Can Long-Run Risks Explain the Distress Puzzle?

Juliusz F. Radwański*

Vienna Graduate School of Finance (VGSF)

November 1, 2010

The distress puzzle refers to the empirical regularity that firms with high measures of default likelihood earn anomalously low returns, despite having relatively high CAPM betas. This paper shows it is possible to qualitatively explain this anomaly using a consumption-based asset pricing model related to the literature on long-run risks. Apart from the usual assumptions of time varying, persistent conditional mean and volatility of consumption and dividend growth rates, I assume a natural cointegration between the levels of consumption and aggregate dividends. To model distress, I employ a simple intensity-based model of default. Distressed firms have short expected lifetimes, so they do not covary with the long-run risk factors, and thus earn low expected returns. Healthy firms are long-lived in expectation, and in the presence of the cointegrating relation, their prices do not respond strongly to the innovations in aggregate dividends, which lowers their betas. The model is calibrated to match the standard set of stock market moments and reproduces them well.

Keywords: Distress, long-run risks, cointegration, CAPM.

JEL Classification Numbers: G12, G33.

*Address: Vienna Graduate School of Finance (VGSF), c/o WU (Vienna University of Economics and Business), Heiligenstädter Strasse 46-48, DG 1.21, 1190 Vienna, Austria, telephone: +43-1-31336-6322, e-mail: *juliusz.radwanski@vgsf.ac.at*. The author is grateful to Hanno Lustig, Jacob Sagi and Josef Zechner for advice and suggestions, and to Michael Brandt, Thomas Dangl, René Garcia, Alois Geyer, Ron Giammarino, Christian Julliard, Colin Mayer, Gordon Phillips, Philipp Schnabl, Ilya Strebulaev, Oren Sussman and seminar participants at the VGSF Brown Bag Seminar and the EFA Doctoral Tutorial in Frankfurt (2010) for helpful comments. The first version (dated January 15, 2010) was titled *Long-Run Risk, Cointegration, and the Distress Puzzle*.

1 Introduction

The term *distress puzzle* refers to the well-established asset pricing phenomenon that stocks with high measures of distress earn anomalously low returns. This would be of little surprise if their low returns could be explained by some risk measures. It turns out however, that portfolios formed of distressed stocks co-move with the aggregate market more than portfolios of healthy firms, so that the anomaly is indeed a puzzle from the point of view of the traditional CAPM.[1]

The problem can be re-stated in the following way: if aggregate default risk is feared by market participants and hard to diversify away, stocks with high default probabilities or other measures of distress should earn positive abnormal returns, and a factor constructed on the basis of the difference between returns of the most and the least distressed portfolios of firms should bear a positive price of risk. Instead, it has been documented that the opposite holds, and the market price of distress risk has a puzzling negative sign. If anything, distressed stocks are perceived safer or the risk-return paradigm is violated.

In this paper, I show that the anomaly can be at least partly understood within a consumption-based asset pricing framework that features (1) persistent components in expected consumption growth, as in the long-run risks framework of Bansal and Yaron [2004], (2) a representative agent with Epstein-Zin [1989] utility, and (3) a natural cointegrating relation between the levels of aggregate consumption and dividends.[2] In other words, I show that the long-run risks model with the above features replicates the failure of the CAPM at the short end of maturity spectrum of equity, which is necessary and sufficient to explain the distress puzzle within the long-run risks paradigm.

Intuitively, distressed firms have short expected maturities of cash-flows, so that their prices do not respond strongly to innovations in highly persistent conditional moments of

[1]Distressed portfolios also have high loadings in the Fama-French [1996] three-factor model, which makes the puzzle even more difficult to explain.

[2]The model of Bansal and Yaron [2004] features two types of long-run risk factors – persistent conditional mean and volatility of consumption growth. Both are present in my model, although the role of volatility is of smaller importance.

consumption growth, which is exactly what makes them safer from the point of view of the representative agent. On the other hand, the existence of cointegration between consumption and dividends assures that the covariation of long-maturity asset returns with the dividend growth is low, because all dividend surprises above realized consumption growth are expected to be brought down to zero in the long run. The latter effect results in lower betas for the long-maturity assets. The assumption of cointegration is very natural, because it guarantees that the ratio of dividends to consumption is stationary, or that consumption and dividends do not dominate each other in the long run.

In my model, market betas are thus determined by two offsetting effects, that work with different strength at different maturities. The first effect makes the betas increase with maturity, because of the existence of long-run risks. On the other hand, the betas are decreasing with maturity, if cointegration is present in the model. In my calibration, market beta is higher for those stocks, that have their payouts shifted towards the present, including the distressed ones. I find, that this effect can be obtained even if the deviations from cointegration are very persistent. Moreover, inclusion of cointegration does not essentially damage the ability of the model to match the standard set of asset pricing moments.

The model is calibrated to match the means, volatilities and first-order autocorrelations of consumption and dividend growth rates, the equity premium, the risk-free rate and the price-dividend ratio. My calibration is close to Bansal and Yaron [2004] and to Bansal, Kiku and Yaron [2007] in terms of the choice of model parameters. The exception is cointegration, not parametrized in their papers.

To directly test the ability of the calibrated model to explain the distress puzzle, I employ a simple intensity-based framework, in which firms are indexed by probabilities of default per year. I construct ten hypothetical groups of firms with increasing average default likelihoods, which follows Campbell, Hilscher and Szilagyi [2008], who document the puzzle using such a sort (see below). My model reproduces the anomaly quite well. It produces the annualized difference of returns between the most and the least distressed groups of about three percentage points, which is not accounted for by the difference in CAPM betas. Unfortunately, the calibration is not able to match the magnitude of the puzzle in the data. One can conclude that the long-run risks framework adopted here provides only partial understanding of the puzzle, if taken literally.[3]

The distress anomaly is a well-documented phenomenon. In a seminal paper, Dichev [1998] constructs a distress factor based on models of bankruptcy prediction of Altman [1968] and Ohlson [1980], showing that high bankruptcy risk is not rewarded by high returns. Moreover, in his post-1980 subsample, the relation is significantly negative.[4] In a more recent study, Campbell, Hilscher and Szilagyi [2008] perform a similar analysis, sorting firms into groups according to fitted default probabilities from a probit regression model. The results are even more dramatic. The most distressed group earns an average annual return of minus 16%.[5] Qualitatively similar results are reported by Griffin and Lemmon [2002], Garlappi, Shu and Yan [2006], and Breig and Elsas [2007]. The latter study the anomaly in a sample of German firms.[6]

My study fits into the growing literature that attempts to explain the distress anomaly theoretically. Griffin and Lemmon [2002] relate it to indicators of informational asymmetry. Campbell et al. [2008] consider two explanations: markets may be irrational or inefficient. Garlappi et al. [2006] consider the possibility that equity holders possess an American option to a fixed portion of firm assets which is exercised at bankruptcy, resulting in the violation of absolute priority of the debt claim. George and Hwang [2009] use a model with market frictions to show that when financial distress is costly and firms make optimal capital structure decisions, low leverage firms will be endogenously exposed to high systematic risk.

A different explanation is proposed by Von Kalckreuth [2006], who argues that the incen-

[3]Economic models are usually over-simplified, which makes it difficult to interpret their imperfect quantitative predictions as signs of fundamental mis-specification.

[4]This result holds despite positive correlation of default likelihood with size and book-to-market ratio. Fama and French [1992] conjectured that the book-to-market effect is due to distress risk. The paper of Dichev contradicts this intuition.

[5]These results are robust to the survivorship bias, which works in the opposite direction.

[6]Interestingly, Altman [1963] reports anomalously low returns for distressed corporate bonds.

tives to withdraw resources from firms as private benefits are larger for distressed companies. These benefits are included in the price of equity before they are extracted. This amounts to re-defining dividends to include the private benefits, and the true returns are then much higher than returns that are measured.[7]

In an independent study, Avramov, Cederburg and Hore [2010] use a continuous time version of the long-run risks model in an attempt to explain three anomalies, including the distress puzzle.[8] They argue as well, that distressed firms have low-duration cash flows, which makes them less prone to the changes in the long-run expected economic growth. However, their model does not imply that distressed firms have higher CAPM beta, which makes the explanation incomplete. In their model, firms are distinguished from each other by the shares in the aggregate dividend, which always sum to one and follow a multivariate process similar to the one in Menzly, Santos and Veronesi [2004]. The shares also determine, which firms are of lower duration (distressed). Since the shares are conditionally uncorrelated to the aggregate dividend growth, the latter affects all firms in exactly the same way, persistently scaling the expected path of future dividends, irrespectively of duration. The betas of all firms with respect to the current dividend growth are thus all the same, while long duration firms still have higher betas with respect to the long-run risk shocks, so that taken together, healthy firms have higher betas. In my model, cointegration makes the prices of long-maturity equity less responsive to the aggregate dividend growth, which offsets the effect of higher exposure to the long-run risks. The cointegration seems to be necessary to generate higher CAPM betas for distressed equity.

My work is related to the vast literature on the role of long-run risks in asset pricing. Important examples are Bansal and Yaron [2004], Hansen, Heaton and Li [2008], Bansal and Shaliastovich [2009], and Bhamra, Kuehn and Strebulaev [2010]. The latter paper is an interesting application of the long-run risks paradigm to corporate finance.

[7] I consider this explanation as complementary to mine.

[8] The other two puzzles refer to the negative relations between returns and idiosyncratic volatility, found by Ang, Hodrick, Xing and Zhang [2006, 2009], and between returns and dispersion in earnings forecasts, reported by Diether, Malloy and Scherbina [2002].

I also rely on the literature that studies the term structure of equity. The papers close to mine in terms of modeling approach are Drechsler [2006], and Lettau and Wachter [2007, 2010]. Drechsler [2006] considers the term structure of equity in a long-run risks model, focusing on long maturity equity claims, and concludes that the traditional CAPM performs quite well for them. My work focuses on shorter maturities, and explains how the CAPM can fail in the presence of cointegration. Lettau and Wachter [2007, 2010] provide a duration-based explanation of the value premium, not focusing on distressed stocks.

2 The long-run risks model

The exposition of the long-run risks model in this section is standard, and similar to the one in Bansal and Yaron [2004] or Beeler and Campbell [2009]. I start with the description of the representative consumer's utility function and the equilibrium stochastic discount factor. Then, I characterize the price-consumption ratio for the aggregate consumption claim and the price-dividend ratio for the aggregate market.[9] The two ratios are central to the derivation of asset pricing relations. Next, I show how to decompose the market price-dividend ratio to account for differences in payout horizons of single *equity strips*, that is, I solve for the prices of aggregate dividends paid at fixed time horizons. Finally, I solve for the excess returns and market betas.

2.1 The stochastic discount factor

A representative agent maximizes her lifetime utility, given by the recursive formulation,

$$V_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(E_t\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}},$$

[9]The price-consumption ratio can also be referred to as wealth-consumption ratio, since the price in the numerator is the risk-adjusted discounted stream of future consumption.

where $0 < \delta < 1$ is the subjective time-discount factor, C_t is consumption at time t, γ is the coefficient of relative risk aversion (RRA), and θ is defined as $(1-\gamma)/(1-1/\psi)$, where ψ is the elasticity of intertemporal substitution (EIS). V_t depends on future values of expected utility recursively. The term $\left(E_t\left[V_{t+1}^{1-\gamma}\right]\right)^{1/\theta}$ is the certainty equivalent at time t of the random utility at $t+1$ which depends on the agent's aversion towards atemporal risk. The certainty equivalent is combined with time-t consumption through the time aggregator, a CES function with elasticity of substitution given by the EIS.[10]

It is assumed that the wealth of the agent evolves according to the budget constraint

$$W_{t+1} = (W_t - C_t)R_{c,t+1},$$

where W_t is her wealth and $R_{c,t+1}$ is the return on the claim to aggregate consumption. Since consumption and aggregate dividend are modeled as separate processes (as will be seen below), it has to be implicitly assumed that the wealth has a component that is the difference between the value of aggregate consumption and aggregate dividend. It is also assumed that the risk-free asset is in zero supply so that all wealth that is not consumed must be invested in the risky asset that earns return $R_{c,t+1}$.

The consumption process is given exogenously (see below). The equilibrium marginal rate of substitution between date t and $t+1$ can be shown to be[11]

$$M_{t+1} = \delta^{\theta}(C_{t+1}/C_t)^{-\theta/\psi}R_{c,t+1}^{\theta-1}.$$

Since the model turns out to have a convenient log-normal structure, it is more practical to

[10]This form of utility, proposed by Epstein and Zin [1989], is designed to make a distinction between coefficients of risk aversion and elasticity of intertemporal substitution, which is helpful in resolving the equity premium puzzle. For a very nice exposition on this and other forms of non-standard utility functions and their applications in economics, see Backus, Routledge and Zin [2004].

[11]For the derivation, see Cochrane [2008], the appendix.

work with the logarithm of M_{t+1}. I define $m_{t+1} \equiv \log(M_{t+1})$,

$$m_{t+1} = \theta \log(\delta) - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{c,t+1}, \tag{1}$$

where $r_{c,t+1} = \log(R_{c,t+1})$ and Δc_{t+1} is a logarithmic growth rate of aggregate consumption. I will refer to m_{t+1} as log stochastic discount factor (log SDF). Every asset with generic logarithmic (continuously compounded) return r^x_{t+1} must satisfy the standard no-arbitrage condition,

$$E_t[e^{m_{t+1}+r^x_{t+1}}] = 1. \tag{2}$$

Although the recursive formulation of Epstein-Zin utility does not show up in equation (1), it is there implicitly, because of the term $r_{c,t+1}$. The representative consumer defines "good times" not only in terms of the strength of consumption growth one period ahead, but also through the return on her wealth. This opens up a channel for non-zero price of risk associated with any shock that is orthogonal to *current* consumption growth, but affects *future* consumption growth or its *riskiness*. All such shocks are reflected in the innovations to $r_{c,t+1}$. Note also, that when $\gamma = 1/\psi$, then $\theta = 1$ and $r_{c,t+1}$ is not present in the log SDF formula anymore. In this case, the Epstein-Zin utility function reduces to the standard power utility.

