe ratios on all claims are equal.) The model also displays ‘excess’ volatility for stock returns in that stock return volatility is higher than long-term dividend volatility by a factor of $1/(1 - e^l) \approx 1.6$. Long term dividend volatility is similar to aggregate cash-flow volatility, which is unaffected by changes in capital structure. Excess returns on leveraged equity (see equation (67)) have two sources of predictability: leverage and the aggregate market price of risk. Together, the last two sections show that when focusing on short term dividend strips (or, indeed, on short term dividend swaps as proposed in BBK) to distinguish different asset pricing models, one has to be careful in accounting for the effects of stationary leverage ratios on the relative riskiness of short versus long-term dividend claims. Irrespective of the framework (i.e., whether there is predictability in cash-flows as in BY, or in risk-premia as in CC), we find that the term structures of expected excess returns and volatilities on dividend claims are decreasing in maturity, and that their term structure of Sharpe ratios are flat.

5 Conclusion

In the long run, EBIT and dividends must be cointegrated, and thus must share long run growth rates and volatility. But in the short run, it is more accurate to think of dividends as a leveraged security. We explain this puzzle by endogenously deriving dividend dynamics when the firm maintains stationary leverage ratios. We argue that when a firm rebalances its debt levels over time to maintain a stationary leverage process, shareholders are being forced to divest (invest) when the firm does well (poorly). In turn, this explains two “puzzling” properties of asset prices. First, we show that this approach generates stock return volatility that is significantly higher than long-term dividend volatility, consistent with Shiller (1981) and LeRoy and Porter (1981). Second, we show that properly accounting for the leveraged nature of dividend dynamics automatically generates a term structure of expected returns and volatilities for dividend strips that are decreasing in maturity, consistent with the empirical findings of Binsbergen, Brandt and Kojen (BBK, 2011).²⁰

Our ‘structural approach’ to dividend dynamics is able to explain the peculiar fact that dividends are “leveraged” in the short run but cointegrated with EBIT in the long run. In contrast, the “reduced form” approach of Abel (1999) ignores the fact that dividends and consumption are cointegrated in the long run.

²⁰See also Binsbergen, Hueskes, Kojen and Vrugt (2011).

6 Appendix

6.1 Proof of Equation(10)

Since

$$M_1^T(t, x_t, y_t) \equiv \mathbb{E}_t [e^{y_T}] \quad (69)$$

is a P-martingale, it follows that

$$0 = M_t - \kappa_x x M_x + M_y \left(g + x - \frac{\sigma_y^2}{2} \right) + \frac{\sigma_y^2}{2} M_{yy} + \frac{\sigma_x^2}{2} M_{xx} + \sigma_y \sigma_{x1} M_{xy}. \quad (70)$$

Since state vector dynamics are exponentially affine, it is well known that the solution is of the form:

$$M^T(t, x_t, y_t) = e^{y_t + A_0(T-t) + x_t A_1(T-t)}. \quad (71)$$

Defining $\tau = (T - t)$, we find initial conditions via

$$\mathbb{E}_T [e^{y_T}] = e^{y_T + A_0(\tau=0) + x_t A_1(\tau=0)}, \quad (72)$$

and hence

$$\begin{aligned} A_0(\tau = 0) &= 0 \\ A_1(\tau = 0) &= 0. \end{aligned} \quad (73)$$

Plugging equation (71) into equation (70), collecting terms linear and independent of x , we obtain the ODE's

$$\begin{aligned} A_1' &= 1 - \kappa_x A_1 \\ A_0' &= g + \frac{\sigma_x^2}{2} A_1^2 + \sigma_y \sigma_{x1} A_1. \end{aligned} \quad (74)$$

The solutions are

$$\begin{aligned} A_1(\tau) &= \frac{1}{\kappa_x} (1 - e^{-\kappa_x \tau}) \\ A_0(\tau) &= g\tau + \left(\frac{\sigma_y \sigma_{x1}}{\kappa_x} \right) (\tau - A_1(\tau)) + \left(\frac{\sigma_x^2}{2\kappa_x^2} \right) \left(\tau - A_1(\tau) - \left(\frac{\kappa_x}{2} \right) A_1^2(\tau) \right) \end{aligned} \quad (75)$$

6.2 Proof of Equation(11)

Since

$$M^T(t, x_t, y_t) \equiv \mathbf{E}_t [e^{2y_T}] \quad (76)$$

is a P-martingale, it follows that

$$0 = M_t - \kappa_x x M_x + M_y \left(g + x - \frac{\sigma_y^2}{2} \right) + \frac{\sigma_y^2}{2} M_{yy} + \frac{\sigma_x^2}{2} M_{xx} + \sigma_y \sigma_{x1} M_{xy}. \quad (77)$$

