JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 58, No. 7, Nov. 2023, pp. 3153–3194

doi:10.1017/S0022109023000212

Fast-Moving Habit: Implications for Equity Returns

Anthony W. Lynch
New York University Stern School of Business and NBER, Asset Pricing Group
alynch@stern.nyu.edu (corresponding author)

Oliver Randall
University of Melbourne, Faculty of Business and Economics
oliver.randall@unimelb.edu.au

Abstract

We find that the Campbell–Cochrane external-habit model can generate a value premium if the persistence of the consumption surplus is sufficiently low. Such low persistence is supported by micro evidence on consumption. If the mean and conditional volatility of consumption growth are highly persistent, as in the Bansal–Yaron long-run risk model, then fast-moving habit can also generate, without eroding the value premium: i) empirically sensible long horizon return predictability; and ii) a price–dividend ratio for market equity that exhibits the high autocorrelation found in the data. Fast-moving habit also delivers several empirical properties of market-dividend strips.

I. Introduction

Several models have been developed that attempt to explain the moments of the market equity price–dividend ratio, the market equity return, and the risk-free rate. The two leading models are i) models with habit preferences and ii) long-run risk models with Epstein–Zin–Weil preferences. Early habit articles include Sundaresan (1989), Abel (1990), and Constantinides (1990). Assuming independent and identically distributed (IID) consumption and a representative agent with external habit preferences, Campbell and Cochrane (CC) (1999) allow the conditional volatility of the log surplus (consumption in excess of habit, scaled by consumption) to vary inversely with the log surplus in such a way that the risk-free rate is constant, and price-of-risk shocks are close to perfectly negatively

The authors thank an anonymous referee, Greg Duffee, Jarrad Harford (the editor), Stijn Van Nieuwerburgh, and participants at two NYU seminars, the Wharton Brown Bag Macro Seminar, a session of the 2011 American Finance Association Meetings, and seminars at Australian National University, Erasmus University, Tilburg University, the University of Washington, the University of Western Australia, and the University of Melbourne for helpful comments and suggestions. The authors also thank Ralph Koijen for generously providing S&P 500-strip portfolios data used in van Binsbergen, Brandt, and Koijen (2012), and Matthieu Gomez for making publicly available code that solves the model in Garleanu and Panageas (2015). All remaining errors are of course the authors' responsibility. This article was previously circulated under the title "Why Surplus Consumption in the Habit Model May be Less Persistent than You Think."

correlated with consumption growth shocks. CC specify the log surplus to be a highly persistent process so that the market price–dividend ratio is highly persistent and able to forecast long-horizon market returns, as in U.S. data. However, the high persistence renders CC's model unable to generate the value premium seen in the data. In Bansal and Yaron's (BY) (2004) long-run risk model, the market's risk premium and return volatility, the autocorrelation of the market price–dividend ratio, and the average risk-free rate are all close to the data. Neither CC nor BY can generate all the empirical properties of market-dividend strips.

Lettau and Wachter (LW) (2007) 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 i) the state variable driving the price of risk is highly persistent, ii) the mean of consumption growth is a slowly mean-reverting process as in BY, and iii) the cash flows of value stocks have shorter durations than those of growth stocks as documented in Dechow, Sloan, and Soliman (2004) and Da (2009) (cf. Chen (2017), who argues the difference may be smaller than previously documented). LW find that large negative correlation between the shocks to the price of risk and consumption growth generates a growth premium for expected excess return and CAPM alpha (expected excess return and CAPM alpha are higher for the extreme growth than the extreme value portfolio). This finding is in contrast to the value premium for expected excess return and CAPM alpha (expected excess return and CAPM alpha are higher for the extreme value than the extreme growth portfolio) found in U.S. data. To produce a value premium, they set this correlation to 0. Their finding raises the question whether habit preferences can generate a value premium as in U.S. data.

A highly persistent log surplus implies a very slow-moving habit. When the log surplus is as persistent as in CC and LW, the two most recent years of consumption contribute less than 26% to the agent's habit level (leaving all earlier consumption to contribute more than 74%), which is much too low to be consistent with the micro evidence. For this reason, our article examines how a less persistent price of risk, which would be implied by a less persistent log surplus, affects the moments of the market price–dividend ratio and return, the differentials in expected return and in CAPM alpha between value and growth stocks, and the properties of returns on market-dividend strips.

Matching the data Sharpe ratio and expected price–dividend ratio for market equity at least as well as LW, we find that when the persistence of the price of risk is low, a large negative correlation between the shock to the price of risk 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. We also find that, so long as the conditional mean of log consumption growth is allowed to be slowly mean-reverting as parameterized by LW and BY based on U.S. data, the market price–dividend ratio exhibits high first-order autocorrelation comparable to that in U.S. data even when the persistence of the price of risk is low. This is because the expression for the price–dividend ratio of market-dividend strips suggests that the autocorrelation of the aggregate market's price–dividend ratio is approximately a weighted average of the autocorrelations of the price-of-risk and the conditional mean of log consumption growth, and the latter is more slowly mean-reverting than the market price–dividend ratio in the data.

Why can low habit persistence allow habit models to deliver a value premium in expected excess return and in CAPM alpha? In LW and our model, a positive shock to the price of risk causes a negative shock to the price–dividend ratios for dividend strips, which in turn causes negative shocks to their returns. Because the price-of-risk shock is negatively correlated with the consumption shock in habit models, the shocks to dividend-strip returns caused by the price-of-risk shock are positively correlated with the consumption shock, and so increase risk premia on the dividend strips. This increase is hump-shaped in maturity. When habit persistence is very high as in CC, the hump occurs far in the future, causing a growth premium in expected excess return since empirical evidence suggests that value stocks have shorter duration cash flows than growth stocks. Lowering habit persistence causes the hump to occur earlier, which is the key piece of intuition to understand why low habit persistence can deliver a value premium in expected excess return: if the hump occurs early enough, the risk premia for the shorter-maturity cash flows of value stocks become sufficiently high relative to the risk premia for the longer-maturity cash flows of growth stocks to generate a value premium in expected excess return. Because the CAPM betas for the extreme value and growth deciles are very similar, the same intuition also delivers a value premium in CAPM alpha.

Our baseline specification (Base case) follows LW and assumes that the consumption process and the market dividend process are the same by calibrating both to the market dividend process for U.S. stocks. It is unable to generate the market return volatility found in the data. We also consider a specification (Wedge case) that allows the consumption process to differ from the dividend process, by calibrating the consumption process to data and leaving the dividend process the same. This specification moves market return volatility much closer to the data, and, relative to the Base case, is able to generate a value premium in expected excess return that is considerably larger, and a value premium in CAPM alpha that is similar in magnitude.

Unfortunately, in contrast to LW, the Base and Wedge cases, with their implicit assumption that log consumption growth is homoscedastic, are unable to replicate the strong predictability of long-horizon market returns found in the data using the market price–dividend ratio as the predictor. To understand this result better, let x be the price of risk and ϕ_x be its persistence level, and consider a setting in which log consumption growth, Δd, is a homoscedastic process equal to log market dividend growth, Δd^m, as in LW and the Base case. Comparative static analysis confirms that the ϕ_x that delivers a 10-year return predictability R^2 equal to the 30% data value is declining as the conditional correlation between Δd and x, $\rho_{d,x}$, becomes more negative, and this decline in ϕ_x is typically associated with a declining value premium in raw return. As a consequence, it is extremely difficult in this setting to simultaneously obtain a value premium in raw return and market return predictability of the magnitude observed in the data when ϕ_x is low and $\rho_{d,x}$ is close to -1.

Consequently, we allow Δd to not only differ from Δd^m, but also, in the spirit of BY's long-run risk in volatility model, exhibit very persistent AR(1) volatility. This specification (LRR-Vol case) delivers long-horizon return predictability of a magnitude much closer to that in the data, because future expected excess market returns depend on future conditional consumption growth volatility. Consistent with this intuition, we find that the long-horizon return predictability in the LRR-Vol case is completely driven by the predictive ability of current conditional consumption growth volatility. The LRR-Vol case also delivers value premia in expected excess return and CAPM alpha that are both larger than for the Base or Wedge cases, though still lower than for LW. Moreover, the value premium in expected excess return is closer to the one found in the data using a BM sort than the one obtained by LW. This specification also generates market return volatility that is within 15% of the data value that LW matches almost exactly. Long-run risk in consumption volatility also helps match the data along several other dimensions.

Our fast-moving habit model also delivers many of the empirical results for market-dividend strips that have recently been documented by van Binsbergen, Brandt, and Koijen (BBK) (2012) and van Binsbergen and Koijen (BK) (2017). BBK find that the means, volatilities, and Sharpe ratios for monthly returns on two portfolios of S&P 500 dividend strips with average maturities around 1 year are all larger than for the S&P 500 index itself. BBK also find higher R^2s for the regressions that forecast the monthly returns on their two strip portfolios using their own price–dividend ratios than for the regression that forecasts the monthly return on the S&P 500 using its price–dividend ratio. Our three fast-moving habit models deliver all these results for quarterly returns, while in contrast, LW only delivers three of the four results (not the volatility result) and BBK find that neither the external habit model with slow-moving habit nor the long-run risk model are able to produce any of the first three results. When market model regressions are run for the CAPM and a 2-factor model with market excess return and Fama–French's HML zero-cost portfolio as the factors, the LRR-Vol model is able to get quite close to the CAPM alphas for short-maturity market-dividend strips in the data, and closer than LW to both the CAPM and 2-factor alphas for these strips in the data. BK present some new empirical evidence about the expected excess returns on market-dividend strips which they obtain using monthly returns on S&P 500 dividend-strip futures. They report that the expected monthly dividend-strip spot return in excess of the market return is increasing in strip maturity going from 1-year strips to 5-year strips. This empirical result is much more consistent with the dividend-strip expected excess return curve for the LRR-Vol case, which is hump-shaped, than the curve for LW, which is downward-sloping.

In summary, 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 log surplus consumption ratio is allowed to be low rather than set to a value close to one. Simultaneously allowing the conditional mean of 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. Our model with all three features can also deliver some, but not all, of the empirical results for market-dividend strips that have recently been documented.

The question arises as to why the persistence of the log surplus might be low in the representative-agent external-habit model. There are two potential answers. First, micro evidence on habit preferences suggests agents do not have slow-moving habit in the spirit of CC, which could be a micro-foundation for a representative-agent external habit model with fast-moving habit, so long as agents exhibit fast-moving habit rather than no habit. Because slow-moving habit is like a subsistence level, it implies that risk-asset holding as a fraction of financial wealth increases in response to wealth increases. In contrast, Brunnermeier and Nagel (2008) do not find such a positive relation empirically, or any other evidence supportive of slow-moving habit. However, their evidence does not rule out fast-moving habit since the relation need not be positive if habit is moving sufficiently fast (see Appendix A.1 in Brunnermeier and Nagel (2006)). Using quarterly credit card purchases as a measure of consumption, Ravina (2019) estimates a log-linearization of the Euler equation for households with habit preferences and finds that the coefficient of lagged own consumption (internal habit) is 0.5 while the coefficient on current household city consumption (external habit) is 0.29. These coefficients are too high to be consistent with slow-moving habit which implies 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 0. However, Ravina obtains similar values to Dynan once she omits household-specific financial control variables not available in Dynan's data set.

Second, economies inhabited by agents with Epstein–Zin–Weil preferences and heterogeneous risk aversion exhibit stochastic aggregate risk aversion just like economies with a representative agent that has external habit preferences (e.g., Chabakauri (2013), (2015), Bhamra and Uppal (2014), and Garleanu and Panageas (2015)). So heterogeneity in risk aversion as a micro-foundation for external habit models like CC can be another micro-foundation for the habit being fast-moving, so long as the consumption shares of agents exhibit low persistence. In particular, Garleanu and Panageas consider an overlapping-generations economy with two types of Epstein–Zin–Weil agents who differ in terms of risk aversion and possibly intertemporal elasticity of substitution. Because their model delivers a nondegenerate stationary equilibrium, the consumption shares have stationary distributions. While these shares are highly persistent in their baseline calibration, the persistence can be reduced somewhat by simultaneously increasing the population share of the less risk-averse agents, decreasing the fraction of output paid out as earnings, and increasing the risk-aversion heterogeneity of the agents (i.e., by simultaneously decreasing the risk aversion of the less risk-averse agents and increasing the risk aversion of the more risk-averse agents). However, it is necessary to also increase the birth/death rate substantially to lower the consumption-share persistence enough to match the value assumed for our log surplus.

While our analysis shows that increasing risk-aversion heterogeneity can reduce the persistence of consumption to a level more in line with the micro data, Gomez (2022) highlights that it also increases wealth inequality, since the more risk-tolerant agents hold riskier assets which tend to outperform the safer portfolios held by the less risk-tolerant agents. He shows that his model, adapted from Garleanu and Panageas, cannot match the excess volatility of asset prices unless preference heterogeneity is increased so much that the right tail of the wealth distribution is much thicker than in the data. Gomez resolves this tension by augmenting his model with purely redistributive wealth shocks between the two investor types and "live-fast-die-young" risk-tolerant investors who transition to more risk-averse investors after a short period of time. Interestingly, both these add-ons are similar mechanisms to increasing the birth/death rate in Garleanu and Panageas, which we find to be the most effective way to reduce the persistence of the consumption shares. While the birth/death rate needed to reduce the persistence to that assumed for our log surplus is unrealistically high, intuition suggests that the persistence could be further reduced by extending their model to allow for "live-fast-die-young" risk-tolerant investors, purely redistributive wealth shocks between the two investor types, and idiosyncratic wealth and labor-income shocks.

Long-run risk models have had some success generating a value premium in excess return and CAPM alpha. Calibrating the BY model by targeting aggregate moments, Kiku (2007) and Bansal, Kiku, and Yaron (2016) find the value portfolio to be more sensitive to long-run consumption risks than the growth portfolio, which generates a value premium in raw returns and CAPM alpha. Using a Bayesian approach to estimate the long-run risk model, Schorfheide, Song, and Yaron (2018) find the data moments related to the risk premium (the market price–dividend ratio) to be roughly at the center (in the right tail) of the posterior distributions for the model.

There are several other related articles. Consistent with LW, Santos and Veronesi (2010) find that habit preferences deliver a growth premium rather than a value premium when value firm cash flows have shorter durations than growth firm cash flows, unless greater cross-firm cash flow heterogeneity than found in the data is introduced. Bekaert and Engstrom (2017) consider an extension to the CC external-habit model in which the log surplus continues to be very persistent, but 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. Croce, Lettau, and Ludvigson (2015) 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. Finally, Hansen, Heaton, and Li (2008) report that the cash flows of value but not growth stocks exhibit positive comovement with macroeconomic risks in the long run.

Section II describes the model, while Section III presents the calibration details. Results are in Section IV, and Section V concludes.

II. The Model

We consider a model that generalizes LW along two dimensions: i) the consumption and dividend processes are allowed to differ; and ii) the conditional volatility of log consumption growth is assumed to follow a highly persistent AR(1) process. All proofs are in the appendices or Supplementary Material.

A. Specification

The model has five shocks which are assumed to be multivariate normal and independent over time: a shock to consumption growth ε^d_{t+1}, a separate shock to

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

dividend growth ε^u_{t+1}, a shock to conditional expected dividend growth ε^z_{t+1}, a shock to the conditional volatility of consumption growth ε^w_{t+1}, and a shock to the price of risk variable ε^x_{t+1}. Define $\sigma_i^2 \equiv \sigma^2[\varepsilon^i]$, $\sigma_{i,j} \equiv \sigma[\varepsilon^i, \varepsilon^j]$, and $\rho_{i,j} \equiv \rho[\varepsilon^i, \varepsilon^j]$ for $i,j = d,z,x,u,w$.

