Asset Pricing Implications from Wealthy Shareholder Consumption and Net Payout

Robert S. Goldstein\textsuperscript{2} \hspace{1cm} Fan Yang\textsuperscript{3}

This Version: May 3, 2015

\textsuperscript{1}We thank Hengjie Ai, David Bates, Frederico Belo, Pierre Collin-Dufresne, Darrell Duffie, Bjorn Eraker, Michael Johannes, Xiaoji Lin, Toby Moskowitz, Annette Vissing-Jorgensen and Stan Zin for their many helpful comments. All remaining errors are our own.
\textsuperscript{2}University of Minnesota and NBER, golds144@umn.edu
\textsuperscript{3}University of Hong Kong, fanyang@hku.hk
Abstract

This paper investigates the consumption-based asset pricing implications of a model calibrated to i) the consumption process of wealthy investors, and ii) net payout, because aggregate consumption and cash dividend processes significantly underestimate the risks that wealthy investors and corporations face. We generalize the standard Epstein/Zin framework to permit agents to exhibit increasing aversion to those risks associated with larger wealth losses. By matching the empirically observed downward sloping term structures of consumption and net payout volatilities, our framework predicts that longer-duration assets are less risky, consistent with many empirical findings.
1 Introduction

This paper investigates the consumption-based asset pricing implications of a model calibrated to i) the consumption process of wealthy investors, and ii) net payout, defined as the sum of dividends, interest and net repurchases of equity and debt. It also generalizes the standard Epstein/Zin framework to permit agents to exhibit increasing aversion to those risks associated with larger wealth losses. Despite moderate levels of risk aversion, the model generates high and volatile equity returns, low and smooth risk free rates, and captures many other salient features of asset prices well. In particular, by matching the empirically observed downward sloping term structures of consumption and net payout volatilities, our framework predicts that longer-duration assets are less risky, generating low long-horizon discount rates (e.g., Binsgergen et al (2012), Giglio et al (2014)), and a decreasing (in absolute value) term structure of Sharpe ratios for straddle returns (e.g., Dew-Becker et al (2014), Andries et al (2014)). We use the remainder of the introduction to motivate the underlying assumptions of our model and to argue why we feel they are important.

This paper builds on the literature which argues that there is little economic justification for interpreting the aggregate consumption process as an optimal decision of a “representative agent”, because the underlying assumptions required for aggregation to be valid (e.g., complete markets, participation in asset markets by all households) do not hold in practice. These facts bring into question the validity of the so-called “equity premium puzzle” of Mehra and Prescott (1985), which is mostly driven by the smoothness of the aggregate consumption process. Indeed, this puzzle is significantly attenuated by the findings of Mankiw and Zeldes

\footnote{Many authors have argued against the representative agent construction (e.g., Hartley (1997), Chang, Kim, and Schorfheide (2011)). For example, Kirman (1992) concludes that the representative agent framework is “both unjustified and leads to conclusions which are usually misleading and often wrong.”}

\footnote{More correctly, the equity premium puzzle is due to both the low volatility of the aggregate consumption process and its low correlation with stock returns.}
(1991), who report that the consumption processes of stockholders are both more volatile and more correlated with stock returns. An important implication is that any model which interprets aggregate consumption as an optimal decision of a representative agent is forced to introduce additional sources of risk to explain the observed equity premium that may not accurately reflect the risks that actual shareholders face. Indeed, often these exaggerated sources of risk are imposed at longer horizons, generating counterfactual implications for the relative riskiness of short-horizon versus long-horizon assets.

We emphasize that researchers are free to aggregate the consumption processes of any subset of households (so long as each satisfies its first-order optimality conditions, among other conditions) in order to identify a stochastic discount factor (SDF). In this paper we focus on the subset of wealthy households who invest significantly in the stock market. This choice is motivated by our belief that these households are closer to satisfying their first order conditions than are poorer households, and are more likely to be subject to smaller (in proportion to their wealth) idiosyncratic shocks to wealth and income. Moreover, we see little justification for using total aggregate consumption (which combines the consumption of stockholders and non-stockholders alike) to identify a stochastic discount factor.

There is strong empirical support for limiting aggregation to the subset of wealthy investors to identify a SDF. Building on the findings of Mankiw and Zeldes (1991), Malloy, Moscowitz and Vissing-Jorgensen (MMV, 2009) investigate micro-level household data from the Consumer Expenditure Survey over the period 1982-2004 and find that the consumption

---

3We define “wealthy investors” as the top one-third wealthiest investors in the data set of Malloy, Moscowitz and Vissing-Jorgensen. These households held at least $18,955 (in 1982 dollars) in stocks in the first half of their sample, and $39,813 in the second half of their sample. That is, most of these agents were not all that wealthy.

4We also suspect that focusing on wealthy investors reduces the sensitivity to model misspecification. For example, if all agents are subject to a subsistence level, a misspecification of agent’s preferences which ignores this aspect will not materially impact the SDF generated from wealthy investor consumption, but would materially impact the SDF implied from agents whose endowment is barely above subsistence.
process of wealthy investors can explain the observed equity premium and value premium with a risk aversion coefficient as low as six. In contrast, they find that the aggregate consumption process requires risk aversion levels on the order of 50-100 to explain asset returns. These results are not too surprising given that, using the data from MMV, we estimate consumption volatility of wealthy investors to be 12.5%, compared to approximately 1% aggregate consumption volatility over this same (“Great Moderation”) time period.

Although focusing on a subset of households with such volatile consumption paths allows us to capture a large equity premium with relatively small levels of risk aversion, it is not at all obvious a priori that the SDF produced from the consumption processes of wealthy households can explain other salient features of asset prices. For example, while the precautionary savings term associated with the risk free rate is small for most preference specifications when calibrated from a consumption process with 2.5% volatility, this term becomes large (and potentially very volatile) when calibrated from a consumption process with 12.5% volatility, which in turn may make it difficult to explain low interest rate volatility. The purpose of this paper is to demonstrate that combining our proposed utility function with the consumption process of wealthy investors produces an SDF that captures many salient features of asset prices well.

Another departure that this paper makes from much of the extant literature is to specify corporate cash flows using the net payout process rather than the cash dividend process. We do so because there is a growing literature that has questioned the practicality of using the cash dividend process to price stocks.\(^5\) Indeed, it is generally accepted that firms smooth their cash dividend process through several channels, including equity repurchases, equity issuances and changes to their capital structure.\(^6\) Using a variance decomposition analysis,

\(^6\) See, for example, Boudkoukh et al. (2007), Chen (2009), Chen and Priestly (2012).
Larrain and Yogo (LY, 2008) conclude that the majority of variation in asset prices is due to changes in expected cash flows rather than to changes in discount rates, and that assuming a constant discount rate is a reasonable first-order approximation. That is, their findings basically resolve the “excess volatility puzzle” of Shiller (1981) and LeRoy and Porter (1981).

Building on this literature, we argue that the cash dividend dynamics specified by most of the extant literature cannot be used to price stocks because their specifications ignore equity repurchases and issuances. More accurately, we demonstrate that in order to price the claim to stocks correctly, one must account for the time variation in the outstanding share process. In particular, we use an argument akin to Miller and Modigliani (1961) to demonstrate that the parameters controlling the cash dividend process have no impact on the prices of corporate securities, whereas the parameters controlling the net payout process are paramount for pricing these securities. Indeed, we show that an all-equity firm can in fact commit to any cash dividend policy, even one which is perfectly deterministic. The implication is that the volatility of the cash dividend process itself may have little to say about the risks associated with the firm.

In addition to attempting to explain large and volatile stock returns, and a low and smooth risk free rate, the literature has also attempted to capture other empirical regularities observed in asset prices. For example, the price dividend ratio tends to be positively correlated with expected consumption growth and negatively correlated with consumption volatility. This finding is difficult for standard preferences (i.e., constant relative risk aversion) to capture, and has motivated the use of utility functions that exhibit preference for early resolution of uncertainty (Bansal and Yaron (BY, 2004)). To capture this observation, we follow BY’s lead by assuming that wealthy shareholders have recursive preferences (Kreps and Porteus (1978), Epstein and Zin (1989)). However, in order to capture the notion that
agents may exhibit increasing aversion to those risks associated with larger wealth losses, we generalize the standard framework by characterizing the agent with two (rather than one) risk aversion coefficients. Such preferences are similar in spirit to the “crash aversion” preferences of Bates (2008), and therefore may provide a better framework for capturing option prices compared to standard recursive preferences.

Additional empirical moments that we match include the (mostly) downward sloping term structures of volatilities for both consumption (Beeler and Campbell (2013)) and corporate cash flows. This contrasts with the assumptions made in the long-run risk literature, where upward sloping term structures are required in order to explain the equity premium. One important mechanism that drives the downward sloping term structure of consumption volatility is that large drops in consumption (e.g., depressions) are associated with low levels of both employment and capital utilization. Eventually, labor and capital get reallocated, and there is a recovery (Gourio (2009)). Indeed, Nakamura et al (2013) estimate that approximately one-half of the consumption drop associated with disasters is subsequently reversed. In other words, there is a strong negative correlation between shocks to consumption growth and shocks to long-run expected consumption growth, which therefore provides a natural hedge to consumption risk over long horizons, making long-horizon consumption less volatile than short-horizon consumption.\footnote{Although we do not investigate, we suspect that if we model “value firms” as those with short duration cash flows and “growth firms” as those with long-duration cash flows, then our SDF will be able to explain a “value premium”. In this sense, our framework captures some of the intuition implicit in the exogenous SDF economy proposed by Lettau and Wachter (2007).}

We demonstrate that calibrating our model to match downward sloping term structures of consumption volatility tend to generate downward sloping term structures of expected returns and volatilities for risky cash flow strips, consistent with the findings of Giglio et al (2013) and Binsbergen et al (2012). Intuitively, downward sloping term structures of
consumption volatility imply that consumption in the long run is not so risky. As such, claims to long horizon cash flows will command relatively small excess returns. In addition, our model generates downward sloping term structure of Sharpe ratios for straddle returns, consistent with Dew-Becker et al (2014) and Andries et al (2014).

Our framework builds upon several strands of literature. Mankiw (1986) and Constantinides and Duffie (1996) provide theoretical arguments for why idiosyncratic labor shocks may dramatically impact stock prices, and hence, for why representative agent frameworks calibrated to aggregate consumption may be inappropriate. Brav, Constantinides and Geczy (2002) and Constantinides and Ghosh (CG, 2013) provide empirical evidence that cross sectional variation in consumption across shareholders may be a crucial state variable for explaining asset returns.\(^8\) Although related, this approach differs significantly from ours, since we focus on the consumption of a representative wealthy agent, and hence there is no notion of a cross-section. Interestingly, MMV find that their low estimates for risk aversion coefficients are not due to cross sectional variation among shareholders. Zhang (2014) shows that the pricing kernel generated from the consumption processes of US shareholders explains returns in those foreign markets for which US shareholders invest heavily in.

Other approaches that do not use a representative agent framework include Basak and Cuoco (1998) and Guvenen (2009), who specify two types of agents: those that do and do not participate in the stock market.\(^9\) These papers show how the existence of a large fraction of non-participants can lead to a large equity premium. While our approach is less elegant, it has the practical advantage of not needing to impose non-participation for non-wealthy agents in order to identify a pricing kernel. It also allows us to match the empirically

\(^8\)Note that CG assume aggregate consumption follows an iid process. As such, it is not clear if their framework can generate a sufficiently large equity premium if it were calibrated to match the empirically observed downward sloping term structure of aggregate consumption volatility.

\(^9\)See also Gomes and Michaelides (2008).
observed downward sloping term structure of consumption volatility, which generates many of asset pricing moments that we focus on.

Other empirical work that puts into question the practice of aggregating consumption emphasizes the importance of idiosyncratic labor income shocks, especially during recessions. For example, Davis and von Wachter (2011) find that agents who lose their job during recessions tend to lose approximately three years worth of labor income. Kahn (2010) shows that college graduates who enter the labor force during recessions may lose up to 15% of their lifelong wages. Guvenen, Ozkan and Song (2012) show that cross sectional labor income shocks are highly skewed during recessions. We suspect that including these features would help rationalize the consumption and investment decisions of non-wealthy agents (in particular, their low levels of investment in stocks). We emphasize, however, that regardless of the explanation for limited participation, this has no impact on our estimated SDF, which is uniquely determined by the consumption process of wealthy investors and their preferences.

We introduce recursive preferences with two risk-aversion parameters in order to capture an empirical result implicit in option prices: Starting with Pan (2002), several papers have concluded that agents appear to be more risk averse to market crashes than to small, diffusive risks. (e.g, Eraker (2004), Broadie, Chernov and Johannes (2007)). To capture this intuition, Bates (2008) introduces “crash-risk” by directly specifying diffusive and jump sources of risk with different levels of risk aversion. In contrast, our preferences generate a level of risk aversion that increases with the size of the consumption loss without directly specifying source-dependent risk aversion. Indeed, in our framework jumps associated with small changes in consumption are subject to the same level of risk aversion as are diffusions.

While our paper is consumption-based, it has important implications for general equilibrium frameworks as well. For example, because Kaltenbrunner and Lochstoer (2010)
calibrate their GE model to match aggregate consumption volatility, they find that if technology shocks are transient (which in turn generates a downward sloping term structure of consumption volatility), then the only way to generate a sizeable market price of risk is to specify an extremely low value for the parameter which controls the elasticity of intertemporal substitution (EIS). However, a low EIS parameter implies counterfactually high interest rate volatility. In contrast, our framework emphasizes that if a model with transient technology shocks is calibrated to be consistent with the consumption volatility of wealthy investors, then it can generate large market prices of risk in spite of a high parametrization for the EIS parameter. This in turn implies that the model-implied interest rate volatility can be made consistent with observation.

