On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle[1]

Long Chen[2] Pierre Collin-Dufresne[3] Robert S. Goldstein[4]

First Version: July 2003
This Version: March 17, 2006

[1]We thank participants at the Skinance 2005 Conference in Norway, the 2005 Wharton conference on 'Credit Risk and Asset Pricing', the BIS 2004 workshop on credit risk in Basel, the Moody's-KMV MAARC meeting, the AFA 2006 annual meeting, the San Francisco FED, and New York University for insightful comments. We are especially thankful to Monika Piazzesi and Pietro Veronesi for their comments. All remaining errors are our own.

[2]Assistant Professor, Department of Finance, Michigan State University, East Lansing, MI 48824, chen@bus.msu.edu

[3]Associate Professor of Finance at the Haas School of Business, University of California Berkeley, 545 Student Services Building #1900, Berkeley CA 94720-1900, dufresne@haas.berkeley.edu and NBER

[4]Associate Professor at the Carlson School of Management, University of Minnesota, Room 3-122, 321-19th Ave. South, rgoldstein@csom.umn.edu and NBER

On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle

Abstract

We examine whether 'large' historical credit spreads can be explained in the face of low historical default rates within a structural framework. For this to be the case, we show that the pricing kernel must covary strongly and negatively with asset prices – a characteristic which is also needed to explain the equity premium puzzle. As such, we explore whether those pricing kernels that have been successful at capturing historical equity returns (e.g., Campbell and Cochrane (CC 1999) and Bansal and Yaron (BY 2004)) can also explain the 'credit spread puzzle'. We find this to be the case if the risk premia are strongly time-varying and the default boundary is counter-cyclical. These properties are necessary because observed ratios of market volatility to total volatility make it difficult for structural models to generate large spreads. We also investigate the time-series implications of these models by backing out predicted year-by-year credit spreads from both models using macroeconomic data (e.g., historical consumption growth and price-dividend ratio). We find that the predicted credit spreads from CC model fit both the level and dynamics of historical credit spreads rather well.

1 Introduction

It is well-known that standard structural models of default predict counterfactually low credit spreads for corporate debt, especially for investment grade bonds of short maturity. Early work includes Jones, Mason and Rosenfeld (1984), who find that the Merton (1974) model generates yield spreads that fall far below empirical observation for investment grade firms. Although subsequent work (e.g., Eom, Helwege and Huang (2004)) has found that various structural models can generate very diverse predictions for credit spreads, Huang and Huang (HH 2003) demonstrate that once these various models are calibrated to be consistent with historical default and recovery rates, they all produce very similar credit spreads that fall well below historical averages. For example, HH report that the theoretical average 4-year (Baa-Treasury) spread is approximately 32 basis points (bp) and relatively stable across models. This contrasts sharply with their reported historical average (Baa-Treasury) spread of 158 bp. Similarly, HH find that the theoretical average 4-year Aaa-Treasury spread is about 1 bp, well below their reported historical average of 55 bp.

The typical 'explanation' for the large discrepancy between observed and theoretically predicted spreads is that these theoretical models only account for credit risk. That is, these models choose to ignore other factors that affect corporate bond prices, such as taxes, call/put/conversion options and the lack of liquidity in the corporate bond markets.[1] However, assuming that the component of the credit spread due to these issues is of similar magnitude for Aaa and Baa bonds, then the (Baa-Aaa) spread should be mostly due to credit risk.[2] Note, however, that the HH results reported above imply a predicted (Baa-Aaa) spread of (32 - 1) 31 bp, far short of the observed (158 - 55) 103 bp. As such, the findings of HH suggest that expected returns on a portfolio that is long Baa bonds and short Aaa bonds are rather large compared to the underlying risks involved. We refer to this result as the 'credit spread puzzle'.

We note that this 'credit spread puzzle' is reminiscent of the so-called equity premium puzzle in that the historical returns on equity also appear to be too high for the risks involved. Now, since corporate bonds and equities are both claims to the same firm value, they clearly share many of the same systematic risk sources. As such, it seems natural to ask whether these two puzzles are related. This question is the focus of our paper.

[1] Several papers have investigated the decomposition of spreads into various components. See, for example, Elton et al. (2001), Geske and Delianedis (2003), Driessen (2005) and Feldhutter and Land (2005)

[2] Admittedly, the call feature on Baa bonds may be more valuable than the call feature on Aaa bonds since, while the value of the call options on both will be increased by a market-wide drop in interest rates, the call option on the Baa bonds may also benefit by an increase in credit quality. We suspect that this difference is small, however.

To motivate our analysis, consider a defaultable discount bond that promises to pay one dollar at date-T. Its price, under some relatively weak no-arbitrage restrictions (see, e.g., Cochrane (2001) or Duffie (1996)), satisfies the following relation:

$$
\begin{aligned}
P &= \mathrm{E}\left[\Lambda\,(1-\mathbf{1}_{\{\tau\le T\}}L_\tau)\right] \\
&= \mathrm{E}\,[\Lambda]\ \mathrm{E}\left[1-\mathbf{1}_{\{\tau\le T\}}L_\tau\right] + \mathrm{Cov}\left[\Lambda\,,(1-\mathbf{1}_{\{\tau\le T\}}L_\tau)\right] \\
&= \frac{1}{R^f}\left(1-E\left[\mathbf{1}_{\{\tau\le T\}}L_\tau\right]\right) - \mathrm{Cov}\left[\Lambda\,,\mathbf{1}_{\{\tau\le T\}}L_\tau\right]. \qquad (1)
\end{aligned}
$$

Here, Λ is the pricing kernel, τ is the time of default, R_f is the risk free gross return, and L_τ is the loss given default. By calibrating expected default and recovery rates, HH force all models to agree on the expected future cash flows $E\left[1-\mathbf{1}_{\{\tau\le T\}}L_\tau\right]$ (i.e., the first term on the RHS). Hence, in order to predict lower prices for risky bonds (and thus higher spreads) consistent with the historical expected loss rate, a model must generate a strong

1) positive covariance between the pricing kernel (Λ_t) and the default time $\left(\mathbf{1}_{\{\tau\le T\}}\right)$.
2) positive covariance between the pricing kernel (Λ_t) and loss rates (L_τ).[3]

Within a structural framework of default, condition 1) can be broken down further into two components. In particular, structural models typically assume that default is triggered the first time an asset value process $\{V_t\}$ crosses a default boundary $\{B_t\}$ (which is typically related to the level of outstanding liabilities of the firm). Hence, the default time τ is defined as:

$$\tau := \inf\{t\ :\ V_t \le B_t\}.$$

Thus, in order for a structural model to generate lower bond prices (conditional on a given expected historical loss rate), it must generate a strong

1a) negative covariance between the pricing kernel (Λ_t) and asset prices (V_t),
1b) positive covariance between the pricing kernel (Λ_t) and the default boundary (B_t),
2) positive covariance between the pricing kernel (Λ_t) and loss rates (L_τ).

Interestingly, we note that channel 1a) is precisely the one researchers pursue in order to explain the equity premium puzzle. Motivated by this finding, below we investigate whether pricing kernels that have been engineered to explain the equity premium puzzle can also explain the credit spread puzzle. We focus on two such models: the habit formation model of Campbell and Cochrane (CC 1999) that focuses on time varying risk premia, and the model of Bansal and Yaron (BY 2004) that emphasizes long-run cash flow risk. We also explore what roles channels 1b) and 2) might play in capturing the credit spread puzzle.

Besides attempting to explain the 'credit spread puzzle', our exercise is meaningful for several other reasons. First, by linking credit spreads to the equity premium, we can provide a justification for the common practice of using credit spreads to estimate the equity premium (e.g., Chen, Roll, and Ross (1986), Keim and Stambaugh (1986), Campbell (1987), Fama and French (1989, 1993), Ammer and Campbell (1993), and Jagannathan and Wang (1996)). Second, our investigation may help discriminate between different explanations of the equity premium puzzle. That is, data on credit spreads can be seen as an out-of-sample test of the equity models of CC and BY. Third, while the equity premium is not directly observable, credit spreads are. As a result, while prior studies have focused on fitting the mean and volatility of the equity premium, here we generate model-implied credit spreads using macro variables (e.g., consumption growth) and compare them with actual spreads year by year. We find that, in addition to fitting average (Baa - Aaa) spreads very well, we also obtain excellent time series agreement between actual spreads and those predicted by the CC model calibrated to equity data.

Our main findings are as follows. First, none of the models can explain either the average level or the time-variation of the short maturity Aaa-Treasury spread. Simply put, the historical default frequencies are too low to be explained from a credit perspective. This result is consistent with interpreting both the level and the time variation of the (Aaa-Treasury) spread to be mostly non-default related.[4] Interestingly, since there is a strong positive correlation between the (Aaa-Treasury) spread and the (Baa-Aaa) spread, this result may suggest that (taking the credit spreads predictions at face value) liquidity, defined as the non-default component of spreads, moves with the business cycle.[5] If so, then during recessions, firms may need to issue bonds at yield spreads that are higher than fair compensation for credit risk. This in turn could justify why default boundaries are counter-cyclical.[6] We use this interpretation as one motivation for analyzing default boundaries that move with the business cycle.

Second, the CC model with a *constant* default boundary generates (Baa-Aaa) spreads that fit historical values better than the benchmark case, but it still falls well short of historical values. Further, this model predicts counterfactual pro-cyclical default probabilities. The in-

[4]Several papers have argued that the Treasury rate should not be the right 'risk-free rate' benchmark due to taxes and time-varying liquidity, e.g., Grinblatt (2000), Collin-Dufresne and Solnik (2001), He (2001), Longstaff (2003), Hull and White (2004)

[5]Of course, an alternative explanation is a 'Peso' problem in the bond market, i.e., the fact that the market accounts for the possibility of a so-far unobserved event where many investment grade firms would default jointly.

[6]There is empirical evidence supporting the fact that in downturns financing constraints tighten (e.g., Gertler and Gilchrist (1993), Kashyap Stein and Wilcox (1993)). This also has an impact on firms leverage decisions, e.g., Korajczyk and Levy (2003), Hennessy and Levy (2005).

tuition for this result is straightforward: in the CC model, the expected return is lowest in 'good times', implying that the probability of future default is higher than in bad times if the location of the default boundary is independent of the state of the economy. Interestingly, however, if we calibrate the model to match the historical relation between spreads and default rates by imposing counter-cyclical default boundary (i.e., channel 1b), then the model captures both the average level and volatility of (Baa-Aaa) spreads. We further show that the counter-cyclical default boundary cannot be interpreted as due to counter-cyclical leverage ratios in that, empirically, leverage ratios are not sufficiently counter-cyclical within the credit-refreshed rating groups. Interestingly, we note that the existence of a counter-cyclical default boundary predicts that those macroeconomic factors that covary with the default boundary should possess additional explanatory power for credit spreads even after controlling for all factors (e.g., leverage, firm value, volatility, etc.) suggested by standard structural models. This prediction is consistent with the empirical findings of Collin-Dufresne, Goldstein and Martin (2001), Elton et al. (2001), Cheyette et al. (2003), and Shaefer and Strebulaev (2004) who document that market wide (e.g., Fama-French factors, VIX) factors are economically and statistically significant for predicting changes in credit spreads even after controlling for these other variables.

Third, we investigate how well the BY model performs in explaining the credit spread puzzle. We note that in the BY model there are three interacting forces at work: (1) a persistent shock to cash flow growth rate; (2) a state variable driving stochastic volatility; and (3) a time-varying risk premium. The way we solve the BY model enables us to isolate the contributions of each separately. We find that both versions of the BY model that consider a *constant* risk premium are unable to explain much more of the credit spread than the simple 'benchmark model.' This suggest that time varying risk-premia are an essential feature for a pricing kernel to explain both equity returns and credit spreads. Although the time-varying risk premium model can explain significantly more (though still not all) of the observed (Baa - Aaa) spread, it appears this model cannot at the same time match the level and volatility of spreads and their covariation with future default rates, even if we were to allow for a counter-cyclical default boundary. We interpret these findings as implying that, as calibrated in the BY paper, their model does not generate a sufficiently strong time-varying Sharpe ratio in order to capture historical (Baa-Aaa) spreads.

Finally, we back out model-implied credit spreads using observable macro variables. We first show that the historical consumption surplus ratio - the key driver of equity premium in the CC model - provides a striking inverse image to the historical credit spreads for 1919-2004.

Subsequently, the simulated credit spread (backed out from consumption surplus ratio) fits the mean and variation of historical (Baa - Aaa) spread quite well. The simulated and actual credit spreads are 72% correlated for the whole sample period; their changes are 46% correlated for the 1919-1945 period and 58% for the 1946-2004 period.

The rest of the paper is as follows. In Section 2 we report historical data on the level and time variation of credit spreads, leverage and default probabilities. In Section 3 we first investigate a simple binomial model to provide a transparent framework for identifying why it is so difficult to capture historical credit spreads within a structural framework. We then demonstrate that the insights gleaned from this binomial setting hold within a Black-Cox (1976) framework, the results of which are used as a benchmark model to compare our results against. In Section 4 we review the pricing kernel of CC and present its implications for credit spreads. In Section 5 we present a continuous time version of the BY model and present its implications for credit spreads. In Section 6 we examine the long run relation between spreads and their predicted values from state variables backed out from consumption data. We conclude in Section 7. In the Appendix, we review some of the numerical predictions of the CC model.

2 Historical data, summary statistics, and benchmark

In this section, we report summary statistics related to macroeconomic variables and default risk. As reported in Panel A of Table 1, we find the price dividend (P/D) ratio to be 27.77 for the 1919-2001 period, and 23.40 for the 1919-1997 period.[7] Using the Moody's 2005 annual report, we find the average 4-year future cumulative default rate for Baa rated bonds to be 1.55% with a standard deviation of 1.04% for the 1970-2001 period. Using data from the Federal Reserve, we estimate the composite (Baa-Aaa) spread to be 1.09% with a standard deviation of 0.41% during this period. A longer dataset that includes the depression era provides similar results for the average (Baa-Aaa) spread but with significantly higher default rates. However, as in CC, who calibrate their model to postwar data (during which time the equity premium was significantly higher than in the longer dataset), we attempt to capture the statistics of this shorter data set for two reasons: First, current prices may reflect a belief that there is a better understanding of the economy so that it is unlikely that the US will ever again experience a depression with such severity. Second, some of the data used to calibrate the model only

[7]The data used is obtained from Shiller's website. Note that the price-dividend ratio does not consider equity repurchases. As such, our PD ratio is biased upward (See, e.g., Boudoukh, Michaely, Richardson, and Roberts (2004)).

go back to 1970. In particular, we match the regression coefficient of the four-year forward cumulative default rate on the (Baa-Aaa) spread, which yields a significant coefficient of 0.86.

We consider three different proxies for the leverage ratio. The first proxy is book leverage (BLV), calculated as the ratio of book debt (obtained from COMPUSTAT) to (book debt + market equity). The second proxy is market leverage (MLV), defined as the ratio of market debt to (market debt + market equity). Here, we estimate the market value of debt by first determining the market value of debt *per dollar of face value* for each firm-year (from the Lehman Brothers fixed income dataset), and then scaling this number by the book debt . The third proxy is the inverse distance to default (IDD), which is defined as the ratio of (0.5*long term book debt + short term book debt) to (market debt + market equity). This last measure is similar to that used by KMV in their implementation of the Black-Scholes-Merton model to estimate their expected default frequencies (EDF).

All measures cover the 1974-1998 period due to restrictions of the Lehman Brothers fixed income dataset. We only report the leverage ratios of Baa rated bonds. IDD is on average 28%, much lower than BLV (45%) and MLV (44%). We present the correlation matrix in Panel B. In addition to the above variables, we also include consumption growth rate, defined as the growth rate of real per capita consumption. The following patterns can be observed. First, the (Baa - Aaa) spread is counter-cyclical: it covaries negatively with both the P/D ratio and the consumption growth. In addition, the 4-year future default rate is significantly positively related to (Baa - Aaa) spread. Furthermore, the three leverage ratio measures appear to be counter-cyclical because they are significantly negatively related to the P/D ratio and positively related to the (Baa - Aaa) spread. Among the three measures, MLV is the least counter-cyclical. This is most likely due to the fact that the comovement of both market debt and equity partially offset each other. On the other hand, IDD is the most counter-cyclical, and this is mostly likely due to the fact that the comovements of the market values of both debt and equity, which are combined to determine the denominator, reinforce each other. We plot in Figure 1 the three leverage ratios of Baa rated bonds as well as (Baa - Aaa) spread for the 1975-1998 period. We refer to a particular year as a recession year if there are at least five months in that year that are defined as being in recession by NBER. It is clear that during the two recession periods the three leverage ratios go up (at least during the first half of the recession), reflecting the fact that market equity values go down more than debt values and/or firms are not cutting debt levels sufficiently fast to maintain a constant leverage throughout the business cycle.

Let's summarize some important properties:

Panel A:	Summary statistics			
Variable	Mean	Std.	Min	Max
P/D ratio	27.77	14.54	10.12	85.42
(Baa - Aaa) spread (%)	1.09	0.41	0.60	2.33
4-year default probability (%)	1.55	1.04	0.00	3.88
Book leverage of Baa	0.45	0.09	0.27	0.62
Market leverage of Baa	0.44	0.08	0.29	0.59
Inverse of the DD of Baa	0.28	0.07	0.16	0.42

Panel B:	Correlation matrix of some benchmark variables						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
P/D ratio (1)	1.00						
Consumption growth (2)	0.14	1.00					
	0.22						
(Baa - Aaa) spread (3)	-0.37	-0.32	1.00				
	0.00	0.00					
4-year default probability (4)	0.19	0.21	0.34	1.00			
	0.30	0.24	0.05				
Book leverage of Baa (5)	-0.70	-0.26	0.57	0.10	1.00		
	0.00	0.20	0.00	0.63			
Market leverage of Baa (6)	-0.61	-0.16	0.49	0.06	0.97	1.00	
	0.00	0.45	0.01	0.79	0.00		
Inverse of the DD of Baa (7)	-0.71	-0.40	0.60	0.14	0.96	0.87	1.00
	0.00	0.05	0.00	0.52	0.00	0.00	

Panel C:	Regressions of default probability on (Baa - Aaa) spread		
Dependent variable	Intercept	(Baa - Aaa)	adj. R-square
4-year default rate	0.57	0.86	8.80
(t-stat)	(1.01)	(2.07)	

Table 1: Summary statistics. The statistics of different variables cover different periods in Panel A. The P/D ratio covers the 1919-2001 period. The 4-year ahead cumulative default probability and the (Baa - Aaa) spread cover the 1970-2001 period. The three leverage measures cover the 1974-1998 period. Among them, book leverage is defined as the ratio of book debt to (book debt + market equity); market leverage is defined as the ratio of market debt to (market debt + market equity); the inverse of the distance to default (DD) is defined as the ratio of (0.5*long term book debt + short term book debt) to (market debt + market equity). In panel B, the first (second) row is the correlation (p-value). The correlation statistics use the maximum common sample size between two series. In Panel C the first row is the OLS regression coefficients. On the second row Newey-West t-statistics are reported, where 4 lags are chosen for 4-year default probability.