Let D^m_t denote market dividends at time t, and define $d^m_t \equiv \log\left(D^m_t\right)$, which evolves as follows:

$$\Delta d^m_{t+1} = g^m + z^m_t + \varepsilon^m_{t+1} \tag{1}$$

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

$$z^m_{t+1} = \phi_z z^m_t + \varepsilon^z_{t+1} \tag{2}$$

with $0 \leq \phi_z < 1$. Let D_t denote consumption at time t, and setting $g \equiv g^m/\delta^m$ and $z_t \equiv z^m_t/\delta^m$, define $d_t \equiv \log(D_t)$, which evolves as follows:

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

where

$$\sigma_{t+1} = \overline{\sigma} + \phi_\sigma(\sigma_t - \overline{\sigma}) + \varepsilon^w_{t+1} \tag{4}$$

and $\varepsilon^m_{t+1} \equiv \delta^m \varepsilon^d_{t+1} + \varepsilon^u_{t+1}$, with ε^w_{t+1} and ε^u_{t+1} uncorrelated with each other and the other shocks. This specification allows separation between the market dividend and consumption, with log dividend growth a levered version of log consumption growth as in Abel (1999) and δ^m the leverage parameter. Finally, both the conditional mean, $g + z_t$, and the conditional volatility, σ_t, of log consumption growth are highly persistent AR(1) processes. This consumption specification is closely related to BY, who specify that the variance, not the volatility, is an AR(1).[1]

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

$$x_{t+1} = (1 - \phi_x)\overline{x} + \phi_x x_t + \varepsilon^x_{t+1} \tag{5}$$

with $0 \leq \phi_x < 1$.

We specify that only the shock to consumption growth is priced, and that the stochastic discount factor takes the form

$$M_{t+1} = \exp\left\{ a + bz_t - \frac{1}{2}(x_t\sigma_t)^2 - x_t\sigma_t \frac{\varepsilon^d_{t+1}}{\sigma_d} \right\}. \tag{6}$$

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

$$x_t\sigma_t \approx \overline{x\sigma} + \overline{x}(\sigma_t - \overline{\sigma}) + \overline{\sigma}(x_t - \overline{x}), \tag{7}$$

[1]We also calibrated a model 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.

$$(8) \qquad = \bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma} x_t,$$

which gives us the following stochastic discount factor:

$$(9) \qquad M_{t+1} = \exp\left\{ a + bz_t - \frac{1}{2}\left(\bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma} x_t\right)^2 - \left(\bar{x}(\sigma_t - \bar{\sigma}) + \bar{\sigma} x_t\right)\frac{\varepsilon^d_{t+1}}{\sigma_d} \right\}.$$

The conditional log-normality of M_{t+1} implies that the log of the risk-free rate from time t to $t+1$ is given by

$$(10) \qquad r^f_t \equiv -a - bz_t.$$

If $b \neq 0$, the riskless rate is time varying. Since the most relevant articles to ours, LW and CC, both assume that the risk-free rate is constant, we also assume that $b = 0$, so we can directly compare our results to theirs.

B. Price–Dividend Ratio and Expected Return for Market-Dividend Strips and Market Equity

Let $P^m_{n,t}$ be the time-t price of a market-dividend strip, paying off in n periods. Following LW, it can be shown that $P^m_{n,t}$ takes the following recursive form for this model:

$$(11) \qquad \frac{P^m_{n,t}}{D^m_t} = F\left(x_t, \sigma_t, z^m_t, n\right) = \exp\left\{A(n) + B_x(n)x_t + B_\sigma(n)(\sigma_t - \bar{\sigma}) + B_z(n) z^m_t\right\}.$$

The log risk premium on a market-dividend strip can be shown to depend on B_x, B_σ, B_z, x_t, σ_t, the variance of the consumption shock, and its covariances with the other shocks:

$$\begin{aligned}(12) \qquad \log\left(E_t\left[\frac{R^m_{n,t+1}}{R^f_t}\right]\right) &= E_t\left[r^m_{n,t+1} - r^f_t\right] + \frac{1}{2}\sigma^2_t\left[r^m_{n,t+1}\right] \\ &= \left(\delta^m \sigma^2_d + \sigma_{d,u} + B_x(n-1)\sigma_{d,x} + B_\sigma(n-1)\sigma_{d,w} + B_z(n-1)\sigma_{d,z}\right) \\ &\quad \left(\frac{\bar{\sigma}}{\sigma_d} x_t + \frac{\bar{x}}{\sigma_d}(\sigma_t - \bar{\sigma})\right).\end{aligned}$$

Market equity is the claim to all future market dividends. By the law of one price, a claim to market equity is equal in price to the sum of the prices of market-dividend strips over all future horizons. Dividends are paid at a quarterly frequency, so we can calculate the annual market price–dividend ratio as follows:

$$(13) \qquad \frac{P^m_t}{\sum_{\tau=0}^{3} D^m_{t-\tau}} = \sum_{n=1}^{\infty} \frac{P^m_{n,t}}{\sum_{\tau=0}^{3} D^m_{t-\tau}}.$$

Quarterly market returns can be calculated as a function of the quarterly market price–dividend ratio and dividend growth:

$$R_{t+1}^m = \frac{P_{t+1}^m + D_{t+1}^m}{P_t^m} \tag{14}$$

$$= \left(\frac{P_{t+1}^m / D_{t+1}^m + 1}{P_t^m / D_t^m}\right)\left(\frac{D_{t+1}^m}{D_t^m}\right). \tag{15}$$

We simulate at a quarterly frequency, and we calculate annual returns by compounding quarterly returns, which is equivalent to reinvesting dividends at the end of each quarter.[2]

C. Relation to the CC Model

CC assume that a representative agent maximizes the utility function:

$$E\left[\sum_{t=0}^{\infty} \delta^t \frac{(D_t - H_t)^{1-\gamma} - 1}{1-\gamma}\right], \tag{16}$$

where H_t is the level of external habit at time t and δ is the subjective discount factor. Suppose the representative agent again maximizes the habit specification in equation (16) and both the conditional mean and volatility of consumption growth are slowly mean-reverting, as in equations (2)–(4). Define the log of the surplus-consumption ratio at time t to be $s_t \equiv \log\left(\frac{D_t - H_t}{D_t}\right)$ as in CC. 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 the consumption surplus evolves as follows:

$$s_{t+1} = (1-\phi_s)\bar{s} + \phi_s s_t + \lambda(\bar{s}) z_t + \lambda(s_t)\sigma_t \varepsilon_{t+1}^d, \tag{17}$$

where $\varepsilon^d \sim N\left(0, \sigma_d^2\right)$ and $\lambda(.)$ is the same sensitivity function as used by CC.[3]

This specification for the log consumption surplus in equation (17) implies that the risk-free rate does not depend on s_t when $\sigma_t = \bar{\sigma}$. It also implies the same desirable properties for the habit process that CC's setup delivers: at the consumption surplus's steady state, log habit is predetermined only by an exponentially weighted sum of past lagged log consumption (see Section II.D); and, habit next period moves positively with consumption next period irrespective of the consumption surplus this period.

This specification implies the following stochastic discount factor:

[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 article.

[3]CC 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 $\bar{s} \equiv \log\left(\bar{S}\right)$ and $s_{\max} = \bar{s} + \frac{1}{2}\left(1 - \left(\bar{S}\right)^2\right)$. We set $\bar{S} \equiv \bar{\sigma}\sigma_d\sqrt{\frac{\gamma}{1-\phi_s}}$.

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

Matching coefficients in the stochastic discount factor we see that the risk-free rate is affine in z_t and σ_t^2. So our model can be used to approximate CC with persistent conditional mean and volatility of consumption growth by first using the same approximation in equation (7) applied to $\gamma(1+\lambda(s_t))$ and σ_t, and then setting $a = \log(\delta) - \gamma g + \frac{\gamma(1-\phi_s)}{2}$, $b = -\gamma(1+\lambda(\bar{s}))$, $\delta^m = 1$, $\sigma_u = 0$, $x_t = \gamma\sigma_d(1+\lambda(s_t))$, and σ_t equal to itself. Our model approximates $\gamma\sigma_d(1+\lambda(s_t))$, a heteroscedastic AR(1) process, with x_t, a homoscedastic AR(1) process. So as long as the sensitivity function is rarely 0, $\rho_{d,x} \approx -1$ and $\phi_x \approx \phi_s$.[4]

D. 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 Section II.C. Defining $h_t \equiv \log(H_t)$, we can apply a log-linear approximation to the definition of s_t:

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

Substituting this into the law of motion for s described in equation (17), and utilizing the imposed restriction that h_{t+1} is predetermined at the steady state, we can show that:

$$h_{t+1} \approx \overline{h-d} + (1-\phi_s)\sum_{j=0}^{\infty}(\phi_s)^j d_{t-j} + \frac{g}{1-\phi_s}, \tag{18}$$

which is the same expression derived in an earlier version of CC for IID consumption growth. 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 (17).

In equation (18), the coefficient on log lagged consumption, d_t, is $(1-\phi_s)$: so when ϕ_s is close to 1, as in CC, this coefficient is close to 0. Thus, equation (18) highlights a point made in the introduction, namely that when the consumption surplus ratio is very persistent with ϕ_s close to 1, recent consumption contributes very little to current habit. It also shows clearly how the large coefficient on lagged own consumption obtained by Ravina (2019) is consistent with a consumption surplus ratio that's not very persistent.

E. Specifying the Share Process and Forming the Value/Growth Deciles

Following LW, we specify that the market is made up of 200 firms whose dividends aggregate to the market dividend. The share of the market dividend produced by each firm is set deterministically. Let $\underline{s}$ be the minimum share of

[4]CC specify consumption to be IID, and our model approximates CC by setting $a = \log(\delta) - \gamma g + \frac{\gamma(1-\phi_s)}{2}$, $\delta^m = 1$, $\sigma_u = 0$, $\sigma_z = 0$, $\sigma_w = 0$, $\bar{\sigma} = 1$, and $x_t = \gamma\sigma_d(1+\lambda(s_t))$.

any firm, and without loss of generality, suppose firm 1 produces this share initially. Following LW, we choose a growth rate of 5.5% per quarter for the share process so that the cross-sectional distribution of dividend growth rates in the model matches that in the data. Given this growth rate, firm 1's share increases by 5.5% a quarter for 100 quarters to a maximum 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 market dividend, pays a dividend $s^i_t D_t$ at time t.

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

$$P^i_t = \sum_{n=1}^{\infty} s^i_{t+n} P^m_{n,t}. \tag{19}$$

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.

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 $P^i_t / \sum_{\tau=0}^{3} D^i_{t-\tau}$ for firm i. We calculate various annualized moments for the decile excess quarterly returns by simulating the model at a quarterly frequency.

III. Calibration

We calibrate four versions of the model described in Section II.A. LW's model does not distinguish between consumption and market dividends, specifying a stochastic discount factor of the form:

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

where r^f is the log of the risk-free rate, and is constant over time. LW's model also assumes that consumption growth is homoscedastic. Our model nests LW by setting $a = -r^f$, $b = 0$, $\delta^m = 1$, $\sigma_u = 0$, $\sigma_w = 0$, and $\overline{\sigma} = 1$. Our LW case replicates LW by having the shocks to x and d be uncorrelated ($\rho_{d,x} = 0$), 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 $\rho_{d,x} = -0.99$, as implied by the habit specification used in CC.[5] Our Wedge case resembles our Base case except that the consumption process is calibrated to data rather than matched to the market-dividend process: δ^m is no longer set equal to 1, and σ_u is allowed to be positive. Allowing the consumption process to differ

[5]In the LW and Base cases, we set log consumption growth equal to log market dividend growth by setting $\delta^m = 1$ and $\sigma_u = 0$.

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

TABLE 1

Model Parameters

Table 1 lists the parameter values used by LW (first column) and for our Base, Wedge, and LRR-Vol cases (final three columns). All parameters are as defined in Section II. The model is quarterly, but the mean of the log dividend growth g^m and the log risk-free 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
g^m	Annual	2.28%	2.28%	2.28%	2.28%
$r^f = -a$	Annual	1.93%	1.93%	1.93%	1.93%
b or b^m		0	0	0	0
$\overline{x}$	Quarterly	0.625	0.25	0.28	0.365
ϕ_z	Annual	0.91	0.91	0.91	0.91
ϕ_x	Annual	0.865	0.14	0.14	0.14
σ_m	Quarterly	0.0724	0.0724	0.0724	0.0724
σ_d	Quarterly	0.0724	0.0724	0.0160	0.0164
σ_z	Quarterly	0.00165	0.00165	0.00165	0.00165
σ_x	Quarterly	0.1225	0.16	0.3305	0.29
σ_u	Quarterly	0	0	0.0435	0.037
$\rho_{m,z} = \rho_{d,z}$	Quarterly	−0.82	−0.82	−0.82	−0.82
$\rho_{d,x}$	Quarterly	0	−0.99	−0.99	−0.99
$\rho_{z,x}$	Quarterly	0	0.81	0.81	0.81
$\rho_{d,u}$	Quarterly	–	–	−0.30	−0.30
$\rho_{z,u}$	Quarterly	–	–	0.0037	0.15
$\rho_{x,u}$	Quarterly	–	–	0.30	0.30
δ^m	Quarterly	1	1	4.54	4.54
$\overline{\sigma}$	Quarterly	1	1	1	0.918
ϕ_σ	Quarterly	–	–	–	0.994
σ_w	Quarterly	0	0	0	0.037

from the market-dividend process, as is the case in the data, delivers higher excess market return volatility that is closer to the data value. Our LRR-Vol case resembles our Wedge case except that the conditional volatility of consumption growth is allowed to be slowly mean-reverting as in BY: σ_w is no longer set equal to 0. Allowing the conditional volatility of consumption growth to be slowly mean-reverting delivers considerably higher predictability of long-run excess market return using the market price–dividend ratio that is nontrivial and much closer to that in the data.

All four cases use exactly the same calibration used by LW for the z^m process, Δd^m process, and r^f. Table 1 reports the parameters used by the four cases. All the parameters for the LW case are the same as those in LW, so the discussion of parameter choices below pertains to the other three cases.

Our Base, Wedge, and LRR-Vol cases all depart from LW in the calibration of the parameters of the price of risk, the x process. LW calibrate the autocorrelation of x to equal the data autocorrelation of the log price–dividend ratio for the aggregate market (0.87 annually). However, as discussed above, since the persistence of our price of risk x is approximately equal to the persistence of the log surplus s in the CC model, there are good theoretical reasons for why the x process might not be very persistent. In particular, 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 consumption in the recent past than consumption in the distant past. For example, Table 2 uses the expression in equation (18) 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 other three cases. At the LW

TABLE 2

Contribution of Lagged Consumption to Habit: Model

Table 2 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 II 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:

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

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

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 given the micro evidence.

The external habit model of CC implies a value close to -1 for $\rho\left[\varepsilon^d, \varepsilon^x\right]$, but LW set this correlation equal to 0 so as to be able to generate a value premium for both expected return and CAPM alpha. However, one of the main goals of our article 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 is not too persistent (see Section IV.B.1 for intuition). For this reason, we set this correlation to -0.99 in the Base, Wedge, and LRR-Vol cases.[6]

In the Base, Wedge, and LRR-Vol cases, we choose our $\bar{x}$ and σ_x to produce a Sharpe ratio and an expected market price–dividend ratio with mean absolute errors relative to the data values that are no larger than those obtained by LW. Table 3 reports moments for excess annual market return, and annual market price–dividend ratio, for the data and the four model cases, with one column for each. Note that $(P^m/D^m)_t = P_t^m/\sum_{\tau=0}^{3} D_{t-\tau}^m$ is the annual market price–dividend ratio, $p^m - d^m \equiv \log(P^m/D^m)$, $SHARPE^m$ is the unconditional Sharpe ratio for the excess annual market return, and AC denotes autocorrelation. 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, while the Wedge and LRR-Vol cases both do a better job than LW matching the former and a similar job to LW matching the latter.

[6]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.

