The time-varying equity premium and the S&P 500 in the twentieth century

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article was published in the Journal of Financial Research [© The Southern Finance Association and the Southwestern Finance Association] and the definitive version is available at: http://dx.doi.org/10.1111/j.1475-6803.2011.01288.x

Metadata Record: https://dspace.lboro.ac.uk/2134/15033

Version: Accepted for publication

Publisher: Wiley © The Southern Finance Association and the Southwestern Finance Association

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/

Mark C. Freeman
Bradford University

The author is grateful for comments and suggestions made by participants at the Inquire UK, Equity Premium Puzzle (Exeter), and Financial Management Association (Europe) conferences as well as seminar participants at the Universities of Auckland, Cambridge, Leeds, Loughborough, Nottingham, Warwick, and York. He thanks Winfried Hallerbach for kindly making his data available and John Powell for advice on timing secular bull and bear markets. He also thanks George Bulkley, Steve McCorriston, Abhijit Sharma, Mark Tippett, and an anonymous referee for detailed comments. All errors, sadly, remain the author's own.
Abstract

I present a new hindcast stock market index for the United States over the twentieth century. This is constructed by calibrating a rational asset pricing model that allows for a time-varying equity premium driven by heteroskedasticity in consumption growth. By incorporating this variation in risk, the mean squared error of the generated index series, when compared to the observed levels of the S&P 500, is significantly reduced. The model also explains well the broad magnitudes and timings of the major bull and bear markets of the twentieth century, particularly before 1973, and the excess volatility puzzle is largely resolved.

JEL Classifications: E21, G12, N21, N22.
I. Introduction

In this article, I argue that much of the low frequency behavior of the S&P 500 during the twentieth century can be explained through variations in the ex ante equity premium driven by changes in the consumption risk faced by stockholders. In particular, the model has high explanatory power for the period up to the early 1970s. The model performs less well for the stock market crash of 1973–1974, but average returns from the mid-1960s to the end of the century are well explained by the model.

This article joins a stream of literature that examines whether observed low frequency market movements are consistent with a model of discounted dividends and rational investor behavior, where perhaps the most prominent puzzle has been the excess volatility anomaly of Shiller (1981) and LeRoy and Porter (1981). To add to this debate, it is necessary to consider how the expected future growth rate of dividends and the cost of equity capital vary over time. In this article, I focus on the second of these effects and my work is particularly related to the study by Lettau, Ludvigson, and Wachter (2008). There has been a growing body of evidence, starting with Blanchard (1993) and subsequently examined by Jagannathan, McGrattan, and Scherbina (2000), Claus and Thomas (2001), and Fama and French (2002), among others, that the ex ante equity premium is considerably lower now than its long-term historic average. Lettau, Ludvigson, and Wachter (2008) observe that there was a downward shift in consumption volatility in the second half of the twentieth century. As investors learned about this, the ex ante equity premium fell, leading to the very strong stock market performance of the 1990s. This article considers this argument over a much longer time frame by examining whether a changing equity premium, caused by time-varying consumption risk,
helped drive a number of the most significant market movements during the twentieth century.

An insight into the intuition of this article is given in Lettau, Ludvigson, and Wachter (2008) and Bansal, Khatchatrian, and Yaron (2005). They respectively note that GDP growth volatility and consumption growth volatility declined in the post-World War II (WWII) period and that this was associated with a rise in the price-dividend ratio at the time. De Santis (2007) also observes that the ex ante equity premium fell in the post-WWII period, which was a period when macroeconomic risk was falling. However, as stressed by Lettau et al., this apparent reduction in risk may purely result as a consequence of more accurate data measurement after WWII.

To overcome this problem, I use income data from households with highest income to estimate changes in the macroeconomic risk faced by stockholders from just after World War I to 1998. These data are available from Piketty and Saez (2003), and in contrast to aggregate consumption, consistency of measurement error is no longer an issue. This enables me to look at the relation between macroeconomic risk and market returns over a much longer horizon than most previous studies without the concern that the results are driven by measurement error effects. This advantage is also exploited by Buranawityawut, Freeman, and Freeman (2006).

This article follows a similar method to Lettau, Ludvigson, and Wachter (2008). Consumption and income growth volatility are tested for a structural break over the whole data set. Although such a transition become clear to the econometrician with the benefit of hindsight, investors could not have perfectly observed it contemporaneously. Therefore, within a regime-switching environment, I estimate when investors would have
learned about changes in macroeconomic risk. Contingent on this learning process, “rational” asset prices can be calculated and compared with actual market behavior.

The model is able to capture much of the low-frequency behavior of the S&P 500 over the twentieth century. When looking at the index in levels, the mean squared error is significantly reduced, in both an economically and statistically significant sense, when the model incorporates a time-varying equity premium. The timings and magnitudes of the major booms and busts of the stock market, particularly before the early 1970s, can also be significantly better explained by allowing the equity premium to vary. The 1966–1981 bear market remains anomalous, but the 1982–1999 period brings the market back to levels consistent with the model. The retrodicted annual standard deviation of market returns over the century is not different from the realized value at standard levels of statistical significance, largely resolving the excess volatility puzzle at this frequency.

II. Changes in Aggregate Consumption Risk

Income and Consumption Data

My objective is to provide a rational market explanation for the low frequency behavior of the S&P 500 during the twentieth century. I place particular emphasis on explaining the main stock market booms and busts as major changes in perceived consumption risk should significantly affect market levels within the framework of this article. Bordo and Wheelock (2007) also call on a macroeconomic argument to explain such events, but their focus is on changes in monetary policy. Barsky and De Long (1990) and White (1990) provide some qualitative perspectives on this issue. Other empirical studies focusing on specific episodes include: (1) demographic arguments for the post-WWII and
1990s bull markets (e.g.; Bakshi and Chen 1994; Poterba 2004); (2) the arrival of new technology explaining the bear market of the 1970s and subsequent bull market of the 1980s (Hobijn and Jovanovic 2001), and (3) dividend growth rate uncertainty accounting for the Nasdaq “bubble” (Pástor and Veronesi 2006).

In this article, the principle driver of low frequency market returns is persistent changes in the equity premium. The ex ante cost of equity is assumed to be driven by the conditional volatility of aggregate consumption growth. This assumption is motivated by the Consumption Capital Asset Pricing Model (CCAPM). Let $R_r, R_f$ denote, respectively, the simple real return to a risky and a risk-free asset over the interval $[t-1, t]$. The CCAPM states that the ex ante risk premium, $E_{t-1}[R_r - R_f]$ is approximated by:

$$E_{t-1}[R_r - R_f] \approx \gamma \rho_{r \sigma} (R_r, r_c) \sigma_{t-1} (R_r) \sigma_{t-1} (r_c)$$

where $\gamma$ is the coefficient of relative risk aversion; $r_c = \ln(c_t/c_{t-1})$; $c_t$ is aggregate per-capita real consumption; $\rho(\cdot, \cdot)$ is the correlation coefficient; and $\sigma(\cdot)$ is the standard deviation operator. As $\sigma_{t-1} (r_c)$ changes, so does the equity premium, ceteris paribus.

To look for empirical evidence of changes in aggregate consumption risk, aggregate real consumption for nondurables and services data are taken directly from Robert Shiller’s website (http://www.econ.yale.edu/~shiller/data.htm). Initial,  

---

1 There is a significant literature that considers the role that changing conditional market volatility might play in determining the equity premium; see, for example, Kim, Morley, and Nelson (2004, 2005) and the references therein. Gordon and St-Amour (2000), amongst others, look at time variation in the coefficient
unreported, exponentially weighted moving average (EWMA) estimates of conditional variance clearly demonstrate a one-off major decline in risk occurring in the immediate aftermath of WWII. However, as stressed by Lettau, Ludvigson, and Wachter (2008), this feature may result purely from a reduction in measurement error.

To eliminate this potential problem, the annual income data described in detail in Piketty and Saez (2003) are also examined from 1913 to 1998. Although it is known that income is a poor proxy for consumption, I show in the Appendix that drawing inferences about the equity premium from income growth volatility is robust within a permanent income hypothesis setting. The main advantage of this approach is that it enables us to look at the relation between macroeconomic risk and market returns over a much longer horizon than previous studies without the concern that the results are driven by measurement error effects. As Piketty and Saez (2003, page 12) note: “The fact that the top shares are very smooth after 1945 and bumpy before is therefore not an artefact of an increase in the accuracy of the data (in fact, the data are more detailed before World War II than after), but reflects real changes in economic conditions.”

Piketty and Saez (2003) take income from annual tax return statistics compiled by the IRS for different tax units (married couples and single adults with dependents). Data are for all income shown on tax returns, gross of personal tax, and excluding capital gains. Average incomes by fractiles in real terms for each year are reported. P90-95, for example, represents the average income from households who were in the top 10%, but not the top 5%, of households by income in that tax year. Descriptive statistics for the logarithmic growth in real income for each income fractile and aggregate real per-capita
consumption are given in Table 1.

Unreported EWMA estimates for the conditional standard deviation of logarithmic growth in real income for various income fractiles again provide evidence of a “high risk” state in the pre-WWII years and a “low risk” state after the middle of the 1960s. In this case, though, for the highest income households (P99–99.5 and above), there is an apparent reversion to the high risk state in the late 1980s. This period saw a significant relative redistribution of wealth with those with highest incomes doing exceptionally well. For example, in the years 1987–1988 fractile P99.99–100 had a real rise in income of over 76%. The rise in conditional volatility in the late 1980s therefore appears to reflect a permanent change in the level of income rather than its risk. This view is supported by the facts that: (1) there is no observable increase in risk at this time for lower income households. It seems unlikely that a source of economic risk should emerge that would only affect the rich but have no significant impact on the poor, (2) from 1989–1998, income growth of the highest fractiles was relatively smooth, and (3) 1987–88 were very strong years for the stock market, suggesting that investors viewed this period positively.

Identifying Regime Shifts in Consumption Risk

I take a similar approach to identifying changes in consumption and income growth volatility to that described in Lettau, Ludvigson, and Wachter (2008). First, structural
break in variance tests are run to help formalize the stylized results presented above; that macroeconomic risk was substantially higher before WWII than afterwards. I then turn to a regime-switching environment to consider how investors might have learned in real time about changes in risk.