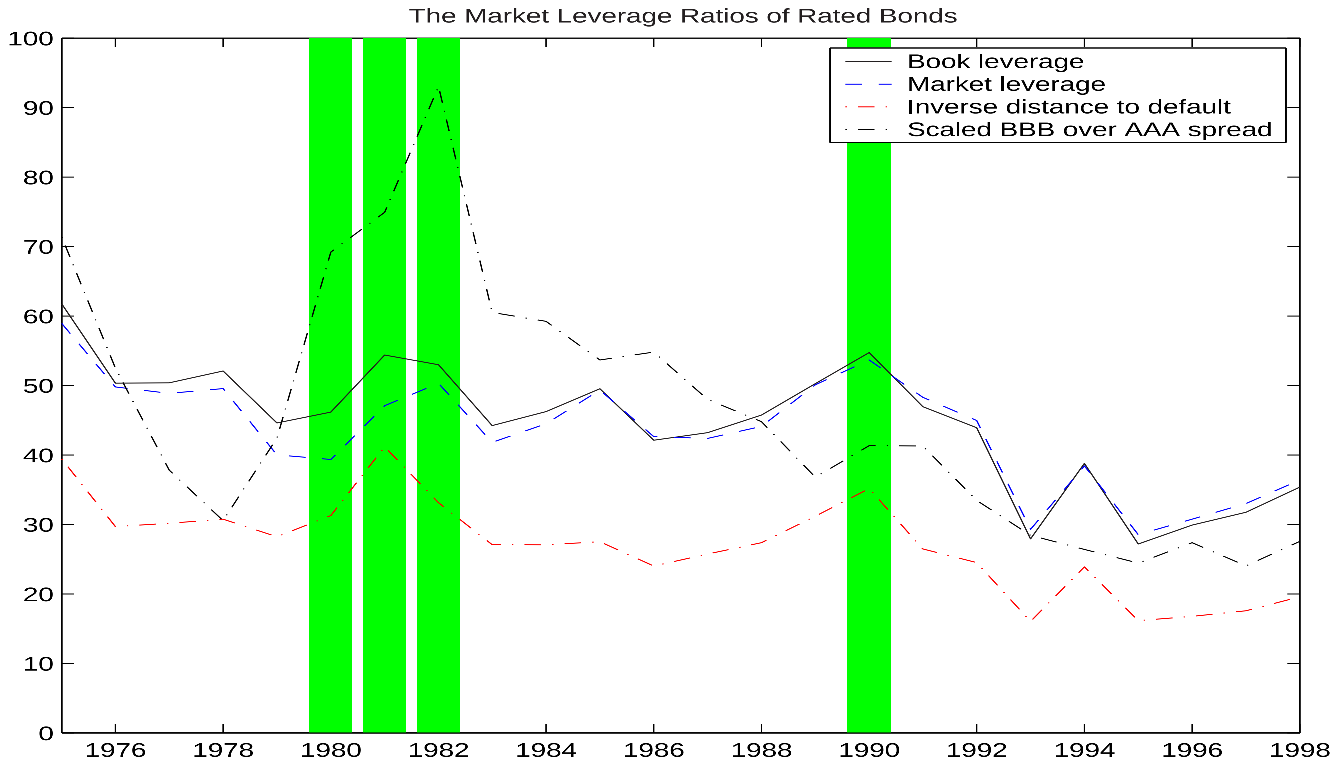

Figure 1: Time series of leverage for Baa rated firms.

- (Baa - Aaa) spreads are high on average (109 bp) and rather volatile (41 bp standard deviation).
- Baa Default rates are low on average (1.55 percent four-year cumulative default probabilities) and volatile.
- Forward default rates are counter-cyclical in that the regression coefficient of forward default rates on spreads is 0.86 and statistically significant.
- Leverage ratios are counter-cyclical (both in terms of P/D ratios and consumption growth) and positively related to credit spreads.

3 Identifying the Causes of the Credit Spread Puzzle

In this section, we first investigate a simple binomial framework in order to identify the reasons why it is so difficult for standard structural models to explain historical credit spreads. We then demonstrate that the insights gleaned from this binomial example hold in a more formal Black-Cox (1976) framework.

Consider a zero-coupon bond that pays at date-T either \$1 if no default has occurred, or $\$F$ if default has occurred. The probability of default is π. If we define $X(T)$ as the random payoff, it follows that the expected payoff is

$$\begin{aligned} \mathrm{E}_0\left[X(T)\right] &= (\pi)\,(F) + (1-\pi)\,(1) \\ &= 1 - \pi\,(1-F)\,. \end{aligned} \tag{2}$$

The payoff variance is

$$\mathrm{Var}_0\left[X(T)\right] = (1-F)^2\left(\pi - \pi^2\right). \tag{3}$$

Since we are considering small values of π, the standard deviation is approximately

$$\sigma\left[X(T)\right] \approx (1-F)\sqrt{\pi}. \tag{4}$$

Now, we specify the dynamics of the pricing kernel Λ with constant interest rates r and constant volatility θ.

$$\frac{d\Lambda}{\Lambda} = -r\,dt - \theta\,dz_{\Lambda}\,. \tag{5}$$

It is well-known that the price of the risky bond can be written as

$$B^X = \mathrm{E}\left[\Lambda(T)\,X(T)\right].$$

Thus, defining the yield to maturity y through $B^X = e^{-yT}$, we get

$$\begin{aligned} e^{-yT} &= e^{-rT}\,E[X] + \sigma_{\Lambda}\sigma_X\,\rho_{\Lambda,X} \\ &= e^{-rT}\,\left[1 - \pi(1-F)\right] + \sigma_{\Lambda}\left[(1-F)\sqrt{\pi}\right]\rho_{\Lambda,X}\,. \end{aligned} \tag{6}$$

Below we will specify that the returns of individual firms have two sources of risk: i) market risk, and ii) idiosyncratic risk that is uncorrelated with other sources of risk in the economy.[8] As such, the correlation between the cash flows of this bond and the pricing kernel can be written as the product

$$\rho_{\Lambda,X} = \rho_{\Lambda,V}\,\rho_{V,X}\,. \tag{7}$$

Here, the subscript V refers to the market portfolio.

[8] We are also implicitly assuming that the idiosyncratic volatility is either a constant, or has dynamics that are driven by idiosyncratic risk.

Further, we assume that the stock market is integrated with the bond market. As such, we can estimate the volatility of the pricing kernel from the instantaneous Sharpe ratio of stocks. In particular, if we specify aggregate stock returns as

$$\frac{dV}{V} = \mu_V \, dt + \sigma_V \, dz, \tag{8}$$

we find

$$\begin{aligned} \sigma_\Lambda &\equiv \left(\mathrm{Var}\left[\Lambda(T)\right] \right)^{\frac{1}{2}} \\ &= e^{-rT} \left[e^{\theta^2 T} - 1 \right]^{\frac{1}{2}} \\ &\approx e^{-rT} \theta \sqrt{T}. \end{aligned} \tag{9}$$

Further, defining the instantaneous Sharpe ratio $\left(\kappa \equiv \frac{\mu_V - r}{\sigma_V}\right)$ and using $V(0) = \mathrm{E}_0\left[V(T)\,\Lambda(T)\right]$, we find

$$\kappa = -\theta \, \rho_{\Lambda,V}. \tag{10}$$

Together, equations (9) and (10) imply

$$\sigma_\Lambda \, \rho_{\Lambda,V} = -\kappa \sqrt{T} \, e^{-rT}. \tag{11}$$

Finally, using equations (7) and (11) we can write equation (6) as

$$e^{-yT} = e^{-rT} \left[1 - \pi(1-F)\right] - e^{-rT} \kappa \sqrt{T} \left[(1-F)\sqrt{\pi}\right] \rho_{V,X}, \tag{12}$$

or equivalently

$$(y - r) = -\left(\frac{1}{T}\right) \log\left(1 - \pi(1-F) - \kappa(1-F)\sqrt{\pi\, T} \rho_{V,X}\right) \tag{13}$$

Since default rates π are low, this can be approximated as

$$(y - r) \approx \left(\frac{1-F}{T}\right)\pi + \left(\frac{1-F}{T}\right)\sqrt{\pi\, T}\, \kappa \, \rho_{V,X}. \tag{14}$$

The first term can be interpreted as yield spread due to expected losses, and the second term as yield spread due to risk premia.

To calibrate this model, we first write $\rho_{V,X} = \rho_{V,P}\, \rho_{P,X}$, where P denotes the returns on the asset value that the risky bond is written on. We assume that the average stock has a beta of 1, the volatility of the market is .16, and the average volatility of the stock is .32. As such, $\rho_{V,P} = \frac{1}{2}$. We also approximate $\rho_{P,X} \approx 0.3$.[9] Consistent with CC, we choose a

[9] To motivate the estimate $\rho_{P,X} \approx 0.3$, assume that P is a normally distributed $N(0,1)$ variable, and that x pays F if P is in the 'default range', and \$1 otherwise.

Sharpe ratio of $\kappa \approx 0.43$. Finally, using the Moody's 2005 default report, we set the recovery rate to $F = 0.449$, the Baa-default probability rate to $\pi_{Baa} = 0.0155$ and the Aaa-default probability rate to $\pi_{Aaa} = 0.0004$. Using these parameters in equation (14), we estimate the (Baa - Treasury) spread to be approximately 43bp, and the (Aaa - Treasury) spread to be approximately 4bp, broadly consistent with the findings of both HH and our benchmark case below.

Inspection of equation (14) suggests that, taking expected default rates as given, one way to increase credit spreads is to increase the correlation between a given corporate bond's cash flows and the aggregate stock return.[10] However, this is not so simple to do within a traditional structural framework since they predict that $\rho_{P,X}$ is determined mechanically and $\rho_{V,P}$ is set by the ratio of a firm's 'market volatility' to its 'total volatility', and this ratio is easily measured empirically. One way to circumvent this restriction is to assume a counter-cyclical default boundary – that is, to use channel 1b) discussed above. To see this, note that the asset value dynamics for a $\beta = 1$ firm is assumed to follow

$$\begin{aligned} \frac{dP}{P} &= \frac{dV}{V} + \sigma_{idio} dz_{idio} \\ &= \mu_V \, dt + \sigma_V \, dz_V + \sigma_{idio} dz_{idio}. \end{aligned} \tag{15}$$

Once again, empirical estimates places the ratio of market volatility to total volatility at approximately

$$\begin{aligned} \frac{\sigma_V}{\sigma_{Tot}} &= \frac{\sigma_V}{\sqrt{\sigma_V^2 + \sigma_{idio}^2}} \\ &\approx \frac{1}{2}. \end{aligned} \tag{16}$$

However, what is crucial is not the dynamics of firm value *per-se*, but rather the dynamics of the so-called 'distance-to-default' $P^* \equiv \frac{P}{P_B}$. Now, if we assume that the default boundary is counter-cyclical:

$$\frac{dP_B}{P_B} \sim -slope \, dz_V, \tag{17}$$

then the distance to default dynamics become

$$\frac{dP^*}{P^*} \sim (\cdot) \, dt + (\sigma_V + slope) \, dz_V + \sigma_{idio} dz_{idio}, \tag{18}$$

which will increase the correlation ρ_{VX} since the ratio of 'effective market volatility' σ_V^{eff} to

[10] We note that another channel that can be used to increase spreads is to assume that the recovery rate F is not a constant but rather is countercyclical. We also examine this channel below.

total volatility increases from $\frac{1}{2}$ to

$$\frac{\sigma_V^{eff}}{\sigma_{Tot}^{eff}} = \frac{\sigma_V + slope}{\sqrt{(\sigma_V + slope)^2 + \sigma_{idio}^2}}. \tag{19}$$

We also note that the calibration above assumed a constant instantaneous Sharpe ratio θ. Below, we will argue that time-varying Sharpe ratios can also help explain observed credit spreads. Intuitively, a highly skewed pricing kernel implies that the prices of certain Arrow-Debreu securities are very expensive. For example, CC specify a pricing kernel that explains the high equity premium during recessions, arguing that the representative agent is not so risk averse *per-se*, but rather extremely risk-averse to recessions. Now, we note in practice that most corporate defaults occur during recessions. Hence, a portfolio that is long Treasuries and short a well-diversified portfolio of corporate bonds will pay almost zero in good times, but handsomely in bad times. That is, such a portfolio is long the expensive A/D securities, and hence is quite expensive. This, in turn implies large spreads.

In summary, then, we can increase spreads within a structural framework by i) considering a pricing kernel that has strongly time-varying Sharpe ratios, and ii) imposing a countercyclical default boundary. Interestingly, we demonstrate below that the CC model can not only support such a stochastic default boundary, but in fact it *requires* such a boundary in order to avoid making the counterfactual prediction that forward default rates are pro-cyclical. In contrast, even the constant boundary in the BY model already overshoots the relation between credit spreads and future default rates, and thus imposing a countercyclical default boundary would only make situations worse. We believe this occurs because the CC model as calibrated generates a much more time varying Sharpe ratio over the business cycle than does the BY model.

3.1 A benchmark model

Here we investigate whether the approximations made in the binomial model above hold in a more rigorous setting. To maintain tractability, we investigate a Black and Cox (1976) economy. In particular, we assume that the underlying firm asset value has the following return dynamics:

$$\frac{dP(t)}{P(t)} = (r - \delta)\, dt + \sigma_P \left[\rho_{PM} dz_M^Q(t) + \sqrt{1-\rho_{PM}^2}\, dz_i^Q(t)\right] \tag{20}$$

$$= (r - \delta + \kappa\sigma_P\rho_{PM})\, dt + \sigma_P \left[\rho_{PM} dz_M(t) + \sqrt{1-\rho_{PM}^2}\, dz_i(t)\right]. \tag{21}$$

We assume default is triggered first time that $P(t)$ reaches the default boundary P_B. At bankruptcy, bond holders receive a fraction of the face value of the bond.

For the benchmark case, we set the parameters to their historical counterparts: $r = 0.04$, $\delta = 0.05$, $\sigma_P = 0.2$ (recall, this is an asset volatility), $\kappa = 0.43$ (consistent with CC). We determine the default boundary, P_B, so that expected default rates match historical default rates, namely, $\pi_4^{Baa} = 0.0155$, $\pi_{10}^{Baa} = 0.0489$, $\pi_4^{Aaa} = 0.0004$ and $\pi_{10}^{Aaa} = 0.0063$. This is in the spirit of the calibration of HH (though HH choose to fix the default boundary exogenously and calibrate the volatility to match default rates). We choose to match historical default rates by choosing the default boundary rather than volatility, because the latter is easier measured than the former (e.g., Davydenko (2005)) and because, as illustrated with the previous example, the ratio of idiosyncratic to market volatility is an important input of the spread puzzle. To investigate, the role of systematic versus idiosyncratic risk we consider two cases: $\rho_{PM} = 1$ and $\rho_{PM} = 0.5$.

In the event of bankruptcy, we assume bond holders recover a fraction of the face value of the bond. The recovery rate is set equal to 44.9% to match the average reported in Moody's (2005) report.

To investigate the sensitivity to the recovery assumption we consider two types of coupon/default payments. The first assumes coupon payments paid semiannually equal to 100bp above the risk free rate, and a recovery rate of 0.449 paid at the default event. The second assumes zero coupon payments and a recovery rate of 0.449 at *maturity*, or equivalently, a recovery rate of $0.449 * e^{-r(T-\tau)}$ at the default date.

To estimate P_B, it is convenient to define $v \equiv \log\left(\frac{P}{P_B}\right)$. Note that, by definition, default occurs the first time $v = 0$. Using Ito's lemma, we find

$$\begin{aligned} dv &= \left(r - \delta - \frac{1}{2}\sigma_P^2\right) dt + \sigma_P \left[\rho_{PM} dz_M^Q(t) + \sqrt{1-\rho_{PM}^2}\, dz_i^Q(t)\right] \\ &\equiv \mu_P\, dt + \sigma_P \left[\rho_{PM} dz_M^Q(t) + \sqrt{1-\rho_{PM}^2}\, dz_i^Q(t)\right]. \end{aligned} \tag{22}$$

Using well-known results, the P-probability that default occurs before maturity T is

$$\pi_0^P\left[\tilde{\tau} > T\right] = \mathrm{N}\left[\frac{v(0) + \mu_P T}{\sqrt{\sigma_P^2 T}}\right] - e^{-\frac{2v(0)\mu_P}{\sigma_P^2}}\, \mathrm{N}\left[\frac{-v(0) + \mu_P T}{\sqrt{\sigma_P^2 T}}\right] \tag{23}$$

Setting the LHS of this equation to historical values, we use this to determine $v(0)$, and hence $P_B = P(0)e^{-v(0)}$ (since we set $P(0) = 1$ without loss of generality). Using a similar formula for the risk-neutral default probability we can price the risky coupon (CB) and zero-coupon (zero) bond in closed-form, and determine the credit spreads (BBB-Treasury), (Aaa - Treasury), and hence (Baa - Aaa). There are 8 cases overall, depending upon 4-year vs. 10-year, $\rho_{MP} = 0.5$, 1, and coupon/early payment (CB) vs. no coupon/payment (zero) at maturity.

Results are presented in the following table.

