

Why Surplus Consumption in the Habit Model May be Less Persistent than You Think ^{*}

Anthony W. Lynch[†]
New York University and NBER

Oliver Randall[‡]
New York University

First Version: 18 March 2009

This Version: 22 October 2009
Very preliminary. Comments welcome.

Do not cite or quote.

^{*}The authors would like to thank Stijn Van Nieuwerburgh and participants at two seminars at NYU for helpful comments and suggestions. All remaining errors are of course the authors' responsibility.

[†]Stern School of Business, New York University, 44 West Fourth Street, Suite 9-190, New York, NY 10012-1126, alynch@stern.nyu.edu, (212) 998-0350.

[‡]Stern School of Business, New York University, 44 West Fourth Street, Suite 9-190, New York, NY 10012-1126, orandall@stern.nyu.edu, (212) 998-0329.

Why Surplus Consumption in the Habit Model May be Less Persistent than You Think

Abstract

In U.S. data, value stocks have higher expected excess returns and higher CAPM alphas than growth stocks. This paper finds the external-habit model of Campbell and Cochrane (1999) can generate a value premium in both CAPM alpha and expected excess return when the log surplus-consumption ratio is allowed to be not very persistent. In contrast, Lettau and Wachter (2007) find that when the log surplus-consumption ratio is assumed to be highly persistent as in Campbell and Cochrane (by assuming that the price-of-risk state variable is highly persistent), the external-habit model generates a growth premium in expected excess return. However, there is a good economic reason for why the persistence of the log surplus-consumption ratio is likely to be low, and the micro evidence also favors a less persistent log surplus-consumption ratio. In particular, the high persistence assumed by Lettau and Wachter's specification implies that the contribution of the most recent 5 years of log consumption to log habit is just a little over 50% and so the contribution of log consumption more than 5 years ago is almost 50%, which seems very high. We choose a value for this persistence which is sufficiently low that the most recent 2 years of log consumption contribute over 98% of all past consumption to log habit, which is a much more reasonable number than the 25% contribution generated by the Lettau-Wachter value. In our specification, expected consumption is slowly mean-reverting, as in the long-run risk model of Bansal and Yaron (2004), which is why our model is able to generate a price-dividend ratio for aggregate equity that exhibits the high autocorrelation found in the data, despite the very low persistence of the price-of-risk state variable. Our results suggest that external-habit preferences and long-run risk consumption may both play important roles in explaining the time-series and cross-sectional properties of equity returns and prices. The one important dimension of equity return behavior that low persistence of the price-of-risk state variable cannot replicate is the predictability of long-horizon equity return using the price-dividend ratio for aggregate equity.

1 Introduction

A number of papers have considered how habit preferences impact the moments of aggregate equity price-dividend ratios and return and the moments of the riskfree rate. Early papers by Constantinides (1990) and Sundaresan (1989) showed how habit preferences could generate a higher equity premium for a given curvature parameter, γ . One issue with habit preferences has been its impact on the volatility of the riskfree rate: many specifications generate too much relative to what we see in U.S. data. Campbell and Cochrane (1999), hereafter CC, consider an economy with i.i.d. consumption and a representative agent with external habit preferences, and model the habit process in such a way as to produce a constant riskfree rate. They specify a process for the log consumption surplus which is defined to be the log of consumption in excess of habit scaled by consumption. The conditional volatility of the log surplus is specified to vary inversely with the log surplus so that the effect of variation in the log surplus on the riskfree rate due to the intertemporal substitution motive is exactly offset by its effect on the riskfree asset due to the precautionary saving motive. The implication is that the shock to the price of risk is close to perfectly negatively correlated with the shock to consumption growth in their specification. CC allow the log surplus to be a highly persistent process so that in their economy the price-dividend ratio is also highly persistent and long-horizon stock returns are forecastable using the price-dividend ratio. Both are features of U.S. data.

Recently Lettau and Wachter (2007), hereafter LW, consider how the correlation between the shock to the price of risk and the shock to log consumption growth affects the expected return differential between value and growth stocks, when the state variable driving the price of risk is highly persistent and the mean of consumption growth is a slowly mean-reverting process as in Bansal and Yaron (2004). They find that large negative correlation between the shock to the price-of-risk state variable and the shock to the consumption growth generates a growth premium for raw returns, in contrast to the value premium found in U.S. data. To produce a value premium, they set this correlation to zero. This finding raises the question whether habit preferences can generate a value premium as in U.S. data.

When the log surplus is as persistent as in CC and LW, the two most recent years of consumption contribute a much smaller fraction to the agent's habit level (the fraction is less than 26%) than all past consumption from more than two years ago, which seems counterintuitive and is likely counterfactual. The last two years of consumption would be expected to make a much larger

contribution to the agent's habit level than the sum of the contributions to the habit level by consumption from more than two years ago. Moreover, the 4 most recent years of consumption still contribute less to the agent's habit level than all past consumption from more than 4 years ago, even though making the cutoff 4 years would be expected to make the contribution of the more recent years even larger relative to a 2-year cutoff. Motivated by this intuition, our paper examines how a less persistent state variable for the price of risk, which would be implied by a less persistent log surplus ratio, affects the moments of the aggregate equity price-dividend ratio and return, and the expected return differential and CAPM-alpha differential between value and growth stocks. Roughly matching the data Sharpe ratio and expected price dividend ratio for aggregate equity, we find that when the persistence of the price-of-risk state variable is low, a large negative correlation between the shock to the price-of-risk state variable and the shock to log consumption growth can generate a value premium for expected excess returns and for CAPM alpha, consistent with U.S. data and in contrast to LW's findings when the persistence of the price-of-risk state variable is high. We also find that so long as the conditional mean of consumption growth is allowed to be slowly mean-reverting, as parameterized by LW and Bansal and Yaron based on U.S. data, the price dividend ratio exhibits first order autocorrelation comparable to that in U.S. data even when persistence of the price-of-risk state variable is low. This is because the expression for the price-dividend ratio for zero-coupon aggregate equity (which pays the aggregate market dividend at a given point in the future) suggests that the autocorrelation of the aggregate market's price dividend ratio is likely to be a weighted average of the autocorrelations of the conditional mean of log consumption growth and price-of-risk processes, and the mean of log consumption growth is still slowly mean-reverting.

When we force the consumption process to be the same as the aggregate dividend process and calibrate both to the dividend process for U.S. stocks, as in LW, we are unable to generate the aggregate equity return volatility found in the data. When we allow the consumption process to be different from the dividend process, we calibrate the consumption process to data while leaving the dividend process the same. This allows us to generate an even larger value premium in both expected excess returns and CAPM alpha relative to the case with consumption and aggregate dividends set equal, if we continue to allow the aggregate equity return volatility to be lower than that in the data. Even when we match aggregate equity return volatility to data, we are able to generate a value premium in expected excess return that is considerably larger than in the case with consumption and aggregate dividends set equal, and a value premium in CAPM alpha that is similar in magnitude to the case with consumption and aggregate dividends set equal.

One important dimension of equity return behavior that low persistence has difficulty replicating is the predictability of long-horizon equity returns using the price-dividend ratio. However, allowing the consumption and aggregate dividend to be different processes leads to equity return predictability that is qualitatively similar to that in the data but much smaller in magnitude. The low persistence also generates negative autocorrelation in aggregate equity return, especially when matching the equity return volatility to data, which is counterfactual.

The paper is also closely related to a recent paper by Santos and Veronesi (2008) which finds, like LW, that habit preferences and firm cash flows which are fractions of aggregate consumption flows, with value firms receiving larger fractions of these flows in the near future and growth firms receiving larger fractions in the distant future, deliver a growth premium rather than a value premium. Santos and Veronesi introduce cash flow heterogeneity across firms to obtain a value premium but find that the heterogeneity needed is too high relative to that found in the data. Also related is a paper by Bekaert and Engstrom (2009) that considers an economy whose representative agent has persistent external habit preferences. Their innovation is that log consumption growth is comprised of positively-skewed "good environment" shocks and negatively-skewed "bad environment" shocks, which allows them to match higher moments of the time series of asset returns. The paper has very little to say about the cross-section of expected returns.

The micro evidence in support of slow-moving habit is quite weak. Brunnermeier and Nagel (2006) test an implication of slow-moving habit that risky asset holdings as a fraction of financial wealth increase in response to wealth increases and find very little evidence in support of this hypothesis. In contrast, when habit moves rapidly in response to recent consumption, the hypothesized increase in risky asset holdings is much reduced, so this evidence does not contradict the presence of a habit that moves rapidly in response to recent consumption. The idea behind the hypothesis is the following. When habit is slow-moving, it is like a subsistence level. When utility is CRRA with subsistence level, the agent puts the present value of future subsistence levels into the riskless asset and the rest into the CRRA-optimal portfolio. When wealth increases, the entire increase is placed in the CRRA-optimal portfolio, causing the agent's risky asset holding as a fraction of financial wealth to increase. If habit is fast-moving, it will increase as consumption adjusts to the wealth increase. Consequently, the agent will only put a fraction of the wealth increase in the CRRA-optimal portfolio because the agent will be compelled to put a fraction of the wealth increase in the riskless asset to cover the habit increase. Hence, the increase in the agent's risky asset holding as a fraction of wealth in response to a wealth increase is much smaller when the habit is fast-moving rather than slow-moving in response to recent consumption.

With access to a unique credit-card panel data set, Ravina (2007) uses quarterly credit card purchases as a measure of quarterly consumption and then estimates a habit model in which a household's internal habit depends on its own consumption last quarter and external habit depends on current and last quarter's consumption in the city that the household lived. Testing a version of the habit model in which internal and external habit are subtracted directly from consumption in the utility function, Ravina finds that the coefficient of lagged own consumption in internal habit is 0.5 and the coefficient on current household city consumption in external habit is 0.29. Both these numbers are way too high to be consistent with the slow-moving habit assumed by CC, since slow-moving habit means that last period's consumption has very little effect on this period's habit. Dynan (2000) uses a similar methodology to Ravina but a different data set, namely annual PSID data, and finds coefficients on lagged own consumption that are insignificantly different from zero. However, Ravina's data set allows her to use household-specific financial information as controls in the estimation. Once Ravina omits these controls from the estimation, the coefficient on lagged own consumption drops to 0.10, a value similar to that obtained by Dynan.

Section 2 describes the model while section 3 presents the calibration details. Results are in section 4, and section 5 concludes.

2 The Model

The model has 4 shocks: a shock to dividend growth, a shock to expected dividend growth, a shock to the price of risk variable, and a shock to consumption growth. These shocks are assumed to be multivariate normal, and independent over time. Let D_t^m denote aggregate dividends at time t , and $d_t^m \equiv \log(D_t^m)$. It evolves as follows:

$$\Delta d_{t+1}^m = g^m + z_t^m + \varepsilon_{t+1}^m \quad (1)$$

with a time-varying conditional mean, $g^m + z_t^m$, where z_t^m follows an AR(1) process:

$$z_{t+1}^m = \phi_z z_t^m + \varepsilon_{t+1}^z \quad (2)$$

with $0 \leq \phi_z < 1$. Log consumption growth evolves as follows:

$$\Delta d_{t+1} = g + z_t + \varepsilon_{t+1}^d \quad (3)$$

where $g \equiv \frac{g^m}{\delta^m}$ and $z_t \equiv \frac{z_t^m}{\delta^m}$. The shock to dividend growth is composed of a levered version of the shock to consumption growth plus an additional shock: $\varepsilon_{t+1}^m \equiv \delta^m \varepsilon_{t+1}^d + \varepsilon_{t+1}^u$. This specification

allows separation between the aggregate dividend and aggregate consumption, with log dividend growth a levered version of log consumption growth as in Abel (1999). In the base case, we set log consumption growth equal to log market dividend growth by setting $\delta^m = 1$ and $\varepsilon^u = 0$. Let $\sigma_i^2 \equiv \sigma[\varepsilon^i]$ for $i = d, z, x, u$, and $\sigma_{i,j} \equiv \sigma[\varepsilon^i, \varepsilon^j]$ and $\rho_{i,j} \equiv \rho[\varepsilon^i, \varepsilon^j]$ for $i, j = d, z, x, u$.

The stochastic discount factor is driven by a single state variable x_t which also follows an AR(1) process:

$$x_{t+1} = (1 - \phi_x)\bar{x} + \phi_x x_t + \varepsilon_{t+1}^x \quad (4)$$

with $0 \leq \phi_x < 1$. We specify that only the shock to consumption growth is priced, and that the stochastic discount factor takes the form:

$$M_{t+1} = \exp \left\{ a + bz_t - \frac{1}{2}x_t^2 - \frac{x_t}{\sigma_d}\varepsilon_{t+1}^d \right\}. \quad (5)$$

Since the conditional log-normality of M_{t+1} implies that $E_t(M_{t+1}) = \exp\{a + bz_t\}$, the log of the risk-free rate from time t to $t + 1$ is given by:

$$r_t^f \equiv -a - bz_t \quad (6)$$

If $b \neq 0$, the riskless rate is time varying. Since the most relevant papers to ours, LW and CC, both assume that the risk-free rate is constant, in order to be able to compare to them we assume this too, i.e. $b = 0$.

2.1 Price-Dividend Ratio and Expected Returns for Zero-coupon Equity

Let $P_{n,t}^m$ be the time- t price of a claim to zero-coupon market equity, paying off in n periods. Following LW, it can be shown that $P_{n,t}^m$ takes the following recursive form:

$$\frac{P_{n,t}^m}{D_t^m} = F(x_t, z_t^m, n) = \exp\{A(n) + B_x(n)x_t + B_z(n)z_t^m\} \quad (7)$$

Using the boundary condition $P_{0,t}^m = D_t^m$ we see $A(0) = B_z(0) = B_x(0) = 0$, and proceeding by induction on n , we can show the following recursive relationships hold:

$$\begin{aligned} A(n) &= A(n-1) + a + g^m + B_x(n-1)\bar{x}(1 - \phi_x) + \frac{1}{2}(\delta^m)^2\sigma_d^2 + \delta^m\Sigma_{d,\varepsilon}(G_{n-1}^m)' \\ &\quad + \frac{1}{2}G_{n-1}^m\Sigma_{\varepsilon,\varepsilon}(G_{n-1}^m)' \end{aligned} \quad (8)$$

$$B_x(n) = \phi_x B_x(n-1) - \delta^m\sigma_d - \frac{1}{\sigma_d}\Sigma_{d,\varepsilon}(G_{n-1}^m)' \quad (9)$$

$$B_z(n) = \frac{(1+b)(1 - \phi_z^n)}{1 - \phi_z} \quad (10)$$

where $G_n^m \equiv [1 \ B_x(n) \ B_z(n)]$, $\Sigma_{d,\varepsilon} \equiv \sigma[\varepsilon^d, \varepsilon]$, and $\Sigma_{\varepsilon,\varepsilon} \equiv \sigma^2[\varepsilon]$ where $\varepsilon \equiv [\varepsilon^u \ \varepsilon^x \ \varepsilon^z]$.