2.2 Aggregate consumption and dividends

Define Δc_{t+1} and Δd_{t+1} as logarithmic growth rates of, respectively, aggregate consumption and dividend. They are assumed to have the following dynamics

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + \psi_c x_t + \phi_c y_t + \sigma_t \eta_{t+1} \\
\Delta d_{t+1} &= \mu_d + \psi_d x_t + \phi_d y_t + \varphi \sigma_t u_{t+1} \\
x_{t+1} &= \rho x_t + \varphi_x \sigma_t \varepsilon_{t+1} \\
y_t &\equiv d_t - c_t \\
y_{t+1} &= (\mu_d - \mu_c) + (\psi_d - \psi_c) x_t + (1 + \phi_d - \phi_c) y_t + \sigma_t(\varphi u_{t+1} - \eta_{t+1}) \\
\sigma^2_{t+1} &= (1-\nu)\sigma^2 + \nu \sigma^2_t + \sigma_w w_{t+1} \\
corr(\eta, u) &= \alpha; \quad \phi_d < 0 < \phi_c
\end{aligned} \tag{3}$$

If consumption growth were i.i.d., then the returns on the wealth portfolio $r_{c,t+1}$ in (1) would be perfectly correlated with innovations to consumption growth. The log SDF would trivialize and attach non-zero price of risk only to current consumption shocks. In the above formulations, this is not true because there are three state variables whose innovations potentially affect the return on the wealth portfolio without directly affecting current consumption growth.

The variable x_t has the interpretation of a small but persistent component in the growth rates of consumption and dividends. Similarly, σ^2_t is proportional to conditional variances of innovations to all variables other than itself.[12] Finally, y_t is defined as the discrepancy from a unit cointegrating relation between consumption and dividends.[13] The assumption $\phi_d < 0 < \phi_c$ is necessary to assure stationarity of this discrepancy series, given that cointegration

[12] It is actually possible to replace the constant σ_w with a square root of σ^2_t (scaled by a constant). Time-varying variance would then provide volatility for its own innovation, which would be likely to change the properties of volatility risk in a potentially interesting way. The solution of the model would require solving a quadratic equation at the stage of undetermined coefficient method for the price-consumption ratio.

[13] Unit cointegration means that in the long run all dividends have to be consumed and only dividends can be consumed. This is not restrictive here, because adding a constant to the cointegrating relation does not change dynamic properties of y_t.

does matter in the model. It also has an economic interpretation – if consumption is far below dividends, its growth is expected to accelerate. If dividends are above consumption – their growth will be lower in the future.[14]

In the whole system, there are four shocks $(\eta, u, \varepsilon, w)$, among which only the first two are correlated, with coefficient α. The constants have rather intuitive interpretations. μ_d and μ_c are the fixed parts of expected consumption and dividend growth rates. ψ_c and ϕ_c are the sensitivities of consumption growth with respect to the state variables x_t and y_t. Analogous interpretation holds for ψ_d and ϕ_d. ρ and ν are the autocorrelation coefficients of x_t and σ_t^2. φ and φ_x determine the conditional volatilities of the dividend growth Δd_{t+1} and innovation to x_t, relative to the volatility of consumption growth. σ_w governs the conditional volatility of shocks to the variance σ_2. Parameter values are chosen in the next section to match the most important moments of consumption and dividend growth rates, as well as those of the risk-free rate, the market excess return and the price-dividend ratio.

The system of equations (3) is very similar to that studied by Bansal and Yaron [2007] and Drechsler [2006]. Bansal and Yaron specify the process of the discrepancy from the cointegrating relation directly, rather than deriving it from consumption and dividend processes. They also do not consider a potential feedback from the cointegrating relation on consumption growth, setting $\phi_c = 0$. Drechsler [2006] assumes a richer correlation structure between the shocks, which is not necessary for the main point of my paper.

3 The solution

The details on how to solve the model defined by (1) and (3) can be found in Appendix A, and all propositions referred to in this section can be found there. Here, I only summarize the the most important results.

Since consumption and dividends are modeled separately, there are two aggregate valuation ratios in the economy, the price-consumption and the price-dividend ratio. They have

[14]This assumption also eliminates the possibility that $\phi_d = \phi_c$, which would imply the unit root in y.

convenient log-affine forms in the state variables x_t, y_t, and σ_t^2. The affine property is the consequence of return log-linearizion borrowed from Campbell and Shiller [1988].

The price-consumption ratio has the form

$$z_{c,t} = A_{c,0} + A_{c,1}x_t + A_{c,2}y_t + A_{c,3}\sigma_t^2, \quad (4)$$

and the price-dividend ratio is

$$z_{d,t} = A_{d,0} + A_{d,1}x_t + A_{d,2}y_t + A_{d,3}\sigma_t^2, \quad (5)$$

where the constants are summarized in Propositions 2, and 5.

The stochastic discount factor is

$$m_{t+1} = -m_0 - m_1 x_t - m_2 y_t - m_3 \sigma_t^2 - \Lambda_\eta \sigma_t \eta_{t+1} - \Lambda_u \sigma_t u_{t+1} - \Lambda_\varepsilon \sigma_t \varepsilon_{t+1} - \Lambda_w \sigma_w w_{t+1}, \quad (6)$$

where the coefficients m_i and prices of risk Λ_j are given in Proposition 3.

The risk-free rate is

$$r_{t+1}^f = r_0^f + r_1^f x_t + r_2^f y_t + r_3^f \sigma_t^2, \quad (7)$$

and the reader is referred to Proposition 4 for the details. The risky return on the market (dividend claim) is summarized in Proposition 6, and has the following form

$$r_{t+1}^m = r_0^m + r_1^m x_t + r_2^m y_t + r_3^m \sigma_2 - R_\eta^m \sigma_t \eta_{t+1} - R_u^m \sigma_t u_{t+1} - R_\varepsilon^m \sigma_t \varepsilon_{t+1} - R_w^m \sigma_w w_{t+1}. \quad (8)$$

The focus of this paper is on the risk-return properties of distressed firms. I will argue later on in the paper that the measures of distress can be understood in terms of the expected duration of cash flows,[15] it will therefore be important to understand the risk-return properties of the so-called *equity strips*.[16] They are defined as claims to aggregate dividends

[15]The same kind of argument is in Avramov, Cederburg and Hore [2010].

[16]The same term is used in Drechsler [2006], and Binsbergen, Brandt and Koijen [2010].

paid by the market portfolio at fixed maturities. It is possible to decompose the aggregate price-dividend ratio into an infinite sum of price-dividend ratios of the equity strips,

$$\frac{P_t}{D_t} = \sum_{n=1}^{\infty} \frac{P_t^n}{D_t}.$$

The same log-affine functional form that is true for the aggregate P/D ratio is also valid for individual strips. One conjectures

$$\log(\frac{P_t^n}{D_t}) = Z_{0,n} + Z_{1,n}x_t + Z_{2,n}y_t + Z_{3,n}\sigma_t^2, \tag{9}$$

and the coefficients can be determined by solving a set of Riccati difference equations (23), stated in Appendix A. Proposition 7 describes the realized returns on individual equity strips, in terms of underlying shocks

$$r_{t+1}^n = r_0^n + r_1^n x_t + r_2^n y_t + r_3^n \sigma_t^2 - R_\eta^n \sigma_t \eta_{t+1} - R_u^n \sigma_t u_{t+1} - R_\varepsilon^n \sigma_t \varepsilon_{t+1} - R_w^n \sigma_w w_{t+1}, \tag{10}$$

where the coefficients r_i^n describe how conditional expected returns depend on the state variables, while R_i^n tell us how the unexpected parts of the returns depend on the innovations.

4 Calibration

Following the standard approach in the literature, I calibrate the model to the yearly moments of the data.[17] The parameters are summarized in Table 1, which also shows the calibrations of Bansal and Yaron (BY) [2004] and Bansal, Kiku and Yaron (BKY) [2007, 2009], for comparison. The latter are usually considered as benchmarks for long-run risks literature.

[17] See Kandel and Stambaugh [1991], Campbell and Cochrane [1999] or Bansal and Yaron [2004].

4.1 Parameter values

The preference parameters $\gamma = 15$ and $\psi = 1.5$ are the same for all three calibrations. I chose the time preference parameter $\delta = 0.9989$, the same as BKY. The parameters governing the consumption process are identical to those in BY. In particular, I restrict the consumption-related coefficient on the deviation from the cointegration relation ϕ_c to be zero, which is a simplifying assumption and means that in the long run consumption provides a trend for dividends, and not the other way round.The properties of the persistent component of consumption growth x_t are the same as in BY. $\rho = 979$ implies a half life of shocks to x_t of about 2.7 years.

The calibration of the dividend process is slightly different from both BY and BKY, although closer to the latter. I choose the 'leverage' parameter at $\phi_d = 2$, which means that long-run component x_t affects expected dividend growth two times stronger than consumption growth.I slightly raise the parameter φ that governs conditional volatility of dividend growth, from 6.5 of BKY to 8.0, and set the correlation coefficient between consumption and dividend growth to $\alpha = 0.5$, compared to 0.4 in BKY.

The cointegration relation in my model operates through dividends. I set $\phi_d = -0.001$, implying that any deviation from the cointegration is very persistent.[18] This makes the model seemingly similar to the one without cointegration, but as I show below, it is needed to capture the pattern of asset betas needed for the main result. Intuitively, since the long-run risk component that produces the positively-sloped pattern of betas is very small, a small impact of cointegration is enough to undo its effect.

Finally, the set of parameters governing the conditional variance process is identical to that of BY. In the calibration of BKY, the mean σ_2 is slightly lower, while its conditional volatility σ_w higher. Coupled with a much higher persistence parameter ν, these assumptions imply much higher unconditional variability of σ_t^2 around the (slightly lower) unconditional mean. As pointed out by Beeler and Campbell [2009], this implies higher probability that

[18] The autoregressive coefficient of the deviation from the cointegration relation is 0.999.

the variance process hits zero at some point.[19] In the calibration of BY (and mine) such events are virtually absent. Moreover, extreme persistence of consumption volatility implied by BKY (the half-life of shocks of 58 years) seems to be exaggerated.

4.2 Calibration outcomes

The set of moments to be matched is standard and includes the dynamics of consumption growth, the dividend growth, the market excess return, the risk-free rate and the price-dividend ratio. For each of them, I consider the mean and the standard deviation of the growth rate, and the yearly first-order autocorrelation. Table 2 shows that the model does a decent job in matching these moments.

The mean consumption growth and its autocorrelation are close to their real-world counterparts. The only exception is consumption growth volatility, which seems to be too high. Bansal, Kiku and Yaron [2009] argue that the value in the data is within the 95% confidence interval around their model's population value, which implies that the latter is not unusual.

The moments of the dividend growth are noticeably different from the empirical counterparts, especially the volatility and autocorrelation seem to be too high. However, the moments are better fitted to the data if one compares them to those of the *earnings*. For the sample between 1930 and 2006, average earnings growth was 2.11%, their volatility amounted to 16.97% and autocorrelation equaled 0.29. These numbers are much closer to the model-implied population values.[20]

The equity premium, its volatility, and autocorrelation are matched by the model very well. This is no surprise, given that the model of Bansal and Yaron [2004] was designed to resolve the equity premium puzzle. The moments of the real interest rate are different than in the data, but paradoxically this can be considered a strength of the model[21] – the high volatility of the real interest rate in the real world is due to frequent changes in inflation.

[19] In such a case it is replaced with a small positive number.

[20] The data on dividends and earnings can be downloaded on Robert Shiller's website, http://www.econ.yale.edu/~shiller/.

[21] This point was made by Beeler and Campbell [2009].

This also explains the relatively low persistence of the empirical value.

Finally, the model matches the moments of the price-dividend ratio relatively well, better than the calibrations of BY and BKY. It slightly understates the volatility, but this moment is known to be difficult to match by the long-run risks models.[22]

5 The results

5.1 Sensitivity of prices w.r.t. changes in the state variables

I start the overview of the model predictions with an analysis of the impact of the state variables on stock prices.

Figure 1 shows the price-dividend ratios of the equity strips (9) as a function of maturity when the state variables x_t, y_t and σ_t^2 are set to their long-run means. The claim with maturity zero has scaled price of unity (immediate payout), analogous to the price of a zero-maturity bond (cash). On the other hand, the claim to the aggregate dividend paid 20 years in the future is worth roughly one-half of the value of the dividend paid with certainty today. The prices decline in maturity, as the claims with longer maturities have higher duration.