Since state vector dynamics are exponentially affine, it is well known that the solution is of the form:

$$M^T(t, x_t, y_t) = e^{2y_t + A_2(T-t) + x_t A_3(T-t)}. \quad (78)$$

Defining $\tau = (T - t)$, we find initial conditions via

$$\mathbf{E}_T [e^{2y_T}] = e^{2y_T + A_2(\tau=0) + x_t A_3(\tau=0)}, \quad (79)$$

and hence

$$\begin{aligned} A_2(\tau = 0) &= 0 \\ A_3(\tau = 0) &= 0. \end{aligned} \quad (80)$$

Plugging equation (78) into equation (77), collecting terms linear and independent of x , we obtain the ODE's

$$\begin{aligned} A_3' &= 2 - \kappa_x A_3 \\ A_2' &= 2g + \sigma_y^2 + \frac{\sigma_x^2}{2} A_3^2 + 2\sigma_y \sigma_{x1} A_3. \end{aligned} \quad (81)$$

The solutions are

$$\begin{aligned} A_3(\tau) &= \frac{2}{\kappa_x} (1 - e^{-\kappa_x \tau}) \\ A_2(\tau) &= \left(2g + \sigma_y^2 \right) \tau + \left(\frac{2\sigma_y \sigma_{x1}}{\kappa_x} \right) (2\tau - A_3(\tau)) + \left(\frac{\sigma_x^2}{2\kappa_x^2} \right) \left(4\tau - 2A_3(\tau) - \left(\frac{\kappa_x}{2} \right) A_3^2(\tau) \right). \end{aligned} \quad (82)$$

6.3 EBIT strip in the BY economy

Here we derive the solution to Equation (18):

$$P^T(t, x_t, y_t) = e^{-r(T-t)} \mathbf{E}_t^Q [e^{y_T}]. \quad (83)$$

Note that $e^{-rt}P^T(t, x_t, y_t)$ is a Q-martingale, implying that

$$\begin{aligned} 0 &= \mathbb{E}^Q [d(e^{-rt}P^T(t, x_t, y_t))] \\ &= -rP + P_t + (g^Q + x)yP_y + \kappa_x(\bar{x}^Q - x)P_x + \frac{1}{2}y^2\sigma_y^2P_{yy} + \frac{\sigma_x^2}{2}P_{xx} + \sigma_y\sigma_{x1}yP_{xy}, \end{aligned} \quad (84)$$

where we have defined $\sigma_x^2 \equiv \sigma_{x1}^2 + \sigma_{x2}^2$. Since the state vector dynamics are affine, it is well known (see, for example, Duffie and Kan (1996)) that the solution takes the exponential affine form:

$$P^T(t, x_t, y_t) = y_t e^{F(T-t) + G(T-t)x_t}, \quad (85)$$

Plugging this functional form into equation (84) and then collecting terms linear and independent of x , we find that the deterministic functions $F(\tau)$ and $G(\tau)$ (where $\tau \equiv (T - t)$) satisfy the Ricatti equations:

$$G_\tau = 1 - \kappa_x G \quad (86)$$

$$F_\tau = (g^Q - r) + G(\sigma_y\sigma_{x1} + \kappa_x\bar{x}^Q) + \frac{\sigma_x^2}{2}G^2, \quad (87)$$

with boundary conditions $G(0) = 0$, $F(0) = 0$. The solutions are

$$\begin{aligned} G(\tau) &= \frac{1}{\kappa_x} (1 - e^{-\kappa_x\tau}) \\ F(\tau) &= (g^Q - r)\tau + \left(\frac{\sigma_y\sigma_{x1}}{\kappa_x} + \bar{x}^Q\right) \left(\tau - G(\tau)\right) + \left(\frac{\sigma_x^2}{2\kappa_x^2}\right) \left(\tau - G(\tau) - \frac{\kappa_x}{2}G^2(\tau)\right). \end{aligned} \quad (88)$$

6.4 Enterprise Value in BY Economy

Here we derive the constants F and G used in equation (22). Enterprise value can be determined via

$$P(x_t, y_t) = \mathbb{E}_t^Q \left[\int_t^\infty ds e^{-r(s-t)} e^{y_s} \right]. \quad (89)$$

The exact solution is

$$P(x_t, y_t) = \int_t^\infty P^s(t, x_t, y_t) ds$$

This explicit solution can be approximated very accurately, by an expression of the form:

$$P(x_t, y_t) \approx e^{y_t + F + Gx_t} \quad (90)$$

where the constants F, G can be derived by some local Taylor expansion argument as in Campbell-Shiller, or by minimizing some global error metric as in the appendix of Chen,

Collin-Dufresne and Goldstein (2008). Since in our case, the closed-form solution is known, we simply minimize the mean-square error between the approximation and the closed-form solution. Results are shown in the main text. The approximation error (difference between exact and approximate solution) in absolute terms is less than 0.002, and in relative terms less than 0.005.