TABLE 3

Aggregate Moments: Data and Model

Table 3 reports moments for market return, market price–dividend ratio, and market dividend growth. The first column reports moments for the data, the second reports simulated moments for the LW case, and the final three report simulated moments for the Base, Wedge, and LRR-Vol cases. The data are the same as that in LW and are annual from 1890 to 2002. Price–dividend ratios are aggregated to annual frequencies: so $(P^m/D^m)_t = P_t^m / \sum_{\tau=0}^{3} D_{t-\tau}^m$ and $(p^m - d^m)_t \equiv \log((P^m/D^m)_t)$. $R^m - R^f$ is the excess annual market return. $SHARPE^m$ is the unconditional Sharpe ratio of the annual market return, and AC is the autocorrelation.

Moment	Data	LW	Base	Wedge	LRR-Vol
$E[P^m/D^m]$	25.55	20.06	30.97	23.02	25.15
$\sigma[p^m - d^m]$	0.38	0.382	0.260	0.259	0.480
$AC[p^m - d^m]$	0.87	0.884	0.897	0.839	0.945
$E[R^m - R^f]$	6.33	8.096	4.484	6.183	6.824
$\sigma[R^m - R^f]$	19.41	19.42	10.69	14.69	16.54
$AC[R^m - R^f]$	0.03	−0.04	−0.13	−0.27	−0.10
$SHARPE^m$	0.33	0.417	0.419	0.421	0.413

To ensure the covariance matrix of $(\varepsilon^d, \varepsilon^z, \varepsilon^x)$ is positive definite, we specify $\sigma[\varepsilon^x, \varepsilon^z]$ in all four cases so that ε^x and ε^z are correlated only through their correlations with ε^d. Details are in the appendices. With $\rho\left[\varepsilon^d, \varepsilon^z\right]$ set equal to -0.82 for all cases, the assumed value for $\rho[\varepsilon^x, \varepsilon^z]$ is 0.81.

While all four cases use the same Δd^m process, only the Wedge and LRR-Vol cases allow log dividend growth to be a levered version of log consumption growth. For these cases, we set $\rho\left[\varepsilon^d, \varepsilon^z\right] = \rho[\varepsilon^m, \varepsilon^z]$, and following the literature (e.g., Abel (1999)), we set δ^m equal to $\frac{\sigma_m}{\sigma_d}$. The annual correlation of log consumption growth with log dividend growth is 0.55 in BY's sample period, but we choose a larger correlation than in the data, 0.82 at a quarterly frequency, so that the price–dividend ratio converges for a range of $\bar{x} > 0$. Using the methods of Stambaugh (1997) and Lynch and Wachter (2013), we estimate the unconditional volatility of annual log consumption growth in the LW sample period to be 3.18% in the data; we match this value in the Wedge and LRR-Vol cases. As with ε^x and ε^z, ε^x and ε^u are correlated only through their correlations with ε^d.

When we calibrate the volatility process for Δd in the LRR-Vol case, σ_t, we nail down the value for ϕ_σ by assuming that the conditional variance of monthly log consumption growth follows the AR(1) specification used in Bansal, Kiku, and Yaron (2007). However, we scale our volatility process σ_t to preserve the unconditional second moment of the shock to log consumption growth from the Wedge case, i.e., $E\left[\left(\sigma_t \varepsilon_{t+1}^d\right)^2\right]$ in the LRR-Vol case equals $E\left[\left(\varepsilon_{t+1}^d\right)^2\right]$ in the Wedge case. Following BY, we assume that ε^w is uncorrelated with all other shocks. The σ_t process in the LRR-Vol case is negative less than 0.3% of the time.[7]

IV. Results

This section reports the results for U.S. data and for the four model cases. The aggregate data are the same as that in LW, annual from 1890 to 2002, and the data for the value and growth portfolios are quarterly from Kenneth

[7]See the appendices for more details of the calibration of Δd for the Wedge and LRR-Vol cases.

French's website (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) for the period 1952 to 2016, and the data for the market-dividend strips are quarterly from 1996 to 2009. Quarterly returns are annualized by multiplying means by 4, volatilities by $\sqrt{4}$, and Sharpe ratios by $4/\sqrt{4}$. Whenever one of the models is simulated to obtain particular statistics of interest, it is simulated at a quarterly frequency for at least 4 million quarters, and until those statistics have converged; that is, until an additional 4 million simulated quarters causes the values of all these moments and statistics to change by less than small prespecified tolerances.

A. 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 consumption process is assumed to be equal to the market dividend process, which is calibrated to data.

1. Aggregate Moments

Table 3 reports aggregate moments. The parameters of Δd^m and z are chosen by LW to match the data, and they show that their chosen parameters, which we use, are able to match the volatility and autocorrelation of log annual market dividend growth in the data. However, the Base case does deliver an expected excess market return and a market excess return volatility that are 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. While the LW case matches the data log annual price–dividend ratio volatility quite closely, the Base case delivers a value of 0.260 that is much lower than the data value of 0.38. Equations (11) and (15) suggest that these two volatilities are much lower in the Base case than the LW case because the unconditional volatility of the price-of-risk variable x and magnitude of the B_x function are much lower in the Base case (see Table 1 and Section IV.B.1, respectively).

Because parameter values were not chosen specifically to match the autocorrelation of the market price–dividend ratio in the data, it is impressive that the Base-case value of this autocorrelation, 0.897, is higher than, but close to, both the data value of 0.87 and the LW value of 0.884. This correlation is high in the Base case even though the autocorrelation of x is low because equation (11) suggests that it is a weighted average of the autocorrelations of x and z^m, where the latter is more persistent than the market price–dividend ratio in the data.

2. Predictive Regressions

Tables 4 and 5 report results for market predictive regressions of the following form:

$$\sum_{i=1}^{H} Y_{t+i} = \beta_0 + \beta_1 \mathrm{INFO}_t + \varepsilon_{t+H}, \tag{20}$$

where Y_{t+i} equals future market excess log return $\left(r^m_{t+i} - r^f_{t+i-1}\right)$ in Table 4, and future changes in log market dividend Δd^m_{t+i} in Table 5. In Panel A of both tables,

TABLE 4

Predictive Regressions for Future Market Excess Log Return: Data and Model

Table 4 reports results for the following regression:

$$\sum_{i=1}^{H}\left(r_{t+i}^{m}-r_{t+i-1}^{f}\right)=\beta_0+\beta_1 \mathrm{INFO}_t+\varepsilon_{t+H},$$

where INFO_t equals the log annual market price–dividend ratio at time t, $(p^m-d^m)_t \equiv \log\left(P_t^m/\sum_{\tau=0}^{3}D_{t-\tau}^m\right)$, in Panel A, and equals each of the three drivers of $(p^m-d^m)_t$ in Panel B: z_t^m, x_t, and σ_t. $r_{t+i}^m-r_{t+i-1}^f$ is the excess log market return for the year ending at time $t+i$. Results are reported for the data and the four model cases: LW, Base, Wedge, and LRR-Vol. The data are annual, spanning 1890–2016. The log consumption-market dividend ratio at t for annual data is used to measure z_t^m in the data. Results are reported for return horizons (H) of 1 year and 10 years for the data and the models. Newey–West t-statistics are reported for a lag length of $(H-1)$ and the maximum-likelihood-optimal lag length. Adj R^2 is the adjusted regression R^2.

	Horizon H (Years)	Data	t-Stat (NW) Lags $H-1$	t-Stat (NW) Lags MLE-Opt	LW	Base	Wedge	LRR-Vol	LW	Base	Wedge	LRR-Vol	LRR-Vol
Panel A													
INFO_t:		$(p^m-d^m)_t$											
β_1	1	−0.24	−4.10	−3.92	−0.12	−0.01	−0.04	−0.06					
	10	−1.19	−3.67	−3.27	−0.68	−0.05	−0.14	−0.49					
Adj R^2	1	0.09			0.07	0.00	0.01	0.03					
	10	0.27			0.30	0.00	0.02	0.27					
Panel B													
INFO_t:		z_t^m							x_t				σ_t
β_1	1	2.82	1.27	1.26	−0.61	0.63	2.02	1.01	0.12	0.10	0.14	0.10	0.09
	10	22.35	2.17	2.20	−4.18	−0.73	−1.08	−2.45	0.67	0.09	0.14	0.09	0.78
Adj R^2	1	0.01			0.00	0.00	0.01	0.00	0.10	0.04	0.18	0.06	0.04
	10	0.08			0.00	0.00	0.00	0.00	0.43	0.00	0.04	0.01	0.33

INFO_t equals the log annual market price–dividend ratio at time t, $(p^m-d^m)_t$, and in Panel B, INFO_t equals each of the three drivers of $(p^m-d^m)_t$ separately: z_t^m; x_t; and σ_t. The data are annual, spanning 1890–2016. The log consumption-market dividend ratio at time t for annual data is used to measure z_t^m in the data.[8] Results are reported for return horizons (H) of 1 year and 10 years for the data and the models. Newey–West t-statistics are reported for a lag length of $(H-1)$ and the maximum-likelihood-optimal lag length. Adj R^2 is the adjusted regression R^2.

Perhaps the most glaring inability of the Base case to match data moments concerns the log excess market return predictability regressions that use $(p^m-d^m)_t$ as the predictor, which are reported in Panel A of Table 4. The data and the LW case deliver R^2s and negative predictability coefficients that are both larger in absolute value at a 10-year return horizon than at a 1-year return horizon. However, 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. In contrast, the LW-case R^2 for the 10-year return horizon is 0.30, which is very close to the 0.27 found in the data. In the LW and Base cases, the log market price–dividend ratio at time t is determined by the price of risk,

[8]Lettau and Ludvigson (2005) show that if consumption growth follows a random walk, then the consumption-market dividend ratio, when stationary, is a linear transformation of z_t^m. We prove in the Supplementary Material that the same result holds in our model. We demean the consumption-market dividend ratio and endow it with the standard deviation of z_t^m as in LW.

TABLE 5

Predictive Regressions for Future Changes in Log Market Dividend: Data and Model

Table 5 reports results for the following regression:

$$\sum_{i=1}^{H} \Delta d_{t+i}^{m} = \beta_0 + \beta_1 \mathrm{INFO}_t + \varepsilon_{t+H},$$

where INFO_t equals the log annual market price–dividend ratio at time t, $(p^m - d^m)_t \equiv \log\left(P_t^m / \sum_{\tau=0}^{3} D_{t-\tau}^m\right)$, in Panel A, and equals each of the three drivers of $(p^m - d^m)_t$ in Panel B: z_t^m, x_t, and σ_t. Δd_{t+i}^m is the change in the log annual market dividend for the year ending at time $t+i$. Results are reported for the data and the four model cases: LW, Base, Wedge, and LRR-Vol. The data are annual, spanning 1890–2016. The log consumption-market dividend ratio at t for annual data is used to measure z_t^m in the data. Results are reported for dividend-change horizons (H) of 1 year and 10 years for the data and the models. Newey–West t-statistics are reported for a lag length of $(H-1)$ and the maximum-likelihood-optimal lag length. Adj R^2 is the adjusted regression R^2.

	Horizon H (Years)	Data	t-Stat (NW) Lags $H-1$	t-Stat (NW) Lags MLE-opt	LW	Base	Wedge	LRR-Vol	LW	Base	Wedge	LRR-Vol	LRR-Vol
Panel A													
INFO_t:		$(p^m - d^m)_t$											
β_1	1	0.01	0.24	0.24	0.04	0.10	0.09	0.03					
	10	−0.08	−0.59	−0.58	0.30	0.69	0.64	0.19					
Adj R^2	1	−0.01			0.02	0.05	0.04	0.01					
	10	0.00			0.09	0.20	0.18	0.05					
Panel B													
INFO_t:		z_t^m							x_t				σ_t
β_1	1	4.58	2.76	3.23	3.64	3.64	3.64	3.64	0.00	0.05	0.02	0.03	0.00
	10	23.48	3.98	3.60	24.70	24.70	24.70	24.70	0.00	0.32	0.16	0.18	0.00
Adj R^2	1	0.06			0.06	0.06	0.06	0.06	0.00	0.01	0.01	0.01	0.00
	10	0.34			0.23	0.23	0.23	0.23	0.00	0.03	0.03	0.03	0.00

x_t, and conditional mean log quarterly consumption growth, z_t^m. Panel B of Table 4 shows that the ability of $(p^m - d^m)_t$ in the LW case to forecast future log excess market return is solely due to the predictive ability of x_t: the predictive coefficients for x_t are positive for both future return horizons, and their magnitude, as well as that of the adjusted R^2s, increase going from the 1- to the 10-year horizon, with the latter topping out at 0.43 for the 10-year horizon. In contrast, the adjusted R^2 for x_t in the Base case declines from 0.04 to 0.00 going from the 1- to the 10-year return horizon, because the persistence of the x process is much smaller for the Base than the LW case by construction. When z_t^m is used as the predictor, the adjusted R^2s at both horizons are indistinguishable from 0 for both the LW and Base cases.

Turning to Panel A of Table 5, both the LW and Base cases do a poor job of reproducing the lack of predictability of 1- and 10-year log dividend growth found in the data using the log market price–dividend ratio, especially for a future return horizon of 10 years: the adjusted R^2 is 0.00 for the data and at least 0.09 for both cases. All the cases considered (LW and Base, as well as Wedge and LRR-Vol) 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 in Panel B of Table 5, all the cases considered here replicate LW's ability to roughly match the predictive coefficients and R^2s of the regressions found in the data when log annual consumption-market dividend ratio at time t is used as a proxy for z^m.

TABLE 6

Value vs. Growth Portfolios: Data and Model

Table 6 reports results for the extreme growth decile (portfolio 1), the extreme value decile (portfolio 10), and the portfolio long portfolio 10 and short portfolio 1 (10 M1 portfolio). Panel A reports expected excess return, the volatility of excess return, and the unconditional Sharpe ratio. Panel B reports CAPM alpha, beta, and adjusted regression R^2. Panel C reports alpha, beta (the slope on the market factor), gamma (the slope on the HML factor), and adjusted regression R^2 from a 2-factor regression. Data are from Ken French's website and span 1952–2016. The HML factor is long value deciles 8–10 (H) and short growth deciles 1–3 (L) in the models, and is the Fama–French HML factor in the data. The first three columns report results for the data: deciles formed by sorting on earnings over market price (E/P), cash flow over market price (C/P), and book value over market price (B/M). The remaining columns report results for the four model cases: LW, Base, Wedge, and LRR-Vol. 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 $P_t^i/\sum_{\tau=0}^{3} D_{t-\tau}^i$ for firm i. Returns are quarterly, but the results are annualized by multiplying the expected excess return and alphas by 4, the volatility of the excess return by $\sqrt{4}$, and the Sharpe ratio by $4/\sqrt{4}$.