How do the prices of equity strips change when the state variables deviate from their long-run values? Figure 2 shows the multiplicative effect that the changes in each of the three state variables by two unconditional standard deviations have on the term structure of P/D ratios. The solid line shows the price responsiveness to the movements in conditional consumption growth x_t. The changes do not affect very short-maturity claims at all. The largest effect on prices occurs at maturities of around 180 months (15 years), which is due to the high persistence of x_t. The effect for very long maturities is gradually and very slowly declining, because higher accumulated dividend growth leads to higher gap between dividends and

[22] See Beeler and Campbell [2009]. Bansal and Shaliastovich [2008] made some progress in this respect, by introducing a third channel for asset price variability – time-varying confidence risk. In their model, the long run component is unobservable by the representative agent, who only observes a set of signals. The precision of the signals (the *confidence*) is time-varying.

consumption, which works on future expected dividend growth in a direction opposite to the initial increase in x_t (effect of cointegration). The dashed line shows how prices are affected by the changes in the deviation from the unit cointegrating relation between dividends and consumption $y_t \equiv d_t - c_t$. Because of high volatility of dividends relative to consumption, the changes in this variable are primarily caused by dividend innovations, so that an increase in y_t is mechanically associated with a fall of the P/D ratios for all maturities. In other words, when y_t is high, today's dividend is high, and dividends in the far future are expected to be relatively low, resulting in decreased prices for long-maturity claims. Finally, the dotted line represents the change in prices caused by higher conditional variance of consumption. Its effect is negative for all maturities, and under my calibration much smaller than the one for x_t. This effect dies out for extremely long maturities (not seen in the graph), because even if the process describing the variance is close to the unit-root, it's effects on dividends must die out in the long run in the presence of cointegration.

5.2 The effect of payout horizon on expected returns

Equation (10) can be used to define expected returns on the equity strips,

$$er_{t+1}^n \equiv \log(E_t[e^{r_{t+1}^n}]),$$

which allows a closed form solution because of the log-linear structure. The expected excess returns on the strips are then

$$rx_{t+1}^n \equiv er_{t+1}^n - r_{t+1}^f, \tag{11}$$

where r_{t+1}^f is given in (7). I annualize the expected excess returns of all strips across maturity prior to plotting them in Figure 3. The solid line corresponds to the 'long-run' situation, in which all conditioning variables are set to zero. The dashed (dotted) line shows expected excess returns when the volatility σ_t^2 is two unconditional standard deviations above (below) its mean. All three lines peak at maturity around 150 months, where the solid line reaches

6.18%, the dashed line 9.00% and the dotted line 3.38%. At the shortest maturity of one month, the three values are much lower: 2.92%, 4.29%, and 1.55%, respectively. Since 150 months is a relatively long time (12.5 years), one can already see that if most distressed firms will be expected to live shorter than that, their expected excess returns will have to be lower. In the long-run situation of the solid line, the spread between the lowest and highest excess returns is about three percent per annum. The actual differences will of course be lower, because even long-lived firms must pay portions of the shortest equity strips as well.[23]

Another interesting question is what are the sources of the differences between excess returns of strips with different maturities. In other words, which risks are priced at those maturities? To answer this question, I first use the standard no-arbitrage condition $e^{m_{t+1}+r^n_{t+1}} = 1$, to derive the risk-return relation

$$rx^n_{t+1} = -cov_t(r^n_{t+1}, m_{t+1}),$$

and using formula (6) for the SDF, I can decompose this covariance into four components

$$\begin{aligned} -cov_t(m_{t+1}, r^n_{t+1}) &= \sigma_t cov_t(r^n_{t+1}, \Lambda_\eta \eta) + \sigma_t cov_t(r^n_{t+1}, \Lambda_u u) \\ &+ \sigma_t cov_t(r^n_{t+1}, \Lambda_\epsilon \epsilon) + \sigma_w cov_t(r^n_{t+1}, \Lambda_w w). \end{aligned}$$

Using (10), this can be written further

$$\begin{aligned} -cov_t(m_{t+1}, r^n_{t+1}) &= R^n_\eta(\Lambda_\eta + \Lambda_u \alpha)\sigma^2_t + R^n_u(\Lambda_u + \Lambda_\eta \alpha)\sigma^2_t + R^n_\epsilon \Lambda_\epsilon \sigma^2_t + R^n_w \Lambda_w \sigma^2_w \\ &= rx^{n,\eta}_{t+1} + rx^{n,u}_{t+1} + rx^{n,\epsilon}_{t+1} + rx^{n,w}_{t+1}, \end{aligned}$$

where the terms in the second row are exactly the portions of the expected excess returns earned by the strips due to their covariances with individual risk factors.

[23]I do not consider other conditioning variables than σ^2_t, because they have zero effect on the expected returns. This can be seen in equations (6) and (10). The volatilities of the innovations to the SDF and to the realized strip returns depend on σ_t only.

I plot the results of the decomposition above (in annualized terms) in Figure 4. Only two components of returns are seen to be of special importance, one due to the exposure to the long-run mean consumption growth shocks ϵ (thick dotted line), and the other due to the exposure to the i.i.d. shocks to the dividend growth u (thick dashed line).[24] At the shortest horizons, only innovations to the dividend growth matter. This is because these innovations are correlated with consumption growth shocks (correlation of α), which directly affects marginal utility of the representative consumer.[25] For longer horizons the importance of this component diminishes, which is caused by cointegration between dividends and consumption in levels: an innovation to the dividend growth is expected to be reversed in the long run, so that the prices of long-maturity assets do not respond strongly to it. However, longer-maturity assets are far more exposed to shocks to the persistent conditional mean of dividend growth. The latter do not affect immediate consumption growth, but have an effect on marginal utility by increasing the agent's wealth. The small magnitude and very high persistence of those shocks are the reasons of why short-maturity asset prices are not affected by them. Intuitively, if the dividend will be paid tomorrow, its price will be independent on what will happen in ten years.

In Figure 4, the effect of x_t is the largest at maturities of around 150 months. Without cointegration, the line would slope up without bound with increasing maturity, which would be the case in the economy of Bansal and Yaron [2004]. Here, the cointegration makes long-term assets safer by reducing the long-run risks for extremely large maturities.

5.3 The effect of the payout horizon on market betas

I move on to analyze market betas of individual equity strips. The natural question that arises is whether the low returns of short-maturity assets can be explained by low covariation

[24]The other two components are far less important, because they have smaller magnitude under my calibration, and because they mainly affect extremely long maturities.

[25]The SDF for the economy with an Epstein-Zin agent can be decomposed in two parts: one that depends on the realized consumption growth, and one that depends on the realized capital gain on wealth. At the shortest maturities, dividend strips covary mainly with the former.

with the market return, or equivalently whether the CAPM works well within the model. The conditional market beta for an n-month equity strip is

$$\beta_t^n = \frac{Cov_t(r_{t+1}^n, r_{t+1}^m)}{Var_t(r_{t+1}^m)}. \tag{12}$$

Figure 5 depicts market betas of the equity strips as a function of maturity, with σ_t^2 set to its long-run value (actually, only σ_t^2 matters, since the other variables do not affect the covariances of returns with the market).[26] It can be seen that under my calibration the CAPM can work well for long maturities, as the decreasing pattern of betas fits decreasing expected returns, shown in Figure 3. This is, however, not the case for the shortest maturities, where expected returns *increase* with maturity, which cannot be explained by the pattern of betas, which are generally downward sloping.[27] It is interesting to note that the betas actually go up slightly for the shortest horizons, which cannot turn the previous conclusion around.[28]

Why are the betas generally decreasing in maturity? To answer this question, I decompose conditional market betas of the equity strips into parts that come from covariation of market return with sources of variation in individual strips, in a way analogous to the decomposition of returns in the previous section. Using (10), I re-write (12) as

$$\begin{aligned} \beta_t^n &= \frac{Cov_t(-R_\eta^n \sigma_t \eta_{t+1} - R_u^n \sigma_t u_{t+1} - R_\varepsilon^n \sigma_t \varepsilon_{t+1} - R_w^n \sigma_w w_{t+1}, r_{t+1}^m)}{Var_t(r_{t+1}^m)} \\ &\equiv \beta_{\eta,t}^n + \beta_{u,t}^n + \beta_{\varepsilon,t}^n + \beta_{w,t}^n. \end{aligned}$$

The resulting components are plotted in Figure 6. The largest portion of betas comes from the exposure to i.i.d. dividend innovations u_{t+1} (dashed line). This component is the only one that is relevant for the shortest-maturity assets, and is also responsible for

[26]Moreover, it turns out that changing σ_t^2 has virtually no effect on betas, so that results can be equally well interpreted as referring to the *unconditional* ones.

[27]This result is in contrast with the findings of Drechsler [2006], who uses a similar model, but under calibrations very different from mine. He does not concentrate on the shortest-maturity equity.

[28]This increasing pattern at the shortest maturities is the effect of high curvature of the dotted line in Figure 6 in that region. See below, for more details on how the figure is constructed.

the generally decreasing overall pattern of betas for longer maturities. The portion that comes from the covariation with the innovations to the long-run risk variable ϵ_{t+1} is much smaller, and it increases with maturity (thick dotted line). The effects from direct exposure to innovations in consumption growth η_{t+1} (thin solid line) and conditional volatility σ_t^2 (thin dotted line) are negligible.

How does the presence of cointegration affect market betas? To provide the answer to this question, I plot the decomposition of betas again, this time for a model in which I assume no cointegration at all. The result is shown in Figure 7. Now, the pattern is very different, the betas do not decrease with maturity, showing a slight upward slope instead. To conclude, the effect that comes from the cointegration is important if one wants to explain empirically higher CAPM betas for short maturity assets.

5.4 Implications for distressed equity

In this section I perform an experiment designed to test the ability of the model to explain the distress puzzle, as observed in empirical studies. I take the hard path and compare the results of my model to the findings of Campbell, Hilscher and Szilagyi [2008], who suggest that the magnitude of the puzzle in the data is especially pronounced. This study suggests a difference of roughly fifteen percentage points per year between the excess returns earned by the least distressed, and the most distressed firms.

5.4.1 Modeling distress

I model distress in a reduced-form fashion. The firms are parametrized by a constant probability of default.[29] Each firm pays a constant (possibly very small) portion of the aggregate dividend until it defaults. In the event of default at time t, the firm will pay the dividend at

[29]Within this model, and under the assumptions listed below, it is hard to distinguish firms and portfolios of firms, since a well-diversified portfolio of firms with a given mean expected default time T will earn exactly the same expected excess return as a single firm with the same T. Therefore, I use the terms 'firms' and 'portfolios of firms' interchangeably, with the hope that no confusion arises.

$t+1$, but nothing in all periods that follow. Finally, I assume that default events are related to firm-specific reasons only.[30]

Finally, I have to make an assumption of what happens with the portion of the aggregate dividend paid by a given firm after it defaults. I adopt a convention that at the default time an identical firm emerges, and continues dividend payments to different stockholders. At any given point in time, only firms that are alive can be traded, which implies that not all dividend claims are available for purchase in the stock market. This does not imply market incompleteness, since the firms that are not yet born can be assumed to be part of the present value of the labor income of the representative agent.

5.4.2 The challenge

I consider ten portfolios with different probabilities of default per year $p \in \{0.0011\%, 0.014\%, 0.018\%, 0.024\%, 0.036\%, 0.057\%, 0.109\%, 0.192\%, 0.340\%, 0.803\%\}$, which correspond exactly to the fitted probabilities of default for ten groups of firms from Campbell, Hilscher and Szilagyi [2008].[31] The groups are based on the percentile cutoffs. The first one contains 5% of the healthiest firms, and the second 5% of firms with slightly higher default probabilities. The third entry corresponds to firms between the 10th and the 20th percentile, and the three following numbers to groups of 20% each. The seventh value is related to firms between 80% and 90% of the distribution, and the last three numbers refer to the most distressed groups, of size 5%, 4% and 1%, respectively.

In Figures 8 and 9, I replicate the Figures 2 and 3 from Campbell, Hilscher and Szilagyi [2008], focusing on unconditional mean excess returns, CAPM betas and alphas. Figure 8 presents the unconditional mean excess returns (solid line) and CAPM alphas (dotted line)

[30]This is similar to the assumption in a related study by Avramov, Cederburg and Hore [2010]. Equivalently, it could be assumed that default events that matter for the aggregate are related only to short-run transitory business fluctuations. Bansal, Kiku and Yaron [2010] show that within the class of long-run risks models, the market price of such temporary risk is very small if the elasticity of intertemporal substitution (EIS) of the representative agent is above one, as is the case here. Most long-run risk models therefore implicitly assume transitory shocks away.

[31]These probabilities correspond to the following expected lifetimes in years: 91, 71, 56, 42, 28, 18, 9, 5, 3, 1.

for ten distress-sorted portfolios (the most distressed on the right). It is clear from there that the extremely low returns earned by the most distressed stocks provide a tough challenge for most asset pricing models. Importantly, CAPM alphas are below the unconditional returns, which means that correcting for risk using this model only makes things worse. Figure 9 makes this point differently by showing that the CAPM market betas or the loadings on the MKT factor from the Fama-French model are increasing in distress, which is at the core of the anomaly.[32] It is immediately clear that these patterns must be difficult to replicate within a consumption-based model, because of the unusually high negative magnitudes of the pricing errors.

5.4.3 Cumulative equity strips

Under the assumptions above, the expected returns implied by my model are pinned down by the distribution of random default times, together with the riskiness of individual equity strips paid up to the moments of default. Before characterizing expected excess returns and market betas of firms with random default times, it is useful to start with a hypothetical firm with deterministic time of default $d > t$, whose future payouts are composed of constant fractions of aggregate dividend between times $t+1$ and $d+1$. The expected excess return of such a firm can be very well approximated (see Appendix B) by the value-weighted average of expected excess returns of the individual strips, that constitute its total payout.[33]

For an asset with fixed default time d, one can define the sequence of weights

$$w^d_{i,t} \equiv \frac{P^i_t}{\sum_{j=1}^{d+1} P^j_t} = \frac{(P^i_t/D_t)}{\sum_{j=1}^{d+1}(P^j_t/D_t)},$$

[32]Campbell et al. do not report CAPM betas. My dotted line is implied from their unconditional returns and CAPM alphas, together with the assumption of mean equity premium of 5.75%.