The rest of the paper is as follows. In Section 2 we review the empirical evidence on the term structures of volatilities for both aggregate consumption and for wealthy shareholder consumption. We calibrate a model that matches the hump shaped term structure of volatilities in order to show that the positive autocorrelation for aggregate consumption at short horizons provides no support for the BY explanation of the equity premium, as their mechanism requires variance ratios to increase with horizon for many decades, not just a few years. We also investigate the term structure of volatilities for net payout. In Section 3 we use an argument akin to Miller and Modigliani (1961) to show that cash dividends can be chosen to be arbitrarily smooth, and therefore can mask the risks inherent in corporate cash flows. Moreover, we show that the price of one share of stock is independent of the cash dividend process, but very sensitive to the net payout process, implying that most of the extant literature is misspecifying the cash flow process needed to price corporate securities. In Section 4 we introduce generalized recursive preferences which are associated with two parameters that govern risk aversion, and show how this captures the notion of “crash risk” similar to that of Bates (2008). We then specify the consumption process for wealthy
investors and identify the pricing kernel. The net payout process is then specified, and is used to price the claim to corporate assets. We show that the calibration matches salient asset pricing moments well. We conclude in Section 5. Proofs are found in the Appendix.

2 Empirical Evidence

Our data includes multiple time series of consumption and aggregate financial flows. The majority of the data comes from the Flow of Funds Accounts (Table F.102) and National Income and Product Accounts (NIPA) of the United States.

We use three sets of consumption data. Real per capita consumption from 1889 to 2009 is obtained from Robert J. Shiller’s website. Real per capita consumption from 1929 to 2013 is determined from the sum of real nondurable consumption (Line 8) and service (Line 13) from NIPA Table 2.3.3, scaled by population (Line 40) from NIPA Table 2.1. Real per capita consumption for the one-third wealthiest shareholders comes from MMV. The data, which is constructed using the Consumer Expenditure Survey (CEX) from 1982 to 2004, includes monthly estimates of quarterly consumption growth from three different tranches (i.e., sets of households that are on either a January/April/July/October or February/May/August/November or March/June/September/December schedule).\textsuperscript{10}

We extend the time series of aggregate net payout for nonfinancial corporate business from Larrain and Yogo (LY, 2008) to 2013 (the original LY data set went from (1927-2004)). Net payout is defined as the sum of net equity payout and net debt payout. Net equity payout is net dividend paid (Line 3) minus increase in corporate equity (Line 39). Net debt payout

\textsuperscript{10}Since households are in the survey for only four quarters, any (idiosyncratic) consumption autocorrelation within a family will be muted beyond this horizon. In this sense, the term structure of consumption volatility may be biased toward a flat structure, making it even more surprising to find a downward sloping curve.
is net interest and miscellaneous payments (Line 25 in NIPA Table 1.14) minus the sum of increase in commercial papers (Line 41) and corporate bonds (Line 43). We deflate aggregate net payout by the December value of the consumer price index (CPI) for all urban consumers from the Bureau of Labor Statistics. For robustness, we also construct a related measure by accounting for changes in aggregate cash holdings, since firms can also manipulate payout using this channel. However, we find that the two time series are reasonably similar, so we focus on net payout below. Details are provided in the Appendix.

We investigate short- and long-run consumption risk by estimating the term structure of consumption volatility. Specifically, we use the unbiased estimator of Lo and Mackinlay (1988) to estimate the term structure of variances. For log consumption time series $c_t$, where $t = 0, 1, 2, ..., N$, we estimate the consumption variance at horizon $T$ as

$$\text{Var}(T) = \frac{1}{m} \sum_{k=T}^{N} (c_k - c_{k-T} - T \hat{\mu})^2,$$

(1)

where $\hat{\mu} = (c_N - c_0)/N$ is average log consumption growth and

$$m = T(N - T + 1) \left(1 - \frac{T}{N}\right).$$

(2)

Since the available time series of consumption is relatively short, it is important to use an unbiased estimator when estimating consumption volatilities at longer horizons with overlapping data.

Figure 1 reports the term structures of consumption volatility for three different time series. The dashed curve is constructed from aggregate consumption using data from 1929-2013, which is used by both BKY and Beeler and Campbell (2013). The dotted curve is constructed from a longer time series (1887 - 2009), which is available on Robert J.
Shiller’s website. We include this longer sample because, even though the data is known to be noisier, both Shiller and Perron (1985) and BKY argue that a long span of data is required for accurately measuring low frequency features of the data. The solid curve uses data from MMV and reflects consumption volatility of the wealthiest one-third of agents in their data set from 1982-2004. Because this time series is so short, we limit the horizon reported to 5 years. There are two important features that we wish to emphasize. First, as we will discuss in more detail below, all three term structures are generally decreasing with maturity, in conflict with the predictions of BY and BKY. This generally downward sloping term structure is consistent, however, with there being predictability in consumption growth, as noted previously in Hansen, Heaton and Li (2008), BKY, and Schorfheide, Song and Yaron (2014).

Figure 1: **Empirical term structures of consumption volatility.** This figure displays three different term structures of consumption volatility. The dashed curve is constructed from aggregate consumption from 1929-2013. The dotted curve is constructed from aggregate consumption from 1887-2009. The solid curve is constructed from consumption of the top one-third wealthiest shareholders from 1982-2004.

11 Incidentally, point estimates for the term structure of consumption volatility for wealthy agents is monotonically decreasing in years six through ten.
The second feature to note from Figure 1 is that at the 1 year horizon, aggregate consumption volatility is approximately 2.5% from the BKY data, while aggregate wealthy consumption volatility is approximately 12.5%. Moreover, if we limit our attention to the same 22 year interval that spans the data for wealthy agents, the aggregate consumption volatility drops even lower, namely, to 0.95%. Hence, the consumption volatility for a representative wealthy shareholder may be more than 5 times larger than volatility of aggregate consumption. Note that 12.5% volatility is probably an underestimate of the consumption volatility for wealthy households from which optimal consumption and investment decisions are made. Indeed, the working paper version of Wachter and Yogo (2010) report that consumption volatility for wealthy individuals in the age groups 36-65 is approximately 20%. However, because there is concern that this estimate is noisy, we focus on the consumption process of the aggregate wealthy shareholder.

It is worth emphasizing that both Mankiw and Zeldes (1991) and MMV (2009) find the aggregate consumption process of shareholders to be not only more volatile, but also more correlated with stock returns. In addition, MMV report show that the consumption process of an aggregated wealthy shareholder explains not only the equity premium but also the value premium with levels of risk aversion as low as six. For these reasons, rather than focusing on aggregate consumption, which combines the consumption processes of shareholders and non-shareholders alike, we identify the SDF using the aggregate consumption processes of wealthy agents.

We now turn our attention to Figure 2, where four more term structures of consumption volatility are reported. The solid curve (again) reports the term structure of aggregate consumption volatility using the empirical data from 1929-2013. The dot-dash curve is created from generated data when the log-consumption process is specified to be iid normal
Figure 2: **Calibrated term structures of consumption volatility.** This figure displays four different term structures of consumption volatility. The solid curve is constructed from aggregate consumption from 1930-2009. The dot-dash curve is created from generated data assuming that log-consumption is iid over one month interval, and then time aggregated over one year. The dashed curve is generated using the BKY calibration. The dotted curve is generated from our model given in equations (5).

over a one month interval, and then time-aggregated over one year. As first noted by Working (1960), time aggregation generates positive autocorrelation over consecutive periods, and in turn, a term structure of volatility which increases monotonically, with long-horizon volatility reaching approximately 125% of its one-year volatility.

The dashed curve is generated from the BKY model, who specify log-consumption dynamics as driven by two persistent variables \((x_t, \sigma_t)\) which control expected consumption growth and consumption volatility respectively:

\[
\Delta c_{t+1} = \mu_c + x_t + \sigma_t \tilde{\epsilon}_{x,t+1}
\]

\[
x_{t+1} = \rho_x x_t + \nu_x \sigma_t \tilde{\epsilon}_{x,t+1}
\]

\[
\sigma^2_{t+1} = \sigma^2 + \rho_\sigma \left( \sigma^2 - \tilde{\sigma}^2 \right) + \nu_\sigma \tilde{\epsilon}_{\sigma,t+1}.
\]

\[ (3) \]
Note that the BKY paradigm requires that the term structure of aggregate consumption volatility increases monotonically, whereas empirically it is hump-shaped – increasing up to year 3, and then dropping monotonically after that. To emphasize how poorly the theoretical fit is to the empirical one, we compare the average slopes of the term structure of aggregate consumption variances in the data to that predicted by BKY. In particular, we define the average slope as:

\[
slope = \left( \frac{1}{7} \right) \sum_{n=4}^{10} \left( \frac{Var(n)/n}{Var(3)/3} - 1 \right).
\]  

Results are shown in Table 1. The empirical point estimate is -0.36. We then calculate the slope for 1,000,000 generated economies from BKY. We find that only about 1 out of 1,000 generated economies have a point estimate that is as low as this empirical estimate. Hence, the slope provides very strong evidence against the upward sloping prediction of BKY beyond three years.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>BKY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>-0.36</td>
<td>Mean 5% 3% 1% 0.2% 0.1%</td>
</tr>
<tr>
<td>Estimate</td>
<td></td>
<td>0.30 -0.09 -0.14 -0.24 -0.33 -0.37</td>
</tr>
</tbody>
</table>

Table 1: This table presents the average slopes for the term structures of aggregate consumption variances in the data and BKY simulated economies. The average slope is defined as \( \left( \frac{1}{7} \right) \sum_{n=4}^{10} \left( \frac{Var(n)/n}{Var(3)/3} - 1 \right) \). The estimate in the data is computed using NIPA log aggregate consumption growth from 1929-2013. The right panel reports the average slope and the distribution of the slope generated by BKY economies. For example, the last column reports that only 0.1% of BKY generated economies have a slope estimate less than -0.37.

The third curve in Figure 2 comes from our extension of the BKY model, and is calibrated to match the hump shape from the empirical data. In particular, we assume that \( x_t \) mean-reverts to \( \bar{\pi}_t \), which itself is an AR1 process. In addition, we consider a more general
correlation structure:

\[
\begin{align*}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \tilde{\epsilon}_{c,t+1} \\
x_{t+1} &= (1 - \rho_x) \bar{x}_t + \rho_x x_t + \nu_x \sigma_t \left[ \alpha_{x} \tilde{\epsilon}_{x,t+1} + \sqrt{1 - \alpha_{x}^2 \tilde{\epsilon}_{x,t+1}^2} \right] \\
\bar{x}_{t+1} &= \rho_x \bar{x}_t + \nu_x \sigma_t \alpha_{x} \tilde{\epsilon}_{x,t+1} + \sqrt{1 - \alpha_{x}^2 \tilde{\epsilon}_{x,t+1}^2} \\
\sigma_{x,t+1}^2 &= \sigma_x^2 + \rho_x \left( \sigma_x^2 - \sigma_t^2 \right) + \nu_x \tilde{\epsilon}_{x,t+1}.
\end{align*}
\]  

(5)

We have intentionally chosen a fit that tends to overestimate the empirical point estimates of volatility in order to provide an upper bound for the risk implicit in this model. Even so, as demonstrated in the Appendix, this model generates a maximum (annualized) Sharpe ratio of approximately 0.2, which is way too low to explain the equity premium. Indeed, this estimate is considerably lower than if consumption dynamics had been specified to be iid, since in this case the BKY calibration generates a maximum Sharpe ratio of \( \gamma \sigma_c = (10)(0.025) = 0.25 \).

An implication of this model is that a term structure of consumption volatility that is upward sloping for only a few years, and then downward sloping after that, provides no support for the mechanism that BY uses to explain the equity premium. Intuitively, this is because BY specify an agent to have preference for early resolution of uncertainty\textsuperscript{12}, and the generally downward sloping term structure of consumption volatility implies that the agent in fact receives early resolution. This example emphasizes why, when investigating the empirical support for the BY paradigm, one should focus on the term structure of volatilities and not on autocorrelations, which, due to the hump shape, mask how poorly the BY calibration matches the term structure of consumption volatilities.

Figure 3 reports three term structures of corporate cash flow volatility. The dashed

\textsuperscript{12}In particular, BY specify the inverse of the parameter which controls agent’s elasticity of intertemporal substitution to be lower than the parameter which controls the agent’s risk aversion: \( \left( \frac{1}{\psi} \right) < \gamma \).
Figure 3: Term structures of corporate cash flow volatility. This figure displays three different term structures of corporate cash flow volatility. The dashed curve is constructed from cash dividends from 1930-2010. The dotted curve is constructed from the BKY calibration. The solid curve is constructed from the net payout time series from 1930-2010.

curve is constructed from the empirical cash dividend process used by BKY. The dotted curve is generated from the BKY model. The solid curve is constructed from the net payout series from 1927-2013. Once again, we find that the empirical term structures of corporate cash flow volatilities are downward sloping, while the BY framework specifies a cash flow process which generates upward sloping term structure. Moreover, in spite of the fact that net payout is an “unleveraged” cash flow (since it is cash flow going to both equity holders and bond holders), its volatility at short horizons is approximately three times larger than the short horizon volatility of cash dividends (a “leveraged” cash flow). Indeed, if the net payout volatility is multiplied by the “leverage factor” \( \left( \frac{1}{1-L} \right) \) with a leverage ratio of \( L = 0.4 \), then the implied volatility for the leveraged cash flow would be five times higher than the volatility of cash dividends. Rather than follow this approach, however, we will instead focus on the unleveraged claim to corporate assets (i.e., equity plus corporate debt), and calibrate
the model to match its empirical estimate for asset premium of 4.8%.

In summary then, the term structures of volatility for aggregate consumption, consumption of wealthy shareholders, cash dividends, and net payout are all downward sloping (at least, beyond the three-year horizon). In contrast, the BY mechanism needs these term structures to be upward sloping in order to capture the equity premium. Below, rather than focus on aggregate consumption, we calibrate our model to an aggregate wealthy investor and show that our model fits both these consumption and net payouts term structures of volatilities and salient asset pricing moments well. First, however, we provide an argument for why net payout rather than cash dividends should be the corporate cash flow that is focused on in these types of calibrations.