Portfolio		E/P	C/P	B/M	LW	Base	Wedge	LRR-Vol
Panel A. $R^i - R^f$								
$E\left[R^i - R^f\right]$	1	6.25	6.01	6.14	5.27	3.38	3.24	3.88
	10	12.30	12.11	11.30	10.79	5.29	7.07	8.24
	10 M1	6.05	6.10	5.15	5.51	1.91	3.83	4.36
$\sigma\left[R^i - R^f\right]$	1	20.80	20.31	18.99	19.53	9.90	10.51	16.16
	10	20.29	19.33	23.20	17.65	11.47	15.73	16.51
	10 M1	15.17	14.80	18.16	8.82	4.46	8.25	8.58
$SHARPE_i$	1	0.30	0.30	0.32	0.27	0.34	0.31	0.24
	10	0.61	0.63	0.49	0.61	0.46	0.45	0.50
	10 M1	0.40	0.41	0.28	0.63	0.43	0.46	0.51
Panel B. $R^i - R^f = \alpha_i + \beta_i(R^m - R^f) + \varepsilon^i$								
α_i	1	−2.60	−2.68	−2.01	−2.72	−0.66	−0.82	−2.43
	10	4.23	4.53	2.66	3.71	0.54	0.48	1.58
	10 M1	6.83	7.20	4.67	6.43	1.20	1.30	4.01
β_i	1	1.18	1.16	1.09	0.99	0.90	0.66	0.93
	10	1.08	1.01	1.15	0.87	1.06	1.07	0.98
	10 M1	−0.11	−0.15	0.06	−0.11	0.16	0.41	0.05
Adj R_i^2	1	0.86	0.86	0.87	0.96	0.95	0.84	0.90
	10	0.75	0.73	0.65	0.92	0.97	0.99	0.96
	10 M1	0.01	0.03	0.00	0.06	0.14	0.53	0.01
Panel C. $R^i - R^f = \alpha_i + \beta_i(R^m - R^f) + \gamma_i R^{\text{HML}} + \varepsilon^i$								
α_i	1	−0.96	−0.98	−0.32	0.06	0.00	0.19	0.00
	10	2.27	2.75	−1.15	0.06	0.00	0.09	−0.01
	10 M1	3.23	3.73	−0.83	0.00	0.00	−0.10	−0.01
β_i	1	1.14	1.12	1.05	0.93	0.99	0.95	0.96
	10	1.13	1.06	1.25	0.95	0.99	0.95	0.96
	10 M1	−0.01	−0.06	0.21	0.02	0.00	0.00	0.00
γ_i	1	−0.37	−0.38	−0.38	−0.18	−0.23	−0.29	−0.24
	10	0.44	0.40	0.85	0.23	0.18	0.11	0.16
	10 M1	0.81	0.78	1.23	0.41	0.42	0.40	0.40
Adj R_i^2	1	0.90	0.91	0.92	1.00	1.00	0.99	1.00
	10	0.81	0.78	0.84	1.00	1.00	1.00	1.00
	10 M1	0.40	0.40	0.62	1.00	1.00	1.00	1.00

3. Value Versus Growth Portfolios

Table 6 reports results for the extreme growth decile (portfolio 1), the extreme value decile (portfolio 10), and the portfolio long portfolio 10 and short portfolio 1 (10 M1 portfolio). Panel A reports expected excess return, the volatility of excess return, and the unconditional Sharpe ratio. Panel B reports CAPM alpha, beta, and adjusted regression R^2. Panel C reports alpha, beta (the slope on the market factor), gamma (the slope on the HML factor), and adjusted regression R^2 from a 2-factor regression. Data are from Ken French's website and span 1952–2016. The HML factor is long value deciles 8–10 (H) and short growth deciles 1–3 (L) in the models, and is the Fama–French HML factor in the data. The first three columns report results for the data: deciles formed by sorting on earnings over market price (E/P),

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

cash flow over market price (C/P), and book value over market price (B/M). The remaining columns report results for the four model cases: returns are quarterly, but the results are annualized.

Table 6 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 5.15% versus the 1.91% per annum delivered by the Base case. Similarly, the CAPM-alpha spread for these two extreme B/M deciles is 4.67% per annum in the data, but only 1.20% per annum in the Base case. Moreover, both the B/M deciles and the Base case deliver a CAPM-alpha spread between the extreme value and the growth deciles that is smaller than the expected excess return spread, while the converse is true for the E/P and C/P deciles and the LW case. The reason is that the CAPM beta for the extreme value decile is higher than for the extreme growth decile for the B/M deciles and the Base case, while the converse is true for the E/P and C/P deciles and the LW case, as the rows labeled β_i in Panel B of Table 6 show.

In Panel A of Table 6, return volatility relative to the data is in the same ballpark for the LW case but much smaller for the Base case with respect to the extreme deciles, and is much smaller for both the LW and Base cases with respect to the 10 M1 portfolio. These results suggest that the returns on the extreme deciles are much more correlated for the two model cases than for the data, which is not surprising given that the market dividend shares received by the firms in the model are deterministic. Both the LW and Base cases do a good job matching the Sharpe ratio for the extreme growth decile in the data, while the LW case does a better job for the extreme value decile, and the Base case does a much better job for the 10 M1 portfolio.

Turning to the results of the 2-factor pricing model in Panel C of Table 6, both the LW and Base cases produce counterfactual results: αs that are close to or equal to 0, and adjusted R^2s that are indistinguishable from 1, for both extreme deciles and the 10 M1 portfolio. In contrast, sorting on E/P and C/P in the data produces αs that are negative for the extreme growth decile and positive for the extreme value portfolio, while sorting on any of the three measures produces adjusted R^2s that are no more than 0.92 for the extreme growth decile, 0.84 for the extreme value decile, and 0.62 for the 10 M1 portfolio. Consistent with the data, the LW and Base cases both produce 2-factor βs that are close to 1 for the extreme deciles and close to or equal to 0 for the 10 M1 portfolio, and 2-factor γs that are negative for the extreme growth decile and positive for the extreme value decile, though the 2-factor γs in the data are much larger in magnitude than for the LW or Base cases.

B. Intuition for the Base Case Results

This subsection provides intuition for the Base case results described in Section IV.A.

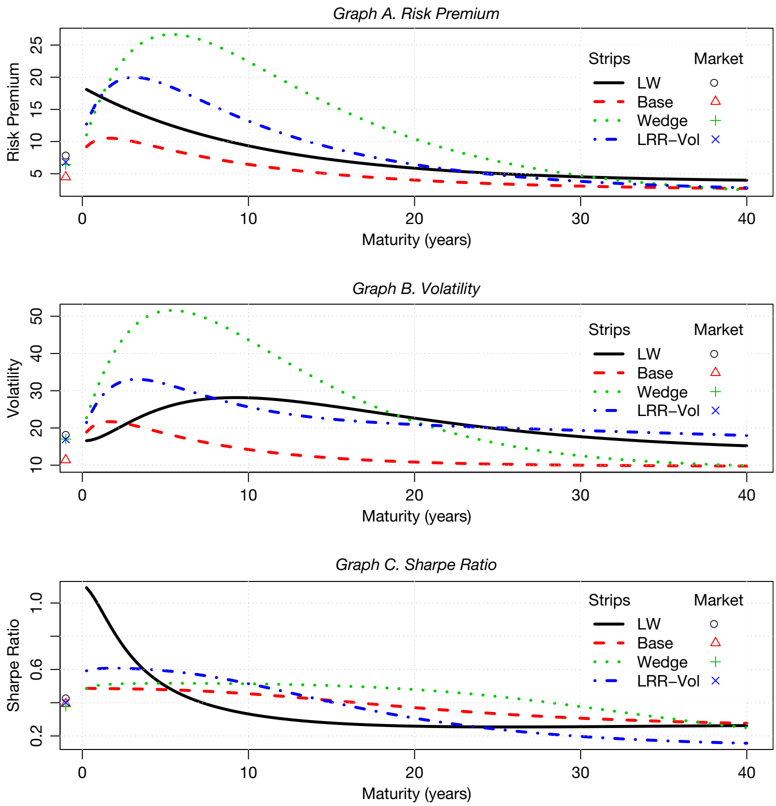
1. How the Base Case Delivers a Value Premium

One of the main messages of Section IV.A is that the Base case can deliver a value premium both in expected excess return and CAPM alpha. To better

FIGURE 1

Returns on Market and Market-Dividend Strips: Model

Figure 1 plots the unconditional expected excess return (risk premium), the unconditional excess return volatility, and the unconditional Sharpe ratio, all for quarterly returns, in Graph A, Graph B, and Graph C, respectively, for market-dividend strips as a function of maturity. The points to the left of each of the graphs represent values for the market portfolio. For each graph, the circle to the left and the solid line are for the LW case, the triangle to the left and the dashed line are for the Base case, the plus to the left and the dotted line are for the Wedge case, and the cross to the left and dot-dashed line are for the LRR-Vol case. The risk premium, volatility, and Sharpe ratio are annualized by multiplying by 4, $\sqrt{4}$, and $4/\sqrt{4}$, respectively.



understand why the Base case delivers a value premium in expected excess return despite a conditional correlation between consumption growth and the price of risk that is close to -1, we now turn our attention to the market-dividend strips described in Section II. Figure 1 plots the unconditional risk premium, the unconditional return volatility, and the unconditional Sharpe ratio, all for quarterly returns, in Graph A, Graph B, and Graph C, respectively, for the market portfolio as points on the left-hand side of the graphs, and for market-dividend strips, as a function of maturity, on the right-hand side of the graphs. In each graph, the solid line is the LW case, the dashed line is the Base case, the dotted line is the Wedge case, and the dot-dashed line is the LRR-Vol case. The risk premium, volatility, and Sharpe ratio are all annualized.

It is worth noting that the excess return on the market portfolio is a weighted average of the excess returns on the market-dividend strips, where all the weights

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

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. Graph A of Figure 1 shows that in the LW case, the expected excess return on the market-dividend strip 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. Graph B shows that excess return volatility is lower for the Base than the LW case for strip maturities longer than about 3 years, and is hump-shaped in both cases, though the hump occurs much earlier for the Base case. Graph C shows that the Sharpe ratio declines monotonically for both cases, though the relation is strongly convex for the LW case at all but the very shortest maturities, and concave at maturities out to about 20 years for the Base case.

The expression for the log risk premium on market-dividend strips in equation (13) simplifies to the following expression in the LW, Base, and Wedge cases by setting $\varepsilon^{w}_{t+1}=0$ and $\overline{\sigma}=1$:

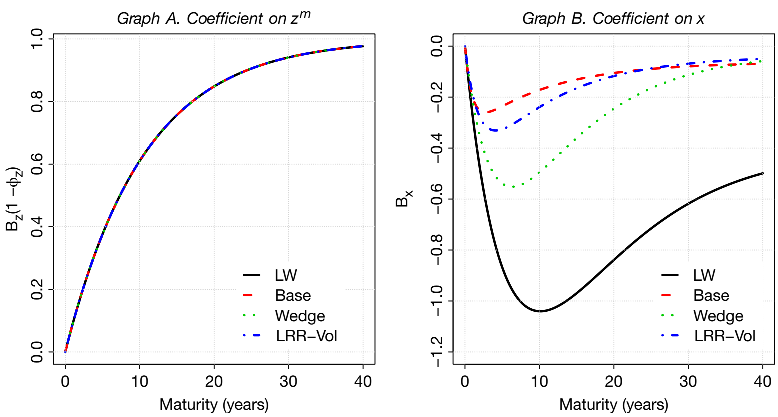
$$
\text{(21)} \quad \log\left(E_t\left[\frac{R^m_{n,t+1}}{R^f_t}\right]\right) = E_t\left[r^m_{n,t+1}-r^f_t\right]+\frac{1}{2}\sigma^2_t\left[r^m_{n,t+1}\right]
$$
$$
= \left(\delta^m\sigma^2_d+\sigma_{d,u}+B_x(n-1)\sigma_{x,d}+B_z(n-1)\sigma_{z,d}\right)\frac{1}{\sigma_d}x_t.
$$

Using equation (21), the shapes of $B_z(n)$ and $B_x(n)$ as functions of n can be used to better understand the relation between the expected excess return on a market-dividend strip and its maturity plotted in Graph A of Figure 1. Figure 2 plots

FIGURE 2

Market-Dividend Strips Log Price–Dividend Ratio Coefficients $B_z(n)$ and $B_x(n)$: Model

Figure 2 plots $B_z(n)(1-\phi_z)$ and $B_x(n)$ in Graph A and Graph B, respectively, for market-dividend strips with n quarters to maturity, as a function of $n/4$ (maturity in years). In each graph, the solid line is the LW case, the dashed line is the Base case, the dotted line is the Wedge case, and the dot-dashed line is the LRR-Vol case. For all four cases, $B_z(n)$ and $B_x(n)$ are, respectively, the coefficient on z^m and the coefficient on x in equation (11) of Section II for the quarterly log price–dividend ratio for the market-dividend strip paying out in n quarters. $B_z(n)$ is multiplied by $(1-\phi_z)$.



$B_z(n)(1-\phi_z)$ and $B_x(n)$ in Graphs A and B, respectively, for market-dividend strips with n quarters to maturity, as a function of $n/4$ (maturity in years). In each graph, the solid line is for the LW case and the dashed line is for the Base case. $B_z(n)$ and $B_x(n)$ are, respectively, the coefficient on z^m and the coefficient on x in equation (11) for the price–dividend ratio of market-dividend strips paying out in n quarters.

As an expression for $\log\left(E_t\left[R^m_{n,t+1}/R^f_t\right]\right)$, the right-hand side of equation (21) is likely to be highly correlated with $\log\left(E_t\left[R^m_{n,t+1}-R^f_t\right]\right)$, and so can be used to understand how the unconditional expected excess return on a market-dividend strip $\left(E\left[R^m_{n,t+1}-R^f_t\right]\right)$ varies with maturity, by evaluating it at the unconditional mean for x, $\bar{x}$, which is positive. Because the z^m process is the same for all four cases, $B_z(n)$ is also the same for all four cases. Figure 2 shows that $B_z(n)$ is always positive and increasing in n, and Table 1 shows that $\rho_{d,z}$ is negative and the same value for all four cases. So using equation (21), the conditional risk premium evaluated at $\bar{x}$ is declining in n whenever the covariance between shocks to x and d is assumed to be 0. Since $\rho_{x,d}$ is 0 in the LW case, it follows that $E_t\left[R^m_{n,t+1}-R^f_t\right]$ 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 6 because value stocks have shorter cash flow durations than growth stocks.

Figure 3 illustrates how, for the LW case, the shape of $\log\left(E_t\left[R^m_{n,t+1}/R^f_t\right]\right)$ as a function of maturity depends on the shape of $B_z(n)\sigma_{z,d}$. Graphs A–D of Figure 3 show the decomposition of equation (12), the log risk premium on market-dividend strips as a function of maturity, n. The equation is evaluated at $(x_t,\sigma_t)=(\bar{x},\bar{\sigma})$ for all cases, and each graph contains the decomposition for one case (see the graph label). In each graph, the solid "Total" line represents the total log risk premium, while the dashed "Constant" line represents the part which is constant in n. The three lines, dotted "B_x," dot-dashed "B_z," and long-dashed "B_σ," represent the contributions from $B_x(n)$, $B_z(n)$, and $B_\sigma(n)$, respectively. The log risk premium is for quarterly returns, as in equation (12), and there is no annualization. Graph A is for the LW case and shows that the $B_x(n)$ term in equation (21) is 0 for all n, while the shape of the total log risk premium is completely driven by the shape of the $B_z(n)$ term.

Since $B_z(n)$ is positive for any n, a positive shock to z_{t+1} causes a positive shock to $P^m_{n,t+1}/D^m_{t+1}$ which causes a positive shock to $R^m_{n,t+1}$. Taking $\rho_{d,z}$ to be negative as in all four model cases discussed in this article, this positive shock to $R^m_{n,t+1}$ is typically associated with a negative shock to d_{t+1} which makes a market-dividend strip a hedge against shocks to consumption. Consequently, when $\rho_{d,x}=0$ as in the LW case,[9] the strip's conditional premium evaluated at $\bar{x}$ as a function of n is flat for $\rho_{d,z}=0$ and becomes more negatively sloped as $\rho_{d,z}$ becomes more negative, which implies a stronger contribution to a value premium. Fixing $\rho_{d,z}$ and increasing ϕ_z causes $B_z(n)$ to be higher for all n and since $B_z(n)$ is increasing in n for any given ϕ_z, the strip's conditional premium evaluated at $\bar{x}$ as a function of n becomes more negatively sloped as ϕ_z increases,[10] which also implies a stronger

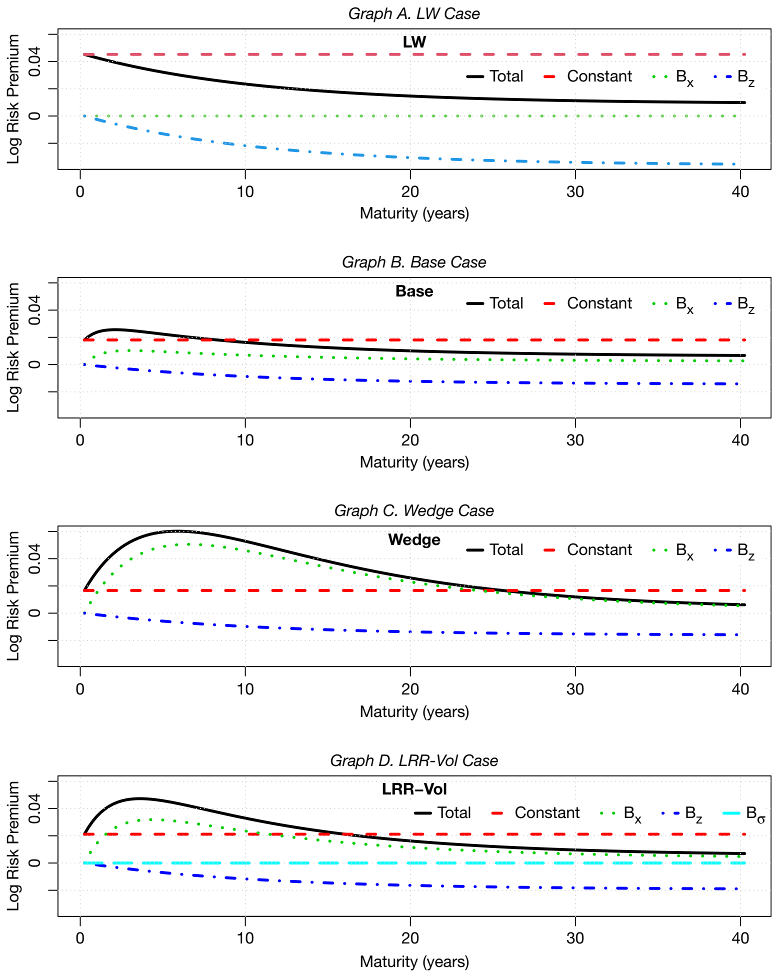
[9]Note that when $\rho_{d,x}=0$, the strip's conditional premium evaluated at $\bar{x}$ is always the same for a maturity $n=1$ irrespective of ϕ_z or $\rho_{d,z}$.