[33]The approximation would not be necessary if one considered simple returns, as opposed to continuously compounded returns analyzed here.

which can be used to derive the return on such an asset,

$$r_{t+1}^{d,cum} = \sum_{k=1}^{d+1} w_{k,t}^{d} r_{t+1}^{k}. \tag{13}$$

The expected return is

$$er_{t+1}^{d,cum} = \sum_{k=1}^{d+1} w_{k,t}^{d} er_{t+1}^{k}, \tag{14}$$

and the expected excess return

$$rx_{t+1}^{d,cum} = er_{t+1}^{d,cum} - r_{t+1}^{f}. \tag{15}$$

Equation (15) defines a term structure of what I name *cumulative equity strips.* Figure 10 compares the excess returns of cumulative strips (solid line) to those of individual strips (dotted line). The maximum return is now lower and shifted to the right, since cumulative returns contain returns of all earlier maturities with relatively large weights.

Similarly, one can think of the unconditional market beta of the firm that is certain to default at $d > t$. From (13), it follows immediately that this beta is a weighted average of individual strip betas up to d,

$$\beta_t^{d,cum} = \frac{Cov_t(r_{t+1}^{d,cum}, r_{t+1}^{m})}{Var_t(r_{t+1}^{m})} = \frac{Cov_t(\sum_{k=1}^{d+1} w_{k,t}^{d} r_{t+1}^{k}, r_{t+1}^{m})}{Var_t(r_{t+1}^{m})} = \sum_{k=1}^{d+1} w_{k,t}^{d} \beta_t^{k}. \tag{16}$$

The difference between the cumulative and individual betas is illustrated in Figure 12. The former are decreasing more slowly with maturity, because they are in fact weighed averages of the latter, starting from maturity zero up to a given point.

5.4.4 Replicating the distress puzzle

The following proposition justifies a straightforward way to compute expected excess returns of firms with given default probabilities, and their market betas. Below, I use the result to compute these characteristics for the ten portfolios of Campbell, Hilscher and

Szilagyi [2008], discussed in one of the previous subsections.

Proposition 1 *Consider a firm with a constant default probability* p *per period of time. Let* T_p *denote the implied (random) time of default, uncorrelated with the shocks to the state variables, and* r^p_{t+1} *the return realized over one period.*

1. *The (conditional) expected return on this firm is*

$$rx^p_{t+1} = E_t(rx^{T_p,cum}_{t+1}).$$

2. *The (conditional) market beta of the firm is*

$$\beta^p_t = E_t(\beta^{T_p,cum}_m).$$

The proposition says that both expected returns and betas of a firm with random default time can be computed simply by taking expectations of cumulative expected strip returns and cumulative betas, under the distribution of the random time of default T. The proof is given in Appendix B. The second part hinges upon the assumption that the time of default is uncorrelated with the state variables, and therefore also with the market.

It is now easy to compute the expected excess returns, betas and alphas for the ten firms of interest (with exactly the same probabilities of default as in Campbell et al.). The results are in Figures 13 through 15. The model replicates the decreasing pattern of expected returns well. The returns on the first five (the healthiest portfolios) are relatively flat as a function of the measure of distress. For the most distressed ones, the graph steepens sharply. This is a qualitative advantage of the model, since the empirical pattern shown in Figure 8 is very similar to the one implied by the model.

Similarly, the model replicates the pattern of market betas in Figure 10 well. They are generally increasing in maturity, falling slightly for the firms with the two highest probabilities of default. This is again strikingly similar to the empirical pattern, reported in figure

9. Taking the results together, the conclusion is that under my calibration the long-run risks model (1) reproduces the failure of the CAPM in pricing short-maturity assets, and (2) provides a basis for qualitative understanding of the distress puzzle.

Quantitatively though, the model is not able to reproduce the magnitude of the negative returns for the most distressed stocks. The remarkable minus 16% in annualized terms for the most distressed firms reported by Campbell et al. [2008] provide a real challenge.

5.4.5 Re-defining the market

In Figures 13 to 15, one is able to notice that (1) all expected excess returns across the distress portfolios are relatively low, as compared to the equity premium of 5.75% reported in Table 2, (2) the CAPM betas are all above unity, and (3) CAPM alphas are uniformly below zero. Notably, even the healthiest firms seem to deliver expected excess returns below the market portfolio. This inconvenient feature of the model comes from the fact, that the dividends are co-integrated with consumption, which makes very long-term assets relatively safe, as compared to the setting without cointegration. Those assets do not loose value very quickly with maturity, which results in relatively large portion of the market value coming from long-term equity.

Does the issue raised above invalidate the ability of the model to explain the distress puzzle? I argue that it does not. I perform a robustness check to see whether another definition of market return can turn the results around. Starting with the distressed groups analyzed in the paper , one may define the return on the market to be the weighted average of the expected excess returns earned by the ten portfolios. Then the CAPM betas can be re-defined in terms of covariances of the ten groups with the newly defined market measure, divided by the variance of the latter.

The results are plotted in Figures 16 to 18. The expected excess market return is now 4.76%, still well within the range of plausible values. More importantly, the figures show the same qualitative patterns that we have already seen, with the only exception of more intuitive magnitudes of the alphas and betas.

6 Conclusions

This paper proposes a consumption-based explanation of the empirical fact that distressed firms earn low expected returns, despite having higher loadings on the aggregate market factor (CAPM beta). I employ a long-run risks framework similar to that of Bansal and Yaron [2004], with the additional assumption of cointegrated consumption and dividend processes, as in Drechsler [2006] or Bansal and Yaron [2007].

Firms with high default likelihoods are intuitively those that have short expected duration of cash flows, and pay the bulk of their lifetime dividends in the very near future. They are relatively immune to the persistent economic shocks, which makes them less risky. Concerning betas, there are two opposing effects. Persistent shocks to expected dividend growth makes the term structure of betas slightly upward sloping, while the effect of the i.i.d. dividend shocks is that the betas decrease with maturity. Overall, the second effect dominates, but only if the cointegration between dividends and consumption is present in the model. The latter ingredient seems necessary to explain the puzzle in a complete way.

I calibrate my economy successfully to match the most important asset pricing moments. Despite introducing cointegration, not present in Bansal and Yaron [2004], I find that staying close to their calibration does not lead to a deterioration of the ability of the model to fit the data. The effect of cointegration, although not particularly strong from the purely technical point of view, plays an important role in the replication of the stylized facts associated with the distress puzzle, most importantly the failure of the CAPM in the explanation of low expected excess returns on firms that pay most of their dividends in a near future.

The model does a very good job on matching the distress puzzle qualitatively. Nevertheless, the same cannot be said about its quantitative ability to explain the magnitude of the anomaly in the data. I leave it for future work to develop a consumption-based model capable of doing that.

Although I model default in a reduced-form fashion, the mechanism presented here must also be at work in any fully-specified structural model of default, provided that a long-run risk framework is used to provide macroeconomic background (the path pursued by Bhamra, Kuehn and Strebulaev [2010]). To the extent that distressed firms have short expected lifetimes because of some purely idiosyncratic risks, they must also have lower expected returns *ceteris paribus*, if long-run risks are important for asset pricing. It is an interesting task for future research to learn the extent to which the events of default are idiosyncratic. If defaults are related to the asset pricing factors, then it is important to know the sign of the correlation. For example, in the framework of this paper one could easily imagine situations in which the innovations to the expected consumption growth are correlated either positively or negatively with the aggregate default rate. Obviously, when the long-term prospects for economic growth are worsening, we should expect some events of default, but we should also expect defaults when there are radical technological improvements, as in the Schumpeterian effect of creative destruction.

Appendix A – Solving the long-run risks model

This appendix briefly outlines how to solve the long-run risks model summarized in equations (1) and (3). The reader may also consult Bansal, Yaron [2004], Drechsler [2006] or Bansal and Yaron [2007].

Log-linear approximations

The model permits a nice log-linear solution, if one uses the Campbell and Shiller [1988] approximation for realized returns. Let D_s, $s \in \{t,t+1,...\}$ be the stream of aggregate dividends, with ex-dividend price P_t. Let $z_{d,t}$ denote the log price-dividend ratio $\log(P_t/D_t)$,

Δd_{t+1} the dividend growth rate, and $r_{d,t+1}$ the log return. Then,

$$r_{d,t+1} = \kappa_{d,0} + \Delta d_{t+1} + \kappa_{d,1} z_{d,t+1} - z_{d,t}, \tag{17}$$

where $\kappa_{d,0}$ and $\kappa_{d,1}$ are constants of linearization, given by[34]

$$\begin{aligned} \kappa_{d,1} &= \frac{\exp(\bar{z}_d)}{1+\exp(\bar{z}_d)}, \\ \kappa_{d,2} &= \log(1+\exp(\bar{z}_d)) - \kappa_{d,1}\bar{z}_d, \end{aligned} \tag{18}$$

and $\bar{z}$ is the mean log price-dividend ratio. Campbell [1993] and Campbell and Koo [1997] find this approximation highly accurate for models with constant volatility, provided that an iteration is used to find a fixed point for $\bar{z}$ (as is the case here, see below). Kiku [2006] and Bansal and Shaliastovich [2009] compare approximate and numerical solutions of their long-run risks models (with stochastic volatility), concluding that the log-linearization errors are negligible. I therefore rely on the approximate solutions.

An approximation similar to (17) can be applied to the aggregate consumption claim that appears in the log-stochastic discount factor (1). With analogous notation, we have

$$r_{c,t+1} = \kappa_{c,0} + \Delta c_{t+1} + \kappa_{c,1} z_{c,t+1} - z_{c,t}, \tag{19}$$

$$\begin{aligned} \kappa_{c,1} &= \frac{\exp(\bar{z}_c)}{1+\exp(\bar{z}_c)}, \\ \kappa_{c,2} &= \log(1+\exp(\bar{z}_c)) - \kappa_{c,1}\bar{z}_c, \end{aligned} \tag{20}$$

[34] To derive (17), write

$$R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t} = \frac{\left(\frac{P_{t+1}}{D_{t+1}}+1\right)D_{t+1}}{P_t} = \left(\frac{P_{t+1}}{D_{t+1}}+1\right)\frac{D_t}{P_t}\frac{D_{t+1}}{D_t} = (\exp(z_{d,t+1})+1)\frac{D_t}{P_t}\frac{D_{t+1}}{D_t},$$

take logarithms and linearize $\log(\exp(z_{d,t+1})+1)$ around the mean price-dividend ratio $\bar{z}_d$, using the first-order Taylor expansion.

where $z_{c,t}$ is the log of the price-consumption ratio, and $\bar{z}_c$ its long-run mean.

The constants $\kappa_{c,0}$ and $\kappa_{c,1}$ for the price-consumption ratio can be treated as known at this stage. They will be found numerically, as explained below. The former is slightly above zero, and the latter slightly below one. Identical comments and properties apply to the constants related to the price-dividend ratio below, $\kappa_{d,0}$ and $\kappa_{d,1}$.

With the formulas for realized returns to aggregate claims (17) and (19), it is easy to obtain analytical expressions for the expected returns of all assets within the model. In principle, this can be accomplished using the no-arbitrage condition (2), together with the log SDF expression (1). The only elements missing are the valuation ratios $z_{c,t}$ and $z_{d,t}$, as unknown functions of the state variables x_t, y_t, σ_t^2), and the constants in (18) and (20).

The price-consumption ratio

Since the price-consumption ratio is necessary to obtain the log SDF, one needs to find it first. Then, given the SDF, it is possible to solve for the market price-dividend ratio and the valuation ratios of other claims. The following proposition gives $z_{c,t}$ as a function of the state variables and model parameters.

Proposition 2 *The log price-consumption ratio $z_{c,t}$ is affine in the state variables,*

$$z_{c,t} = A_{c,0} + A_{c,1}x_t + A_{c,2}y_t + A_{c,3}\sigma_t^2$$

and

$$
\begin{aligned}
A_{c,2} &= \left(1 - \frac{1}{\psi}\right) \frac{\phi_c}{1 - \kappa_{c,1}(1 + \phi_d - \phi_c)}, \\
A_{c,1} &= \frac{1}{1 - \kappa_{c,1}\rho} \left(1 - \frac{1}{\psi}\right) \left[\psi_c + \kappa_{c,1} s(\psi_d - \psi_c)\right], \\
A_{c,3} &= \frac{1}{2}(1-\gamma)(1-\tfrac{1}{\psi})\left[\tfrac{1+2\kappa_{c,1}s(\varphi\alpha-1)+(\varphi^2-2\alpha\varphi+1)(\kappa_{c,1}s)^2+(\varphi_x\kappa_{c,1}p)^2}{1-\kappa_{c,1}\nu}\right], \\
A_{c,0} &= \frac{1}{1-\kappa_{c,1}}\left[\log(\delta)+(1-\tfrac{1}{\psi})\mu_c+\kappa_{c,0}+\kappa_{c,1}A_{c,2}(\mu_d-\mu_c)+\kappa_{c,1}A_{c,3}(1-\nu)\sigma^2+\tfrac{1}{2}(\theta\kappa_{c,1}A_{c,3}\sigma_w)^2\right], \\
s &\equiv \frac{\phi_c}{1 - \kappa_{c,1}(1 + \phi_d - \phi_c)}, \\
p &\equiv \frac{\psi_c + \kappa_{c,1} s(\psi_d - \psi_c)}{1 - \kappa_{c,1}\rho}.
\end{aligned}
$$

Proof. Assume that the functional form of the log price-consumption ratio log-affine, as claimed in the proposition. The aim is to verify that the return to the consumption claim satisfies (2), when the coefficients $A_{c,i}$ (for $i \in \{0,..,3\}$) are set to the values in the proposition.

As a notational convention, in the proof I suppress the subscripts c. Thus, for example, $r_{c,t+1}$ is just r_{t+1}.