For many of the income series used in this study there are only approximately 70 observations. As Lettau, Ludvigson, and Wachter (2008) note, the estimation of regime-switching models based on the method of Hamilton (1989) requires a large number of data points. As a consequence, alternative estimation methods are used for the main calibrations in this article, which are subsequently supported by the use of regime-switching MSVAR models and detailed sensitivity analysis.

To retrospectively identify a single structural change in variance, I use the method of Inclán and Tiao (1994). This technique has been used in a finance context by, for example, Aggarwal, Inclan, and Leal (1999) and Wang and Moore (2009). I also use the procedure of Chen and Gupta (1997), which has similar power. These tests are applied to the aggregate consumption series and eight income series in Table 1. For log consumption growth the time interval is 1890–2000. For income series P0-90, P90-95 and P95-99 the time period is 1918–1998. This is the total sample in Piketty and Saez (2003). For all other income series, the time interval is 1914–1986, apart from P95–100 which starts in 1918. I truncate these series in 1986 to avoid the “false” apparent rise in risk at the end of the 1980s discussed in the previous subsection.

Panel A of Table 2 presents evidence for a single structural change in volatility based on the real logarithmic growth of each consumption and income series. In each case, $p$-values are given associated with rejecting the null of no structural break.
However, as shown in Table 1, there is considerable autocorrelation in the growth rate of many of the series. As both the Inclán and Tiao (1994) and Chen and Gupta (1997) tests assume that the data have zero autocorrelation and a constant mean, residuals from Autoregressive Moving Average (ARMA) models for the real log growth rate of each series are also tested for structural breaks in volatility and the results are presented in Panel B of Table 2. Brief details of the ARMA model used in each case are given in Table 1.

[Insert Table 2 around here]

In all cases in Panels A and B of Table 2, the null hypothesis can be rejected with statistical significance of less than 1%. Although there are some disagreements about the exact timing of the break, all series and both tests place it within the interval 1936–1959.

Panel C of Table 2 provides information on the standard deviation of the ARMA residuals before and after the identified date of the structural break, which is taken to be the average of the two dates identified by the Inclán and Tiao (1994) and Chen and Gupta (1997) tests based on these data. Although taking an average in this way is somewhat ad hoc, the low sensitivity of the main results to this assumption is discussed below. The sharp decline in risk between the first and second halves of the century is clearly revealed.

Panel D of Table 2 reports the structural breaks in volatility identified by the MSVAR package, written by Krolzig in Ox, that uses a maximum likelihood EM
algorithm. This is run on the ARMA residuals using the MSIAH(2)-AR(0) model. These results support the earlier findings. For all series except P95–99 the findings are similar to the previous three panels. For P95–99, the MSVAR approach identifies multiple structural breaks in volatility.

In this article, following Veronesi (1999), it is assumed that investors cannot perfectly identify the underlying state that prevails at the time. To consider how investors would have drawn inferences about the current state, I turn to a Bayesian learning regime-switching model. The estimation technique is closely related to that described by Moore and Schaller (2002) and Brandt, Zeng, and Zhang (2004). Investors know that the ARMA residuals of the real logarithmic growth rate for any given income or consumption series, $v_{ct}$, are described by:

$$v_{ct} = \sum_{s=1}^{2} (\mu_s + \epsilon_{st}) \delta_{st}$$  \hspace{1cm} (2)

where $\mu_s$ denotes the mean residual of income or consumption growth in state $s$; and $\epsilon_{st} \sim N(0, \sigma^2_{st})$. There are assumed to be two states with indicator dummy variable $\delta_{st} \in \{0,1\}$, where state 1 (2) is the high (low) risk state and $\delta_{2t} = 1 - \delta_{1t}$ for all $t$. Investors also know that the transition probability matrix is given by:

$$\begin{bmatrix}
    \zeta & 1 - \zeta \\
    1 - \zeta & \zeta
\end{bmatrix}$$  \hspace{1cm} (3)

It is assumed that investors have always known the five parameter values $\mu_s$, $\sigma^2_s$, and $\zeta$. For estimation purposes $\mu_s$ and $\sigma^2_s$ are taken from the mean and variance of $v_{ct}$ before and after the identified structural break date given in Panel C of Table 2, where the values of $\sigma_s$ are also reported. All values of $\mu_s$ are close to zero as expected given
that $\nu_{ct}$ are residuals.

For this Markov switching process, the expected time of first transition from one state to the other is $\sum_{t=0}^{\infty} \zeta^{t-1}(1-\zeta) = (1-\zeta)^{-1}$. For the base calibration the expected time to the next change in state is 50 years, which appears consistent with one major shift in macroeconomic volatility over the last century, setting $\zeta = 98\%$. This value is also supported by the MSVAR maximum likelihood estimates of the transition parameters. For P95–100, which is the income series used in the baseline calibrations below, the estimated values are $\zeta = 0.9712$ for state 1 and $\zeta = 1.000$ for state 2. For consumption data the values are $\zeta = 0.9802$ and $\zeta = 1.000$ for states 1 and 2 respectively. Although state 2 appears to be absorbing, empirically estimating $\zeta$ in this state when there has only been one observed switch in the data (from state 1 to state 2) is clearly problematic. As a consequence, setting $\zeta = 0.98$ for both states is broadly consistent with these estimates. In the sensitivity analysis described below, other values of $\zeta$ are investigated.

The variables that investors cannot observe are $\delta_{st}$. Therefore they have to draw inferences about these variables. If $\Omega_{t-1}$ denotes the information that investors have available to them at time $t-1$ then $P_{t-1}(s) = \text{Prob}(\delta_{st} = 1|\Omega_{t-1})$; the probability that an investor assigns at time $t-1$ to being in state $s$ at time $t$. As there are only two states, $P_{t-1}(2) = 1 - P_{t-1}(1)$. According to Bayes rule:

$$\text{Prob}(\delta_{st} = 1|\Omega_t) = \frac{P_{t-1}(s)f(\nu_{ct} | \delta_{ct} = 1)}{P_{t-1}(1)f(\nu_{ct} | \delta_{ct} = 1) + P_{t-1}(2)f(\nu_{ct} | \delta_{ct} = 1)}$$

where $f(\nu_{ct} | \delta_{ct} = 1)$ is the probability density function for a normal distribution.
\[ N(\mu_s, \sigma_s^2) \] and:

\[ P_t(s) = \zeta \text{Prob}(\delta_{st} = 1|\Omega_t) + (1 - \zeta)\text{Prob}(\delta_{st} = 0|\Omega_t) \]

For aggregate consumption, the starting probability of being in the high risk state \( P_{1889}(1) = 0.5 \); the analysis is neutral about whether the economy starts in a high risk or low risk state. This setting contrasts with the type of learning process described by Cogley and Sargent (2008). They argue that the aftermath of the Great Depression caused investors to have an exaggerated fear of future depressions. This fear has slowly declined over time to more rational levels, leading to a reduction in the market price of risk. They, therefore, deliberately “inject an initial dose of pessimism” into their framework; something that is not required here as the results presented in this article are highly insensitive to the choice of prior. Within five years, and before any results are reported, the estimated value of \( P_{1894}(1) = 97.99\% \), which is almost the highest possible value of 98%.

For the income series, which all begin after the aggregate consumption series, the starting value of \( P_{\text{start}}(1) \) is set equal to the Bayesian probability from the aggregate consumption data for that year. The probabilities of being in the high-risk state for the aggregate consumption series and three income series are presented in Figure I.

\[ \text{[Insert Figure I around here]} \]

The intuition behind this article is clear from this figure. The consumption data quickly recognize that the start of the twentieth century was a time of high risk. All series then identify a reduction in risk during the 1920s, which helps explain the
magnitude of the 1920s bull market. Perceived macroeconomic risk then rises sharply
during the Great Depression, offsetting the decline in the cost of equity in the previous
decade and amplifying the effect of the bear market at this time. Finally, all series
recognize the reduction in risk in the period around the middle of the century. This
explains the strong secular bull market of the post-WWII period.

In unreported results, the values of $P_t(1)$ as estimated for P95–100 and the
aggregate consumption series are shown to be highly similar to the smoothed regime
probabilities estimated from the MSVAR approach. The values of $P_t(1)$ used in the
asset pricing model of the next section are from the P95–100 series for the years
1917–1986. This series is preferred because, particularly in the first two-thirds of the
last century, those with highest income held the greatest proportion of stocks (see, in
particular, Blume, Crockett, and Friend 1974; Poterba and Samwick 1995). Also, by
inspection of Figure I, it appears to be neither an early nor a late indicator of the decline
in risk in the post-WWII period. For the period before 1917 and after 1986,
probabilities from aggregate consumption are taken.²

III. The Asset Pricing Model

For the calibrations, three data series are used for annual stock returns and dividends over
the period 1872–1999. The main calibrations are based on data from Wilson and Jones
(2002), henceforth W&J, and from Robert Shiller's website. For further robustness, data
are also taken from Global Financial Data (GFD). For W&J and GFD, end of year data

² By construction, the data splice perfectly in 1917. In 1986, there is only a small difference between the
values of $P_t(1)$ for P95-100 and aggregate consumption and no attempt is made to adjust for this.
are taken. Shiller's annual data, though, are for January. I therefore consider the lagged series “Shiller+1”, when stock price data for January 1950, for example, is used for the year 1949, but the dividend for 1950 is kept in year 1950. For the W&J and GFD series, the price and dividend series are deflated using the Consumer Price Index (CPI) deflator from GFD. For Shiller, real prices and dividends are taken directly from his spreadsheet. There are very high correlations between the real logarithmic total returns calculated for the three series over the period 1897–1999 (99.5% between W&J and GFD, 96.5% between W&J and Shiller+1). The correlation between the real logarithmic dividend growth series over the sample is also high (89.9% between W&J and GFD, 86.5% between W&J and Shiller+1).