[10]Note that $B_z(1)$ is the same irrespective of ϕ_z.

FIGURE 3

Decomposition of Market-Dividend Strip Log Risk Premium: Model

Graphs A–D in Figure 3 show the decomposition of equation (12), the log risk premium on market-dividend strips, as a function of maturity, n. The equation is evaluated at $(x_t, \sigma_t) = (\bar{x}, \bar{\sigma})$ for all cases (LW, Base, Wedge, and LRR-Vol), and each graph contains the decomposition for one case (see the graph title). In each graph, the solid "Total" line represents the total log risk premium, while the short-dashed "Constant" line represents the part which is constant in n. The three lines, dotted "B_x," dot-dashed "B_z," and long-dashed "B_σ," represent the contributions from $B_x(n)$, $B_z(n)$, and $B_\sigma(n)$, respectively. The log risk premium is for quarterly returns, as in equation (12), and there is no annualization.



contribution to a value premium. This analysis suggests that a less negative $\rho_{d,z}$ or a lower ϕ_z, holding all else equal, makes it more difficult to obtain a value premium.

When the covariance between shocks to x and d is negative, its effect on the conditional risk premia for n-period market-dividend strips depends on the sign and magnitude of $B_x(n)$. If $B_x(n)$ is negative, which Figure 2 shows is so for all four model cases, and 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 market-dividend strips increases in the absolute value of $B_x(n)$ for all n. Moreover, the relation between the conditional risk premia for the n-period market-dividend strip 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 quarters in the future. Consequently, the absolute value of $B_x(n)$ increases monotonically for many quarters 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 quarters into the future. Consequently, the absolute value of $B_x(n)$ increases monotonically for only a few quarters 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.

In the Base case, the persistence of x is sufficiently low that this intuition causes negative-valued $B_x(n)$ to have an inverted hump shape as shown in Figure 2. Consequently, $B_x(n-1)\sigma_{x,d}$ in equation (21) evaluated at $\bar{x}$ is hump-shaped, and the implication is that $E_t\left[R^m_{n,t+1} - R^f_t\right]$ can be hump-shaped, as reported in Figure 1. Hence, the low persistence of the price of risk variable x is able to deliver a value premium in expected excess return in the Base case as reported in Table 6, despite a CC-habit conditional correlation between x and d of close to -1. Graph B of Figure 3 is for the Base case and shows that the hump-shape as a function of maturity for the total log risk premium on market-dividend strips is indeed being driven by the $B_x(n)$ term in equation (21) being hump-shaped as a function of maturity n with an early turning point, around 3 years, consistent with our intuition.

2. Trade-Off Between Value Premium and Market Return Predictability in Models Without Long-Run Risk in Volatility

The other main message of Section IV.A is that the Base case is unable to deliver the ability of the log market price–dividend ratio to predict future log market excess returns that is found in the data. The reason is the assumed low persistence of the price of risk variable x in the Base case. Equation (21) shows that in the LW and Base cases, the time-series variation in the log risk premia for market-dividend strips is driven by time-series variation in x. Since the market is a portfolio of the market-dividend strips, it follows that the time-series variation in the log risk premium for the market is also driven by time-series variation in x, and so the level of predictability of future excess market returns using today's information depends on ϕ_x, the persistence of x. Because ϕ_x is low in the Base case by construction, it follows that the log market price–dividend ratio must have little predictive ability for future excess market returns.

But it is precisely the low persistence of the price of risk variable x that allows the Base case to deliver a value premium despite a $\rho_{d,x}$, the conditional correlation of d with x, of -0.99. LW argue that when $\rho_{d,x} = -0.99$ and ϕ_x is high, as in the CC habit model, the economy produces a growth premium rather than a value premium. Moreover, LW set ϕ_x high, which delivers market return predictability, but then are forced to set $\rho_{d,x} = 0$ to obtain a value premium. For this reason, it is interesting to analyze the trade-off between the value premium and market return predictability as a function of ϕ_x for various values of $\rho_{d,x}$. The trade-off for values of $\rho_{d,x}$ between -1 and 0 is also of interest because some recent pricing models allow $\rho_{d,x}$ to lie in this range (e.g., the external habit model of Bekaert, Engstrom, and Grenadier (2010)).

To this end, Figure 4 shows how the value premium in annual raw return, and the predictability of the market excess log return using the market portfolio's log price–dividend ratio, vary with ϕ_x and $\rho_{d,x}$. The trade-off likely depends on the values of ϕ_z and $\rho_{d,z}$, since the analysis in Section IV.B.1. indicates that a less negative $\rho_{d,z}$ or a lower ϕ_z, holding all else equal, makes it more difficult to obtain a value premium. Consequently, Figure 4 looks at the trade-off for various values of these two parameters in the LW model. The measure of log return predictability is taken to be the regression R^2 at a 10-year horizon. The value premium and price–dividend ratio are both annual. The benchmark at the top sets $\phi_z = 0.91$, and $\rho_{d,z} = -0.82$, as in the LW case. Graphs B and D, and then C and E, show how the results vary as ϕ_z declines and $\rho_{d,z}$ becomes less negative, respectively. On each graph there are five lines, showing how the results vary with $\rho_{d,x}$.[11] The points marked on each line show the ϕ_x values for which the adjusted $R^2 = 0.30$, the LW value. For lower ϕ_x values, the R^2 is lower than the LW value; for higher ϕ_x values, it is higher. The thick-line $\rho_{d,x}$ value is the cutoff $\rho_{d,x}$ value such that the value premium equals 0 at the ϕ_x value that delivers $R^2 = 0.30$. So for any $\rho_{d,x}$ value that is more negative than this critical $\rho_{d,x}$ value, there is no ϕ_x value such that the value premium is positive and $R^2 = 0.30$. The closer to 0 is this critical $\rho_{d,x}$ value for a given ϕ_z and $\rho_{d,z}$ pair, the more difficult it is to get a value premium when the adjusted R^2 needs to be 0.30.

Graph A of Figure 4 shows a cutoff $\rho_{d,x}$ of -0.4455 for the cases in which $\phi_z = 0.91$, and $\rho_{d,z} = -0.82$ as in the LW case: at this cutoff $\rho_{d,x}$, a persistence parameter for x of 0.49 annualized delivers an adjusted R^2 of 0.30 and a value premium of 0. For any $\rho_{d,x}$ more negative than this cutoff, there does not exist a ϕ_x value that delivers an adjusted R^2 of at least 0.30 that is also able to deliver a value premium. In contrast, for the LW value for $\rho_{d,x}$ of 0, the cutoff ϕ_x value at which the adjusted R^2 is 0.30 is the LW value of 0.87, which delivers a large positive value premium as reported in Table 6.

Going from Graph A to B to D of Figure 4, ϕ_z declines from 0.91 in Graph A to 0.455 in Graph B to 0 in Graph D, and we see a corresponding move toward 0 by the cutoff $\rho_{d,x}$ from -0.4455 to -0.15 to -0.04, which implies that, as ϕ_z declines, it becomes more difficult to get a value premium when the adjusted R^2 needs to be 0.30. Going from Graph A to C to E, $\rho_{d,z}$ moves from -0.82 in Graph A to -0.41 in Graph B to 0 in Graph E, and we see a corresponding move toward 0 by the cutoff $\rho_{d,x}$ from -0.4455 to -0.22 to -0.03, which implies that, as $\rho_{d,z}$ becomes less negative, it becomes more difficult to get a value premium when the adjusted R^2 needs to be 0.30. These two findings are consistent with the analysis in Section IV.B.1 that suggests that a less negative $\rho_{d,z}$ or a lower ϕ_z, holding all else equal, makes it more difficult to obtain a value premium.

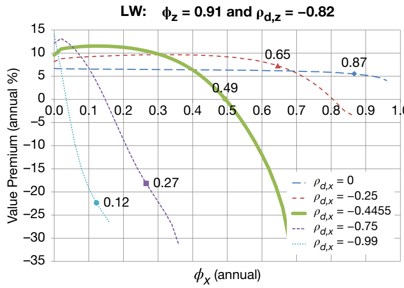
[11] σ_x and σ_z are adjusted to keep the unconditional variances of x and z constant across all cases plotted.

FIGURE 4

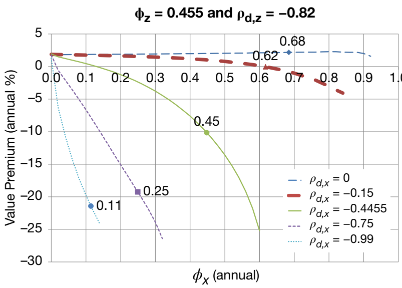
Value Premium and Log Return Predictability as a Function of ϕ_x and $\rho_{d,x}$: Model

Figure 4 shows how (1) the value premium, and (2) the predictability of the log market excess return using the market portfolio's annual log price–dividend ratio, vary with ϕ_x and $\rho_{d,x}$ for different ϕ_z and $\rho_{d,z}$ values in the LW model. The measure of log return predictability is the adjusted regression R^2 at a 10-year horizon. The value premium is the average annual return differential between the value and the growth deciles. The benchmark at the top is the LW case, where $\phi_z = 0.91$, and $\rho_{d,z} = -0.82$. The left and right columns of graphs show how the results vary with ϕ_z and $\rho_{d,z}$, respectively. On each graph there are five lines, showing how the results vary with $\rho_{d,x}$. σ_x and σ_z were adjusted to keep the unconditional variances of x and z constant across the cases. The points marked on each line show the ϕ_x values where $R^2 = 0.30$, the LW value. For lower ϕ_x values, the R^2 is lower than the LW value; for higher ϕ_x values, it is higher. The thick-line $\rho_{d,x}$ value is the $\rho_{d,x}$ value such that the value premium equals 0 at the ϕ_x value that delivers $R^2 = 0.30$. So for any $\rho_{d,x}$ value (i.e., more negative than this thick-line $\rho_{d,x}$ value), there is no ϕ_x value such that the value premium is positive and R^2 is at least 0.30.

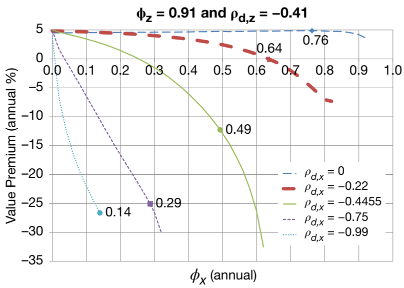
Graph A. $\phi_z = 0.91$ and $\rho_{d,z} = -0.82$ (LW Case)



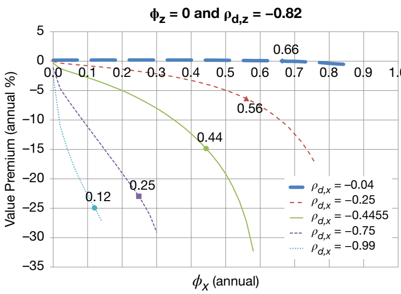
Graph B. $\phi_z = 0.455$ and $\rho_{d,z} = -0.82$



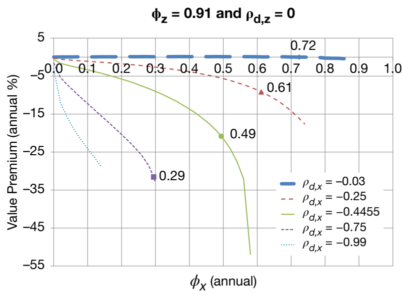
Graph C. $\phi_z = 0.91$ and $\rho_{d,z} = -0.41$



Graph D. $\phi_z = 0$ and $\rho_{d,z} = -0.82$



Graph E. $\phi_z = 0.91$ and $\rho_{d,z} = 0$



C. Wedge Case

This subsection discusses the results for the Wedge case, comparing them to those for the LW and Base cases and the data. While the Base case takes the

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

consumption process to be the market dividend process, the Wedge case instead calibrates its consumption process to aggregate consumption data. By allowing the process for consumption to differ from that for the market dividend, the Wedge case does a better job than the Base or LW cases matching the data value for the expected excess market return, and is able to deliver an excess market return volatility that is much closer to the data value than the very low value obtained by the Base case, though still not as close as the value delivered by the LW case: Table 3 shows that the volatility increases from 10.69 to 14.69 to 19.41 going from the Base case to the Wedge case to the data. The magnitude of the B_x function and the unconditional volatility of the price-of-risk variable x both increase going from the Base to the Wedge case, which causes the autocorrelation of the log market price–dividend ratio to decline, using equation (11), the low persistence of x relative to that of z^m for both cases, and the reasoning given for the high autocorrelation in the Base case. Given that the volatility of the log market price–dividend ratio is unchanged going from the Base to the Wedge case (see Table 3), equation (15) and log-linearizations show that a key driver of the higher log excess market return volatility in the Wedge case is this lower autocorrelation of the log market price–dividend ratio (see the appendices for details). Table 3 also shows that the Wedge case does a poorer job than the Base case matching the positive autocorrelation of the excess annual market return found in the data, delivering a value that is even more negative than the value obtained for the Base case.

Turning to the predictability results for future market excess log return and future changes in log market dividend in Tables 4 and 5, respectively, all the results for the Wedge case are qualitatively similar to those for the Base case. So just like the Base case, the Wedge case is also unable to generate the predictability of the future market excess log return found in the data, particularly at long horizons, using the log market price–dividend ratio. Focusing on the 10 M1 portfolio (which is long the extreme value decile and short the extreme growth decile), Table 6 shows that, as compared to the Base case, the Wedge case does a better job of matching the data unconditional mean and volatility of excess quarterly return, but does a similar job of matching the data Sharpe ratio for quarterly return and the value premium in CAPM alpha found in the data. The Wedge case's volatility of excess quarterly return for the 10 M1 portfolio is larger than that for the Base case, but comparable to that for the LW case, which means that it is lower than in the data.

In summary, the Wedge case shows that allowing a wedge between the market dividend and consumption processes as in the data helps generate return volatility for the market and the 10 M1 portfolio that is closer to the data, without destroying the value premium in CAPM alpha obtained in the Base case. However, the wedge does not remedy the inability of the log market price–dividend ratio to predict future market excess log returns in the Base case.

D. 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 Wedge cases. Like the Wedge case, the LRR-Vol case allows the consumption process to differ from the market dividend process. But unlike the Wedge case, the LRR-Vol case allows the volatility as well as the mean of consumption growth to be slowly mean-reverting and calibrated to the data.