The no-arbitrage condition for the consumption claim is

$$E_t[e^{m_{t+1}+r_{t+1}}] = 1.$$

For convenience, I rewrite the formulas for the log SDF (1) and the return on wealth (19), which appear in the exponent,

$$
\begin{aligned}
m_{t+1} &= \theta \log(\delta) - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{t+1}, \\
r_{t+1} &= \kappa_0 + \Delta c_{t+1} + \kappa_1 z_{t+1} - z_t.
\end{aligned}
$$

The sum of the two is therefore

$$m_{t+1} + r_{t+1} = \theta \log(\delta) + (1 - \frac{1}{\psi})\Delta c_{t+1} + \theta(\kappa_0 + \kappa_1 z_{t+1} - z_t).$$

The following steps are tedious but straightforward. The expressions for z_{t+1} and z_t can be replaced using the conjectured log-affine functional form. Now, the sum is expressed only in terms of the dynamics of consumption plus the state variables at times t and $t+1$. One can use the specification (3) to write everything in terms of state variables x_t, y_t, σ_t^2 and four normally distributed shocks $(\eta, u, \varepsilon, w)$. Now it can be seen that the exponent is a Gaussian random variable. Taking the conditional expectation and noting that the coefficients at all three state variables plus the constant term must all be zero to ensure no-arbitrage, one obtains four equations that can be solved for the four coefficients $A_{c,i}$. ■

The log price-consumption ratio in Proposition 2 depends on the linearization constants $\kappa_{c,0}$ and $\kappa_{c,1}$ in (20). Since they depend on the mean price-consumption ratio, one needs to compute them recursively:[35]

1. Start with an initial guess of $\bar{z}_c$.

2. Compute $\kappa_{c,0}$ and $\kappa_{c,1}$ using (20).

3. Calculate the coefficients $A_{c,i}$ in Proposition 2.

4. Obtain a new value of $\bar{z}_c$, using[36]

$$\bar{z}_c = A_{c,0} + A_{c,1}\bar{x} + A_{c,2}\bar{y} + A_{c,3}\sigma^2 = A_{c,0} + A_{c,2}\frac{\mu_d - \mu_c}{\phi_c - \phi_d} + A_{c,3}\sigma^2.$$

5. With the new value of $\bar{z}_c$, repeat steps (2)–(4) until convergence.

[35] Alternatively, nonlinear solver can be used.

[36] Note that the long-run mean of x_t is zero. When the mean growth rates of consumption and dividends are equal, then also the unconditional mean of y_t is zero. If the model is restricted in a way that the cointegration relation has no impact on consumption and dividend growth rates ($\phi_c = \phi_d = 0$), the formula has to be modified to $\bar{z}_c = A_{c,0} + A_{c,3}\sigma^2$.

This algorithm converges quickly (usually 10-20 iterations). The same method can be used to find $\kappa_{d,0}$ and $\kappa_{d,1}$ in the price-dividend ratio below.

Looking at proposition 2, it is important to note how the parameter ψ that measures the EIS affects the price-consumption ratio. Assume that $\gamma > 1$ and $\phi_c > 0$ (when dividends are above consumption, its growth is likely to accelerate). When $\psi < 1$, the expressions for $A_{c,1}$ and $A_{c,2}$ are negative, while the one for $A_{c,3}$ is positive. Intuitively, this means that the consumer values her consumption *more*, when expected consumption growth is *low*, the deviation from the cointegration relation is *low* (dividends below consumption) or the conditional volatility of consumption is *high*. The explanation is straightforward – for low EIS, when the consumer expects high growth in future consumption (either as an effect of higher conditional growth rate, or dividends being above consumption and pulling it up in expectations), she wants to borrow from the future to smooth her path of utility. This raises the interest rate to the extent such that the higher future consumption is heavily discounted, with negative effect on wealth. Similarly, when the volatility of consumption growth is higher, she is willing to save more because of the precautionary motive, which decreases the interest rate and has a positive effect on wealth. However, with EIS close to $1/\gamma$, the utility function is more similar to power utility, and the model is not able to replicate high equity premium. This implies that in the long-run risks model of Bansal and Yaron [2004] (and its variations), one needs EIS above one. Whether this is the case in the data is still an open question.[37]

The stochastic discount factor and the risk-free rate

Having solved for the price-consumption ratio, I can characterize the stochastic discount factor for the economy.

[37]There have been many attempts to estimate the EIS. Hansen and Singleton [1982], Vissing-Jorgensen [2002], Vissing-Jorgensen and Attanasio [2003] and Guvenen [2006] report values above one. Hall [1988] and Campbell [1996] argue for values much closer to zero.

Proposition 3 *The log stochastic discount factor takes the form*

$$m_{t+1} = -m_0 - m_1 x_t - m_2 y_t - m_3 \sigma_t^2 - \Lambda_\eta \sigma_t \eta_{t+1} - \Lambda_u \sigma_t u_{t+1} - \Lambda_\varepsilon \sigma_t \varepsilon_{t+1} - \Lambda_w \sigma_w w_{t+1},$$

with the coefficients and prices of risk given by

$$\begin{aligned}
m_0 &= -\theta \log(\delta) + \gamma \mu_c + (1-\theta)\{\kappa_{c,0} + \kappa_{c,1}[A_{c,0} + A_{c,2}(\mu_d - \mu_c) + A_{c,3}(1-\nu)\sigma^2)] - A_{c,0}\}, \\
m_1 &= \gamma \psi_c + (1-\theta)[\kappa_{c,1}(A_{c,1}\rho + A_{c,2}(\psi_d - \psi_c) - A_{c,1})], \\
m_2 &= \gamma \phi_c + (1-\theta) A_{c,2}[\kappa_{c,1}(1 + \phi_d - \phi c) - 1], \\
m_3 &= (1-\theta) A_{c,3}(\kappa_{c,1}\nu - 1); \\
\Lambda_\eta &= \gamma - (1-\theta)\kappa_{c,1} A_{c,2}, \\
\Lambda_u &= (1-\theta)\kappa_{c,1} A_{c,2} \varphi, \\
\Lambda_\varepsilon &= (1-\theta)\kappa_{c,1} A_{c,1} \varphi_x, \\
\Lambda_w &= (1-\theta)\kappa_{c,1} A_{c,3}.
\end{aligned}$$

The parameters Λ_x, $x \in \{\eta, u, \varepsilon, w\}$ measure the representative agent's risk aversion associated with the four shocks. With $\psi > 1$ we have $1 - \theta > 0$. The price of direct consumption risk (η_{t+1}) is equal to the coefficient of relative risk aversion, decreased by the amount to which the cointegrating relation hedges against shocks to the current consumption (if there is strong consumption growth, it is likely to partly revert in the future). The price of shocks to the dividend (u_{t+1}) is non-zero only because of the presence of cointegration – if $A_{c,2} = 0$, this price is zero as well. The shocks to the persistent conditional growth rate (ε_{t+1}) and conditional volatility (w_{t+1}) have positive prices of risk. Note that in this model, as in Bansal and Yaron [2004], the prices of risk are constant.

What makes the market risk premia time-varying in the model, is the *amount* of risk given by σ_t. This is one of the most important differences between the long-run risks models and the models with habit formation,[38] in which consumption process is assumed to be homoskedastic and the *market price of risk* varies.

Using the proposition 3, it is straightforward to obtain the continuously compounded risk-free rate r_t^f through

$$e^{-r_t^f} = E_t\left[e^{m_{t+1}}\right]. \tag{21}$$

Proposition 4 *The continuously compounded risk-free rate is*

$$r_{t+1}^f = r_0^f + r_1^f x_t + r_2^f y_t + r_3^f \sigma_t^2,$$

$$\begin{aligned}
r_0^f &= m_0 - \frac{1}{2}\Lambda_w^2 \sigma_w^2, \\
r_1^f &= m_1, \\
r_2^f &= m_2, \\
r_3^f &= m_3 - \frac{1}{2}\Lambda_\eta^2 - \frac{1}{2}\Lambda_u^2 - \Lambda_\eta \Lambda_u \alpha - \frac{1}{2}\Lambda_\varepsilon^2.
\end{aligned}$$

The price-dividend ratio and the market return

To price the market portfolio (the claim to all future dividends), one needs the log-linearized market return (17) and the price-dividend ratio $z_{d,t}$.

Proposition 5 *The log price-dividend ratio* $z_{d,t}$ *is affine in the state variables,*

$$z_{d,t} = A_{d,0} + A_{d,1} x_t + A_{d,2} y_t + A_{d,3} \sigma_t^2,$$

[38]For example, Campbell and Cochrane [1999], or Wachter [2006].

with the coefficients given by

$$\begin{aligned}
A_{d,2} &= \frac{\phi_d - m_2}{1-\kappa_{d,1}(1+\phi_d-\phi_c)}, \\
A_{d,1} &= \frac{\psi_d + \kappa_{d,1}A_{d,2}(\psi_d-\psi_c) - m_1}{1-\kappa_{d,1}\rho}, \\
A_{d,3} &= \frac{1}{1-\kappa_{d,1}\nu}\left[\tfrac{1}{2}(\Lambda_\eta+\kappa_{d,1}A_{d,2})^2+\tfrac{1}{2}(\Lambda_u-\varphi(\kappa_{d,1}A_{d,2}+1))^2+\tfrac{1}{2}(\Lambda_\varepsilon-\kappa_{d,1}A_{d,1}\varphi_x)^2\right] \\
&\quad + \frac{1}{1-\kappa_{d,1}\nu}\left[(\Lambda_\eta+\kappa_{d,1}A_{d,2})(\Lambda_u-\varphi(\kappa_{d,1}A_{d,2}+1))\alpha - m_3\right], \\
A_{d,0} &= \frac{1}{1-\kappa_{d,1}}\left[\kappa_{d,0}+\mu_d+\kappa_{d,1}A_{d,2}(\mu_d-\mu_c)+\kappa_{d,1}A_{d,3}(1-\nu)\sigma^2\right] \\
&\quad + \frac{1}{1-\kappa_{d,1}}\left[\frac{1}{2}\sigma_w^2(\Lambda_w-\kappa_{d,1}A_{d,3})^2 - m_0\right].
\end{aligned}$$

Proof. Almost identical to that of Proposition 2. ■

Define $r^m_{t+1} = \log(D_{t+1}+P_{t+1})/P_t$ to be the continuously compounded market return. The following proposition characterizes it in terms of the four underlying innovations.

Proposition 6 *The continuously compounded return on the aggregate market is*

$$r^m_{t+1} = r^m_0 + r^m_1 x_t + r^m_2 y_t + r^m_3 \sigma_2 - R^m_\eta \sigma_t \eta_{t+1} - R^m_u \sigma_t u_{t+1} - R^m_\varepsilon \sigma_t \varepsilon_{t+1} - R^m_w \sigma_w w_{t+1},$$

and the coefficients are

$$
\begin{aligned}
r_0^m &= \kappa_{d,0} + \mu_d - A_{d,0} + \kappa_{d,1}\left[A_{d,0} + A_{d,2}(\mu_d - \mu_c) + A_{d,3}(1-\nu)\sigma^2\right], \\
r_1^m &= \psi_d - A_{d,1} + \kappa_{d,1}\left[A_{d,1}\rho + A_{d,2}(\psi_d - \psi_c)\right], \\
r_2^m &= \phi_d - A_{d,2}(1 - \kappa_{d,1}(1 + \phi_d - \phi_c)), \\
r_3^m &= -A_{d,3}(1 - \kappa_{d,1}\nu); \\
R_\eta^m &= \kappa_{d,1}A_{d,2}, \\
R_u^m &= -\varphi(1 + \kappa_{d,1}A_{d,2}), \\
R_\varepsilon^m &= -\kappa_{d,1}A_{d,1}\varphi_x, \\
R_w^m &= -\kappa_{d,1}A_{d,3}.
\end{aligned}
$$

The term structure of equity

The focus of this paper is to analyze risk properties of assets that pay their dividends at various maturities. I therefore decompose the aggregate price-dividend ratio into the infinite sum of price-dividend ratios for equity claims (*equity strips*)[39] that pay the aggregate dividend in the economy at fixed future time points. Define P_t^n/D_t as the price of aggregate dividend paid by the economy at date $t+n$, normalized by currently paid dividend. It is straightforward to observe that

$$\frac{P_t}{D_t} = \sum_{n=1}^{\infty} \frac{P_t^n}{D_t}.$$

Define also R_{t+1}^n to be the gross return on an equity strip of maturity n. For $n = 1$, this return is only composed of the dividend payout at $t+1$, while for $n > 1$ there are only capital gains, as in the case of long-maturity bonds. The return can be written

$$R_{t+1}^n = \frac{P_{t+1}^{n-1}}{P_t^n} = \frac{\frac{P_{t+1}^{n-1}}{D_{t+1}} D_{t+1}}{\frac{P_t^n}{D_t} D_t} = \frac{\frac{P_{t+1}^{n-1}}{D_{t+1}}}{\frac{P_t^n}{D_t}} \frac{D_{t+1}}{D_t}.$$

[39] In this nomenclature, I follow Drechsler [2006]. Similar decomposition of equity is also applied in Lettau and Wachter [2007], [2010] and Croce, Lettau and Ludvigson [2006], among others.