Let $R_{n,t}^m$ be the time- t price of a claim to zero-coupon market equity, paying off in n periods, and define $r_{n,t}^m \equiv \log(R_{n,t}^m)$. It can be shown that (see LW):

$$r_{n,t+1}^m = E_t[r_{n,t+1}^m] + \delta^m \varepsilon_{t+1}^d + \varepsilon_{t+1}^u + B_z(n-1)\varepsilon_{t+1}^z + B_x(n-1)\varepsilon_{t+1}^x \quad (11)$$

$$\sigma_t^2[r_{n,t+1}^m] = C_{n-1}^m \Sigma (C_{n-1}^m)' \quad (12)$$

where $C_{n-1}^m \equiv [\delta^m, 1, B_z(n-1), B_x(n-1)]$ and $\Sigma = \sigma^2[\varepsilon^d \ \varepsilon^u \ \varepsilon^z \ \varepsilon^x]$.

We can show that the risk premium on a zero-coupon claim depends on B_z , B_x , x , the variance of the consumption shock and its covariances with the other shocks:

$$\begin{aligned} \log \left(E_t \left[\frac{R_{n,t+1}^m}{R_t^f} \right] \right) &= E_t[r_{n,t+1}^m - r_t^f] + \frac{1}{2} \sigma_t^2[r_{n,t+1}^m] \\ &= \left(\frac{\delta^m \sigma_d^2 + \sigma_{d,u} + B_z(n-1)\sigma_{z,d} + B_x(n-1)\sigma_{x,d}}{\sigma_d} \right) x_t \end{aligned} \quad (13)$$

Since $B_z(n)$ is positive for all n , it follows that the the conditional risk premium for n -period zero-coupon market equity increases monotonically with the covariance between shocks to z and d for all n . Moreover, $B_z(n)$ is increasing in n . So taking the covariance between shocks to z and d to be negative, the conditional risk premium evaluated at the unconditional mean of x_t is declining in n whenever the covariance between shocks to x and d is assumed to be zero. As reported in LW, this generates a value premium in expected excess returns because value stocks have shorter cash flow durations than growth stocks. Since $B_z(n)$ is positive for any n , a positive shock to z_{t+1} causes a positive shock to $P_{n,t+1}^m/D_{t+1}^m$ which causes a positive shock to $R_{n,t+1}^m$. When $\rho_{d,z}$ is taken to be negative, this positive shock to $R_{n,t+1}^m$ is typically associated with a negative shock to d_{t+1} which makes the zero-coupon market equity a hedge against shocks to aggregate consumption and causes its conditional premium to be lower than when $\rho_{d,z}$ is taken to be zero.

Turning to the covariance between shocks to x and d , its effect on the conditional risk premia for n -period zero-coupon market equity depends on the sign of $B_x(n)$. If $B_x(n)$ is negative, which is usually the case, then it follows that the conditional risk premium for n -period zero-coupon market equity decreases monotonically with the covariance between shocks to x and d for all n . If the correlation between shocks to x and d is close to -1, as the CC external habit model implies, the conditional risk premium for n -period zero-coupon market equity increases in the absolute value of $B_x(n)$ for all n . Moreover, the relation between the conditional risk premia for the n -period

zero-coupon market equity and its maturity n depends on how $B_x(n)\sigma_{x,d}$, which is positive, and $B_z(n)\sigma_{z,d}$, which is negative, vary with n . We have already seen that $B_z(n)\sigma_{z,d}$ is decreasing in maturity. Whether there is still a value premium when the correlation between shocks to x and d is close to -1 depends on how $B_x(n)\sigma_{x,d}$ varies with n . When the persistence of x is high, a shock to x today impacts the value of x for many periods in the future. Consequently, the absolute value of $B_x(n)$ increases monotonically for many periods into the future which causes a growth premium rather than a value premium. However, when the persistence of x is low, a shock to x today only affects the value of x for a few periods into the future. Consequently, the absolute value value of $B_x(n)$ increases monotonically for a few periods into the future before starting to decline. If the persistence of x is sufficiently low, this turning point can be sufficiently early that there is still a value premium in expected excess return. This intuition explains why the almost perfect negative correlation between shocks to x and d in our base and wedge cases is still able to generate a value premium due to the assumed low persistence of x .

2.2 Aggregate Equity Price Dividend Ratios and Returns

Aggregate equity is the claim to all future aggregate dividends. By the law of one price, a claim to aggregate equity is equal in price to the sum of the prices of zero-coupon market equity over all future horizons, from which we can calculate the price-dividend ratio as follows:

$$\frac{P_t^m}{D_t^m} = \sum_{n=1}^{\infty} \frac{P_{n,t}^m}{D_t^m} \quad (14)$$

Market returns can be calculated as a function of the market price-dividend ratio and dividend growth:

$$R_{t+1}^m = \frac{P_{t+1}^m + D_{t+1}^m}{P_t^m} \quad (15)$$

$$= \left(\frac{P_{t+1}^m/D_{t+1}^m + 1}{P_t^m/D_t^m} \right) \left(\frac{D_{t+1}^m}{D_t^m} \right) \quad (16)$$

Note that we simulate at a quarterly frequency, and we calculate annual returns by compounding quarterly returns. This approach is equivalent to reinvesting dividends at the end of each quarter and can be contrasted with the calculation of annual returns using annual price-dividend ratios, which is equivalent to assuming that dividends earn a zero net return within a year.¹

¹We reproduced all our tables using the return calculation that sums dividends within a year and the results that we obtained were very similar to the ones we report in the paper.

2.3 Relation to other models

This specification is related to a number of other models.

2.3.1 LW

LW don't distinguish between consumption and dividends and specify a stochastic discount factor of the form:

$$M_{t+1} = \exp \left\{ -r^f - \frac{1}{2} x_t^2 - \frac{x_t}{\sigma_d} \varepsilon_{t+1}^d \right\}$$

where r^f is the log of the risk-free rate, and is constant over time. Notice that our model nests LW by setting $a = -r^f$, $b = 0$, $\delta^m = 1$, and $\sigma_u = 0$.

2.3.2 CC with i.i.d. Consumption Growth

CC assume that a representative agent maximizes the utility function:

$$E \sum_{t=0}^{\infty} (\delta^{\text{cc}})^t \frac{(D_t^{\text{cc}} - H_t^{\text{cc}})^{1-\gamma^{\text{cc}}} - 1}{1 - \gamma^{\text{cc}}} \quad (17)$$

where D^{cc} is consumption, H^{cc} the level of external habit, and δ^{cc} is the subjective discount factor. Defining $d_t^{\text{cc}} \equiv \log(D_t^{\text{cc}})$ and $s_t^{\text{cc}} \equiv \log\left(\frac{D_t^{\text{cc}} - H_t^{\text{cc}}}{D_t^{\text{cc}}}\right)$, the log of the surplus-consumption ratio, they specify dynamics:

$$\begin{aligned} \Delta d_{t+1}^{\text{cc}} &= g^{\text{cc}} + \varepsilon_{t+1}^d \\ s_{t+1}^{\text{cc}} &= (1 - \phi_s^{\text{cc}}) \bar{s}^{\text{cc}} + \phi_s^{\text{cc}} s_t^{\text{cc}} + \lambda^{\text{cc}}(s_t^{\text{cc}}) \varepsilon_{t+1}^d \end{aligned}$$

for sensitivity function $\lambda^{\text{cc}}(s)$, where $\varepsilon^d \sim N(0, \sigma_d^2)$. This implies a stochastic discount factor equivalent to:

$$M_{t+1} = \exp \left\{ -\gamma^{\text{cc}} g^{\text{cc}} + \log(\delta^{\text{cc}}) + \gamma^{\text{cc}} (1 - \phi_s^{\text{cc}}) (s_t^{\text{cc}} - \bar{s}^{\text{cc}}) - \gamma^{\text{cc}} (1 + \lambda^{\text{cc}}(s_t^{\text{cc}})) \varepsilon_{t+1}^d \right\}$$

They specify the form of the sensitivity function as:

$$\lambda^{\text{cc}}(s_t^{\text{cc}}) = \begin{cases} \frac{1}{\bar{S}^{\text{cc}}} \sqrt{1 - 2(s_t^{\text{cc}} - \bar{s}^{\text{cc}})} - 1 & s_t^{\text{cc}} \leq s_{\text{max}}^{\text{cc}} \\ 0 & s_t^{\text{cc}} \geq s_{\text{max}}^{\text{cc}} \end{cases}$$

where $\bar{S}^{\text{cc}} \equiv \sigma_d \sqrt{\frac{\gamma^{\text{cc}}}{1 - \phi_s^{\text{cc}}}}$, $\bar{s}^{\text{cc}} \equiv \log(\bar{S}^{\text{cc}})$, and $s_{\text{max}}^{\text{cc}} = \bar{s}^{\text{cc}} + \frac{1}{2}(1 - (\bar{S}^{\text{cc}})^2)$.

Notice that our model approximates CC by setting $a = \log(\delta^{\text{cc}}) - \gamma^{\text{cc}}g^{\text{cc}} + \frac{\gamma^{\text{cc}}(1-\phi_s^{\text{cc}})}{2}$, $g = g^{\text{cc}}$, $\delta^m = 1$, $\sigma_u = 0$, and $\sigma_z = 0$, so z is a zero process. This implies $x_t = \gamma\sigma_d(1 + \lambda^{\text{cc}}(s_t^{\text{cc}}))$ which we approximate in our model as a homoskedastic AR(1) process. As long as the sensitivity function is rarely zero it follows that $\rho_{d,x} \approx -1$ and $\phi_x \approx \phi_s^{\text{cc}}$.

2.3.3 CC with Persistent Conditional Mean Consumption Growth

With $\sigma_z \neq 0$, the model approximates CC with persistent conditional mean consumption growth. Suppose the representative agent again maximizes the habit specification in equation (17) but the conditional mean of aggregate consumption growth is slowly mean-reverting, following equations (2) and (3). We extend the dynamics of the log consumption surplus when there is long run risk so that the sensitivity function still loads on the innovation to log consumption growth above its mean, $\Delta d_{t+1} - g$, which is equal to $z_t + \varepsilon_{t+1}^d$ when we include long run risk. Specifically we assume the consumption surplus evolves as follows:

$$s_{t+1} = (1 - \phi_s)\bar{s} + \phi_s s_t + \lambda^{\text{cc}}(\bar{s})z_t + \lambda^{\text{cc}}(s_t)\varepsilon_{t+1}^d \quad (18)$$

where $\lambda^{\text{cc}}(\cdot)$ is the same sensitivity function as used by CC which is described in the previous subsection.

By approximating the sensitivity function's loading on z_t by its steady state value, this extension allows the risk-free rate to depend only on z_t . It also allows habit to be predetermined only by an exponentially-weighted sum of past lagged consumption when evaluated at the consumption surplus's steady state (see section 2.4 below), a desirable feature. This specification implies the following stochastic discount factor:

$$M_{t+1} = \exp\{-\gamma^{\text{cc}}g + \log(\delta^{\text{cc}}) - \gamma^{\text{cc}}(1 + \lambda^{\text{cc}}(\bar{s}))z_t + \gamma^{\text{cc}}(1 - \phi_s)(s_t - \bar{s}) - \gamma^{\text{cc}}(1 + \lambda^{\text{cc}}(s_t))\varepsilon_{t+1}^d\}$$

Matching coefficients in the stochastic discount factor we see that the risk-free rate is affine in z_t . So our model approximates CC with persistent conditional mean consumption growth by setting $a = \log(\delta^{\text{cc}}) - \gamma^{\text{cc}}g + \frac{\gamma^{\text{cc}}(1-\phi_s)}{2}$, $b = -\gamma^{\text{cc}}(1 + \lambda^{\text{cc}}(\bar{s}))$, $\delta^m = 1$, and $\sigma_u = 0$.

2.3.4 Power Utility with Persistent Mean Consumption Growth

When $\sigma_x = 0$, we see $x_t \equiv \bar{x}$, and the model reduces to a representative agent with power utility:

$$E \sum_{t=0}^{\infty} \delta^t \frac{(D_t)^{1-\gamma}}{1-\gamma}$$

Again, the conditional mean of aggregate consumption growth is slowly mean-reverting, following equations (2) and (3). This specification implies the following stochastic discount factor.

$$M_{t+1} = \exp\{-\gamma g + \log(\delta) - \gamma z_t - \gamma \varepsilon_{t+1}^d\}$$

2.4 Relation between External Habit and Past Consumption

Following an earlier version of CC, and extending their model such that consumption growth exhibits a persistent conditional mean, we can show that log habit is approximately a moving average of lagged log consumption. Define $h_t \equiv \log(H_t)$, and apply a log-linear approximation to the definition of s_t :

$$\begin{aligned} s_t &= \log\left(1 - e^{h_t - d_t}\right) \\ &\approx \log\left(1 - e^{\overline{h-d}}\right) + [(h_t - d_t) - (\overline{h-d})] \left(\frac{-e^{\overline{h-d}}}{1 - e^{\overline{h-d}}}\right) \end{aligned}$$

Substituting this in to the law of motion for s described in (18), and setting $\lambda(s_t) \approx \lambda(\bar{s}) = \frac{e^{\overline{h-d}}}{1 - e^{\overline{h-d}}}$, so that h_{t+1} is predetermined in the sense that it depends on d_t, d_{t-1}, \dots but not d_{t+1} , we can show that:

$$h_{t+1} \approx \overline{h-d} + (1 - \phi_s) \sum_{j=0}^{\infty} (\phi_s)^j d_{t-j} + \frac{g}{1 - \phi_s} \quad (19)$$

This is precisely the same expression derived in an earlier version of CC, in which consumption growth is assumed i.i.d.. Almost by definition, habit should only depend on lagged consumption so this is an attractive feature of the specification for s_t given in equation (18) when consumption growth has a persistent conditional mean as in Bansal and Yaron. We can also derive an expression for the innovation to habit, which is a function of how far consumption is above habit:

$$h_{t+1} - h_t \approx g + (1 - \phi_s) [(d_t - h_t) - \overline{d-h}] \quad (20)$$

The lower the persistence of the surplus-consumption ratio, the more impact the most recent consumption has on habit.

These expressions highlight a point made in the introduction, namely that when habit is slow-moving with ϕ_s close to 1, recent consumption contributes very little to current habit. The coefficient on log lagged consumption, d_t , in the expression for log habit, h_{t+1} , in equation (19) is $(1 - \phi_s)$. So when ϕ_s is close to 1, as in CC, this coefficient is close to 0. This expression for habit

shows clearly how the large coefficient on lagged own consumption obtained by Ravina is consistent with habit being fast-moving in response to recent consumption.

