WHY SURPLUS CONSUMPTION IN THE HABIT MODEL MAY BE LESS PERSISTENT
THAN YOU THINK

Anthony W. Lynch
Oliver Randall

Why Surplus Consumption in the Habit Model May be Less Persistent than You Think
Anthony W. Lynch and Oliver Randall
NBER Working Paper No. 16950
April 2011
JEL No. D91,E21,G12

ABSTRACT

In U.S. data, value stocks have higher expected excess returns and higher CAPM alphas than growth stocks. We find the external-habit model of Campbell and Cochrane (1999) can generate a value premium in both CAPM alpha and expected excess return so long as the persistence of the log surplus-consumption ratio is not too high. 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, the external-habit model generates a growth premium in expected excess return. However, the micro evidence favors a less persistent log surplus-consumption ratio. 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 model, 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 an external habit model in the spirit of Campbell and Cochrane can deliver an empirically sensible value premium once the persistence of the surplus consumption ratio is calibrated to the micro evidence rather than set to a value close to one. When we allow the conditional volatility of consumption growth to also be slowly mean reverting as in the long-run risk model of Bansal and Yaron, our model is also able to generate empirically sensible predictability of long-horizon returns using the price-dividend ratio, without eroding the value premium.

Anthony W. Lynch
New York University
44 W. 4th Street, #9-190
New York, NY 10012
and NBER
alynch@stern.nyu.edu

Oliver Randall
New York University
44 W. 4th Street, #9-190
New York, NY 10012
orandall@stern.nyu.edu

1 Introduction

A number of papers have considered how habit preferences impact the moments of the aggregate equity price-dividend ratio, the aggregate equity return, and the riskfree rate. Early papers by Constantinides (1990), and Sundaresan (1989) show how preferences with internal habit can generate a higher equity premium for a given curvature parameter, γ, while Abel (1990) obtains a similar result using external habit. One issue with habit preferences is 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 in such a way 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 rate 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 consumption growth generates a growth premium for expected excess 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 (less than 26%) than all past consumption from more than two years ago, which seems counterintuitive and appears to be inconsistent

with the micro evidence. 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.

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 the 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 approximately 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.

Our baseline specification follows LW and assumes that the aggregate consumption process and the aggregate dividend process are the same by calibrating both to the aggregate dividend process for U.S. stocks. It is unable to generate the aggregate equity return volatility found in the data. In addition to this specification, we also consider two specifications that allow the consumption process to differ from the dividend process, by calibrating the consumption process to data and leaving the dividend process the same. The first of these specifications continues to allow the aggregate equity return volatility to be much lower than that in the data, and generates an even larger value premium in both expected excess returns and CAPM alpha relative to the baseline specification that sets aggregate consumption and dividend equal. The second of these specifications moves aggregate equity return volatility much closer to the data, but still is able to generate a value premium in expected excess return that is considerably larger than that in the baseline specification, and a

3

value premium in CAPM alpha that is similar in magnitude to that in the baseline specification.

Unfortunately, these three specifications, with their implicit assumption that log consumption growth is homoscedastic, are unable to replicate the strong predictability of long-horizon equity returns found in the data when the price-dividend ratio is used as the predictor. However, when the consumption and dividend processes are specified to be different, we are able to obtain long-horizon return predictability of a magnitude much closer to that in the data, and without drastically reducing the value premia, by allowing the conditional volatility of consumption growth to also be slowly mean reverting. This specification with slowly mean reverting consumption volatility delivers value premia in both expected excess return and CAPM alpha that are larger than for any of the other three specifications: in fact, we are able to generate a value premium in expected excess return that is very close to the one found in the data using a book-to-market sort. This specification also comes very close to matching the volatility of aggregate equity return found in the data. Allowing the conditional volatility of consumption growth to be an autoregressive process is in the spirit of the second model in Bansal and Yaron (2004), which allows the conditional variance of consumption growth to be an autoregressive process. Long run risk in consumption volatility also helps match the data along several other dimensions: for example, while all our specifications counterfactually deliver negative market return autocorrelation, the specification with long run risk in consumption volatility, by producing the least negative autocorrelation, comes closest to matching the positive autocorrelation in the data.

Thus, our results suggest that an external habit model in the spirit of CC can deliver an empirically sensible value premium once the persistence of the surplus consumption ratio is calibrated to the micro evidence rather than set to a value close to one. Simultaneously allowing the conditional mean consumption growth to be slowly mean reverting delivers a log price-dividend ratio that exhibits empirically sensible persistence, without eroding the value premium. Also allowing the conditional volatility of consumption growth to be slowly mean reverting gives rise to empirically sensible predictability of long-horizon returns using the price-dividend ratio, again without eroding the value premium.

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, but 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

4

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 a 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 where 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.

Our paper is closely related to a recent paper by Santos and Veronesi (2008) which, like LW, finds that when firm cash flows 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, habit preferences 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 heterogenity 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 focuses on the time-series, rather than the cross-section, of expected returns. Kroce, Lettau and Ludvigson (2010) examine how incorporating limited information in a long-run risk model can result in short-duration assets having higher expected returns than long-maturity assets, as in the data. Using the long-run risk model of Bansal and Yaron for aggregate consumption growth together with Epstein-Zin preferences, Kiku (2006) documents how value stocks have relatively higher exposures to long-run consumption shocks while growth firm are more exposed to short-lived consumption fluctuations, and then shows how these different exposures lead to a value premium in expected return, CAPM alpha, and consumption-CAPM alpha. Finally, Hansen, Heaton, and Li (2008) report that the cash flows of value stocks but not growth stocks exhibit positive comovement with macroeconomic risks in the long run, and then examine how equilibrium pricing depend on investor preferences and the cash flow horizon.

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

2 The Model

We consider two versions of a model that is in the spirit of LW.

2.1 Model with One Price of Risk Variable

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 define $d_t^m \equiv \log(D_t^m)$. It evolves as follows:

$$\Delta d_{t+1}^m = g^m + z_t^m + \varepsilon_{t+1}^m \tag{1}$$

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

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

with $0 \leq \phi_z < 1$. Let D_t denote aggregate consumption at time t, and define $d_t \equiv \log(D_t)$. Log aggregate consumption growth evolves as follows:

$$\Delta d_{t+1} = g + z_t + \varepsilon^d_{t+1} \tag{3}$$

where $g \equiv \frac{g^m}{\delta^m}$ and $z_t \equiv \frac{z^m_t}{\delta^m}$. The shock to dividend growth is composed of a levered version of the shock to consumption growth plus an additional shock: $\varepsilon^m_{t+1} \equiv \delta^m \varepsilon^d_{t+1} + \varepsilon^u_{t+1}$. 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 $\varepsilon^u = 0$. Define $\sigma^2_i \equiv \sigma^2[\varepsilon^i]$ 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:

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

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:

$$M_{t+1} = \exp\left\{a + bz_t - \frac{1}{2}x^2_t - \frac{x_t}{\sigma_d}\varepsilon^d_{t+1}\right\}. \tag{5}$$

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

$$r^f_t \equiv -a - bz_t \tag{6}$$

If $b \neq 0$, the riskless rate is time varying. Since the most relevant papers to ours, LW and CC, both assume that the riskfree rate is constant, we assume this too, i.e. that $b = 0$, so we can directly compare our results to theirs.

We consider four cases using this version of model. We examine a case, the LW case, that essentially replicates LW by having the shocks to x and d be uncorrelated ($\rho_{x,d} = 0$), the x process highly persistent (ϕ_x close to 1), and a consumption process that matches the dividend process which has been calibrated to data. Our base case also sets the consumption process equal to the calibrated dividend process, but allows the x process to be less persistent, as suggested by recent evidence about the persistence of habit, and $\rho_{x,d} = -0.99$, as implied by the habit specification used in CC. We also examine two wedge cases that resemble our base case except that the consumption process is calibrated to data rather than matched to the dividend process. Further details of the calibrations follow in section 3.