3 Net Payout versus Cash Dividends

There is a large and growing literature which argues that cash dividends are smoothed by management, and therefore do not accurately reflect the risks faced by corporations (e.g., Longstaff and Piazzesi (2004), Boudoukh et al. (2007), Larrain and Yogo (2008), Chen (2009)). In this section, we provide a simple example which demonstrates why net payout should provide a more economically meaningful time series than cash dividends. The argument is akin to the logic in Miller and Modigliani (1961), and captures the intuition that, in a world without financial frictions, managers cannot create (or destroy) value by manipulating the cash dividend process in ways that have no impact on the underlying operations of the firm.

Consider an all-equity firm whose (log) net payout process $x_t = \log X_t$ follows arithmetic
Brownian motion:

\[ dx = \left( g - \frac{\nu^2}{2} \right) dt + \nu \, dz. \]  

(6)

We specify the pricing kernel dynamics to follow

\[ \frac{d\Lambda}{\Lambda} = -r \, dt - \theta \, dz. \]  

(7)

All parameters are specified as constants. Hence, risk neutral dynamics are

\[ dx = \left( g^Q - \frac{\nu^2}{2} \right) dt + \nu \, dz^Q, \]  

(8)

where \( g^Q = (g - \nu \theta) \). The date-t value of the claim to all payout takes the Gordon growth model form:

\[
V(x_t) = \mathbb{E}_t^{Q} \left[ \int_t^\infty du \, e^{-r(u-t)} e^{x_u} \right] = \frac{e^{x_t}}{r - g^Q}.
\]  

(9)

We specify the number of shares outstanding at date \( t \) to be \( N_t \). Hence, agent-\( i \) who owns one share at date-\( t \) has a claim that is worth:

\[
V_i(x_t, N_t) = \left( \frac{1}{N_t} \right) \left( \frac{e^{x_t}}{r - g^Q} \right).
\]  

(10)

We now demonstrate that the value of her claim is unimpacted by the equity issuance/repurchase choices of management, although these choices can dramatically impact cash flow volatilities.

As noted previously, there is strong empirical evidence suggesting that firms smooth cash
dividends through equity issuance and repurchase decisions, and also through cash holdings and capital structure decisions. As such, here we assume (and in the next few paragraphs, prove that this assumption is self-consistent) that management chooses the log-cash dividend policy $y_t$ to follow:

$$
\begin{align*}
    dy &= \alpha \, dt + \beta \, dz \\
    &= (\alpha - \beta \theta) \, dt + \beta \, dz^Q.
\end{align*}
$$

(11)

The reason that it can commit to this process is that, at any date-$t$, no matter how low $x_t$ is, it is greater than minus infinity, implying that firm value $V(x_t) = \frac{e^{x_t}}{r - g^Q}$ is non-infinitesimal, whereas the amount of cash dividend owed, $e^{y_t} \, dt$ is infinitesimal. Hence, even if the firm is doing “poorly” in that the cash dividend promised is larger than net payout (i.e., $(y_t > x_t)$), the firm is of sufficient value that management will be able to issue enough new equity to maintain their dividend strategy.

To continue with our example, at date-$t$ the firm generates net payout of $e^{x_t} \, dt$, but has committed to pay $e^{y_t} \, dt$. Thus, the firm must raise\(^{13}\) $(e^{y_t} - e^{x_t}) \, dt$. Since the firm is worth $V(x_t) = \frac{e^{x_t}}{r - g^Q}$, it follows that new shareholders purchasing $dN_t$ of shares would own a fraction $\left(\frac{dN_t}{N_t + dN_t}\right)$ of the firm, and hence would have a claim whose value is

$$
\left(\frac{dN_t}{N_t + dN_t}\right) \left(\frac{e^{x_t}}{r - g^Q}\right).
$$

(12)

Setting this equal to the amount being raised, we find

$$
\frac{dN_t}{N_t + dN_t} = (r - g^Q) \left(e^{y_t - x_t} - 1\right) \, dt.
$$

(13)

\(^{13}\)if $x_t > y_t$, then the firm is actually repurchasing shares with the excess funds.
Since the $dN_t$ process is locally deterministic, Ito's lemma implies that it simplifies to

$$\frac{dN_t}{N_t} = (r - g^Q) (e^{y_t-x_t} - 1) \, dt. \quad (14)$$

In summary, the firm can commit to a cash-dividend policy as in equation (11) for arbitrary choices of parameter values $(y_0, \alpha, \beta)$ if it chooses its equity issuance/repurchase policy to follow equation (14).

Here we demonstrate that the value of the claim to a shareholder who owns one share at date-$t$ is unimpacted by the equity issuance/repurchase strategy of the firm. Since at any future date-$s$ she has claim to the fraction $\left(\frac{1}{N_s}\right)$ of the firm, she will receive a cash dividend of $\left(\frac{1}{N_s}\right) e^{y_s}$. As such, the value of her claim is:

$$V(x_t, y_t, N_t) = E^Q_t \left[ \int_t^\infty ds \, e^{r(s-t)} \left(\frac{1}{N_s}\right) e^{y_s} \right]. \quad (15)$$

This implies that

$$e^{-rt}V(x_t, y_t, N_t) + \int_0^t ds \, e^{-rs} \left(\frac{1}{N_s}\right) e^{y_s} = E^Q_t \left[ \int_0^\infty ds \, e^{-rs} \left(\frac{1}{N_s}\right) e^{y_s} \right] \quad (16)$$

is a $Q$-martingale, implying that is satisfies

$$0 = \frac{1}{dt} E^Q_t [dV] - rV + \frac{e^{y_t}}{N_t}. \quad (17)$$

It is straightforward to show that $V(x_t, y_t, N_t) = \left(\frac{1}{N_t}\right) \left(\frac{e^{x_t}}{r-g^Q}\right)$ satisfies this condition, which equals the value found in equation (10). That is, the agent owns a fraction $\left(\frac{1}{N_t}\right)$ of the firm, whose value is equal to the Gordon Growth model value $\left(\frac{e^{x_t}}{r-g^Q}\right)$ independent of its dividend policy $(y_t, \alpha, \beta)$. 

20
An implication of equation (15) is that, while it is possible to price one share of stock by determining the present value of future cash dividends paid to this one share, necessary inputs to do so include both the number of shares process \( N_t \) and the net payout process \( x_t \). Moreover, the value of the claim to one share of stock ends up independent of the cash dividend process. For example, the firm is free to commit to a perfectly deterministic cash dividend process by choosing \( \beta = 0 \), and this will have no impact on the value of equity.

For simplicity, we have focused on an all-equity firm. But the intuition that management is free to smooth the cash dividend process holds also for the more general situation where the firm has outstanding debt. Even in this case, however, in order to price equity, a necessary input will be the net payout process. The bottom line is that the cash dividend process is basically irrelevant for pricing equity, whereas the net payout process is paramount.

### 4 Generalized Recursive Preferences

In this section, we specify the preferences of the representative wealthy shareholder as a generalized form of recursive preferences (Kreps and Porteus (1978), Epstein and Zin (EZ, 1989)). In order to focus on the intuition, all proofs are relegated to the Appendix.

The agent’s preferences are characterized by two risk aversion parameters \( (\gamma_L, \gamma_H) \) with relative “weights” \( (w_L, w_H) \) which are positive and sum to unity: \( w_L + w_H = 1 \):

\[
U_t = \left\{ \delta \, dt \, C_t^{1-\rho} + (1 - \delta \, dt) \left[ E_t \left( U_{t+dt}^{1-\gamma_L} \right) \right] \left( \frac{w_L}{1-\gamma_L} \right) \left[ E_t \left( U_{t+dt}^{1-\gamma_H} \right) \right] \left( \frac{w_H}{1-\gamma_H} \right) \right\}^{\frac{1}{1-\rho}} . \tag{18}
\]
For the special case $\rho = 1$, this reduces to

$$u_t = \delta dt \, c_t + (1 - \delta dt) \left\{ \left( \frac{w_L}{1 - \gamma_L} \right) \log \left[ \mathbb{E}_t \left( e^{(1-\gamma_L)u_{t+\Delta t}} \right) \right] + \left( \frac{w_H}{1 - \gamma_H} \right) \log \left[ \mathbb{E}_t \left( e^{(1-\gamma_H)u_{t+\Delta t}} \right) \right] \right\},$$

where we have defined $c_t \equiv \log C_t$ and $u_t \equiv \log U_t$. Just as in the standard EZ framework, this special case leads to analytic solutions if the dynamics for the state vector that controls log-consumption are affine. To see this intuitively, let us guess that $u_{t+\Delta t}$ is affine in $c_{t+\Delta t}$ (and other variables that drive the log-consumption dynamics). In that case, as is well known (e.g., Duffie and Kan (1996)), the expectation $\mathbb{E}_t \left[ e^{(1-\gamma)u_{t+\Delta t}} \right]$ will take on an exponential-affine form. This in turn implies that $\left( \log \mathbb{E}_t \left[ e^{(1-\gamma)u_{t+\Delta t}} \right] \right)$ will be affine in the state vector.

The implication is that all three terms on the right hand side of the formula are affine in the state vector, implying that their sum, $u_t$, is affine in the state vector. Hence, our original guess is self-consistent.

The pricing kernel associated with the preferences described in equation (18) can be expressed as

$$M_{t,t+\Delta t} = e^{-\delta dt - \rho dc_t} \left( \frac{U_{t+\Delta t}^{1-\gamma_L}}{\mathbb{E}_t \left[ U_{t+\Delta t}^{1-\gamma_L} \right]} \right)^{w_L(p-1)} \left( \frac{U_{t+\Delta t}^{1-\gamma_H}}{\mathbb{E}_t \left[ U_{t+\Delta t}^{1-\gamma_H} \right]} \right)^{w_H(p-1)} \times \left[ w_L \left( \frac{U_{t+\Delta t}^{1-\gamma_L}}{\mathbb{E}_t \left[ U_{t+\Delta t}^{1-\gamma_L} \right]} \right) + w_H \left( \frac{U_{t+\Delta t}^{1-\gamma_H}}{\mathbb{E}_t \left[ U_{t+\Delta t}^{1-\gamma_H} \right]} \right) \right]. \quad (19)$$

To provide some intuition for how these generalized preferences differ from the standard recursive ones, consider the special case where consumption growth is iid:

$$\frac{dC}{C} = \mu dt + \sqrt{V} \, dz_t + \left( e^{-\eta_c} - 1 \right) (dq - \lambda dt). \quad (20)$$
Here, $\mu$ is expected consumption growth, $V$ is consumption variance, $dz_c$ is a Brownian motion, $(e^{-\eta_c} - 1)$ is the jump size, and $dq$ is a Poisson jump with intensity $\lambda$. Pricing kernel dynamics can be shown to take the form:

$$
\frac{d\Lambda}{\Lambda} = -r \, dt - \sqrt{V} \, dz_c + (dq - \lambda \, dt) \left[ w_L e^{\eta_c \gamma_L} + w_H e^{\eta_c \gamma_H} - 1 \right],
$$

where we have defined a weighted-average risk aversion coefficient

$$
\gamma \equiv w_L \gamma_L + w_H \gamma_H.
$$

It follows that $\gamma$ can be interpreted as the effective risk aversion coefficient associated with diffusions, whereas the effective risk aversion coefficient $\hat{\gamma}$ associated with jumps depends on the size of the jump, and is defined implicitly via:

$$
e^{\eta_c \hat{\gamma}} \equiv w_L e^{\eta_c \gamma_L} + w_H e^{\eta_c \gamma_H}.
$$

Note from equation (20) that large values of $\eta_c$ are associated with large negative jumps in consumption. In Figure 4, we plot the implied risk aversion $\hat{\gamma}$ as a function of $\eta_c$. It demonstrates that $\hat{\gamma}$ increases from $\gamma_L$ for large positive consumption jumps to $\gamma_H$ for large negative consumption jumps. In this sense, these preferences capture the notion that agents are more risk averse to bets associated with larger wealth losses. Interestingly, for a jump size near zero, the effective risk aversion is equal to $\gamma$. This property differs significantly from implications of source-dependent utility functions.
Figure 4: Effective Level of Risk Aversion as a Function of Jump Size. This figure plots the effective level of risk aversion $\hat{\gamma}$ as a function of a parameter which controls jump size $\eta_c$. The actual jump size in consumption growth is $(e^{-\eta_c} - 1)$. Parameter values are $\gamma_L = 2$, $\gamma_H = 10$, $w_L = \frac{5}{8}$. Note how the effective level of risk aversion varies from $\gamma_L$ to $\gamma_H$ as the (negative) jump size increases. Also note that $\hat{\gamma} = \tilde{\gamma}$ when $\eta_c = 0$.

4.1 Consumption Process

We specify the log-endowment process of the representative wealthy shareholder to be:

\[
dc_t = \left( \mu_t - \frac{1}{2} V_t + \alpha_c \lambda_t \right) dt + \sqrt{V_t} dz_c - (\tilde{\eta}_c dq - \alpha_c \tilde{\eta}_v \lambda_t \, dt)
\]

\[
d\mu_t = \kappa_\mu (\bar{\mu} - \mu_t) \, dt - \sigma_\mu \sqrt{V_t} \, dz_\mu - \sigma_\mu \sqrt{V_t} \, dz_\mu + (\tilde{\eta}_\mu dq - \alpha_\mu \tilde{\eta}_v \lambda_t \, dt)
\]

\[
dV_t = \kappa_V (V - V_t) \, dt - \sigma_v \sqrt{V_t} \, dz_\epsilon - \sigma_v \sqrt{V_t} \, dz_\mu - \sigma_v \sqrt{V_t} \, dz_\nu + (\tilde{\eta}_v dq - \tilde{\eta}_v \lambda_t \, dt)
\]

\[
\lambda_t = \lambda_0 + \lambda_V V_t.
\]
where the \(\{dz\}\) are Brownian motions and \(dq\) is a jump process. We have defined \(\alpha_\lambda = -\left(\frac{\alpha^2 \eta_1^2}{1 + \alpha \eta_2} \right)\) so that \(\mu\) can be interpreted as the expected consumption growth rate:

\[
\frac{dC}{C} = \mu_i dt + \sqrt{V_i} dz_c + (e^{-\xi} - 1) dq + \frac{\alpha_c \eta_1}{1 + \alpha_c \eta_2} \lambda_i dt. \tag{25}
\]

In the spirit of Eraker (2004) and Benzoni et al. (2009), we specify the jumps in \((c_t, \mu_t, V_t)\) to be correlated while guaranteeing that variance remains positive. In order to reduce the number of free parameters in the name of parsimony, the jump size variables \((\eta_c, \eta_\mu, \eta_v)\) are drawn from the distribution

\[
\pi (\eta_c, \eta_\mu, \eta_v) = \frac{1}{\eta_v} e^{-\frac{\eta_v}{\eta_v}} \delta (\eta_\mu - \alpha_\mu \eta_v) \delta (\eta_c - \alpha_c \eta_v), \tag{26}
\]

where \(\delta(\cdot)\) is the Dirac Delta function. This implies that, conditional on the volatility jump size, the jumps in \(\mu\) and \(c\) are determined.