Benchmark Results

Baa Spread ($\pi_4^{Baa} = 0.0155$ and $\pi_{10}^{Baa} = 0.0489$)

	4 year maturity			10 year maturity		
	P_B	zero	CB	P_B	zero	CB
$\rho = 1$	0.459	126.68	122.96	0.441	205.22	199.81
$\rho = 0.5$	0.397	56.58	54.38	0.320	89.74	82.71

Aaa Spread ($\pi_4^{Aaa} = 0.0004$ and $\pi_{10}^{Aaa} = 0.0063$)

	4 year maturity			10 year maturity		
	P_B	zero	CB	P_B	zero	CB
$\rho = 1$	0.298	8.18	7.76	0.287	64.52	58.63
$\rho = 0.5$	0.255	2.32	2.19	0.199	18.4	16.13

We find that ignoring idiosyncratic risk has a very large impact on spreads (up to 110bp for 10 year Baa bonds). Also, we find that ignoring coupon payments has a small impact on spreads, but that this impact is more pronounced for longer maturity bonds. (2bp - 10bp). The main message of this table is that the credit spread puzzle is closely tied to the ratio of idiosyncratic to total volatility. If all of the idiosyncratic volatility were systematic risk, then there would be no puzzle (or if anything the puzzle would be that average spreads are too low!).

We next proceed to a few robustness checks.

3.2 The 'Convexity Effect'

David (2006) argues that the calibration of HH suffers from a very large 'convexity effect' due to the time variation in leverage ratios. In particular, he argues that using the average leverage ratio to evaluate spreads leads to a very large underestimation of credit spreads. Here we argue that this 'convexity effect' is small and in fact goes in the *opposite direction* than documented in David (2006). The source this discrepancy is the following: David (2006) first calibrates the location of his default boundary to match historical default rates, and then obtains an average credit spread. Then, while *maintaining the same boundary*, he determines the credit spread for a firm starting at the average leverage ratio. We argue, however, that this is not the correct way to estimate the 'convexity effect' because, as calibrated, if all firms start at the average leverage ratio, then the expected default rate is *significantly lower* than the historical average. Instead, we argue that the proper way of estimating the convexity effect is to recalibrate the default boundary so that, assuming all firms start at the average leverage ratio, expected default rates match historical default rates.

Here, we quantify this convexity effect by performing the following experiment. We assume that one-half of all firms have an initial leverage ratio of $(0.4328+\epsilon)$, and one-half of all firms have an initial leverage ratio of $(0.4328-\epsilon)$, where $\epsilon = 0.09$ is chosen to capture the standard deviation of leverage given in Table 1. We then assume that default occurs at $P_B^+ = \beta(0.4328+\epsilon)$ for the high-leveraged firms and at $P_B^- = \beta(0.4328-\epsilon)$ for the low-leveraged firms. β is endogenously chosen so that expected default rates match historical ones. That is, β is chosen so that, for the 4-year, Baa case:

$$\begin{aligned}0.0155 &= \frac{1}{2}\left[\mathrm{N}\left[\frac{-\log(P_B^+)+\mu_P T}{\sqrt{\sigma_P^2 T}}\right] - e^{-\frac{2v(0)\mu_P}{\sigma_P^2}}\mathrm{N}\left[\frac{\log(P_B^+)+\mu_P T}{\sqrt{\sigma_P^2 T}}\right]\right] \\ &\quad +\frac{1}{2}\left[\mathrm{N}\left[\frac{-\log(P_B^-)+\mu_P T}{\sqrt{\sigma_P^2 T}}\right] - e^{-\frac{2v(0)\mu_P}{\sigma_P^2}}\mathrm{N}\left[\frac{\log(P_B^-)+\mu_P T}{\sqrt{\sigma_P^2 T}}\right]\right].\end{aligned}$$

We then estimate the spread and we report the results in the following table for the Baa zero-coupon four-year maturity spreads. We show the results for both cases $\rho = 1$ and $\rho = 0.5$.

'Convexity effect'

Convexity effect for Baa zero Spread 4-year mat					
	E[CS]	σ[CS]	CS(lev = $\overline{lev}$)	def rate(lev = $\overline{lev}$)	recalibrated E[CS(lev = $\overline{lev}$)]
$\rho = 1$	109.73	92.42	75.43	0.76%	126.68
$\rho = 0.5$	52.49	46.32	32.21	0.80%	56.58

What David (2006) refers to as the 'convexity effect' would equal be (109.7 - 75.4 = 34.3bp) for the $\rho = 1$ case. Note, however, that the expected default rate (.76%) is only about one-half the historical rate (1.55%).

Instead, taking the results from our base case on the previous page, we actually see that there is a slight concavity effect of (109.7 - 126.7 = -17bp) for the $\rho = 1$ case and (52.5 - 56.6 = -4.1bp) for the $\rho = 0.5$ case.

The intuition for why there is a concavity effect can be understood by noting that most of the Baa credit spread is due to risk premia, not expected losses. If there is a significant dispersion in leverage ratios, then most of the defaults will be due to those firms with high initial leverage ratios. However, for these firms to default, the market portfolio does not have to perform so badly. Hence, such defaults are more idiosyncratic, and hence do not deserve as much compensation in terms of a high spread.

We note, however, that it is likely that for a rating agency to give a high-leverage firm the same rating as a low-leverage firm, then on average we can expect the high-leverage firm to have lower asset volatility levels, and vice-versa. This will reduce any concavity effect even further. As such, below we follow HH and investigate spreads using only an average initial leverage ratio.

In the next two sections, we investigate how well pricing kernels that have been engineered to match historical equity premium fare in predicting credit spreads.

4 The CC Habit Formation Model

Slightly modifying their notation, Campbell and Cochrane (1999) specify the utility function of the representative agent in an exchange economy as

$$U(C_t, \widehat{C}_t, t) = e^{-\alpha t} \frac{\left(C - \widehat{C}\right)^{1-\gamma} - 1}{1-\gamma}, \tag{24}$$

where $\widehat{C}$ is an exogenous habit. CC define the surplus consumption ratio as $S \equiv \left(\frac{C-\widehat{C}}{C}\right)$, and for convenience, also define $s \equiv \log S$, $c \equiv \log C$. Since there are no investment opportunities, and since the dividend is perishable, it follows that in equilibrium consumption equals the dividend payment. Further, the pricing kernel is equal to the marginal utility of the representative agent:

$$\begin{aligned} \Lambda_t &= U_C(C_t, \widehat{C}_t, t) \\ &= e^{-\alpha t} \left(C - \widehat{C}\right)^{-\gamma} \end{aligned}$$

$$= \quad e^{-\alpha t}\, e^{-\gamma\, s}\, e^{-\gamma\, c}. \tag{25}$$

CC specify the log-consumption and log-dividend processes as

$$\Delta c \quad = \quad g_c\, \Delta t + \sigma_c\, \Delta z_c \tag{26}$$

$$\Delta d \quad = \quad g_d\, \Delta t + \sigma_d\, \left(\rho_{cd}\, \Delta z_c + \sqrt{1 - \rho_{cd}^2}\, \Delta z_d \right). \tag{27}$$

Finally, CC specify the log surplus consumption ratio dynamics as[11]

$$\Delta s \quad = \quad \begin{cases} \kappa(\overline{s} - s)\Delta t + \sigma \left[\frac{1}{\overline{S}} \sqrt{1 - 2(s - \overline{s})} - 1 \right] \Delta z & \text{for} \quad s \leq s_{max} \\ \\ \kappa(\overline{s} - s)\Delta t & \text{for} \quad s > s_{max}, \end{cases}$$

where

$$\overline{S} \quad \equiv \quad \sigma \sqrt{\frac{\gamma}{\kappa}} \tag{28}$$

$$s_{max} \quad \equiv \quad \overline{s} + \frac{1}{2} \left(1 - \overline{S}^2 \right). \tag{29}$$

This specification generates an economy with a constant real risk free rate (for $s < s_{max}$):

$$r_f = \alpha + \gamma g_c - \frac{1}{2} \gamma \kappa. \tag{30}$$

The price-consumption ratio for the claim to consumption can be written as

$$\left(\frac{P(t)}{C(t)} \right) \quad = \quad \mathrm{E}_t \left[\frac{\Lambda(t+1)}{\Lambda(t)} \frac{C(t+1)}{C(t)} \left(1 + \frac{P(t+1)}{C(t+1)} \right) \right] \tag{31}$$

$$= \quad \mathrm{E}_t \left[\sum_{j=1}^{\infty} \frac{\Lambda(t+j)}{\Lambda(t)} \frac{C(t+j)}{C(t)} \right]. \tag{32}$$

An analogous formula holds for the price-dividend ratio. While their framework does not provide analytic solutions for the price-consumption ratio, equations (31) and (32) suggest two numerical schemes for estimating this ratio. In particular, equation (31) can be estimated by using a recursive scheme to obtain a self-consistent solution for $\frac{P}{C}$. Alternatively, equation (32) can be estimated using Monte-Carlo methods. Unfortunately, both methods are vulnerable to certain types of errors, as discussed in the Appendix.[12] Indeed, there we demonstrate that their estimated price-consumption ratio (which generates all of their later results) differs significantly from our estimate.

[11] We use the parameter κ instead of $(1 - \phi)$ because κ, which has units of inverse-time, can be easily 'annualized' if first measured using a different frequency. In contrast, annualizing ϕ is more involved.

[12] A third approach based on the continuous time version of the model, and which relies on solving a partial differential equation, is discussed in the appendix as well.

4.1 Estimating Credit spreads in the CC Framework

Following CC, we calibrate the consumption dynamics $g_c = 0.0189$ and $\sigma_c = 0.015$ to match their historical averages. Further, the historical average real risk free rate $r_f = 0.0094$ is used to calibrate $\alpha = 0.133$ via equation (30). Finally, $\kappa = 0.138$ is chosen to match the serial correlation of the log price-dividend ratio. We then choose g_d, σ_d, ρ_{cd} and γ to best match historical data on equity. The higher growth rate on dividends compared to consumption captures the leveraged nature of equity (Abel (1999, 2005), Goldstein (2006)). The results are given in Table 2. We see that the model does a good job at capturing historical levels and volatilities of both the price dividend ratio and excess returns, as well as the historical Sharpe ratio.

We note that structural models of default take the firm value process (i.e., the claim to dividends and interest payments) as the fundamental state variable, and not the equity value process (i.e., the claim to dividends), which instead is determined (along with the debt claim) endogenously within a structural model . As such, we define the firm's 'output' as the sum of payments made to dividends plus interest, and then specify the log aggregate output process $o(t) = \log O_t$ as

$$\Delta o = g_o \, \Delta t + \sigma_o \, \left(\rho_{co} \, \Delta z_c + \sqrt{1 - \rho_{co}^2} \, \Delta z_o \right). \tag{33}$$

With γ determined from the equity data,[13] and $g_o = 0.0189$ chosen to match the consumption growth rate[14] we choose σ_o and ρ_{co} to best match historical moments. These results are also given in Table 2. Historical values were estimated assuming historical weighted averages of debt and equity returns, where the weights came from historical leverage ratios.

With this calibration in place, we now estimate credit spreads. First, we determine the aggregate price-output ratio $PO(s_t)$ as a function of the lone state variable s_t using the method described in the appendix. A plot of the price-output ratio as a function of s_t is given in Figure (4.1). We then determine aggregate firm value $V(O_t, s_t)$ by noting that price equals output times the price/output ratio:

$$V(O_t, s_t) = O_t \, PO(s_t). \tag{34}$$

Given the dynamics of aggregate output O_t in equation (33) and the estimated functional form for the price-output ratio $PO(s_t)$, it is straightforward to demonstrate that the dynamics of

[13]We choose γ to best match equity data, since this is the most easily estimated and most studied. Note that the other parameters of the dividend process are not used for the analysis below.

[14]Here, we are thinking of the claim to output as a non-leveraged security, and hence should have a growth rate equal to that of consumption.

Panel A:	Parameter Inputs			
CF type	g_d	σ_d	γ	ρ_{cd}
Dividends	.040	.080	2.45	.60
Output	.0189	.063	2.45	.48

Panel B:	Model Outputs:				
CF type	$\exp(\mathrm{E}[p-d])$	$\sigma(p-d)$	Sharpe	$\mathrm{E}[r-r_f]$	$\sigma(r-r_f)$
Claim to Dividends	24	.21	.44	.073	.17
Historical equity	25	.26	.43	.067	.16
Claim to output	23	.15	.44	.053	.12
Historical (debt + equity)	19	.20	.43	.050	.10

Table 2: Panel A: Parameter Calibrations for dividend and output processes. Panel B: Sample Moments of claims to dividends vs. historical values; and claims to (dividends plus interest) vs. historical values.

aggregate firm value under both the P and Q measures take the forms

$$\frac{\Delta V(t)}{V(t)} = \Big(\theta(s_t) + r - \delta(s_t)\Big)\,\Delta t + \sigma(s_t)\,\Delta z_V(t) \tag{35}$$

$$\frac{\Delta V(t)}{V(t)} = \Big(r - \delta(s_t)\Big)\,\Delta t + \sigma(s_t)\,\Delta z_V^Q(t). \tag{36}$$

Here, the risk-premium $\theta(s_t)$, the dividend yield (which equals the inverse price-output ratio) $\delta(s_t)$, and volatility $\sigma(s_t)$ are all functions of s_t and independent of O_t. That is, as noted by CC, $s(t)$ is the only state variable driving asset return dynamics.

In the spirit of, for example, a CAPM framework, we then assume that the return dynamics for a typical firm follows[15]

$$\frac{\Delta P(t)}{P(t)} = \frac{\Delta V(t)}{V(t)} + \sigma_{idio}\,\Delta z_{idio}(t). \tag{37}$$

Without loss of generality, we set initial firm value $P(0) = 1$.

4.1.1 Constant default boundary case

Following HH, we set the default boundaries to be a constant. In particular, for Baa firms, we choose $P_{def}^{Baa} = (0.6)(0.4328) \approx 0.26$. The 0.4328 comes from the average leverage ratio for Baa used by HH, and the (.6) accounts for the fact that firm value can drop well below

[15] We fully note that this model assumes that idiosyncratic risk is independent across firms and that all firms load on a single 'market factor'. A more general model could permit, e.g., industries to have correlated idiosyncratic risks. For example, in the CC model we could model idiosyncratic volatility as a function of the surplus variable, i.e., $\sigma_{idio}(s_t)$. This might be consistent with the recent evidence in Campbell and Taksler (2005). We save this interesting question for future work.

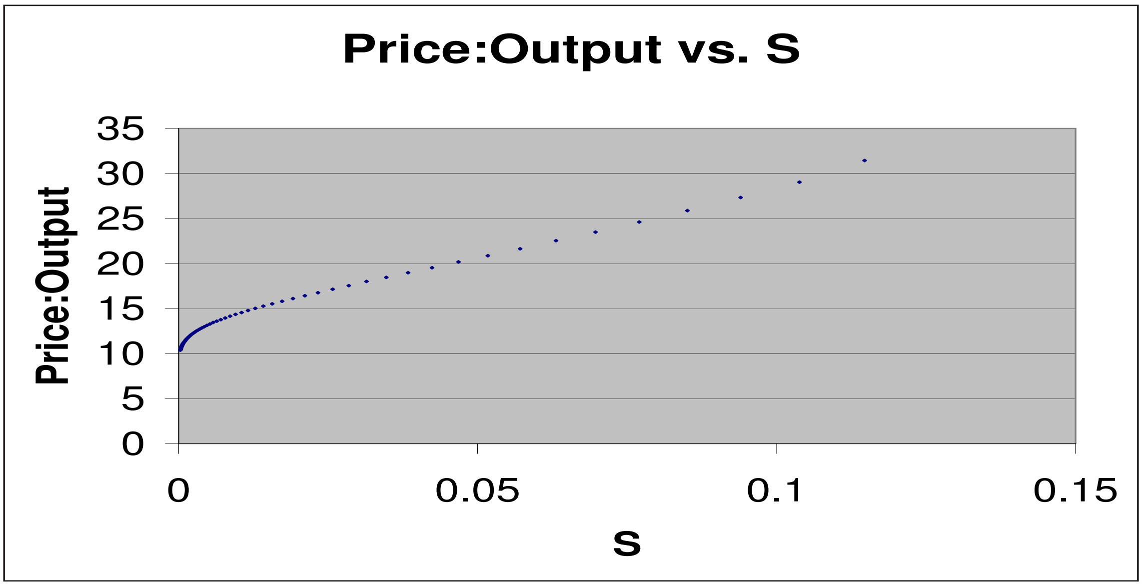

Figure 2: Price-Output as a function of S.

initial book value of debt before defaulting. This number is consistent with recovery rates of approximately 50%, and bankruptcy costs of approximately 15%, which is broadly consistent with the empirical estimates of Andrade and Kaplan (1998).[16] Analogously, for Aaa firms, we choose $P_{def}^{Aaa} = (0.6)(0.1308) \approx 0.078$, where again the (0.1308) matches the calibration of HH.

Since the CC model is calibrated in real terms, and since corporate bonds are written in nominal terms, we need to account for these correctly. For simplicity, we assume a constant inflation rate of 3%. Hence, the nominal growth rate of output is $g_o = .03 + .0189 = .0489$. The coupon rate for the Treasury bond is set equal to the sum of the real risk free rate[17] plus inflation. The coupon rate on the corporate bond is set equal to the real risk free rate plus

[16]Consistent with the findings of HH, we find the credit spread estimates to be very robust to changes in default boundary location, since in order to match historical default rates, a higher boundary, for example, implies a lower volatility, which tends to cancel most of the effect on credit spreads. As noted in the introduction, only changing the *covariance* of the pricing kernel with default and recovery rates will produce significantly different results.

[17]recall that $r(s)$ is a constant for all values of $s < s_{max}$, and that, in discrete time, s can actually be greater than s_{max}. Hence, $r(s)$ is stochastic in the CC framework. However, in the continuous time version of the CC model discussed in the appendix, interest rates are truly constant since s_{max} constitutes a natural boundary.