1. Aggregate Moments

Table 3 shows that, by allowing the consumption process to have a different generating process from the market dividend, both the Wedge and LRR-Vol cases do a better job on the mean, but a poorer job on the volatility, of the excess market return than the LW case: both generate lower volatility than in the data, with the value for LRR-Vol case closer to the data value. While the Base and Wedge cases severely understate the volatility of the market price–dividend ratio, the LRR-Vol case has the opposite problem, producing a volatility of 0.480 that is much higher than the data value of 0.38. This occurs, despite the B_x function and σ_x being smaller in magnitude for the LRR-Vol case than the Wedge case, because the volatility of σ_t in the LRR-Vol case generates volatility in both the log market price–dividend ratio and the excess market return in Table 3, as equations (11) and (15) and some log-linearizations show.

While the autocorrelation of the excess log annual market return is a problem for the Wedge case, -0.27 versus 0.03 in the data, its value in the LRR-Vol case decreases in magnitude to -0.10 which is much closer to the data value. However, the higher excess log annual market return volatility in the LRR-Vol case can only explain about 34% of this decline in magnitude; the remainder is due to a reduction in the magnitude of the autocovariance. The volatility of σ_t in the LRR-Vol case is also a key driver of this reduction, as equations (11) and (15) and some log-linearizations show.

As would be expected given the high persistence of the σ_t process, equation (11), and the reasoning provided for the Base case, the autocorrelation of the price–dividend ratio is higher than in the Base and Wedge cases, taking a value that is even higher than the data value. In summary, the LRR-Vol case does a comparable job to LW and a better job than the Base and Wedge cases of matching the aggregate moments reported in Table 3.

2. Predictive Regressions

Tables 4 and 5 report predictive regression results and the last column of each reports results for the LRR-Vol case. Table 4 shows that the LRR-Vol case remedies the most glaring weakness of the Base and Wedge cases, namely their inability to generate the excess market log return predictability observed in the data using the log market price–dividend ratio. The LRR-Vol case is able to produce much more negative predictability coefficients and much larger R^2s for these regressions than those cases. For the 10 year return horizon regression, Panel A of Table 4 reports that the R^2 is 0.27 for the data and the LRR-Vol case, but only 0.02 or less for these other cases. The implication is that allowing consumption growth to have volatility that is slowly mean-reverting can help external habit models with fast-moving habit to generate the return predictability in the data.

Panel B of Table 4 makes it clear that the log market price–dividend ratio's predictive ability for future market excess log return is being driven by the persistence in the σ_t process. Of the drivers of the market price–dividend ratio, virtually all the predictive ability for future market excess return is coming from σ_t: the adjusted

R^2 for the 10-year return horizon regression is 0.33 for σ_t, but is counterfactually 0 for z_t^m as for the other three cases, and only 0.01 for x_t.

Equation (13) shows that time-series variation in the time-t log risk premia on market-dividend strips, and hence in the time-t market log risk premium, is driven by variation in x_t and σ_t in the LRR-Vol model, but is totally driven by variation in x_t in the other three cases. Given the chosen parameters for x_t and σ_t for the LRR-Vol case, 88.7% of the time-series variation in the time-t log risk premia on market-dividend strips is driven by time-series variation in x_t, while only 11.3% is driven by time-series variation in σ_t (see the appendix A for details). The reason that the LRR-Vol case can generate similar 10-year market excess log return predictability to that generated by the LW case despite σ_t driving such a small fraction of the variation in the time-t market log risk premium is the higher persistence of σ_t relative to that of x_t in the LW case (0.994 vs. 0.964 using quarterly data).

Panel A of Table 5 shows that the predictability of future market dividend growth using the log market price–dividend ratio is much reduced for the LRR-Vol case as compared to the Base and Wedge cases, so much so that the R^2s are lower than for LW and closer to the data R^2s which are indistinguishable from 0. Turning to the drivers of log market price–dividend ratio in Panel B of Table 5, x_t has similar predictive ability as in the Base and Wedge cases: small, but nonzero. As to be expected since $\text{cov}\left[z_t^m, \sigma_t\right] = 0$, σ_t has no predictive ability for future log market dividend growth. Using equation (11), this explains why, in Panel A, the adjusted R^2 using log market price–dividend ratio as the predictive variable is considerably lower than for the Base and Wedge cases.

3. Value Versus Growth Portfolios

The last column of Table 6 shows that the LRR-Vol case generates value premia in both expected excess quarterly return and CAPM quarterly return alpha that are higher than those obtained for the Base and Wedge cases, but still a little lower than those for the LW case: the expected excess quarterly return spreads between the extreme value and the extreme growth portfolio is 4.36% annualized which is quite close to the data value of 5.15% annualized when sorting on B/M, while the CAPM-alpha value spread is 4.01% annualized which is closer to the 4.67% annualized in the data for the B/M sort than the 6.43% annualized in the LW case. Like the Base and Wedge cases, the LRR-Vol case delivers a CAPM-alpha value spread that is smaller than the expected excess return value spread, which is consistent with the results for the B/M data sort but not the other two data sorts.

Graph A of Figure 1 shows that the unconditional expected excess quarterly return on the market-dividend strips for the LRR-Vol case is hump-shaped in maturity with the hump occurring at a maturity less than 6 years, same as for the Base and Wedge cases. This hump-shape delivers a value premium in expected excess quarterly return for the same reason that the hump-shape for the Base and Wedge cases also delivered such a premium: 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. The unconditional expected excess quarterly return on the market-dividend strips is hump-shaped in maturity because $B_x(n)$ plotted in Figure 2 as a function of strip maturity n is u-shaped for the LRR-Vol case just as it is for the Base and Wedge cases, and this result is illustrated in Graph D of Figure 3. For all n in Figure 2, the $B_x(n)$ curve for the LRR-Vol case is smaller in magnitude than the $B_x(n)$ curve for the Wedge case, but $\bar{x}$ for the LRR-Vol case is higher than for the Wedge case. Relative to the Wedge case, the lower $|B_x(n)|$ curve lowers the dividend strip expected excess return curve, while the higher $\bar{x}$ moves it up.

The shape of $B_\sigma(n)$ does not matter for the value premium in expected excess return, because the LRR-Vol case imposes $\rho_{d,w}=0$, and equation (13) shows that $\rho_{d,w}=0$ means that $B_\sigma(n)$ does not matter for unconditional mean excess market returns. Graph D of Figure 3 confirms this, showing that the $B_\sigma(n)\sigma_{d,w}$ term in equation (13) is 0 at all maturities for the LRR-Vol case.

The zero correlation between Δd_t and σ_t explains why, for the LRR-Vol case, the dividend strip expected excess return curve remains hump-shaped and is able to deliver value premia in both expected excess quarterly return and CAPM quarterly return alpha. So the role of the conditional consumption growth volatility process, σ_t, in the LRR-Vol model is similar to that of x_t in LW. In LW, the high persistence of x_t generates equity return predictability, while its zero conditional correlation with log aggregate consumption growth, Δd, leaves unaffected the value premium delivered by the conditional mean process for Δd assumed by LW. Similarly, the high persistence of σ_t in the LRR-Vol case generates equity return predictability, while its zero conditional correlation with Δd ensures that the value premium delivered by our habit model with fast-moving habit and homoscedastic Δd is not compromised.

LRR-Vol's excess-return volatility numbers for the two extreme deciles and the 10 M1 portfolio, though lower than the data values, are higher than the numbers produced by the Base and Wedge cases, and in the same ballpark as the numbers produced by the LW case. Averaging the absolute Sharpe ratio errors relative to the three data sorts, the LRR-Vol case does a poorer matching job relative to the LW case for the extreme deciles, but a much better job for the 10 M1 portfolio.

As with the LW case, the adjusted R^2 of the CAPM market model regression in the LRR-Vol case is slightly higher than the data adjusted-R^2 values for the growth decile, but much larger than the data adjusted-R^2 values for the value decile. While the LW case produces an adjusted R^2 for 10 M1 portfolio that is quite close to the adjusted-R^2 values for the three data sorts, the LRR-Vol case produces an even closer value, one that is also much lower and closer to the data values than is produced by the Base or Wedge case. Turning to the 2-factor HML regressions, both the LW and LRR-Vol cases do a disappointing job replicating the data along a number of dimensions. While both cases, like the Base and Wedge cases, deliver a 2-factor alpha for the 10 M1 portfolio that is close to 0, this alpha is more than 3% annualized for the data E/P and C/P sorts, and equal to -0.83% for the data B/M sort. For the two extreme deciles and the 10 M1 portfolio, the magnitudes of the HML γ coefficients are higher, and the adjusted R^2s are lower, for all three data sorts than for any of the cases, including LRR-Vol and LW.

4. Return Properties of Market-Dividend Strips

Recent empirical work by BBK examines the properties of short-horizon returns on two portfolios of S&P 500-dividend strips with average maturities of around 1 year. An interesting question is whether the LW case or any of our fast-

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

TABLE 7

Return Performance for van Binsbergen et al. Two Market-Dividend Strip Portfolios: Data

Table 7 shows the quarterly return performance of the S&P 500 and two portfolios of S&P 500-dividend strips from van Binsbergen et al. (2012). Strategy 1 is long a portfolio of short-maturity S&P 500-dividend strips, while Strategy 2 is long a portfolio of dividend steepeners; a steepener pays out all dividends received between two specified short-maturity horizons. While Strategy 2 has a longer average dividend maturity than Strategy 1, each strategy pays dividends whose average maturity is around 1 year. Panel A reports expected excess return, the volatility of excess return, and the unconditional Sharpe ratio. Panel B reports CAPM alpha, beta, and adjusted regression R^2. Panel C reports alpha, beta (the slope on the market factor), gamma (the slope on the HML factor), and adjusted regression R^2 from a 2-factor regression. The data span Apr. 1996–Sept. 2009. Returns are quarterly, but the results are annualized by multiplying the expected excess return and alphas by 4, the volatility of the excess return by $\sqrt{4}$, and the Sharpe ratio by $4/\sqrt{4}$.

	Strategy 1	Strategy 2	S&P 500
Panel A. $R^i - R^f$			
$E[R^i - R^f]$	8.09	6.53	3.48
$\sigma[R^i - R^f]$	19.74	22.60	18.33
$SHARPE_i$	0.41	0.29	0.19
Panel B. $R^i - R^f = \alpha_i + \beta_i(R^m - R^f) + \varepsilon^i$			
α_i	7.03	4.81	
β_i	0.30	0.49	
Adj R_i^2	0.08	0.16	
Panel C. $R^i - R^f = \alpha_i + \beta_i(R^m - R^f) + \gamma_i R^{\text{HML}} + \varepsilon^i$			
α_i	7.66	5.11	
β_i	0.28	0.48	
γ_i	−0.11	−0.05	
Adj R_i^2	0.09	0.16	

moving habit cases can replicate the empirical properties that BBK document for these strip portfolios. Table 7 shows the quarterly return performance of the S&P 500 and BBK's two portfolios of S&P 500-dividend strips over their data period from Apr. 1996 to Sept. 2009. Strategy 1 is long a portfolio of short-maturity S&P 500-dividend strips, while Strategy 2 is long a portfolio of dividend steepeners; a steepener pays out all dividends received between two specified short-maturity horizons. Strategy 2 has a longer average dividend maturity than Strategy 1. Panel A reports expected excess return, the volatility of excess return, and the unconditional Sharpe ratio, and shows that all three are higher for the two short-maturity strip strategies than for the S&P index itself, consistent with results in BBK for monthly returns. In Figure 1, comparing the strips with a maturity of 1 year to the market points on the left of the graphs shows that all our fast-moving habit cases are able to deliver these same three results for quarterly returns, while the LW case only delivers two of the three results (not the excess-return volatility result). Moreover, BBK find that neither the external habit model with slow-moving habit nor the long-run risk model can produce any of the three results.

When BBK uses the log annual price–dividend ratio of Strategy 1 to forecast its monthly return, and log annual price–dividend ratio for the S&P 500 to forecast its monthly return, they find a higher R^2 and larger slope coefficient in absolute value for the Strategy 1 regression than for the S&P regression. The LW case and all our fast-moving habit models are able to deliver an equivalent result. Figure 5 plots, as a function of maturity, the results from the regression of future excess log quarterly returns of market-dividend strips on their own log annual price–dividend

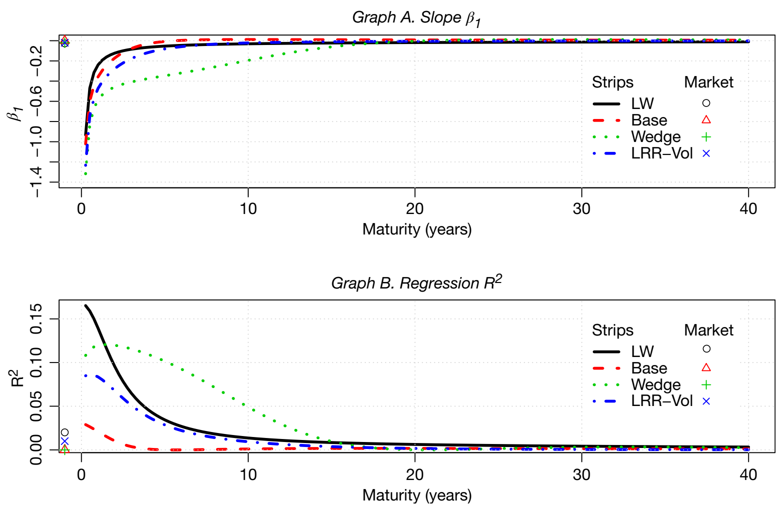
FIGURE 5

Predictability of Market and Market-Dividend Strips: Model

Figure 5 plots, as a function of maturity, the results from the regression of future excess log return of market-dividend strips on their own log price–dividend ratio today:

$$r^m_{n,t+1} - r^f = \beta_0 + \beta_1 \left(p^m_n - d^m\right)_t + \varepsilon_{n,t+1}.$$

The return is calculated as the quarterly return from holding the market-dividend strip with n quarters to maturity at the start of the quarter. The log price–dividend ratio $\left(p^m_n - d^m\right)_t \equiv \log\left(P^m_{n,t} / \sum_{\tau=0}^{3} D^m_{t-\tau}\right)$ is the annual log price–dividend ratio of the market-dividend strip with n quarters to maturity at the start of the return period. The regression coefficient β_1 and R^2 are reported in Graph A and Graph B, respectively. The points to the left of each of the graphs represent values for the regression of future excess log quarterly market return from time t to time $t+1$, r^m_{t+1}, on the log annual market price–dividend ratio at time t, $\left(p^m - d^m\right)_t \equiv \log\left(P^m_t / \sum_{\tau=0}^{3} D^m_{t-\tau}\right)$. For each graph, the circle to the left and the solid line are for the LW case, the triangle to the left and the dashed line are for the Base case, the plus to the left and the dotted line are for the Wedge case, and the cross to the left and dot-dashed line are for the LRR-Vol case.