6.5 Solution for dividend strips V_t^T in the BY economy

Recall that

$$V^T(t, \ell_t, x_t, y_t) = \left[\frac{\partial}{\partial T} W^T(t, \ell_t, x_t, y_t) + P^T(t, x_t, y_t) \right].$$

where

$$W^T(t, \ell_t, x_t, y_t) = \mathbb{E}_t^Q \left[e^{-r(T-t)} e^{\ell_T + y_T + F + Gx_T} \right]$$

Therefore all we need is an analytic expression for $W^T(t)$. As show in the main text

$$W^T(t, \ell_t, x_t, y_t) = e^{y_t + H(T-t) + I(T-t)\ell_t + J(T-t)x_t}, \quad (91)$$

where, the functions H, I, J satisfy the PDE (??). After defining time to maturity $\tau \equiv (T - t)$, we find that they solve the following system of equations:

$$\begin{aligned} I'(\tau) &= -\kappa_\ell I \\ J'(\tau) &= \alpha\kappa_\ell I - \kappa_x J + 1 \\ H'(\tau) &= -r + \kappa_\ell \bar{\ell}^Q I + \kappa_x \bar{x}^Q J + g^Q + \frac{\sigma_\ell^2}{2} I^2 + \frac{\sigma_x^2}{2} J^2 - [\sigma_{x1} (G\sigma_{x1} + \sigma_y) + G\sigma_{x2}^2] IJ \\ &\quad - \sigma_y (G\sigma_{x1} + \sigma_y) I + \sigma_{x1} \sigma_y J. \end{aligned} \quad (92)$$

The initial conditions are

$$\begin{aligned} H(0) &= F \\ I(0) &= 1 \\ J(0) &= G. \end{aligned} \quad (93)$$

The solutions are

$$\begin{aligned} I(\tau) &= e^{-\kappa_\ell \tau} \\ J(\tau) &= \left[G - \frac{1}{\kappa_x} - \left(\frac{\alpha\kappa_\ell}{\kappa_x - \kappa_\ell} \right) \right] e^{-\kappa_x \tau} + \frac{1}{\kappa_x} + \left(\frac{\alpha\kappa_\ell}{\kappa_x - \kappa_\ell} \right) e^{-\kappa_\ell \tau} \\ H(\tau) &= F + (g^Q - r) \tau + \left[\kappa_\ell \bar{\ell}^Q - \sigma_y (G\sigma_{x1} + \sigma_y) \right] \int_0^\tau ds I(s) + \left[\kappa_x \bar{x}^Q + \sigma_{x1} \sigma_y \right] \int_0^\tau ds J(s) \\ &\quad + \frac{\sigma_\ell^2}{2} \int_0^\tau ds I(s)^2 + \frac{\sigma_x^2}{2} \int_0^\tau ds J(s)^2 - [\sigma_{x1} (G\sigma_{x1} + \sigma_y) + G\sigma_{x2}^2] \int_0^\tau ds I(s)J(s) \end{aligned} \quad (94)$$

where

$$\begin{aligned}
\int_0^\tau ds I(s) &= \frac{1}{\kappa_\ell} (1 - e^{-\kappa_\ell \tau}) \\
\int_0^\tau ds J(s) &= \frac{1}{\kappa_x} \left[G - \frac{1}{\kappa_x} - \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) \right] (1 - e^{-\kappa_x \tau}) + \frac{1}{\kappa_x} \tau + \left(\frac{\alpha \kappa_\ell^2}{\kappa_x - \kappa_\ell} \right) (1 - e^{-\kappa_\ell \tau}) \\
\int_0^\tau ds I^2(s) &= \frac{1}{2\kappa_\ell} (1 - e^{-2\kappa_\ell \tau}) \\
\int_0^\tau ds J^2(s) &= \frac{1}{2\kappa_x} \left[G - \frac{1}{\kappa_x} - \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) \right]^2 (1 - e^{-2\kappa_x \tau}) + \left(\frac{\tau}{\kappa_x^2} \right) + \left(\frac{1}{2\kappa_\ell} \right) \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right)^2 (1 - e^{-2\kappa_\ell \tau}) \\
&\quad + \frac{2}{\kappa_x^2} \left[G - \frac{1}{\kappa_x} - \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) \right] (1 - e^{-\kappa_x \tau}) + \left(\frac{2}{\kappa_x \kappa_\ell} \right) \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) (1 - e^{-\kappa_\ell \tau}) \\
&\quad + 2 \left[G - \frac{1}{\kappa_x} - \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) \right] \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) \left(\frac{1}{\kappa_x + \kappa_\ell} \right) (1 - e^{-(\kappa_x + \kappa_\ell) \tau}) \\
\int_0^\tau ds I(s)J(s) &= \left(\frac{1}{\kappa_x + \kappa_\ell} \right) \left[G - \frac{1}{\kappa_x} - \left(\frac{\alpha \kappa_\ell}{\kappa_x - \kappa_\ell} \right) \right] (1 - e^{-(\kappa_x + \kappa_\ell) \tau}) + \left(\frac{1}{\kappa_x \kappa_\ell} \right) (1 - e^{-\kappa_\ell \tau}) \\
&\quad + \left(\frac{\alpha}{2(\kappa_x - \kappa_\ell)} \right) (1 - e^{-2\kappa_\ell \tau}). \tag{95}
\end{aligned}$$