The proposed dynamics are reminiscent of Bansal and Yaron (2004). However, as we demonstrate below, the calibration is completely different. Indeed, in the BY calibration, the majority of the equity premium comes from the assumption that innovations to consumption growth and expected consumption growth are uncorrelated, which leads to the term structure of consumption volatilities to be *increasing* with horizon. However, as we demonstrated in the previous section, the empirical term structure of consumption volatility is downward sloping beyond three years. Further, we showed that if we calibrate a model to match this empirical fact, their mechanism cannot capture a large equity premium. One reason why consumption variance ratios are downward sloping, and hence, why the long run risk explanation of the equity premium is counterfactual, is that large drops in consumption (e.g., recessions) are associated with low levels of both employment and capital utilization.
Eventually, labor and capital get reallocated, and there is a *recovery* (Gourio (2009)). In other words, there is a negative correlation between shocks to consumption growth and shocks to long-run expected consumption growth. These recessions and recoveries combine to make consumption volatility decreasing in horizon.

To proceed, we follow an approach similar to that of Bansal and Yaron (2004) by approximating the log utility-consumption ratio to be affine in the state vector \((\mu_t, V_t)\) (It can be shown that this relation is exact in the case \(\rho = 1\)):

\[
\log \left( \frac{U_t}{C_t} \right) \approx u_o + u_v \mu_t - u_V V_t. \tag{27}
\]

Under this assumption, pricing kernel dynamics follow:

\[
\frac{d\lambda}{\lambda} = -r_t \, dt - \theta_e \sqrt{V_t} \, dz_{e} - \theta_\mu \sqrt{V_t} \, dz_{\mu} - \theta_V \sqrt{V_t} \, dz_{V} + \left( \hat{\Gamma}_t - 1 \right) dq - \left( E_t \left[ \hat{\Gamma}_t \right] - 1 \right) \, \lambda_t \, dt; \tag{28}
\]

where

\[
\theta_e = \rho + (\bar{\sigma} - \rho) \left[ 1 - \sigma_{\mu} u_{\mu} + \sigma_{V} u_{V} \right] \tag{29}
\]

\[
\theta_\mu = (\bar{\sigma} - \rho) \left[ \sigma_{\mu} u_{\mu} - \sigma_{\mu} u_{\mu} \right] \tag{30}
\]

\[
\theta_V = (\bar{\sigma} - \rho) \left[ \sigma_{V} u_{V} \right]. \tag{31}
\]

It follows that pricing kernel variance is

\[
\frac{1}{dE} E_t \left[ \left( \frac{d\lambda}{\lambda} \right)^2 \right] = V_t \left( \theta_e^2 + \theta_\mu^2 + \theta_V^2 \right) + \lambda_t E_t \left[ \left( \hat{\Gamma}_t - 1 \right)^2 \right]. \tag{32}
\]
The market price of consumption-risk \( \theta_c \) given in Equation (29) can be interpreted as follows. Note that, from equation (24), the parameter \( \sigma_{\mu c} \) controls the correlation between innovations in consumption and expected consumption growth, since, ignoring jumps, \( \left( \begin{array}{c} \frac{1}{\pi} \end{array} \right) \left[ dc_t \: d\mu_t \right] = -\sigma_{\mu c} V_t \). In the BY framework, this parameter has been set to zero. However, in order to capture the generally downward sloping term structures of consumption volatility, we need to choose \( \sigma_{\mu c} \) to be positive. Since \( u_\mu \) is positive (that is, utility is increasing in expected consumption growth), it follows from Equation (29) that increasing the value of \( \sigma_{\mu c} \) tends to reduce the market price of consumption-risk when the agent has preference for early resolution of uncertainty (i.e., when \( \bar{\gamma} - \rho > 0 \)). Intuitively, this is because the negative correlation between changes in consumption growth and changes in expected consumption growth provides the agent with early resolution of uncertainty (i.e., downward sloping term structure of consumption volatility). As such, risk premia are reduced. Indeed, recall in the previous section when we calibrated a model to fit the hump shape term structure of volatility for consumption dynamics, we obtained a maximum Sharpe ratio smaller than had consumption been iid. That result was directly due to specifying a negative correlation between changes in consumption and expected long-horizon consumption growth. We emphasize however that, whereas matching the generally downward sloping term structure of consumption volatility eliminates any possibility for the BY paradigm to explain the equity premium, our framework can still capture the equity premium in spite of this feature because aggregate consumption volatility for wealthy investors is so much higher at short horizons (12.5% vs. 2.5%).

Within this log-linear approximation, the risk free rate is affine in the state vector

\[
 r_t = r_o + \rho \mu_t - r_v V_t, \tag{33}
\]
where \( r_o \) and \( r_v \) are constants. The instantaneous interest rate volatility can be identified by noting that

\[
\begin{align*}
\left. dr \right|_{\text{stoch}} &= \sqrt{V_t} dz_c (r_v \sigma_{c} - \rho \sigma_c) + \sqrt{V_t} dz_\mu (r_v \sigma_{\mu} - \rho \sigma_\mu) + \sqrt{V_t} dz_v (r_v \sigma_v) \\
&\quad + d\rho (\rho \hat{\eta}_\mu - r_v \hat{\eta}_v),
\end{align*}
\]

(34)

and hence

\[
\frac{1}{dt} E_t [dr^2] = V_t \left( (r_v \sigma_{c} - \rho \sigma_c)^2 + (r_v \sigma_{\mu} - \rho \sigma_\mu)^2 + (r_v \sigma_v)^2 \right) + 2 \lambda \hat{\eta}_v^2 (\rho \sigma_\mu - r_v)^2.
\]

(35)

In order for our framework to capture low interest rate volatility, we use the insights of both BY (who specify \( \rho \) to be small) and Campbell and Cochrane (1999) (who specify their habit model so that the intertemporal substitution term moves in the opposite direction from the precautionary savings term). This latter effect is accomplished by assuming a positive correlation between changes in \( \mu \) and changes in \( V \). Intuitively, bad shocks to consumption are associated with a recovery (i.e., an positive shock to long run expected consumption growth) and to high consumption volatility.

### 4.2 Wealth Process

As with standard Epstein/Zin preferences, because equation (18) is homogeneous of degree one in consumption, it follows that the log wealth-consumption ratio is linearly related to the log utility-consumption ratio:

\[
\log \delta + \log \left( \frac{W_t}{C_t} \right) = (1 - \rho) \log \left( \frac{U_t}{C_t} \right).
\]

(36)
Note that by combining equations (36) and the log-linear approximation (27), we obtain

\[ W_t = e^{c_t - \log \delta + (1 - \rho)(u_0 + u_\mu \mu_t + u_\nu V_t)}. \]  

(37)

Note that for \( \rho < 1 \), consistent with observation for the price-dividend ratio on stocks, the wealth-consumption ratio is increasing in expected consumption growth \( \mu_t \) and decreasing in consumption volatility \( V_t \), since both \( u_\mu \) and \( u_\nu \) are positive. We find the same relation holds for the price:net payout ratio.

Applying Ito’s lemma (extended for jumps) to equation (37) generates a solution of the form:

\[ \frac{dW}{W} = \mu_w dt + \theta_{w,\nu} \sqrt{V} dz + \theta_{w,\mu} \sqrt{V} dz + \theta_{w,v} \sqrt{V} dz + dq (\Gamma - 1), \]  

(38)

where the coefficients are given in the Appendix.

In order to identify the parameters \( (u_0, u_\mu, u_\nu) \), note that wealth is the claim to consumption, and therefore satisfies

\[ 0 = \frac{1}{dt} E_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dW}{W} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dW}{W} \right) \right] + e^{-z_t}, \]  

(39)

where \( z_t = \log \left( \frac{W_t}{C_t} \right) \) is the log wealth-consumption ratio. As demonstrated by Bansal and Yaron (2004), log-linearization provides an excellent approximation:

\[ 0 \approx \frac{1}{dt} E_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dW}{W} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dW}{W} \right) \right] + e^{-z_t} [1 - (z_t - \bar{z})], \]  

(40)
where equation (37) implies that

\[ z_t \approx -\log \delta + (1 - \rho) \left( u_\mu + u_\nu \mu_t - u_\nu V_t \right), \quad (41) \]

and hence

\[ \bar{z} \approx -\log \delta + (1 - \rho) \left( u_0 + u_\nu \bar{\mu} - u_\nu \bar{V} \right). \quad (42) \]

Plugging equations (28) and (38) into equation (40) and then collecting terms linear in \( \mu_t, V_t \), and terms independent of the state vector, we find the three equations which, combined with equation (42) identify the parameters \((u_0, u_\mu, u_\nu, \bar{z})\): For example, we find

\[ u_\mu = \frac{1}{e^{-\bar{z} + \kappa_\mu}}. \quad (43) \]

The other roots do not possess analytic solutions, but are easily found via numerical root finding. We note that there are in general two roots for \( u_\nu \). However, following the insights of Tauchen (2011), the root whose value remains finite in the limit of the volatility and intensity parameters going to zero is the economically relevant one.

### 4.3 Pricing Net Payout Strips

We specify the dynamics of log net payout \( x_t = \log X_t \) to be stationary with log-consumption:

\[
\begin{align*}
    dx &= \left[ \kappa_z (c_t + \phi_x - x) + \phi_\mu (\mu - \bar{\mu}) + \phi_\nu (V - \bar{V}) \right] dt + \sigma_x \sqrt{V_t} dz_x + \sigma_{x\nu} \sqrt{V_t} dz_{x\nu} \\
    &\quad + \sigma_{x\nu} \sqrt{V_t} dz_{x\nu} + \sigma_x \sqrt{V_t} dz_x - \eta x d\xi + \alpha_x \eta_x \lambda_t dt.
\end{align*}
\]

\[(44)\]
where the jumps are conditionally correlated

\[ \pi(\eta_c, \eta_m, \eta_r, \eta_v) = \frac{1}{\eta_v} e^{-\eta_v} \mathbf{1}_{(\eta_v>0)} \delta(\eta_m - \alpha_c \eta_r) \delta(\eta_c - \alpha_c \eta_r) \delta(\eta_r - \alpha_r \eta_v). \tag{45} \]

The date-t price \( P^T(t, c_t, x_t, \mu_t, V_t) \) of the claim to \( e^{x_T} \) satisfies the equation

\[ \Lambda_t P^T(t, c_t, x_t, \mu_t, V_t) = \mathbb{E}_t[\Lambda_T e^{x_T}]. \tag{46} \]

This implies that \( \Lambda_t P^T(t, c_t, x_t, \mu_t, V_t) \) is a \( P \)-martingale, implying that its expected change is zero:

\[ 0 = \frac{1}{dt} \mathbb{E}_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dP}{P} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dP}{P} \right) \right]. \tag{47} \]

Given that state vector dynamics are affine, the solution takes the form:

\[ P^T(t, c_t, x_t, \mu_t, V_t) = e^{F_0(T-t) + F_c(T-t)c_t + F_x(T-t)x_t + F_\mu(T-t)\mu_t - F_v(T-t)V_t}. \tag{48} \]

As shown in the Appendix, we can solve analytically for \( F_x(\tau), F_c(\tau), F_\mu(\tau) \):

\[ F_x(\tau) = e^{-\kappa_x \tau}. \tag{49} \]

\[ F_c(\tau) = 1 - e^{-\kappa_c \tau}. \]

\[ F_\mu(\tau) = \left( \frac{1 - \rho}{\kappa_\mu} \right) \left( 1 - e^{-\kappa_\mu \tau} \right) + \left( \frac{\phi_\mu - 1}{\kappa_\mu - \kappa_x} \right) (e^{-\kappa_x \tau} - e^{-\kappa_\mu \tau}). \]

\( F_0(\tau) \) and \( F_v(\tau) \) must be solved numerically. The claim to all net payout, that is, the value of (equity plus corporate debt), is equal to the sum (more correctly, the integral) of these net payout strips.
4.4 Calibration

We calibrate the model to jointly capture the salient features of asset prices and the term structures of volatilities for both net payout and wealthy shareholder consumption. Parameter values are given in Table 2. We separate the parameters into three categories: preferences, consumption dynamics, and net payout dynamics. Preference parameters include the time preference parameter $\delta = 0.043$, which is useful for matching the average risk-free rate. As in BY, we set the elasticity of intertemporal substitution (EIS) to be greater than 1, which helps capture three features of the data: a positive correlation between price-dividend and expected consumption growth, a negative correlation between price-dividend and consumption volatility, and a smooth risk free rate process. Specifically, we set the reciprocal of EIS $\rho = 0.8$.

![Table 2: Calibration Parameters](image)

We specify the two risk aversion parameters $\gamma_L = 2$ and $\gamma_H = 10$. We also set $\omega_L = 5/8$, implying that the effective risk aversion for diffusion is $\tilde{\gamma} \equiv (\omega_L \gamma_L + \omega_H \gamma_H) = 5$. While Mehra and Prescott (1985) suggest that a reasonable upper bound for risk aversion is 10, many economists feel this number is way too high, so we consider this relatively low value for $\tilde{\gamma}$ to be a nice feature of our framework. Note that, since $\tilde{\gamma}$ is higher than $\rho$, the agent prefers early resolution of uncertainty.