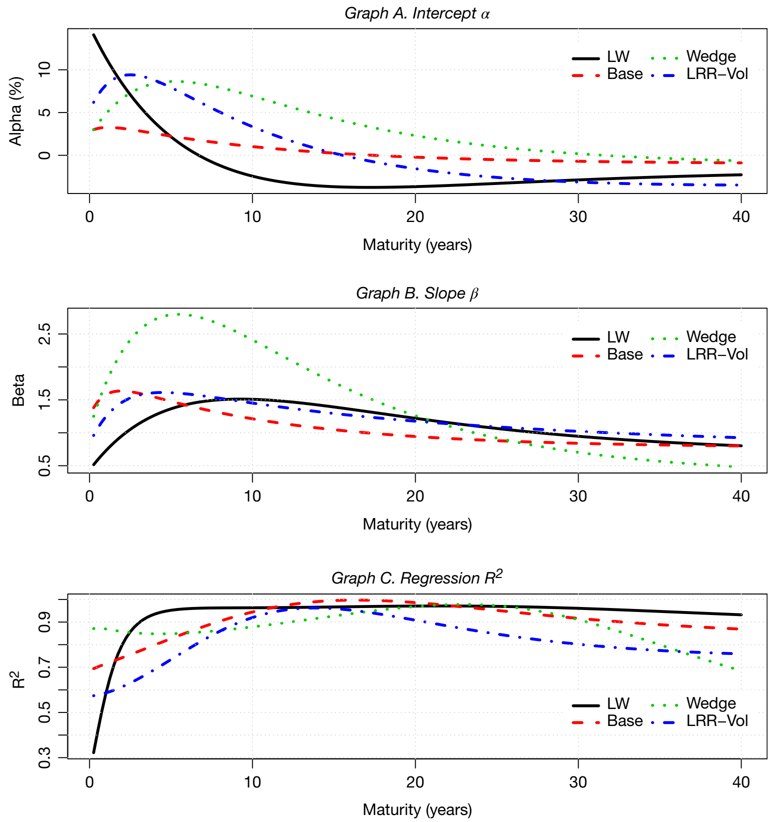
ratios today as a function of maturity. The regression coefficient and R^2 are reported in Graphs A and B, respectively. The points to the left of each of the graphs represent the values for the regression of future excess log quarterly market return on its log annual price–dividend ratio today. For the LW case and all our fast-moving habit cases, both the R^2 and the magnitude of the slope coefficient are much larger for the 1-year market-dividend strip (see Figure 5 at a maturity of 1 year) than for the market (see the points in Figure 5).

BBK also run CAPM market-model regressions for monthly returns on the two short-maturity S&P 500-dividend-strip portfolios and find very low CAPM betas of around 0.5. Panel B of Table 7 reports similar values for quarterly returns, though the CAPM beta for Strategy 1 is even lower at 0.3. Figure 6 plots CAPM market-model regression results for quarterly market-dividend strip returns as a function of maturity, for our four model cases. For all four model cases, the 1-year strips have higher CAPM betas and R^2s than those for BBK's two short-maturity S&P 500-dividend-strip portfolios, with the LW and LRR-Vol cases producing CAPM betas and R^2s closest to those for the data portfolios. From Table 7, the average quarterly CAPM alpha for the two BBK strategies is 5.92% annualized, while Figure 6 shows that all four cases produce quarterly CAPM alphas for the

FIGURE 6

CAPM Regressions for Market-Dividend Strips: Model

Figure 6 shows results from time-series regressions of excess market-dividend strip returns on excess market returns, as a function of maturity, for our four model cases. The intercept α, slope β, and R^2 are in Graph A, Graph B, and Graph C, respectively. In each graph, the solid line is the LW case, the dashed line is the Base case, the dotted line is the Wedge case, and the dot-dashed line is the LRR-Vol case. Returns are quarterly, but α is annualized by multiplying by 4.



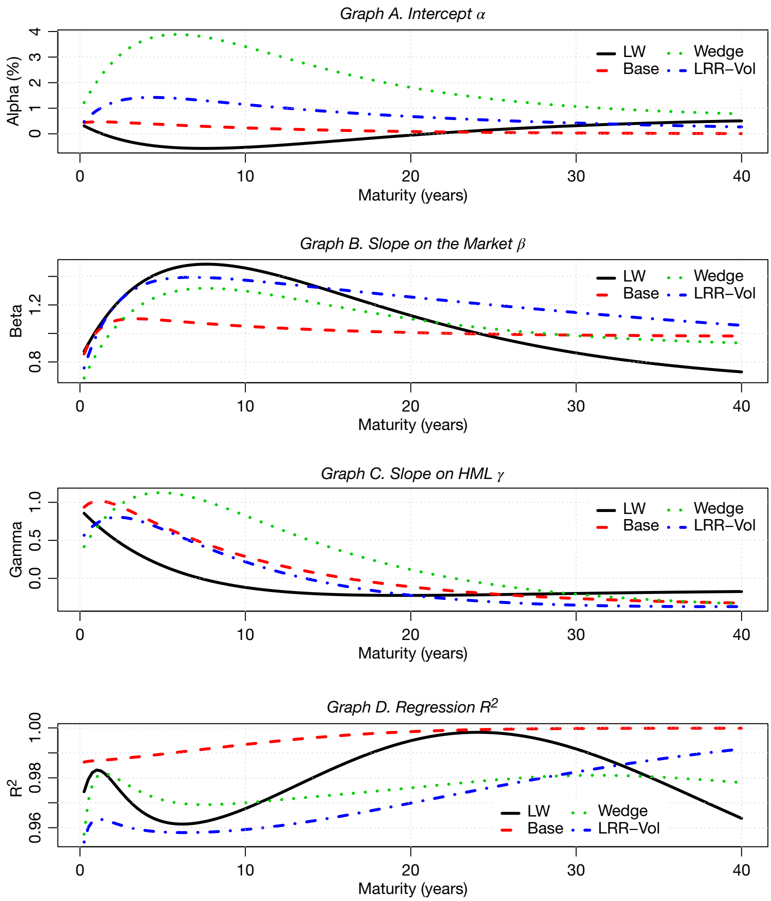
1-year market-dividend strip that are positive, with the alpha for the LRR-Vol case being closest to the alphas for the data strategies.

Panel C of Table 7 reports 2-factor market-model regressions for quarterly returns on BBK's two short-maturity S&P 500-dividend-strip portfolios, using Fama and French's HML portfolio as the second factor over BBK's data period, Apr. 1996 to Sept. 2009. The two BBK strategies have 2-factor alphas, market betas, and adjusted R^2s that are comparable to those for the CAPM market-model regressions, while the slope coefficient on HML, γ, is negative for both. Figure 7 shows 2-factor market-model regression results for quarterly market-dividend strip returns as a function of maturity, for our four model cases. Again, the second factor is the HML portfolio, which is constructed by going long the three extreme value deciles, equal-weighted, and short the three extreme growth deciles, also equal-weighted. For all four model cases, the 1-year strips have higher 2-factor betas and R^2s than those for BBK's 2 short-maturity S&P 500-dividend-strip portfolios, while

FIGURE 7

Two-Factor (Market and HML) Regressions for Market-Dividend Strips: Model

Figure 7 shows results from time-series regressions of excess market-dividend strip returns on excess market and HML returns, as a function of maturity, for our four model cases. The HML portfolio is constructed by going long the three extreme value deciles, equal-weighted, and short the three extreme growth deciles, also equal-weighted. The intercept α, slope on the market β, slope on HML γ, and R^2 are plotted in Graph A, Graph B, Graph C, and Graph D, respectively. In each graph, the solid line is the LW case, the dashed line is the Base case, the dotted line is the Wedge case, and the dot-dashed line is the LRR-Vol case. Returns are quarterly, but α is annualized by multiplying by 4.



γ is counterfactually positive for all four model cases. All four model cases produce 2-factor alphas for the 1-year market-dividend strip that are positive, but much smaller than the 2-factor alphas for the two data strategies, with the alpha for the LRR-Vol case being closer to the alphas for the data strategies than the alpha for the LW case.

Finally, BK present some new empirical evidence about the expected excess returns on market-dividend strips which they obtain using monthly returns on S&P 500 dividend-strip futures. They report that, in the United States, the expected monthly dividend-strip spot return in excess of the market return is increasing in

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

strip maturity going from 1-year strips to 5-year strips. According to Graph A in our Figure 1, this empirical result is much more consistent with the hump-shaped dividend strip expected excess return curve for the LRR-Vol case than the downward sloping curve in the LW case.

In sum, the results described in this subsection show that models with fast-moving habit can deliver some, but not all, of the empirical properties of market-dividend strips that have been recently documented. Moreover, the LRR-Vol model better fits the empirical properties than LW along several dimensions: excess return volatility is higher for 1-year market dividend strips than for the market in the data and the LRR-Vol model, but not in LW, while compared to LW, the LRR-Vol model delivers alphas for the CAPM and 2-factor market-model regressions that are much closer to the data values for the two BBK dividend-strip portfolios. However, both the LRR-Vol model and LW do a poor job matching the data values for the risk loadings and R^2s for these regressions. Since parameters for the LRR-Vol case are not chosen to match the recently documented empirical properties of market dividend strips, it is encouraging that the LRR-Vol case can match some empirical properties that LW cannot, and understandable that neither are able to match all of them.

5. Time-Series Behavior of the Implied x_t and σ_t Processes and Macroeconomic Variables

Variation in conditional expected equity returns in LW's model is driven by variation in the price-of-risk state variable x_t. Consequently, in LW's model, information variables that are able to forecast future equity returns empirically would be expected to covary with the price-of-risk state variable x_t. Variation in conditional expected equity returns in our LRR-Vol model is driven by variation in the price-of-risk state variable x_t and the consumption-volatility state variable σ_t. Consequently, in our LRR-Vol model, information variables that are able to forecast future equity returns empirically would be expected to covary with the price of risk state variable x_t and the consumption-volatility state variable σ_t.

Table 8 reports the results from regressing, over the period 1947 to 2002, three macroeconomic variables (Lustig and Van Nieuwerburgh's (2005) *my*; Lettau and Ludvigson's (2001) *cay*; and Piazzesi, Schneider, and Tuzel's (2007) *alpha*) on state variables imputed for the LW and LRR-Vol cases. For the LRR-Vol case, quarterly price–dividend ratios for the market, $\log\left(P_t^m/D_t^m\right)$, and value decile, $\log\left(P_t^v/D_t^v\right)$, are used to impute x_t and σ_t using log-linearizations of equation (11) applied to the market and the value decile, while for the LW case, $\log\left(P_t^m/D_t^m\right)$ is used to impute x_t using equation (11) applied to the market. Regression slopes, labeled as βs, are reported for *my*, *cay*, and *alpha*. For each variable, LW's x_t process is the independent variable in the top row, LRR-Vol's x_t and σ_t processes are the independent variables in the second and third rows, respectively, and LRR-Vol's x_t and σ_t processes are the independent variables of a joint regression in the last row. In addition to OLS t-statistics, Newey–West t-statistics for the maximum-likelihood-optimal lag length are also reported. The far-right column shows the adjusted R^2 for each row's regression.

TABLE 8

Implied x_t and σ_t Processes and Macroeconomic Variables

Table 8 shows the results of regressing three macroeconomic variables (*my*, *cay*, and *alpha*) on the state processes implied by Lettau–Wachter's (2007) model (LW) and our LRR-Vol model. Regression betas (βs) are reported for Lustig and Van Nieuwerburgh's (2005) *my*, Lettau and Ludvigson's (2001) *cay*, and Piazzesi et al.'s (2007) *alpha*. For each panel, LW's x_t process is the independent variable in the top row, LRR-Vol's x_t and σ_t processes are the independent variables in the second and third rows, respectively, while LRR-Vol's x_t and σ_t processes are the independent variables of a joint regression in the last row. All variables are standardized by subtracting their mean and dividing by their standard error. OLS and Newey–West t-statistics for the maximum-likelihood-optimal lag length are both reported. The right column shows the adjusted R^2 for each row's regression. The data covers the period 1947 to 2002. For the LRR-Vol case, quarterly log price–dividend ratios for the market ($\log(P_t^m/D_t^m)$) and value decile ($\log(P_t^v/D_t^v)$) are used to impute x_t and σ_t using log-linearizations of equation (11) applied to the market and the value decile, while for the LW case, $\log(P_t^m/D_t^m)$ is used to impute x_t using equation (11) applied to the market. The log price–dividend ratios are sampled at an annual frequency at the end of each year, since the imputation requires z_t^m, which is only available at an annual frequency from annual consumption data.

	LW			LRR-Vol						Adj
	x_t			x_t			σ_t			R^2
	β	t-Stat		β	t-Stat		β	t-Stat		
		OLS	NW		OLS	NW		OLS	NW	
	0.74	8.13	3.62							0.55
my				0.16	1.18	0.69				0.01
							−0.11	−0.82	−0.53	−0.01
				5.94	4.60	2.47	5.80	4.49	2.52	0.30
	0.60	5.52	4.03							0.35
cay				0.34	2.60	2.01				0.10
							−0.28	−2.15	−1.87	0.06
				6.63	5.74	13.71	6.32	5.47	13.02	0.44
	0.58	5.23	1.88							0.33
alpha				0.33	2.57	1.42				0.09
							−0.31	−2.36	−1.37	0.08
				3.08	2.20	1.15	2.76	1.97	1.08	0.17

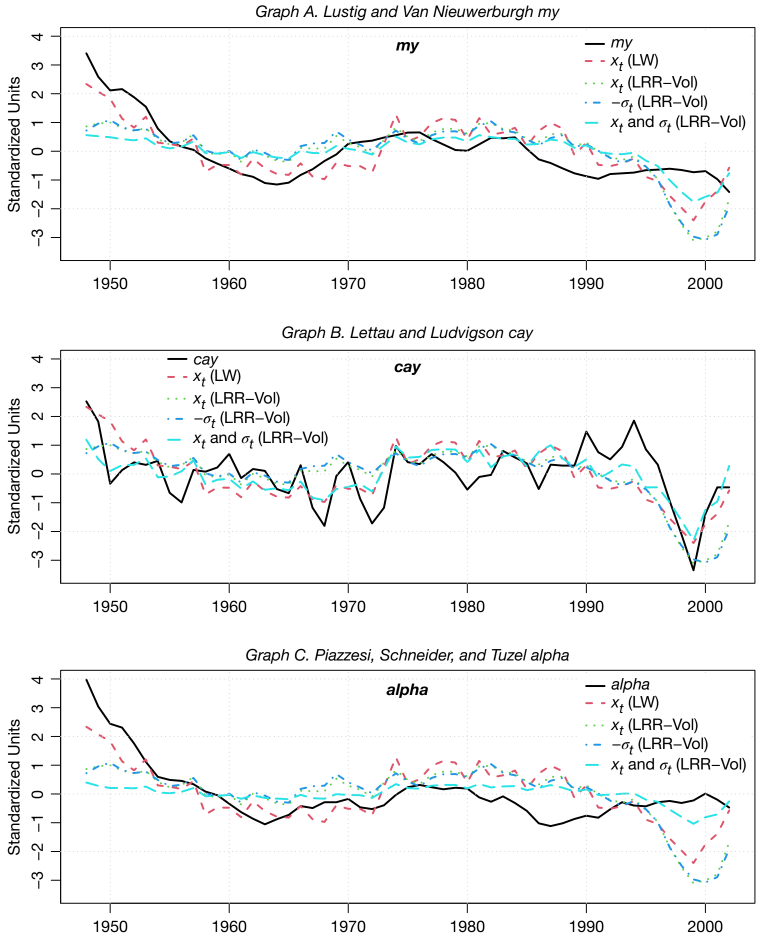
Table 8 shows that LRR-Vol's state variables x_t and σ_t together have strong forecasting power for *my* and *cay*, as does LW's state variable x_t. With *my* as the dependent variable, the adjusted R^2 is higher when LW's x_t is the independent variable than when LRR-Vol's x_t and σ_t are the independent variables, though in the latter regression, the adjusted R^2 is still 0.30. In contrast, with *cay* as the dependent variable, the adjusted R^2 is higher when LRR-Vol's x_t and σ_t are the independent variables than when LW's x_t is the independent variable: an adjusted R^2 of 0.44 for the former compared to 0.35 for the latter.

Figure 8 shows time-series graphs of the implied state variables from the LW model and our LRR-Vol model, the three macroeconomic variables, and forecasts of the three macroeconomic variables using the two implied state variables from the LW model. Graph A plots *my*, Graph B plots *cay*, and Graph C plots *alpha*. In all three graphs, the x_t process implied by the LW model is labeled "x_t (LW)," the x_t process implied by our LRR-Vol model is labeled "x_t (LRR-Vol)," the σ_t process implied by our LRR-Vol model is labeled "σ_t (LRR-Vol)," and the fitted value from regressing the graph's macroeconomic variable on x_t and σ_t simultaneously is labeled "x_t and σ_t (LRR-Vol)." In Figure 8, as in Table 8, all variables are standardized by subtracting their mean, and dividing by their volatility. Graph C of Figure 8 confirms one of the main results in Table 8: The best linear forecast of *cay* using LRR-Vol's x_t and σ_t tracks *cay* better than LW's x_t.