$s(0)$	Steady State Distribution	Baa Spread over Treasury	Baa Q-Default Rate	Baa P-Default Rate	Aaa Spread over Treasury	Aaa Q-Default Rate	Aaa P-Default Rate	(Baa - Aaa) Spread
-3.66	0.011	75.7	5.5	0.95	4.7	0.30	0.04	71.0
-3.56	0.014	75.2	5.4	0.96	4.6	0.30	0.04	70.6
-3.46	0.017	75.1	5.4	1.02	4.6	0.29	0.04	70.5
-3.36	0.023	73.9	5.3	1.05	4.5	0.29	0.04	69.4
-3.26	0.029	73.2	5.3	1.09	4.5	0.29	0.05	68.7
-3.16	0.036	72.3	5.2	1.16	4.5	0.29	0.05	67.8
-3.06	0.046	71.3	5.1	1.20	4.5	0.28	0.05	66.8
-2.96	0.057	70.1	5.0	1.27	4.5	0.28	0.06	65.7
-2.86	0.072	68.9	5.0	1.34	4.3	0.27	0.06	64.6
-2.76	0.090	67.6	4.9	1.44	4.1	0.25	0.07	63.5
-2.66	0.109	65.3	4.7	1.54	4.0	0.25	0.07	61.3
-2.56	0.128	62.2	4.5	1.66	3.8	0.24	0.07	58.4
-2.46	0.147	59.1	4.3	1.79	3.4	0.20	0.08	55.7
-2.36	0.144	53	3.8	1.94	2.9	0.17	0.08	50.1
-2.27	0.038	45.6	3.3	2.10	2.9	0.16	0.10	42.7
Average		63.48	4.58	1.55	3.83	0.24	0.062	59.7
Std. Dev.		7.66	0.55	0.31	0.59	0.04	0.01	7.1

Table 3: Model generated 4-year Baa and Aaa credit spreads when the nominal default boundary is a constant (equal to (0.6)(0.4328) $\approx$ .26 for Baa firms and (0.6)(.1308) $\approx$ 0.078 for Aaa firms). The idiosyncratic risk needed to match historical default rate for Baa (Aaa) of 1.55% (0.06%) is $\sigma_{idio}^{Baa} = 0.268$ ($\sigma_{idio}^{Baa} = 0.345$).

inflation plus 100bp. As such, both bonds are issued near par value.[18]

Following HH, we calibrate the value of σ_{idio} to match the historical 4-year default frequency. We do this by determining the 4-year conditional default frequency as a function of the state variable s, and then weight these results by the steady state distribution π_{ss}. We find $\sigma_{idio}^{Baa} = 0.268$ and $\sigma_{idio}^{Aaa} = 0.345$.[19]

Upon default, we assume that the agent immediately receives a recovery of 0.449, consistent with the recovery rate of Moody's 2005 report. All future promised coupon payments receive zero recovery. We then estimate the (Baa - Treasury) spread as a function of $s(0)$. The results are tabulated in Table 3.

The model generates an average (Baa - Aaa) spread of 59.7 bp with a standard deviation of 7.1 bp. These results fall far short of the historical level of 109 bp and the historical volatility of 41 bp. Further, this model predicts that 4-year forward default rates are strongly pro-cyclical.

[18] As expected, we find that credit spreads generated from this model are extremely insensitive to the specification of the coupon rate.

[19] We emphasize that it lower the levels of σ_{idio} may be obtained by, for example, assuming the default boundary is located at 80% of average leverage ratios rather than at 60% leverage ratios as we did above. Further, as in Collin-Dufresne and Goldstein (2001), we can specify the debt outstanding to have a deterministic trend, especially for Aaa debt.

$s(0)$	Steady State Distribution	Baa Spread over Treasury	Baa Q-Default Rate	Baa P-Default Rate	Aaa Spread over Treasury	Aaa Q-Default Rate	Aaa P-Default Rate	(Baa - Aaa) Spread
-3.66	0.011	175.8	30.82	2.34	42.5	8.62	0.36	133.3
-3.56	0.014	168.5	29.60	2.29	39.6	8.08	0.52	128.9
-3.46	0.017	164.1	28.80	2.48	38.1	7.66	0.20	126.0
-3.36	0.023	158.2	27.88	2.98	35.2	7.11	0.35	123.0
-3.26	0.029	155.3	27.37	2.89	33.7	6.74	0.41	121.6
-3.16	0.036	149.4	26.34	3.38	30.8	6.31	0.36	118.6
-3.06	0.046	143.6	25.52	3.40	29.3	5.97	0.32	114.3
-2.96	0.057	139.2	24.63	3.79	27.8	5.68	0.51	111.4
-2.86	0.072	133.3	23.69	4.31	26.4	5.38	0.44	106.9
-2.76	0.090	130.4	23.03	4.50	26.4	5.15	0.48	104.0
-2.66	0.109	123.0	21.89	4.83	24.9	4.84	0.56	98.1
-2.56	0.128	118.7	21.05	5.35	22.0	4.44	0.68	96.7
-2.46	0.147	111.3	20.04	6.01	22.0	4.22	0.79	89.3
-2.36	0.144	104.0	18.61	6.65	19.0	3.79	0.89	85.0
-2.27	0.038	95.2	17.29	7.68	16.1	3.34	1.07	79.1
Average		126.8	22.54	4.89	25.5	5.10	0.63	101.3
Std. Dev.		20.6	3.49	1.40	6.3	1.30	0.21	14.4

Table 4: Model generated 10-year Baa and Aaa credit spreads when the nominal default boundary is a constant (equal to 0.6*0.4328). The idiosyncratic risk needed to match historical default rate for Baa (Aaa) of 4.89% (0.63%) is $\sigma_{idio}^{Baa} = 0.223$ ($\sigma_{idio}^{Baa} = 0.285$).

That is, the 4 year forward probability of default *increases* with the initial value s_0. This occurs because in good times, expected returns are low – indeed, low enough to more than compensate for the lower payout ratio and lower volatility. Hence, if the default boundary is specified as a constant, then the CC economy predicts that there is a greater probability for default in the near future when the economy is in a boom rather than in a recession. To quantify this result, we estimate the theoretical regression coefficient for the 4-year future default rates on spreads via:

$$\begin{aligned}
\beta_{theory} &= \frac{\text{cov}_{ss}(\text{def rate, spread})}{\text{var}_{ss}(\text{spread})} \\
&= \frac{\sum_j \pi_{ss}(s_j)\left(\text{def rate}_j - \text{E}_{ss}[\text{def rate}]\right)\left(\text{spread}_j - E_{ss}[\text{spread}]\right)}{\sum_{s_j} \pi_{ss}\left(\text{spread}_j - E_{ss}[\text{spread}]\right)^2} \\
&= -2.78 \qquad (38)
\end{aligned}$$

This result contrasts significantly with the empirical result of $\beta = +0.86$ reported in the previous section.

In addition to the four-year results, we repeat the same procedure on ten-year spreads in Table 7. We find (Baa - Aaa) spreads to be 101.4 bp, short of the 131 bp empirical estimate reported by HH. Also, the volatility of 14.4 bp is well below empirical observation.

Hence, compared with historical data, the CC model with a constant initial leverage ratio generates a predicted (Baa - Aaa) spread that is: i) too low, ii) not sufficiently volatile, and iii) varies negatively with 4-year forward default rates. Although within the CC framework the representative agent is willing to pay a large premium for securities that pay off in bad states, the constant boundary specification suggests that defaults happen too often in the near future when the current state of nature is good.

Previously, we identified three channels that can be used to increase credit spreads while matching historical default rates. Channel 1a) has been captured by using the CC pricing kernel. We now investigate channel 1b), namely, time-varying default boundaries, to see if this property can capture the historical credit spread data better.

4.1.2 Counter-cyclical default boundary case

Arguably the most vague aspect of structural models of default is the link between outstanding debt and the location of the default boundary. Some models (e.g., Leland (1994)) endogenously determine the default boundary to be a constant fraction of the level of debt outstanding. Other models (e.g., Longstaff and Schwartz (1995), Collin-Dufresne and Goldstein (2001)) exogenously specify the default boundary location to be either a constant or to generate stationary leverage ratios, but do not make a connection between the default boundary and debt level. The dynamics and the location of the default boundary are not observable, so researchers must specify the default boundary based on indirect information.

As noted previously, Collin-Dufresne, Goldstein and Martin (2001), Cheyette et al. (2003), and Shaefer and Strebulaev (2004) all document that market wide (e.g., Fama-French) factors have additional predictive power for changes in credit spreads even after controlling for all variables that 'standard' structural models claim are sufficient to determine spreads. One way to capture this empirical fact is to assume that default boundaries are dynamic and are affected by economic conditions. As such, and given that a constant default boundary assumption generates pro-cyclical default rates, in this section we specify the default boundary to be linearly decreasing in the current value of S_t:[20,21]

$$P^{Baa}_{def}(s) \quad = \quad (0.6)(0.4328)\left(1 - slope * (S - \overline{S})\right) \tag{39}$$

[20] We choose the boundary to be linear in S rather than $s = \log S$ because S is bounded by $S \in (0, S_{max} \approx 0.1)$ whereas s has no minimum.

[21] We discuss the empirical implications for leverage and other possible proxies below.

Bond Maturity (Years)	slope	σ_{idio}^{Baa}	σ_{idio}^{Aaa}	Baa-Treasury Spread	Aaa-Treasury Spread	(Baa - Aaa) Spread
4	12.5	0.230	0.302	147.8	6.6	141.2
				$\pm$73.6	$\pm$3.0	$\pm$70.6
4	7.0	0.252	0.320	112.2	5.4	106.8
				$\pm$43.7	$\pm$1.9	$\pm$41.8
10	12.5	0.191	0.261	239.5	52.2	187.2
				$\pm$62.3	$\pm$17.1	$\pm$45.6
10	7.0	0.208	0.272	194.1	41.2	153.0
				$\pm$45.5	$\pm$13.5	$\pm$31.0

Table 5: Model generated Baa and Aaa credit spreads when the nominal default boundary is specified as in equations (39) and (40). The choice of slope = 12.5 matches the historical point estimate of the regression coefficient between spreads and forward defaults. The choice of slope = 7 matches the historical (Baa - Aaa) population volatility of 41.6bp

$$P_{def}^{Aaa}(s) \quad = \quad (0.6)(0.1308)\left(1 - slope * (S - \overline{S})\right). \tag{40}$$

Now, recall from Table 1 we reported two features of the data that would be desirable to match: i) the estimated historical regression coefficient (and standard error) between 4-year future default probabilities and (Baa - Aaa) spreads is $\beta \sim .86 \pm .42$, and ii) the standard deviation of the unconditional (Baa - Aaa) distribution is 41bp. As such, we consider two different values for the slope parameter: the first value (*slope* = 12.5) perfectly matches the point estimate of the regression coefficient. The second value *slope* = 7 matches the unconditional (Baa - Aaa) volatility. The results are given in Table 5. We see that for slope = 12.5 both the predicted 4-year $(Baa - Aaa)$ spread levels and volatility $(141.2, 70.6)$ are *above* historical levels. For *slope* = 7, which is calibrated to perfectly match observed (Baa - Aaa) unconditional volatility of 41.8bp, we see that the predicted 4-year $(Baa - Aaa)$ spread level (106.8) captures the historical value extremely well. Furthermore, this estimate for slope generates a regression coefficient of 0.64 – well within one standard error of the point estimate. As such, we consider *slope* = 7 to be our benchmark calibration.

We note, however, that our results cannot explain either the average level or the time variation of the Aaa-Treasury spread. Taking at face value the prediction of the model, this seems to suggest (very much in line with HH) that much of the Aaa-Treasury spread is due to factors independent of credit risk.

In summary, the CC model (which is calibrated to match many properties associated with equity returns) can successfully capture the (Baa - Aaa) spread, volatility, and the correlation between spreads and future default rates if the default boundary is modeled as counter-cyclical. This is all accomplished while calibrating the model to match historical default and average recovery rates.

Finally, from Table 5 we note that the model also does fairly well at the 10-year maturity level. In particular, for *slope* = 7 we find the average (Baa - Aaa) spread to be 153bp, broadly consistent with the empirical findings of HH.

4.1.3 Counter-cyclical default boundary due to leverage changes

It remains to be seen whether we can tie the variation of the implied default boundary of the previous section to variation in empirical leverage ratios, as suggested in Table 1. To tackle this issue, we first regress the market leverage ratio (MLV) of Baa rated bonds on the exponential surplus consumption ratio and obtain the follow relation:

$$MLV_{Baa}(s) = 0.52 - .61S. \tag{41}$$

Note that the coefficient 0.61 multiplying S is approximately one-third the size of our benchmark case (.6)(.4328)(7) = 1.82. Turning this around, when we choose *slope* = 2.367 so that the default boundary specified in equation (39) has the same sensitivity to S as does leverage, we find that the predicted (Baa - Aaa) spread level and volatility are (77.0,18.8). Furthermore, the theoretical regression coefficient of default rate on credit spread is -0.65, versus the empirically observed estimate of +.86. These findings suggest that time-varying leverage alone is not sufficient to generate the appropriate level of credit spread; nor can it induce counter-cyclical default rate. Therefore, the default boundary appears to be more counter-cyclical than what can be captured solely by the counter-cyclical nature of leverage. Again, this result is consistent with the fact that the Fama French factors have predictive power even after all traditional structural form factors have been controlled for.

4.1.4 Pro-cyclical Recovery Rates

So far, we have investigated how channels 1a) and 1b) discussed in the introduction can be used to help explain the 'credit spread puzzle'. Recently, several papers (e.g., Altman, Resti, and Sironi (2004, 2005) and Acharya, Bharath, and Srinivasan (2005)) have noted that recovery rates are pro-cyclical. Here we report that this channel can also have a significant impact on credit spreads. Interestingly, we find that pro-cyclical recovery rates can only be considered a partial 'substitute' for counter-cyclical default boundary in that, if one reduces the parameter of *slope* used in equation (39) and then considers a pro-cyclical recovery rate that matches the historical unconditional recovery rate, then one obtains very similar predictions for the (Baa - Aaa) spread level and standard deviation. In particular, we specified the *slope* in equation

(39) to equal 5.0. We then specified the recovery rate as[22]

$$\text{Recovery}(S) = .35 + 2.5S \tag{42}$$

We found that this matched the unconditional recovery rate of 0.449. Further, it generated a (Baa - Aaa) spread of 109bp with a unconditional standard deviation of 40bp. These results are very similar to that obtained in our base case with slope = 7 above and a constant recovery rate.

It is important to note that a counter-cyclical default boundary is necessary even in the presence of pro-cyclical recovery rate. This is because a pro-cyclical recovery rate only induces higher credit spreads, but does not affect default probability. Put differently, in the case of pro-cyclical recovery rate but constant default boundary, we shall still obtain counterfactual pro-cyclical default rate in the CC model.

5 The Bansal-Yaron long-run risk model

In this section we consider the implication for credit spreads of an alternative model, that of Bansal and Yaron (BY 2004). BY's model is very successful in explaining many features of equity data, and in particular: the average equity premium and its volatility, the average price dividend ratio and its volatility, the average risk-free rate and its volatility. In contrast to CC's model, which explains all these features with iid consumption but time varying risk-aversion generated by the habit process, BY's model has standard (Epstein-Zin type) utility function with constant risk-aversion but modifies the consumption process. It allows both its growth and volatility to follow highly persistent mean-reverting stochastic processes. BY argue that in finite sample their consumption process cannot statistically be distinguished from the iid consumption process assumed in CC. However, it helps explain the equity premium puzzle using what appears to be a quite different mechanism than CC. Namely, one based on consumption/cash-flow risk as opposed to the risk-premium/discount rate risk in CC. Implementing the BY model and studying its implications for spreads is thus interesting for two reasons. First, it provides a potential alternative explanation for credit spread level and variation. The results of the calibration can help us sort out which components are more important for spreads. Second, looking at the implications of this model for credit spreads, provides an out-of-sample test of the two explanations of the equity premium: cash-flow risk versus risk-aversion. It seems natural to expect that a model that can explain many features of equity

[22] At first blush, this calibration would seem to overestimate the recovery rate of .449 since the average value of S is approximately 0.09. However, most defaults occur for values of S well below the mean.

prices should also be able to explain corporate bond prices accurately. A failure along that dimension might indicate that the model is misspecified or that bond and equity markets are segmented. Admittedly both models are highly stylized and may illustrate two different mechanism both of which may be at work in the data. In fact, our results show that time varying cash flow risk alone cannot match either level or time-series properties of spreads. It generates too low average spread level with too high a covariance with default probabilities. Time varying risk-premia appears thus crucial to explain spreads. On the other hand, the results suggest that adding stochastic volatility to a model with time varying risk-aversion may drastically improve the time series properties of predicted spreads (time varying risk premia raise allow for a larger wedge between P and Q measure default probability; stochastic volatility induces counter-cyclical default probabilities).

We first describe the continuous time version of the BY model we implement, then the calibration and the results.

5.1 The continuous time 'BY' model

BY propose a complex model based on Epstein-Zin preferences and a consumption process with mean reverting consumption growth and volatility. To solve their model they use several approximations. First, they use a log-linearization of gross returns effectively twice: (i) to express the log pricing kernel as an affine function of the state variables, and (ii) to obtain an affine log price/dividend ratio (for both the consumption and dividend claims). Second, they assume that the variance of log consumption is normally distributed (i.e., possibly negative). Below we propose a continuous time version of their model with similar state variables dynamics (we choose affine dynamics for the consumption, its growth and volatility to match unconditional and conditional moments of BY's model, keeping a positive variance). We follow BY in approximating the pricing kernel as affine,[23] but derive explicit solutions for the price of the consumption and dividend claims (i.e., we do not use the log-linearization approximation to solve for price/dividend ratios). As we document below, the model essentially matches most of the empirical facts of dividend claim as shown in BY. So, in principle, the BY model could perform as well as the CC model in explaining spreads once calibrated to successfully fit equity returns. We turn to this in the next section.

Our version of the 'BY' model is:

$$dc_t = (\mu + x_t)dt + (v_t + \bar{v})\, dZ_c(t) \tag{43}$$

[23] We use the approach proposed in Collin-Dufresne and Goldstein (2005) to improve upon the log-linearization of Campbell and Shiller to solve the model.

$$dd_t = (\mu_d + \phi x_t)dt + \sigma_d (v_t + \bar{v})\, dZ_c(t) \quad (44)$$

$$dx_t = -\kappa x_t dt + \sigma_x (v_t + \bar{v})\, dZ_x(t) \quad (45)$$

$$dv_t = \nu(\bar{v} - v_t)dt + \sigma_v\, dZ_v(t) \quad (46)$$

where c, d are the log consumption and dividend process respectively and (Z_c, Z_x, Z_v) are independent Brownian motions.

The representative agent has recursive utility of the Epstein-Zin-Kreps-Porteus type, i.e., maximizes a utility index of the form:

$$J(t) = \mathrm{E}_t \left(\int_t^T f(C_s, J(s)) + J(T) \right) . \quad (47)$$

where the so-called 'normalized' aggregator function is given by:

$$f(C,J) = \begin{cases} \frac{\beta u_\rho(C)}{((1-\gamma)J)^{1/\theta - 1}} - \beta\theta J & \gamma, \rho \neq 1 \\ (1-\gamma)\beta J \log(C) - \beta J \log((1-\gamma)J) & \gamma \neq 1, \rho = 1 \\ \frac{\beta u_\rho(C)}{e^{(1-\rho)J}} - \frac{\beta}{1-\rho} & \gamma = 1, \rho \neq 1 \,. \end{cases} \quad (48)$$

for $u_\rho(c) = \frac{c^{1-\rho}}{1-\rho}$.