We calibrate the consumption dynamics to closely match the term structure of consump-
tion volatility for the wealthiest one-third of shareholders. Recall, however, that the MMV investigate micro-level household data from the Consumer Expenditure Survey over the period 1982-2004 when aggregate consumption volatility was about 1% – that is, less than one-half of its 1930-2004 estimate. Hence, we consider it very likely that our estimate using the 1982-2004 data significantly underestimates the average level of volatilities. Moreover, since most of the downward slope in the term structure of aggregate consumption is due to the depression followed by recovery, we suspect that if we had the consumption data of wealthy individuals over this longer horizon, not only would the levels be higher, but most likely, the term structure would be even more downward sloping than what is presented in Figure 1.

In addition to setting $V = (0.17)^2$, we set $\kappa_v = 0.5$, which corresponds to a monthly persistence of consumption variance $\exp(-0.5/12) = 0.959$. This value falls between the calibration value in BY and the empirical estimate from Bansal, Kiku, Shaliastovich, and Yaron (2014). We set $\sigma_{vc} = 0.05$ and $\sigma_{mc} = 0.3$ so that consumption shocks are negatively correlated with both shocks to expected consumption growth and consumption volatility. This specification captures the notion that both consumption variance and expected consumption growth tend to be higher after a recent drop in consumption. As noted previously, a positive value for $\sigma_{mc}$ helps capture the downward sloping term structures of consumption volatility. As can be seen in equation (35), a positive value for $\sigma_{vc}$ helps reduce interest rate volatility since in this case variation in the second term (intertemporal substitution) and third term (precautionary saving) of the risk free rate in equation (33) partially offset each other. Note that this mechanism is reminiscent of that used by Campbell and Cochrane (1999), where their definition of habit was chosen to make these two terms cancel exactly, leading to a constant risk free rate.
Other parameters specifications include $\kappa = 0.4$, which corresponds to a monthly persistence of long-run expected growth risk $\exp(-0.4/12) = 0.967$. This choice contrasts with the BY and BKY specifications, which require higher persistence levels in order to generate a sizeable equity premium. In our calibration, however, the representative wealthy investor faces so much more consumption risk that these additional channels are not needed to obtain a large market price of risk. $\bar{\mu} = 0.034$ is chosen to match average log consumption growth.

For the parameters of jump risk, we set $\alpha_c = 3.0$ and $\bar{\eta}_\nu = 0.01$. These values yield an average consumption jump size $\alpha_c \bar{\eta}_\nu = 0.03$, which is quite small given that short horizon consumption diffusion volatility is over 12%. We set the average jump size in long-run expected growth risk to $\alpha_e \bar{\eta}_\nu = 0.0005$. Jumps in consumption are perfectly negatively correlated with jumps in both expected consumption growth and consumption variance. This enhances the negative correlation between innovations in consumption and expected consumption, which in turn helps generate a decreasing term structure of consumption volatility. We set jump intensity parameters $\lambda_\nu = 0$ and $\lambda_\eta = 0.1$ so that average interval between jumps is approximately five years.

Combining the probability structure of the jump process in equation (26) with the definition of $\Gamma_\lambda$ in the pricing kernel dynamics (equation (28)), we find that the effective risk aversion associated with a given jump size $\eta_\nu$ can be expressed as

$$e^{-\gamma(-\alpha_c + u_\mu \alpha_e - u_\nu)} = w_L e^{-\gamma_L(-\alpha_c + u_\mu \alpha_e - u_\nu)} + w_H e^{-\gamma_H(-\alpha_c + u_\mu \alpha_e - u_\nu)}.$$  (50)

With this specification, we find that the effective risk aversion for an average jump is approximately 5.3, and approximately 6.1 for a jump two standard deviations above the mean. Hence, the levels of risk aversion implicit in our framework are moderate even for large negative consumption jumps.
We calibrate the net payout dynamics to match the term structure of aggregate net payout volatility. We set $\kappa_x = 0.45$, which is equivalent to a monthly persistence of $\exp(-0.45/12) = 0.96$, to match the slope of the term structure of net payout volatilities. $\phi_x = 1.2$ is chosen to match average log net payout growth. We set $\phi_\mu = 0$ and $\phi_V = 0$.

We also set $\sigma_{\mu \mu} = 0$, $\sigma_{xc} = -1.8$, $\sigma_{xV} = 0$, and $\sigma_x = 2.7$ to match net payout volatilities. We set $\alpha_x = 2$, which closely matches the jump in consumption of wealthy investors.

The implied moments are generated from 10,000 economies, each of which is 80 years long. Clearly, the model does an excellent job of matching many different moments. Further, as Figures 5 - 6 show, our model also captures well the downward slope of the term structures of volatilities for both aggregate consumption of wealthy shareholders and net payout.

![Figure 5: Term Structure of Net Payout Volatility.](image)

This figure reports term structure of net payout volatility. The solid curve is constructed from the time series data from 1927 to 2013. The dashed curve is generated from 10,000 model-simulated economies.
Figure 6: **Term Structure of Wealthy Consumption Volatility.** This figure reports term structure of wealthy consumption volatility. The solid curve is constructed from the consumption process of the top one-third wealthiest shareholders from 1982 to 2004. The dashed curve is generated from 10,000 model-simulated economies.

It is worth noting that while BKY also match several moments well, there are a few examples where their model fails to match observation. Paramount is the prediction that the term structure of consumption volatility is upward sloping when empirically it is downward sloping after 3 years, as shown in Figure 2. In addition, BKY estimate the volatility of the risk-free rate to be 2.86%, yet their 95% confidence interval is the range (0.54, 1.59). If we take this range to equal four standard deviations, then their median value is approximately 8 standard deviations from the empirical point estimate. Similarly, BKY estimate the standard deviation of log price-dividend to be 0.45, yet their 95% confidence interval is the range (0.11, 0.30) Again, if we take this range to equal four standard deviations, then their median estimate is approximately 6 standard deviations from the empirical point estimate.
### Table 3: Sample moments estimated from 10,000 economies, each 80 years long.

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset risk prem.</td>
<td>0.048</td>
<td>0.048</td>
</tr>
<tr>
<td>Vol. of asset return</td>
<td>0.122</td>
<td>0.118</td>
</tr>
<tr>
<td>Average real risk-free rate</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>Vol. of risk-free rate</td>
<td>0.029</td>
<td>0.030</td>
</tr>
<tr>
<td>Average log consumption growth</td>
<td>0.019</td>
<td>0.019</td>
</tr>
<tr>
<td>Std. log consumption growth</td>
<td>0.126</td>
<td>0.123</td>
</tr>
<tr>
<td>Average log net payout growth</td>
<td>0.038</td>
<td>0.034</td>
</tr>
<tr>
<td>Std. log net payout growth</td>
<td>0.384</td>
<td>0.389</td>
</tr>
<tr>
<td>Average log(P/D)</td>
<td>4.249</td>
<td>3.788</td>
</tr>
<tr>
<td>Std. log(P/D)</td>
<td>0.605</td>
<td>0.589</td>
</tr>
</tbody>
</table>

#### 4.5 The Term Structures of Assets

In addition to matching the observed equity premium, recently the literature (e.g. Binsbergen et al. (2012)) has begun to investigate the risk premia associated with different horizons. For example, while stock is the claim to all payouts to shareholders, one can ask what is the price and expected return associated with each individual payout strip. This term structure of risk premia generates important cross sectional implications for expected returns. For example, if we interpret value firms as firms with relatively short duration cash flows, and growth firms as firms with long duration cash flows, then the existence of a value premium would suggest that short-horizon cash flows are riskier, consistent with our framework. This is because a downward sloping term structure of consumption volatility implies that long-horizon is not as risky as short-horizon cash flows. And in general equilibrium, less quantity of aggregate risk implies less price of risk. With this interpretation in mind, here we examine several other assets impacted by horizon effects.
4.5.1 The Term Structure of Consumption Strips

A consumption strip with maturity-\( T \) is a claim which pays at date-\( T \) the level of consumption \( C_T \). The date-\( t \) price \( P^T(t,c_t,x_t,\mu_t,V_t) \) of a consumption strip can be determined via:

\[
\Lambda_t P^T(t,c_t,x_t,\mu_t,\gamma_t) = E_t \left[ \Lambda_T e^{\gamma_T} \right],
\]

where \( \Lambda_t \) is the pricing kernel. We can solve for the price of a consumption strip using the same set of equations as those for pricing the net payout strip, but with a slightly different set of initial conditions. In particular, the initial conditions for consumption strips are \( A_{c,\gamma}(\tau = 0) = \beta \), and all other \( A_{i,\gamma}(\tau = 0) = 0 \). In addition, we can also define the yield on a consumption strip \( y_{c,t}^T \) as follows

\[
e^{-y_{c,t}^T(T-t)} E_t \left[ e^{\gamma_T} \right] = P^T(t,c_t,x_t,\mu_t,V_t).
\]

Figure 7 summarizes the model’s predictions. Consistent with the results of Binsbergen et al. (2012) for dividend strips, our model generates long-horizon consumption strips that are associated with lower yields, lower risk premia, and lower volatility than short term consumption strips. The loadings of consumption strips on consumption risks over many horizons are exactly one. Given that we match the downward sloping term structure of consumption volatility, it is not surprising that the volatility of consumption strips is also downward sloping.

In contrast, because the BKY model assumes an upward sloping consumption term structure of volatility, their model predicts an upward sloping volatility for consumption strips.
which in turn generates a higher risk premium in the long term than in the short term.

4.5.2 The Term Structure of Straddles on Wealth

A straddle is a portfolio of options that takes a long position in both an at-the-money European put and an at-the-money European call option with the same maturity. This strategy is widely used for volatility trading. Since the payoff of a straddle is the absolute value of the change in the price of the underlying asset over the maturity, a straddle mainly exposes to volatility risk. And the risk loading is positive. Therefore, the risk premium earned by holding a straddle is largely determined by the price of volatility risk. Using the S&P 500 index option data, Andries et al. (2014) find that i) the risk premiums of straddles are negative; ii) the Sharpe ratio (in absolute value) of a straddle is decreasing its maturity. Our model naturally generates these two patterns.

Here we use the model to price a straddle written on the agent’s wealth at \( T \). The price of this straddle is determined by

\[
ST^T (t, c_t, x_t, \mu_t, V_t) = \mathbb{E}_t \left[ \frac{\Lambda_T}{\Lambda_t} \left( \max(W_T - W_t, 0) + \max(W_t - W_T, 0) \right) \right], \tag{53}
\]

where \( W_t \) is determined by the wealth-consumption ratio given by Equation 37. We use the Monte Carlo simulation to price these straddles with several maturities around the steady state.

Figure (8) reports the term structures of the prices, risk premia, volatilities, and Sharpe ratios of these straddles. Consistent with the data, the model predicts i) negative risk premiums for straddles; and ii) the absolute value of the Sharpe ratio of a straddle is decreasing in maturity. This is closely related to the composition of the asset risk premium in the model.
Since the risk premia are largely due to exposure to short horizons in our model, the price of long-horizon volatility risk is also less than the price of short-horizon volatility risk.

### 4.5.3 The Term Structure of Real Interest Rates

In addition to matching the mean and volatility of the risk-free rate, the model also predicts a upward sloping term structure of real interest rates. The solution of the yield curve in our model is described in the appendix. Figure 9 presents the term structure of real interest rates predicted by the model. Over the long end of the curve (100 years), the real interest rate converges to a finite level less than 3.5%.

The term structure of real interest rates predicted by our model is upward sloping. This is in contrast to BKY, whose model predicts a downward sloping term structure. Intuitively, our model is associated with a decreasing term structure of consumption volatility, which in turn generates a precautionary savings term which decreases with horizon. This in turn creates an upward sloping term structure of real interest rates. In contrast, the BKY model is associated with an increasing term structure of consumption volatility, which in turn generates a precautionary savings term that increases with horizon. this in turn leads to a term structure of real interest rates which decreases with horizon. Empirically, the data is mixed on this issue.
Figure 7: Model-predicted term structures of the yields, risk premia, and volatilities of consumption strips
Figure 8: The term structures of the price, risk premia, volatilities, and Sharpe ratios associated with straddles on wealth.
Figure 9: The Term Structure of Real Interest Rate
5 Conclusion

This paper investigates the consumption-based asset pricing implications of a model calibrated to the consumption process of wealthy investors and net payout, defined as the sum of dividends, interest and net repurchases of equity and debt. One motivation for focusing on wealthy investors is that we feel they are more likely to satisfy the conditions necessary for aggregation to hold (e.g., satisfying first order optimality conditions, subject to minimal idiosyncratic shocks), and that the aggregated consumption process of any subset of investors that satisfy these conditions can be used to identify a stochastic discount factor (SDF) for pricing risky assets.

A more contentious motivation is that we feel there is little justification for interpreting the aggregate consumption process (which combines the consumption processes of shareholders and non-shareholders alike) as an optimal decision of some “representative agent.” As such, any attempt to explain the equity premium using the aggregate consumption process with its 2.5% volatility is forced to introduce additional sources of risk to explain the observed equity premium that may not accurately reflect the risks that actual shareholders face. Indeed, often these exaggerated sources of risk are imposed at longer horizons, generating counterfactual implications for the relative riskiness of short-horizon versus long-horizon assets.

Using micro-level household data from the Consumer Expenditure Survey over the period 1982-2004 from Malloy, Moscowitz and Vissing-Jorgensen (2009), we estimate consumption volatility of wealthy investors to be 12.5%, compared to approximately 1% aggregate consumption volatility over this same time period. This high level of volatility at the one-year horizon allows us to capture the equity premium in spite of calibrating the model to match the empirically observed downward sloping term structure of consumption volatilities, in
contrast to the upward sloping assumption which drives most of the equity premium in the long-run risk literature (Bansal and Yaron (2004)). The downward sloping term structure of consumption volatilities implies that our framework predicts that longer-duration assets are less risky, generating low long-horizon discount rates (e.g., Binsgergen et al (2012), Giglio et al (2014)), and a decreasing (in absolute value) term structure of Sharpe ratios for straddle returns (e.g., Dew-Becker et al (2014), Andries et al (2014)).