A further adjustment is made to the dividend series. Because of the rapid rise in share repurchases towards the end of the century, for every year after 1972 (the first year for which aggregate repurchase data are available) the dividend series is multiplied by a repurchase adjuster. This is calculated from Grullon and Michaely (2002).³

The asset pricing model used in this article is based upon Ang and Liu (2004). Full details are given in the Appendix, where it is also shown that similar results, both in terms of price levels and returns, are obtained when using the log-linear model of Campbell and Shiller (1988) in place of the one described here.

From Ang and Liu (2004), the net present value, $p_t$, of an asset with dividend

³ For each year $t$ after 1971, the adjuster is defined by $(\text{Stock Repurchases} + \text{Dividends}) / \text{Dividends}_t$. The dividend series is then multiplied by $\text{Adjuster}_t / \text{Adjuster}_{1972}$. It should be noted that there is some disagreement about the absolute level of aggregate stock repurchases. See, for example, Weston & Siu (2003).
stream $d_t$ is given by

$$
p_t = \sum_{j=1}^{\infty} E_t \left( d_{t+j} \exp \left( -\sum_{k=1}^{j} m_{t+k} \right) \right)
$$

$$
\exp (m_{t+k}) = E_{t+k-1} \left( \frac{p_{t+k} + d_{t+k}}{p_{t+k-1}} \right)
$$

The expected simple single period return is comprised of a risk-free component $r_{ft+k}$ and a risk premium, $\rho_{t+k}$, so that $m_{t+k} = \ln(1 + r_{ft+k} + \rho_{t+k})$. Taking a first order Taylor's series approximation of this around $1 + \rho_{t+k}$ gives $m_{t+k}$

$$
\approx \ln(1 + \rho_{t+k}) + r_{ft+k}/(1 + \rho_{t+k}) = \ln(1 + \rho_{t+k}) + r_{ft+k}
$$

In addition, $d_{t+j} = d_t \exp(\sum_{k=1}^{j} g_{t+k})$

where $g_{t+k} = \ln(d_{t+k}/d_{t+k-1})$, so that:

$$
p_t \approx d_t \sum_{j=1}^{\infty} E_t \left( \exp \left( \sum_{k=1}^{j} g_{t+k} - \ln(1 + \rho_{t+k}) - r_{ft+k} \right) \right)
$$

This is a standard discounted dividend model with a time-varying cost of capital.

I assume that the risk premium is independent of $g_{t+k} - r_{ft+k}$, including at all leads and lags.

Table 3 presents summary statistics for $r_{ft}$, $g_t$, and $g_t - r_{ft}$. Data for both the nominal risk-free rate (the Treasury bills return index) and CPI are taken from Global Financial Data to determine the real risk-free rate, $r_{ft}$. This variable could, in principle, be treated endogenously within the type of theoretical framework that forms the basis for this article. Here, though, the observed real risk-free rate is used. In addition, dividend growth and the risk-free rate could be modeled separately. However, because real dividend growth is considerably more volatile than the real risk-free rate, $g_{t+k} - r_{ft+k}$ has
a correlation of 93% with $g_{t+k}$. Therefore, for simplicity, $g_{t+k} - r_{f+k}$ is modeled as a single series, and unreported tests suggest that the results I present have low sensitivity to this choice. Although the chi-squared statistics in Table 3 indicate that this is not the case, it is also assumed for tractability that this series is normally distributed. Under these assumptions:

$$p_t \approx d, \sum_{j=1}^{\infty} \exp \left(1_j m_{tj} + \frac{1}{2} 1_j S_{tj} 1_j \right) \rho_{jt}$$

$$\rho_{jt} = E_t \prod_{k=1}^{j} (1 + \rho_{t+k})^{-1}$$

(8)

$1_j$ is a $j$-vector of 1s. $m_{tj}$ is a $j$-vector with elements $E_t(g_{t+k} - r_{f+k})$ and $S_{tj}$ is a $j \times j$ variance-covariance matrix with elements $\text{Cov}_t(g_{t+k} - r_{f+k}, g_{t+k} - r_{f+k'})$ for $k, k' \in [1, j]$.

[Insert Table 3 around here]

An important proposed resolution for the excess volatility paradox is that dividend growth is highly persistent; see, for example, Barsky and De Long (1993). Donaldson and Kamstra (1996) also use a model with persistent dividend growth rates to get a very close fit for the observed stock market behavior of the 1920s and 1930s. However, the integrated dividend growth process of Barsky and De Long (1993) has come under criticism both for the lack of empirical support (Campbell, Lo and MacKinlay 1997) and theoretical inconsistency (Pagès 1999). This article follows Bansal and Lundblad (2002) by allowing for dividend growth predictability using an ARMA process.

Augmented Dickey-Fuller tests, reported in Table 3, reject the null hypothesis of
there being a unit root in the series $g_{t+k} - r_{t+k}$ with levels of significance well below 1%.

For all three data sets, the dividend growth minus risk-free rate series can be well modeled by an MA(1) process:

$$g_t - r_t = \alpha_0 + \beta \eta_{t-1} + \eta_t$$ \hspace{1cm} (9)

Estimates of $\hat{\beta}$ provide statistically significant evidence of autocorrelation in the dividend growth process for the W&J and GFD data but not for Shiller's data. Even in the case of W&J and GFD, the autocorrelation is not particularly strong from an economic perspective with $\hat{\beta} < 0.25$. This conclusion is also confirmed by the first-order autocorrelation statistics reported in Table 3. The finding that dividend growth rates have only limited predictability over one year and almost no predictability more than one year ahead is also supported by the multivariate regression results reported in Fama and French (2002). Under this MA(1) process, it is straightforward to prove that, if $\Gamma_j$ is defined by:

$$\Gamma_j = \exp\left(j\left(\alpha_0 + 0.5\sigma_{\eta}^2\right) + (j-1)\sigma_{\eta}^2(0.5\beta^2 + \beta_j)\right)$$ \hspace{1cm} (10)

then equation (8) becomes:

$$p_t \approx d_t \exp(\beta \eta_i) \sum_{j=1}^{\infty} \Gamma_j \rho_{jt}$$ \hspace{1cm} (11)

In the case where there is both constant dividend growth $\text{Var}(\eta_i) = \sigma_{\eta}^2 = 0$, so $\Gamma_j = \exp(j\alpha_0)$, and a constant equity premium, $\rho$, so $\rho_{jt} = (1+\rho)^{-j}$, then $p_t = d_t \exp(\alpha_0)/(1+\rho - \exp(\alpha_0))$. This, as expected, is the standard static Gordon Growth model. In the case when there is a constant equity premium, $\rho_{jt} = (1+\rho)^{-j}$, but there is non-constant dividend growth, $\text{Var}(\eta_i) = \sigma_{\eta}^2 \neq 0$, then it is straightforward to show
that:

\[ p_t = d_t \frac{\exp\left(\beta_t \eta_t + \alpha_t + 0.5\sigma_t^2\right)}{1 + \rho - \exp\left(\alpha_0 + \sigma_t^2(0.5(\beta_1^2 + 1) + \beta_1)\right)} \]  \hspace{1cm} (12)

The $\beta_t \eta_t$ term incorporates into the price the one-period ahead forecastability of dividend growth. The $\sigma_t^2$ terms capture Itô effects from $\frac{1}{2} \mathbf{1}_j S_q \mathbf{1}_j$ in equation (8). This version of the model will be referred to as the “partial model”.

When there is both time-varying dividend growth and equity premium, referred to as the “full model”, no simple closed form solution for $p_t$ is available. The Appendix describes in detail how $\sum_{j=1}^{\infty} \Gamma_j \rho_{\beta}$ can be empirically estimated in this case. In particular, it is proved that, to high accuracy, this variable can be modeled as linear in $P_t(1)$. As a consequence:

\[ p_t \approx d_t \exp\left(\beta_t \eta_t \left(x_2 + P_t(1)(x_1 - x_2)\right)\right) \]  \hspace{1cm} (13)

where $x_1$ and $x_2$ are constants that can be broadly interpreted as average price / dividend ratios in the two states. The values of these variables depend on the MA(1) parameters for dividend growth, the transition probability $\zeta$, and particularly $\rho(s)$, which are the values that the single-period equity premium would take if the true state next period were revealed immediately to investors. For calibration purposes, the values of $x_1$ and $x_2$ are estimated by Monte Carlo simulation and, from there, determining rational asset prices follows immediately from equation (13).

To complete the calibration of $x_1$ and $x_2$ it is necessary to estimate values of the single period equity premium in the two states, $\rho(s)$. For the base calibrations, I use
\( \rho(1) = 8\% \), which is broadly consistent with, for example, Cornell (1999) who suggests that a risk premium of 7\% over bonds and 9\% over bills is a good rough approximation of a forward looking equity premium. I set \( \rho(2) = 3.5\% \), which is low when compared with current survey estimates of the ex ante equity premium of around 5.5\% reported in Welch (2008), but consistent with studies such as Claus and Thomas (2001). These two values of the equity premium can also be compared with the realized returns over the periods 1897–1951 and 1959–1987. These dates are chosen as the estimates of \( P_t(1) \) are relatively steady during these periods. 1952–1958 represents the strong shift in risk in the post-WWII period. 1988–1999 captures the third, even lower, risk state described in section V of this article. These periods provide estimates of \( \rho(1) = 7.3\% \) and \( \rho(2) = 4.2\% \). Again, this is clear evidence of a reduction in the equity premium, albeit less than the 8.0\% to 3.5\% used in the baseline calibration of the article. Sensitivity analysis is presented for \( \rho(1) \) and \( \rho(2) \) below as there is considerable disagreement in the literature concerning estimates of the pre-break and post-break equity premia. In particular, this will focus on \( \rho(1) = 7.0\% \) and \( \rho(2) = 4.5\% \), which are close to the ex post estimates.