2.1.1 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:

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

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:

$$
\begin{aligned}
A(n) &= A(n-1) + a + g^m + B_x(n-1)\bar{x}(1-\phi_x) + \frac{1}{2}(C_{n-1}^m)'\Sigma_{\varepsilon,\varepsilon}C_{n-1}^m \\
B_x(n) &= \phi_x B_x(n-1) - \frac{1}{\sigma_d}\Sigma_{d,\varepsilon}C_{n-1}^m \\
B_z(n) &= \frac{(1+b/\delta^m)(1-\phi_z^n)}{1-\phi_z}
\end{aligned}
\tag{8}
\tag{9}
$$

where $C_n^m \equiv [\delta^m \ 1 \ B_x(n) \ B_z(n)]'$, $\varepsilon \equiv [\varepsilon^d \ \varepsilon^u \ \varepsilon^x \ \varepsilon^z]'$, $\Sigma_{d,\varepsilon} \equiv E[\varepsilon^d \varepsilon']$, and $\Sigma_{\varepsilon,\varepsilon} \equiv E[\varepsilon\varepsilon']$.

Let $R_{n,t+1}^m$ be the return from time t to $t+1$ of a claim to zero-coupon market equity paying off at time $t+n$, and define $r_{n,t+1}^m \equiv \log(R_{n,t+1}^m)$. It can be shown that (see LW):

$$r_{n,t+1}^m = E_t[r_{n,t+1}^m] + (C_{n-1}^m)'\varepsilon_{t+1} \tag{10}$$

$$\sigma_t^2[r_{n,t+1}^m] = (C_{n-1}^m)'\Sigma_{\varepsilon,\varepsilon}C_{n-1}^m. \tag{11}$$

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:

$$
\begin{aligned}
\log\left(E_t\left[\frac{R_{n,t+1}^m}{R_t^f}\right]\right) &= E_t[r_{n,t+1}^m - r_t^f] + \frac{1}{2}\sigma_t^2[r_{n,t+1}^m] \\
&= \left(\delta^m\sigma_d^2 + \sigma_{d,u} + B_x(n-1)\sigma_{x,d} + B_z(n-1)\sigma_{z,d}\right)\frac{1}{\sigma_d}x_t
\end{aligned}
\tag{12}
$$

2.1.2 Implications for the value/growth premium

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

8

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 when the persistence of x is assumed to be low.

2.2 Model with heteroscedasctic log consumption growth

The base and wedge cases fail to match the price-dividend ratio's ability to predict the returns of long-horizon equity we see in the data. To match this feature, we consider a second model for which the conditional volatility of log consumption growth, σ_t, is a highly persistent AR(1) process, i.e.

log consumption growth evolves as:

$$\Delta d_{t+1} = g + z_t + \sigma_t \varepsilon_{t+1}^d \tag{13}$$

where

$$\sigma_{t+1} = \bar{\sigma} + \phi_\sigma(\sigma_t - \bar{\sigma}) + \varepsilon_{t+1}^w \tag{14}$$

and $\varepsilon_{t+1}^w \sim N(0, \sigma_w^2)$, uncorrelated with the other shocks. This specification is closely related to Bansal and Yaron (2004), who specify that the variance, not the volatility, is an AR(1). The stochastic discount factor of our base and wedge cases becomes:

$$M_{t+1} = \exp\left\{ a + bz_t - \frac{1}{2}(x_t\sigma_t)^2 - x_t\sigma_t \frac{\varepsilon_{t+1}^d}{\sigma_d} \right\} \tag{15}$$

Using a first order Taylor approximation, we can approximate the price of risk as follows:

$$x_t\sigma_t \approx \bar{x}\bar{\sigma} + \bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma}(x_t - \bar{x}) \tag{16}$$

$$= \bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma}x_t, \tag{17}$$

which gives us the following stochastic discount factor

$$M_{t+1} = \exp\left\{ a + b^m z_t^m - \frac{1}{2}(\bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma}x_t)^2 - (\bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma}x_t)\frac{\varepsilon_{t+1}^d}{\sigma_d} \right\} \tag{18}$$

We consider one case using this version of model, a long run risk in volatility case (the LRR-vol case), that resembles the two wedge cases described above, except that log consumption growth's conditional volatility, as well as its conditional mean, is allowed to be mean reverting. As mentioned above, the aggregate consumption process considered in the LRR-vol case is in the spirit of Bansal and Yaron (2004), except that in the LRR-vol case, aggregate consumption growth's conditional volatility is an AR(1) process, while in Bansal and Yaron, its conditional variance is an AR(1) process.[1] Further details of the calibration of this case are provided in section 3.

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

It can be shown that $P_{n,t}^m$ takes the following recursive form for this model:

$$\frac{P_{n,t}^m}{D_t^m} = F(x_t, \sigma_t, z_t^m, n) = \exp\{A(n) + B_x(n)x_t + B_\sigma(n)(\sigma_t - \bar{\sigma}) + B_z(n)z_t^m\} \tag{19}$$

---

[1]We also ran a case where σ_t^2 is an AR(1), using a first order Taylor expansion for $x_t\sqrt{\sigma_t^2}$ in the stochastic discount factor, and the results were qualitatively the same.

Using the boundary condition $P^m_{0,t} = D^m_t$ we see $A(0) = B_z(0) = B_x(0) = B_\sigma(0) = 0$, and proceeding by induction on n, we can show the following recursive relationships hold:

$$A(n) = A(n-1) + a + g^m + B_x(n-1)\bar{x}(1-\phi_x) + \frac{1}{2}(C^m_{n-1})'\Sigma_{\varepsilon,\varepsilon}C^m_{n-1}$$

$$B_x(n) = \phi_x B_x(n-1) - \frac{\bar{\sigma}}{\sigma_d}\Sigma_{d,\varepsilon}C^m_{n-1} \tag{20}$$

$$B_\sigma(n) = \phi_\sigma B_\sigma(n-1) - \frac{\bar{x}}{\sigma_d}\Sigma_{d,\varepsilon}C^m_{n-1} \tag{21}$$

$$B_z(n) = \frac{(1+b/\delta^m)(1-\phi^n_z)}{1-\phi_z} \tag{22}$$

where $C^m_n \equiv [\delta^m\ 1\ B_x(n)\ B_\sigma(n)\ B_z(n)]'$, $\varepsilon \equiv [\varepsilon^d\ \varepsilon^u\ \varepsilon^x\ \varepsilon^w\ \varepsilon^z]'$, $\Sigma_{d,\varepsilon} \equiv E[\varepsilon^d\varepsilon']$, and $\Sigma_{\varepsilon,\varepsilon} \equiv E[\varepsilon\varepsilon']$.

It can be shown that $r^m_{n,t+1}$ can be written as follows in this model:

$$r^m_{n,t+1} = E_t[r^m_{n,t+1}] + (C^m_{n-1})'\varepsilon_{t+1} \tag{23}$$

$$\sigma^2_t[r^m_{n,t+1}] = (C^m_{n-1})'\Sigma_{\varepsilon,\varepsilon}C^m_{n-1}. \tag{24}$$

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

$$\log\left(E_t\left[\frac{R^m_{n,t+1}}{R^f_t}\right]\right)$$
$$= E_t[r^m_{n,t+1} - r^f_t] + \frac{1}{2}\sigma^2_t[r^m_{n,t+1}] \tag{25}$$
$$= \left(\delta^m\sigma^2_d + \sigma_{d,u} + B_x(n-1)\sigma_{x,d} + B_\sigma(n-1)\sigma_{w,d} + B_z(n-1)\sigma_{z,d}\right)\left(\frac{\bar{\sigma}}{\sigma_d}x_t + \frac{\bar{x}}{\sigma_d}(\sigma_t - \bar{\sigma})\right)$$

In the LRR-vol case, ε^w is uncorrelated with all other shocks. So we can see that for this case, the first expression in parentheses on the right hand side of equation (25) has the same terms as the first expression in parentheses on the right hand side of equation (12). Now $\bar{x}$ is always positive, and $\bar{\sigma}$ is positive in the LRR-vol case. So holding σ^2_d, $\sigma_{x,d}$ and $\sigma_{z,d}$ fixed, the shape of $E[E_t[r^m_{n,t+1} - r^f_t] + \frac{1}{2}\sigma^2_t[r^m_{n,t+1}]]$ as a function of n in the first model and in the LRR-vol case depends on the shapes of $B_x(n-1)$ and $B_z(n-1)$ as functions of n. So if the shapes of $B_x(n-1)$ and $B_z(n-1)$ as functions of n remain similar once σ_t is allowed to be slowly mean-reverting rather than constant, then allowing σ_t to be slowly mean reverting rather than constant won't affect the ability of the CC model with low persistence of the surplus consumption ratio to deliver a value premium in expected excess return.