We specify corporate cash flows using net payout rather than cash dividends because there is strong empirical evidence suggesting that firms smooth cash dividends in order to mask the true risks associated with corporate cash flows. Building on this literature, we argue that the cash dividend dynamics specified by most of the extant literature cannot be used to price stocks because their specifications ignore equity repurchases and issuances. More accurately, we demonstrate that in order to price the claim to stocks correctly, one must account for the time variation in the outstanding share process. In particular, we use an argument akin to Miller and Modigliani (1961) to demonstrate that the parameters controlling the cash dividend process have no impact on the prices of corporate securities, whereas the parameters controlling the net payout process are paramount for pricing these securities. Indeed, we show that an all-equity firm can in fact commit to any cash dividend policy, even one which is perfectly deterministic. The implication is that the volatility of the cash dividend process itself may have little to say about the risks associated with the firm.

By specifying the agent to have a preference for early resolution of uncertainty (Kreps and Porteus (1978), Epstein and Zin (1989)), we also capture the empirically observed positive correlation between the price-dividend ratio and expected consumption growth, and the negative correlation between the price-dividend ratio and consumption volatility. In addition, in order to capture the notion that agents are more risk averse to large negative consum-
tion shocks than to small shocks (e.g., Pan (2002)), we generalize the standard recursive preferences by endowing agents with two risk aversion coefficients. These preferences are reminiscent of the “crash aversion” preferences of Bates (2008), but do not directly specify risk aversion as source-dependent. Indeed, in our framework jumps associated with small changes in consumption are subject to the same level of risk aversion as are diffusions.
References


A Appendix

A.1 Cash Holdings and the Term Structure of Net Payout Volatility

Other than equity and debt, cash holdings are also important components of firms’ capital structure. Since firms can also use cash to manipulate their payout policy, it is important to investigate whether adding cash changes the term structure of net payout volatility. Hence, we construct a measure of net payout plus increase in aggregate cash and compute the term structure of its volatility. It turns out that the term structure is even more downward sloping with cash.

We construct aggregate increase in cash using Table F.102 in the Flow of Funds data following Eisfeldt and Muir (2014). Increase in cash is defined as net acquisition of financial assets (Line 16) minus commercial paper as assets (Line 23) minus total mortgage as assets (Line 27) minus trade receivable as assets (Line 30) minus 2/3 times unidentified miscellaneous assets (Line 32).

Figure 10 reports the term structure of net payout volatility in comparison with the volatility term structure of net payout plus increase in cash. The term structure is even more downward sloping if we including cash. There is more short run cash flow risk and less long run cash flow risk if we include change in cash holdings.
Figure 10: Term Structure of Net Payout Volatility. This figure reports term structure of net payout volatility. The first one is the term structure of net payout volatility from 1927 to 2013. The second one is the volatility term structure of net payout plus increase in cash holdings from 1948 to 2013.

A.2 Calibration of the Extension of BKY

We rewrite the long run expected growth risk process as in Equation 5 to match the hump shape term structure of consumption volatility. Table 4 reports the calibrated parameter values. Among these parameters, a positive $\alpha_{cs}$ and a negative $\alpha_{c3}$ are crucial to generate the hump shape volatility term structure. In the model, we break consumption risk into short term, medium term, and long term components which are represented by $\Delta c_{t+1}$, $x_{t+1}$, and
$\bar{x}_{t+1}$. A positive correlation between short term and medium term consumption risk $\alpha_{es} = 1$ generates a upward sloping term structure of consumption volatility from short to medium horizons. Meanwhile, a negative correlation between short term and long term consumption risk $\alpha_{e\bar{x}} = -1$ generates a downward sloping term structure of consumption volatility from medium to long horizons. In addition, we choose $\mu_c$ to match average log consumption growth. $\bar{x}_{t+1}$ ($\rho_x = 0.95$) is set to be more persistent than $x_{t+1}$ ($\rho_x = 0.92$) since it is a longer term shock. We choose $\bar{\sigma} = .003$ to match average aggregate consumption volatility, $\rho_x = 0.999$ is from BKY. We set $\nu_x = 0.175$, $\nu_x = 0.1$, $\alpha_{e\bar{x}} = 0$, and $\nu_x = 1.26 \times 10^{-6}$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\mu_c$</th>
<th>$\rho_x$</th>
<th>$\nu_x$</th>
<th>$\alpha_{es}$</th>
<th>$\rho_{e\bar{x}}$</th>
<th>$\nu_{e\bar{x}}$</th>
<th>$\alpha_{e\bar{x}}$</th>
<th>$\bar{\sigma}$</th>
<th>$\rho_x$</th>
<th>$\nu_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.0015</td>
<td>0.92</td>
<td>0.175</td>
<td>1</td>
<td>0.95</td>
<td>0.1</td>
<td>-1</td>
<td>0.003</td>
<td>0.999</td>
<td>1.26 $\times 10^{-6}$</td>
</tr>
</tbody>
</table>

Table 4: Parameter Values for the Extension of BKY.

A.3 Generalized Recursive Preferences

We specify the preferences of the representative shareholder as a generalized form of Epstein and Zin (1989). In particular, the agent’s preferences are characterized by two risk aversion parameters ($\gamma_L$, $\gamma_H$) with relative “weights” ($w_L$, $w_H$) which are positive and sum to unity: $w_L + w_H = 1$:

$$U_t = \left\{ (1 - \beta)C_t^{1-\rho} + \beta \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_L} \left( \frac{w_L(1-\rho)}{1-\gamma_L} \right) \right] \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_H} \left( \frac{w_H(1-\rho)}{1-\gamma_H} \right) \right] \right)^{\frac{1}{1-\rho}} \right\}. \quad (54)$$
Here, $\rho = \frac{1}{\psi}$ is the inverse of the elasticity of intertemporal substitution. The special case $\rho = 1$ can be identified by first rewriting equation (54) as

$$e^{(1\rho) \log(U_t)} = (1 - \beta) e^{(1\rho) \log C_t} + \beta e^{\left( \frac{w_L (1\rho)}{1 - \gamma_L} \log \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right) \right] \right)} e^{\left( \frac{w_H (1\rho)}{1 - \gamma_H} \log \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right) \right] \right)},$$

(55)

and then taking a partial derivative with respect to $\rho$ before setting $\rho = 1$ to find:

$$\log(U_t) = (1 - \beta) \log C_t + \beta \left\{ \left( \frac{w_L}{1 - \gamma_L} \log \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right) \right] \right) \right\} + \left\{ \left( \frac{w_H}{1 - \gamma_H} \log \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right) \right] \right) \right\},$$

(56)

In order to identify a pricing kernel consistent with a given (optimal) consumption process, consider a security with current price $P_t$ and cash flows next period $X(\omega_{t+\Delta t})$. As is well known, along the optimal consumption path $\{C_t, C(\omega_{t+\Delta t})\}$, the following condition holds

$$0 = \frac{\partial}{\partial \epsilon} U_t \left( C_t - \epsilon P_t, \{C(\omega_{t+\Delta t}) + \epsilon X(\omega_{t+\Delta t})\} \right) \bigg|_{\epsilon = 0},$$

(57)

which can be rewritten as:

$$P_t = E_t \left[ M_{t,t+\Delta t} X_{t+\Delta t} \right]$$

(58)

where the pricing kernel is

$$M_{t,t+\Delta t} = \left( \frac{1}{\pi(\omega_{t+\Delta t})} \right) \frac{\partial U_i / \partial C(\omega_{t+\Delta t})}{\partial U_i / \partial C_t} = \left( \frac{1}{\pi(\omega_{t+\Delta t})} \right) \left( \frac{\partial U_i / \partial C(\omega_{t+\Delta t})}{\partial U_i / \partial C_t} \right).$$

(59)
From Equation (54), we determine

$$\frac{\partial U_t}{\partial C_t} = (1 - \beta) \left( \frac{U_t}{C_t} \right)^\rho \tag{60}$$

$$\frac{\partial U_t}{\partial (\omega_{t+\Delta t})} = U_t^\rho \beta \pi(\omega_{t+\Delta t}|\omega_i) \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right) \right] \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right) \right] \left( \frac{U_{t+\Delta t}^{1-\gamma_L}}{E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right)} \right) \left( \frac{U_{t+\Delta t}^{1-\gamma_H}}{E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right)} \right) \cdot \frac{\partial U_t}{\partial C_t} + \frac{\partial U_t}{\partial C_t} \left[ \frac{U_{t+\Delta t}^{1-\gamma_L}}{E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right)} \right] + \frac{\partial U_t}{\partial C_t} \left[ \frac{U_{t+\Delta t}^{1-\gamma_H}}{E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right)} \right] \tag{61}$$

Note that equation (60) holds for all dates:

$$\frac{\partial U(\omega_{t+\Delta t})}{\partial C(\omega_{t+\Delta t})} = \frac{\partial U(\omega_{t+\Delta t})}{\partial C(\omega_{t+\Delta t})} \tag{62}$$

Plugging in, we find that the pricing kernel is

$$M_{t, t+\Delta t} = \beta \left( \frac{C(\omega_{t+\Delta t})}{C_t} \right)^{-\rho} \left( \frac{U(\omega_{t+\Delta t})}{E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right) \left[ E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right) \right] \left( \frac{U_{t+\Delta t}^{1-\gamma_L}}{E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right)} \right) \left( \frac{U_{t+\Delta t}^{1-\gamma_H}}{E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right)} \right) \cdot \frac{\partial U_t}{\partial C_t} + \frac{\partial U_t}{\partial C_t} \left[ \frac{U_{t+\Delta t}^{1-\gamma_L}}{E_t \left( U_{t+\Delta t}^{1-\gamma_L} \right)} \right] + \frac{\partial U_t}{\partial C_t} \left[ \frac{U_{t+\Delta t}^{1-\gamma_H}}{E_t \left( U_{t+\Delta t}^{1-\gamma_H} \right)} \right] \right) \tag{63}$$

It is convenient to note that eq (54) is homogeneous of degree one in that if you scale all consumptions by a factor \(\lambda\), then \(U_t\) will increase by the same factor:

$$\lambda \times U_t(\{C(\omega_{t+\Delta t})\}) = U_t(\{\lambda \times C(\omega_{t+\Delta t})\}) \tag{64}$$

Taking a derivative wrt \(\lambda\), and then setting \(\lambda = 1\), it thus follows that

$$U_t = \sum \sum \left( \frac{\partial U_t}{\partial C(\omega_{t+\Delta t})} \right) C(\omega_{t+\Delta t}) \tag{65}$$
Dividing this last equation by \( \frac{\partial U_t}{\partial C_t} \) and using eqs. (63), we find

\[
\frac{U_t}{\partial U_t/\partial C_t} = \sum_s \sum_{\omega_{t+s}} \left( \frac{\partial U_t}{\partial C_t} \right) C(\omega_{t+s}) \\
= \sum_s E_s\left[ M_{t,t+s} C_{t+s} \right] \\
= W_t
\]

(66)

Using eq. (60), eq. (66) can be rewritten as

\[
\frac{U_t}{U_t^\rho (1-\beta) C_t^{-\rho}} = W_t, \quad (67)
\]

or equivalently,

\[
(1-\beta) \left( \frac{W_t}{C_t} \right) = \left( \frac{U_t}{C_t} \right)^{(1-\rho)}. \quad (68)
\]

That is, the log wealth-consumption ratio is linearly related to the log utility-consumption ratio:

\[
\log(1-\beta) + \log \left( \frac{W_t}{C_t} \right) = (1-\rho) \log \left( \frac{U_t}{C_t} \right). \quad (69)
\]

Consumption dynamics are given in equations 24. To proceed, we follow an approach similar to that of Bansal and Yaron (2004) by approximating the log utility-consumption ratio to be affine in the state vector \((\mu_t, V_t)\) (It can be shown that this relation is exact in the case \(\rho = 1\)):

\[
\log \left( \frac{U_t}{C_t} \right) \approx u_0 + u_\mu \mu_t - u_v V_t. \quad (70)
\]

58
Equivalently, this can be expressed as:

\[ U_t \approx e^{c_t + u_0 + u_\mu \mu_t - u_V V_t}. \] (71)

To identify the pricing kernel, we first determine

\[
E_t \left[ U^{(1-\gamma)}_{t+dt} \right] = U_t^{(1-\gamma)} e^{(1-\gamma) \left[ \frac{1}{2} V_t + \alpha_t \lambda_t + \alpha_t \pi_t \lambda_t + u_\mu (\kappa_t (\pi - \mu) - \sigma_t \nu_t \lambda_t) - u_V (\kappa_t (\pi - V_t) - \pi_t \lambda_t) \right] dt \\
\times E_t \left[ e^{(1-\gamma) d\tilde{q}} (-\hat{\eta}_t + u_\mu \tilde{\eta}_t - u_V \tilde{\eta}_V) \right] E_t \left[ e^{(1-\gamma) V d\mu (1 - \sigma_\mu u_\mu + \sigma_{Vt} u_V)} \right] \\
\times E_t \left[ e^{(1-\gamma) V d\mu (\sigma_{Vt} u_V - \sigma_\mu u_\mu)} \right] E_t \left[ e^{(1-\gamma) V (\sigma_\mu u_\mu + \sigma_{Vt} u_V)} \right] \\
\times e^{\lambda dt \left[ 1 + (1-\gamma) \pi_t (\alpha_\mu + u_\mu - u_\mu \alpha_\mu) \right]^{-1} - 1} \\
\times e^{\frac{1}{2} (1-\gamma)^2 V_t dt \left[ (1 - \sigma_\mu u_\mu - \sigma_{Vt} u_V)^2 + (\sigma_\mu u_\mu - \sigma_{Vt} u_V)^2 + (\sigma_{Vt} u_V)^2 \right]}. \] (72)

We note that the utility is finite only if the following condition holds:

\[ [1 + (1 - \gamma) \pi_t (\alpha_\mu + u_\mu - u_\mu \alpha_\mu)] > 0. \] (73)

Our parameter vector is chosen to satisfy this condition (and similar conditions noted below)

It follows that

\[
\frac{U^{(1-\gamma)}_{t+dt}}{E_t \left[ U^{(1-\gamma)}_{t+dt} \right]} = 1 + (1 - \gamma) \sqrt{V_t} d\mu \left( 1 - \sigma_\mu u_\mu + \sigma_{Vt} u_V \right) + (1 - \gamma) \sqrt{V_t} d\mu \left( \sigma_{Vt} u_V - \sigma_\mu u_\mu \right) \\
+ (1 - \gamma) \sqrt{V_t} d\mu \sigma_{Vt} u_V + dq \left[ e^{(1-\gamma) \left( -\hat{\eta}_t + u_\mu \tilde{\eta}_t - u_V \tilde{\eta}_V \right) - 1} \right] \\
- \lambda_t dt \left[ 1 + (1 - \gamma) \pi_t (\alpha_\mu + u_\mu - u_\mu \alpha_\mu) \right]^{-1} - 1. \] (74)
It is convenient to define the diffusive-average risk-aversion coefficient \( \varpi \equiv (w_L \gamma_L + w_H \gamma_H) \).