The baseline calibration values of \( \rho(1) = 8\% \) and \( \rho(2) = 3.5\% \) can also be compared with the equity premium that would be justified by the CCAPM as given in equation (1). Based on estimates of \( \rho(R_{it}, r_{ct}) \), \( \sigma(R_{it}) \), and \( \sigma(r_{ct}) \) for the periods 1897–1951 and 1959–1987, \( \rho(1) \) and \( \rho(2) \) imply that the coefficient of relative risk aversion, \( \gamma \), is equal to 17 for state 1 and 38 for state 2. These values of \( \gamma \) would be viewed by most economists as being above realistic estimates of the coefficient of
relative risk aversion for the majority of investors. This is a re-statement of the equity premium puzzle of Mehra and Prescott (1985). If, though, an economist did view \( \gamma = 17 \) as a realistic estimate, then this would theoretically justify setting \( \rho(2) = 1.61\% \). This gives some credence to the view that the calibrated change in the equity premium between the two states is being underestimated in the baseline calibration. The focus of this article remains on explaining observed low frequency stock market movements based on standard empirical, as opposed to theoretically justified, estimates of the equity premium.

IV. Results

The S&P 500 in Levels

Figure II presents the observed real level of the stock market from 1889–2000 in logs and the retrodicted values from five rational asset pricing models. This replicates Shiller (2003, Figure 1) with hindcast values from both the full and partial models superimposed. For the partial model, \( \rho = \rho(1) = 8\% \). Shiller (2003) presents three series of rational prices. The first, “PDV, Constant Discount Rate”, is derived from his original 1981 model. The second, “PDV, Interest Rate”, varies the discount rate in line with changes in the one-year real interest rate. The final series is derived from the model of Grossman and Shiller (1981), where the discount rate varies with the marginal rate of substitution of aggregate consumption, which is assumed to be perfectly forecastable.

[Insert Figure II around here]
The high explanatory power of the full model, particularly in the period before 1973, is clearly revealed. Although the Grossman and Shiller (1981) model captures the Wall Street Crash with some accuracy, not one of Shiller's series comes close to reflecting the observed volatility of the market over the total sample; this is the puzzle he sets. In contrast, the full model broadly captures all the major movements, both in terms of magnitude and timing, in the period before the 1973–1974 stock market crash.

Figure III compares the price / dividend (p/d) ratio of the full and partial models against the observed values, based on Shiller's data.

[Insert Figure III around here]

For the partial model, the retrodicted p/d ratio is almost flat. This is because there is very little dividend forecastability based on annual data and aggregate dividends are smooth. As a consequences, \( \exp(\beta_i \eta_i) \approx 1 \) for all \( t \) and the p/d ratio, from equation (12), is approximately constant. The correlation between the observed and retrodicted p/d ratio from the partial model is only 2% over the total sample period 1897–1999. For the full model, the p/d ratio is approximately equal to \( x_2 + P_t(1)(x_1 - x_2) \). Again, the ability of the full model to explain low-frequency returns, particularly in the 1920s, 1930s, 1950s and 1960s, is clearly revealed, as is its inability to explain the period following the 1973 stock market crash. The p/d ratio of the full model has a correlation of 61% with the observed p/d ratio over the period 1897–1999.

Figure III also includes a “perfect knowledge” retrodicted p/d series that shows the importance to this article of Veronesi's (1999) assumption that investors cannot
perfectly identify the current state. In contrast to the full model, the “perfect knowledge” case assumes that investors could instantaneously identify the structural break in 1949 reported in Panel C of Table 2 for the P95–100 series without the need to draw inferences using Bayes rule; \( \text{Prob} \left( \delta_t = 1 \mid \Omega_t \right) = 1 \) before 1950 and \( \text{Prob} \left( \delta_t = 1 \mid \Omega_t \right) = 0 \) thereafter. As can be seen from Figure III, this retrodicted p/d series has a sharp, instantaneous rise, from 18.9 in 1949 to 32.7 in 1950. This contrasts with the full model where investors take several years to fully learn about the change in states and therefore the p/d ratio does not rise above 30 until 1959 which better matches the observed data. The full model also performs noticeably better than the perfect knowledge model early in the century as the correlation between the observed p/d ratio and the retrodicted p/d ratio for the perfect knowledge series is 52% for the total sample period 1897–1999. Lettau, Ludvigson, and Wachter (2008) also exploit this advantage of gradual discovery in a highly related context. The use of Bayesian learning in this article should therefore be contrasted with its more common application in financial econometrics, where it is generally assumed that investors have full contemporaneous knowledge about the underlying state but that econometricians have access to a much more limited set of information. Bayes rule is then used as the foundation for a number of empirical techniques that generate ex post estimates of the regimes.

To determine whether the information contained in the conditional volatility of consumption and income growth is statistically significant in helping to explain the level of the S&P 500 index over the twentieth century, a Monte Carlo approach is taken over 1000 iterations.

For each year 1889–2000 within each iteration, a value is taken at random, with
replication, from the observed time series of consumption growth. Income data are not used in the Monte Carlo process as they are only available for part of the sample. When randomizing the data, splicing income and consumption estimates is problematic. For each randomly generated consumption series, the Inclán and Tiao and Chen and Gupta tests are used to test for a single structural break in volatility. The identified date of the break for each iteration is the average of the two dates identified by the two tests irrespective of the level of statistical significance. If the identified break is in either the first or last decade, the iteration is discarded and replaced. Conditional on the break date, values of $\mu_s$, $\sigma_s$ and $P_t(s)$ are calculated as before. The single period equity premium in each state is given by $\rho(s) = 0.007408038 + 1.80182851 \sigma_s$.\footnote{This formula ensures that when $\sigma_1 = 4.03\%$ and $\sigma_2 = 1.53\%$, which represent the standard deviation of consumption growth in the two states as given in Table 2, then $\rho(1) = 8\%$ and $\rho(2) = 3.5\%$. It also means that the greater the change in consumption growth volatility between the two states, the greater the difference in the equity premia.} Equation (13) can then be used to generate 1000 Monte Carlo price series based on the randomized consumption growth data. To establish the accuracy of each price series, the error $\varphi_t = \ln(p_t^{\text{estimated}}) - \ln(p_t^{\text{observed}})$ is calculated for each year. A mean squared error, $\sum_{t=1889}^{2000} \varphi_t^2 / 112$, can then be determined for each price series and a $p-$value for the full model established.

An advantage of this method of estimating $p-$values is that it helps overcome a potential look-ahead bias. In the calibration process, it has been assumed that the main parameters ($\mu_z$, $\sigma_z^2$, $\zeta$, $\rho(s)$, $\alpha_0$, $\beta_1$, and $\sigma_\eta^2$), while estimated ex post, have...
always been in the investors' information set. As the same method is used to generate each of the Monte Carlo price series as the full model, this assumption applies in all cases. A statistically significant $p$-value for the full model cannot, then, be explained away by the look-ahead bias.

The full model in Figure II has a lower mean squared error than 99.9% of the 1000 Monte Carlo price series.\(^5\) This shows to a very high degree of statistical significance that the conditional volatility of income and consumption growth is informative for understanding low frequency movements in the S&P 500.

**Bull and Bear Markets**

If changes in the ex ante equity premium are truly driving low frequency movements in the S&P 500, it follows that large increases and decreases in $P(1)$ should have driven the major stock market busts and booms of the twentieth century. As a consequence, the full model should have excellent ability to explain average returns over individual historic secular bull and bear markets.

To test this hypothesis, it is necessary to identify the dates when each market cycle began and ended. There are a number of previous studies that examine the timing

\(^5\) Two different sets of tests of statistical significance were also run. In the first, the value of $\rho(s)$ is always set equal to 8% in the high risk state and 3.5% in the low risk state. In the second, the mean squared error is calculated by $\sum_{t=1889}^{2000} \left( \varphi_t - \bar{\varphi} \right)^2 / 112$ where $\bar{\varphi} = \sum_{t=1889}^{2000} \varphi_t / 112$. This adjusts for systematic over- or under-valuation by the model. In all cases, the full model outperforms at least 99.6% of the random series. When the data is truncated at 1973, the full model always outperforms all the random series.
of bull and bear markets, including Pagan and Sossounov (2003) and Gonzalez et al. (2005). The tendency, though, is to use monthly data in the analysis. As a consequence, the average durations of bull and bear markets are too short to be relevant for a study based on annual data.

To decide on the timings of secular bull and bear markets using annual data, I use the counts cusum technique described in Gonzalez et al. (2006). There is considerable agreement between the W&J, GFD and Shiller+1 series over when the bull and bear markets occur and these timings agree with other sources, such as the dates given by Michael Alexander (http://www.safehaven.com/article-68.htm). Bull markets for the twentieth century are taken to run from 1897–1905, 1922–1928, 1949–1965 and 1982–1999, with bear markets making up the intervening periods.

Observed and retrodicted average returns over each major market cycle are presented in Table 4. Panel A presents the base parameterization based on Wilson and Jones (2002) data. There are seven columns. The first gives the observed average real annual logarithmic capital gain to the stock market during each identified secular bull and bear market. Dividends are excluded as they are an input into the model. The next six columns present the retrodicted average real annual logarithmic capital growth based on three variants of the model together with associated \( p \)-values. The partial model is based on equation (12), when the equity premium is fixed at 8%. The full model has a time-varying equity premium and is based on equation (13). The “Consump.” model is also based on this equation, but values of \( P_t(1) \) are estimated from consumption data alone as shown in Figure I. \( p \)-values are calculated from the Monte Carlo process described in the previous subsection where the consumption growth data have been
randomized. The values represent the percentage of the 1000 Monte Carlo price series that have a lower absolute error than the model in explaining observed average returns for each market cycle.

[Insert Table 4 around here]

Unsurprisingly, the partial model does not perform particularly well, with \( p \) values being close to 50\% in all cases. By contrast, the full model shows strong explanatory power. For all seven booms and busts, it explains average returns better than the partial model. The power of the full model is particularly revealed in the period before 1966. For example, for the 1922–1928 bull market, when observed real growth averaged 16.2\% per year, the full model hindcasts 12.0\% while the partial model only retrodicts 7.8\%. Similarly, for the 1949–1965 bull market, the retrodicted average annual real capital gains from the full and partial models are 6.9\% and 3.9\% against an observed figure of 8.6\%. That the model has explanatory ability for average returns is established at 10\% significance for the 1906–1921 bear market, 2\% significance for the 1897–1905 bull market and 1\% significance for the bear market of 1929–1948 and bull markets of 1922–1928 and 1949–1965.\(^6\)

The second part of the table also shows that the Wall Street Crash of 1929–1932

\[ p(s) = 0.007408038 + 1.80182851 \sigma_s \]

\( p \) values were also calculated in the case when the single period equity premium always equals 8\% in the high risk state and 3.5\% in the low risk state. In this case, the full model is significant at the 10\% level for 1897–1905, the 5\% level for 1929–1948 and the 1\% level for 1922–1928 and 1949–1965.
is very well explained by the full model, when the observed, full model and partial model average annual real capital gains are -26.5%, -19.0% and -10.8% respectively. The fall in dividends and expected future dividend growth is enhanced by a sharp rise in the equity premium caused by the associated macroeconomic risk of the Great Depression.