2.5 Specifying the Share Process

We follow LW and specify that the market is made up of 200 firms that generate dividends which aggregate to the market dividend. The share of the aggregate produced by each firm is set deterministically. Let \underline{s} be the minimum share of any firm. Without loss of generality suppose firm 1 produces this share initially. Its share increases by 5.5% a quarter for 100 quarters to a maximum share of $1.055^{100}\underline{s}$, then declines at the same rate for 100 quarters such that its share after 200 quarters exactly equals its initial share. Firm 2 starts at the second point in the cycle, and so on, so that each firm is at a different point in the cycle at any time. Here \underline{s} is set so that the shares of the 200 firms add up to 1 at all times. So firm i , with share s^i of the aggregate dividend, pays a dividend $s_t^i D_t$ at time t . LW choose a growth rate of 5.5% per quarter for the share process so that the cross-sectional distribution of dividend growth rates in the model matches that in the sample.

The law of one price determines that firm i 's ex-dividend price equals:

$$P_t^i = \sum_{n=1}^{\infty} s_{t+n}^i P_{n,t}^m \quad (21)$$

Quarterly returns for individual firms can be calculated similarly to the market, as a function of the firm's quarterly price-dividend ratio and quarterly dividend growth. Annual returns are calculated as described above, by compounding the quarterly returns.

2.6 Forming the Value/Growth Deciles

Following LW, we specify a period in the the model to be a quarter. At the start of each year, we sort firms into deciles from value to growth based on their annual price-dividend ratios, which are given by $P_t^j / \sum_{\tau=0}^3 D_{t-\tau}^j$ for firm j . We calculate moments for the decile excess annual returns and annual CAPM α by simulating the model at a quarterly frequency and then compounding the quarterly firm returns to obtain annual firm returns, as described above.

3 Calibration

As a comparison point, we first implement the calibration in LW using the parameter values provided in the paper.² Both the LW case and our base case assume the aggregate consumption process is the same as the aggregate dividend process. We also consider two wedge cases in which the aggregate consumption process is allowed to differ from the aggregate dividend process. Table 1 reports the parameters used by LW and in the calibrations for our base case and our two wedge cases, which all use exactly the same calibration for the z^m process, Δd^m process and r^f as used by LW.

Our base case and two wedge cases depart from LW in the calibration of the parameters of the price-of-risk state variable, the x process. The external habit model of CC implies a value close to -1 for $\rho[\varepsilon^d, \varepsilon^x]$ but LW show that at their chosen parameter values, a large negative value for this correlation generates a growth rather than value premium in expected return. For this reason, they set this correlation equal to 0 and are able to generate a value premium for both expected return and CAPM alpha. However, one of the main goals of our paper is to show that a value premium is possible for both expected return and CAPM alpha when this correlation is close to -1 so long as the price-of-risk state variable is not too persistent. For this reason, we set this correlation to -0.99 in the base case and the two wedge cases.³

While the model is quarterly, the log riskfree rate r^f is converted into an annual number by multiplying by a factor of 4. We express the persistence parameters ϕ_x and ϕ_z at annual frequencies by raising each of them to the power of 4.

3.1 Calibration of the base case: the x process

To ensure the covariance matrix of $(\varepsilon^d, \varepsilon^z, \varepsilon^x)$ is positive definite, we specify $\sigma[\varepsilon^x, \varepsilon^z]$ so that ε^x and ε^z are correlated *only* through their correlations with ε^d , i.e. $\sigma[\varepsilon^x, \varepsilon^z]$ is calculated as follows: 1) Regress ε^d on ε^z , yielding $\varepsilon^d = \beta_{d,z}\varepsilon^z + u^d$ where $\rho[\varepsilon^z, u^d] = 0$; and, 2) Regress ε^x on ε^d , yielding

²The values reported in LW are likely subject to rounding which explains why our parameter values are slightly different from those reported in LW.

³Choosing -0.99 instead of 1 seems unimportant since the base case results are unaffected by setting this correlation to -0.995 or -0.999.

$\varepsilon^x = \beta_{x,d}\varepsilon^d + u^x$ where $\rho[\varepsilon^d, u^x] = 0$. The following expression can be derived:

$$\begin{aligned}\sigma[\varepsilon^x, \varepsilon^z] &= \sigma\left[\beta_{x,d}\beta_{d,z}\varepsilon^z + \beta_{x,d}u^d + u^x, \varepsilon^z\right] \\ &= \left(\rho[\varepsilon^d, \varepsilon^x]\rho[\varepsilon^d, \varepsilon^z] + \left[1 - \rho[\varepsilon^d, \varepsilon^x]^2\right]^{\frac{1}{2}}\rho[u^x, \varepsilon^z]\right)\sigma_z\sigma_x\end{aligned}\quad (22)$$

When $\rho[\varepsilon^d, \varepsilon^x] = -0.99$, the chosen value for $\rho[u^x, \varepsilon^z]$ does not much affect $\rho[\varepsilon^x, \varepsilon^z]$ or $\sigma[\varepsilon^x, \varepsilon^z]$, so we use (22) with $\rho[u^x, \varepsilon^z] = 0$ to calculate $\sigma[\varepsilon^x, \varepsilon^z]$. Notice this specification has the attractive property that when $\sigma[\varepsilon^d, \varepsilon^x]$ is set equal to 0, $\sigma[\varepsilon^x, \varepsilon^z]$ is set equal to 0 as well. Since $\rho[\varepsilon^d, \varepsilon^z]$ is set equal to -0.82, the assumed value for $\rho[\varepsilon^x, \varepsilon^z]$ is 0.82. Note that this correlation measures the correlation of the shock to future expected returns with the shocks to future expected consumption growth and future expected dividend growth.

The next parameter of the x process to be calibrated is the persistence parameter. LW calibrate the autocorrelation of x to equal the data autocorrelation of the log price-dividend ratio for the aggregate market (0.87 annually), arguing that since the variance of expected dividend growth ($g^m + z_t^m$) is small, the autocorrelation of the log price-dividend ratio is primarily driven by the autocorrelation of x . However, the expression for the price-dividend ratio for zero-coupon aggregate equity (7) in section 2 suggests that the autocorrelation of the aggregate market's price dividend ratio is likely to be a weighted average of the autocorrelations of the z and x processes. So the fact that the z process is highly persistent, with an annualized autocorrelation of 0.91, means that it may be possible to have an x process that is not very persistent and still have a log price-dividend ratio for the aggregate market with an annualized autocorrelation of 0.87.

Moreover, there are good theoretical reasons for why the x process might not be very persistent. In particular, it is easy to show that the CC model implies that the persistence of our price-of-risk state variable x is approximately equal to the persistence of the log surplus s in their model. While CC themselves use a very large value for the autocorrelation of the log surplus in their model, the use of such a large value implies that habit depends much less on the consumption in the recent past than consumption in the distant past. For example, Table 2 uses the expression in (19) that relates log habit to past log consumption in CC to calculate the contribution of lagged log consumption to log habit when x 's persistence parameter is set equal to the LW annualized value of 0.87 and to the value in our base and wedge cases. At the LW value, the contribution of the most recent 5 years is just a little over 50% and so the contribution of log consumption more than 5 years ago is almost 50% which seems very high. We choose an annualized value for ϕ_x of 0.14 which is sufficiently low that the most recent 2 years of log consumption contribute over 98% of all past consumption to log habit, which is a much more reasonable number than the 25% contribution generated by the

LW value. Section 2.1 discussed the intuition for why a value premium can be generated by an x -variable whose shock is highly negatively correlated with the d -variable shock so long as it is not too persistent.

The remaining parameters of the x process left to calibrate are its mean \bar{x} and its conditional volatility σ_x . LW calibrate \bar{x} such that the maximum conditional quarterly Sharpe ratio $\sqrt{e^{\bar{x}^2} - 1}$ equals 0.70, which corresponds to $\bar{x} = 0.625$. They calibrate σ_x to match the volatility of the price-dividend ratio for aggregate equity. When choosing \bar{x} and σ_x , we concentrate on matching the mean rather than the volatility of the price-dividend ratio for aggregate equity, in addition to the unconditional Sharpe ratio for aggregate equity in the data. Both the unconditional Sharpe ratio and the expected price-dividend ratio for aggregate equity move positively with both \bar{x} and σ_x . We choose our \bar{x} and σ_x to produce a Sharpe ratio that roughly corresponds to the 4.1 value obtained by LW and an expected price-dividend ratio for aggregate equity whose mean absolute error relative to the data value is similar to that obtained by LW. The data value of the Sharpe ratio, at 0.33, is a little lower than the values obtained by LW and our base case. While the LW value for the expected price-dividend ratio is about 5.5 lower than the data value of 25.55, the value obtained by our base case is about 5.5 higher than the data value.

3.2 Calibration of the wedge cases: distinguishing between consumption and dividends

In the base case we do not make a distinction between dividend growth and consumption growth, i.e. $\delta^m = 1$ and $\sigma_u^2 = 0$. In the two wedge cases we do make this distinction, and consider log dividend growth to be a levered version of log consumption growth. The two wedge cases both have the same processes for log dividend growth and log consumption growth but each has a different specification for the x process.

We keep the volatility of ε^m and the covariance of ε^m with ε^z the same as in the base case, matching the following to LW's data moments:

$$\sigma[\varepsilon^m] = (\delta^m)^2 \sigma_d^2 + \sigma_u^2 + 2\delta^m \sigma_{d,u} \quad (23)$$

$$\sigma[\varepsilon_{t+1}^m, \varepsilon_{t+1}^z] = \delta^m \sigma_{d,z} + \sigma_{u,z} \quad (24)$$

We can get a closed-form expression for the annual covariance of log consumption and dividend growth:

$$\begin{aligned}
& \sigma \left[\sum_{i=1}^4 \Delta d_{t+i}, \sum_{i=1}^4 \Delta d_{t+i}^m \right] \\
&= \frac{1}{(1-\phi_z)^2 \delta^m} \left((1-\phi_z^4)(1+\phi_z^2) + 3 - \frac{2\phi_z(1-\phi_z^3)}{1-\phi_z} + \frac{\phi_z^2(1-\phi_z^6)}{1-\phi_z^2} \right) \sigma_z^2 \\
& \quad + \frac{2}{(1-\phi_z)^2} (3-4\phi+\phi^4) \sigma_{z,d} + \frac{1}{\delta^m(1-\phi_z^2)} (3-4\phi_z+\phi_z^4) \sigma_{z,u} + 4\delta^m \sigma_d^2 + 4\sigma_{d,u}
\end{aligned} \tag{25}$$

The annual correlation of log consumption growth with log dividend growth is 0.55 in Bansal-Yaron's sample period. This value for the annual correlation requires $\bar{x} < 0$ for the price-dividend ratio to converge, which is a problem since the x process is positive in CC. The correlation of log consumption growth with log dividend growth at a quarterly frequency is a simple expression: $\rho[\varepsilon_{t+1}^m, \varepsilon_{t+1}^d] = \frac{\delta^m \sigma[\varepsilon_{t+1}^d]}{\sigma_m \sigma_d} + \frac{\sigma_u \rho[\varepsilon_{t+1}^d, \varepsilon_{t+1}^u]}{\sigma_m}$. Simulations suggest that the annual and quarterly correlations are very similar, at least for the range of parameter values we consider, so we focus on the quarterly number because its expression is much simpler. Since the x process is positive in CC, we instead chose a larger correlation than in the data, 0.82 at a quarterly frequency, for which the price-dividend ratio converges for a range of $\bar{x} > 0$.

Using the methods of Stambaugh (1997) and Lynch and Wachter (2008) and given the volatility of annual log consumption and dividend growth and their correlation in the Bansal-Yaron sample period (1929-1998), and the volatility of annual log dividend growth for the LW sample period (1890-2002), we can estimate the volatility of annual log consumption growth in the LW sample period. The Bansal-Yaron moments allow us to regress annual log consumption growth on annual log dividend growth, estimating the regression coefficient and the variance of the residuals. Using these and the volatility of annual log dividend growth for the LW sample period, we can back out an estimate for the volatility of annual log consumption growth for this period. This comes out to be 3.18%, and we matched this to our analytical expression for the variance of annual log consumption growth:

$$\begin{aligned}
& \sigma^2 \left[\sum_{i=1}^4 \Delta d_{t+i} \right] \\
&= \left(\frac{1}{(1-\phi_z) \delta^m} \right)^2 \left((1-\phi_z^4)(1+\phi_z^2) + 3 - \frac{2\phi_z(1-\phi_z^3)}{1-\phi_z} + \frac{\phi_z^2(1-\phi_z^6)}{1-\phi_z^2} \right) \sigma_z^2 \\
& \quad + \frac{2}{\delta^m(1-\phi_z)^2} (3-4\phi_z+\phi_z^4) \sigma_{z,d} + 4\sigma_d^2
\end{aligned} \tag{26}$$

Typically in the literature δ^m is set equal to $\frac{\sigma_m}{\sigma_d}$. Our set-up allows δ^m to be different from this, but we chose this value as the natural point of departure. Since LW calibrate their dividend/consumption process to U.S dividend data, we keep the joint $\{z^m, \Delta d^m\}$ process the same, i.e. $\phi_z, \sigma_z, \sigma_m, g^m, \rho_{m,z}$ are unchanged from the base case. We set $\rho[\varepsilon^d, \varepsilon^z] = \rho[\varepsilon^m, \varepsilon^z]$ which has

a couple of attractive features in our setting. First, there is an asset in the wedge cases with the same cash-flows and price as produced by the market dividend in the base case. Second, given σ_z and σ_m fixed and $\delta^m = \frac{\sigma_m}{\sigma_d}$, then as $\rho[\varepsilon^d, \varepsilon^m]$ tends to 1, the pricing implications for the two wedge cases, in which aggregate consumption and dividends are allowed to differ, converge to those for our base case in which the two are the same.

Given a δ^m value and $\rho[\varepsilon^d, \varepsilon^z] = \rho[\varepsilon^m, \varepsilon^z]$, the system of equations defined in (20)-(23) yields $\sigma_d, \sigma_u, \sigma_{d,u}$ and $\sigma_{z,u}$. The resulting σ_d can be used to calculate $\frac{\sigma_m}{\sigma_d}$, which becomes the new δ^m value. We iterate until convergence, namely, until the obtained $\frac{\sigma_m}{\sigma_d}$ value equals the δ^m used to obtain it.

Turning to the x process, $\rho[\varepsilon^d, \varepsilon^x] = -0.99$ as in the base case, in the spirit of CC. The covariances $\sigma[\varepsilon^x, \varepsilon^z]$ and $\sigma[\varepsilon^x, \varepsilon^u]$ are set to ensure that the covariance matrix of $(\varepsilon^d, \varepsilon^z, \varepsilon^x, \varepsilon^u)$ is positive definite. As in the base case, $\sigma[\varepsilon^x, \varepsilon^z]$ is defined in equation (22), while $\sigma[\varepsilon^x, \varepsilon^u]$ is calculated similarly:

$$\sigma[\varepsilon^x, \varepsilon^u] = \rho[\varepsilon^d, \varepsilon^x] \rho[\varepsilon^d, \varepsilon^u] \sigma_u \sigma_x$$

4 Results

This section reports the results for the LW, base and two wedge cases, as well as some comparative statics. We also report results for U.S. data. The data for all but the value and growth portfolios are the same as that in LW and are annual from 1890-2002. The data for the value and growth portfolios are monthly from Ken French's website and span 1952 to 2002: means are annualized by multiplying by 12 and volatilities by $12^{0.5}$. The models are simulated at a quarterly frequency for 4 million quarters, by which point the reported moments and statistics for the aggregate market, the predictive regressions and the value and growth portfolios have converged, i.e. the first and second halves of the simulation yield sufficiently similar moments. When calculating the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return on zero-coupon equity with n years to maturity, the models are simulated at a quarterly frequency for 4 million quarters, or until convergence, i.e. until the first and second halves of the simulation yield sufficiently similar moments, whichever comes later.