2.3 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. We specify that dividends are paid at a quarterly frequency, so we can calculate the annual price-dividend ratio as follows:

$$\frac{P_t^m}{\sum_{\tau=0}^3 D_{t-\tau}^m} = \sum_{n=1}^{\infty} \frac{P_{n,t}^m}{\sum_{\tau=0}^3 D_{t-\tau}^m}. \tag{26}$$

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

$$R_{t+1}^m = \frac{P_{t+1}^m + D_{t+1}^m}{P_t^m} \tag{27}$$

$$= \left(\frac{P_{t+1}^m/D_{t+1}^m + 1}{P_t^m/D_t^m}\right)\left(\frac{D_{t+1}^m}{D_t^m}\right). \tag{28}$$

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.[2]

2.4 Relation to other models

These two specifications are related to a number of other models.

2.4.1 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 riskfree rate, and is constant over time. Notice that our first model nests LW by setting $a = -r^f$, $b = 0$, $\delta^m = 1$, and $\sigma_u = 0$.

---

[2]We 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.

2.4.2 CC with i.i.d. Consumption Growth

CC assume that a representative agent maximizes the utility function:

$$E \sum_{t=0}^{\infty} \delta^t \frac{(D_t - H_t)^{1-\gamma} - 1}{1 - \gamma} \tag{29}$$

where H_t is the level of external habit at time t and δ is the subjective discount factor. Defining $s_t \equiv \log\left(\frac{D_t - H_t}{D_t}\right)$, the log of the surplus-consumption ratio at time t, they specify the following dynamics:

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

where $\varepsilon^d \sim N(0, \sigma_d^2)$ and $\lambda(.)$ is a sensitivity function. They specify the sensitivity function to be:

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

where $\bar{S} \equiv \sigma_d \sqrt{\frac{\gamma}{1-\phi_s}}$, $\bar{s} \equiv \log(\bar{S})$, and $s_{\max} = \bar{s} + \frac{1}{2}(1 - (\bar{S})^2)$. These dynamics imply a stochastic discount factor equal to:

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

Our first model approximates CC by setting $a = \log(\delta) - \gamma g + \frac{\gamma(1-\phi_s)}{2}$, $\delta^m = 1$, $\sigma_u = 0$, $\sigma_z = 0$, and $x_t = \gamma\sigma_d(1 + \lambda(s_t))$. The model approximates the heteroskedastic process for $\gamma\sigma_d(1 + \lambda(s_t))$ in CC by specifying x_t as a homoskedastic AR(1) process. As long as the sensitivity function is rarely zero, it follows that our first model can approximate CC when $\rho_{d,x} \approx -1$ and $\phi_x \approx \phi_s$.

2.4.3 CC with Persistent Conditional Mean Consumption Growth

CC with persistent conditional mean consumption growth can be approximated by the first model when $\sigma_z \neq 0$. Suppose the representative agent again maximizes the habit specification in equation (29), but the conditional mean of aggregate consumption growth is slowly mean-reverting, following equations (2) and (3). We extend the dynamics for the log consumption surplus in CC to the case in which there is long run risk in mean consumption growth by assuming that CC's sensitivity function loads on the innovation to log consumption growth above its conditional mean, $\Delta d_{t+1} - g - z_t$,

which is equal to ε_{t+1}^d. We also assume that the log consumption surplus depends linearly on z_t with coefficient $\lambda(\bar{s})$. Putting these two together, the consumption surplus evolves as follows:

$$s_{t+1} = (1 - \phi_s)\bar{s} + \phi_s s_t + \lambda(\bar{s})z_t + \lambda(s_t)\varepsilon_{t+1}^d \qquad (30)$$

where $\lambda(.)$ is the same sensitivity function as used by CC and described in the previous subsection.

Using the same sensitivity function as in CC, and setting z_t's loading to be the sensitivity function evaluated at the steady state surplus consumption value, allows the riskfree rate to depend only on z_t. The specification in (30) also implies the following desirable properties for the habit process: at the consumption surplus's steady state, log habit is predetermined only by an exponentially-weighted sum of past lagged log consumption (see section 2.5 below); and, habit next period moves positively with consumption next period irrespective of the consumption surplus this period. Note that CC's specification for surplus consumption implies that their habit process satisfied these same two properties, given their assumptions about the consumption growth process.

This specification implies the following stochastic discount factor:

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

Matching coefficients in the stochastic discount factor we see that the riskfree rate is affine in z_t. So we can approximate CC with persistent mean consumption growth using our first model by setting $a = \log(\delta) - \gamma g + \frac{\gamma(1 - \phi_s)}{2}$, $b = -\gamma(1 + \lambda(\bar{s}))$, $\delta^m = 1$, $\sigma_u = 0$ and $x_t = \gamma\sigma_d(1 + \lambda(s_t))$. As in the previous subsection, our first model uses x_t, a homoskedastic AR(1) process, to approximate $\gamma\sigma_d(1 + \lambda(s_t))$, a heteroskedastic AR(1) process, and so, as long as the sensitivity function is rarely zero, $\rho_{d,x} \approx -1$ and $\phi_x \approx \phi_s$.

2.4.4 CC with Persistent Conditional Mean and Volatility of Consumption Growth

CC with persistent conditional mean and volatility of consumption growth can be approximated by our second model when σ_z and σ_w are both strictly positive. Suppose the representative agent again maximizes the habit specification in equation (29) but both the conditional mean and volatility of aggregate consumption growth are slowly mean-reverting, following equations (2), (13) and (14). We extend the dynamics for the log consumption surplus in CC to the case in which there is long run risk in the mean and volatility of consumption growth, by assuming that CC's sensitivity function loads on the innovation to log consumption growth above its conditional mean, $\Delta d_{t+1} - g - z_t$,

which becomes equal to $\sigma_t \varepsilon_{t+1}^d$. As in the previous subsection, we assume that the log consumption surplus depends linearly on z_t with coefficient $\lambda(\bar{s})$. Specifically we assume the consumption surplus evolves as follows:

$$s_{t+1} = (1 - \phi_s)\bar{s} + \phi_s s_t + \lambda(\bar{s})z_t + \lambda(s_t)\sigma_t \varepsilon_{t+1}^d \tag{31}$$

where $\lambda(.)$ is the same sensitivity function as used by CC which is described in subsection 2.4.2.

Using the same sensitivity function as in CC, and setting z_t's loading to be the sensitivity function evaluated at the steady state surplus consumption value, allows the riskfree rate to depend only on z_t and σ_t^2. The specification in (31) also delivers the same two desirable properties for the habit process that we obtained in the previous subsection.

This specification implies the following stochastic discount factor:

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

Matching coefficients in the stochastic discount factor we see that the riskfree rate is affine in z_t and σ_t^2. So our second model can be used to approximate CC with persistent conditional mean and volatility of consumption growth by first using the same approximation in (16) applied to $\gamma(1 + \lambda(s_t))$ and σ_t, and then setting $a = \log(\delta) - \gamma g + \frac{\gamma(1 - \phi_s)}{2}$, $b = -\gamma(1 + \lambda(\bar{s}))$, $\delta^m = 1$, $\sigma_u = 0$ and $x_t = \gamma\sigma_d(1 + \lambda(s_t))$ and σ_t equal to itself. As in the previous subsections, our second model approximates $\gamma\sigma_d(1 + \lambda(s_t))$, a heteroskedastic AR(1) process, with x_t, an homoskedastic AR(1) process. So again, as long as the sensitivity function is rarely zero, $\rho_{d,x} \approx -1$ and $\phi_x \approx \phi_s$.

2.4.5 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.

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

2.5 Relation between External Habit and Past Consumption

Following an earlier version of CC, we can show that log habit is approximately a moving average of lagged log consumption, for the specification of log consumption growth and the log surplus consumption ratio in subsection 2.4.4, which allows the conditional mean and volatility of consumption growth to be slowly mean reverting as in Bansal and Yaron (2004). Define $h_t \equiv \log(H_t)$, and apply a log-linear approximation to the definition of s_t:

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

Substituting this in to the law of motion for s described in (31), and utilizing the imposed restriction that h_{t+1} is predetermined at the steady state, we can show that:

$$
h_{t+1} \approx \overline{h-d} + (1 - \phi_s)\sum_{j=0}^{\infty}(\phi_s)^j\, d_{t-j} + \frac{g}{1 - \phi_s} \tag{32}
$$

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 property of the specification for s_t given in equation (31) when consumption growth has a persistent conditional mean and volatility 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:

$$
h_{t+1} - h_t \approx g + (1 - \phi_s)\left[(d_t - h_t) - \overline{d-h}\right] \tag{33}
$$

The lower the persistence of the surplus-consumption ratio, the more impact the most recent consumption has on habit. Notice that these expressions also hold for the specification of log consumption growth and the log surplus consumption ratio in subsection 2.4.3, which only allow the conditional mean of consumption growth to be slowly mean reverting. This follows because the specification in subsection 2.4.4 nests the one in 2.4.3.