Hence,

$$\begin{aligned}
V^T(t, \ell_t, x_t, y_t) &= \frac{\partial}{\partial T} e^{y_t + H(T-t) + I(T-t)\ell_t + J(T-t)x_t} + P^T(t, x_t, y_t) \tag{96} \\
&= W^T(t, \ell_t, x_t, y_t) \left[-\kappa_\ell I(T-t)\ell_t + x_t (\alpha \kappa_\ell I(T-t) - \kappa_x J(T-t) + 1) - r + \kappa_\ell \bar{\ell}^Q I(T-t) \right. \\
&\quad \left. + \kappa_x \bar{x}^Q J(T-t) + g^Q + \frac{\sigma_\ell^2}{2} I^2(T-t) + \frac{\sigma_x^2}{2} J^2(T-t) - I(T-t)J(T-t) [\sigma_{x1}(G\sigma_{x1} + \sigma_y) + G\sigma_{x2}^2] \right. \\
&\quad \left. - I(T-t)\sigma_y(G\sigma_{x1} + \sigma_y) - J(T-t)\sigma_{x1}\sigma_y \right] + P^T(t, x_t, y_t). \tag{97}
\end{aligned}$$

6.6 Identification of Ω_1 and Ω_2

Applying Itô's lemma to $V^T(t) \equiv V^T(t, \ell_t, x_t, y_t)$ we find

$$dV^T(t) = rV^T(t)dt + (V_\ell^T(t)\sigma_{\ell 1} + V_x^T(t)\sigma_{x1} + V_y^T(t)\sigma_y) dZ_1^Q(t) + (V_\ell^T(t)\sigma_{\ell 2} + V_x^T(t)\sigma_{x2}) dZ_2^Q(t)$$

where we have defined

$$\sigma_{\ell 1} = -(G\sigma_{x1} + \sigma_y) \tag{98}$$

$$\sigma_{\ell 2} = -G\sigma_{x2} \tag{99}$$

Therefore

$$\Omega_1 = \frac{(V_\ell^T(t)\sigma_{\ell 1} + V_x^T(t)\sigma_{x1} + V_y^T(t)\sigma_y)}{V^T(t)} \tag{100}$$

$$\Omega_2 = \frac{(V_\ell^T(t)\sigma_{\ell 2} + V_x^T(t)\sigma_{x2})}{V^T(t)} \tag{101}$$

All derivatives can be computed in closed form based on the closed-form expression for V above.

6.7 Solution to the EBIT strip price in the CC economy

Here we derive the solution to the EBIT strip price in the CC economy:

$$P^T(t, \theta_t, y_t) = e^{-r(T-t)} \mathbb{E}_t^Q [e^{y_T}]. \quad (102)$$

From equation (50), we see that $e^{-rt} P^T(t, x_t, y_t)$ is a \mathbb{Q} -martingale, implying that

$$\begin{aligned} 0 &= \mathbb{E}^Q [d(e^{-rt} P^T(t, x_t, y_t))] \\ &= -rP + P_t + (g - \frac{\sigma_y^2}{2} - \sigma_y \theta_t) P_y + \kappa^Q (\bar{\theta}^Q - \theta) P_\theta + \frac{\sigma_y^2}{2} P_{yy} + \frac{\nu^2}{2} P_{\theta\theta} - \sigma_y \nu P_{\theta y}, \end{aligned} \quad (103)$$

Since the state vector dynamics are affine, it is well known (see, for example, Duffie and Kan (1996)) that the solution takes the exponential affine form:

$$P^T(t, \theta_t, y_t) = y_t e^{F(T-t) - G(T-t) \theta_t}. \quad (104)$$

Plugging this functional form into equation (103) and then collecting terms linear and independent of θ , we find that the deterministic functions $F(\tau)$ and $G(\tau)$ (where $\tau \equiv (T - t)$) satisfy the Ricatti equations:

$$G_\tau = \sigma_y - \kappa^Q G \quad (105)$$

$$F_\tau = (g - r) + G(\sigma_y \nu - \kappa^Q \bar{\theta}^Q) + \frac{\nu^2}{2} G^2, \quad (106)$$

with boundary conditions $G(0) = 0$, $F(0) = 0$. The solutions are

$$G(\tau) = \frac{\sigma_y}{\kappa^Q} (1 - e^{-\kappa^Q \tau}) \quad (107)$$

$$F(\tau) = (g - r) \tau + \left(\frac{\sigma_y \nu - \kappa^Q \bar{\theta}^Q}{\kappa^Q} \right) (\sigma_y \tau - G(\tau)) + \left(\frac{\nu^2}{2 \kappa^Q} \right) \left[\left(\frac{\sigma_y}{\kappa^Q} \right) (\sigma_y \tau - G(\tau)) - \frac{1}{2} G^2(\tau) \right].$$