4.1 Base Case

This subsection discusses the results for the base case and compares them to the results from the data and for the LW case. As in LW, the aggregate consumption process is assumed to be equal to the market dividend process, which is calibrated to data. The market dividend process used in the base case is the same as that used by LW. The correlation between the change in log consumption and the price-of-risk variable x is 0 in LW but is set to -0.99 in the base case, consistent with habit preferences. As discussed in section 3, the persistence of the x process is set to 0.14 annualized, which is a much lower value than the 0.87 annualized used by LW based on CC. Recall from section 3 that in the base case, the remaining x parameters are chosen to match the mean of the equity price-dividend ratio and its unconditional Sharpe ratio.

4.1.1 Aggregate Moments

Table 3 reports moments for aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth. The first column reports moments for the data, the second reports simulated moments for the LW case, and the third reports simulated moments for the base case. Returns, dividends, and price-dividend ratios are aggregated to annual frequencies: so $P_t^m/D_t^m = P_t^m / \sum_{\tau=0}^3 D_{t-\tau}^m$ and $p_t^m - d_t^m \equiv \log(P_t^m/D_t^m)$. $Sharpe^m$ is the unconditional Sharpe ratio of the aggregate equity, and AC is the autocorrelation.

Since the parameters of Δd^m were chosen to match the data, it is to be expected that the base case is able to closely match the autocorrelation and unconditional volatility of Δd^m . And as discussed above, the expected price dividend ratio obtained from the base case is as close to the data value as the LW value and the base-case unconditional Sharpe ratio is virtually the same as the LW value. But these close matches are to be expected and are not evidence of the base case's ability to match data moments since the parameters of the x process not nailed down by the habit specification were chosen to match these parameters. However, because parameter values were not chosen specifically to match the autocorrelation of the price-dividend ratio in the data, it is impressive the base-case value of this autocorrelation, 0.896, is higher than, but close to, both the data value of 0.87 and the LW value of 0.882.

While the base case is calibrated to match the unconditional Sharpe ratio, it delivers an expected excess market return and a market excess return volatility that is too low relative to the data and

LW values. The delivered volatility of 10.69% is particularly low relative to the data volatility of 19.41% which LW does a good job matching. The base case also delivers a price-dividend ratio volatility, 0.260, that is much lower than the data value of 0.38. Again, the LW case matches the data moment quite closely, delivering a value of 0.381. Excess market returns also exhibit negative autocorrelation of -0.13 which is counterfactual: the data value is 0.03. The LW case also delivers excess market returns that are negatively autocorrelated, though the magnitude is much smaller at -0.04.

4.1.2 Predictive Regressions

Table 4 reports results for the predictive regressions. The top panel is the regression of the future log excess return on the aggregate equity on the log aggregate equity price-dividend ratio today. The middle panel is the regression of future changes in log aggregate equity dividend on the log aggregate equity price-dividend ratio today. The bottom panel is the regression of future changes in log aggregate equity dividend on the consumption-aggregate equity dividend ratio for the data and z^m today for the models. The first column reports results for the data, the second reports results for the LW case, and the third reports results for the Base case. Log returns, log dividend growth, log price-dividend ratios and log consumption-aggregate dividend ratios are all aggregated to annual frequencies, as described above. Results are reported for horizons, H , of 1 and 10 years. R^2 is the regression R^2 .

Perhaps the most glaring inability of the base case to match data moments concerns the excess market return predictability regressions. The data and the LW case deliver R^2 s and negative predictability coefficients that are both larger in absolute value at a 1 year return horizon than at a 10 year return horizon. While the base case is able to produce negative predictability coefficients, their magnitudes are much smaller than those observed in the data, and the base-case R^2 s at horizons of 1 and 10 years are both negligible. As in LW, the base case does a poor job of reproducing the predictability of 1 or 10 year log dividend growth found in the data using either the log price-dividend ratio or a proxy for z^m , especially at the 10 year horizon for log dividend growth.

4.1.3 Value vs Growth Portfolios

Table 5 reports results for the extreme growth decile (portfolio 1), the extreme value decile (portfolio 10), and the portfolio which is long portfolio 10 and short portfolio 1 (the HML portfolio). The top panel reports expected excess annual return, the volatility of excess annual return and the unconditional Sharpe ratio for annual return. The bottom panel reports CAPM alpha, CAPM beta and regression R^2 using annual returns. The first three columns report results for the data, the fourth reports results for the LW case, and the fifth reports results for the base case. For the data, the first column obtains the extreme portfolios by sorting on earnings yield (E/P), the second column by sorting on equity cash flow to market value (C/P), and the third column by sorting on equity book-to-market value (B/M). For the models, we sort the 200 firms into deciles at the start of each year from value to growth based on their annual price-dividend ratios, which are given by: $P_t^j/D_t^j = P_t^j / \sum_{\tau=0}^3 D_{t-\tau}^j$ for firm j . $Sharpe^i$ is the unconditional Sharpe ratio for portfolio i 's annual return while R^2 is the regression R^2 .

Table 5 shows that the base case can generate a positive value premium in both excess return and CAPM alpha, though the magnitudes of the two are less than those found in the data or delivered by the LW case. In the data, using B/M to sort stocks into deciles, the excess return spread between the value and the growth portfolio is 4.88% versus the 1.91% per annum delivered by the base case. Similarly, the CAPM-alpha spread for these two extreme book-to-market deciles is 5.63% per annum in the data, but only 1.20% per annum in the base case. Moreover, both the data and LW deliver a CAPM-alpha spread between the extreme value and the growth deciles that is larger than the excess return spread, while the converse is true for the base case. The reason is that the CAPM beta for the extreme value decile is lower than for the extreme growth decile for the data and LW, while the converse is true for the base case, as the rows labeled β_i in Table 5 show. The base case delivers excess return volatility that is higher for the extreme value decile than for the extreme growth decile, which is consistent with the data when B/M is used to construct the extreme deciles, but inconsistent when E/P or C/P is used. The volatility numbers for the two extreme deciles and especially for HML are much lower for the base case than for the data. The LW case also delivers lower volatility for HML than the data, though not as low as delivered by the base case. The implication is the returns on the extreme deciles are much more correlated for the two cases, especially the base case, than for the data. Consistent with this observation, the R^2 s of the CAPM market model regressions for the two extreme deciles are typically largest for the base case and smallest for the data, with the LW case in the middle. The R^2 s of the CAPM market

model regressions for HML are below 15% for all the data sorts and the base and LW cases. The unconditional Sharpe ratio is largest for the extreme value decile, for the two cases and all the data sorts, and is lowest for the extreme growth decile, for all but the LW case.

The main message of Table 5 is that the base case can deliver a value premium both in excess return and CAPM alpha. To better understand why the base case delivers a value premium in excess return, we now turn our attention to the zero-coupon market dividend claims described in section 2. Figure 1 plots, as a function of maturity for zero-coupon equity with n years to maturity, the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return in the top, middle and bottom graphs respectively. In each graph, the solid line is for the LW case and the dot-dashed line is for the base case. The quarterly return on zero-coupon equity with n years to maturity is calculated as the return from holding zero-coupon equity with n years to maturity at the start of the quarter. The annual return on zero-coupon equity with n years to maturity is then obtained by rolling over these quarterly returns for 4 quarters.

It is worth noting that the excess return on the market equity portfolio is a weighted average of the weights on the zero-coupon equity claims, where all the weights are positive. Further, the firms in the extreme value decile receive fractions of the market dividend that are relatively larger in the near future than in the far future. The converse is true for the firms in the extreme growth decile. The top graph of Figure 1 shows that in the LW case, the expected excess return on the zero-coupon equity claim is declining in the claim's maturity which explains why this case delivers a value premium in excess return. For the base case, it is hump-shaped as a function of maturity, but the hump occurs at a sufficiently short maturity to still deliver a value premium in excess return, as discussed in section 2.1. The middle graph shows that excess return volatility is hump-shaped in both cases, though the hump occurs much earlier for the base case. The bottom graph shows that the Sharpe ratio declines monotonically for both cases, though the relation is strongly convex for the LW case and concave at most maturities for the base case.

Using the expressions for the excess return on zero-coupon equity and its first two moments in equations (11)-(13), the shapes of $A(n)$, $B_z(n)$ and $B_x(n)$ as functions of n can be used to better understand the relations plotted in the figure 1, especially the relation between the expected excess return on zero-coupon equity and its maturity plotted in the top graph. Figure 2 plots, as a function of maturity for zero-coupon equity with n years to maturity, $A(n)$, $B_z(n)(1 - \phi_z)$ and $B_x(n)$ in the top left, top right, and bottom left graphs respectively. In each graph, the solid

line is for the LW case and the dot-dashed line is for the base case. $A(n)$, $B_z(n)$ and $B_x(n)$ are, respectively, the constant coefficient, the coefficient on z^m and the coefficient on x in equation (7) for the zero-coupon aggregate equity price-dividend ratio paying out in n periods. $B_z(n)$ is multiplied by $(1 - \phi_z)$. Note that $B_z(n)$ is the same for both these cases, and in all cases considered except those in which $\sigma_z = 0$. For those cases in which $\sigma_z = 0$, the zero-coupon aggregate market price-dividend ratio paying out in n periods does not depend on z^m and so $B_z(n)$ is not defined.

As was pointed out in section 2.1, $B_z(n)$ is always positive and increasing in n . Moreover, $B_z(n)$ is the same for the LW and base cases. Equation (13) shows that $E_t[R_{n,t+1}^m/R_t^f]$ depends on the right hand side of (13) and since $E_t[R_{n,t+1}^m/R_t^f]$ is likely to be highly correlated with $E_t[R_{n,t+1}^m - R_t^f]$, this right hand side can be used to understand how the unconditional expected excess return on a zero-coupon equity claim ($E[R_{n,t+1}^m - R_t^f]$) varies with maturity. This analysis is performed in section 2.1 by exploiting the fact that the unconditional mean of x is positive. Since $\sigma_{x,d}$ is zero in the LW case, the shape of $E_t[R_{n,t+1}^m/R_t^f]$ as a function of its maturity n depends on the shape of $B_z(n)$ and the sign of $\sigma_{z,d}$. Since $\sigma_{z,d}$ is negative and $B_z(n)$ positive and increasing, $E_t[R_{n,t+1}^m/R_t^f]$ is decreasing in n as reported in Figure 1. Hence the LW case delivers a value premium in excess return as reported in Table 5.

Now $\sigma_{x,d}$ is negative in the base case so the shape of $B_x(n)$ matters for the shape of $E_t[R_{n,t+1}^m/R_t^f]$. Figure 2 shows that $B_x(n)$ is negative and has an inverted hump shape, consistent with the observation of section 2. Consequently, $B_x(n-1)\sigma_{x,d}$ in equation (13) is hump-shaped and the implication is that $E_t[R_{n,t+1}^m/R_t^f]$ can be hump-shaped, as reported in Figure 1. Hence, the base case is able to deliver a value premium in excess return as reported in Table 5.

4.2 Wedge Cases

This subsection discusses the results for the two wedge cases and compares them to the results from the data and for the LW and base cases. In contrast to the LW and base cases, the aggregate consumption process is allowed to differ from the market dividend process in the two wedge cases. This process is the same for both wedge cases and is calibrated to the data, while the market dividend process in both wedge cases is the same as that used in the base case. As in the base case, the correlation between the change in log consumption and the price-of-risk variable x is set to -0.99 in the 2 wedge cases, consistent with habit preferences. The persistence of the x process in the two wedge cases is the same low value used in the base case. As discussed in section 3, the

x parameters in wedge case 1 are chosen to match the autocorrelation of the equity price-dividend ratio and its unconditional Sharpe ratio, as in the base case. For wedge case 2, the x parameters are chosen to match the unconditional volatility of the market equity excess return as well.

4.2.1 Aggregate Moments

The last two columns of Table 3 contain moments for aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth for the two wedge cases. Both wedge cases are calibrated to match the unconditional equity Sharpe ratio and the volatility of its price dividend and Table 3 shows that they are both able to match both moments. Wedge case 1 does not attempt to match the volatility of the market equity excess return and produces a much lower value than in the data, just as the base case does. On the other hand, wedge case 2 does attempt to match this volatility and the last column of Table 3 shows that it does so successfully. Because wedge case 2 is calibrated to match the equity Sharpe ratio and the volatility of the equity excess return, it also matches the data expected excess market return. Since wedge case 2 successfully matches the volatility of equity excess return, it is somewhat surprising that it suffers the same fate as wedge case 1 and the base case of substantially understating the volatility of the equity price-dividend ratio. Again, because parameter values for the 2 wedge cases were not chosen specifically to match the autocorrelation of the price-dividend ratio in the data, it is impressive that this autocorrelation is above 0.89 for wedge case 1 and still above 0.78 for wedge case 2. Finally, as with the base case, excess market returns exhibit negative autocorrelation for both wedge cases, especially the second, which is counterfactual.

4.2.2 Predictive Regressions

Table 4 reports results for the predictive regressions and the last two columns report results for the 2 wedge cases. The most glaring weakness of the base case was its inability to generate the excess market return predictability observed in the data. Unfortunately, the two wedge cases do not perform much better than the base case along this dimension. While the 2 wedge cases are also able to produce the data's negative predictability coefficients, their magnitudes are again much smaller than those observed in the data, and the predictive regression R^2 s at horizons of 1 and 10 years are still negligible. The 2 wedge cases also do a poor job of reproducing the predictability of 1 or 10 year log dividend growth found in the data using either the log price-dividend ratio or a

proxy for z^m , especially at the 10 year horizon for log dividend growth.

4.2.3 Value vs Growth Portfolios

Table 5 reports results for the extreme growth and value deciles as well as HML and the last two columns report results for the 2 wedge cases. The last two columns of Table 5 show that the 2 wedge cases can also generate a value premium in both excess return and CAPM alpha. Moreover, the premia in excess return generated by the two cases, while still less than in the data, is much closer than the premium generated by the base case. The same is true for the premium in CAPM-alpha for wedge case 1 while the premium in CAPM-alpha for wedge case 2 is similar in magnitude to that obtained from the base case. In wedge cases 1 and 2, the excess return spreads between the value and the growth portfolio are 4.18% per annum and 3.78% per annum respectively, which are much closer to the data value of 4.88% per annum when sorting on B/M than the 1.91% per annum delivered by the base case. For wedge case 1, the CAPM-alpha spread for these two extreme book-to-market deciles is 2.90% per annum which is also closer to the 5.63% per annum in the data than the 1.20% per annum in the base case. However, wedge case 2 delivers a CAPM-alpha spread of 1.21% per annum which is comparable to the value generated by the base case. Figure 1 plots, as a function of maturity, important statistics for the annual return on zero-coupon equity with n years to maturity, with the dotted lines representing wedge case 1 and the dashed lines representing wedge case 2. The top graph in Figure 1 shows that the unconditional expected annual return on the zero-coupon equity is hump-shaped in maturity with the both humps occurring at maturities less than 10 years. Recall that the firms in the extreme value decile receive fractions of the market dividend that are relatively larger in the near future than in the far future, while the converse is true for the firms in the extreme growth decile. Consequently, these hump-shapes in the top graph of Figure 1 are consistent with the two wedge cases delivering value premia in excess returns, just as the base case delivering the same hump-shape is consistent with it also delivering a value premium in excess return. Moreover, $B_x(n)$ plotted in Figure 2 as a function of n is u-shaped for the two wedge cases just as it is for the base case.