As is well-known, in the special case where $\gamma = \rho$ this reduces to the standard time separable constant relative risk-aversion utility. The pricing kernel is given by (e.g., Duffie and Skiadas (1999)):

$$\Lambda(t) = e^{\int_0^t f_J(C_s, J_s) ds} f_C(C_t, J_t) \quad (49)$$

Further, given the affine dynamics of the state variables, it can be shown that the dynamics of the pricing kernel can be approximated (Collin-Dufresne and Goldstein (2005)) as follows:

$$\frac{d\Lambda_t}{\Lambda_t} = -r_t dt - (\lambda_{c0} + \lambda_{c1} v_t) dZ_c(t) - (\lambda_{v0} + \lambda_{v1} v_t) dZ_v(t) - (\lambda_{x0} + \lambda_{x1} v_t) dZ_x(t) \quad (50)$$

$$r_t = \alpha_0 + \alpha_x x_t + \alpha_v (v_t + \bar{v})^2 \quad (51)$$

These equation can be compared with equation (A1) and (A10) in the appendix of BY. The two models are identical for the case where volatility is constant and equal to its long-term mean (Case I in BY), and differ only slightly in the case where volatility is stochastic (Case II in BY).[24] We note that all the parameters of the affine pricing kernel (α_i, λ_j) are endogenous. We shall report the values of these parameters (obtained using a continuous time version of a Campbell-Shiller approximation proposed by Collin-Dufresne and Goldstein (2005)) below.

[24]The only difference between our model relative to case II in BY is that we assume that volatility of consumption growth and (not variance as in BY) follows a Gaussian AR1 process. This avoids the issue of negative variances.

In this model (once we adopt the approximate pricing kernel dynamics given in (50) above) we can solve explicitly for the price dividend ratio and all relevant quantities. Under the risk-neutral measure the processes are given by:

$$
\begin{aligned}
dc_t &= (\mu + x_t)dt + (v_t + \bar{v})\ \left(dZ_c^Q(t) - (\lambda_{c0} + \lambda_{c1} v_t)dt\right) && (52)\\
dd_t &= (\mu_d + \phi x_t)dt + \sigma_d (v_t + \bar{v})\ \left(dZ_c^Q(t) - (\lambda_{c0} + \lambda_{c1} v_t)dt\right) && (53)\\
dx_t &= -\kappa x_t dt + \sigma_x (v_t + \bar{v})\ \left(dZ_x^Q(t) - (\lambda_{x0} + \lambda_{x1} v_t)dt\right) && (54)\\
dv_t &= \nu(\bar{v} - v_t)dt + \sigma_v\ \left(dZ_v^Q(t) - (\lambda_{v0} + \lambda_{v1} v_t)dt\right) && (55)
\end{aligned}
$$

We first derive the price of the dividend claim which solves:

$$
\begin{aligned}
P^d(t) &= \mathrm{E}_t^Q[\int_t^\infty e^{-\int_t^s r(u)du + d(s)} ds] && (56)\\
&= D(t) \int_t^\infty \Psi_{a,b,c,d}(\tau, x_t, v_t) d\tau && (57)
\end{aligned}
$$

where we have defined:

$$
\Psi_{a,b,c,d}(T-t, x_t, v_t) = \mathrm{E}_t^Q\left[e^{-\int_t^T (a + b x_s + c v_s + d v_s^2) ds}\right] \qquad (58)
$$

and the constants:

$$
\begin{aligned}
a &= \alpha_0 + \alpha_v \bar{v}^2 - \mu_d + \sigma_d \bar{v} \lambda_{c0} - \frac{1}{2}\sigma_d^2 \bar{v}^2 && (59)\\
b &= \alpha_x - \phi && (60)\\
c &= 2\alpha_v \bar{v} + \sigma_d * (\lambda_{c0} + \lambda_{c1} * \bar{v}) - \sigma_d^2 * \bar{v} && (61)\\
d &= \alpha_v + \sigma_d * \lambda_{c1} - \sigma_d^2/2 && (62)
\end{aligned}
$$

We find:

$$
\Psi_{a,b,c,d}(\tau, x, v) = \exp(A(\tau) + B(\tau)x + C(\tau)v + D(\tau)v^2) \qquad (63)
$$

where A, B, C, D satisfy a standard system of ODE (the approach follows Duffie and Kan (1996)) that can be solved in closed-form partially and numerically using standard solvers (e.g., Mathematica). For simplicity we do not report the system of ODE.[25]

Note that the consumption claim is obtained similarly

$$
\begin{aligned}
P^c(t) &= \mathrm{E}_t^Q[\int_t^\infty e^{-\int_t^s r(u)du + c(s)} ds] && (64)\\
&= C(t) \int_t^\infty \Psi_{a',b',c',d'}(\tau, x_t, v_t) d\tau && (65)
\end{aligned}
$$

[25]They are available upon request.

where a', b', c', d' are obtained from a, b, c, d with $\mu_d = \mu_c, \phi = 1, \sigma_d = 1$.

Finally we can get all the quantities we need for calibration, using a two step procedure. First we compute all the conditional moments, then we integrate the conditional moments with respect to the unconditional distribution of the state variables. We obtain the unconditional distribution of x and v by simulation. (We denote by $f(x, v)$ the corresponding joint density). Define $\frac{P^d}{D} = Y(x, v)$ then we compute the following moments:

1. Unconditional mean of log price-dividend ratio:

$$E[\log \frac{P^d}{D}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \log Y(x, v) f(x, v) dx dv \tag{66}$$

2. Unconditional variance of log price-dividend ratio:

$$V[\log \frac{P^d}{D}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\log Y(x, v) - E[\log \frac{P^d}{D}] \right)^2 f(x, v) dx dv \tag{67}$$

3. Unconditional average volatility of log price dividend ratio:

$$E[\left(d \log \frac{P^d}{D} \right)^2]/dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left((\frac{Y_x}{Y})^2 \sigma_x^2 \bar{v}^2 + (\frac{Y_v}{Y})^2 \sigma_v^2 \right) f(x, v) dx dv \tag{68}$$

4. Unconditional average Risk-premium:

$$E[\frac{dP^d + Ddt}{P^d} - rdt]/dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\lambda_{c0} \sigma_d v^2 + \frac{Y_x}{Y} \lambda_{x0} \sigma_x \bar{v} + \frac{Y_v}{Y} \lambda_{v0} \sigma_v v \right) f(x, v) dx dv \tag{69}$$

5. Unconditional average volatility of excess return:

$$E[\left(\frac{dP^d + Ddt}{P^d} - rdt \right)^2]/dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\sigma_d^2 v^2 + (\frac{Y_x}{Y})^2 \sigma_x^2 \bar{v}^2 + (\frac{Y_v}{Y})^2 \sigma_v^2 \right) f(x, v) dx dv \tag{70}$$

6. Unconditional mean short rate:

$$\mathrm{E}[r] = \alpha_0 + \alpha_x \mathrm{E}[x] + \alpha_v \mathrm{E}[(v + \bar{v})^2] \tag{71}$$

7. Unconditional variance of short rate:

$$V[r] = V[\alpha_x x + \alpha_v (v + \bar{v})^2] \tag{72}$$

Note that $Y(x,v) = \frac{P}{D} = \int_t^\infty \Psi_{a,b,c}(\tau,x,v)d\tau$ so that

$$Y_x = \int_t^\infty B(\tau)\Psi_{a,b,c}(\tau,x,v)d\tau \tag{73}$$

$$Y_v = \int_t^\infty \left(C(\tau)+2vD(\tau)\right)\Psi_{a,b,c}(\tau,x,v)d\tau \tag{74}$$

for appropriate parameters a, b, c. Similarly, we obtain the equivalent quantities for the consumption claim.

We set the parameters of the consumption and dividend process identical to those of BY except for the volatility parameters, since in our case volatility follows an AR1 process, whereas in BY variance does. We choose the parameters of the volatility process so that the unconditional means and variances of the variance process in our model coincide both under physical and risk-neutral measures with that of BY's model. Following BY's table IV we thus obtain:

$$\mu = \mu_d = 0.0015 * 12, \;\; \phi = 3, \;\; \sigma_d = 4.5, \;\; \kappa = (1-0.979) * 12, \;\; \sigma_x = 0.044 * 12,$$

$$\bar{v} = 0.027, \;\; \nu = (1-0.987) * 6, \;\; \sigma_v = 0.00177.$$

Further, we choose preference parameters similar to BY: a constant relative risk-aversion parameter of $\gamma = 6$ and an intertemporal rate of substitution of 1.5 which implies $\rho = 1/1.5$. Lastly we set the time preference parameter $\beta = 0.009$.

The preference parameters (α, ρ, γ) and the parameters of the consumption process jointly determine all the parameters of the pricing kernel (α_i and λ_{ij}) of the affine approximation given in (50) above. For example, we obtain $\lambda_{c1} = \gamma$ and $\alpha_x = \rho$.[26]

With these parameters we confirm BY's result that the model matches post-war equity data very well. It generates (unconditional average):[27]

- an equity premium of 6.5%,
- a risk-free rate of 1.1% with a volatility of 1.3%,

[26]For the other parameters the mapping is non-linear and we solve for them numerically as described in Collin-Dufresne and Goldstein (2005).

[27] More precisely, in our affine approximation of the economy the interest rate is

$$r = 0.011 + 0.667x - 0.468v - 8.317v^2$$

Further, using the approach of Collin-Dufresne and Goldstein, the price dividend ratio can be accurately approximated by:

$$\frac{P}{D} = \exp(3.141 + 7.895x - 21.318v - 182.024v^2)$$

In our case the price-dividend ratio obtains in closed-form as an integral of exponential affine functions. Therefore the approximation of the price-dividend ratio is not necessary. However, it is so accurate on a domain encompassing plus or minus 4 standard deviations of the state variables that we use it for simplicity in the extraction of the state variables below.

- a price-dividend ratio (PD) such that $\exp(\mathrm{E}\log PD) = 23.13$ with a volatility of the log-PD equal to 11.87%,

- a volatility of equity return of 16.9%.

We therefore turn to the implications of this model for corporate bond spreads.

5.2 Implications for Credit Spreads

As before we suppose that the return generating process is that of an average firm with additional idiosyncratic risk. We follow exactly the same procedure as described above for the CC model. We solve for the claim to aggregate output.[28] We price nominal bonds assuming that inflation is constant and equal to 3%.

Further, we note that while the emphasis of BY is on cash-flow risk from two sources (growth rate risk and volatility risk), their model also exhibits time variation in risk-premia. Our parametrization allows us to analyze the implications of each component separately. In particular, to distinguish the impact of growth rate risk, volatility and risk-premia on predicted spreads, we distinguish three cases:

- Case I: growth rate risk (i.e., we set the volatility and risk-premia to be constant: $\sigma_v = 0, v_0 = \bar{v}$ and $\lambda_{j1} = 0$, $j = c, v, x$).

- Case II: growth rate and volatility risk (i.e., we set the risk-premia to be constant: $\lambda_{j1} = 0$, $j = c, v, x$)).

- Case III: growth rate and volatility risk as well as time varying risk-premia.

For each case, we report the level of idiosyncratic volatility needed to generate an unconditional average four year default probability of 1.55%. The prediction of a constant (nominal) default Barrier Black-Cox (1976), Longstaff and Schwartz (1995) model are given in table 6 below.

The table clearly illustrates the implications of cash flow risk for credit spreads. Cash-flow risk alone generated by time varying growth rate risk without time variation in risk-premia (which also corresponds to case I in BY) cannot explain the spread puzzle. The results in that

[28] We assume that the process for the log of aggregate output o_t is identical to that of aggregate dividends:

$$do_t = (\mu_d + \phi x_t)dt + \sigma_o (v_t + \bar{v})\, dZ_c(t)$$

with a lower volatility scaled to match panel A of table 3 panel A, i.e., $\sigma_o = \sigma_d * \frac{0.063}{0.09}$.

CASE I							
Stochastic growth - Constant volatility and risk-premia							
P def prob	P-DP Std Dev	Q def prob	Q-DP Std Dev	Average Spread	Std Dev. Spread	Reg Coef	σ_{idio}^{Baa}
0.0155	0.0031	0.036	0.0066	0.0045	0.00082	4.04	0.245
(0.0005)	(0.0009)						
CASE II							
Stochastic growth and volatility - Constant risk-premia							
P def prob	P-DP Std Dev	Q def prob	Q-DP Std Dev	Average Spread	Std Dev. Spread	Reg Coef	σ_{idio}^{Baa}
0.0153	0.0057	0.0394	0.0145	0.0046	0.0017	3.2589	0.242
(0.0004)	(0.0006)						
CASE III							
Stochastic growth and volatility - Time varying risk-premia							
P def prob	P-DP Std Dev	Q def prob	Q-DP Std Dev	Average Spread	Std Dev. Spread	Reg Coef	σ_{idio}^{Baa}
0.0155	0.0049	0.0492	0.0204	0.0058	0.0025	1.853	0.241
(0.0004)	(0.0007)						

Table 6: Estimated values of P and Q default probabilities as well as the unconditional mean and variance of the credit spread for four year to maturity Baa firms. Standard errors of estimates are in parenthesis. Parameters of the typical Baa firm are as defined above for the output claim. The spread is simulated within a structural model which assumes a constant nominal default boundary at 60% of the average Baa leverage ratio ($K = 0.6 * 0.4328$. Upon default bond recover constant fraction of face value corresponding to average historical Baa recovery rate 51.31%. Simulations are run with 100000 runs for each price estimation (conditional on state), with standard antithetic variance reduction.

case (presented in the first row) are only marginally different from our benchmark constant coefficient case.

We note that relative to the benchmark case our case I only introduces a time varying payout ratio. The risk-premium induces a larger average payout under the risk-neutral measure than under the historical measure, which explains the slightly larger credit spread relative to the benchmark case, but is not sufficient to explain the data. This contrasts with the equity results. Indeed, BY show that even with constant risk-premia, long-run cash flow risk can have a sizable effect on the equity premium. This is because equity are claims to long lived cash-flows which may be subject to highly persistent shocks and thus are effectively much riskier than they might appear (i.e., in contrast to an iid consumption world) in the long run. However, this effect does not appear to have a sizable impact on short maturity corporate bonds. Effectively, stochastic growth rate of dividends does not induce sufficient variation in default rates to explain sizable spreads given constant risk-premia.

Case II shows that adding stochastic volatility only marginally increases the average spread relative to Case I (the spreads increases by 1 bp to 46 bp). However, adding stochastic volatility increases the volatility of spreads substantially (from 0.8 bp to 17 bp).

Instead, Case III shows the crucial role played by the time variation in risk-premia, even in the cash-flow risk model, to explain the credit spread puzzle. The BY model makes risk-premia countercyclical by linking them to volatility. This effectively makes high volatility states ‘more

	4-year Baa			4-year Aaa			
	Spread over Treasury	Q-Default Rate	P-Default Rate	Spread over Treasury	Q-Default Rate	P-Default Rate	(Baa - Aaa) Spread
Average	0.0058	0.0492	0.0155	0.0006	0.0052	0.0004	0.0052
Std Dev.	0.0025	0.0204	0.0049	0.0007	0.0001	0.0000	0.0018
	10-year Baa			**10-year Aaa**			
	Spread over Treasury	Q-Default Rate	P-Default Rate	Spread over Treasury	Q-Default Rate	P-Default Rate	(Baa - Aaa) Spread
Average	0.0118	0.240	0.0487	0.0051	0.1154	0.0063	0.0055
Std Dev.	0.0036	0.0636	0.0123	0.0028	0.0588	0.0044	

Table 7: Model generated 4-year and 10-year Baa and Aaa credit spreads. The idiosyncratic risk needed to match the 4-year historical default rate for Baa (Aaa) of 1.55% (0.04%) is $\sigma^{Baa}_{idio} = 0.241$ ($\sigma^{Baa}_{idio} = 0.110$). The idiosyncratic risk needed to match the 10-year historical default rate for Baa (Aaa) of 4.89% (0.63%) is $\sigma^{Baa}_{idio} = 0.188$ ($\sigma^{Baa}_{idio} = 0.078$).

expensive.' Since defaults occur more frequently when volatility is high, stochastic volatility linked to time-varying risk-premia effectively shifts default events from good to bad states, the combined effects imply an increased spread for the same average loss rate. The model can explain a larger fraction of the spread 58 bp as opposed to the 46 bp in Case II. Also, case III further increases the predicted volatility in spreads. Note that the BY model because of the cash flow risk induced counter-cyclicality in defaults does not face the problem of the CC model of generating pro-cyclical P-measure default probabilities. In addition, when time varying risk-premium are introduced, the regression coefficient of future default probabilities on spreads decreases from 3.26 to 1.85. This is consistent with the notion that adding time varying risk-premia generates higher variation in credit spreads given the variation in default rates.

The next table looks at the (Baa - Aaa) spread generated by the model across maturities for case III, i.e., with all three BY channels at work. The table shows that the BY model predicts similar size credit spread for 4-year maturity bonds across rating classes as the CC model with constant boundary. The BY model also predicts similar magnitude spread for the 10-year Baa spread (around 118 bp as opposed to 126 bp for the CC model). However, the BY model predicts a spread of 55 bp for the 10-year Aaa-Treasury which is twice as large as the spread predicted by the CC model and much closer to historical levels. As a result, the BY model predicts a substantially flatter credit spread curve than the CC model. The latter predicts a 50% increase in spreads when going from 4 to 10 year maturity vs. less than 10% increase for the BY model.

Overall, we conclude that to explain the size and time variation in credit spreads time varying risk-aversion is essential, even within a model with 'long run cash flow risk' as studied by BY.[29]

On the positive side, the BY model generates the correct prediction that in bad (i.e., high volatility and/or low growth) states the default probability is high. This is because most variation in expected default probabilities across states is due to the variation in expected cash-flows. The latter are low in bad states and high in good states, inducing counter-cyclical default probabilities in the BY model. The offsetting effect on expected default rates due to the counter-cyclical variation in risk-premia does not dominate the cash-flow effect.