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 (32) 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.

2.6 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 dividend 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. LW choose a growth rate of 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. Following LW subject to rounding, we choose a growth rate of 5.5% per quarter for the share process. With this growth rate choice, firm 1's 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.

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

$$P_t^i = \sum_{n=1}^{\infty} s_{t+n}^i P_{n,t}^m \tag{34}$$

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.

2.7 Forming the Value/Growth Deciles

Recall that we specify a period in the model to be a quarter as in LW. 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 $P_t^i / \sum_{\tau=0}^{3} D_{t-\tau}^i$ for firm i. We calculate moments for the decile excess annual returns and annual CAPM alpha 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 their parameter values.[3] 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 and a LRR-vol case in which the aggregate consumption process is allowed to differ from the aggregate dividend process. Consumption growth is homoscedastic in the two wedge cases, and heteroscedastic in the LRR-vol case. Table 1 reports the parameters used by these 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, two wedge and LRR-vol 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, two wedge and LRR-vol cases.[4]

While the model is quarterly, the log riskfree rate r^f is converted into an annual number in Table 1 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. That is, $\sigma[\varepsilon^x, \varepsilon^z]$ 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

---

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

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

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

$$
\begin{aligned}
\sigma[\varepsilon^x, \varepsilon^z] &= \sigma\left[\beta_{x,d}\beta_{d,z}\varepsilon^z + \beta_{x,d}u^d + u^x, \varepsilon^z\right] \\
&= \left(\rho[\varepsilon^d, \varepsilon^x]\rho[\varepsilon^d, \varepsilon^z] + \left[1 - \rho[\varepsilon^d, \varepsilon^x]^2\right]^{\frac{1}{2}}\rho[u^x, \varepsilon^z]\right)\sigma_z\sigma_x \quad (35)
\end{aligned}
$$

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 (35) with $\rho[u^x, \varepsilon^z] = 0$ to calculate $\sigma[\varepsilon^x, \varepsilon^z]$. Notice this specification has the attractive property that when $\sigma[\varepsilon^d, \varepsilon^z]$ 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 0.81. Note that this correlation measures the correlation of the shock to future expected returns with the shock to future expected consumption growth (which is also the shock to future expected dividend growth in the base case).

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, equation (7) in section 2, suggests that the autocorrelation of the aggregate market's price-dividend ratio is approximately 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 (32) 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. Subsection 2.1.2 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 0.41 value obtained by LW (0.42 in our simulation of 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. we set $\delta^m = 1$ 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:

$$\sigma^2[\varepsilon^m] = (\delta^m)^2 \sigma_d^2 + \sigma_u^2 + 2\delta^m \sigma_{d,u} \tag{36}$$

$$\sigma[\varepsilon_{t+1}^m, \varepsilon_{t+1}^z] = \delta^m \sigma_{d,z} + \sigma_{u,z} \tag{37}$$

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

$$\sigma \left[\sum_{i=1}^{4} \Delta d_{t+i}, \sum_{i=1}^{4} \Delta d_{t+i}^m \right]$$

$$= \frac{1}{(1-\phi_z)^2 \delta^m} \left((1-\phi_z^4)(1+\phi_z^2) + 3 - \frac{2\phi_z(1-\phi_z^3)}{1-\phi_z} + \frac{\phi_z^2(1-\phi_z^6)}{1-\phi_z^2} \right) \sigma_z^2$$
$$+ \frac{2}{(1-\phi_z)^2}(3 - 4\phi + \phi^4)\sigma_{z,d} + \frac{1}{\delta^m(1-\phi_z^2)}(3 - 4\phi_z + \phi_z^4)\sigma_{z,u} + 4\delta^m \sigma_d^2 + 4\sigma_{d,u}$$

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:

$$\rho[\varepsilon_{t+1}^m, \varepsilon_{t+1}^d] = \frac{\delta^m \sigma_d}{\sigma[\varepsilon^m]} + \frac{\sigma_{d,u}}{\sigma[\varepsilon^m]\sigma_d}. \tag{38}$$

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 $\bar{x} > 0$ in the base case.

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 square this and match it to our analytical expression for the variance of annual log consumption growth:

$$\sigma^2 \left[\sum_{i=1}^{4} \Delta d_{t+i} \right] = \left(\frac{(1+\phi_z+\phi_z^2+\phi_z^3)^2}{1-\phi_z^2} + 1 + (1+\phi_z)^2 + (1+\phi_z+\phi_z^2)^2 \right) \frac{\sigma_z^2}{(\delta^m)^2}$$
$$+ 4\sigma_d^2 + \frac{2\sigma_{d,z}(3+2\phi_z+\phi_z^2)}{\delta^m} \tag{39}$$

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 (36)-(39) 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, we set $\rho[\varepsilon^d, \varepsilon^x] = -0.99$ as in the base case, in the spirit of CC, and $\phi_x = 0.14$ as in the base case, in the spirit of the micro evidence concerning the persistence of the log consumption surplus. 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 obtained using equation (35) with $\rho[u^x, \varepsilon^z] = 0$, while $\sigma[\varepsilon^x, \varepsilon^u]$ is calculated similarly using

$$\begin{aligned} \sigma[\varepsilon^x, \varepsilon^u] &= \sigma\left[\beta_{x,d}\beta_{d,z}\varepsilon^u + \beta_{x,d}u^d + u^x, \varepsilon^u\right] \\ &= \left(\rho[\varepsilon^d, \varepsilon^x]\rho[\varepsilon^d, \varepsilon^u] + \left[1 - \rho[\varepsilon^d, \varepsilon^x]^2\right]^{\frac{1}{2}}\rho[u^x, \varepsilon^u]\right)\sigma_z\sigma_x, \end{aligned} \quad (40)$$

with $\rho[u^x, \varepsilon^u]$ set equal to 0 to calculate $\sigma[\varepsilon^x, \varepsilon^u]$.

Both wedges cases choose the $\bar{x}$ and σ_x values to match the mean of the price-dividend ratio and the unconditional Sharpe ratio. Table 3 shows that both wedge cases are able to match both these aggregate moments about as well as the LW and base cases. However, the first wedge case, like the base case, does not try to match the unconditional volatility of the excess return on aggregate equity, understating it by a magnitude comparable to the base case, while the second wedge case tries to match this unconditional volatility and does a much better job than the first of doing so.

3.3 Calibration of the LRR-Vol case: making the conditional volatility of log consumption growth stochastic

When we calibrate the LRR-vol case, ϕ_x is set at 0.14 annually and $\rho_{d,x}$ is set equal to -0.99, as in the base and wedge cases. The expressions for $\sigma^2[\varepsilon^m]$ and $\sigma[\varepsilon^m_{t+1}, \varepsilon^z_{t+1}]$ are the same as in the

wedge cases. The expression for the annual variance of consumption growth becomes

$$\sigma^2 \left[\sum_{i=1}^{4} \Delta d_{t+i} \right] = \left(\frac{(1 + \phi_z + \phi_z^2 + \phi_z^3)^2}{1 - \phi_z^2} + 1 + (1 + \phi_z)^2 + (1 + \phi_z + \phi_z^2)^2 \right) \frac{\sigma_z^2}{(\delta^m)^2}$$
$$+ 4\sigma_d^2 E[\sigma_t^2] + \frac{2\sigma_{d,z}(3 + 2\phi_z + \phi_z^2)}{\delta^m} \bar{\sigma}, \tag{41}$$

and the expression for the quarterly correlation between the shocks to dividend growth and consumption growth becomes

$$\rho[\varepsilon_{t+1}^m, \sigma_t \varepsilon_{t+1}^d] = \frac{(\delta^m \sigma_d^2 + \sigma_{d,u})\bar{\sigma}}{\sigma_m \sigma_d \sqrt{E[\sigma_t^2]}}. \tag{42}$$

We choose the values of δ^m, $\rho_{d,u}$, g^m, ϕ_z and σ_z as in the wedge cases, so that the joint distribution of z^m and d^m is the same as in the base and wedge cases.