The two wedge cases generate results that are similar to the results for the base case in a number of other respects. First, both wedge cases deliver a CAPM-alpha spread between the extreme value and the growth deciles that is smaller than the excess return spread, while the converse is true for the data and LW cases. Again, the reason is that the CAPM beta for the extreme value decile is higher than for the extreme growth decile in both wedge cases, while the converse is true for the

data and LW, as the rows labeled β_i in Table 5 show. Second, both wedge cases deliver excess return volatility that is higher for the extreme value decile than for the extreme growth decile, which is consistent with the data when B/M is used to construct the extreme deciles, but inconsistent when E/P or C/P is used. Third, the volatility numbers in the two wedge cases for the two extreme deciles and especially for HML, though not as low as in the base case, are much lower than in the data, except for the value decile in wedge case 2, whose excess return volatility is close to that in the data.

For the two wedge cases, the R^2 s of the CAPM market model regressions are similar to the data R^2 values for the growth decile, but, as in the base case, are larger than the data R^2 values for the value decile and HML. As for all the data sorts and the base case, the extreme growth decile has the lowest Sharpe ratio of the two extreme portfolios and HML for both wedge cases. The extreme value decile has the highest Sharpe ratio for all the data sorts and the base case, while HML has the highest Sharpe ratio for both of the wedge cases.

4.3 Comparative Static Analyses

The section examines the effects of varying the persistence of the price-of-risk variable ϕ_x , of setting σ_x equal to zero, and of setting σ_z equal to zero to make the conditional mean of the aggregate equity dividend growth equal to a constant. The comparative static analyses are performed by perturbing ϕ_x , σ_x and σ_z in the LW and base cases, which means the aggregate consumption and dividend processes are set equal to each other for all perturbations. In the base perturbations, $\rho_{x,d}$ is set equal to -0.99, consistent with external habit preferences while in the LW perturbations, it's set equal to 0. Thus, by perturbing ϕ_x , σ_x and σ_z relative to their values in these two cases, we can assess how the chosen value for $\rho_{x,d}$ affects the comparative static analyses for ϕ_x , σ_x and σ_z .

4.3.1 Persistence of the Price-of-risk Variable ϕ_x and Setting x Equal to a Constant

Table 6 shows the effects of changing the persistence of the x process, and of setting σ_x equal to zero, holding all other parameters fixed at either the base case (the first 4 columns) or the LW (the last 4 columns) parameter values. For each set of 4 cases, ϕ_x is monotonically declining going from the case in the first column of the set to the third column, with ϕ_x set equal to zero in the third column of each set. So the base case is the second column of the first set while the LW case is the first column of the second set. Across the first three columns of each set, the unconditional volatility of

the quarterly x process is held fixed by varying σ_x appropriately as ϕ_x is varied. The fourth column of each set fixes $\sigma_x = 0$, which means that $x_t = \bar{x}$. The panel labeled Parameters lists the parameter values used in the 8 cases. The panel labeled Simulated Moments reports simulated moments for the aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth as in Table 3. The panel labeled Predictive Regressions has 3 subpanels and reports the same results for the same regressions as Table 4. The panel labeled Value vs Growth Portfolios has 2 subpanels which correspond to the 2 panels in Table 5.

Figure 3 shows the effects on the returns of zero-coupon equity of varying ϕ_x for 2 pairs of cases. The 2 cases in the first pair are plotted in the 3 left-hand side graphs and both cases have all their other parameters equal to the base case. The 2 cases in the second pair are plotted in the 3 right-hand side graphs and both cases have all their other parameters equal to those for the LW case. In each graph for a pair, the solid line is for the high- ϕ_x case and the dot-dashed line is for the low- ϕ_x case. For the left-hand side graphs, the solid line corresponds to the base+ $\Delta\phi_x$ case in Table 6 and the dot-dashed line to the base case, while for the right-hand side graphs, the solid line corresponds to the LW case and the dot-dashed line to the LW- $\Delta\phi_x$ case in the same table. For zero-coupon equity, the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return are all plotted as a function of maturity in the top two, middle two, and bottom two graphs respectively.

Figure 4 shows the effects on $A(n)$ and $B_x(n)$ for zero-coupon equity of varying ϕ_x for 2 pairs of cases. The 2 cases in the first pair are plotted in the 2 left-hand side graphs and both cases have all their other parameters equal to those for the base case. The 2 cases in the second pair are plotted in the 2 right-hand side graphs and both cases have all their other parameters equal to those for the LW case. In each graph for a pair, the solid line is for the high- ϕ_x case and the dot-dashed line is for the low- ϕ_x case. For the left-hand side graphs, the solid line corresponds to the base+ $\Delta\phi_x$ case in Table 6 and the dot-dashed line to the base case, while for the right-hand side graphs, the solid line corresponds to the LW case and the dot-dashed line to the LW- $\Delta\phi_x$ case in the same table. The figure plots $A(n)$ and $B_x(n)$ in the top 2 and bottom 2 graphs respectively. $A(n)$ and $B_x(n)$ are, respectively, the constant coefficient and the coefficient on x in equation (7) of section 2 for the zero-coupon aggregate equity price-dividend ratio paying out in n periods.

The Simulated Moments panel of Table 6 shows that the effects of varying ϕ_x and of setting $\sigma_x = 0$ on the expected equity price-dividend ratio and the expected equity excess return depend on the value of $\rho_{x,d}$. As ϕ_x declines to zero holding the unconditional volatility of x fixed and then σ_x is

set to 0, the expected equity price-dividend ratio increases monotonically and the expected equity excess return decreases monotonically for the base perturbations, while the converse is true for the LW perturbations. Further, the equity Sharpe ratio varies with ϕ_x and σ_x in the same manner as the expected equity excess return.

The Predictive Regressions panel of Table 6 shows that for the predictive return regressions, the coefficients, the β_1 s, are always zero or negative and their magnitudes and the regression R^2 s are monotonically increasing in ϕ_x irrespective of the return horizon or the case. This finding suggests that it may be difficult to generate return predictability without at least one price-of-risk variable that is persistent. For the regressions of future long-horizon log dividend growth on the price-dividend ratio, the regressions R^2 s are monotonically decreasing in ϕ_x irrespective of the return horizon or the case. The same is true for the coefficients, the β_1 s, in the LW case, but they exhibit a hump-shaped pattern as a function of ϕ_x in the base case. For both these regressions and both cases, the regression results for $\sigma_x = 0$ are virtually the same as for $\phi_x = 0$, which suggests that when $\phi_x = 0$, varying σ_x does not affect the ability of the price dividend ratio to predict long horizon equity returns or log dividend growth rates. Since the z^m and Δd^m processes are the same across all cases and perturbations, the results for the regressions of future long-horizon log dividend growth on z^m are the same for all the columns.

The Value vs Growth Portfolios panel of Table 6 shows that the effect of varying ϕ_x on the expected excess annual return differential between the two extreme deciles depends on the value of $\rho_{x,d}$, with the effect being very different between the base and LW cases. In the LW case, $\rho_{x,d}$ is equal to zero so equation (13) says that quarterly $E_t \left[\frac{R_{n,t+1}^m}{R_t^f} \right]$, which is for the zero-coupon claim on the market dividend that matures in n quarters, does not depend on $B_x(n-1)$. Consequently, for each n , quarterly $E_t \left[\frac{R_{n,t+1}^m}{R_t^f} \right]$ is the same for all the LW perturbations, because $B_z(n-1)$, $\delta^m \sigma_d^2$, $\sigma_{d,u}$, $\sigma_{z,d}$ and σ_d are the same. However, equation (7) shows that the quarterly price-dividend ratio for a zero-coupon claim on the market dividend still depends on the current x value, so the composition of the extreme deciles in terms of the zero-coupon market equity claims is still likely to vary across the LW perturbations as the ϕ_x and σ_x are varied. The implication is that the extreme decile quarterly excess returns and the extreme decile return differential would likely be similar but not identical for all the LW perturbations. Further, the Value vs Growth Portfolios panel of Table 6 reports the expected excess return differential between the two extreme deciles for annual rather than quarterly returns. The extreme decile excess returns and differentials reported in Value vs Growth Portfolios panel of Table 6 are similar across the LW perturbations as would be expected. The expected excess annual return differential between the two extreme deciles ranges

across the LW perturbations from 5.49% per annum for the LW case up to 6.79% per annum for the LW perturbation with $\sigma_x = 0$. Consistent with this observation, Figure 3 shows that the unconditional expected annual return on zero-coupon equity is similar for the LW case and the low- ϕ_x LW perturbation.

Turning to the base perturbations, the Value vs Growth Portfolios panel of Table 6 shows that the high- ϕ_x base perturbation delivers a growth premium in expected excess annual return of 12.77%. As per LW, Figure 4 shows that when ϕ_x is large, $B_x(n)$ is negative and declining in maturity, which together with the base $\rho_{x,d}$ value of -0.99 causes the right-hand side of equation (13) to increase with maturity, as demonstrated by the top left graph of Figure 3. The inability to obtain a value premium in expected excess return when the price of risk variable x is highly persistent and its shock is highly negatively correlated with the aggregate dividend growth shock was first documented in LW. Going from the base case to the $\phi_x = 0$ base perturbation causes the value premium to decline from 1.91% per annum to 1.14% per annum while setting $\sigma_x = 0$ causes a further decline to 0.68% per annum.

Turning to the CAPM-alpha differential between the two extreme deciles, there is a value premium for all the base and LW perturbations. This premium decreases as ϕ_x declines for the base perturbations, being at its lowest for the base perturbation with $\sigma_x = 0$. For the LW perturbations, the LW case, which has the highest ϕ_x value, also has the the highest value premium.

4.3.2 Constant versus Persistent Conditional Mean Aggregate Consumption Growth

Table 7 shows the effect of making the conditional mean of the aggregate consumption/equity dividend growth equal to a constant by setting $\sigma_z = 0$, while keeping the unconditional quarterly volatility of Δd^m fixed, for the base and LW cases, which set aggregate consumption growth and equity dividend growth equal. All other parameters are held fixed at either the base case (the first 2 columns) or the LW case (the last 2 columns) parameter values. For each pair of cases, σ_z is set equal to zero in the second column of the pair. So the base case is the first column of the first set while the LW case is the first column of the second set. Across the two columns of each pair, the unconditional volatility of the quarterly Δd^m process is held fixed by varying σ_d appropriately as σ_z is varied. The panel labeled Parameters lists the parameter values used in the 4 cases. The panel labeled Simulated Moments reports simulated moments for the aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth as in Table

3. The panel labeled Predictive Regressions has 3 subpanels and reports the same results for the same regressions as Table 4. The panel labeled Value vs Growth Portfolios has 2 subpanels which correspond to the 2 panels in Table 5. Note that in these calibrations with $\sigma_z = 0$, β_1 and R^2 in the third predictive regression are not identified.

Figure 5 shows the effects on the returns of zero-coupon equity of long-run risk for 2 pairs of cases. The 2 cases in the first pair are plotted in the 3 left-hand side graphs and both cases have all their other parameters equal to the base case. The 2 cases in the second pair are plotted in the 3 right-hand side graphs and both cases have all their other parameters equal to those for the LW case. In each graph for a pair, the solid line is for the case where long-run risk is incorporated into the consumption growth process, and the dashed line is for the case where long-run risk is turned off. For the left-hand side graphs, the solid line corresponds to the base case in Table 7 and the dashed line to the base ($\sigma_z = 0$) case, while for the right-hand side graphs, the solid line corresponds to the LW case and the dashed line to the LW ($\sigma_z = 0$) case in the same table. For zero-coupon equity, the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return are all plotted as a function of maturity in the top two, middle two, and bottom two graphs respectively. The quarterly return on zero-coupon equity with n years to maturity is calculated as the return from holding zero-coupon equity with n years to maturity at the start of the quarter. The annual return on zero-coupon equity with n years to maturity is then obtained by rolling over these quarterly returns for 4 quarters. The models are simulated at a quarterly frequency for 4 million quarters, or until convergence, i.e. the first and second halves of the simulation yield sufficiently similar moments, whichever comes later.

The Simulated Moments panel of Table 7 shows that the effects of setting $\sigma_z = 0$ on the expected equity price-dividend ratio, expected equity excess return and equity Sharpe ratio don't depend on the value of $\rho_{x,d}$. When σ_z is set to 0, the expected equity price-dividend ratio declines and both the expected equity excess return and the equity Sharpe ratio increase for both the base and LW perturbations. The increases in expected excess return are consistent with the $B_z(n-1)\sigma_{z,d}$ term in equation (13) being negative in both cases when $\sigma_z > 0$. In contrast, setting $\sigma_z = 0$ causes the autocorrelation of the price dividend ratio to decline dramatically from 0.896 to less than 0.1 in the base perturbation, but to decline by less than 0.05 to 0.852 in the LW perturbation. This finding is to be expected since the price dividend ratio depends on z and x , and x is much more persistent in the LW case than the base case, which causes the price dividend ratio to remain much more persistent in the LW case than the base case when σ_z is set equal to 0. This finding suggests that it may be difficult to generate price-dividend ratio autocorrelation without at least one price-of-risk

variable that is persistent.

The Predictive Regressions panel of Table 7 shows that for the predictive return regressions, the coefficients, the β_1 s, are always negative and their magnitudes and the regressions R^2 s always increase when σ_z is set to 0, irrespective of the return horizon or the case. These increases are likely driven by the increases in the unconditional Sharpe ratio resulting from setting $\sigma_z = 0^4$ allowing conditional expected excess return variation to make a larger contribution to the total variance of equity return. When future log dividend growth is regressed on the price-dividend ratio, the coefficients and the regression R^2 s are zero when $\sigma_z = 0$ because future log dividend growth is i.i.d. and hence unpredictable using the price-dividend ratio. When $\sigma_z = 0$, z^m is a constant, the regressions of future log dividend growth on z^m is not defined, and so the regression coefficient and R^2 are not reported for this regression.