With the exception of the 1929–1948 bear market, similar results are achieved whether we base estimates of $P_t(1)$ on income or consumption data. However, the consumption-based model performs poorly for this bear market, with a $p$–value of over 80%. The second part of the table shows that the consumption data can explain the Wall Street Crash, but not the falling market of 1937–1948. This is because, under the consumption parameters, macroeconomic risk was perceived to be falling at this time; see Figure I.

Panels B and C present similar results for the full model based on different data sets. Some even more supportive findings for the time-varying equity premium model appear here. For example, based on the GFD data, the observed and retrodicted values for the 1922–1928 bull market are 17.3% and 15.4% respectively.

The final part of the table examines the volatility of the market over the whole sample and two fairly even sub-samples divided by the end of the 1929–1948 secular bear market. Over the total sample, the observed standard deviation is 19.5%, while the full model hindcasts 17.4%. An F-test fails to reject the null hypothesis that the variance of the full model's capital returns is equal to the variance of observed capital returns at the one-sided 10% level. The same null hypothesis is rejected at the one-sided 2% level for the partial model, where the standard deviation is 15.9% over the total sample. There is therefore only limited evidence of excess market volatility when
persistent changes in the equity premium are recognized.

**Sensitivity Analysis**

I now demonstrate that the economic conclusions of this article have low sensitivity to large changes in input values.

Two of the driving variables for this model are the assumed levels of the equity premium in the two states, \( \rho(s) \), and the persistency of those states, captured by \( \zeta \). Panel A of Table 5 presents results for \( \zeta = 0.9, 0.95 \) and 0.99. These correspond to an expected time to a first change in state of 10, 20 and 100 years respectively. All other values are as for the base case parameterization. The greater the value of \( \zeta \), the better the model is at explaining the magnitude of the secular bull and bear markets, particularly in the 1922–1965 period. This finding is consistent with previous literature that shows that persistence, be it either in dividend growth rates or changes in the cost of capital, is crucial for generating the extreme movements observed in stock prices. For all values of \( \zeta \) considered, the full model has a lower absolute forecasting error for all seven of the secular bull and bear markets of the twentieth century than the partial model.

*Panel B of Table 5 presents results for \( \rho(1) \in \{7\%, 9\%\} \) and \( \rho(2) \in \{2.5\%, 4.5\%\} \) with \( \zeta = 0.98 \). The model performs best in this case with the greatest difference in equity premium between the two states; \( \rho(1) = 9\% \) and \( \rho(2) = 2.5\% \). The average level of equity premium has an important impact on the*
estimated dividend yields of the market in the two states, \( x_1 \) and \( x_2 \), but a much smaller effect on hindcast returns. Again, for all values of \( \rho(1), \rho(2) \) considered, the full model has a lower absolute forecasting error for all seven secular bull and bear markets than the partial model.

The case \( \rho(1) = 7\% \) and \( \rho(2) = 4.5\% \) is of particular interest as these values are close to the ex post estimates of the equity premium over the intervals 1897–1951 and 1959–1987 respectively. The Monte Carlo simulations were rerun in this case to see if the full model continues to have statistically significant explanatory power.

When retrodicting the S&P 500 in levels, the model has a lower mean squared error under this calibration than under the baseline calibration. This is because, with the equity premium in state 2 now raised, the model overpredicts market values by less in the period after 1973. As a consequence the full model performs better than all 1,000 of the Monte Carlo simulations when the index is examined in levels.

When explaining low frequency returns, the full model performs more poorly under this calibration than under the baseline calibration, but still clearly has explanatory power that is significant in both a statistical and economic sense. The \( p \)-values are below 1% for the bull markets of 1922–1928 and 1949–1965 and the bear market of 1929–1948, below 5% for the bull market of 1897–1905 and below 10% for the bear market of 1906–1921.\(^7\)

\(^7\) In the Monte Carlo iterations, \( \rho(s) = 0.02967113 + 1.001016\sigma_s \), giving \( \rho(1) = 7\% \) and \( \rho(2) = 4.5\% \) when substituting in the observed level of consumption growth volatility in the two states. \( p \)-values were also calculated in the case when the single period equity premium always equals 7% in the high risk state and 4.5% in the low risk state. In this case, the \( p \)-value is significant at the 5% for
Next, I consider how the estimation process for $P_t(1)$ might have influenced the results. This parameter is a function of $\mu_s$ and $\sigma_s$. These are empirically estimated by the pre-break and post-break mean and standard deviation of the ARMA residuals of the real log growth rate in the consumption or income series. Therefore, $P_t(1)$ depends on the assumed date of the structural break. A somewhat ad-hoc process was used to determine this in Panel C of Table 2, where the average of the dates identified by the Inclán and Tiao (1994) and Chen and Gupta (1997) break tests was chosen. To test for robustness, values for $P_t(1)$ were re-estimated for series P95–100 based on the structural break being in 1939 and 1959. These dates are chosen to be one decade earlier and later respectively than the date used in the main calibration. The estimated values of $P_t(1)$ are found in unreported results to be highly insensitive to the assumed date of the break. Instead, as shown in Figure I, these values are much more sensitive to the data series used in the analysis. The asset pricing model was also re-calibrated using the values of $P_t(1)$ from two different income series; P0–90 and P90–95. In general, the full model is slightly less accurate for both of these series than when $P_t(1)$ is estimated using P95–100. Despite this, both versions of the model have a lower absolute forecasting error than the partial model for six of the seven secular bull and bear markets. The exception is 1966–1981, but this is a period that even the main calibration fails to accurately capture.

Finally, I fit the series $g_{t+1} - r_{f_{t+1}}$ using AR(1), MA(2) and ARMA(1,1) processes as opposed to the MA(1) process in the main calibration. The results are shown, in

unreported results, to be highly insensitive to the way in which the dividend growth series is ARMA modeled.

V. 1966–1999

This study has, so far, focused on movements in the cost of equity capital in the early- to mid-twentieth century. Lettau, Ludvigson, and Wachter (2008), amongst others, have argued that the equity premium declined much more recently. I now extend the full model to allow for the possibility of a third, even lower risk, state between 1966 and 1999. The enhanced model can be calibrated using higher frequency data as measurement error in consumption data is of less concern in this period.

For this post-war period, seasonally adjusted real monthly data for the consumption of nondurables and services are taken from the Federal Reserve Bank of St. Louis from January 1959 to December 1999. Unreported EWMA models of the consumption growth series reveal fairly steady conditional volatility from 1959–1985 and then a sharp decline from 1985–1999.

To test for a structural break in the volatility of real consumption growth, the series is ARMA(3,1) modeled and the residuals tested using the methods of Inclán and Tiao (1994) and Chen and Gupta (1997). Both tests report very clear evidence (p-values <1%) of a structural break in volatility. The Chen & Gupta test identifies the structural break in January 1992, while Inclán and Tiao points to March 1985. Again, the average of these two dates is taken as the basis for estimating $\mu_1, \sigma_1, \mu_2, \sigma_2$; the pre- and post-break parameters of the normal distributions that describe the ARMA residuals of real aggregate consumption growth. The estimated volatility parameters are
\[ \sigma_1 = 0.381\% \text{ and } \sigma_2 = 0.263\%. \] These two states are referred to as “Medium” and “Low” macroeconomic risk respectively. A third state, \( s = 3 \), that represents “High” risk, is also introduced, to reflect the possibility that the economy might return to the conditions experienced during the inter-war years. Because of the measurement error issues discussed above, it is difficult to accurately calibrate \( \mu_3, \sigma_3 \), but the results have low sensitivity to these values. Baseline calibrations are based upon \( \mu_3 = \mu_1 \) and \( \sigma_3 = 2\sigma_1 \).

The transition matrix is:

\[
\begin{bmatrix}
2\zeta - 1 & 1 - \zeta & 1 - \zeta \\
1 - \zeta & \zeta & 0 \\
1 - \zeta & 0 & \zeta
\end{bmatrix}
\]

The average time spent in the Low and High risk states on each visit, at \( 1/(1 - \zeta) \), is twice as long as the average time spent in the Medium risk state on each visit. I set \( 1/(1 - \zeta) = 360 \) months. As visits to the middle state are twice as frequent as to the extreme states, the economy spends an equal amount of time in each state asymptotically. As before, initial probabilities are set equal so as not to impose strong prior beliefs on the data; \( P_{19592}(1) = P_{19592}(2) = P_{19592}(3) = 1/3 \). Figure IV presents the estimated probabilities of being in each of the three states. Again, the reduction in macroeconomic risk over the period 1985–1995 is clearly revealed. Although the precise method used is different, this figure can be compared directly with Figure 4 in Lettau, Ludvigson, and Wachter (2008).

[Insert Figure IV around here]
Monthly share price and dividend series are taken from Hallerbach (2005) which, in turn, are based on the Wilson and Jones (2002) data set. For robustness, in unreported results, I also estimated monthly dividends from the capital-growth only and total returns indices of the NYSE, AMEX and NASDAQ taken from CRSP. Results are largely insensitive to which series is used. The risk-free rate data come from CRSP. These series are deflated using the CPI index from Robert Shiller's website. The share repurchase adjustor to the dividend series from Grullon and Michaely (2002) is converted to monthly values using a cubic spline method. The data are finally converted to quarterly to overcome seasonality issues in dividend payments and then ARMA(5,2) modeled.