When we calibrate the process for σ_t in the LRR-vol case, we want the shock to monthly log consumption growth to match that used by Bansal, Kiku, and Yaron (2009). To do this, we start by simulating the conditional variance of monthly log consumption growth using the AR(1) specification and parameters from Bansal, Kiku, and Yaron. We discard any negative draws, as they do. We approximate the quarterly variance by the sum of the variance for the 3 consecutive months in the quarter. We then compute the quarterly volatility as the square root of this quarterly time series, and fit the volatility to an AR(1) process. This nails down the value for ϕ_σ. Since ε^d in Bansal, Kiku, and Yaron is scaled to have unit variance, while ours is not, we scale our volatility process to preserve the unconditional second moment of the shock to log consumption growth from the wedge cases, i.e. $E[(\sigma_t \varepsilon_{t+1}^d)^2]$ in the LRR-vol case equals $E[(\varepsilon_{t+1}^d)^2]$ in the wedge cases. Combining this with the 4 moment conditions in (36)-(37) and (41)-(42), we solve for the volatility scaling factor, σ_d, σ_u, $\sigma_{d,z}$ and $\sigma_{z,u}$. These values nail down the values for $\bar{\sigma}$ and σ_w. It follows that the unconditional correlation between the shock to log consumption growth and the shock to the mean of log consumption growth is the same as in the wedge cases; i.e. $\rho[\sigma_t \varepsilon_{t+1}^d, \varepsilon_{t+1}^z]$ in the LRR-vol case equals $\rho[\varepsilon_{t+1}^d, \varepsilon_{t+1}^z]$ in the wedge cases. The estimated parameters for $\bar{\sigma}$, ϕ_σ and σ_w are given in Table 1. The calibrated volatility process goes negative less than 0.3% of the time. Finally, we follow Bansal and Yaron (2004) and impose that ε^w is uncorrelated with all other shocks.

Turning to the x process, $\sigma[\varepsilon^x, \varepsilon^z]$ is obtained using equation (35) with $\rho[u^x, \varepsilon^z]$ set equal to zero, and $\sigma[\varepsilon^x, \varepsilon^u]$ is obtained using equation (40) with $\rho[u^x, \varepsilon^u]$ set equal to zero. The reasons for doing this are the same as in the base and wedge cases. Finally, just as with the second wedge cases, $\bar{x}$ and σ_x are chosen to best match the data values for the unconditional volatility of the excess return on aggregate equity, as well as its unconditional Sharpe ratio and its expected price-dividend ratio.

As Table 3 shows, the LRR-vol case matches these three aggregate moments better than either wedge case.

4 Results

This section reports the results for the five cases: LW, base, two wedge, and LRR-vol. We also report results for U.S. data. The data are the same as that in LW: the aggregate data are annual from 1890-2002, while 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 $12^{0.5}$.

Each model is simulated at a quarterly frequency for at least 4 million quarters, and until the mean price-dividend ratio and Sharpe ratio of the aggregate market and the two value premia have converged: i.e., until an additional 10% of simulated quarters causes the values of all these variables to change by less than small prespecified tolerances. The unconditional mean and volatility of annual excess market equity return on zero-coupon equity with n years to maturity are simulated to convergence in the same way.

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

24

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 denotes autocorrelation.

Since the parameters of Δd^m are 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 that the base-case value of this autocorrelation, 0.895, is higher than, but close to, both the data value of 0.87 and the LW value of 0.883.

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 of 0.261 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 0.382. 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 much less so than in our base case, at a value of -0.04.

4.1.2 Predictive Regressions

Table 4 reports results for three predictive regressions. The regression reported in the top panel regresses the future log excess aggregate equity return on today's log aggregate equity price-dividend ratio. The regression reported in the middle panel regresses future changes in log aggregate equity dividend on today's log aggregate equity price-dividend ratio. The regression reported in the bottom panel regresses future changes in log aggregate equity dividend on today's consumption-aggregate equity dividend ratio for the data and today's z^m for the cases. 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 of 1 and 10 years for future return and dividend growth. 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^2s 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^2s 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 the log price-dividend ratio, especially at the 10 year horizon, where the sign of the predictive coefficient is negative for the data and positive for the base case. All the cases considered calibrate the joint process for log market dividend and z^m in exactly the same way as LW. Consequently, when forecasting 1 or 10 year log dividend growth using z^m for the cases and a proxy for z^m in the data, all the cases considered here replicate LW's ability to match the R^2s of the regressions and also their ability to match the sign but not the magnitude of the predictive coefficients.

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: $P_t^i/D_t^i = P_t^i/\sum_{\tau=0}^3 D_{t-\tau}^i$ for firm i. $Sharpe^i$ is the unconditional Sharpe ratio for portfolio i's annual return while R_i^2 is the regression R^2.

Table 5 shows that the base case can generate a positive value premium in both expected 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 expected 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 expected 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, which is not surprising given that the dividend shares received by the firms are deterministic. Consistent with this observation, the R^2s 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^2s 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 lowest for the extreme growth decile for the two cases and all the data sorts, and is highest for the extreme value 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 expected excess return and CAPM alpha. To better understand why the base case delivers a value premium in expected 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 excess 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 black 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 excess returns 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 expected 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 expected 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 (10)-(12), 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 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 black 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 price-dividend ratio of zero-coupon aggregate equity paying out in n periods. $B_z(n)$ is multiplied by $(1 - \phi_z)$. Note that $B_z(n)$ is the same in all cases considered, including these two.

As was pointed out in subsection 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. The right hand side of equation (12) provides an expression for $E_t[R^m_{n,t+1}/R^f_t]$, and since $E_t[R^m_{n,t+1}/R^f_t]$ is likely to be highly correlated with $E_t[R^m_{n,t+1} - R^f_t]$, this right hand side of (12) can be used to understand how the unconditional expected excess return on a zero-coupon equity claim $(E[R^m_{n,t+1} - R^f_t])$ 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 $E_t[R^m_{n,t+1}/R^f_t]$ 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, $E_t[R^m_{n,t+1}/R^f_t]$ is decreasing in n as reported in Figure 1. Hence the LW case delivers a value premium in expected

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 $E_t[R_{n,t+1}^m/R^f]$. 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 (12) is hump-shaped and the implication is that $E_t[R_{n,t+1}^m/R^f]$ can be hump-shaped, as reported in Figure 1. Hence, the base case is able to deliver a value premium in expected 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 annualized value of 0.14 used in the base case. As discussed in section 3, the x parameters in wedge case 1 are chosen to match the mean of the equity price-dividend ratio and its unconditional Sharpe ratio, as in the base case. For wedge case 2, the x parameters are also chosen to get close to the unconditional volatility of the market equity excess return.

4.2.1 Aggregate Moments

The two columns of Table 3 labeled "Wedge" 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 unconditional mean of its price-dividend ratio, and Table 3 shows that they both get close to 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 with some success. 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 0.90 for wedge case 1 and still above 0.84 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 two columns labeled "Wedge" 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^2s at horizons of 1 and 10 years are still negligible. As in the base case, the two wedge cases do a poor job of reproducing the predictability of 1 or 10 year log dividend growth found in the data using the log price-dividend ratio, especially at the 10 year horizon, where the sign of the predictive coefficient is negative for the data and positive for the wedge cases.

4.2.3 Value vs Growth Portfolios

Table 5 reports results for the extreme growth and value deciles as well as HML and the two columns labeled "Wedge" report results for the 2 wedge cases. These "Wedge" columns show that the 2 wedge cases can also generate a value premium in both expected excess return and CAPM alpha. Moreover, the premia in expected 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 expected excess return spreads between the value and the growth portfolio are 4.05% per annum and 3.82% 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 2.71% 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 1.30% 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 expected excess returns, just as the base case delivering the same hump-shape is consistent with it also delivering a value premium in expected 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 expected excess return spread. Second, both wedge cases deliver excess return volatility that is higher for the extreme value decile than for the extreme growth decile. 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 extreme value decile in wedge case 2, whose excess return volatility is close to that in the data.

For the two wedge cases, the R^2s 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   LRR-Vol Case

This subsection discusses the results for the LRR-vol case and compares them to the results from the data and for the LW, base and two wedge cases. As in the two wedge cases, the aggregate

consumption process is allowed to differ from the market dividend process in the LRR-vol case. This process has mean and volatility that is slowly mean reverting 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 and the two wedge cases, the correlation between the change in log consumption and the price-of-risk variable x is set to -0.99 in the LRR-vol case, consistent with habit preferences. The persistence of the x process in the LRR-vol case is the same low annualized value of 0.14 used in the base and two wedge cases. As discussed in section 3, the x parameters are chosen to match the unconditional volatility of the market equity excess return as well as the mean of the equity price-dividend ratio and its unconditional Sharpe ratio, as in the second wedge case.