6.8 Enterprise Value in the CC Economy

The enterprise value can be determined via

$$P(\theta_t, y_t) = \mathbb{E}_t^Q \left[\int_t^\infty ds e^{-r(s-t)} e^{y_s} \right] \quad (108)$$

$$(109)$$

The exact solution is:

$$P(\theta_t, y_t) = \int_t^\infty P^s(t, \theta_t, y_t) ds.$$

We approximate this exact solution, very accurately, by an expression of the form:

$$P(\theta_t, y_t) \approx e^{y_t + F - G\theta_t} \quad (110)$$

by simply minimizing the mean-square error between the approximate and the closed-form solution. Results are shown in the main text. The approximation error (difference between exact and approximate solution) in absolute terms is less than 0.06, and in relative terms less than 0.0012.

6.9 Solution for dividend strips V_t^T in the CC economy

Recall that

$$V^T(t, \ell_t, \theta_t, y_t) = \left[\frac{\partial}{\partial T} W^T(t, \ell_t, \theta_t, y_t) + P^T(t, \theta_t, y_t) \right]. \quad (111)$$

where

$$W^T(t, \ell_t, \theta_t, y_t) = \mathbb{E}_t^Q \left[e^{-r(T-t)} e^{\ell_T + y_T + F - G\theta_T} \right]$$

It can be shown that $W^T(t, \ell_t, \theta_t, y_t)$ satisfies the PDE

$$\begin{aligned} 0 = & -rW + W_t + W_\ell \kappa_\ell (\bar{\ell} + \alpha^Q \theta - \ell) + W_y \left(g - \frac{\sigma_y^2}{2} - \sigma_y \theta_t \right) + W_\theta \kappa^Q (\bar{\theta}^Q - \theta_t) \\ & + \frac{\sigma_\ell^2}{2} W_{\ell\ell} + \frac{\sigma_y^2}{2} W_{yy} + \frac{\nu^2}{2} W_{\theta\theta} - \sigma_\ell \sigma_y W_{\ell y} + \nu \sigma_\ell W_{\ell\theta} - \nu \sigma_y W_{y\theta}, \end{aligned} \quad (112)$$

where we have defined

$$\sigma_\ell \equiv (G\nu + \sigma_y). \quad (113)$$

Since the dynamics of the state vector are affine, the solution takes an exponential-affine form:

$$W^T(t, \ell_t, \theta_t, y_t) = e^{y_t + H(T-t) + I(T-t)\ell_t - J(T-t)\theta_t}, \quad (114)$$

To find the functions $(H, I(\tau), J(\tau))$ we plug equation (114) into equation (112) and then collect terms linear in θ , linear in ℓ , and independent of θ and ℓ to obtain three coupled Riccati equations:

$$\begin{aligned} I'(\tau) &= -\kappa_\ell I \\ J'(\tau) &= \sigma_y - \kappa^Q J - \kappa_\ell \alpha^Q I \\ H'(\tau) &= g - r + (\nu \sigma_y - \kappa^Q \bar{\theta}^Q) J + (\kappa_\ell \bar{\ell} - \sigma_y \sigma_\ell) I + \frac{\nu^2}{2} J^2 + \frac{\sigma_\ell^2}{2} I^2 - \nu \sigma_\ell I J. \end{aligned} \quad (115)$$

Initial conditions are $I(0) = 1$, $J(0) = -G$, $H(0) = F$. Solutions are

$$\begin{aligned} I(\tau) &= e^{\kappa_\ell \tau} \\ J(\tau) &= \left(\frac{\kappa_\ell \alpha_Q}{\kappa_\ell - \kappa^Q} \right) e^{-\kappa_\ell \tau} - \left(G + \left(\frac{\kappa_\ell \alpha_Q}{\kappa_\ell - \kappa^Q} \right) + \frac{\sigma_y}{\kappa^Q} \right) e^{-\kappa^Q \tau} + \frac{\sigma_y}{\kappa^Q} \end{aligned} \quad (116)$$

$$\begin{aligned} H(\tau) &= (g - r)\tau + (\nu\sigma_y - \kappa^Q \bar{\theta}^Q) \int_0^\tau J(s) ds + (\kappa_\ell \bar{\ell} - \sigma_y \sigma_\ell) \int_0^\tau I(s) ds \\ &\quad + \frac{\nu^2}{2} \int_0^\tau J^2(s) ds + \frac{\sigma_\ell^2}{2} \int_0^\tau I^2(s) ds - \nu\sigma_\ell \int_0^\tau I(s) J(s) ds. \end{aligned} \quad (117)$$

All integrals can be obtained in closed-form but not reported for the sake of brevity. Finally $V^T(t, \ell_t, \theta_t, y_t)$ can be obtained in closed-form from the expression for $W^T(t)$ and $P(t)$ using equation (111) above.