The Value vs Growth Portfolios panel of Table 7 shows that setting $\sigma_z = 0$ decreases the expected excess annual return differential between the two extreme deciles for the base and LW cases. The implication is that the persistence in the mean of log consumption/dividend growth assumed in both these cases is playing an important role in their ability to both generate a value premium in excess return. In the LW case, $\rho_{x,d}$ is equal to 0 so equation (13) says that quarterly $E_t \left[\frac{R_{n,t+1}^m}{R_t^f} \right]$ is constant for all n when $\sigma_z = 0$ and the $B_z(n-1)\sigma_{z,d}$ term is zero. Consequently, quarterly $E_t \left[\frac{R_{n,t+1}^m}{R_t^f} \right]$ is the same for all n for the LW perturbation with $\sigma_z = 0$. The implication is that the two extreme decile quarterly excess returns are the same and the quarterly excess return differential is zero. However, the Value vs Growth Portfolios panel of Table 7 reports expected excess return differential between the two extreme deciles for annual rather than quarterly returns, which explains why the differential can be 0.73% per annum and not zero. Consistent with this value being close to zero, the top right graph of Figure 5 shows that the unconditional expected annual return on zero-coupon equity as a function of maturity is quite flat for the LW perturbation with $\sigma_z = 0$.

Turning to the CAPM-alpha value premium between the two extreme deciles, setting $\sigma_z = 0$ drives the premium down to 0.13% per annum from 1.20% for the base case but, increases it from 6.41% per annum to 12.78% per annum for the LW case. Thus, the persistence in the mean of log consumption/dividend growth assumed in these cases is important for the CAPM-alpha value premium obtained in the base case but not in the LW case.

⁴See the first panel of Table 7.

5 Conclusion

This paper finds the external-habit model of CC (1999) can generate a value premium in both CAPM alpha and expected excess return when the log surplus-consumption ratio is allowed to be not very persistent. In contrast, LW find that when the log surplus-consumption ratio is assumed to be highly persistent as in CC (by assuming that the price-of-risk state variable is highly persistent), the external-habit model generates a growth premium in expected excess return. However, there are good economic reasons for why the persistence of the log surplus-consumption ratio is likely to be low. In particular, the high persistence assumed by LW's specification implies that the contribution of the most recent 5 years of log consumption to log habit is just a little over 50% and so the contribution of log consumption more than 5 years ago is almost 50%, which seems very high. We choose a value for this persistence which is sufficiently low that the most recent 2 years of log consumption contribute over 98% of all past consumption to log habit, which is a much more reasonable number than the 25% contribution generated by the LW value. The micro evidence also favors a less persistent log surplus-consumption ratio: Brunnermeier and Nagel (2006) examine how risky asset holdings change in response to wealth shocks and reject a persistent habit specification and Ravina (2007) shows that credit card purchases are more in line with a fast-moving habit than a slow-moving habit.

In our specification, expected consumption is slowly mean-reverting, as in the long-run risk model of Bansal and Yaron (2004), which is why our model is able to generate a price-dividend ratio for aggregate equity that exhibits the high autocorrelation found in the data, despite the very low persistence of the price-of-risk state variable. Our results suggest that external-habit preferences and long-run risk consumption may both play important roles in explaining the time-series and cross-sectional properties of equity returns and prices.

When aggregate consumption and market dividend are assumed to be the same and are both calibrated to market dividends, the model is able to match the mean equity price dividend ratio and equity Sharpe ratio, but is unable to generate a sufficiently volatile equity return. Driving a wedge between aggregate consumption and market dividend by calibrating the former to aggregate consumption data allows the model to also match the volatility of the equity return.

One important dimension of equity return behavior that low persistence has difficulty replicating is the predictability of long-horizon equity return using the price-dividend ratio. However, allowing the consumption and aggregate dividend to be different processes leads to equity return predictability

that is qualitatively similar to that in the data but much smaller in magnitude. The low persistence also generates negative autocorrelation in aggregate equity return, especially when matching the equity return volatility to data, which is counterfactual. The implication is that a second price-of-risk state variable that is highly persistent may be needed to generate the autocorrelation and predictability of aggregate equity return equity observed in the data.

References

- [1] Abel, Andrew B., 1999. Risk Premia and Term Premia in General Equilibrium. *Journal of Monetary Economics* 43, 3-33.
- [2] Bansal, Ravi, and Amir Yaron, 2004. Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* 59, 1481-1509.
- [3] Bansal, Ravi, Amir Yaron, Dana Kiku, 2007. Risks For the Long Run: Estimation and Inference. Working Paper, Duke University and University of Pennsylvania.
- [4] Bekaert, Geert, and Eric Engstrom, 2009. Asset Return Dynamics under Bad Environment-Good Environment Fundamentals. Working Paper, Columbia University and the Federal Reserve.
- [5] Brunnermeier, Markus, and Stefan Nagel, 2006. Do Wealth Fluctuations Generate Time-varying Risk Aversion? Micro-Evidence on Individuals; Asset Allocation. Working Paper, Princeton University and Stanford University.
- [6] Campbell, John Y., and John H. Cochrane, 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107, 205-251.
- [7] Constantinides, George M., 1990. Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy* 98, 519-43.
- [8] Dynan, Karen E., 2000. Habit Formation in Consumer Preferences: Evidence from Panel Data. *American Economic Review* 90, 391-406.
- [9] Lettau, Martin, and Jessica Wachter, 2007. Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium. *Journal of Finance* 62, 55-92.
- [10] Lynch, Anthony W., and Jessica Wachter, 2008. Using Samples of Unequal Length in Generalized Method of Moments Estimation. Working Paper, New York University and University of Pennsylvania.
- [11] Ravina, Enrichetta, 2007. Habit Persistence and Keeping Up with the Jones: Evidence from Micro Data. Working Paper, New York University.
- [12] Santos, Tano and Pietro Veronesi, 2008. Habit Formation, the Cross Section of Stock Returns and the Cash-Flow Risk Puzzle. Working paper, Columbia University and University of Chicago.

- [13] Stambaugh, Robert, 1997. Analyzing Investments Whose Histories Differ in Length. *Journal of Financial Economics* 45 285-331.
- [14] Sundaresan, Suresh, 1989. Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth. *Review of Financial Studies* 2, 73-89.

Table 1: **Model Parameters.** This table lists the parameter values used by LW (first column) and for our base and wedge cases (final three columns). All parameters are as defined in section 2. The model is quarterly, but the log riskfree rate r^f is converted into an annual number by multiplying by a factor of 4 and we express the persistence parameters ϕ_x and ϕ_z at annual frequencies by raising each of them to the power of 4.

Variable	Frequency	LW	Base	Wedge	
				1	2
g^m	annual	2.28%	2.28%	2.28%	2.28%
$r^f = -a$	annual	1.93%	1.93%	1.93%	1.93%
b		0	0	0	0
\bar{x}	quarterly	0.625	0.25	0.365	0.27
ϕ_z	annual	0.91	0.91	0.91	0.91
ϕ_x	annual	0.865	0.14	0.14	0.14
σ_m	quarterly	0.0724	0.0724	0.0724	0.0724
σ_d	quarterly	0.0724	0.0724	0.0160	0.0160
σ_z	quarterly	0.00165	0.00165	0.00165	0.00165
σ_x	quarterly	0.1225	0.16	0.3	0.335
σ_u	quarterly	0	0	0.0435	0.0435
$\rho_{m,z} = \rho_{d,z}$	quarterly	-0.82	-0.82	-0.82	-0.82
$\rho_{d,x}$	quarterly	0	-0.99	-0.99	-0.99
$\rho_{z,x}$	quarterly	0	0.81	0.81	0.81
$\rho_{d,u}$	quarterly	-	-	-0.30	-0.30
$\rho_{z,u}$	quarterly	-	-	0.0037	0.0037
$\rho_{x,u}$	quarterly	-	-	0.30	0.30
δ^m	quarterly	1	1	4.54	4.54

Table 2: **Contribution of Lagged Consumption to Habit.** This table shows the percentage contribution of lagged log consumption to log habit in the external habit model of CC for parameters implied by the LW case (see the first column) and our cases (see the second column). Section 2 shows how to back out the implied CC parameters from the models and presents the approximate relation between log habit and lagged log consumption used to calculate the contributions: $h_{t+1} \approx \overline{h-d} + (1-\phi_s) \sum_{j=0}^{\infty} (\phi_s)^j d_{t-j} + \frac{g}{1-\phi_s}$. This table decomposes habit into the proportion from consumption within the last 5 years, and the proportion from more than 5 years before, for LW and our calibrations.

Consumption Lag (yrs)	Habit Contribution (%)	
	LW	Ours
1	13.50	86.00
2	11.68	12.04
3	10.10	1.69
4	8.74	0.24
5	7.56	0.03
1 to 5	51.57	99.99
>5	48.43	0.01

Table 3: **Aggregate Moments.** This table reports moments for aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth. The first column reports moments for the data, the second reports simulated moments for the LW case, and the final three report simulated moments for the base and wedge cases. The data are the same as that in LW and are annual from 1890-2002. The models are simulated at a quarterly frequency for 4 million quarters, by which point the reported moments have converged, i.e. the first and second halves of the simulation yield sufficiently similar moments. Returns, dividends, and price-dividend ratios are aggregated to annual frequencies: so $P^m/D^m = P_t^m / \sum_{\tau=0}^3 D_{t-\tau}^m$ and $p^m - d^m \equiv \log(P^m/D^m)$. $Sharpe^m$ is the unconditional Sharpe ratio of the aggregate equity, and AC is the autocorrelation.

Moment	Data	LW	Base	Wedge	
				1	2
$E[P^m/D^m]$	25.55	20.06	30.95	31.74	21.42
$\sigma[p^m - d^m]$	0.38	0.381	0.260	0.283	0.259
$AC[p^m - d^m]$	0.87	0.882	0.896	0.899	0.789
$E[R^m - R^f]$	6.33	8.120	4.519	4.501	6.983
$\sigma[R^m - R^f]$	19.41	19.41	10.69	11.21	17.58
$AC[R^m - R^f]$	0.03	-0.04	-0.13	-0.16	-0.31
$Sharpe^m$	0.33	0.418	0.423	0.401	0.397
$AC[\Delta d^m]$	-0.09	-0.03	-0.03	-0.03	-0.03
$\sigma[\Delta d^m]$	14.48	14.39	14.39	14.45	14.43

Table 4: **Predictive Regressions.** The top panel is the regression of the future log excess return on the aggregate equity on the log aggregate equity price-dividend ratio today. The middle panel is the regression of future changes in log aggregate equity dividend on the log aggregate equity price-dividend ratio today. The bottom panel is the regression of future changes in log aggregate equity dividend on the consumption-aggregate equity dividend ratio for the data and z^m today for the models. The first column reports results for the data, the second reports results for the LW case, and the final three report results for the base and wedge cases. The data are the same as that in LW and are annual from 1890-2002. The models are simulated at a quarterly frequency for 4 million quarters, by which point the reported moments have converged, i.e. the first and second halves of the simulation yield sufficiently similar moments. Log returns, log dividend growth, log price-dividend ratios and log consumption-aggregate dividend ratios are all aggregated to annual frequencies: so $P_t^m/D_t^m = P_t^m / \sum_{\tau=0}^3 D_{t-\tau}^m$ and $p_t^m - d_t^m \equiv \log(P_t^m/D_t^m)$. Results are reported for horizons, H , of 1 and 10 years. R^2 is the regression R^2 .

	Horizon (yrs)	Data	LW	Base	Wedge	
					1	2
$\sum_{i=1}^H (r_{t+i}^m - r_{t+i-1}^f) = \beta_0 + \beta_1(p_t^m - d_t^m) + \epsilon_t$						
β_1	1	-0.12	-0.12	-0.01	-0.02	-0.07
	10	-1.09	-0.68	-0.04	-0.11	-0.17
R^2	1	0.05	0.07	0.00	0.00	0.01
	10	0.31	0.30	0.00	0.01	0.02
$\sum_{i=1}^H \Delta d_{t+i}^m = \beta_0 + \beta_1(p_t^m - d_t^m) + \epsilon_t$						
β_1	1	0.02	0.05	0.11	0.10	0.10
	10	-0.31	0.33	0.73	0.67	0.66
R^2	1	-0.01	0.02	0.04	0.04	0.03
	10	0.05	0.09	0.22	0.21	0.18
$\sum_{i=1}^H \Delta d_{t+i}^m = \beta_0 + \beta_1 z_t^m + \epsilon_t$						
β_1	1	0.10	3.86	3.86	3.85	3.86
	10	0.68	26.33	26.33	25.98	26.33
R^2	1	0.03	0.04	0.04	0.04	0.04
	10	0.25	0.25	0.25	0.24	0.25

Table 5: **Value vs Growth Portfolios.** The table reports results for the extreme growth decile (portfolio 1), the extreme value decile (portfolio 10), and the portfolio which is long portfolio 10 and short portfolio 1 (the HML portfolio). The top panel reports expected excess annual return, the volatility of annual excess return and the unconditional Sharpe ratio for annual return. The bottom panel reports CAPM α , CAPM β and regression R^2 using annual returns. The first three columns report results for the data, the fourth reports results for the LW case, and the final three report results for the base and wedge cases. Data are monthly from Ken French's website and span 1952 to 2002: means are annualized by multiplying by 12 and volatilities by $12^{0.5}$. For the models, we sort the 200 firms into deciles at the start of each year from value to growth based on their annual price-dividend ratios, which are given by: $P_t^j/D_t^j = P_t^j/\sum_{\tau=0}^3 D_{t-\tau}^j$ for firm j . The models are simulated at a quarterly frequency for 4 million quarters, by which point the reported moments have converged, i.e. the first and second halves of the simulation yield sufficiently similar moments. $Sharpe^i$ is the unconditional Sharpe ratio for portfolio i 's annual return while R^2 is the regression R^2 .