4.3.1 Aggregate Moments

The last column of Table 3 contains moments for aggregate equity return, aggregate equity price-dividend ratio and aggregate dividend growth for the LRR-vol case. The LRR-vol case is calibrated to match the unconditional Sharpe ratio of market equity, the volatility of its excess return and the unconditional mean of its price-dividend ratio, and Table 3 shows that it matches all three moments better than wedge case 2. As with wedge case 2, since the LRR-vol case does a good job matching the unconditional market equity Sharpe ratio and the volatility of the market equity excess return, it also does a good job matching the unconditional market equity excess return.

However, while wedge case 2 successfully matches the volatility of equity excess return, it severely understates the volatility of the equity price-dividend ratio. This is not a problem that the LRR-vol case suffers from. In fact, the volatility value obtained for the LRR-vol case of 0.480 is higher than the data value of 0.38. As would be expected given the high persistence of the σ_t process, the autocorrelation of the price-dividend ratio is higher than in either wedge case, taking a value that is even higher than the data value. One problem for wedge case 2 is the autocorrelation of the excess market equity return, since it's 0.03 in the data but -0.27 for wedge case 2. Allowing σ_t to be slowly mean reverting rather than constant attenuates this problem since the autocorrelation increases to -0.10 in the LRR-vol case.

4.3.2 Predictive Regressions

Table 4 reports results for the predictive regressions and the last column reports results for the LRR-vol case. We see that the LRR-vol case remedies the most glaring weakness of the base

and two wedge cases, namely their inability to generate the excess market return predictability observed in the data using market equity price-dividend ratio. The LRR-vol case is much closer to the data than the base or either wedge case because it is able to produce much more negative predictability coefficients and much larger R^2s for the regressions than those cases. For the 10 year return horizon regression, the R^2 is 0.31 for the data, 0.27 for the LRR-vol case, and less than 0.02 for these other cases. While the R^2s are still a little low relative to the data, they represent a significant improvement over the results for the base and two wedge cases. The implication is that allowing consumption growth to have volatility that is slowly mean reverting can help models with external habit preferences to generate the return predictability in the data. The question (answered in the next subsection) is whether the value premium survives the introduction of long run risk in consumption growth volatility. As in the other cases, the LRR-vol case does a poor job of reproducing the predictability of 1 or 10 year log dividend growth found in the data using the log price-dividend ratio, especially at the 10 year horizon.

4.3.3 Value vs Growth Portfolios

Table 5 reports results for the extreme growth and value deciles as well as HML and the last column reports results for the LRR-vol cases. This last column shows that the LRR-vol case can also generate a value premium in both expected excess return and CAPM alpha. Moreover, the premia in both expected excess return and CAPM alpha generated by the LRR-vol case is higher than those obtained for either wedge case. In the LRR-vol case, the expected excess return spreads between the value and the growth portfolio is 4.39% which is higher than the maximum value of 4.05% obtained by the other non-LW cases and actually quite close to the data value of 4.88% per annum when sorting on B/M. The CAPM-alpha spread for these two extreme book-to-market deciles is 4.03% per annum which is also closer to the 5.63% per annum in the data than the 2.71% per annum in wedge case 1.

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 pale solid line representing the LRR-vol case. The top graph of Figure 1 shows that the unconditional expected annual excess return on the zero-coupon equity is hump-shaped in maturity with the hump occurring at a maturity less than 10 years, same as for all the other cases except LW. Recall once more 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,

33

this hump-shape in the top graph of Figure 1 is consistent with the LRR-vol case delivering value premia in expected excess return, just as the base and wedge cases deliver a value premium in expected excess return by delivering the same hump-shape. Further, across the cases, the top graph of Figure 1 shows that the difference between the expected annual return on the zero-coupon market equity maturing in 3 and 13 years is largest for the LRR-vol case, which is consistent with the value premium in expected excess return being the largest for this case.

Moreover, $B_x(n)$ plotted in Figure 2 as a function of n is u-shaped for the LRR-vol case just as it is for the base and two wedge cases. The $B_x(n)$ curve for the LRR-vol case lies just above the $B_x(n)$ curve for the two wedge cases, but from equation (25) we see this generates a larger value premium in expected excess return, because $\bar{x}$ for the LRR-vol case is larger than for the two wedge cases. The shape of $B_\sigma(n)$, which is plotted in the bottom right-hand graph of Figure 2 as a function of n, does not matter for the value premium in expected excess return, because the LRR-vol case imposes $\rho_{d,w} = 0$, and equation (25) also shows that $\rho_{d,w} = 0$ means $B_\sigma(n)$ does not matter for unconditional mean excess market equity returns.

The LRR-vol case generates results that are similar to the results for the base and two wedge cases in a number of other respects. First, the LRR-vol case delivers a CAPM-alpha spread between the extreme value and the growth deciles that is smaller than the expected excess return spread, which is counterfactual. Second, the LRR-vol case delivers excess return volatility that is higher for the extreme value decile than for the extreme growth decile. However, the volatility numbers produced by the LRR-vol case for the two extreme deciles, though lower, are actually quite close to the data values, which is in contrast to the much lower numbers produced by the base and wedge cases.

Similar to the other non-LW cases, the R^2s of the CAPM market model regressions in the LRR-vol case are slightly higher than the data R^2 values for the growth decile, but much larger than the data R^2 values for the value decile. All the other cases produce R^2s for HML that are much too high, while the LRR-vol case produces a low R^2 that is quite close to the data values. As in all the data sorts and the other cases, the extreme growth decile has the lowest Sharpe ratio of the two extreme portfolios and HML in the LRR-vol case. Across the three portfolios, the extreme value decile has the highest Sharpe ratio for all the data sorts, while HML has the highest Sharpe ratio for the LRR-vol case, just like both of the wedge cases.

5 Conclusion

This paper finds that 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-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, recent micro evidence indicates that the persistence of the log surplus-consumption ratio is likely to be quite low. Brunnermeier and Nagel (2006) examine how risky asset holdings change in response to wealth shocks and reject a persistent habit specification, while Ravina (2007) shows that credit card purchases are more in line with a fast-moving habit than a slow-moving habit. Moreover, the high persistence assumed by Lettau-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 that is more in line with the micro evidence and 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. 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 also allows the model to more closely match the volatility of the equity return.

One important dimension of equity return behavior that low persistence has difficulty replicating when consumption growth is homoscedastic is the predictability of long-horizon equity return using the price-dividend ratio. However, in a setting in which the consumption and dividend processes are specified to be different, we are able to obtain long-horizon return predictability of a magnitude much closer to that in the data, and without destroying the value premium, by allowing the

conditional volatility of consumption growth to also be slowly mean reverting. Our results suggest that external-habit preferences and long-run risk in the mean and volatility of consumption growth may both play important roles in explaining the time-series and cross-sectional properties of equity returns and prices.

An interesting dimension not considered here is the extent to which long horizon returns on the extreme value and growth portfolios are predictable using aggregate equity's price-dividend ratio and other measures of the aggregate state. It would be interesting to see if fast moving habit and aggregate consumption specified to possess long-run risk in mean and volatility and calibrated to data can generate the low return predictability documented for these extreme portfolios by Roussanov (2011) and others. We leave an examination of this question to future research.

References

[1] Abel, Andrew B., 1990. Asset Prices under Habit Formation and Catching Up with the Joneses. *American Economic Review* 80, 38-42.

[2] Abel, Andrew B., 1999. Risk Premia and Term Premia in General Equilibrium. *Journal of Monetary Economics* 43, 3-33.

[3] Bansal, Ravi, and Amir Yaron, 2004. Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* 59, 1481-1509.

[4] Bansal, Ravi, Amir Yaron, Dana Kiku, 2007. Risks For the Long Run: Estimation and Inference. Working Paper, Duke University and University of Pennsylvania.

[5] Bekaert, Geert, and Eric Engstrom, 2009. Asset Return Dynamics under Bad Environment-Good Environment Fundamentals. Working Paper, Columbia University and the Federal Reserve.

[6] 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.

[7] 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.

[8] Constantinides, George M., 1990. Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy* 98, 519-43.