Applying Itô's lemma to $V^T(t)$ we find:

$$dV^T(t) = rV^T(t)dt + (V_y^T \sigma_y - V_\theta^T \nu - V_\ell^T (\nu G + \sigma_y)) dZ_t$$

Thus

$$\Omega(\tau, \ell, \theta, y) = \frac{(V_y^T \sigma_y - V_\theta^T \nu - V_\ell^T (\nu G + \sigma_y))}{V^T(t)}$$

6.10 Accounting for Investment

We have argued previously that, due to the forced investments/divestments imposed by a stationary leverage ratio policy, our model predicts that leverage positively forecasts dividend growth. Intuitively this follows because if, for example, leverage is low today, management will issue debt to push leverage ratios back to their target value. This generates two channels, both of which lower the future growth rate of dividends. First, funds from the newly issued debt allow management to increase the level of dividends paid today. In addition, this debt issuance forces equity holders to divest – that is, it reduces the size of the equity stake, and in turn, reduces future dividend payments. Both of these channels will imply that low leverage today will generate low dividend growth in the future.

Admittedly, one may be concerned that both of these channels disappear when, for example, firms issue debt to pay for investment opportunities. In this case, funds raised by the debt issue do not go toward current dividend payments, and furthermore the size of the equity stake is not reduced (in fact, it increases if the NPV of the projects are positive.) To address this concern, in this section, we attempt to capture investment within our framework. In order to do so as parsimoniously as possible, we will assume all investments are into zero net present value projects²¹ In particular, we model investments as a simultaneous increase in both EBIT

²¹We also investigated positive NPV projects, and obtained similar findings.

y and leverage ℓ . Intuitively, we are modeling an investment that is paid for by debt (hence, an increase in debt), and immediately generates a permanent increase in EBIT. Specifically, recall that enterprise value and the current value of debt are

$$\begin{aligned} P(y_t, x_t) &= e^{y_t + F + Gx_t} \\ B(y_t, x_t, \ell_t) &= e^{\ell_t + y_t + F + Gx_t}. \end{aligned} \quad (118)$$

Now, consider an investment opportunity that will permanently increase EBIT by Δy . In order to pay for this increase, debt is issued, which in turn increases the level of leverage by $\Delta \ell$. It follows that the new enterprise value and outstanding debt levels become:

$$\begin{aligned} P(y_t + \Delta y_t, x_t) &= e^{y_t + \Delta y_t + F + Gx_t} \\ B(y_t + \Delta y_t, x_t, \ell_t + \Delta \ell_t) &= e^{\ell_t + \Delta \ell_t + y_t + \Delta y_t + F + Gx_t}. \end{aligned} \quad (119)$$

For reasons of parsimony, we consider only investments (or disinvestments) into zero NPV projects. As such, enterprise value changes by the same amount that outstanding debt value changes (leaving equity unchanged). Hence,

$$P(y_t + \Delta y_t, x_t) - P(y_t, x_t) = B(y_t + \Delta y_t, x_t, \ell_t + \Delta \ell_t) - B(y_t, x_t, \ell_t) \quad (120)$$

This implies that for a given $\Delta \ell$, the value of Δy is determined via

$$\begin{aligned} \Delta y &= \log \left(\frac{1 - e^{\ell_t}}{1 - e^{\ell_t + \Delta \ell}} \right) \\ &\stackrel{\mathcal{O}(\Delta \ell)^2}{\approx} \left(\frac{e^{\ell}}{1 - e^{\ell}} \right) \Delta \ell + \frac{1}{2} (\Delta \ell)^2 \left[\left(\frac{e^{\ell}}{1 - e^{\ell}} \right)^2 + \left(\frac{e^{\ell}}{1 - e^{\ell}} \right) \right]. \end{aligned} \quad (121)$$

Hence, to account for investments, we modify state vector dynamics for (y_t, x_t, ℓ_t) from equations (8), (9), (??) to:

$$\begin{aligned} dy &= \left(g + x - \frac{\sigma_y^2}{2} \right) dt + \sigma_y dz_1 + \Delta y. \\ dx &= -\kappa_x x dt + \sigma_{x1} dz_1 + \sigma_{x2} dz_2. \\ d\ell_t &= \kappa_\ell \left(\bar{\ell}^P + \alpha x - \ell \right) dt - (G\sigma_{x1} + \sigma_y) dz_1 - G\sigma_{x2} dz_2 + \Delta \ell. \end{aligned} \quad (122)$$

$$(123)$$

We plot the results in Figure (??) for the case $\Delta \ell = 0.06 dz_2$. Interestingly, we find that the slope of the term structure of dividend strip volatilities becomes even more steep.