	Portfolio	E/P	C/P	B/M	LW	Base	Wedge	
							1	2
Expected Excess Return: $E[R^i] - R^f$								
$E[R^i] - R^f$	1	4.71	5.05	5.67	5.30	3.41	1.97	3.83
	10	12.95	11.81	10.55	10.80	5.32	6.16	7.61
	HML	8.25	6.77	4.88	5.49	1.91	4.18	3.78
$\sigma[R^i - R^f]$	1	19.35	18.99	17.77	19.52	9.90	9.58	12.25
	10	18.11	17.24	18.46	17.64	11.47	12.47	18.07
	HML	15.40	14.57	15.15	8.82	4.47	6.26	8.59
$Sharpe^i$	1	0.24	0.27	0.32	0.27	0.34	0.21	0.31
	10	0.72	0.69	0.57	0.61	0.46	0.49	0.42
	HML	0.54	0.46	0.32	0.62	0.43	0.67	0.44
CAPM α : $R_t^i - R_t^f = \alpha_i + \beta_i(R_t^m - R_t^f) + \epsilon_{it}$								
α_i	1	-3.09	-2.70	-1.66	-2.71	-0.66	-1.66	-0.75
	10	6.22	5.34	3.97	3.71	0.53	1.24	0.46
	HML	9.31	8.04	5.63	6.41	1.20	2.90	1.21
β_i	1	1.18	1.17	1.11	0.99	0.90	0.81	0.66
	10	1.02	0.98	1.00	0.87	1.06	1.09	1.02
	HML	-0.16	-0.19	-0.11	-0.11	0.16	0.29	0.37
R_i^2	1	0.83	0.85	0.87	0.96	0.95	0.89	0.89
	10	0.71	0.72	0.65	0.92	0.97	0.97	0.99
	HML	0.02	0.04	0.01	0.06	0.14	0.26	0.57

Table 7: **The Effects of Turning off Long-run Risk.** This table shows the effect of turning off long-run risk in the aggregate equity dividend by setting $\sigma_z = 0$, while keeping the unconditional quarterly volatility of Δd^m fixed. All other parameters are held fixed at either the base case (the first 2 columns) or the LW case (the last 2 columns) parameter values. For each pair of cases, σ_z is set equal to zero in the second column of the pair. So the base case is the first column of the first set while the LW case is the first column of the second set. Across the two columns of each pair, the unconditional volatility of the quarterly Δd^m process is held fixed by varying σ_d appropriately as σ_z is varied. The panel labelled Parameters lists the parameter values used in the 4 cases. The panel labelled Simulated Moments reports simulated moments for the aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth. The panel labelled Predictive Regressions has 3 subpanels. The top subpanel is the regression of the future log excess return on the aggregate equity on the log aggregate equity price-dividend ratio today. The middle subpanel is the regression of future changes in log aggregate equity dividend on the log aggregate equity price-dividend ratio today. The bottom subpanel is the regression of future changes in log aggregate equity dividend on the consumption-aggregate equity dividend ratio for the data and z^m today for the models. The panel labelled Value vs Growth Portfolios has 2 subpanels. The top panel reports expected excess annual return, the volatility of annual excess return and the unconditional Sharpe ratio for annual return. The bottom panel reports CAPM α , CAPM β and regression R^2 using annual returns. The data for all but the value and growth portfolios are the same as that in LW and are annual from 1890-2002. The data for the value and growth portfolios are monthly from Ken French's website and span 1952 to 2002: means are annualized by multiplying by 12 and volatilities by $12^{0.5}$. The models are simulated at a quarterly frequency for 4 million quarters, by which point the reported moments have converged, i.e. the first and second halves of the simulation yield sufficiently similar moments. Note that in the calibrations where $\sigma_z = 0$, β_1 and R^2 in the third predictive regression are not identified.

	Base		LW	
	$\sigma_z > 0$	$\sigma_z = 0$	$\sigma_z > 0$	$\sigma_z = 0$
<u>Parameters</u>				
σ_z	0.00165	0	0.00165	0
$\sigma[d^m]$	0.0728	0.0728	0.0728	0.0728
<u>Simulated Moments</u>				
$E[P^m/D^m]$	30.95	11.59	20.06	12.46
$\sigma[p^m - d^m]$	0.260	0.126	0.381	0.619
$AC[p^m - d^m]$	0.896	0.064	0.882	0.852
$E[R^m - R^f]$	4.519	11.999	8.120	20.416
$\sigma[R^m - R^f]$	10.69	23.39	19.41	44.24
$AC[R^m - R^f]$	-0.13	-0.22	-0.04	-0.03
$Sharpe^m$	0.423	0.513	0.418	0.462
$AC[\Delta d^m]$	-0.03	0.00	-0.03	0.00
$\sigma[\Delta d^m]$	14.39	14.56	14.39	14.56

Table 7: cont

	Horizon (yrs)	Predictive Regressions			
$\sum_{i=1}^H (r_{t+i}^m - r_{t+i}^f) = \beta_0 + \beta_1(p_t^m - d_t^m) + \epsilon_t$					
β_1	1	-0.01	-0.42	-0.12	-0.23
	10	-0.04	-0.48	-0.68	-1.28
R^2	1	0.00	0.07	0.07	0.16
	10	0.00	0.02	0.30	0.59
$\sum_{i=1}^H \Delta d_{t+i}^m = \beta_0 + \beta_1(p_t^m - d_t^m) + \epsilon_t$					
β_1	1	0.11	0.00	0.05	0.00
	10	0.73	0.00	0.33	0.00
R^2	1	0.04	0.00	0.02	0.00
	10	0.22	0.00	0.09	0.00
$\sum_{i=1}^H \Delta d_{t+i}^m = \beta_0 + \beta_1 z_t^m + \epsilon_t$					
β_1	1	3.86	.	3.86	.
	10	26.33	.	26.33	.
R^2	1	0.04	.	0.04	.
	10	0.25	.	0.25	.
	Portfolio	Value vs Growth Portfolios			
Expected Excess Return: $E[R^i] - R^f$					
$E[R^i] - R^f$	1	3.41	12.42	5.30	19.99
	10	5.32	11.29	10.80	20.72
	HML	1.91	-1.13	5.49	0.73
$\sigma[R^i - R^f]$	1	9.90	24.30	19.52	61.09
	10	11.47	21.84	17.64	34.36
	HML	4.47	2.49	8.82	32.99
$Sharpe^i$	1	0.34	0.51	0.27	0.33
	10	0.46	0.52	0.61	0.60
	HML	0.43	-0.45	0.62	0.02
CAPM $\alpha: R_t^i - R_t^f = \alpha_i + \beta_i(R_t^m - R_t^f) + \epsilon_{it}$					
α_i	1	-0.66	-0.05	-2.71	-7.55
	10	0.53	0.08	3.71	5.23
	HML	1.20	0.13	6.41	12.78
β_i	1	0.90	1.04	0.99	1.35
	10	1.06	0.93	0.87	0.76
	HML	0.16	-0.11	-0.11	-0.59
R_i^2	1	0.95	1.00	0.96	0.95
	10	0.97	1.00	0.92	0.95
	HML	0.14	0.98	0.06	0.63

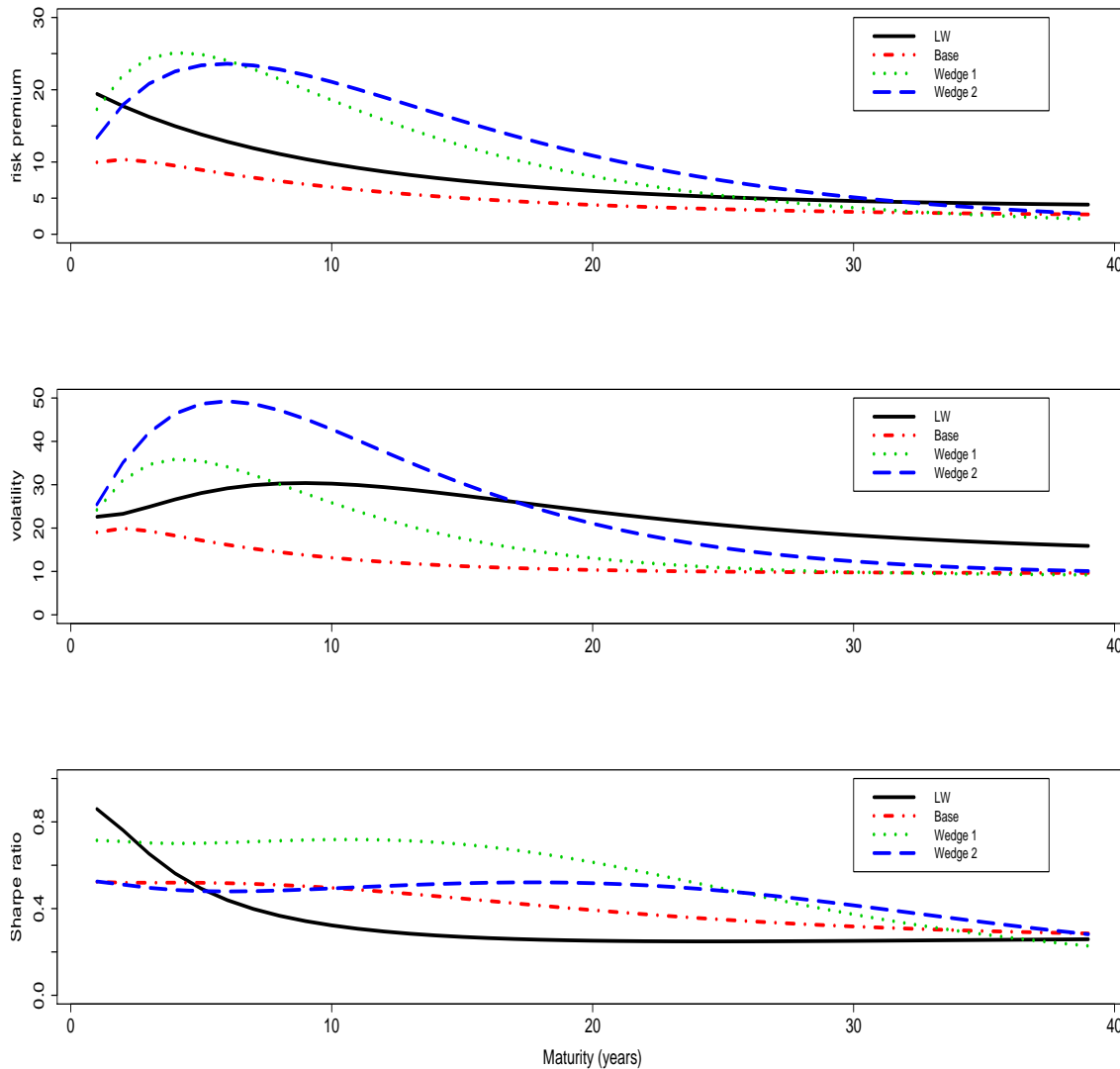


Figure 1: **Returns on Zero-coupon Equity: Base vs LW Case.** The figure plots the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return in the top, middle and bottom graphs respectively. In each graph, the solid line is for the LW case, the dot-dashed line is for the base case, the dotted line is for wedge case 1, and the dashed line is for wedge case 2. The quarterly return on zero-coupon equity with n years to maturity is calculated as the return from holding zero-coupon equity with n years to maturity at the start of the quarter. The annual return on zero-coupon equity with n years to maturity is then obtained by rolling over these quarterly returns for 4 quarters. The models are simulated at a quarterly frequency for 4 million quarters, or until convergence, i.e. the first and second halves of the simulation yield sufficiently similar moments, whichever comes later.

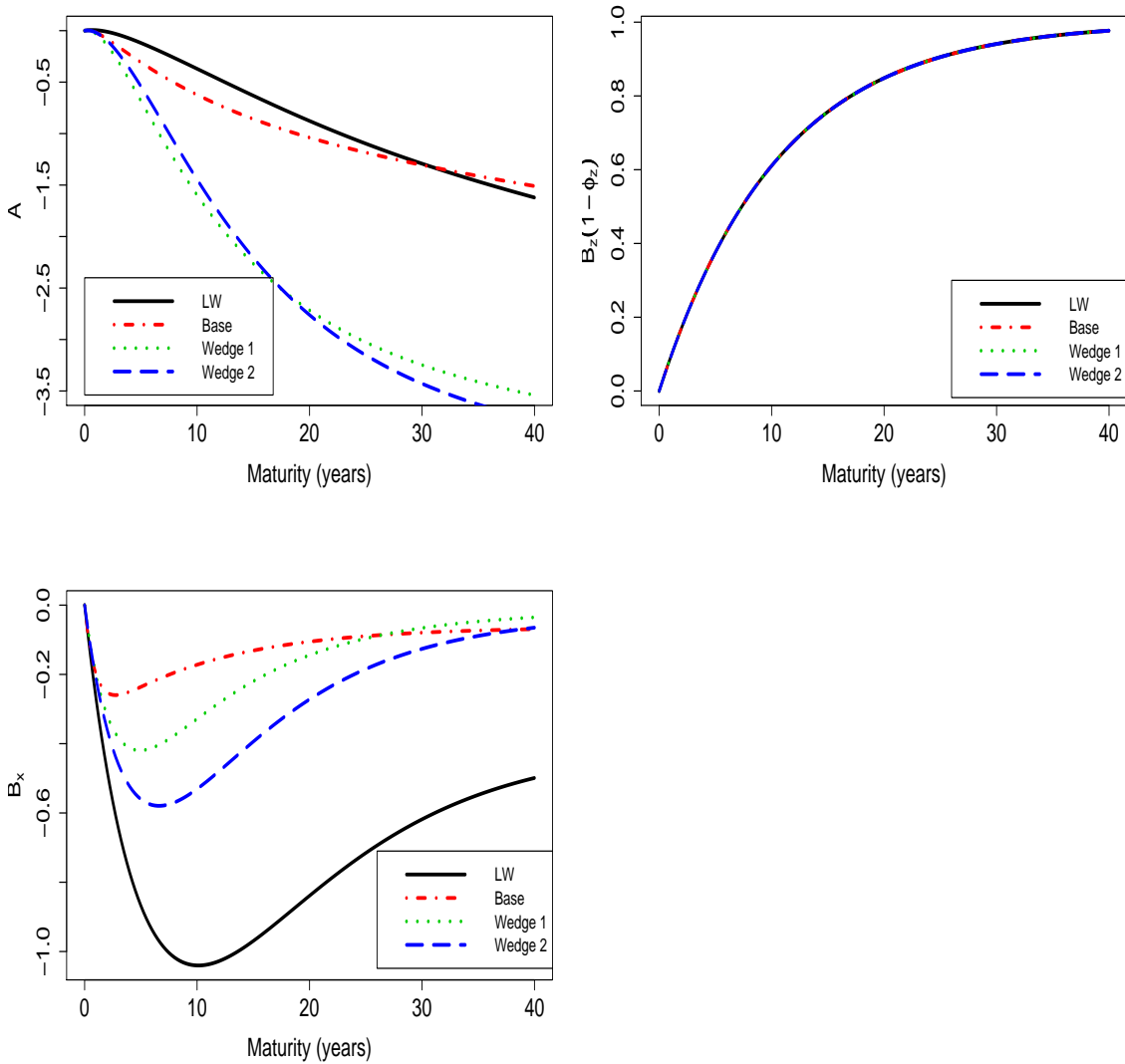


Figure 2: **Zero-coupon Aggregate Equity Log Price-dividend Ratio Coefficients $A(n)$, $B_z(n)$ and $B_x(n)$: Ours vs LW Case.** The figure plots $A(n)$, $B_z(n)(1 - \phi_z)$ and $B_x(n)$ in the top left, top right, and bottom left graphs respectively. In each graph, the solid line is for the LW case, the dot-dashed line is for the base case, the dotted line is for wedge case 1, and the dashed line is for wedge case 2. $A(n)$, $B_z(n)$ and $B_x(n)$ are, respectively, the constant coefficient, the coefficient on z^m and the coefficient on x in equation (7) of section 2 for the zero-coupon aggregate equity price-dividend ratio paying out in n periods. $B_z(n)$ is multiplied by $(1 - \phi_z)$.