The asset pricing model was calibrated in this case using a Monte Carlo method based on quarterly data from 1959:1 to 1999:4:

\[
p_{jt} = d_t \sum_{j=1}^{\infty} E_t \left( \exp \left( \sum_{k=1}^{j} g_{t+k} - r_{jt+k} \right) \right) \rho_{jt}
\]

\( \rho_{jt} \) is the equity premium discount factor applied at time \( t \) for a cash flow with horizon \( j \). A formal definition of this variables for the two-state model is given in equation (A5) in the Appendix; an analogous version is used in this three state version. For each quarter in the sample, the variable is numerically estimated using 5000 Monte Carlo iterations for \( j < 2000 \). Then, for each quarter and each \( j \in [1,2000] \), 5000 iterations are run to calculate the value of \( E_t \left( \exp \left( \sum_{k=1}^{j} g_{t+k} - r_{jt+k} \right) \right) \) with \( g_{t+k} - r_{jt+k} \) simulated using the ARMA(5,2) process with parameters and residuals taken from the observed data. The ex ante equity premium levels are taken to be \( \rho(1) = 4.5\%/4 \),
\[ \rho(2) = 2.5\%/4 \quad \text{and} \quad \rho(3) = 8\%/4. \]

Figure V presents the observed and estimated log price series both when the equity premium is time-varying (the full model) and when it is fixed at \( 4.5\%/4 \) for all states (the partial model). For the combined bear and bull period from 1965:Q3–1999:Q4, the time-varying equity premium model is highly accurate in explaining the average real capital gain to the market. The estimated value is 0.84\% per quarter, which is closer to the observed value of 0.86\% than the partial model estimate of 0.59\%. This results in the retrodicted level of the market being 39\% higher in 1999 when using the full model than the partial model. The improvement comes mainly from the final part of the sample. Allowing for the third equity premium state drives up the retrodicted average annual real capital growth to the market from 4.9\% to 6.9\% for the 1982–1999 secular bull market. The full model, though, does not help improve our understanding of the secular bear market of 1966–1981.

[Insert Figure V around here]

I conjecture that the reason why this bear market occurred is because of the type of low probability disaster event described in an equity premium context by Reitz (1988), Barro (2006, 2009) and Martin (2008). If investors allocate a small probability to a very severe negative event, then this can have a significant impact on the ex ante equity premium and hence the value of the market. The event about which investors worried need not then be observed in the ex post data. The secular bear market of 1966–1981 coincided with much of the Cold War, nuclear weapon proliferation and the oil crisis.
That the perceived level of macroeconomic risk increased during this time seems rational, even if this is not reflected in the subsequent consumption data. When observed consumption risk fell further in the late 1980s, then this may also have been associated with a decline in the perceived risk of future economic disaster. The bull market of 1982–1999 is then driven by three factors; a “catch-up” from the 1970s as investors worried less about possible future economic catastrophe, strong dividend plus stock repurchase performance and a decline in measured macroeconomic volatility.

VI. Conclusion

In this article, I present an asset pricing model that performs well at explaining low frequency movements in the S&P 500 index during the twentieth century. The model is founded on the observation that the level of macroeconomic risk faced by stockholders has changed significantly over time. This, in turn, would have had a major impact on the ex ante equity premium. In particular, the bull markets of 1922–1928 and 1949–1965 occurred during periods of falling risk, while macroeconomic uncertainty rose greatly during the Wall Street Crash. This article also shows that, at least based on annual data, there is little evidence of significant excess volatility in stock returns.

The full model can also help explain other stylized facts about low frequency stock returns. Mehra and Prescott (2003) note that there was a rapid change in the equity premium around 1933, which they attribute to the United States coming off the Gold Standard. An alternative explanation arises from the model in this article. From unreported EWMA estimates of consumption risk, 1932 represents the year of highest perceived macroeconomic risk in the whole of the twentieth century; a feature that is not
surprising given the Great Depression. In contrast, EWMA estimates based on monthly
data from 1959 show that November 1999 had the lowest perceived macroeconomic risk.
In addition, in the three years preceding 1934, real log aggregate dividend growth had
averaged -23% per year. By contrast, in the three years preceding 2000, real aggregate
dividends plus stock repurchases had grown on average by 9.9% a year. Under a
Gordon growth model with time-varying parameters, the discount rate would have been
considerably higher in 1933 than 2000 and the expected future dividend growth rate
much lower. As a consequence, the full model retrodicts average annual real capital
growth of −0.9% for 1890–1932 while the observed value was −1.1%. For the
1933–1999 period, the figures are 3.4% (4.2% when allowing for the three-state
model of Section V) and 4.2% respectively.

If consumption risk has really declined in the second half of the century, then this
should lead to other observable effects in asset pricing. In particular, we might look for
an increase in the real risk-free rate associated with a decline in the precautionary savings
demand. CASUAL inspection reveals this phenomenon. From 1897–1948 the real
risk-free rate averaged 0.8% per annum, while from 1949–1999 it averaged 1.3%. There are, though, many other potential explanations for this effect and this issue
deserves a more detailed study in its own right.

---

8 Indeed, it is at least theoretically possible that, even though the equity premium declined in the second
half of the century, the overall cost of equity could have increased due to an offsetting diminishing
precautionary savings effect of greater magnitude. The small rise in the observed real risk-free rate does
not, though, seem to be sufficiently large to dominate the significant fall in the equity premium.
Appendix - The Asset Pricing Model

This Appendix provides details of the asset pricing model used in this article. The first subsection explains how equation (13) follows from equation (11). It also shows how the variables $x_1$, $x_2$ can be estimated using Monte Carlo methods. The second subsection shows how highly similar results can be derived using the log-linear model of Campbell and Shiller (1988).

The Term Structure of Equity Discount Premia

The baseline calibrations in this article calculate state probabilities, $P_t(s)$, based on income growth volatility, while the equity premium depends on the conditional volatility of consumption, $\sigma(r_{t+1})$. The link between aggregate consumption and income volatility is explored in the literature on consumption smoothness (see, for example, Campbell and Deaton 1989). Realized change in consumption over the interval $[t, t+1]$ will depend on the change in permanent income over the same period and so the conditional volatility of income will directly influence the conditional volatility of consumption. This relation can be clearly seen in the values of $P_t(1)$ presented in Figure I for the period when both income and consumption data are available. Although there is some disagreement over the exact timing of regime changes, particularly following WWII, it is clear that both series give highly similar patterns. As a consequence, the main results of this article, presented in Table 4, are largely insensitive to whether $P_t(1)$ is estimated from income or consumption data apart from in the 1937–1948 bear market subperiod. Because of the higher quality of income data in the first half of the century, this article bases its main calibrations on income based estimates.
of \( P_t(1) \).

To formally establish the relation between \( P_t(1) \) as estimated from income data and \( \sigma_t(r_{t+1}) \), let \( y_t \) denote income of the representative rich household and \( G_{t+\tau} = \ln(y_{t+\tau+1}/y_{t+\tau}) \) log income growth. To capture autocorrelation in income growth in a simple manner, assume that it follows an MA(1) process, \( G_{t+\tau} = a_0 + b_1 e_{t+\tau-1} + e_{t+\tau} \), where \( e_{t+\tau} \) is taken to be independently and normally distributed. From Table 1, this MA(1) process fits a number of the fractile groups' average income growth well, including P99–99.5, which has similar first order autocorrelation and standard deviation to series P95–100. Consistent with the main model in the article, it is assumed that \( e_t \) is heteroskedastic and follows a two-state Markov switching process. For consistency with the P99–99.5 series, \( b_1 = 0.2651, e_{t+\tau} \sim N(0,0.819\%^2) \) for state 1 and \( e_{t+\tau} \sim N(0,3.06\%^2) \) for state 2. However, I set \( \alpha_0 = -0.0035 \), which is well below the empirical estimate \( (\alpha_0 = 0.0115) \).

The reason for this choice is that, with rapidly growing income, the permanent income hypothesis predicts high initial consumption, leading to unrealistic implied borrowing levels. The value of \( \alpha_0 \) is chosen so that initial consumption is close to the level of initial income, making the implicit saving and borrowing decisions more realistic. The results presented below have low sensitivity to this choice of \( \alpha_0 \).

The representative household makes consumption and savings decisions based on its expectation of current and future income. Unfortunately, under constant relative risk aversion, which forms the basis for the Consumption CAPM, there is no currently known solution to the optimal consumption problem (Zeldes, 1989). Therefore, a very simple
permanent income hypothesis model is used where the risk-free rate is constant at $r$.

Optimal consumption is then given by (Deaton, 1992):

$$c_t = \frac{r}{1 + r} \left( W_t + \sum_{\tau=0}^{\infty} (1 + r)^{-\tau} E_t (y_{t+\tau}) \right)$$

$$W_{t+1} = (1 + r)(W_t + y_t - c_t)$$

(A1)

where $W_t$ denotes capital wealth at time $t$. Given the income generation process, an optimal consumption path then follows:

$$c_t = \frac{r W_t}{1 + r} + \frac{r y_t}{1 + r} \left( 1 + \exp(b_t e_t) \sum_{\tau=1}^{\infty} \exp(\pi_0) (1 + r)^{-\tau} \sum_{s=1}^{2} Y_{st} P_s(s) \right)$$

$$Y_{st} = \begin{cases} E_t(\exp(e_{t+1}) | \delta_{st+1} = 1) & \tau = 1 \\ E_t(\exp \left( \sum_{g=1}^{t-1} (1 + b_t) e_{t+g} + e_{t+\tau} \right) | \delta_{st+1} = 1) & \tau > 1 \end{cases}$$

(A2)

To calculate the expectations in $Y_n$ for $\tau \leq 1,000$, a Monte Carlo method is used over 10,000 iterations. As in the main simulations, there is a 2% probability of a change in state each period. The real risk-free rate is taken as 2% and initial capital wealth set equal to zero.

To examine the relation between $P_1(1)$ and $\sigma^2_{\tau-1}(r_t)$, the MA(1) income process is simulated 500 times for 100 years' worth of data with a structural shift in income risk generated at the 50th data point. Based on these simulated MA(1) residuals, I repeat the standard process for estimating a structural break in risk and the probability of being in the two states, $P_s(s)$. For 446 of the 500 simulations, both the Chen and Gupta and Inclán and Tiao tests identify a structural decline in the volatility at a level of statistical significance of less than 1% and place the break in the interval [Year 40, Year 60].