We therefore conjecture that combining CC pricing kernel with time-varying volatility might be very helpful in capturing spread level and dynamics.[30]

6 Predicted time series of credit spreads

This section shows the 'out-of sample' predicted time series of credit spreads using the two models calibrated to equity data and where we extract the state variable and corresponding equity risk-premia from consumption data, risk-free rate and price-dividend ratios.

6.1 CC model

As discussed earlier, we calibrate the CC model to fit moments of historical equity returns. This procedure produces a set of model parameters, and, in addition, a mapping between the surplus consumption ratio, expected equity premium, and simulated (Baa - Aaa) credit spread. With this help, we can then back out the time series of expected equity premium and credit spread using historical surplus consumption ratios. We then compare the simulated credit spread with its historical counterpart to gauge the success of our calibration.

Because the consumption surplus ratio is the key state variable in CC model, we first plot in Figure 3 the historical consumption surplus ratio and the (Baa - Aaa) spread for the 1919-2004 period. The two time series exhibit a striking inverse image of each other. For example, when the credit spread reaches its peak during the great depression, the consumption surplus ratio also bottoms. The figure suggests that the aggregate credit spread is highly systematic;

[29]We emphasize that BY's original model combines both in an intimate way. Our analysis permits to decompose the relative importance of each channel for credit spreads.

[30]We do not pursue the investigation of adding a time varying default boundary in the BY model. Contrary, to the CC model, because it is not likely to improve the predictions of the BY model. Indeed, the BY model already predicts 'too much' amount of covariation between default probabilities and credit spreads. Introducing a countercyclical default boundary will help raise spreads (by the same mechanism discussed for the CC model) but it will also further increase the covariation between default probabilities and credit spreads. So the model even with a time varying boundary cannot match both level and time series properties of spreads.

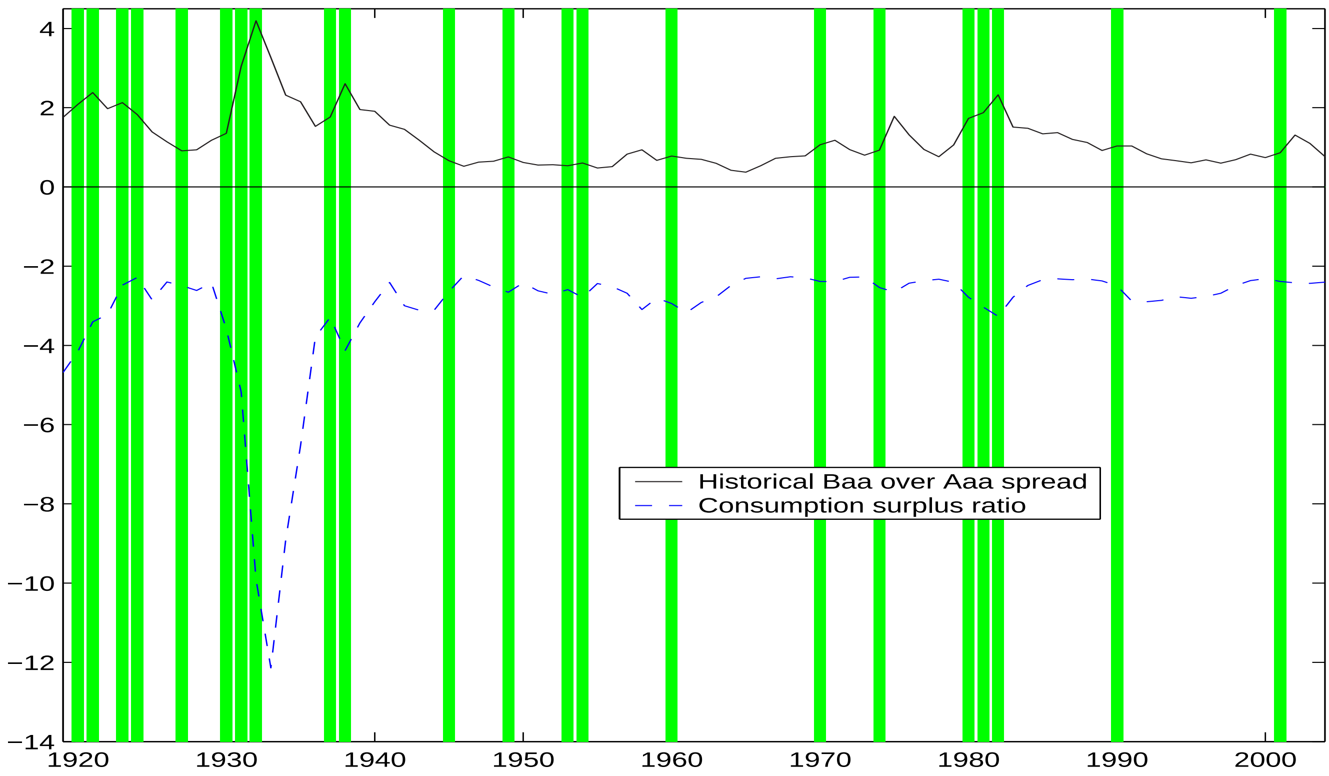

Figure 3: The relation between historical credit spread and consumption surplus ratio.

and that the credit spread backed out using the historical consumption surplus ratio will be able to track its historical counterparty well.

Figure 4 confirms this conjecture. In the upper panel we plot the simulated credit spread and the equity premium as well as the historical credit spread. The equity premium is scaled to fit into the picture. As is clear, the simulated and historical credit spreads exhibit similar time series dynamics throughout the business cycles for the 85 years we examine. In the lower panel we plot the changes of the simulated and historical credit spreads. Again, these changes fit each other successfully. Noteworthy is the fact that the simulated credit spread change tracks its historical counterparty well not only during the great depression period when there are large credit spread movements, but particularly so during the postwar period when there are many relatively small credit spread movements.

Table 8 quantifies what we have learned from the graphs. For the 86 years the simulated credit spread fits well the historical mean (117 bp versus 120 bp), standard deviation (75 bp versus 70 bp), minimum (33 bp versus 37 bp), and the maximum (411 bp versus 420 bp). The

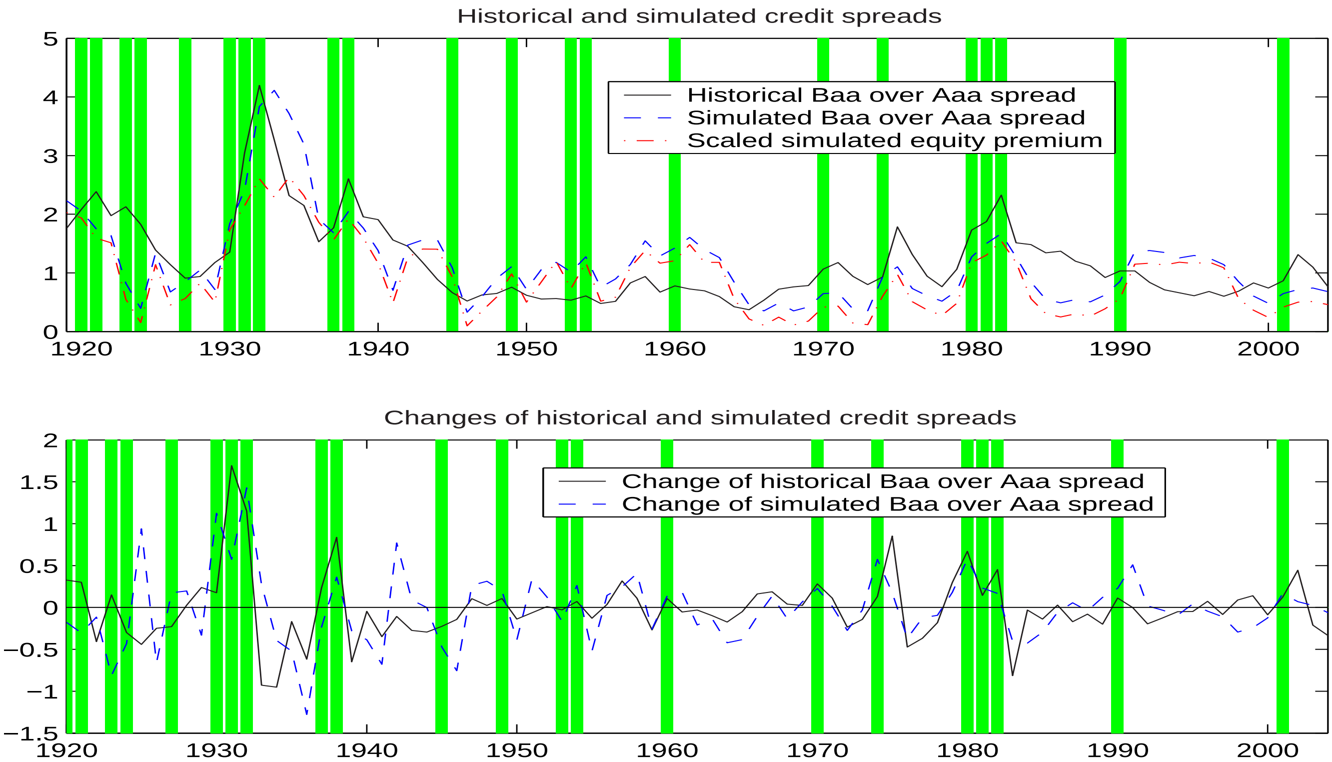

Figure 4: The levels and changes of historical and simulated credit spreads and simulated equity premium.

mean of the simulated and that of the historical credit spread are identical at 90 bp for the postwar period; the corresponding standard deviations are at 38 bp and 39 bp respectively.

For the whole sample period the simulated credit spread is 72% correlated with the historical credit spread; the corresponding number drops to 19% for the postwar period. A careful examination suggests that this drop off of the correlation in the postwar period is driven by the slightly added noise in the last decade of the sample. The simulated and the historical credit spreads still track each other quite well during this period, as shown in Figure 4. Indeed, the correlation of the changes of the simulated and historical credit spreads is 46% for the full sample, but increases to 54% for the postwar period; the dynamics of the two time series intensify in the latter sample. Another noteworthy point is that the historical credit spread correlates much stronger with the simulated credit spread, for both the levels and the changes, than with the price-dividend ratio. It seems to suggest that the consumption surplus ratio might be a superior state variable than the price-dividend ratio in capturing systematic risk -

this result is not surprising given the endogenous nature of corporate dividend policies.

The success of the simulated credit spread to capture both the level and variation of the historical credit spread is nontrivial as it may help to shed light on how closely credit spreads are related to the equity risk premium, a subject still largely unsettled in the current literature. For example, Campbell and Taksler (2004) suggest that idiosyncratic risk can dominate the level and variation of credit spreads. On the other hand, empirical asset pricing papers often use the default spread as an empirical proxy for the equity risk premium. For example, in Jagannathan and Wang (1996), the unobservable market equity risk premium is assumed to be a linear function of the (Baa - Aaa) yield spread alone. Therefore, sorting out whether the yield spread is mainly driven by idiosyncratic or systematic risk seems important. Theoretically, the aggregate yield spread level, which is the average bond yield spread, should not reflect only systematic risk. This is because yield spread is related to total firm asset volatility. A firm with higher idiosyncratic volatility will thus, ceteris paribus, have a higher yield spread. Unlike realized equity returns, this effect will not be diversified away when a portfolio of bonds is constructed. To what extent the idiosyncratic risk components across individual firms have common trends or even factors[31] and thus credit spreads behave differently from equity risk-premia is an empirical issue.

Our results thus confirm two points. First, the historical level and variation of credit spreads can be matched in the CC model successfully where the stochastic component is the variation in aggregate consumption, which is by definition purely driven by systematic risk. Therefore, this suggests that credit spread is likely to be mostly driven by systematic risk. Second, the close resemblance between equity premium and credit spread further suggests that the historical credit spread may be a good proxy for the unobservable equity premium, lending some support this widely adopted practice in the literature. However, the increased discrepancy between the simulated and actual credit spreads in the last decade is also consistent with the analysis of Campbell and Taksler (2004).

6.2 BY model

There are two state variables in the BY model: the persistent component of consumption growth (x) and time-varying consumption volatility (v). Fortunately, there are also two observables from the model: the price dividend ratio and real interest rate. We can thus use the formulas in Footnote 28 to back out x and v for each year. We then simulate (Baa - Aaa)

[31]Note that even though these factors would be common, i.e., market wide, they might still not be priced in that they might be uncorrelated with aggregate consumption - in that sense idiosyncratic.

Panel A: Summary Statistics

	1919-2004				**1946-2004**			
Variable	Mean	Std.	Min	Max	Mean	Std.	Min	Max
SS	1.17	0.75	0.33	4.11	0.90	0.38	0.33	1.67
SEP	7.39	4.99	0.77	21.05	5.55	3.39	0.77	12.37
Spread	1.20	0.70	0.37	4.20	0.90	0.39	0.37	2.33

Panel B: Cross Correlations of Levels

	1919-2004				**1946-2004**			
	CS	SS	SEP	PD	CS	SS	SEP	PD
PD	0.21	-0.30	-0.32		0.22	-0.22	-0.22	
Spread	-0.72	0.71	0.64	-0.31	-0.23	0.19	0.18	-0.13

Panel C: Cross Correlations of the changes

	1919-2004				**1946-2004**			
	CS	SS	SEP	PD	CS	SS	SEP	PD
PD	0.16	-0.25	-0.24		0.35	-0.38	-0.36	
Spread	-0.49	0.46	0.35	-0.20	-0.58	0.54	0.54	-0.28

Table 8: CS is consumption surplus ratio. SS is simulated (Baa - Aaa) spread; SEP is simulated equity premium; Spread is the actual (Baa - Aaa) spread; PD is the actual Price dividend ratio. Panel A reports the summary statistics of the variables. Panel B reports the correlation of the levels of the variables, and Panel C the changes of the variables. Most correlation coefficients are significant at one percent.

spreads conditional on x and v.

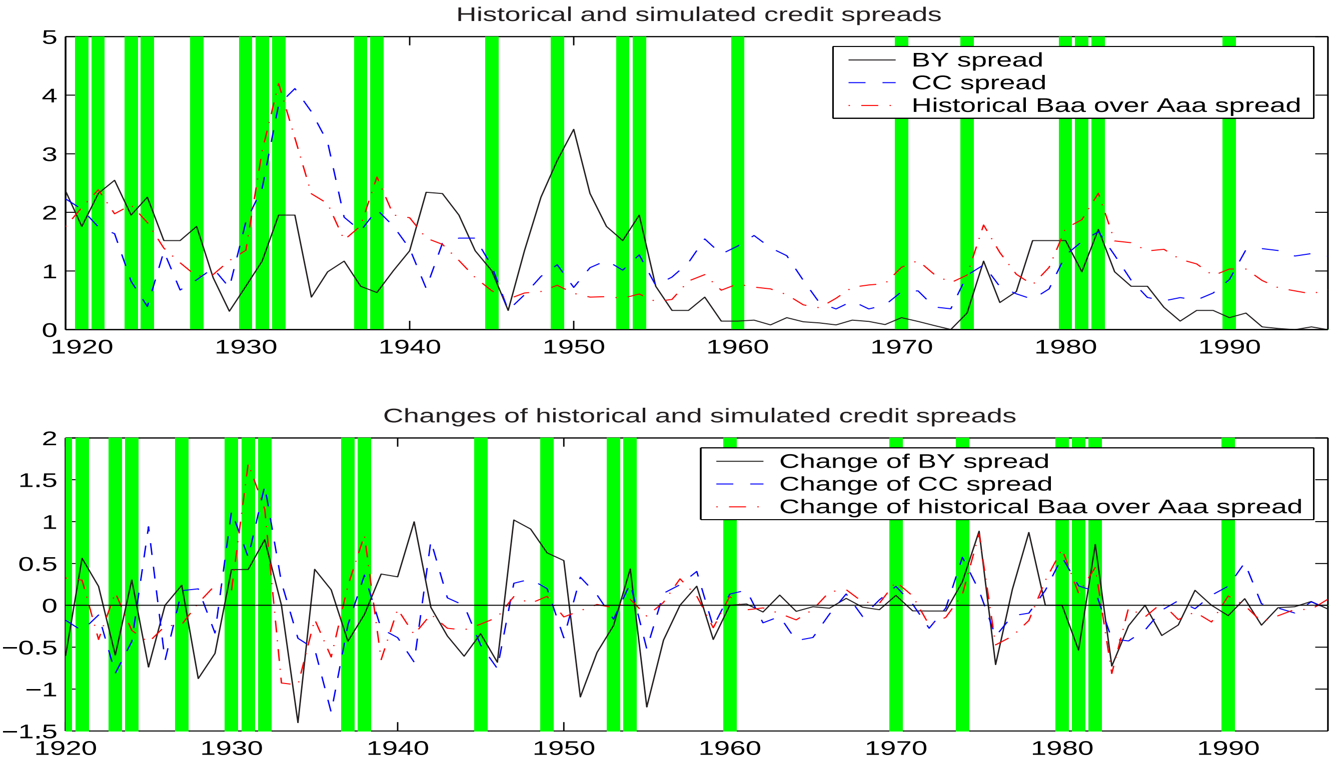

Figure 5: The levels and changes of historical and simulated credit spreads from both model

We plot in Figure 5 the simulated credit spread from the BY model for the 1919-1996 period in comparison with the simulated spread from the CC model as well as the historical spread.[32] Because the BY model uses the PD ratio and the riskfree rate to back out credit spreads, and because we can also use PD ratio to back out credit spreads from the CC model, for comparison, the simulated credit spread from the CC model in Figure 5 is the one backed out from the PD ratio (instead of the surplus consumption ratio).

Noteworthy in the upper panel is that the BY spread misses badly the Great Depression period. While the historical spread reaches its peak in the Great Depreassion period, there are at least three periods during which the simulated credit spreads are higher than those of the Great Depression period. The sharpest contrast is between the Great Depression period and 1950. The historical credit spread during Great Depression is about 6.7 times of that in 1950;

[32]The sample stops in 1996 because no combination of x and v can satisfy the backing-out equations after that.

the BY simulated spread in Great Depression is only 57% of that in 1950. In addition, the BY simulated spreads approach zero since early 1990's.

We observe, however, a similar bad fit from the CC model once the credit spreads are backed out from PD ratio. In other words, the bad fit of the credit spreads in Figure 5, for both the BY model and the CC model, stems from the fact that historical consumption surplus ratio correlates with historical credit spreads much better than historical PD ratio does. For example, the price dividend ratio bottoms in 1950 rather than in the Great Depression, and peaks in the 1990's. Credit spreads backed out using the PD ratio will inherit these properties. Credit spreads backed out using the consumption surplus ratio, as in Figure 4, track the historical spreads much better.