[9] Croce, M., Lettau, M and S.C. Ludvigson, 2010 Investor Information, Long-Run Risk, and the Duration of Risky Cash-Flows. Working paper, New York University.

[10] Dynan, Karen E., 2000. Habit Formation in Consumer Preferences: Evidence from Panel Data. *American Economic Review* 90, 391-406.

[11] Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008. Consumption strikes back? Measuring long-run risk. *Journal of Political Economy* 116, 260-302.

[12] Kiku, Dana, 2006. Is the value premium a puzzle? Working Paper, University of Pennsylvania.

[13] 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.

[14] 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.

[15] Ravina, Enrichetta, 2007. Habit Persistence and Keeping Up with the Jones: Evidence from Micro Data. Working Paper, New York University.

[16] Roussanov, Nikolai, 2011. Composition of Wealth, Conditioning Information, and the Cross-Section of Stock Returns. Working Paper, University of Pennsylvania.

[17] 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.

[18] Stambaugh, Robert, 1997. Analyzing Investments Whose Histories Differ in Length. *Journal of Financial Economics* 45 285331.

[19] 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, wedge and LRR-vol cases (final four columns). All parameters are as defined in section 2. The model is quarterly, but the mean of the log dividend growth g^m and the log riskfree rate r^f are 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		LRR-Vol
				1	2	
g^m	annual	2.28%	2.28%	2.28%	2.28%	2.28%
$r^f = -a$	annual	1.93%	1.93%	1.93%	1.93%	1.93%
b or b^m		0	0	0	0	0
$\bar{x}$	quarterly	0.625	0.25	0.3493	0.28	0.365
ϕ_z	annual	0.91	0.91	0.91	0.91	0.91
ϕ_x	annual	0.865	0.14	0.14	0.14	0.14
σ_m	quarterly	0.0724	0.0724	0.0724	0.0724	0.0724
σ_d	quarterly	0.0724	0.0724	0.0160	0.0160	0.0164
σ_z	quarterly	0.00165	0.00165	0.00165	0.00165	0.00165
σ_x	quarterly	0.1225	0.16	0.3049	0.3305	0.29
σ_u	quarterly	0	0	0.0435	0.0435	0.037
$\rho_{m,z} = \rho_{d,z}$	quarterly	-0.82	-0.82	-0.82	-0.82	-0.82
$\rho_{d,x}$	quarterly	0	-0.99	-0.99	-0.99	-0.99
$\rho_{z,x}$	quarterly	0	0.81	0.81	0.81	0.81
$\rho_{d,u}$	quarterly	-	-	-0.30	-0.30	-0.30
$\rho_{z,u}$	quarterly	-	-	0.0037	0.0037	0.15
$\rho_{x,u}$	quarterly	-	-	0.30	0.30	0.30
δ^m	quarterly	1	1	4.54	4.54	4.54
$\bar{\sigma}$	quarterly	1	1	1	1	0.918
ϕ_σ	quarterly	-	-	-	-	0.994
σ_w	quarterly	0	0	0	0	0.037

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 (the first column) and our cases (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	Habit Contribution (%)	
(yrs)	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 four report simulated moments for the base, wedge and LRR-vol 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 a minimum of 4 million quarters, until the moments have converged: i.e. are sufficiently close to those when an additional 10% of simulated quarters are also included. Returns, dividends, and price-dividend ratios are aggregated to annual frequencies: so $P^m/D^m = P^m_t / \sum_{\tau=0}^{3} D^m_{t-\tau}$ 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		LRR-Vol
				1	2	
$E[P^m/D^m]$	25.55	20.04	30.94	31.09	23.06	25.08
$\sigma[p^m - d^m]$	0.38	0.382	0.261	0.279	0.260	0.480
$AC[p^m - d^m]$	0.87	0.883	0.895	0.900	0.843	0.946
$E[R^m - R^f]$	6.33	8.116	4.497	4.512	6.180	6.835
$\sigma[R^m - R^f]$	19.41	19.44	10.69	10.71	14.68	16.55
$AC[R^m - R^f]$	0.03	-0.04	-0.13	-0.14	-0.27	-0.10
$Sharpe^m$	0.33	0.418	0.421	0.421	0.421	0.413
$AC[\Delta d^m]$	-0.09	-0.03	-0.03	-0.03	-0.03	-0.02
$\sigma[\Delta d^m]$	14.48	14.42	14.38	14.38	14.39	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 four report results for the base, wedge and LRR-vol 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 a minimum of 4 million quarters, until the moments have converged: i.e. are sufficiently close to those when an additional 10% of simulated quarters are also included. 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 | | LRR-Vol |
					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	-0.01	-0.04	-0.06
	10	-1.09	-0.68	-0.05	-0.11	-0.14	-0.49
R^2	1	0.05	0.07	0.00	0.00	0.01	0.03
	10	0.31	0.30	0.00	0.01	0.02	0.27
$\sum_{i=1}^{H}\Delta d_{t+i}^m = \beta_0 + \beta_1(p_t^m - d_t^m) + \epsilon_t$							
β_1	1	0.02	0.05	0.11	0.10	0.10	0.03
	10	-0.31	0.32	0.73	0.68	0.68	0.21
R^2	1	-0.01	0.02	0.04	0.04	0.03	0.01
	10	0.05	0.09	0.22	0.22	0.19	0.06
$\sum_{i=1}^{H}\Delta d_{t+i}^m = \beta_0 + \beta_1 z_t^m + \epsilon_t$							
β_1	1	0.10	3.89	3.83	3.87	3.86	3.88
	10	0.68	26.38	26.02	26.15	26.26	26.42
R^2	1	0.03	0.04	0.04	0.04	0.04	0.04
	10	0.25	0.25	0.25	0.25	0.25	0.24

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 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 final four report results for the base, wedge and LRR-vol 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 a minimum of 4 million quarters, until the moments have converged: i.e. are sufficiently close to those when an additional 10% of simulated quarters are also included. $Sharpe^i$ is the unconditional Sharpe ratio for portfolio i's annual return while R_i^2 is the regression R^2 for portfolio i.