For the 446 out of 500 simulations where a structural break is identified in income growth residuals in the interval [Year 40, Year 60] by both the Chen and Gupta and
Inclán and Tiao tests at the 1% significance level, the same tests are run on the consumption growth data estimated using equation (A2). In 436 of these 446 cases, both tests also recognize a decline in the volatility of consumption growth at the same level of significance and in the same interval. In 88% of these 436 cases, for both tests there is perfect agreement between the consumption and income series over the identified date of the break. New values of \( P_t(1) \) are also calculated based on the consumption series. The average correlation between the \( P_t(1) \) series as generated from the income and consumption data across the 446 simulations is 99.3% and in no case is the correlation less than 90%.

Given the similarity of estimates of \( P_t(1) \) from income and consumption data, it then follows that

\[
\sigma_s^2(r_{t+1}) \approx \sigma_s^2 + P_t(1)(\sigma_1^2 - \sigma_2^2)
\]  

(A3)

where \( \sigma_s \) refers to the consumption (not income) growth volatility in each state. To test this, for each year 1914–1986, the current optimal level of consumption is calculated based on the income level, the MA residual and the estimate of \( P_t(1) \) based on the P99–99.5 data for that period with \( \alpha_0 = -0.0035 \) and with current wealth set equal to zero. Then, for each year, 10,000 Monte Carlo simulations are run for next period's income. With probability \( P_t(1)(1-P_t(1)) \), \( e_{t+1} \) is drawn from a normal distribution with standard deviation 8.19% (3.06%). Based on this drawing, a new estimate of \( P_{t+1}(1) \) can be calculated. Given this, for each of these Monte Carlo runs, the optimal consumption next period can be calculated based on equation (A2). Based on this and next periods' projected consumption, for each year we have 10,000 simulations of \( r_{t+1} \).
Calculation of a time series of \( \sigma_s^2(r_{t+1}) \) over the period 1914–1986 is then straightforward. This series can then be used in the OLS regression, \( \sigma_s^2(r_{t+1}) \approx A + BP_t(1) \), over the interval 1914–1986, where the values of \( P_t(1) \) are again estimated from income data. The \( R^2 \) of this equation is 99.9%, showing the accuracy of the linear approximation. Empirical estimates of \( \sigma_s^2 \) can be derived directly from \( A \) and \( B \).

Within the framework of this article, time variation in the equity premium is driven by heteroskedasticity in consumption growth alone. From equation (1), it then follows that \( \rho_{t+k} = \xi \sigma_{t+k-1}(r_{t+k}) \) for constant \( \xi = \gamma \rho(R_t^r, r_t) \sigma(R_t^r) \). Equation (A3) implies that \( \rho_{t+k} = \xi \sqrt{A + BP_{t+k-1}(1)} \). Define \( \rho(s) = \xi \sigma_s \), which represents the single period equity premium that would currently prevail if next period's state was revealed as \( s \) immediately; \( P_{t+k-1}(s) = 1 \). Then the single period equity premium is given by:

\[
\rho_{t+k} = \sqrt{\rho^2(2) + P_{t+k-1}(1)\left(\rho^2(1) - \rho^2(2)\right)}
\]

(A4) Combining equations (8) and (A4) allows for the term structure of the equity premium discount factor to be calculated:

\[
\rho_{\beta} = E\left(\prod_{k=1}^{j} \frac{1}{1 + \sqrt{\rho^2(2) + P_{t+k-1}(1)\left(\rho^2(1) - \rho^2(2)\right)}}\right)
\]

(A5) As \( \rho(1), \rho(2) \) are constants, \( \rho_{\beta} \) is time dependent only through \( P_{t+k-1}(1) \) making it relatively easy to estimate this equation through Monte Carlo simulation. For all \( P_t(1) \in \{0\%, 1\%, 2\%, ..., 100\%\} \), with probability \( P_t(1) (1 - P_t(1)), e_{t+1} \) is drawn from a normal distribution with standard deviation 8.19% (3.06%). Based on this
drawing, $P_{t+1}(1)$ can be calculated in the usual way. This process is iterated to calculate $P_{t+j}(1)$ and all values of $j \in \{1,1000\}$. By repeating this process 5,000 times, $\rho_{j\mu}$ can be numerically estimated.

For asset pricing purposes, the key variables are not $\rho_{j\mu}$ themselves but the summation $\sum_{j=1}^{\infty} \Gamma_j \rho_{j\mu}$. The Monte Carlo results reveal that, to very high approximation ($R^2 > 99.5\%$) this summation is linear in $P_t(1)$:

$$\sum_{j=1}^{\infty} \Gamma_j \rho_{j\mu} = x_2 + P_t(1)(x_1 - x_2)$$

(A6)

$$x_s = \sum_{j=1}^{\infty} \Gamma_j \rho_{j\mu} | P_t(s) = 1$$

where $x_s$ are the Monte Carlo values from the cases $P_t(1)=1.0$. Intuitively, the accuracy of this approximation is unsurprising as it would hold exactly, through the law of conditional expectations, if $x_s$ were conditioned on $\delta_{s+1} = 1$ rather than $P_t(s) = 1$. As $\delta_{s+1}$ is not observable to investors, the conditioning in equation (A6) is more useful for calibration purposes and the linear approximation is highly precise.

By combining equations (A6) and (11), the main pricing equation; equation (13); follows directly. For calibration purposes, the values of $x_s$ are calculated using the Monte Carlo method described above.

*The Log-linear Model of Campbell & Shiller (1988)*

An alternative approach might involve the use of the log-linear model of Campbell and Shiller (1988):
\[ \ln(p_t) - \ln(d_t) \approx E_t \sum_{j=0}^{\infty} \theta^i (g_{t+j+1} - \phi_{t+j+1}) - \kappa \]

\[ \phi_{t+1} = \ln \left( \frac{p_{t+1} + d_{t+1}}{p_t} \right) \]

\[ \theta \approx \left( \frac{p_t}{p_t + d_t} \right) \]

\[ \kappa = \frac{\ln(\theta)}{1 - \theta} + \ln \left( \frac{1}{\theta} - 1 \right) \]

where \( \bar{()} \) denotes a sample average. It is necessary to be careful about the distinction between \( m_{t+j} \), which is the log of an expected single period simple return and \( \phi_{t+j} \), which is a log return: \( \exp(m_{t+j}) = E_{t+j-1}(\exp(\phi_{t+j})) \). If log returns are normally distributed with constant variance \( \sigma^2 \), \( E_{t+j-1}(\phi_{t+j}) = m_{t+j} - 0.5\sigma^2 \). By the law of iterated expectations, \( E_t(\phi_{t+j}) = E_t(E_{t+j}(\phi_{t+j})) = E_t(m_{t+j}) - 0.5\sigma^2 \), Campbell and Shiller's model can be rewritten as:

\[ \ln(p_t) - \ln(d_t) \approx E_t \sum_{j=0}^{\infty} \theta^i (g_{t+j+1} - m_{t+j+1} + 0.5\sigma^2) - \kappa \] (A7)

As for the main model, approximate \( m_{t+k} \approx \ln(1 + \rho_{t+k}) + r_{t+k} \) and, for calibration purposes, the value of \( \sigma^2 \) is estimated from the total sample. Under the MA(1) process for the dividend growth rate minus the risk-free rate, \( E_t(g_{t+1} - r_{t+1}) = \alpha_0 + \beta_t \eta_t \) and \( E_t(g_{t+j+1} - r_{t+j+1}) = \alpha_0 \) for all \( j > 0 \). Then:

\[ \ln(p_t) - \ln(d_t) \approx \beta_t \eta_t + \frac{\alpha_0 + 0.5\sigma^2}{1 - \theta} - E_t \sum_{j=0}^{\infty} \theta^i \ln(1 + \rho_{t+j+1}) - \kappa \] (A8)

Substituting for \( \rho_{t+j} \) from equation (A4) gives:

\[ \ln(p_t) - \ln(d_t) \approx \beta_t \eta_t + \frac{\alpha_0 + 0.5\sigma^2}{1 - \theta} - E_t \sum_{j=0}^{\infty} \theta^i \ln(1 + \rho_{t+j+1}) - \kappa \] (A9)
\[
\ln(p_t) - \ln(d_t) \approx \beta \eta_t + \frac{\alpha_0 + 0.5 \omega^2}{1 - \theta} - E \sum_{j=0}^{\infty} \theta^j \ln \left(1 + \sqrt{\rho^2(2) + P_{t+j}(1)(\rho^2(1) - \rho^2(2))}\right) - \kappa
\]  
(A10)

For each year in the sample, a Monte Carlo simulation is run to calculate the expectation based on that period’s value of \(P_t(1)\) using 10,000 iterations for \(j < 1,000\). The correlation between this simulated capital growth series and the baseline calibration is 99.7%. Price levels are similar, with the values from the Campbell and Shiller (1998) model being, on average, around 9% higher than those from the Ang and Liu (2004) model. The Ang and Liu (2004) is preferred for baseline calibrations as it is less reliant on historic data averages and so can be seen as more of a pure ex ante model.
References


Hamilton, J. D., 1989, A new approach to the analysis of nonstationary returns and the


Shiller, R.J., 2003, From efficient markets theory to behavioral finance, *Journal of


This table presents descriptive statistics for the real logarithmic growth rate in income for various income fractiles and for aggregate consumption. The income data are taken from Piketty and Saez (2003) and the consumption data from Robert Shiller's website. ***, **, * Significant at the 1%, 5% and 10% levels respectively for the Chi-Squared and Augmented Dickey-Fuller tests.
## TABLE 2: Structural Break in Volatility Tests for Income and Consumption Growth.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>81</td>
<td>81</td>
<td>81</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>69</td>
<td>111</td>
</tr>
</tbody>
</table>

### Panel A: First Identified Structural Break in Volatility; Log Growth
- Chen & Gupta p-value: <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%
- Inclan & Tiao p-value: <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%