Taking logarithms,

$$r^n_{t+1} \equiv \log(R^n_{t+1}) = \log(\frac{P^{n-1}_{t+1}}{D_{t+1}}) - \log(\frac{P^n_t}{D_t}) + \Delta d_{t+1}.$$

I now conjecture the functional form of $\log(P^n_t/D_t)$ to be

$$\log(\frac{P^n_t}{D_t}) = Z_{0,n} + Z_{1,n}x_t + Z_{2,n}y_t + Z_{3,n}\sigma^2_t,$$

which allows me to rewrite the previous equation as

$$\begin{aligned} r^n_{t+1} &= \left[Z_{0,n-1} + Z_{1,n-1}x_{t+1} + Z_{2,n-1}y_{t+1} + Z_{3,n-1}\sigma^2_{t+1}\right] \\ &\quad - \left[Z_{0,n} + Z_{1,n}x_t + Z_{2,n}y_t + Z_{3,n}\sigma^2_t\right] + \Delta d_{t+1}. \end{aligned} \tag{22}$$

It is now straightforward to use the dynamics of the state variables (3) and the no-arbitrage condition (2) applied to (22), to derive a set of Riccati difference equations relating the coefficients $Z_{i,n}$ to $Z_{i,n-1}$; $i \in \{0,..,3\}$:

$$\begin{aligned} Z_{0,n} &= \mu_d - m_0 + Z_{0,n-1} + Z_{2,n-1}(\mu_d - \mu_c) + Z_{3,n-1}(1-\nu)\sigma^2 \\ &\quad + \frac{1}{2}(\Lambda_w - Z_{3,n-1})^2\sigma^2_w, \\ Z_{1,n} &= \psi_d - m_1 + Z_{1,n-1}\rho + Z_{2,n-1}(\psi_d - \psi_c), \\ Z_{2,n} &= \phi_d - m_2 + Z_{2,n-1}(1 + \phi_d - \phi_c), \\ Z_{3,n} &= -m_3 + Z_{3,n-1}\nu + \frac{1}{2}(\Lambda_\eta + Z_{2,n-1})^2 + \frac{1}{2}(\Lambda_u - \varphi(Z_{2,n-1}+1))^2 \\ &\quad + \frac{1}{2}(\Lambda_\varepsilon - Z_{1,n-1}\varphi_x)^2 + (\Lambda_\eta + Z_{2,n-1})(\Lambda_u - \varphi(Z_{2,n-1}+1))\alpha. \end{aligned} \tag{23}$$

The system can be solved recursively, using the set of the initial conditions $Z_{0,0} = Z_{1,0} = Z_{2,0} = Z_{3,0} = 0$, which guarantee that the price of the dividend at time t is exactly D_t

(known at t) so that $P_t^0/D_t = 1$.[40]

Knowing the valuation ratios of all equity strips it is useful for further analysis of their excess returns to write the returns as functions of the shocks to the economy. The proposition below follows from (22) and (23), together with (3).

Proposition 7 *The continuously compounded return on an equity strip that gives the right to the aggregate dividend at time t+n is*

$$r_{t+1}^n = r_0^n + r_1^n x_t + r_2^n y_t + r_3^n \sigma_t^2 - R_\eta^n \sigma_t \eta_{t+1} - R_u^n \sigma_t u_{t+1} - R_\varepsilon^n \sigma_t \varepsilon_{t+1} - R_w^n \sigma_w w_{t+1},$$

where the coefficients are

$$\begin{aligned}
r_0^n &= \mu_d - Z_{0,n} + Z_{0,n-1} + Z_{2,n-1}(\mu_d - \mu_c) + Z_{3,n-1}(1-\nu)\sigma^2, \\
r_1^n &= \psi_d - Z_{1,n} + Z_{1,n-1}\rho + Z_{2,n-1}(\psi_d - \psi_c), \\
r_2^n &= \phi_d - Z_{2,n} + Z_{2,n-1}(1 + \phi_d - \phi_c), \\
r_3^n &= -Z_{3,n} + Z_{3,n-1}\nu; \\
R_\eta^n &= Z_{2,n-1}, \\
R_u^n &= -\varphi(Z_{2,n-1} + 1), \\
R_\varepsilon^n &= -Z_{1,n-1}\varphi_x, \\
R_w^n &= -Z_{3,n-1}.
\end{aligned}$$

[40] Since the focus of this paper is on equity, I do not solve for the zero-coupon bond prices. This is possible using the same set of difference equations, but assuming constant payouts of one at every maturity, instead of the aggregate dividends.

Appendix B – Other Results

Approximation of returns in equations (13) and (14).

As mentioned in the main text, the formulas (13) and (14) are approximate. To justify them, I consider the function

$$f(x_1, \ldots, x_{d+1}) = \log(w_1 e^{x_1} + \cdots + w_{d+1} e^{x_{d+1}}), \tag{24}$$

where I assume that all x_i are close to zero, and the w_i's sum to one. One can approximate this function using the first order multivariate Taylor expansion around $x_1 = x_2 = \cdots = x_{d+1} = 0$, which yields

$$\begin{aligned} f &\approx f(0, \ldots, 0) + f_1(0, \ldots, 0)(x_1 - 0) + \cdots + f_{d+1}(0, \ldots, 0)(x_{d+1} - 0) = \\ &= w_1 x_1 + \cdots + w_{d+1} x_{d+1}. \end{aligned}$$

Now, write the exact version of formula (13):

$$e^{r^{d,cum}_{t+1}} = \sum_{i=1}^{d+1} w^d_{i,t} e^{r^i_{t+1}}.$$

Taking logarithms on both sides, one recognizes the functional form (24), and the approximation yields (13). The formula (13) shows up after applying the same reasoning to the expected return, $er^{d,cum}_{t+1}$.

Proof of Proposition 1.

Throughout the proof, I will use the notation $E_T(.)$, $Cov_T(.)$ and $Var_T(.)$ to denote the expectations and covariances conditional on the information about the random time T_p at time t.

To prove the first part, it is enough to show $er^p_{t+1} = E_t[er^{T_p,cum}_{t+1}]$, since the result will follow by subtracting the risk-free rate from both sides. We have

$$\begin{aligned} er^p_{t+1} &= E_t[r^p_{t+1} + \frac{1}{2}Var_t(r^p_{t+1})] = E_t[E_T(r^p_{t+1} + \frac{1}{2}Var_t(r^p_{t+1}))] \\ &= E_t[E_T(r^{T,cum}_{t+1} + \frac{1}{2}Var_t(r^{T,cum}_{t+1}))] = E_t[r^{T,cum}_{t+1} + \frac{1}{2}Var_t(r^{T,cum}_{t+1})] = \\ &= E_t[er^{T,cum}_{t+1}]. \end{aligned}$$

In words, I condition the expectation upon the information on the possible realizations of the random time T_p.

For the second part, I use the definition of the CAPM beta,

$$\beta^p_t = \frac{Cov_t(r^p_{t+1}, r^m_{t+1})}{Var_t(r^m_{t+1})}. \tag{25}$$

Without loss of generality, assume that the time of default can take only two values, $T_p \in \{s, l\}$, with associated probabilities p_l and p_s. In what follows, I drop the time subscripts t for the clarity of presentation (all expectations and covariances are conditional upon the information set at time t). Equation (25) can be re-written as

$$\frac{Cov(r^p_{t+1}, r^m_{t+1})}{Var(r^m_{t+1})} = \frac{E[Cov_T(r^p_{t+1}, r^m_{t+1})] + Cov(E_T(r^p_{t+1}), E_T(r^m_{t+1}))}{Var(r^m_{t+1})},$$

where I use the law of total covariance for the numerator. The second term in the sum is zero, because $E_T(r^m_{t+1})$ is a constant, equal to (time-t) conditional expected market return (this hinges upon the assumption that T_p is uncorrelated with the market). One can write the numerator further as

$$E[Cov_T(r^p_{t+1}, r^m_{t+1})] = p_s Cov_s(r^{s,cum}_{t+1}, r^m_{t+1}) + p_l Cov(r^{l,cum}_{t+1}, r^m_{t+1}), \tag{26}$$

which leads to

$$\beta_t^p = p_s \frac{Cov_s(r_{t+1}^{s,cum}, r_{t+1}^m)}{Var(r_{t+1}^m)} + p_l \frac{Cov_l(r_{t+1}^{l,cum}, r_{t+1}^m)}{Var(r_{t+1}^m)} = E(\beta_t^{T_p,cum}).$$

References

Altman E. I., *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance 23, (1968), 589-609.

Altman E. I., *Corporate Financial Distress and Bankruptcy*, Wiley, New York, (1993).

Ang A., Hodrick R., Xing Y., Zhang X., *The Cross-Section of Volatility and Expected Returns*, Journal of Finance 61, (2006), 259-299.

Ang A., Hodrick R., Xing Y., Zhang X., *High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence*, Journal of Financial Economics, (2009), 1-23.

Asquith P., Gertner R., Scharfstein D., *Anatomy of Financial Distress: An Examination of Junk-Bond Issuers*, Quarterly Journal of Economics 109, (1994), 625-658.

Avramov D., Cederburg S., Hore S., *Cross-Sectional Asset Pricing Puzzles: An Equilibrium Perspective*, Working Paper, (2010).

Backus D. K., Routledge B. R., Zin S. E., *Exotic Preferences for Macroeconomists*, NBER Macroeconomics Annual 19, (2004).

Bansal R., Dittmar R. F., Lundblad C. T., *Consumption, Dividends and the Cross Section of Equity Returns*, Journal of Finance 60, (2005), 1639-1672.

Bansal R., Kiku D., Yaron A., *Risks for the Long Run: Estimation and Inference*, Working Paper, (2007).

Bansal R., Kiku D., Yaron A., *An Empirical Evaluation of the Long-Run Risks Model for Asset Prices*, NBER Working Paper 15504, (2009).

Bansal R., Kiku D., Yaron A., *Long-Run Risks, the Macroeconomy, and Asset Prices*, Working Paper, (2009).

Bansal R., Shaliastovich I., *Confidence Risk and Asset Prices*, Working Paper, (2008).

Bansal R., Shaliastovich I., *A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets*, Working Paper, (2009).

Bansal R., Yaron A., *Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles*, Journal of Finance 59, (2004), 1481-1509.

Bansal R., Yaron A., *The Asset Pricing-Macro Nexus and Return-Cash Flow Predictability*, Working Paper, (2007).

Beeler J., Campbell J. Y., *The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment*, NBER Working Paper 14788, (2009).

Bhamra H. S., Kuehn L. A., Strebulaev I., *The Aggregate Dynamics of Capital Structure and Macroeconomic Risk*, Review of Financial Studies, forthcoming.

Binsbergen J. H., Brandt M. W., Koijen R. S. J., *On the Timing and Pricing of Dividends*, Working paper, (2010).

Breig C., Elsas R., *The Impact of Default Risk on Equity Returns: Evidence From a Bank-Based Financial System*, Working Paper, (2007).

Campbell J. Y., *Intertemporal Asset Pricing Without Consumption Data*, American Economic Review 83, (1993), 487-512.

Campbell J. Y., *Understanding Risk and Return*, Journal of Political Economy 104, (1996), 298-345.

Campbell J. Y., Cochrane J. H., *By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior*, Journal of Political Economy 107, (1999), 205-251.

Campbell J. Y., Koo H. K., *A Comparison of Numerical and Analytical Approximate Solutions to an Intertemporal Consumption Choice Problem*, Journal of Economic Dynamics and Control 21, (1997), 273-295.

Campbell J. Y., Shiller R. J., *The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors*, Review of Financial Studies 1, (1988), 195-227.

Campbell J., Hilscher J., Szilagyi J., *In Search of Distress Risk*, Journal of Finance 63, (2008), 2899-2939.

Chen L., Da Z., Priestley R., *Dividend Smoothing and Predictability*, Working Paper, (2009).

Cochrane J. H., *Financial Markets and the Real Economy*, Handbook of the Equity Risk Premium, (2008), 237-326.

Croce M. M., Lettau M., Ludvigson S., *Investor Information, Long-Run Risk, and the Duration of Risky Cash Flows*, Working Paper, (2006).

Dichev I. D., *Is the Risk of Bankruptcy a Systematic Risk?*, Journal of Finance 53, (1998), 1131-1147.

Diether K. B., Malloy C. J., Scherbina A., *Differences in Opinion and the Cross Section of Stock Returns*, Journal of Finance 57, (2002), 2113-2141.

Drechsler I., *The Role of Payout Horizon in Determining the Risks and Returns of Claims to Aggregate Cash Flows*, Working Paper, (2006).

Epstein L. G., Zin S. E., *Substitution, Risk Aversion, and the Intertemporal Behavior of Consumption and Asset Returns: A Theoretical Framework*, Econometrica 57, (1989), 937-969.

Fama E. F., French K. R., *The Cross Section of Expected Stock Returns*, Journal of Finance 47, (1992), 427-465.

Fama E. F., French K. R., *Multifactor Explanations of Asset Pricing Anomalies*, Journal of Finance 51, (1996), 55-84.

Fisher E. O., Heinkel R., Zechner J., *Dynamic Capital Structure Choice: Theory and Tests*, Journal of Finance 44, (1989), 19-40.

Garlappi L., Shu T., Yan H., *Default Risk, Shareholder Advantage, and Stock Returns*, Review of Financial Studies 21, (2008), 2743-2778.

George T., Hwang C.-Y., *A Resolution of Distress Risk and Leverage Puzzles in the Cross Section of Stock Returns*, Journal of Financial Economics, (2009), *forthcoming.*

Giesecke K., Longstaff F. A., Schaefer S., Strebulaev I., *Corporate Bond Default Risk: A 150-year Perspective*, Working Paper, (2009).

Griffin J. M., Lemmon M. L., *Book-to-Market Equity, Distress Risk and Stock Returns*, Journal of Finance 57, (2002), 2317-2336.

Guvenen F., *Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective*, Journal of Monetary Economics 53, (2006), 1451-1472.

Hall R. E., *Intertemporal Substitution in Consumption*, Journal of Political Economy 96, (1988), 339-357.

Hansen L. P., Heaton J. C., Li N., *Consumption Strikes Back? Measuring Long-Run Risk*, Journal of Political Economy 116, (2008), 260-302.