We can therefore express the last term of the pricing kernel as

\[
\text{4th term } \equiv w_L \frac{U_{i+dt}^{(1-\gamma_L)}}{E_t \left[ U_{i+dt}^{(1-\gamma_L)} \right]} + w_H \frac{U_{i+dt}^{(1-\gamma_H)}}{E_t \left[ U_{i+dt}^{(1-\gamma_H)} \right]} \\
= 1 + m_i \ dt - \theta_{i,c} \sqrt{V_i} \ dz_c - \theta_{i,\mu} \sqrt{V_i} \ dz_\mu - \theta_{i,v} \sqrt{V_i} \ dz_v + \left( \tilde{\Gamma}_i - 1 \right) \ dg, \quad (75)
\]

where we have defined

\[
m_i = -\lambda_i \left[ w_L \left[ \frac{1 + (1 - \gamma_L)}{\eta_c (\alpha_c + u_c - \alpha_c)} \right] - 1 \right] + w_H \left[ \frac{1 + (1 - \gamma_H)}{\eta_c (\alpha_c + u_c - \alpha_c)} \right] - 1 \right] \\
\theta_{i,c} = (\varpi - 1) \left( 1 - \sigma_{c,\mu} u_\mu + \sigma_{c,v} u_v \right) \\
\theta_{i,\mu} = (\varpi - 1) \left( \sigma_{\mu,v} u_v - \sigma_{\mu,\mu} \right) \\
\theta_{i,v} = (\varpi - 1) \sigma_{v,v} \\
\tilde{\Gamma}_i = w_L e^{(1-\gamma_L)} (-\tilde{\eta}_c + u_c \tilde{\eta}_c - \tilde{\eta}_c \tilde{\eta}_c) + w_H e^{(1-\gamma_H)} (-\tilde{\eta}_c + u_c \tilde{\eta}_c - \tilde{\eta}_c \tilde{\eta}_c). \quad (76)
\]

The second and third terms for \((i = (L, H))\) are

\[
\text{2nd, 3rd term } = \left( \frac{U_{i+dt}^{1-\gamma_i}}{E_t \left[ U_{i+dt}^{1-\gamma_i} \right]} \right)^{w_i (\rho-1)} \\
= 1 + m_i \ dt - \theta_{i,c} \sqrt{V_i} \ dz_c - \theta_{i,\mu} \sqrt{V_i} \ dz_\mu - \theta_{i,v} \sqrt{V_i} \ dz_v + \left( \tilde{\Gamma}_i - 1 \right) \ dg. \quad (77)
\]

60
where

\[ m_i = \frac{1}{2} V_t w_i (\rho - 1) \left[ w_i (\rho - 1) - (1 - \gamma_i) \right] \left[ (1 - \sigma_{\nu^i} u_\nu + \sigma_{\nu^i} u_\nu)^2 + (\sigma_{\nu^i} u_\nu - \sigma_\nu u_\nu)^2 + (\sigma_\nu u_\nu)^2 \right] 
- \lambda_i \left( \frac{w_i (\rho - 1)}{1 - \gamma_i} \right) \left[ 1 + (1 - \gamma_i) \bar{\beta}_V (\alpha_\varepsilon + u_\nu - u_\nu \alpha_\mu) \right]^{-1} - 1 \]

\[ \theta_{i,\varepsilon} = w_i (1 - \rho) \left( 1 - \sigma_{\nu^i} u_\nu + \sigma_{\nu^i} u_\nu \right) \]

\[ \theta_{i,\mu} = w_i (1 - \rho) \left( \sigma_{\nu^i} u_\nu - \sigma_\nu u_\nu \right) \]

\[ \theta_{i,V} = w_i (1 - \rho) \sigma_\nu u_\nu \]

\[ \Gamma_i = e^{w_i (\rho - 1) \left( -\tilde{\varepsilon}_c + u_\mu \tilde{\varepsilon}_c - u_\nu \tilde{\varepsilon}_c \right)}. \quad (78) \]

The first term is

\[ e^{-\rho \tilde{\varepsilon}_c} = 1 + m_\rho \ dt - \rho \sqrt{V_t} \ dz_c + dq \left( e^{m_\rho} - 1 \right). \quad (79) \]

where

\[ m_\rho = -\rho \left( \frac{1}{2} V_t + \alpha_\lambda \lambda_i + \alpha_\varepsilon \bar{\beta}_V \lambda_i - \frac{\rho}{2} V_t \right) \ dt \quad (80) \]

Combining all of these terms, we identify pricing kernel dynamics:

\[ \frac{d\Lambda}{\Lambda} = M_{i,t+dt} - 1 \quad (81) \]

\[ = -r_i \ dt - \theta_c \sqrt{V_t} d\varepsilon_c - \theta_\mu \sqrt{V_t} d\varepsilon_\mu - \theta_\nu \sqrt{V_t} d\varepsilon_\nu + \left( \tilde{\Gamma}_\lambda - 1 \right) dq - \left( E_t \tilde{\Gamma}_\lambda - 1 \right) \lambda_i \ dt. \]
where

\[ \theta_c = \rho + \theta_{L,c} + \theta_{H,c} + \theta_{4,c} \]
\[ = \rho + (\gamma - \rho) \left[ 1 - \sigma_{\mu c} u_\mu + \sigma_{Vc} u_V \right] \]

\[ \theta_\mu = \theta_{L,\mu} + \theta_{H,\mu} + \theta_{4,\mu} \]
\[ = (\gamma - \rho) \left[ \sigma_{\mu V} u_V - \sigma_\mu u_\mu \right] \]

\[ \theta_V = \theta_{L,V} + \theta_{H,V} + \theta_{4,V} \]
\[ = (\gamma - \rho) \left[ \sigma_V u_V \right] \]

\[ \tilde{\Gamma}_\Lambda = e^{\rho \tilde{c}} \tilde{\Gamma}_L \tilde{\Gamma}_H \tilde{\Gamma}_4 \]

\[ E_t \left[ \tilde{\Gamma}_\Lambda \right] = \frac{w_L}{\eta_V} \left( \frac{1}{\eta_V} - (\rho - \gamma_L) u_V - \alpha_c \gamma_L - (\rho - \gamma_L) u_\mu \alpha_\mu \right)^{-1} \]
\[ + \frac{w_H}{\eta_V} \left( \frac{1}{\eta_V} - (\rho - \gamma_H) u_V - \alpha_c \gamma_H - (\rho - \gamma_H) u_\mu \alpha_\mu \right)^{-1} \]

\[ E_t \left[ \tilde{\Gamma}_\Lambda^2 \right] = \frac{w_L^2}{\eta_V} \left( \frac{1}{\eta_V} - 2\alpha_c \gamma_L - 2(\gamma_L - \rho) \left( u_\mu \alpha_\mu - u_V \right) \right)^{-1} \]
\[ + \frac{w_H^2}{\eta_V} \left( \frac{1}{\eta_V} - 2\alpha_c \gamma_H - 2(\gamma_H - \rho) \left( u_\mu \alpha_\mu - u_V \right) \right)^{-1} \]
\[ + \frac{2w_L w_H}{\eta_V} \left( \frac{1}{\eta_V} - \alpha_c (\gamma_L + \gamma_H) + (\gamma_L + \gamma_H - 2\rho) \left( u_\mu \alpha_\mu - u_V \right) \right)^{-1}. \quad (82) \]

and the risk free rate is

\[ r_t = r_0 + \rho \mu_t - r_V V_t, \quad (83) \]
where

\[ r_o = \delta + \rho \lambda_o (\alpha_l + \alpha_c \tilde{\eta}_v) - \lambda_o (\gamma_L - \rho) \left( \frac{w_L}{1 - \gamma_L} \right) \left[ \frac{1 + (1 - \gamma_L)\tilde{\eta}_v (\alpha_c + u_L - u_\mu \alpha_\mu)}{1 - \gamma_L} \right] - 1 \]

\[ - \lambda_o (\gamma_H - \rho) \left( \frac{w_H}{1 - \gamma_H} \right) \left[ \frac{1 + (1 - \gamma_H)\tilde{\eta}_v (\alpha_c + u_H - u_\mu \alpha_\mu)}{1 - \gamma_H} \right] - 1 \]

\[ - \lambda_o \frac{w_L}{\tilde{\eta}_v} \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_L)u_L - \alpha_c \gamma_L - (\rho - \gamma_L)u_\mu \alpha_\mu \right)^{-1} \]

\[ - \lambda_o \frac{w_H}{\tilde{\eta}_v} \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_H)u_H - \alpha_c \gamma_H - (\rho - \gamma_H)u_\mu \alpha_\mu \right)^{-1} + \lambda_o, \] \hspace{1cm} (84) \]

\[ r_v = \frac{\rho}{2} (1 + \rho) - \rho \lambda_v (\alpha_l + \alpha_c \tilde{\eta}_v) + \frac{1}{2} (\rho - 1) \left[ (w_L^2 + w_H^2) (\rho - 1) + \gamma - 1 \right] \]

\[ \times \left[ (1 - \sigma_{\mu} u_\mu + \sigma_{\nu} u_\nu)^2 + (\sigma_{\nu \mu} u_\nu - \sigma_{\mu} u_\mu)^2 + (\sigma_{\nu} u_\nu)^2 \right] \]

\[ + \lambda_v (\gamma_L - \rho) \left( \frac{w_L}{1 - \gamma_L} \right) \left[ \frac{1 + (1 - \gamma_L)\tilde{\eta}_v (\alpha_c + u_L - u_\mu \alpha_\mu)}{1 - \gamma_L} \right] - 1 \]

\[ + \lambda_v (\gamma_H - \rho) \left( \frac{w_H}{1 - \gamma_H} \right) \left[ \frac{1 + (1 - \gamma_H)\tilde{\eta}_v (\alpha_c + u_H - u_\mu \alpha_\mu)}{1 - \gamma_H} \right] - 1 \]

\[ + \rho \theta_{v,c} + \rho \theta_{v,c} + \theta_{L,c \theta_{H,c}} + \theta_{L,c \theta_{H,\nu}} + \theta_{L,c \theta_{H,c}} + \theta_{H,c \theta_{4,c}} \]

\[ + \theta_{L,c \theta_{H,\nu}} + \theta_{L,\nu \theta_{4,\nu}} + \theta_{L,\nu \theta_{H,\nu}} + \theta_{L,\nu \theta_{4,\nu}} + \theta_{H,\nu \theta_{4,\nu}} \]

\[ + \lambda_v \frac{w_L}{\tilde{\eta}_v} \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_L)u_L - \alpha_c \gamma_L - (\rho - \gamma_L)u_\mu \alpha_\mu \right)^{-1} \]

\[ + \lambda_v \frac{w_H}{\tilde{\eta}_v} \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_H)u_H - \alpha_c \gamma_H - (\rho - \gamma_H)u_\mu \alpha_\mu \right)^{-1} - \lambda_v, \] \hspace{1cm} (85) \]

The instantaneous interest rate volatility can be identified by noting that

\[ dr_t \mid_{\text{stoch}} = \sqrt{V} dz_c \left( r_v \sigma_{v,c} - \rho \sigma_{\mu} \right) + \sqrt{V} dz_\mu \left( r_v \sigma_{v,\mu} - \rho \sigma_{\mu} \right) + \sqrt{V} dz_v \left( r_v \sigma_{v,v} \right) + dq \left( \rho \tilde{\eta}_v - r_v \tilde{\eta}_v \right), \]

and hence

\[ \frac{1}{dt} \mathbb{E}_t [dr^2] = V_t \left[ (r_v \sigma_{v,c} - \rho \sigma_{\mu})^2 + (r_v \sigma_{v,\mu} - \rho \sigma_{\mu})^2 + (r_v \sigma_{v,v})^2 \right] + 2 \lambda_v \tilde{\eta}_v^2 \left( \rho \sigma_{\mu} - r_v \right)^2. \] \hspace{1cm} (86)
A.4 Wealth Process

Note that by combining equations (69) and (71), we obtain

\[ W_t = e^{-\delta_t(1-\rho)\left(u_0 + u_\mu, i - u_V V_i\right)} \]

(87)

Ito’s lemma (extended for jumps) gives

\[ \frac{dW}{W} = \mu_W dt + \theta_{w_\mu} \sqrt{V_i} dz_c + \theta_{w_\mu} \sqrt{V_i} dz_\mu + \theta_{w_\upsilon} \sqrt{V_i} dz_\upsilon + dq(G_w - 1) \]