Table 9 reports the statistical properties of the simulated spreads from the BY model. We also report the spreads from the CC model using the surplus consumption ratio. The average BY simulated spread is 98 bp for 1919-1996, compared to 122 bp from the CC model and 123 bp of the historical spred. The high BY simulated spread (though still lower than the historical spread) is driven by several large peaks (e.g., rare events). For example, for 1946-1996, if we exclude 1949 and 1950, then the average BY spread drops from 70 bp to 60 bp.

In Panel B, historical price dividend ratio is 91% correlated with the state variable x, 100% correlated with the state variable v, and 91% with the BY simulated spread. In contrast, the real interest rate is not related to any of these variables. The pattern indicates that, even though we back out x and v from both historical price dividend ratio and the real rate, the price dividend ratio is the main variable that is responsible for the dynamic of the simulated spreads. Given that the consumption surplus ratio relates better to historical credit spreads than the price dividend ratio does, it is not surprising that the CC model provides a much better fit than the BY model. Indeed, as shown in Panels B and C, CC credit spreads are more closely related to historical spreads for both levels and changes than BY spreads do, for both the full sample and the postwar period.

7 Conclusion

In this paper we investigate whether models that are reverse engineered to fit the equity premium can explain the level and time-variation in credit spreads once they are calibrated to equity data. We compare two alternative models: the Campbell and Cochrane habit formation model which explains equity premium with time varying risk-aversion, and the Bansal and Yaron model which explains the equity premium with highly persistent shocks to expected growth and volatility of consumption. Our results suggest that highly time varying risk-premia

Panel A: Summary Statistics

	1919-1996				**1946-1996**			
Variable	Mean	Std.	Min	Max	Mean	Std.	Min	Max
SSCC	1.22	0.77	0.33	4.11	0.92	0.39	0.33	1.67
SSBY	0.98	0.85	0.00	3.42	0.70	0.82	0.00	3.42
Spread	1.23	0.72	0.37	4.20	0.91	0.41	0.37	2.33

Panel B: Cross Correlations of Levels

	1919-2004				**1946-2004**			
	x	v	SSCC	SSBY	x	v	SSCC	SSBY
Real r	0.02	-0.02	0.20	-0.05	-0.11	-0.12	0.36	-0.16
PD	-0.91	-1.00	-0.28	-0.91	-0.89	-1.00	-0.07	-0.90
Spread	0.40	0.43	0.71	0.35	0.15	0.25	0.20	0.09

Panel C: Cross Correlations of the changes

	1919-2004				**1946-2004**			
	x	v	SSCC	SSBY	x	v	SSCC	SSBY
Real R	0.12	0.12	0.21	0.05	0.25	0.30	0.57	0.23
PD	-0.79	-1.00	-0.29	-0.80	-0.74	-1.00	-0.45	-0.76
Spread	0.34	0.30	0.47	0.31	0.43	0.42	0.58	0.42

Table 9: SSCC is the simulated (Baa - Aaa) spread from CC model using historical surplus consumption ratio; SSBY is the simulated (Baa - Aaa) spread from BY model; Spread is the actual (Baa - Aaa) spread; PD is the actual Price dividend ratio; Real r is the real interest rate. Panel A reports the summary statistics of the variables. Panel B reports the correlation of the levels of the variables, and Panel C the changes of the variables.

are essential to explain the level and variation observed in spreads. However, CC's model also generates the inconsistent prediction that default probabilities are pro-cyclical unless the default boundary is allowed to be counter-cyclical. We argue that the latter is consistent with some of the existing empirical evidence (we also offer some empirical support).

Alternatively, combining insights from CC and BY may be necessary to capture adequately the behavior of spreads. Indeed, adding stochastic volatility to the CC model may help solve the pro-cyclical nature of default probabilities. Separately, it may be necessary to reconcile the predictions of the CC model with empirical evidence on the price of the consumption claim. Indeed, in the appendix, we report our estimates of the CC habit formation model, which differ substantially from the results reported in their original paper. It seems that for the CC model to reconcile historical return and volatility patterns of the post-war US economy with the aggregate consumption time series, it may require adding some shocks to growth rate and/or consumption volatility.

References

[1] S. T. B. Acharya, V. and A. Srinivasan. Understanding the recovery rates of defaulted securities. working paper, London Business School, 2005.

[2] A. R. Altman E. and A. Sironi. Default recovery rates in credit risk modelling: A review of the literature and empirical evidence. *Economic Note*, 33:183–208, 2004.

[3] A. R. Altman E. and A. Sironi. The link between default and recovery rates: Effects on the procyclicality of regulatory capital ratios. Working paper, New York University., 2005.

[4] G. Andrade and S. N. Kaplan. How costly is financial (not economic) distress? evidence from highly leveraged transactions that became distressed. *Journal of Finance*, v53 n5:1443–1494, 1998.

[5] R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59:1481–1509, 2004.

[6] F. Black and J. C.Cox. Valuing corporate securities: some effects of bond indentures provisions. *Journal of Finance*, 31:351–367, 1976.

[7] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[8] J. Boudoukh, R. Michaely, M. Richardson, and M. Roberts. On the importance of measuring payout yield: Implications for empirical asset pricing. Forthcoming, Journal of Finance, 2005.

[9] J. Campbell. Stock returns and term structure. *Journal of Financial Economics*, 18:373–399, 1987.

[10] J. Campbell and J. Cochrane. By force of habit: A consumption-based explanation of aggregate stock return behavior. *Journal of Political Economy*, 107:205–251, 1999.

[11] J. Campbell and G. Taksler. Equity volatility and corporate bond yields. *Journal of Finance*, 58:2321–2349, 2003.

[12] J. Y. Campbell and J. Ammer. What moves the stock and bond markets? a variance decomposition for long-term asset returns. *Journal of Finance*, v48 n1 March:3–37, 1993.

[13] N. Chen, R. Roll, and S. Ross. Economic forces and the stock market. *Journal of Business*, 59:383–403, 1986.

[14] J. Cochrane. *Asset Pricing*. Princeton University Press, 2001.

[15] P. Collin-Dufresne and R. S. Goldstein. Do credit spreads reflect stationary leverage ratios? *Journal of Finance*, 56:1929–1957, 2001.

[16] P. Collin-Dufresne, R. S. Goldstein, and J. S. Martin. The determinants of credit spread changes. *Journal of Finance*, 56:2177–2207, 2001.

[17] P. Collin-Dufresne and B. Solnik. On the term structure of default premia in the swap and libor markets. *JF*, 56:1095–1115, 2001.

[18] Y. C. Cosimano, T. and A. Himonas. By force of habit: An exploration of asset pricing models using analytic methods. Working paper, University of Notre Dame., 2004.

[19] A. David. Inflation uncertainty, asset valuations, and credit risk. Working Paper, University of Calgary, 2005.

[20] S. Davydenko and I. Strebulaev. Strategic actions, capital structure, and credit spreads: An empirical investigation. Working paper, London Business School, 2004.

[21] J. Driessen. Is default event risk priced in corporate bonds? *The Review of Financial Studies*, 18:165–196, 2005.

[22] G. R. Duffee. Estimating the price of default risk. *The Review of Financial Studies*, 12:187–226, 1999.

[23] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton: University Press, 1996.

[24] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

[25] D. Duffie and K. Singleton. An econometric model of the term structure of interest-rate swap yields. *Journal of Finance*, LII.no.4:1287–1381, 1997.

[26] E. J. Elton, M. J. Gruber, D. Agrawal, and C. Mann. Explaining the rate spread on corporate bonds. *Journal of Finance*, 56 no. 1:247–277, 2001.

[27] Y. Eom, J. Helwege, and J.-Z. Huang. Structural models of corporate bond pricing. *RFS*, 17:499–544, 2004.

[28] J. Ericsson and O. Renault. Liquidity and credit risk. Working paper, McGill University, 2001.

[29] E. F. Fama and K. R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, v25 n1 November:23–49, 1989.

[30] E. F. Fama and K. R. French. Common risk factors in the returns on stock and bonds. *Journal of Financial Economics*, v33 n1 February:3–56, 1993.

[31] P. Feldhtter and D. Lando. Decomposing swap spreads. Working Paper, Copenhagen Business School, 2005.

[32] G. Gennotte and T. Marsh. Variations in economic uncertainty and risk premiums on capital assets. *European Economic Review*, 37(5):1021–1042, 1993.

[33] R. L. Geske and G. Delianedis. The components of corporate credit spreads: Default, recovery, taxes, jumps, liquidity, and market factors. Working Paper, UCLA, 2003.

[34] M. Grinblatt. An analytical solution for interest rate swap spreads. UCLA Working Paper, 1995.

[35] F. Guvenen. A parsimonious macroeconomic model for asset pricing: Habit formation or cross-sectional heterogeneity? Working paper, University of Rochester, 2003.

[36] J. Huang and M. Huang. How much of the corporate-treasury yield spread is due to credit risk? a new calibration approach. Working paper Stanford University, 2000.

[37] R. Jagannathan and Z. Wang. The conditional capm and the cross-section of expected returns. *Journal of Finance*, 51:3–53, 1996.

[38] P. E. Jones, S. P. Mason, and E. Rosenfeld. Contingent claims analysis of corporate capital structures: An empirical investigation. *Journal of Finance*, v39 n3 July:611–25, 1984.

[39] D. B. Keim and R. F. Stambaugh. Returns and volatility of low-grade bonds: 1977-1989. *Journal of Financial Economics*, v17 n2 December:357–90, 1986.

[40] H. E. Leland. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance*, 49:1213–1252, 1994.

[41] F. A. Longstaff and E. Schwartz. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–821, 1995.

[42] T. Marsh and R. Merton. Dividend variability and variance bounds tests for the rationality of stock market prices. *American Economic Review*, 76:483–448, 1986.

[43] R. C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.

[44] R. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436, 1981.

A Calibration of the CC model

We calibrate the CC model and report the results in Table 9. Columns 2 and 3 correspond to the results in CC, and the last two columns are from the historical data postwar annual data and long term S&P 500 data respectively). A comparison of Column 2 and 8 indicates the success of the original CC study. For example, they choose parameters to match postwar annual consumption growth rate (1.89% versus 1.89%), the volatility of consumption growth (1.50% versus 1.50%), the real riskfree rate (0.94% versus 0.94%), and the Sharpe ratio (0.43 versus 0.43). The 'out-of-sample' success of their model comes from the proximity of the following three statistics that are not calibrated: the average equity premium (6.64% versus 6.69%) and the volatility of equity premium (15.20% versus 15.73%), and the price dividend ratio (18.30 versus 24.70). Overall, CC show that an i.i.d. consumption process, with the aid of an external equity habit, can go a long way in interpreting the equity premium puzzle.

We note that the above simulated statistics in CC are for the consumption claim. The third column in Table A1 reports the similar statistics for the dividend claim. Except for the fact that the Sharpe ratio is a bit low (0.32), other statistics are reasonably close to actual postwar data.

We replicate the results in CC and it appears to us that they do not give enough consideration to the extreme bad states when conducting simulations. Because the representative agent is very risky averse during bad times, it turns out that these extreme bad states, rare as they are, are important in asset valuations. We confirm this point by using both the recursive scheme and the Monte Carlo method (Equations (12) and (13) in the text). For example, in Columns 4 and 5, we use the same parameters and the recursive scheme as in CC, but with more considerations to the extreme bad states. By chance, we find that $\gamma = 2$ still closely matches the observed Sharpe ratio at 0.43 for the consumption claim. However, the level and the standard deviation of the excess return are much lower at 3.70% and 8.56% respectively, only about half of the historical values. Furthermore, the implied price/dividend ratio is much higher (at 36.60) than historical data. The fitting for the dividend claim is equally bad. We vary the parameters and find it difficult to fit the historical statistics simultaneously. Hence, it seems that the CC model cannot capture the post war security returns when calibrated to consumption dynamics.

Because the historical evidence is actually on the dividend claim (instead of the consumption claim), we vary parameters such that we can fit the dividend claim. Indeed, with the parameters as in Table 3, we can fit the historical data reasonably well. As shown in Columns

	CC Est. $\gamma = 2.00$		Our Est: $\gamma = 2.00$		Our Est: $\gamma = 2.45$		Historical Data			
							Cons		Div	
Statistics	Cons	Div	Cons	Div	Cons	Div	Postwar data	Long data	Postwar data	Long data
Calibration										
E[dc]	1.89	1.89	1.89	1.89	1.89	4	1.89	1.72	6.21	3.19
sigma(dc)	1.5	11.2	1.5	11.2	11.2	8	1.5	3.32	11.2	12.93
E[rf]	0.94	0.94	0.94	0.94	0.94	0.94	0.94	2.92	0.94	2.92
E(r-rf)/sigma(r-rf)	0.43	0.32	0.43	0.25	0.49	0.44	NA	NA	0.43	0.22
Out-of-sample										
E(r-rf)	6.64	6.52	3.7	3.73	3.9	7.3	NA	NA	6.69	3.9
sigma(r-rf)	15.2	20	8.56	15.04	8.03	16.7	NA	NA	15.73	17.96
exp[E(p-d)]	18.3	18.7	36.6	35.68	33.98	24.1	NA	NA	24.7	21.16
sigma(p-d)	0.27	0.29	0.14	0.16	0.13	0.21	NA	NA	0.26	0.27

Table 10: The historical data are obtained from CC whenever available. The data that are absent from CC include the dividend growth rates for both the postwar and the long periods. We use the dividend time series (obtained from Robert Shiller) to construct these statistics. The first two columns are orignal estimates from CC. The third and fourth columns are our estimates, using the same parameters as in CC, but allow for wider simulation boundaries. The fifth and sixth columns include our estimates that intend to fit the dividend side of the postwar historical data, in which case the parameter inputs are the same as in Table 3 for the dividend process.

7 and 10, the simulated and actual Sharpe ratios are 0.44 versus 0.43; the corresponding equity premiums are 7.30% versus 6.69%; the standard deviations of the equity return are 16.70% versus 15.73%; the price-dividend ratios are 24.1 versus 24.7; and the standard deviations of the log price-dividend ratio are 0.21 versus 0.26.

We find that the other properties generated in CC are still very robust. For example, in Table 10 we study whether the price/dividend ratios can still predict future equity return, another important fact recovered in many classic empirical studies. We find that, similar to CC, current price/dividend ratio or price/consumption ratio can indeed predict future returns. The predictive power increases with investment horizon. This result is important because it indicates that the primary property of the habit formation model is robust, namely, that the representative agent's risk aversion increases in the bad states, and hence demands a higher equity risk premium in order to bear such risk.

B Discrepancy with the results of Campbell and Cochrane (1999)

As noted in the text, equations (31) and (32) suggest two numerical schemes for determining the price-dividend ratio as a function of s. In particular, equation (31) can be approximated by using a recursive scheme which is iterated until the program converges to a self-consistent

CC parameters (gama=2) but relaxing distribution boundary restrictions								
Horizon	Cons. Claim		Div. Claim		Postwar data		Long data	
(Years)	10*coef.	R^2	10*coef.	R^2	10*coef.	R^2	10*coef.	R^2
1	-1.54	0.06	-1.65	0.03	-2.6	0.18	-1.32	0.04
2	-2.87	0.12	-3.05	0.05	-4.25	0.27	-2.77	0.08
3	-3.97	0.17	-4.23	0.07	-5.37	0.37	-3.48	0.09
5	-5.87	0.25	-6.42	0.1	-9.02	0.55	-6.04	0.18
7	-7.24	0.3	-8.13	0.12	-12.11	0.65	-7.54	0.23
10	-8.92	0.37	-10.43	0.15	-16.37	0.8	-9.25	0.24

Our model								
Horizon	Cons. Claim		Div. Claim		Postwar data		Long data	
(Years)	10*coef.	R^2	10*coef.	R^2	10*coef.	R^2	10*coef.	R^2
1	-1.56	0.06	-1.67	0.04	-2.6	0.18	-1.32	0.04
2	-2.89	0.12	-3.08	0.08	-4.25	0.27	-2.77	0.08
3	-4	0.16	-4.25	0.11	-5.37	0.37	-3.48	0.09
5	-5.91	0.24	-6.36	0.16	-9.02	0.55	-6.04	0.18
7	-7.29	0.3	-7.94	0.2	-12.11	0.65	-7.54	0.23
10	-8.98	0.37	-9.97	0.24	-16.37	0.8	-9.25	0.24

Table 11: We regress future cumulative long-horizon returns on the price/dividend (or price/consumption) ratio. In Panel A we use the same parameters as in CC but relaxing the restriction on the distribution boundaries. In Panel B we use the same parameters as in Table 3 in the text, which are meant to fit the historical evidence on the dividend claim. As is clear, in each panel the predictive power of the price/dividend ratio is retained regardless of the parameter choices.

solution. Alternatively, Monte-Carlo methods can be used to estimate the price-dividend ratio via equation (32). Unfortunately, both methods are vulnerable to certain types of numerical error, as we now demonstrate.