Hansen L. P., Singleton K., *Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models*, Econometrica 82, (1982), 767-804.

Kandel E., Stambaugh R., *Asset returns and Intertemporal Preferences*, Journal of Monetary Economics 27, (1991), 39-71.

von Kalckreuth, U., *Wrecking Behaviour and the Overpricing of Distressed Companies*, Working Paper, (2006).

Kiku D., *Is Value Premium a Puzzle?*, Working Paper, (2006).

Leland H. E., *Corporate Debt Value, Bond Covenants and Optimal Capital Structure*, Journal of Finance 49, (1994), 1213-1252.

Lettau M., Wachter J., *Why is Long-horizon Equity Less Risky? A Duration-based Explanation of the Value Premium*, Journal of Finance 62, (2007), 55-92.

Lettau M., Wachter J., *The Term Structures of Equity and Interest Rates*, Working paper, (2010).

Livdan D., Sapriza H., Zhang L., *A Neoclassical Model of Financially Constrained Stock Returns*, Working Paper, (2005).

Lucas R., *Asset Prices in Exchange Economy*, Econometrica 46, (1978), 1429-1451.

Menzly L., Santos T., Veronesi P., *Understanding Predictability*, Journal of Political Economy 112, (2004), 1-47.

Ohlson J. A., *Financial ratios and the Probabilistic Predicyion of Bankruptcy*, Journal of Accounting Research 18, (1980), 109-131.

Opler T., Titman S., *Financial Distress and Corporate Performance*, Journal of Finance 49, (1994), 1015-1040.

Wachter J., *A Consumption-Based Model of the Term Structure of Interest Rates*, Journal of Financial Economics 79, (2006), 365-399.

Vissing-Jorgensen A., *Limited Asset Market Participation and the Elasticity of Intertemporal Substitution*, Journal of Political Economy 110, (2002), 825-853.

Vissing-Jorgensen A., Attanasio O. P., *Stock Market Participation, Intertemporal Substitution, and Risk Aversion*, American Economic Review, Papers and Proceedings, (2003), 383-391.

(A) Preference Parameters

	γ	ψ	δ
me	10	1.5	.9989
BY	10	1.5	.998
BKY	10	1.5	.9989

(B) Consumption Growth

	μ_c	ψ_c	ϕ_c	ρ	φ_e
me	.0015	1	0	.979	.044
BY	.0015	1	0	.979	.044
BKY	.0015	1	0	.975	.038

(C) Dividend Growth

	μ_d	ψ_d	ϕ_d	φ	α
me	.0015	2.0	-.001	8.0	.5
BY	.0015	3.0	0	4.5	.0
BKY	.0015	2.5	0	6.5	.4

(D) Volatility

	σ	σ_w	ν
me	.0078	.0000023	.987
BY	.0078	.0000023	.987
BKY	.0072	.0000028	.999

Table 1. Calibration of parameters. The first row in each panel corresponds to my calibration, the second to Bansal and Yaron [2004] and the third to Bansal, Kiku and Yaron [2009]. All parameters are monthly.

Model Implied Moments

Moment	me	-std	+std	BY	BKY	Data
$E(\Delta c)$	1.80	1.11	2.49	1.79	1.82	1.95
$\sigma(\Delta c)$	2.92	2.58	3.26	2.92	2.96	2.16
$\rho_1(\Delta c)$	0.52	0.41	0.63	0.51	0.44	0.44
$E(\Delta d)$	1.80	-0.21	3.81	1.66	1.85	1.02^a
$\sigma(\Delta d)$	18.00	16.22	19.78	11.57	16.42	10.69^b
$\rho_1(\Delta d)$	0.27	0.16	0.38	0.40	0.29	0.14^c
$E(r^m - r^f)$	5.75	3.54	7.96	6.62	6.58	6.20
$\sigma(r^m - r^f)$	18.60	16.64	20.56	16.88	21.35	18.34
$\rho_1(r^m - r^f)$	0.00	-0.12	0.12	0.03	0.02	0.04
$E(r^f)$	1.46	1.04	1.88	2.56	0.99	0.99
$\sigma(r^f)$	1.37	1.16	1.58	1.30	1.28	4.28
$\rho_1(r^f)$	0.78	0.70	0.86	0.85	0.86	0.59
$E(p - d)$	3.39	3.19	3.59	3.00	3.04	3.31
$\sigma(p - d)$	0.37	0.32	0.42	0.16	0.26	0.46
$\rho_1(p - d)$	0.91	0.77	1.05	0.77	0.95	0.88

Table 2. Model-implied moments. To compute the population moments, I run a hundred repetitions of 10000 years of simulated data (which gives a total of 12 m months), and average the model-implied moments across the repetitions. The last column in the table is taken from Beeler and Campbell [2009], who use yearly time series for the time-span 1930-2006. For the fields with a-c superscripts, I calculate analogous moments using the data from 1930 to 2006 on real earnings from Robert Shiller's website. The results (for the same time span) are: a–2.11, b–16.97, c–0.29.

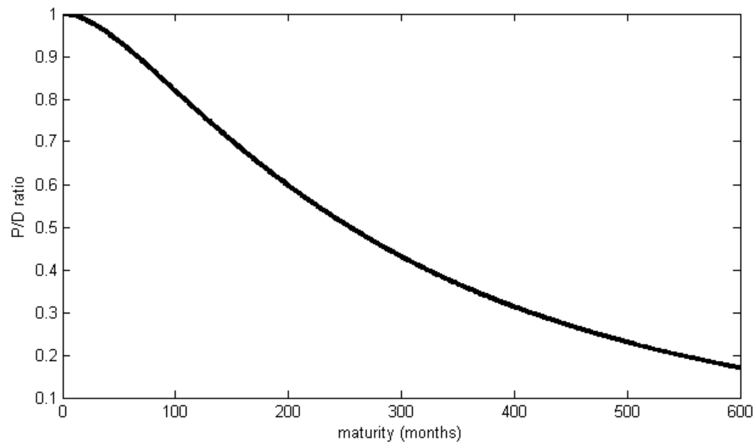

Figure 1. Price-dividend ratios of individual equity strips, when the state variables are set at their long-run mean values $x_t = 0$, $y_t = 0$, and $\sigma_t^2 = \sigma^2$.

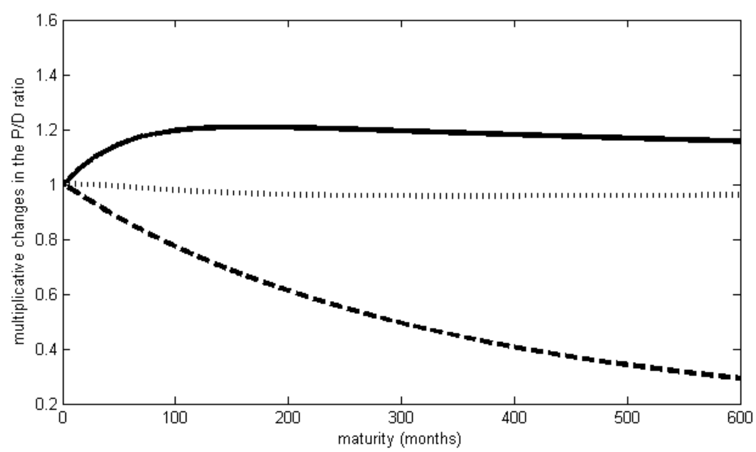

Figure 2. Multiplicative changes in the price-dividend ratios of the equity strips implied by *ceteris paribus* changes of the state variables x_t (solid line), y_t (dashed line), and σ_t^2 (dotted line), by two unconditional standard deviations.

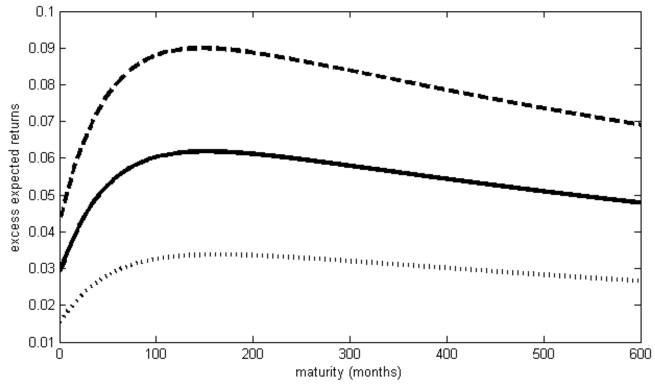

Figure 3. Annualized conditional expected excess returns of the equity strips. The solid line corresponds to a situation in which all state (conditioning) variables are set to their long-run means. The dashed (dotted) line corresponds to a situation in which the conditional variance of consumption growth is two unconditional standard deviations above (below) its mean.

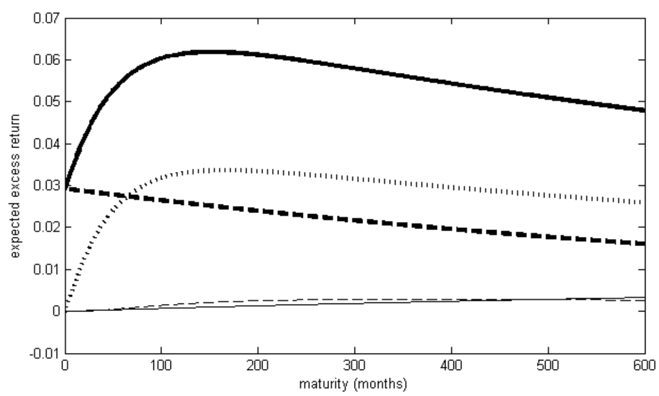

Figure 4. Decomposition of the total conditional expected excess return of the equity strips (thick solid line) into four components, earned due to exposure of each strip to the innovation in consumption growth η_{t+1} (thin solid line), dividend growth u_{t+1} (thick dashed line), conditional mean consumption growth ϵ_{t+1} (dotted line), and conditional variance of the consumption growth w_{t+1} (thin dashed line).

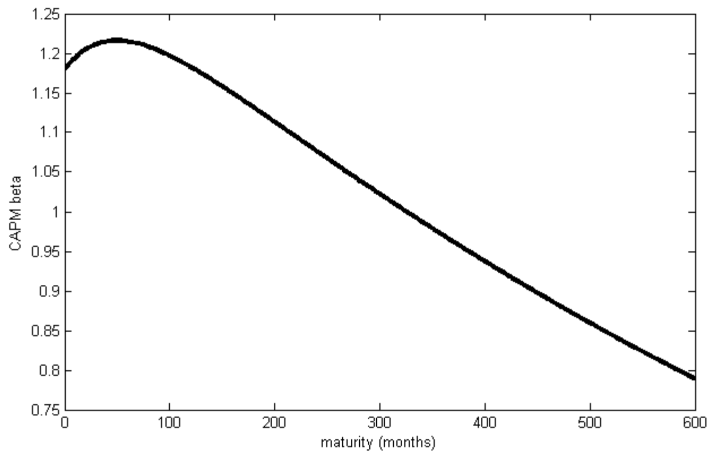

Figure 5. Conditional market (CAPM) betas of the equity strips, when σ^2 is set to its long-run mean value.

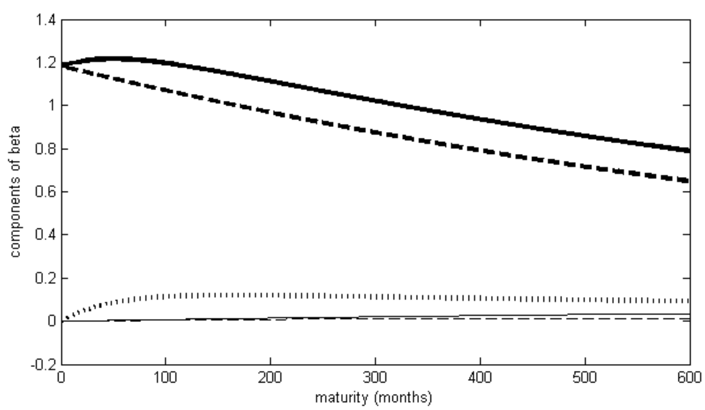

Figure 6. Decomposition of conditional market betas of individual strips (thick solid line) into portions related to the exposure to various risk factors: shocks to the dividend growth (thick dashed line), expected consumption growth (thick dotted line), i.i.d. consumption growth (thin solid line), and conditional variance of consumption growth (thin dotted line). The parameter measuring the strength of cointegration is set to $\phi_d = -0.001$.

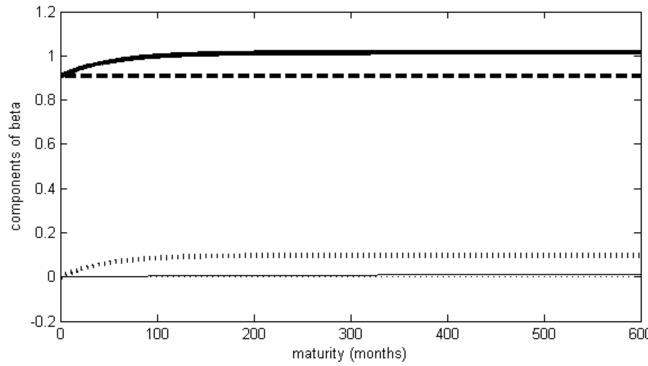

Figure 7. Decomposition of conditional market betas of individual strips (thick solid line) into portions related to the exposure to various risk factors: shocks to the dividend growth (thick dashed line), expected consumption growth (thick dotted line), i.i.d. consumption growth (thin solid line), and conditional variance of consumption growth (thin dotted line). The parameter measuring the strength of cointegration is set to $\phi_d = 0$.