6.11 Guaranteeing positive dividends: Asset Sales

Note that our model allows for both negative dividends and leverage ratios higher than unity, which implies that the assumption of risk-free debt is not really warranted (if there is limited liability). Here we extend the model by introducing asset sales that keep dividends positive and insure that leverage remains smaller than unity, in turn insuring that debt is risk-free. The model can only be solved numerically, but we find that for the parameter set used, the differences between the closed-form presented in the text and solution with asset sales are negligible.

To make the model more realistic, we assume that when the instantaneous dividend approaches zero, the firm sells assets in order to repurchase some outstanding debt, in turn lowering the leverage ratio. Thus, for example, if the current dividend is negative, and the current state vector is (y_t, θ_t, ℓ_t) , it follows that the current enterprise value and current value of debt is

$$\begin{aligned} P(y_t, \theta_t) &= e^{y_t + F - G\theta_t} \\ B(y_t, \theta_t, \ell_t) &= e^{\ell_t + y_t + F - G\theta_t}. \end{aligned} \tag{124}$$

Let us assume that the firm sells assets so that the new enterprise value is

$$\begin{aligned} P(y_t - \Delta y_t, \theta_t) &= e^{y_t - \Delta y_t + F - G\theta_t} \\ B(y_t - \Delta y_t, \theta_t, \ell_t - \Delta \ell_t) &= e^{\ell_t - \Delta \ell_t + y_t - \Delta y_t + F - G\theta_t}. \end{aligned} \tag{125}$$

All funds raised by the asset sale are used to reduce the outstanding debt, implying that

$$P(y_t - \Delta y_t, \theta_t) - P(y_t, \theta_t) = B(y_t - \Delta y_t, \theta_t, \ell_t - \Delta \ell_t) - B(y_t, \theta_t, \ell_t) \tag{126}$$

This implies that for a given $\Delta \ell$, the value of Δy is determined via

$$\Delta y = -\log \left(\frac{1 - e^{\ell_t}}{1 - e^{\ell_t - \Delta \ell}} \right) \tag{127}$$

In figure 11 below we compare the term structures of dividend strip expected returns and volatilities as reported in the main text (for the CC economy), and compare them to the same quantities obtained via simulations, for the model that implements asset sales to keep dividends positive and maintain leverage ratios below one. We see from the figure that the numerical solution are very similar to the closed-form solution, indicating that, for our parameter choices, the probability of dividends going negative does not significantly affect our results. In fact, the simulation also differ from the close-form because, to be closer to our empirical approach,

we also time-aggregate dividends over one year, as we explain next. However, both effects, time aggregation and asset sales, appear to have negligible effects on our results relative to the closed-form solution reported in the text.

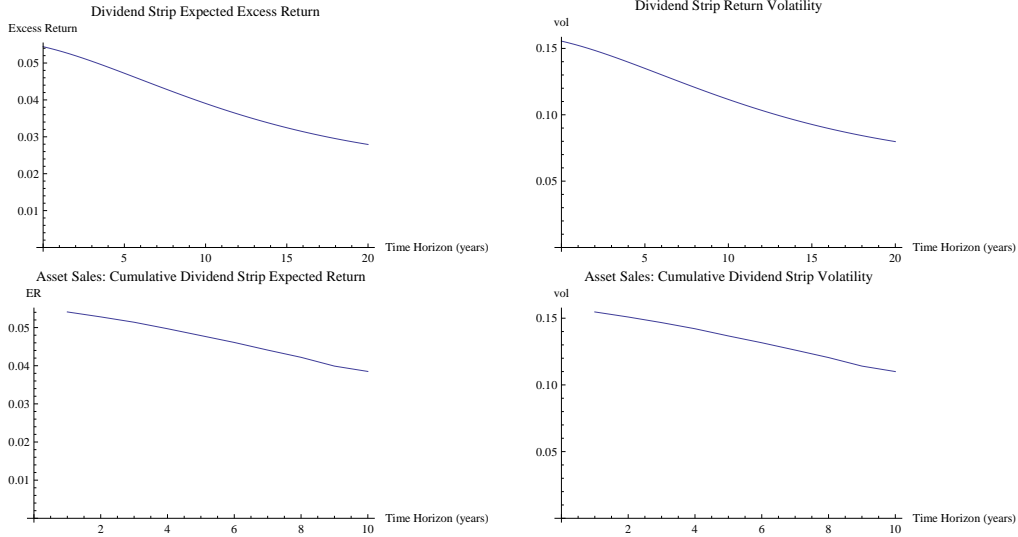


Figure 11: Term structure of Dividend Strip expected returns (equation (64)) and volatilities (equation (65)) for the CC economy versus similar term structure for claims in a model where asset sales are implemented whenever leverage reaches one to maintain positive dividends and guarantee that leverage cannot go above unity. Parameters are set as in table 8 and 10.