Recursive schemes are vulnerable to numerical errors because in order to implement them, one must discretize the state variable s and then specify some set of lower and upper bound cut-offs. While the discretization process can potentially introduce errors, these errors can usually be controlled by obtaining estimates for different sizes of increments, and then extrapolating back to the limit $ds \to 0$. In contrast, specifying finite values for the cut-off values $\underline{s}$ and $\overline{s}$ can lead to errors that are more difficult to control. Typically the best that one can do here is to see whether the estimated solution is converging as one increases the range $s \in (\underline{s}, \overline{s})$. We note that CC use the recursive method to estimate the price-dividend ratio. Using their own code, we demonstrate below that the estimated solution changes dramatically as $\underline{s}$ is lowered. [33]

Monte Carlo methods are also vulnerable to numerical error, as can be seen from this illustrative example. Assume a random variable $\tilde{X}$ can take on only two values: 10^{12} with the probability $p\left(\tilde{X} = 10^{12}\right) = 10^{-12}$, and zero. Clearly, the true expected value is one, and

[33] Cosimano, Chen, and Himonas (2004) also find a similar result.

the variance is approximately 10^{12}. However, if one simulates only, say, 10^6 paths, then it is very likely that the simulation will generate a sample mean and sample variance of zero! This example emphasizes that the implied standard error might be a poor indicator of whether or not a sufficient number of sample paths have been run. It is worth noting that Monte Carlo approaches do not bias the mean estimate, but they may bias the estimate for the 'typical' or 'median' path. As an example more relevant to the current situation, if a given random variable is normal $\tilde{X} \sim (0,1)$, and one attempts to estimate $\mathrm{E}\left[e^{aX}\right]$, the number of sample paths necessary for convergence increases significantly with a, even though the implied standard error may not give an indication that convergence has not been obtained. As we demonstrate below, it is the 'long-tail' of the pricing kernel Λ which makes the Monte Carlo approach fail here, at least for reasonable numbers of sample paths. Indeed, the price of a security can be calculated as the expectation of the product of the pricing kernel and the state-dependent cash flows.

$$
\begin{aligned}
P &= \mathrm{E}\left[M\,X\right] \\
&= \int_{-\infty}^{s_{max}} ds_T\, \pi\left(s_T \,\middle|\, s_t \right) M(s_T)\, X(s_T) \\
&= \int_{-\infty}^{s_{max}} ds_T\, \pi\left(s_T \,\middle|\, s_t \right) e^{-\alpha(T-t)} e^{-\gamma\left[(s_T+C_T)-(s_t+c_t)\right]}\, X(s_T).
\end{aligned}
\tag{B.75}
$$

We claim that, due to the long tail, finite sample Monte Carlo estimation methods bias downward those probabilities $\pi\left(s_T \,\middle|\, s_t \right)$ for large negative s_T, where marginal utility is higher, and bias upward those probabilities $\pi\left(s_T \,\middle|\, s_t \right)$ for less negative values of s_T, where marginal utility is lower, leading to a downward biased estimate for the price (or equivalently, the price-dividend ratio) of a security for the 'typical' Monte Carlo simulation.

Thus, the problem that the Monte Carlo approach runs into can be traced back to the long tail of the distribution of s. However, for this particular model, there is a simple way to circumvent this difficulty. Indeed, here we demonstrate that by transforming from the 'historical measure' to the 'risk-neutral measure', the problems associated with the 'long-tail' of the pricing kernel are eliminated. Indeed, we find that the Monte Carlo approach under the Q-measure generates a solution which is very well behaved, even for a relatively low number of sample paths. Specifically, we can re-write equations (31)-(32) as

$$
\left(\frac{P(t)}{C(t)}\right) = \mathrm{E}_t^Q\left[\left(\frac{1}{R_{t,t+1}}\right)\frac{C(t+1)}{C(t)}\left(1+\frac{P(t+1)}{C(t+1)}\right)\right] \tag{B.76}
$$

$$
= \sum_{j=1}^{\infty} \mathrm{E}_t^Q\left[\left(\frac{1}{\prod_{m=0}^{j-1} R_{(t+m),\,(t+m+1)}}\right)\frac{C(t+j)}{C(t)}\right]. \tag{B.77}
$$

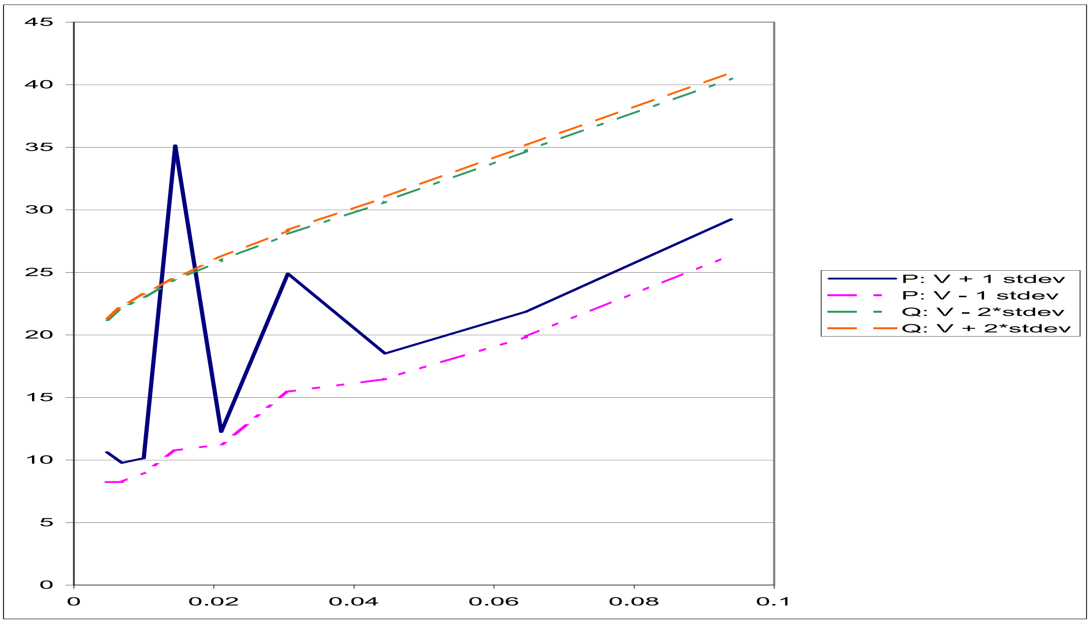

Figure 6: Estimation of the price-consumption ratio using Monte Carlo methods under both the P and Q-measures. Parameter values are $g = 0.0189$, $\gamma = 2$, $\sigma = 0.015$, $\kappa = 0.138457$, and $r = 0.0094$. 100,000 sample paths are used for the P-measure estimates, whereas only 10,000 sample paths are used for the Q-measure estimates.

The reason why this transformation is useful is that, even though the pricing kernel has a long tail, its expectation (for values of $s < s_{max}$) generates a one-period risk-free rate that is constant.

We demonstrate the advantage of this transformation by estimating the price-consumption ratio using Monte Carlo methods under both the P-measure, and the Q-measure. As demonstrated in Figure (6), we see that the Q-measure Monte Carlo estimation using only 10,000 sample paths generates a smooth, monotonic function whose standard error is so low that we need to add plus-or-minus two standard errors in order for the two curves to be distinguishable. In contrast, even after 100,000 sample paths under the P-measure, the estimates are still very noisy and non-monotonic. Interestingly, we note that the P-measure estimates are rather similar to those obtained by CC, whereas the Q-measure estimates generate significantly different values, especially for low values of s. Below, we argue that the Q-measure estimates are in fact

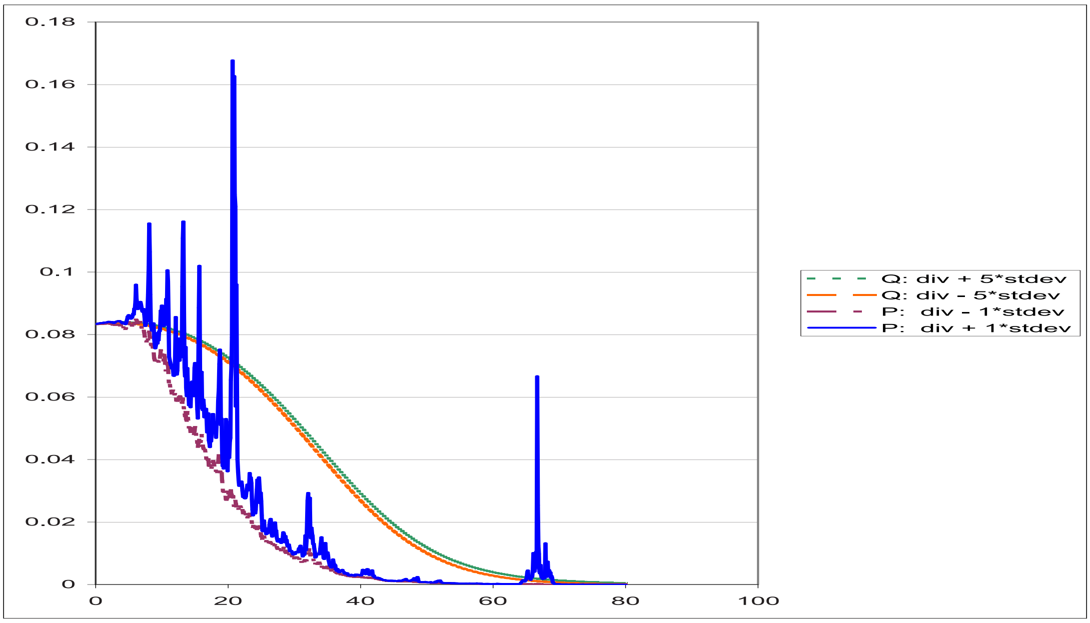

Figure 7: Estimation of the value of the individual consumption claims using Monte Carlo methods under both the P and Q-measures. Parameter values are $g = 0.0189$, $\gamma = 2$, $\sigma = 0.015$, $\kappa = 0.138457$, and $r = 0.0094$. 100,000 sample paths are used for the P-measure estimates, whereas only 10,000 sample paths are used for the Q-measure estimates.

the true price-consumption ratio in the CC economy.

One advantage of the Monte Carlo approach is that one may price each dividend claim separately. In Figure (7), we plot the P- and Q-measure estimates for the prices of the individual consumption claims as a function of maturity. We find that the Q-measure estimate generates a very smooth function. Indeed, we needed to add plus-or-minus five standard errors to create two lines that were distinguishable. In contrast, the P-measure estimates generate a rather noisy function, even though we used ten-times more sample paths to estimate it.

Of course, smoothness alone does not guarantee that our Q-measure estimates are accurate. In order to provide more convincing evidence that the Q-measure estimates are the solution to the CC model, here we examine a one-period model in a CC framework to price both the claim to \$1, that is, a risk-free bond, and a claim to next period's consumption $C(t+1)$. The advantage of a one period model is that, since the CC model is specified to have a log-normal

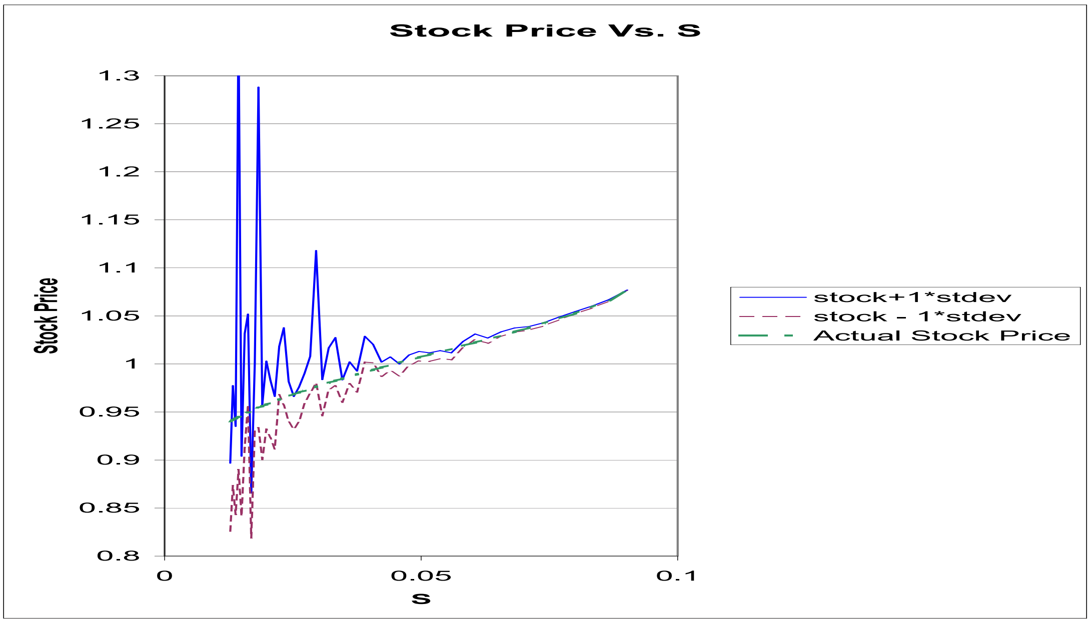

Figure 8: P-estimation for the value of the claim to one period's consumption. The three curves are the true value, and the estimated value plus or minus one standard error. Parameter values are $g = 0.0189$, $\gamma = 2$, $\sigma = 0.015$, $\kappa = 0.138457$, $r = 0.0094$ and $dt = 10$. 100,000 sample paths are used for the P-measure estimates.

distribution, the true solution is known in analytical form.

In Figure 8 we plot P-estimates (plus or minus one standard error) for the value of the one period consumption claim, along with the actual value. We do not plot the Q-value estimates, because even plus or minus 10 standard errors would generate a curve indistinguishable from the actual answer. Admittedly, both the P- and Q-estimates give excellent answers when the time increment is on the order of one month. However, in order to capture the point that pricing a security implies pricing accurately the dividend claim for many decades, we chose $dt = 10$ years for this figure. We emphasize, however, that the one period model is exact regardless of the time increment. This finding strongly suggests that the Q-measure Monte Carlo estimates given above in Figure 6 are provide excellent numerical estimates of the actual value.

Analogously, we priced the one-period bond as a function of s under the P-measure. Note

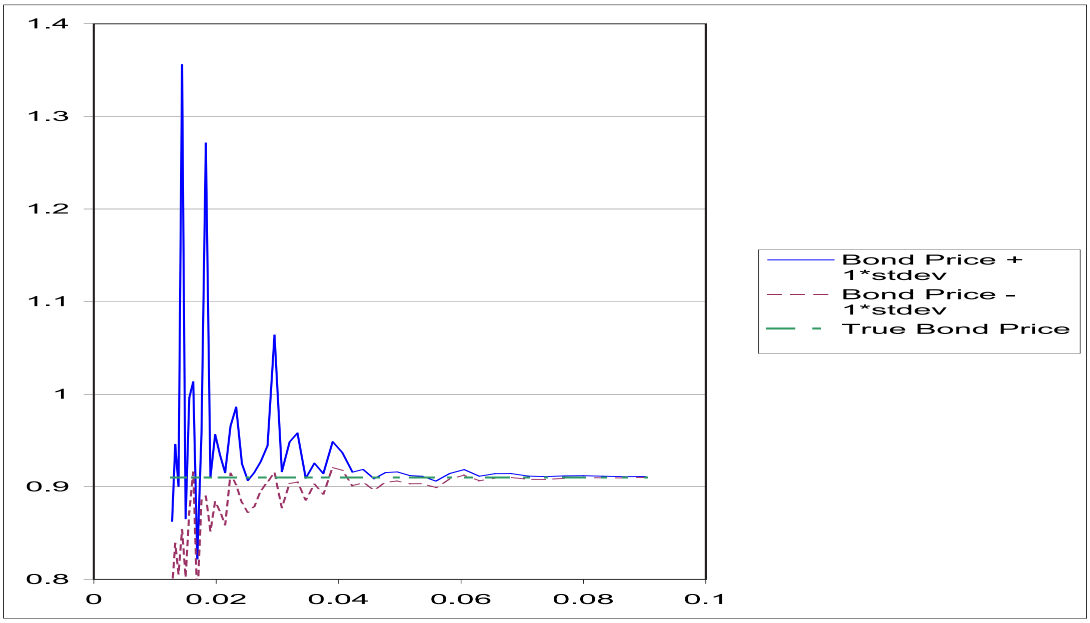

Figure 9: P-estimation for the value of the one-period riskless bond. The three curves are the true value, and the estimated value plus or minus one standard error. Parameter values are $g = 0.0189$, $\gamma = 2$, $\sigma = 0.015$, $\kappa = 0.138457$, $r = 0.0094$ and $dt = 10$. 100,000 sample paths are used for the P-measure estimates.

that under the Q-measure, the solution is exact by construction. Once again, we see that for large negative values of s the Monte Carlo estimates are poor.

Finally, in Figure 10 we plot the estimated value of the price-consumption ratio using the computer program of CC, available on their web page. In the figure we plot the price/dividend ratio and price/consumption ratio using both the cutoff points as in CC and lower cutoff points. It is clear that, in the original version of CC, the level of the ratios are much lower than what we obtain through the Monte Carlo simulation with the risk-neutral pricing formula. On the other hand, when much lower cutoff points are used, the ratios are much higher and very similar to those obtained for the risk-neutral Monte Carlo estimates. This comparison indicates that, when the proper cutoff points are used, the two methods will yield similar results.

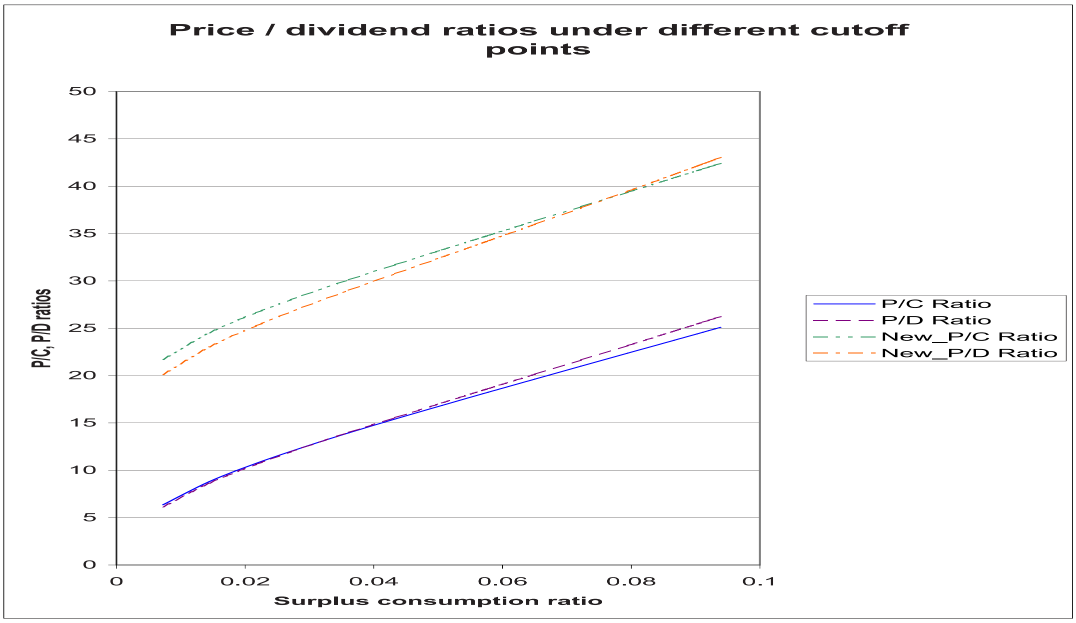

Figure 10: P and Q-estimations using the iterative procedure of CC. Two different values are used for $\underline{s}$: -5 and -50.