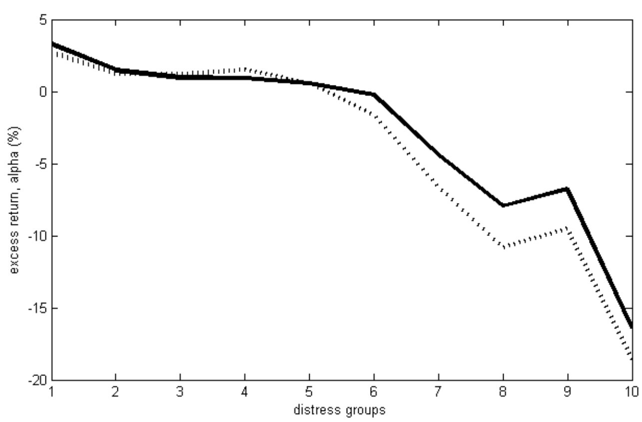

Figure 8. Mean excess returns on ten portfolios sorted with respect to *per year* probability of default (solid line), and CAPM alphas of these portfolios (dotted line). The probabilities of default are (from left to right): $p \in \{0.0011\%, 0.014\%, 0.018\%, 0.024\%, 0.036\%, 0.057\%, 0.109\%, 0.192\%, 0.340\%, 0.803\%\}$. *Source: Campbell et al. [2008].*

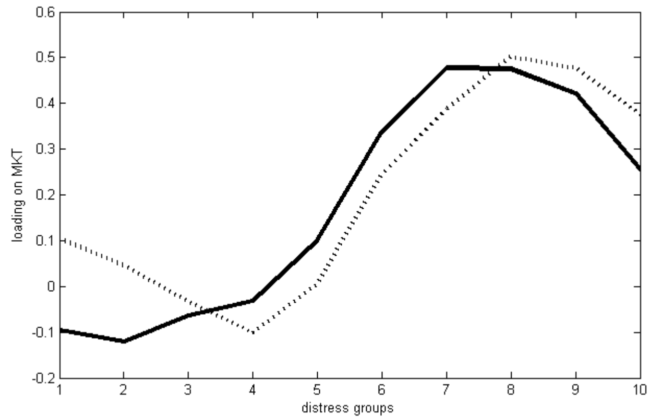

Figure 9. The loadings on the market for ten portfolios of firms sorted with respect to the probability of default *per year*. The probabilities of default are as in Figure 8. The solid line is the loading on MKT in the Fama and French [1996] three factor model, and the dotted line is the loading in on the market in one factor CAPM. *Source: Campbell et al. [2008].*

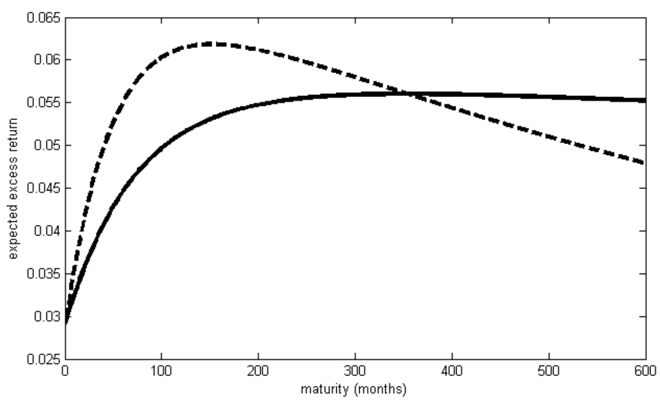

Figure 10. Expected excess returns on *cumulative equity strips* (solid line), defined as cumulative value-weighted means of expected excess returns on individual strips (dotted line).

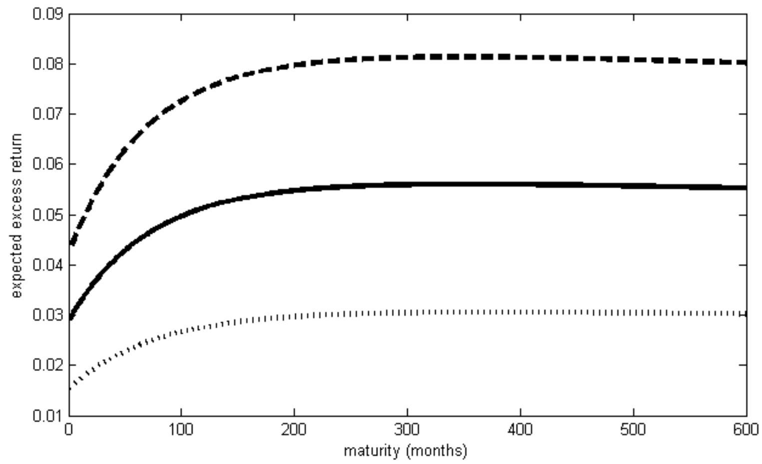

Figure 11. Expected excess returns on *cumulative equity strips*, conditional on σ_t^2 being equal to its long-run mean (solid line), and increased (decreased) by two unconditional standard deviations – dashed (dotted) line.

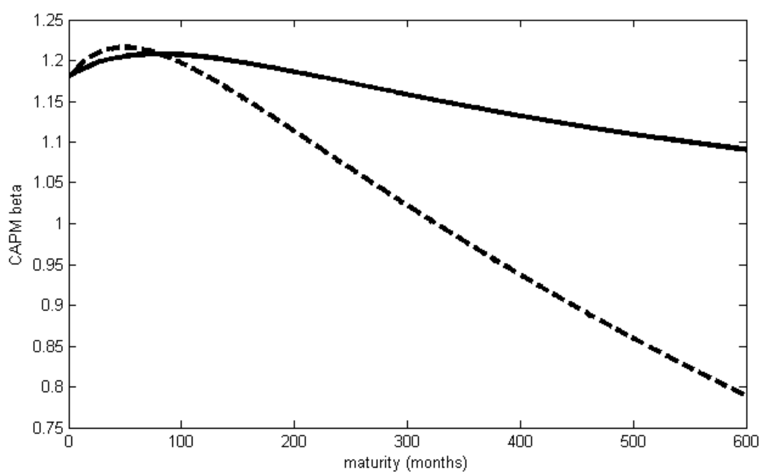

Figure 12. Conditional CAPM betas (with the value of σ_t^2 set to the long-run mean) of cumulative equity strips (solid line), and CAPM betas of single equity strips (dotted line).

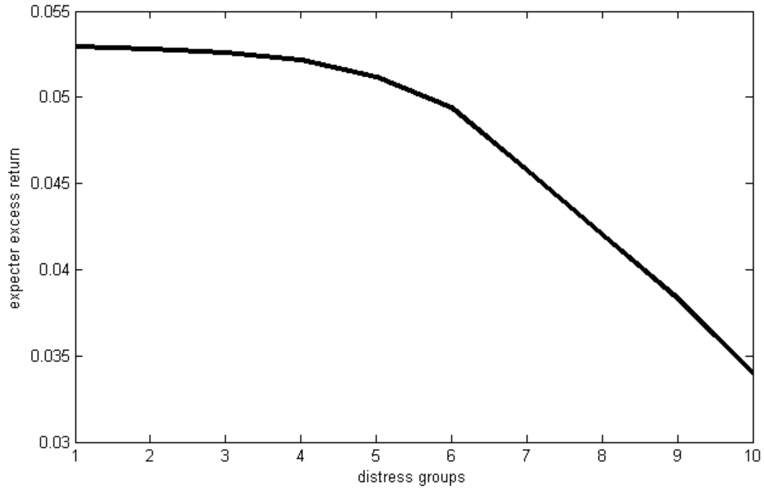

Figure 13. Model-implied expected excess returns on the ten portfolios of firms with probabilities of default as in Figure 8. (the most distressed on the right).

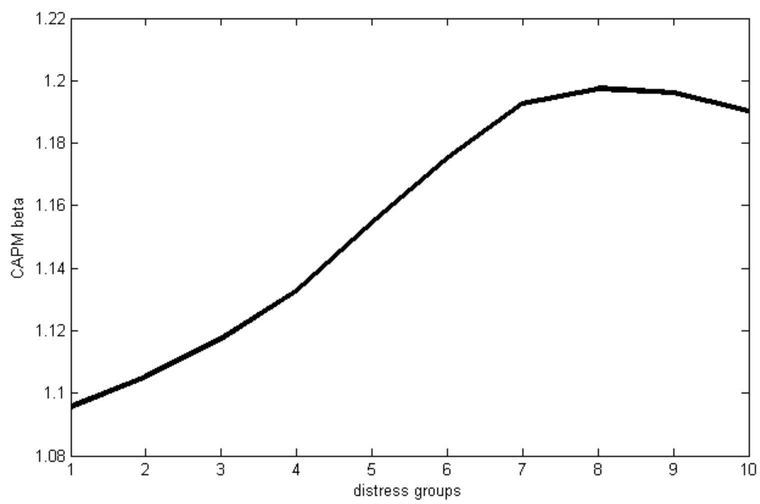

Figure 14. Model-implied market betas of the ten portfolios of firms with probabilities of default as in Figure 8. (the most distressed on the right).

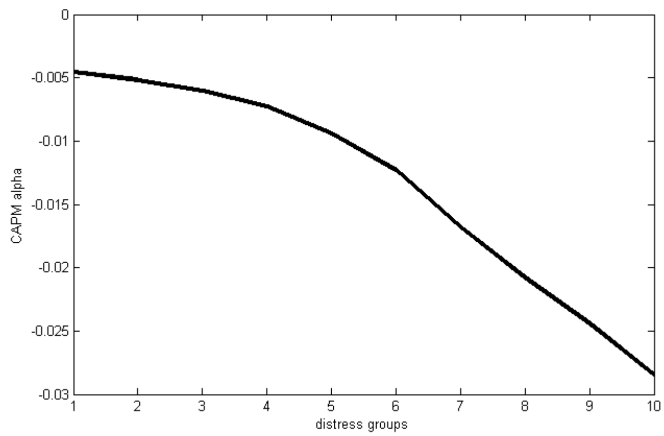

Figure 15. Model-implied CAPM alphas of the ten portfolios of firms with probabilities of default as in Figure 8. (the most distressed on the right).

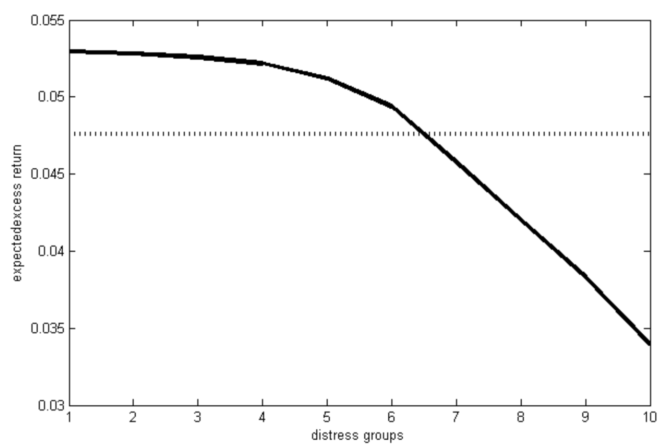

Figure 16. Alternative (see the main text) model-implied expected excess returns (solid line) on the ten portfolios of firms with probabilities of default as in Figure 8., plotted with alternative market equity premium (dotted line).

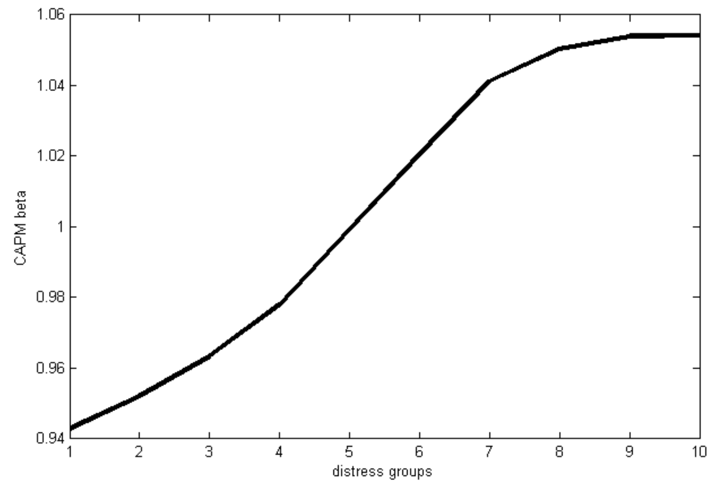

Figure 17. Alternatively computed (see the main text) model-implied market betas of the ten portfolios of firms with probabilities of default as in Figure 8.

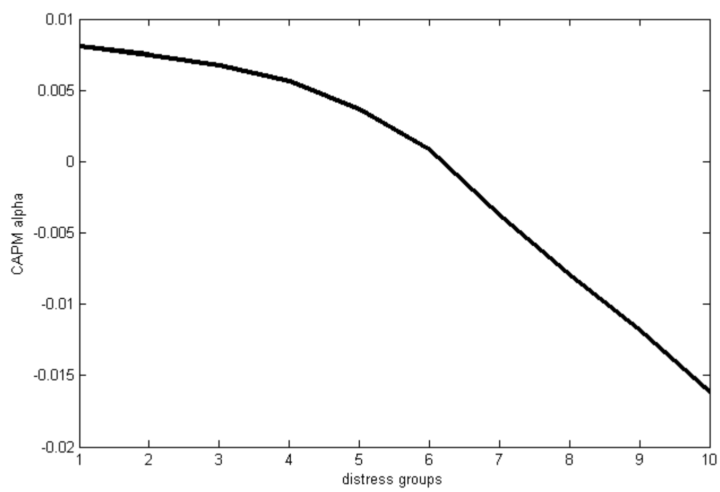

Figure 18. Alternatively computed (see the main text) model-implied CAPM alphas of the ten portfolios of firms with probabilities of default as in Figure 8.