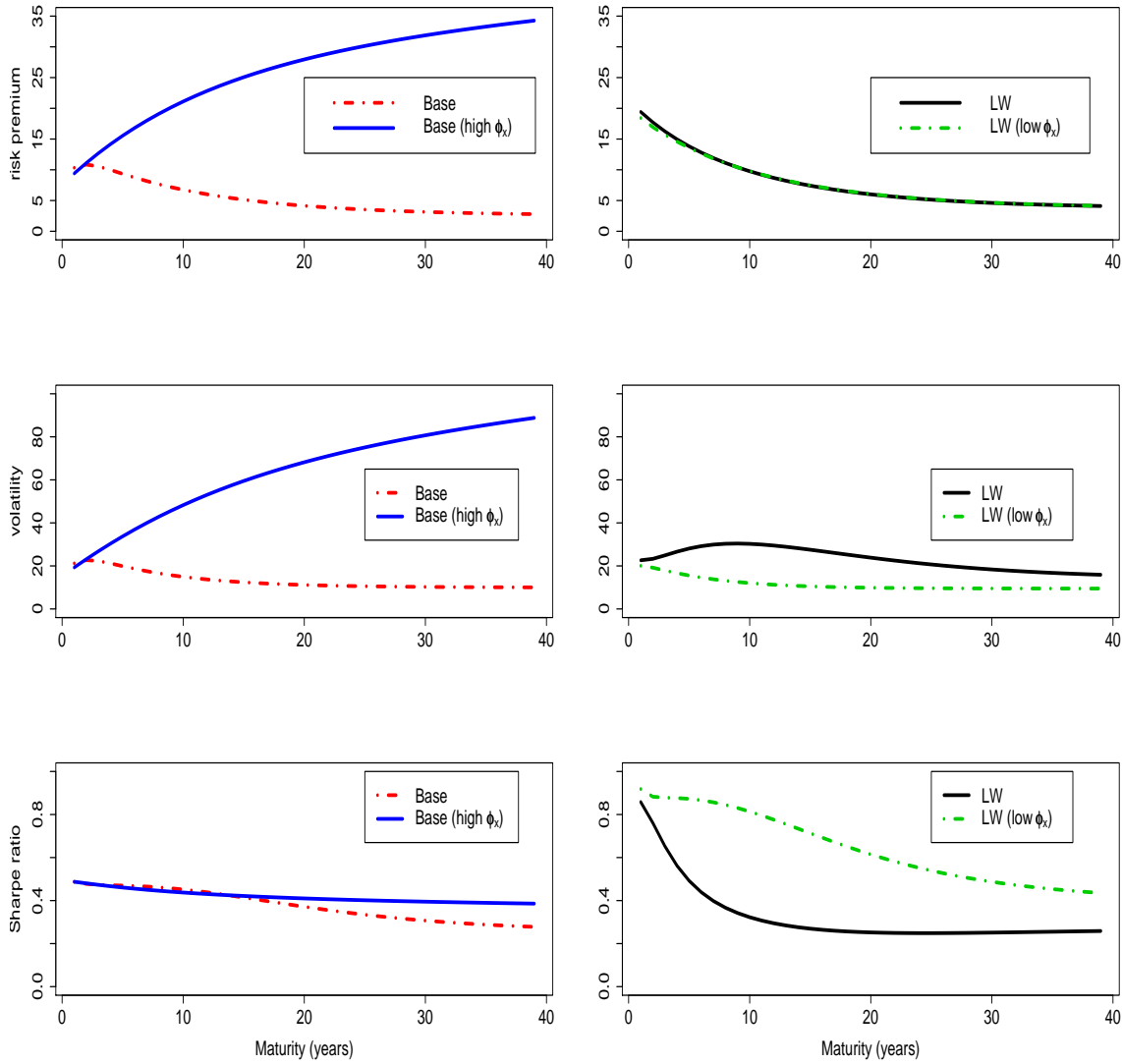


Figure 3: Returns on Zero-coupon Equity: Varying the Persistence of the x Process.

The figure shows the effects on the returns of zero-coupon equity of varying ϕ_x for 2 pairs of cases. The 2 cases in the first pair are plotted in the 3 LHS graphs and both cases have all their other parameters equal to the base case. The 2 cases in the second pair are plotted in the 3 RHS graphs and both cases have all their other parameters equal to those for the LW case. In each graph for a pair, the solid line is for the high- ϕ_x case and the dot-dashed line is for the low- ϕ_x case. For the LHS graphs, the solid line corresponds to the base+ $\Delta\phi_x$ case in Table 6 and the dot-dashed line to the base case, while for the RHS graphs, the solid line corresponds to the LW case and the dot-dashed line to the LW- $\Delta\phi_x$ case in the same table. For zero-coupon equity, the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return are all plotted as a function of maturity in the top two, middle two, and bottom two graphs respectively. The quarterly return on zero-coupon equity with n years to maturity is calculated as the return from holding zero-coupon equity with n years to maturity at the start of the quarter. The annual return on zero-coupon equity with n years to maturity is then obtained by rolling over these quarterly returns for 4 quarters. The models are simulated at a quarterly frequency for 4 million quarters, or until convergence, i.e. the first and second halves of the simulation yield sufficiently similar moments, whichever comes later.

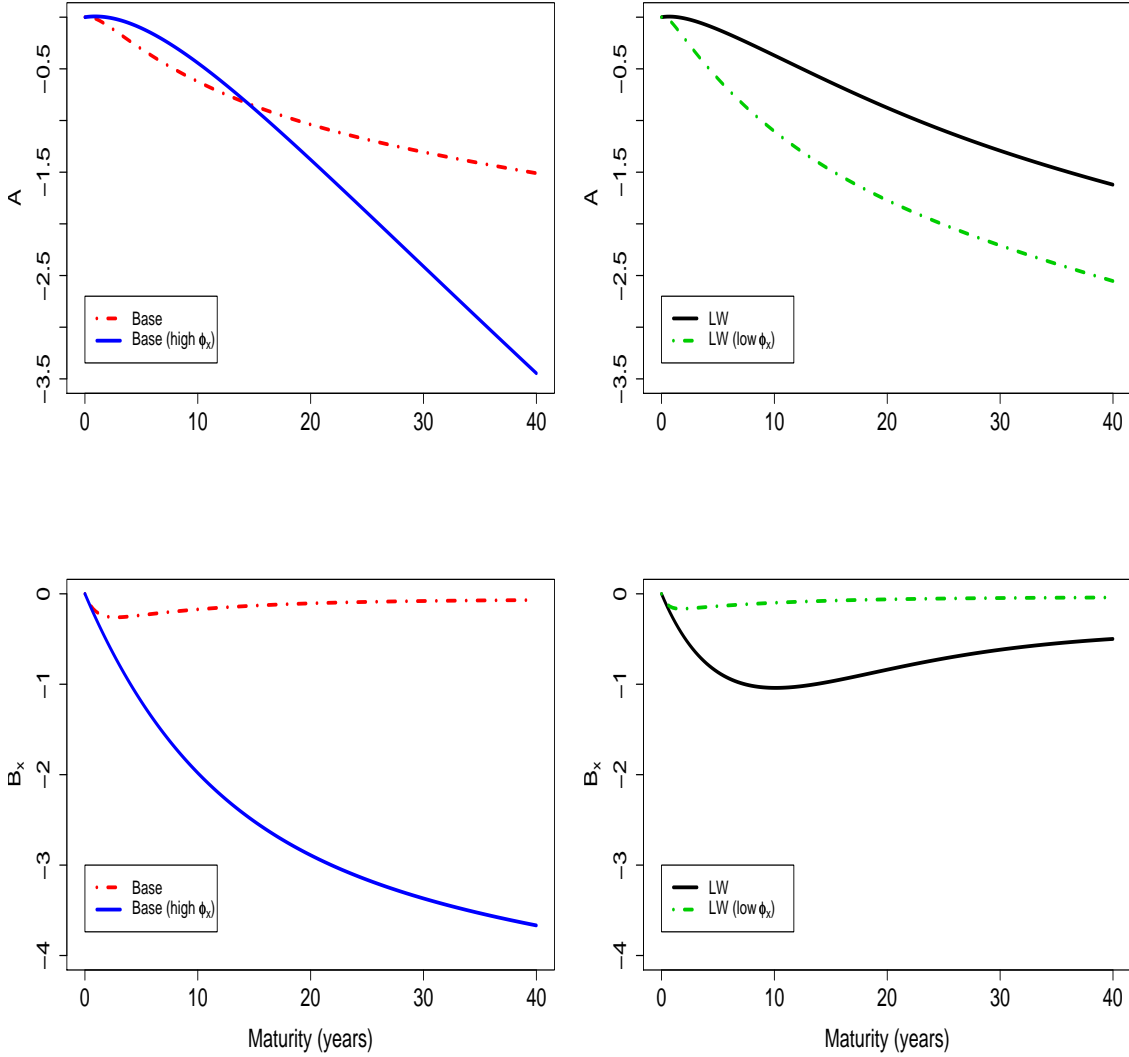


Figure 4: **Zero-coupon Aggregate Equity Log Price-dividend Ratio Coefficients $A(n)$ and $B_x(n)$: Varying the Persistence of the x Process.** The figure shows the effects on $A(n)$ and $B_x(n)$ for zero-coupon equity of varying ϕ_x for 2 pairs of cases. The 2 cases in the first pair are plotted in the 2 LHS graphs and both cases have all their other parameters equal to those for the base case. The 2 cases in the second pair are plotted in the 2 RHS graphs and both cases have all their other parameters equal to those for the LW case. In each graph for a pair, the solid line is for the high- ϕ_x case and the dot-dashed line is for the low- ϕ_x case. For the LHS graphs, the solid line corresponds to the base+ $\Delta\phi_x$ case in Table 6 and the dot-dashed line to the base case, while for the RHS graphs, the solid line corresponds to the LW case and the dot-dashed line to the LW- $\Delta\phi_x$ case in the same table. The figure plots $A(n)$ and $B_x(n)$ in the top 2 and bottom 2 graphs respectively. $A(n)$ and $B_x(n)$ are, respectively, the constant coefficient and the coefficient on x in equation (7) of section 2 for the zero-coupon aggregate equity price-dividend ratio paying out in n periods.

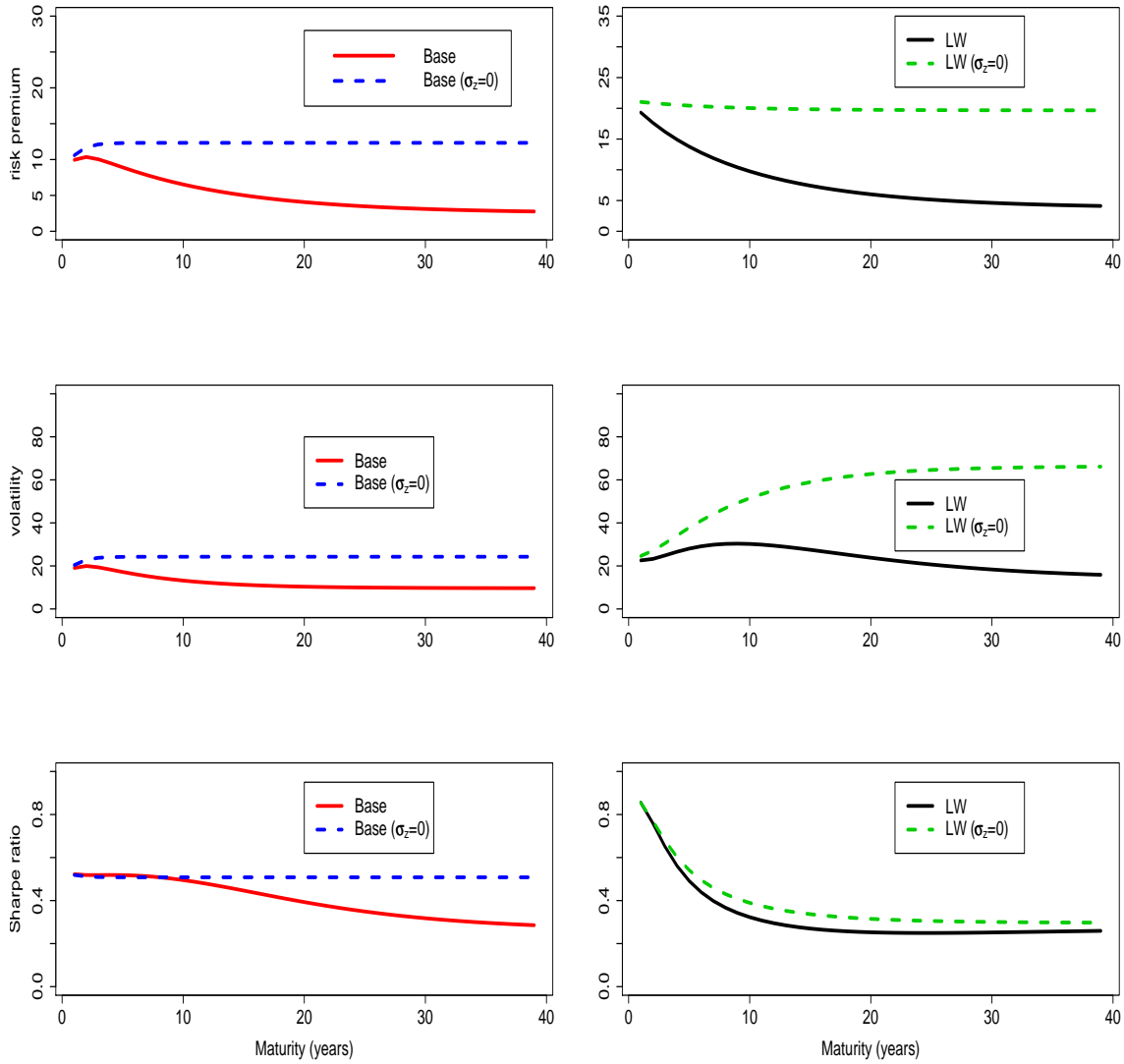


Figure 5: Returns on Zero-coupon Equity: Turning off Long-run Risk. The figure shows the effects on the returns of zero-coupon equity of long-run risk for 2 pairs of cases. The 2 cases in the first pair are plotted in the 3 LHS graphs and both cases have all their other parameters equal to the base case. The 2 cases in the second pair are plotted in the 3 RHS graphs and both cases have all their other parameters equal to those for the LW case. In each graph for a pair, the solid line is for the case where long-run risk is incorporated into the consumption growth process, and the dashed line is for the case where long-run risk is turned off. For the LHS graphs, the solid line corresponds to the base case in Table 7 and the dashed line to the base ($\sigma_z = 0$) case, while for the RHS graphs, the solid line corresponds to the LW case and the dashed line to the LW ($\sigma_z = 0$) case in the same table. For zero-coupon equity, the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return are all plotted as a function of maturity in the top two, middle two, and bottom two graphs respectively. The quarterly return on zero-coupon equity with n years to maturity is calculated as the return from holding zero-coupon equity with n years to maturity at the start of the quarter. The annual return on zero-coupon equity with n years to maturity is then obtained by rolling over these quarterly returns for 4 quarters. The models are simulated at a quarterly frequency for 4 million quarters, or until convergence, i.e. the first and second halves of the simulation yield sufficiently similar moments, whichever comes later.