### Panel B: First Identified Structural Break in Volatility; ARMA Residuals
- Chen & Gupta p-value: <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%
- Inclan & Tiao p-value: <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%, <1%

### Panel C: Pre and Post Break Volatility of ARMA Residuals
- Std before break: 8.88%, 5.33%, 6.18%, 8.19%, 10.81%, 12.87%, 20.09%, 7.58%, 4.03%
- Std after break: 2.90%, 2.24%, 2.61%, 3.06%, 4.43%, 4.69%, 7.19%, 2.63%, 1.53%

### Panel D: MSVAR Regimes for ARMA Residuals
- Break date: 1959, 1951, Multiple, 1951, 1951, 1952, 1941, 1951, 1938
- Std before break: 8.16%, 5.27%, - 7.97%, 10.26%, 12.48%, 18.83%, 7.34%, 3.90%
- Std after break: 2.31%, 2.19%, - 3.01%, 3.66%, 4.43%, 7.02%, 2.46%, 1.51%

This table presents evidence of single structural breaks in volatility for real annual income and aggregate consumption growth rate data. Tests used are Chen & Gupta (1997) and Inclan & Tiao (1994) based on logarithmic real growth and the ARMA residuals of log real growth. p-values for the rejection of the null of no structural break are given together with the date identified as being the most likely for the structural break in volatility. The date taken for the break is the average of the dates identified by the Chen & Gupta and Inclan & Tiao tests. Pre-break and post-break standard deviation estimates are provided. Panel D reports the structural breaks in volatility identified by the MSVAR package, written by Krolzig in Ox, that uses a maximum likelihood EM algorithm.
TABLE 3: Descriptive Statistics for Dividend Growth and the Risk-free Rate.

<table>
<thead>
<tr>
<th>1890-2000</th>
<th>$r_f$</th>
<th>$g$</th>
<th>$g - r_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>W&amp;J</td>
<td>GFD</td>
</tr>
<tr>
<td></td>
<td>1.30%</td>
<td>1.47%</td>
<td>1.61%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.70%</td>
<td>12.82%</td>
<td>12.23%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.55</td>
<td>-0.27</td>
<td>-0.26</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.71</td>
<td>4.18</td>
<td>3.37</td>
</tr>
<tr>
<td>Chi-Squared for normality</td>
<td>31.91***</td>
<td>45.42***</td>
<td>34.70***</td>
</tr>
<tr>
<td>Correlation with $g - r_f$</td>
<td>-31.44%</td>
<td>93.75%</td>
<td>93.09%</td>
</tr>
<tr>
<td>1st order autocorrelation</td>
<td>65.5%</td>
<td>15.2%</td>
<td>21.3%</td>
</tr>
<tr>
<td>Augmented Dickey-Fuller</td>
<td>-5.399***</td>
<td>-7.733***</td>
<td>-8.124***</td>
</tr>
</tbody>
</table>

This table presents descriptive statistics for the real risk-free rate, $r_f$, the real logarithmic growth in dividends, $g$, and the difference between these two variables, $g - r_f$. “W&J” = Wilson and Jones (2002). “GFD”=Global Financial Data. “Shiller” is Robert Shiller's unlagged data. *** Significant at the 1% level for the Chi-Squared and Augmented Dickey-Fuller tests.
### TABLE 4: Observed and Retrodicted Average Annual Growth Rates of the S&P500.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th></th>
<th>Panel A: WJ Data</th>
<th></th>
<th>Panel B: GFD Data</th>
<th></th>
<th>Panel C: Shiller+1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Partial</td>
<td>$p$-value</td>
<td>Full</td>
<td>$p$-value</td>
<td>Observed</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1897 – 1905</td>
<td>7.7%</td>
<td></td>
<td>6.2%</td>
<td>53.6%</td>
<td>7.4%</td>
<td>1.6%</td>
<td>7.4%</td>
</tr>
<tr>
<td>1906 – 1921</td>
<td>-6.5%</td>
<td></td>
<td>-3.6%</td>
<td>46.0%</td>
<td>-4.2%</td>
<td>7.6%</td>
<td>-4.4%</td>
</tr>
<tr>
<td>1922 – 1928</td>
<td>16.2%</td>
<td></td>
<td>7.8%</td>
<td>51.6%</td>
<td>12.0%</td>
<td>0.1%</td>
<td>10.3%</td>
</tr>
<tr>
<td>1929 – 1948</td>
<td>-4.0%</td>
<td></td>
<td>-0.5%</td>
<td>45.5%</td>
<td>-2.1%</td>
<td>0.1%</td>
<td>-0.2%</td>
</tr>
<tr>
<td>1949 – 1965</td>
<td>8.6%</td>
<td></td>
<td>3.9%</td>
<td>48.1%</td>
<td>6.9%</td>
<td>0.0%</td>
<td>5.5%</td>
</tr>
<tr>
<td>1966 – 1981</td>
<td>-5.0%</td>
<td></td>
<td>-1.1%</td>
<td>46.0%</td>
<td>-1.2%</td>
<td>35.6%</td>
<td>-1.1%</td>
</tr>
<tr>
<td>1982 – 1999</td>
<td>10.6%</td>
<td></td>
<td>5.3%</td>
<td>50.6%</td>
<td>5.3%</td>
<td>38.1%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1929 – 1932</td>
<td>-26.5%</td>
<td></td>
<td>-10.8%</td>
<td>40.3%</td>
<td>-19.0%</td>
<td>0.0%</td>
<td>-15.4%</td>
</tr>
<tr>
<td>1933 – 1936</td>
<td>22.1%</td>
<td></td>
<td>13.7%</td>
<td>56.6%</td>
<td>13.7%</td>
<td>56.2%</td>
<td>13.8%</td>
</tr>
<tr>
<td>1937 – 1948</td>
<td>-5.3%</td>
<td></td>
<td>-1.8%</td>
<td>45.0%</td>
<td>-1.8%</td>
<td>51.7%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1897 – 1948</td>
<td>22.3%</td>
<td></td>
<td>19.2%</td>
<td>20.9%</td>
<td>20.3%</td>
<td></td>
<td>22.7%</td>
</tr>
<tr>
<td>1949 – 1999</td>
<td>16.0%</td>
<td></td>
<td>11.6%</td>
<td>12.9%</td>
<td>12.3%</td>
<td></td>
<td>16.3%</td>
</tr>
</tbody>
</table>

Observed and estimated average annual real capital growth rates to the stock market over secular bull and bear markets of the twentieth century. “W&J” = Wilson and Jones (2002). “GFD”=Global Financial Data. “Shiller+1” is Robert Shiller's data lagged one year. “Partial” is retrodicted returns from equation (12). “Full” and “Consump.” are both based on equation (13), with the former being calibrated on a mix of income and consumption data, while the latter is calibrated on the basis of consumption data alone. $p$-values are calculated from Monte Carlo simulations with randomized consumption growth data.

<table>
<thead>
<tr>
<th></th>
<th>Panel A: Transition Probability</th>
<th>Panel B: Equity Premia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed $\zeta = 0.90$</td>
<td>$\zeta = 0.95$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1897 – 1905</td>
<td>7.7%</td>
</tr>
<tr>
<td></td>
<td>1906 – 1921</td>
<td>-6.5%</td>
</tr>
<tr>
<td></td>
<td>1922 – 1928</td>
<td>16.2%</td>
</tr>
<tr>
<td></td>
<td>1929 – 1948</td>
<td>-4.0%</td>
</tr>
<tr>
<td></td>
<td>1949 – 1965</td>
<td>8.6%</td>
</tr>
<tr>
<td></td>
<td>1966 – 1981</td>
<td>-5.0%</td>
</tr>
<tr>
<td></td>
<td>1982 – 1999</td>
<td>10.6%</td>
</tr>
<tr>
<td>Mean</td>
<td>1929 – 1932</td>
<td>-26.5%</td>
</tr>
<tr>
<td></td>
<td>1933 – 1936</td>
<td>22.1%</td>
</tr>
<tr>
<td></td>
<td>1937 – 1948</td>
<td>-5.3%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1897 – 1948</td>
<td>22.3%</td>
</tr>
<tr>
<td></td>
<td>1949 – 1999</td>
<td>16.0%</td>
</tr>
<tr>
<td></td>
<td>1897 – 1999</td>
<td>19.5%</td>
</tr>
</tbody>
</table>

This table is similar to Table 4 for the W&J data set and the full model. Panel A presents results for different transition probabilities between states; $\zeta$. The higher $\zeta$, the more persistent the state. Panel B considers different levels of the equity premium before ($\rho(1)$) and after ($\rho(2)$) the structural shift in macroeconomic volatility. All other variables are as for the base case.
Figure I. Estimated State Probabilities.

This figure provides the Bayesian probabilities of being in the high risk state, $P_t(1)$, next period as estimated from the three income series and the real aggregate consumption growth series.
Figure II. Retrodicted Price Series.

This is a replication of Figure 1 in Shiller (2003) from 1890 with the logarithm of retrodicted prices from the full and partial models added. The four lines from the Shiller exhibit show the logarithm of observed real prices and simulated prices from three models. The “PDV, Constant Discount Rate” model is that of Shiller (1981). The “PDV, Interest Rate” model allows for the discount rate to vary with short-term interest rates. The “PDV, Consumption” model is that of Grossman and Shiller (1981) that allows for the discount rate to vary with the marginal rate of substitution for real per-capita consumption.
Figure III. Retrodicted Price / Dividend Ratios.

This graph compares the retrodicted price / dividend ratios from the full and partial models against the observed values based on Shiller’s data. The graph also presents the “perfect knowledge” retrodicted values, where it is assumed that investors could identify instantaneously the regime change in 1949 without the requirement to undertake Bayesian learning.
Figure IV. Estimated State Probabilities 1959–1999.

The probability of being in the “Medium”, “Low” and High” risk macroeconomic states based on monthly real aggregate consumption growth rates of non-durables and services 1959M2–1999M12.
Figure V. Retrodicted Price Series, 1960–1999.

The observed logarithmic price of the market from 1960 to 1999 and the retrodicted log price based on quarterly consumption and dividend data. The probabilities used to estimate the changing equity premium are given in Figure IV.