6.12 Time Aggregation of Dividends

In the theory part we present the term structure of dividend strips defined as claims to the instantaneous dividends, and estimated at the long-run mean of the state vector. This allows us to make use of the closed form solutions for all the prices and moments of returns. Instead, in empirical work typically, we consider dividends aggregated over one full year.

For example, aggregating over one year, the computation of expected dividend growth and variance ratio becomes:

$$\begin{aligned}
 g_{D,T} &\equiv \left(\frac{1}{T}\right) \log \left(\frac{\mathbb{E}_0 [D_{T+1} - D_T]}{\mathbb{E}_0 [D_1 - D_0]} \right) \\
 \sigma_{D,T}^2 &\equiv \left(\frac{1}{T}\right) \log \left[\frac{\mathbb{E}_0 [(D_{T+1} - D_T)^2]}{(\mathbb{E}_0 [(D_{T+1} - D_T)])^2} \right].
 \end{aligned} \tag{128}$$

Also, the dividend strip claims considered by BBK are typically claims to the sum of all dividends paid out by the index over the year. In this section we compare the effects of

aggregating dividends over one year for various statistics presented in the paper, and show that the difference is not economically significant. In figure 11 above we compare the term structures of dividend strip expected returns and volatilities as reported in the main text, which refer to claims to future instantaneous dividend flows, to the expected returns and volatilities on claims to the **cumulative** dividends aggregated over one year before maturity (so the payout at the maturity T of the dividend strip is $\int_{T-1}^T \mathcal{D}(s)ds$). We see from the figure that the slopes on the cumulative dividend claims are almost indistinguishable from the claims to the instantaneous dividend strips (note that the cumulative dividend picture also differs because the model allows for time-aggregation, but both effects have almost no impact on the predictions of the model).

6.13 The basis between dividend strips extracted from options and dividend swaps

Given the discussion in the literature about dividend swaps (e.g., BBK), here we point out that dividend strip prices extracted from index options and dividend strip prices obtained directly from dividend swap strikes, should not, even in theory, be equal. Indeed, the former are a claim to the dividends as they get paid over the maturity of the option contract, whereas the latter are a claim to the *undiscounted* sum of dividends paid out over the maturity of the swap. Especially for long maturities this can lead to a substantial difference. And, of course, the sum of all dividend swap strikes should not be equal, in equilibrium, to the stock value. More specifically, the dividend strip extracted from options or futures, as the difference between the spot price $V(t)$ and the discounted Futures prices $F^T(t) = E_t^Q[V(T)]$ is, by absence of arbitrage, equal to:

$$V(t) - e^{-r(T-t)}F^T(t) = \int_t^T V^s(t)ds \equiv E^Q[\int_t^T e^{-r(s-t)}\mathcal{D}(s)ds]$$

Instead, dividend swaps pay the **undiscounted** sum of all dividends paid out over the life of the swap, minus the swap strike (say K_t^T) at maturity. Thus the arbitrage free swap strike satisfies:

$$K_t^T = E^Q[\int_t^T \mathcal{D}(s)ds]$$

In turn we see that the option implied (cumulative) dividend strips should always be lower than the dividend swap strike prices. Figure 12 shows that the difference becomes significant at longer horizons.

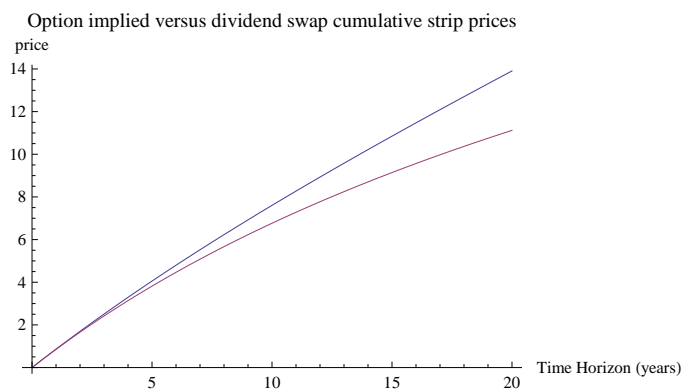


Figure 12: Term structure of dividend Strip prices extracted from options (or futures) compared to the fair dividend swap strike. Because the latter omits discounting it is always higher than the former. Parameters are set as in table 8 and 10.