FIGURE 8

Implied x_t and σ_t Processes and Macroeconomic Variables

Figure 8 shows the time-series graphs of state variables from LW and our LRR-Vol model, three macroeconomic variables, and forecasts of the three macroeconomic variables using two state variables from the LRR-Vol model. Graph A plots Lustig and Van Nieuwerburgh *my*, Graph B plots Lettau and Ludvigson *cay*, and Graph C plots Piazzesi et al. *alpha*. Each is represented by a solid line. In all graphs, the x_t process implied by the LW model is labeled "x_t (LW)" and represented by the short-dashed line, the x_t process implied by the LRR-Vol model is labeled "x_t (LRR-Vol)" and represented by the dotted line, the negative of the σ_t process implied by our LRR-Vol model is labeled "$-\sigma_t$ (LRR-Vol)" and represented by the dot-dashed line, and the fitted value from regressing the graph's macroeconomic variable on x_t and σ_t simultaneously is labeled "x_t and σ_t (LRR-Vol)" and represented by the long-dashed line. All variables are standardized by subtracting their mean, and dividing by their volatility. The data covers the period 1947 to 2002. For the LRR-Vol case, quarterly price–dividend ratios for the market ($\log(P_t^m/D_t^m)$) and value decile ($\log(P_t^v/D_t^v)$) are used to impute x_t and σ_t using log linearizations of equation (11) applied to the market and the value decile, while for the LW case, $\log(P_t^m/D_t^m)$ is used to impute x_t using equation (11) applied to the market. The price–dividend ratios are sampled at an annual frequency at the end of each year.



V. Conclusion

This article finds that the external-habit model of Campbell and Cochrane (1999) can generate the value premium in both CAPM alpha and expected excess return seen in the data, as long as the persistence of the consumption surplus is sufficiently low. In contrast, Lettau and Wachter (2007) find that when it is highly persistent as in Campbell–Cochrane (by assuming that the price of risk is highly persistent), the external-habit model generates, counterfactually, a growth premium in expected excess return.

Recent micro evidence by Brunnermeier and Nagel (2006) and Ravina (2019) rules out slow-moving habit and suggests that the persistence of the consumption surplus is likely to be quite low. Moreover, the high persistence assumed by Lettau–Wachter's specification implies that the contribution to log habit of log consumption from more than 5 years ago is almost 50%, which seems too high. We choose a value for this persistence that is more in line with the micro evidence, and sufficiently low that the most recent 2 years of log consumption contribute over 98% of all past consumption to log habit, which seems more reasonable.

In our specification, expected consumption growth 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 market price–dividend ratio that exhibits the high autocorrelation found in the data, despite the low persistence of our price of risk. When consumption growth is homoscedastic, fast-moving habit has difficulty replicating the ability of the price–dividend ratio to predict long-horizon market return in the data. However, allowing the conditional volatility of consumption growth to be highly persistent, we obtain long-horizon market return predictability of a magnitude much closer to the data, without sacrificing the value premium. Fast-moving habit also delivers several empirical properties of market-dividend strips documented in van Binsbergen et al. (2012) and van Binsbergen and Koijen (2017). Overall, 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.

Appendix A. Correlation of ε^x and ε^z

To ensure the covariance matrix of (ε^d, ε^z, ε^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: i) regress ε^d on ε^z, yielding $\varepsilon^d = \beta_{d,z}\varepsilon^z + u^d$, where $\rho\left[\varepsilon^z, u^d\right] = 0$; and ii) regress ε^x on ε^d, yielding $\varepsilon^x = \beta_{x,d}\varepsilon^d + u^x$, where $\rho\left[\varepsilon^d, u^x\right] = 0$. The following expression can be derived:

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

When $\rho\left[\varepsilon^d, \varepsilon^x\right] = -0.99$, 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 equation (A-1) with $\rho[u^x, \varepsilon^z] = 0$ to calculate $\sigma[\varepsilon^x, \varepsilon^z]$. Notice this specification has the attractive property that when $\sigma\left[\varepsilon^d, \varepsilon^z\right]$ is set equal to 0, $\sigma[\varepsilon^x, \varepsilon^z]$ is set equal to 0 as well. With $\rho\left[\varepsilon^d, \varepsilon^z\right]$ set equal to -0.82, the implied value for $\rho[\varepsilon^x, \varepsilon^z]$ is 0.81.

https://doi.org/10.1017/S0022109023000212 Published online by Cambridge University Press

Appendix B. Distinguishing Between Consumption and Dividends

Since LW calibrate their dividend/consumption process to U.S. dividend data, we keep the joint $\{z^m, \Delta d^m\}$ process unchanged from the Base case for the Wedge and LRR-Vol cases (i.e., ϕ_z, σ_z, σ_m, g^m, $\rho_{m,z}$ remain the same). We match the following to LW's data moments:

$$\sigma^2[\varepsilon^m] = (\delta^m)^2\sigma_d^2 + \sigma_u^2 + 2\delta^m\sigma_{d,u}, \tag{B-1}$$

$$\sigma\left[\varepsilon_{t+1}^m, \varepsilon_{t+1}^z\right] = \delta^m\sigma_{d,z} + \sigma_{u,z}. \tag{B-2}$$

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

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:

$$\rho\left[\varepsilon_{t+1}^m, \sigma_t\varepsilon_{t+1}^d\right] = \frac{\left(\delta^m\sigma_d^2 + \sigma_{d,u}\right)\overline{\sigma}}{\sigma_m\sigma_d\sqrt{E\left[\sigma_t^2\right]}}. \tag{B-3}$$

Simulations suggest that the annual and quarterly correlations are very similar, at least for the range of parameter values we consider, so we focus on the quarterly number because its expression is much simpler. Since the x process is positive in CC, we instead chose a larger correlation than in the data, 0.82 at a quarterly frequency, for which the price–dividend ratio converges for a range of $\bar{x} > 0$.

Using the methods of Stambaugh (1997) and Lynch and Wachter (2013), 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:

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

To ensure that the covariance matrix of $\left(\varepsilon^d,\varepsilon^z,\varepsilon^x,\varepsilon^u\right)$ is positive definite, $\sigma[\varepsilon^x,\varepsilon^u]$ is calculated similarly to $\sigma[\varepsilon^x,\varepsilon^z]$ using

$$\text{(B-5)}\qquad \begin{aligned}\sigma[\varepsilon^x,\varepsilon^u] &= \sigma\left[\beta_{x,d}\beta_{d,z}\varepsilon^u + \beta_{x,d}u^d + u^x,\varepsilon^u\right]\\ &= \left(\rho\left[\varepsilon^d,\varepsilon^x\right]\rho\left[\varepsilon^d,\varepsilon^u\right] + \left[1-\rho\left[\varepsilon^d,\varepsilon^x\right]^2\right]^{\frac{1}{2}}\rho\left[u^x,\varepsilon^u\right]\right)\sigma_z\sigma_x\end{aligned}$$

with $\rho[u^x,\varepsilon^u]$ set equal to 0. Typically in the literature (e.g., Abel (1999)), 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.

For the Wedge case, we set $\overline{\sigma}=1$ and $\sigma_w=0$, which implies $E\left[\sigma_t^2\right]=1$. Given a δ^m value and $\rho\left[\varepsilon^d,\varepsilon^z\right]=\rho[\varepsilon^m,\varepsilon^z]$, the system of equations defined in equations (B-1)–(B-4) 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.

Appendix C. Making the Conditional Volatility of Log Consumption Growth Stochastic

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 (2007). 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 three 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 case (i.e., $E\left[\left(\sigma_t\varepsilon_{t+1}^d\right)^2\right]$ in the LRR-Vol case equals $E\left[\left(\varepsilon_{t+1}^d\right)^2\right]$ in the Wedge case). Combining this with the four moment conditions in equations (B-1)–(B-4), 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 $\overline{\sigma}$ 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 case (i.e., $\rho\left[\sigma_t\varepsilon_{t+1}^d,\varepsilon_{t+1}^z\right]$ in the LRR-Vol case equals $\rho\left[\varepsilon_{t+1}^d,\varepsilon_{t+1}^z\right]$ in the Wedge case). Finally, we follow Bansal and Yaron (2004) and impose that ε^w is uncorrelated with all other shocks.

Appendix D. Variance Decomposition of Equation (13) for LRR-Vol Case

From equation (13), the log risk premium on a market-dividend strip is equal to a constant that depends on the strip's maturity times $[\overline{\sigma}x_t+\overline{x}(\sigma_t-\overline{\sigma})]$, which means that the variance decomposition of the log risk premium on a market-dividend strip is the same for all strip maturities and equal to

$$\begin{aligned}\mathrm{Var}[\overline{\sigma}x_t + \overline{x}(\sigma_t - \overline{\sigma})] &= \overline{\sigma}^2\mathrm{Var}[x_t] + \overline{x}^2\mathrm{Var}[\sigma_t] \\ &= \frac{(\overline{\sigma}\sigma_x)^2}{1-\phi_x^2} + \frac{(\overline{x}\sigma_w)^2}{1-\phi_\sigma^2} \\ &= \underbrace{\frac{(0.918\times 0.29)^2}{1-\left(0.14^{0.25}\right)^2}}_{88.7\%} + \underbrace{\frac{(0.365\times 0.037)^2}{1-0.994^2}}_{11.3\%}.\end{aligned}$$

Appendix E. Variance of Log Annual Market Return

From equation (15), the log quarterly market return from time t to $t+1$ is

$$\begin{aligned}\log\left(R_{t+1}^m\right) &= \log\left(\left(\frac{P_{t+1}^m/D_{t+1}^m + 1}{P_t^m/D_t^m}\right)\left(\frac{D_{t+1}^m}{D_t^m}\right)\right) \\ &\approx pd_{t+1}^m - pd_t^m + \Delta d_{t+1}^m,\end{aligned}$$

where $pd_t^m \equiv \log\left(P_t^m/D_t^m\right)$. The log annual market return from time t to $t+4$ is $\log\left(\prod_{\tau=1}^4 R_{t+\tau}^m\right) = \sum_{\tau=1}^4 \log\left(R_{t+\tau}^m\right)$, and its variance is

$$\begin{aligned}\sigma^2\left[\sum_{\tau=1}^4 \log\left(R_{t+\tau}^m\right)\right] &= \sigma^2\left[pd_{t+4}^m - pd_t^m + \sum_{\tau=1}^4 \Delta d_{t+\tau}^m\right] \\ &\approx 2\sigma^2\left[pd_t^m\right].\left(1-\rho\left[pd_{t+4}^m, pd_t^m\right]\right) \\ &\quad + \sigma^2\left[\sum_{\tau=1}^4 \Delta d_{t+\tau}^m\right] + 2\sigma\left[pd_{t+4}^m - pd_t^m, \sum_{\tau=1}^4 \Delta d_{t+\tau}^m\right].\end{aligned}$$

Supplementary Material

The supplementary material for this article can be found at http://doi.org/10.1017/S0022109023000212.

References

Abel, A. B. "Asset Prices Under Habit Formation and Catching Up with the Joneses." *American Economic Review*, 80 (1990), 38–42.

Abel, A. B. "Risk Premia and Term Premia in General Equilibrium." *Journal of Monetary Economics*, 43 (1999), 3–33.

Bansal, R.; D. Kiku; and A. Yaron. "Risks for the Long Run: Estimation and Inference." Working Paper, Duke University and University of Pennsylvania (2007).

Bansal, R.; D. Kiku; and A. Yaron. "Risks for the Long Run: Estimation with Time Aggregation." *Journal of Monetary Economics*, 82 (2016), 52–69.

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

Bekaert, G., and E. Engstrom. "Asset Return Dynamics Under Bad Environment-Good Environment Fundamentals." *Journal of Political Economy*, 125 (2017), 713–760.

Bekaert, G.; E. Engstrom; and S. Grenadier. "Stock and Bond Returns with Moody Investors." *Journal of Empirical Finance*, 17 (2010), 867–894.

Bhamra, H. S., and R. Uppal. "Asset Prices with Heterogeneity in Preferences and Beliefs." *Review of Financial Studies*, 27 (2014), 519–580.

van Binsbergen, J. H.; M. W. Brandt; and R. S. J. Koijen. "On the Timing and Pricing of Dividends." *American Economic Review*, 102 (2012), 1596–1618.

van Binsbergen, J. H., and R. S. J. Koijen. "The Term Structure of Returns: Facts and Theory." *Journal of Financial Economics*, 124 (2017), 1–21.

Brunnermeier, M., and S. Nagel. "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation." Working Paper, Princeton University and Stanford University (2006).

Brunnermeier, M., and S. Nagel. "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation." *American Economic Review*, 98 (2008), 713–736.

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

Chabakauri, G. "Dynamic Equilibrium with Two Stocks, Heterogeneous Investors, and Portfolio Constraints." *Review of Financial Studies*, 26 (2013), 3104–3141.

Chabakauri, G. "Asset Pricing with Heterogeneous Preferences, Beliefs, and Portfolio Constraints." *Journal of Monetary Economics*, 75 (2015), 21–34.

Chen, H. "Do Cash Flows of Growth Stocks Really Grow Faster?" *Journal of Finance* 72 (2017), 2279–2330.

Constantinides, G. M. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, 98 (1990), 519–43.

Croce, M. M.; M. Lettau; and S. C. Ludvigson. "Investor Information, Long-Run Risk, and the Duration of Risky Cash-Flows." *Review of Financial Studies*, 28 (2015), 706–742.

Da, Z. "Cash Flow, Consumption Risk, and the Cross-Section of Stock Returns." *Journal of Finance*, 64 (2009), 923–956.

Dechow, P.; R. Sloan; and M. Soliman. "Implied Equity Duration: A New Measure of Equity Security Risk." *Review of Accounting Studies*, 9 (2004), 197–228.

Dynan, K. E. "Habit Formation in Consumer Preferences: Evidence from Panel Data." *American Economic Review*, 90 (2000), 391–406.

Garleanu, N., and S. Panageas. "Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing." *Journal of Political Economy*, 123 (2015), 670–685.

Gomez, M. "Asset Prices and Wealth Inequality." Working Paper, Columbia University (2022).

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

Kiku, D. "Is the Value Premium a Puzzle?" Working Paper, University of Pennsylvania (2007).

Lettau, M., and S. C. Ludvigson. "Consumption, Aggregate Wealth, and Expected Stock Returns." *Journal of Finance*, 56 (2001), 815–850.

Lettau, M., and S. C. Ludvigson. "Expected Returns and Expected Dividend Growth." *Journal of Financial Economics*, 76 (2005), 583–626.

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

Lustig, H., and S. Van Nieuwerburgh. "Housing Collateral, Consumption Insurance and Risk Premia: An Empirical Perspective." *Journal of Finance*, 60 (2005), 1167–1219.

Lynch, A. W., and J. Wachter. "Using Samples of Unequal Length in Generalized Method of Moments Estimation." *Journal of Financial and Quantitative Analysis*, 48 (2013), 277–307.

Piazzesi, M.; M. Schneider; and S. Tuzel. "Housing, Consumption, and Asset Pricing." *Journal of Financial Economics*, 83 (2007), 531–569.

Ravina, E. "Habit Persistence and Keeping Up with the Joneses: Evidence from Micro Data." Working Paper, Federal Reserve Bank of Chicago (2019).

Santos, T., and P. Veronesi. "Habit Formation, the Cross Section of Stock Returns and the Cash-Flow Risk Puzzle." *Journal of Financial Economics*, 98 (2010), 385–413.

Schorfheide, F.; D. Song; and A. Yaron. "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach." *Econometrica*, 86 (2018), 617–654.

Stambaugh, R. "Analyzing Investments Whose Histories Differ in Length." *Journal of Financial Economics*, 45 (1997), 285–331.

Sundaresan, S. "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth." *Review of Financial Studies*, 2 (1989), 73–89.