|	Portfolio	E/P	C/P	B/M	LW	Base	Wedge 1	Wedge 2	LRR-Vol
Expected Excess Return: $E[R^i] - R^f$									
$E[R^i] - R^f$	1	4.71	5.05	5.67	5.28	3.38	2.02	3.24	3.87
	10	12.95	11.81	10.55	10.81	5.30	6.07	7.06	8.27
	HML	8.25	6.77	4.88	5.53	1.91	4.05	3.82	4.39
$\sigma[R^i - R^f]$	1	19.35	18.99	17.77	19.55	9.90	9.23	10.51	16.18
	10	18.11	17.24	18.46	17.67	11.46	11.99	15.72	16.54
	HML	15.40	14.57	15.15	8.84	4.47	6.40	8.24	8.62
$Sharpe^i$	1	0.24	0.27	0.32	0.27	0.34	0.22	0.31	0.24
	10	0.72	0.69	0.57	0.61	0.46	0.51	0.45	0.50
	HML	0.54	0.46	0.32	0.63	0.43	0.63	0.46	0.51
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	-2.72	-0.66	-1.66	-0.82	-2.45
	10	6.22	5.34	3.97	3.72	0.54	1.10	0.48	1.59
	HML	9.31	8.04	5.63	6.45	1.20	2.71	1.30	4.03
β_i	1	1.18	1.17	1.11	0.99	0.90	0.80	0.66	0.92
	10	1.02	0.98	1.00	0.87	1.06	1.10	1.07	0.98
	HML	-0.16	-0.19	-0.11	-0.11	0.16	0.30	0.41	0.05
R_i^2	1	0.83	0.85	0.87	0.96	0.95	0.87	0.84	0.90
	10	0.71	0.72	0.65	0.92	0.97	0.97	0.99	0.96
	HML	0.02	0.04	0.01	0.06	0.14	0.25	0.53	0.01
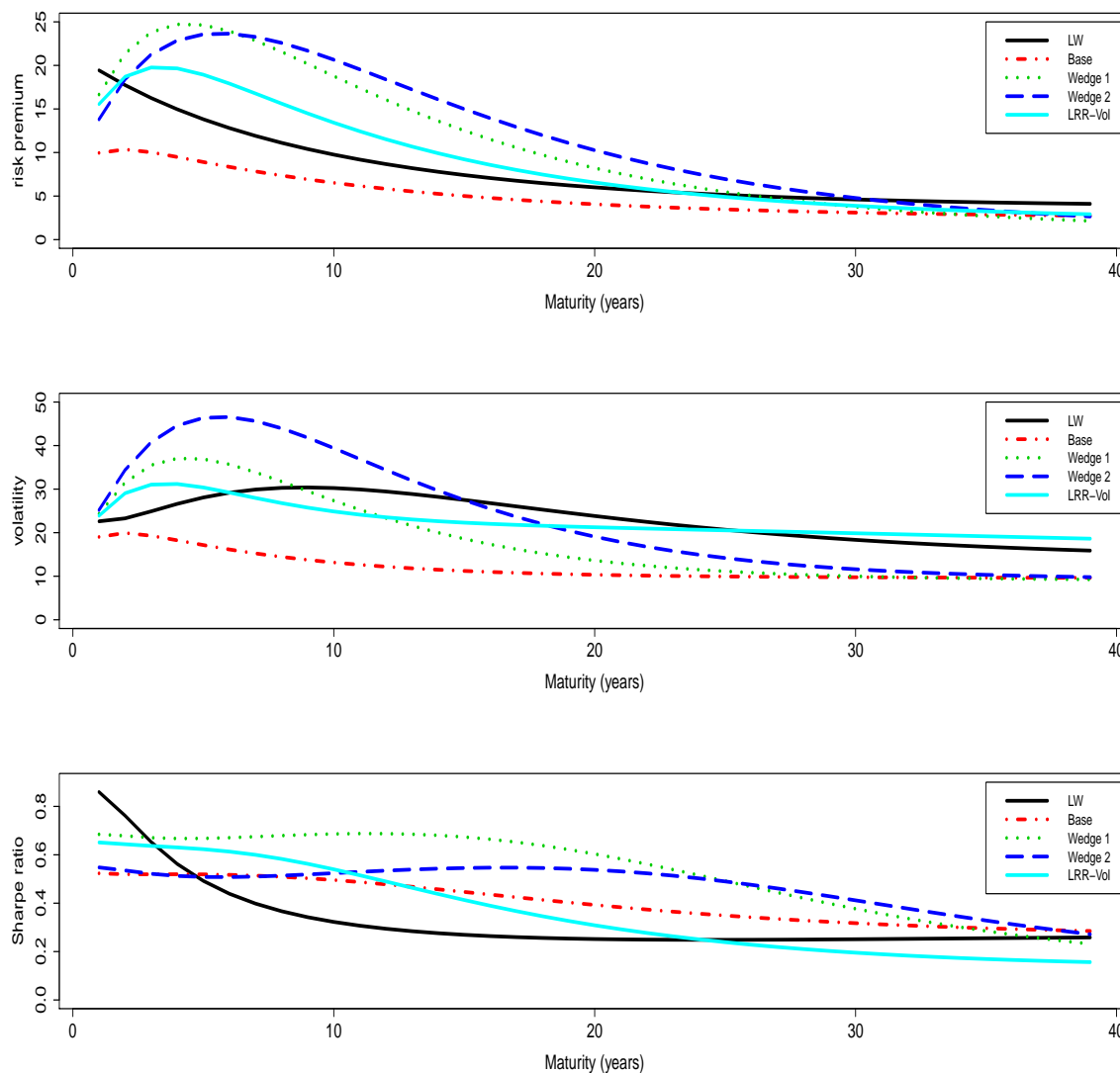
Figure 1: **Returns on Zero-coupon Equity: Base vs LW Case**. The figure plots the unconditional expected annual return, the unconditional volatility of annual return and the unconditional Sharpe ratio for annual return in the top, middle and bottom graphs respectively. In each graph, the solid black line is for the LW case, the dot-dashed line is for the base case, the dotted line is for wedge case 1, the dashed line is for wedge case 2, and the solid pale blue line is for the LRR-vol 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. The models are simulated at a quarterly frequency for a minimum of 4 million quarters, until the moments have converged: i.e. are sufficiently close to those when an additional 10% of simulated quarters are also included.
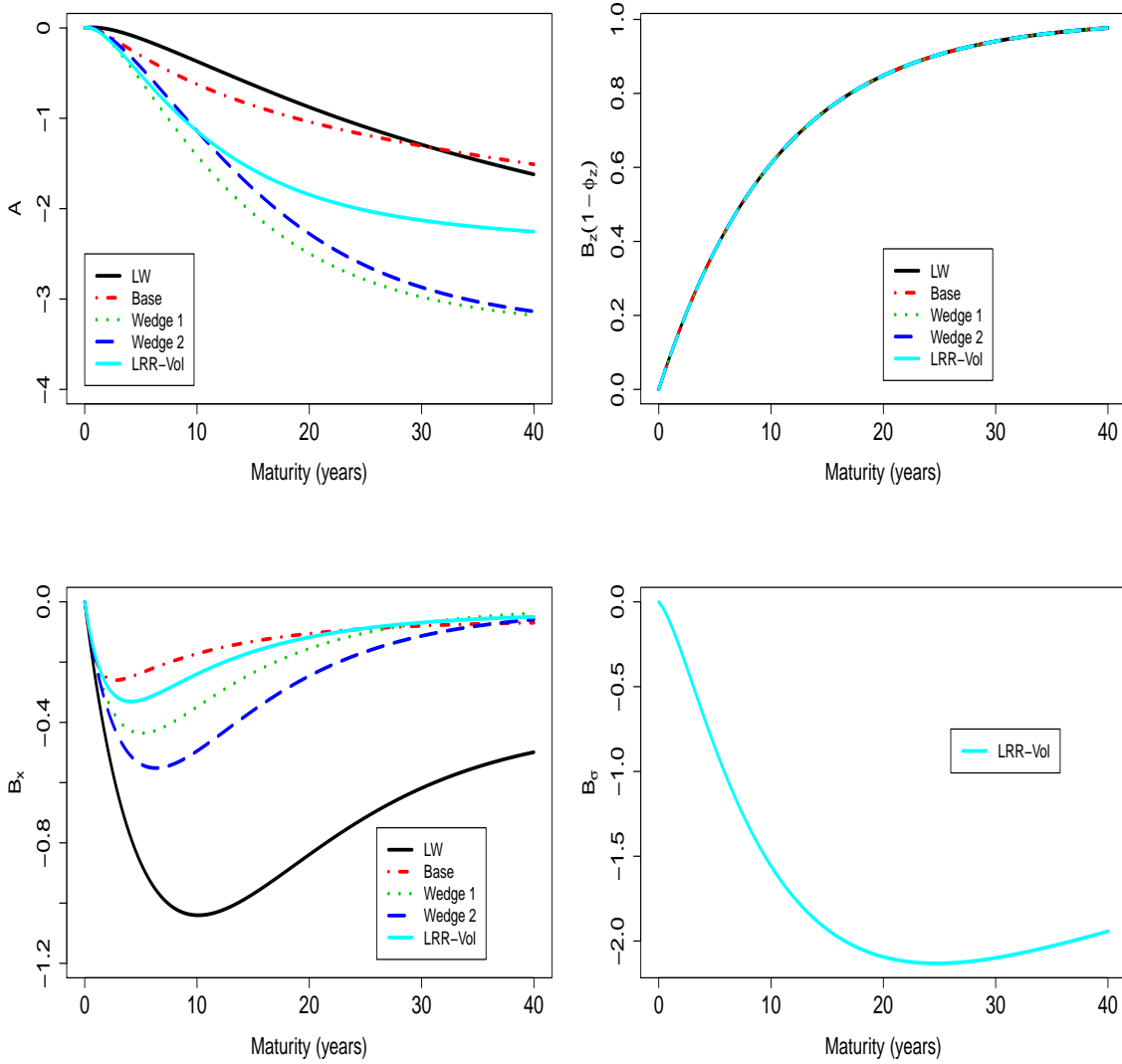
Figure 2: **Zero-coupon Aggregate Equity Log Price-dividend Ratio Coefficients** $A(n)$, $B_z(n)$ **and** $B_x(n)$: **Ours vs LW Case**. The figure plots $A(n)$, $B_z(n)(1 - \phi_z)$ and $B_x(n)$ in the top left, top right, and bottom left graphs respectively. In each graph, the solid black line is for the LW case, the dot-dashed line is for the base case, the dotted line is for wedge case 1, the dashed line is for wedge case 2, and the solid pale blue line is for the LRR-vol 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) of section 2 for the zero-coupon aggregate equity price-dividend ratio paying out in n periods. $B_z(n)$ is multiplied by $(1 - \phi_z)$.