(88)

where we have defined:

\[
\begin{align*}
\mu_W &\equiv \mu_W + \mu_W \mu_i + \mu_W V_i \\
\mu_{W_0} &\equiv \alpha_\lambda \lambda_0 + \alpha_c \eta_\lambda \lambda_0 + (1-\rho) u_\mu\left(\kappa_\mu \bar{\mu} - \alpha_\mu \eta_\lambda \lambda_0\right) - (1-\rho) u_\upsilon \left(\kappa_\upsilon \bar{\upsilon} - \eta_\lambda \lambda_0\right) \\
\mu_W V &\equiv -\frac{1}{2} + \alpha_\lambda \lambda_0 + \alpha_c \eta_\lambda \lambda_0 + (1-\rho) u_\mu\left(-\alpha_\mu \eta_\lambda \lambda_0\right) - (1-\rho) u_\upsilon \left(-\kappa_\upsilon \eta_\lambda \lambda_0\right) \\
&\quad + \frac{1}{2} \left[1 + (1-\rho) u_\nu \sigma_{\nu} - (1-\rho) u_\mu \sigma_{\mu}\right]^2 + \frac{(1-\rho)^2}{2} \left[u_\nu \sigma_{\nu} - u_\mu \sigma_{\mu}\right]^2 + \frac{(1-\rho)^2}{2} \left[u_\nu \sigma_{\nu}\right]^2 \\
\mu_{W_\mu} &\equiv 1 - (1-\rho) u_\mu \kappa_\mu \\
\theta_{W_\mu} &\equiv \left[1 + (1-\rho) u_\nu \sigma_{\nu} - (1-\rho) u_\mu \sigma_{\mu}\right] \\
\theta_{W_\nu} &\equiv (1-\rho) \left[u_\nu \sigma_{\nu} - u_\mu \sigma_{\mu}\right] \\
\theta_{W_\upsilon} &\equiv (1-\rho) u_\nu \sigma_{\nu} \\
\Gamma_w &\equiv e^{-\delta_t(1-\rho) u_\mu \delta_\mu - (1-\rho) u_\nu \delta_\nu} \\
E_t[\Gamma_w] &\equiv \left(\frac{1}{\eta_\upsilon}\right) \left[\frac{1}{\eta_\upsilon} + \alpha_\lambda + (1-\rho) u_\nu - (1-\rho) u_\mu \alpha_\mu\right]^{-1}. \\
\end{align*}
\]

(89)

In order to identify the parameters \((u_0, u_\mu, u_\nu)\), note that wealth is the claim to con-
assumption, and therefore satisfies

\[ 0 = \frac{1}{dt} E_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dW}{W} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dW}{W} \right) \right] + e^{-z_t}, \]  

(90)

where \( z_t = \log \left( \frac{w_t}{c_t} \right) \) is the log wealth-consumption ratio. As demonstrated by Bansal and Yaron (2004), log-linearization provides an excellent approximation:

\[ 0 \approx \frac{1}{dt} E_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dW}{W} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dW}{W} \right) \right] + e^{-\bar{z}} [1 - (z_t - \bar{z})], \]  

(91)

where equation (87) implies that

\[ z_t \approx -\log \delta + (1 - \rho) \left( u_o + u_\mu \mu_t - u_\nu V_t \right). \]  

(92)

and hence

\[ \bar{z} \approx -\log \delta + (1 - \rho) \left( u_o + u_\mu \bar{\mu} - u_\nu \bar{V} \right). \]  

(93)

The individual components of equation 91 are:

\[ \frac{1}{dt} E_t \left[ \frac{d\Lambda}{\Lambda} \right] = -(r_o + \rho \mu_t - r_\nu V_t) \]

\[ \frac{1}{dt} E_t \left[ \frac{dW}{W} \right] = \mu_w + \lambda_t \left( E \left[ \Gamma_w \right] - 1 \right) \]

\[ \frac{1}{dt} E_t \left[ \frac{d\Lambda dW}{\Lambda W} \right] = -V_t \left( \theta_{w_c} \theta_c + \theta_{w_\mu} \mu_t + \theta_{w_\nu} \nu_t \right) + \lambda_t E \left[ \Gamma_w \Gamma_\Lambda - \Gamma_w - \Gamma_\Lambda + 1 \right]. \]  

(94)
where

\[
E \left[ \Gamma_w \Gamma_\lambda \right] = \left( \frac{w_L}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + (\gamma_L - 1) (-\alpha_e - u_v + u_\mu \alpha_\mu) \right]^{-1} + \left( \frac{w_H}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + (\gamma_H - 1) (-\alpha_e - u_v + u_\mu \alpha_\mu) \right]^{-1}. 
\]  

(95)

Plugging equations (81) and (88) into equation (91) and then collecting terms linear in \( \mu_t \), \( V_t \), and terms independent of the state vector, we find the three equations which, combined with equation (93) identify the parameters \((u_o, u_\mu, u_v, z)\):

\[
\mu_t : 0 = -\rho + \mu_\mu - e^{-\bar{z}}(1 - \rho)u_\mu \\
V_t : 0 = r_v + \mu_\nu + \lambda_v \left( \frac{w_L}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + (\gamma_L - 1) (-\alpha_e - u_v + u_\mu \alpha_\mu) \right]^{-1} + \lambda_v \left( \frac{w_H}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + (\gamma_H - 1) (-\alpha_e - u_v + u_\mu \alpha_\mu) \right]^{-1} \left( \theta_w \theta_v + \theta_w \theta_v + \theta_w \theta_v \right) \right. \\

- \lambda_v \left( \frac{w_L}{\tilde{\eta}_v} \right) \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_L)u_v - \alpha_e \gamma_L - (\rho - \gamma_L)u_\mu \alpha_\mu \right)^{-1} \left. + e^{-\bar{z}}(1 - \rho)u_v \right) \\
const : 0 = -r_o + \mu_\omega + \lambda_0 \left( \frac{w_L}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + (\gamma_L - 1) (-\alpha_e - u_v + u_\mu \alpha_\mu) \right]^{-1} \left. + \lambda_0 \left( \frac{w_H}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + (\gamma_H - 1) (-\alpha_e - u_v + u_\mu \alpha_\mu) \right]^{-1} \right. \\

- \lambda_0 \left( \frac{w_L}{\tilde{\eta}_v} \right) \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_L)u_v - \alpha_e \gamma_L - (\rho - \gamma_L)u_\mu \alpha_\mu \right)^{-1} \left. + \lambda_0 \left( \frac{w_H}{\tilde{\eta}_v} \right) \left( \frac{1}{\tilde{\eta}_v} + (\rho - \gamma_H)u_v - \alpha_e \gamma_H - (\rho - \gamma_H)u_\mu \alpha_\mu \right)^{-1} \right. \\

+ e^{-\bar{z}}(1 + \bar{z} + \log \delta - (1 - \rho)u_o). 
\]  

(96)
Plugging in for $\mu_{w_\mu}$, we find

$$ u_\mu = \frac{1}{e^{-\tau} + \kappa_\mu}. \quad (97) $$

The other roots are more complicated, but are easily found via numerical root finding. Indeed, there are in general two roots for $u_\nu$. However, following the insights of ??, the root whose value remains finite in the limit of the volatility and intensity parameters going to zero is the economically relevant one.

### A.5 Variance Ratios

Here we investigate the consumption variance ratios for the wealthy shareholder. Given consumption dynamics in equations 24 are affine, it follows that the date-$t$ expectations of the first two moments of date-$T$ consumption possess exponential-affine forms:

$$ E_t [e^{c_T}] = e^{c_t + A_{0.1}(T-t) + \mu_t A_{\mu,1}(T-t) + \nu_t A_{\nu,1}(T-t)} $$

$$ E_t [e^{2c_T}] = e^{2c_t + A_{0.2}(T-t) + \mu_t A_{\mu,2}(T-t) + \nu_t A_{\nu,2}(T-t)}. \quad (98) \quad (99) $$

where the deterministic coefficients $A_{(\cdot,d)}$ can be identified by noting that $G^3(t, c_t, \mu_t, \nu_t) = E_t [e^{3c_T}]$ is a P-martingale, implying its expected change is zero. Thus we find

$$ 0 = G_t + G_c \left[ \mu - \frac{1}{2} V + \alpha_\lambda \lambda_t + \alpha_\nu \tilde{\nu}_t \lambda_t \right] + G_{\mu} \left[ \kappa_\mu (\bar{\mu} - \mu) - \alpha_\mu \tilde{\nu}_t \lambda_t \right] + G_{\nu} \left[ \kappa_\nu (\bar{V} - V) - \tilde{\nu}_t \lambda_t \right] $$

$$ + \frac{1}{2} G_{\sigma c} V + \frac{1}{2} G_{\sigma \mu} V \left( \sigma_{\mu c}^2 + \sigma_{\mu}^2 \right) + \frac{1}{2} G_{\sigma \nu} V \left( \sigma_{\nu c}^2 + \sigma_{\nu}^2 + \sigma_{\nu \lambda}^2 \right) - \sigma_{\mu c} G_{\sigma \mu} V - \sigma_{\nu c} G_{\sigma \nu} V $$

$$ + G_{\mu V} \left( \sigma_{\mu c} \sigma_{\lambda c} + \sigma_{\mu c} \sigma_{\lambda} \right) + \lambda \tilde{\nu}_t \left[ G(c - \tilde{\nu}_c, \mu + \tilde{\mu}_c, V + \tilde{\nu}_c) - G(c, \mu, V) \right]. \quad (100) $$
Collecting terms linear in $\mu_t$, $V_t$, and independent of $(\mu_t, V_t)$, and defining $\tau \equiv (T - t)$, we identify the three equations

\begin{align}
A'_{\mu,\beta}(\tau) &= \beta - \kappa_\mu A_{\mu,\beta} \\
A'_{V,\beta}(\tau) &= -\frac{1}{2}\beta + \beta \alpha_\lambda \lambda_V + \beta \alpha_\mu \lambda_V - A_{\mu,\beta} \alpha_\mu \lambda_V - \kappa_\lambda A_{V,\beta} - A_{V,\beta} \lambda_V + \frac{\beta^2}{2} \\
&\quad + \frac{1}{2} A^2_{\mu,\beta} \left( \sigma^2_{\mu,\epsilon} + \sigma^2_{\mu} \right) + \frac{1}{2} A^2_{V,\beta} \left( \sigma^2_{V,\epsilon} + \sigma^2_{V} \right) - \beta A_{\mu,\beta} \sigma_{\mu,\epsilon} - \beta A_{V,\beta} \sigma_{V,\epsilon} \\
&\quad + A_{\mu,\beta} A_{V,\beta} \left( \sigma_{\mu,\epsilon} \sigma_{V,\epsilon} + \sigma_{\mu} \sigma_{V} \right) + \lambda_V \left[ \frac{1}{\eta_V} \left( \frac{1}{\eta_V} + \beta \alpha_\epsilon - \alpha_\mu A_{\mu,\beta} - A_{V,\beta} \right)^{-1} - 1 \right] \\
A'_{0,\beta}(\tau) &= \beta \alpha_\lambda \lambda_0 + \beta \alpha_\mu \lambda_0 + \kappa_\mu A_{\mu,\beta} - A_{\mu,\beta} \alpha_\mu \lambda_0 - \kappa_\mu \lambda A_{V,\beta} - A_{V,\beta} \lambda_0 \\
&\quad + \lambda_0 \left[ \frac{1}{\eta_V} \left( \frac{1}{\eta_V} + \beta \alpha_\epsilon - \alpha_\mu A_{\mu,\beta} - A_{V,\beta} \right)^{-1} - 1 \right],
\end{align}

subject to the “initial” conditions $A_{.,\beta}(\tau = 0) = 0$. The only functions with analytic solutions are

\begin{equation}
A_{\mu,\beta} = \left( \frac{\beta}{\kappa_\mu} \right) \left[ 1 - e^{-\kappa_\mu \tau} \right].
\end{equation}

The other functions must be calculated numerically. However, that can be done extremely quickly.

We use these functions for $\beta = (1, 2)$ to identify the term structure of consumption volatilities:

\begin{equation}
\sigma_{c,\tau} = \sqrt{\left( \frac{1}{\tau} \right) \log \left[ \frac{E_0 \left[ e^{2(c_{\tau} - c_0)} \right]}{\left(E_0 \left[ e^{c_{\tau} - c_0} \right] \right)^2} \right]}
\end{equation}
A.6 Net Payout Dynamics

We specify the dynamics of log net payout \( x_t = \log X_t \) to be stationary with log-consumption:

\[
\begin{align*}
dx = & \left[ \kappa_x \left( c_t + \phi_x - x \right) + \phi_x \left( \mu - \mu_t \right) + \phi_v \left( V - \bar{V} \right) \right] dt + \sigma_{xv} \sqrt{V_t} dz_e + \sigma_{xv} \sqrt{V_t} dz_v \\
& + \sigma_{xv} \sqrt{V_t} dz_v + \sigma_x \sqrt{V_t} dz_e - \eta_x dq + \alpha_x \eta_v \lambda_t dt, \\
\end{align*}
\]

(105)

where the jumps are conditionally correlated

\[
\pi \left( \eta_e, \eta_\mu, \eta_v, \eta_x \right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\eta_e^2}{V}} 1_{(\eta_e > 0)} \delta \left( \eta_\mu - \alpha_v \eta_v \right) \delta \left( \eta_e - \alpha_x \eta_x \right) \delta \left( \eta_v - \alpha_v \eta_v \right). \tag{106}
\]

Here we investigate net payout variance ratios. Given that net payout dynamics are affine, it follows that the date- \( t \) expectation possesses an exponential-affine form:

\[
E_t \left[ e^{\beta x_T} \right] = e^{A_0,\beta (T-t) + c_t A_{c,\beta} (T-t) + x_t A_{x,\beta} (T-t) + \mu_t A_{\mu,\beta} (T-t) + V_t A_{V,\beta} (T-t)}. \tag{107}
\]

The deterministic coefficients \( A_{(\cdot,\beta)} \) can be identified by noting that \( G^\beta(t, c_t, x_t, \mu_t, V_t) = E_t \left[ e^{\beta x_T} \right] \) is a P-martingale, implying its expected change is zero. Thus we find

\[
0 = G_t + G_x \left[ \mu - \frac{1}{2} V + \alpha_x \lambda_t + \alpha_v \bar{\eta}_v \lambda_t \right] + G_\mu \left[ \kappa_\mu \left( \mu - \mu_t \right) - \alpha_v \bar{\eta}_v \lambda_t \right] + G_v \left[ \kappa_v \left( V - \bar{V} \right) - \bar{\eta}_v \lambda_t \right] + G_x \left[ \kappa_x \left( c_t + \phi_x - x \right) + \phi_x \left( \mu - \mu_t \right) + \phi_v \left( V - \bar{V} \right) + \alpha_v \bar{\eta}_v \lambda_t \right] + \frac{1}{2} G_{xx} V + \frac{1}{2} G_{xv} V \left( \sigma_{xc}^2 + \sigma_{x\mu}^2 + \sigma_{xv}^2 + \sigma_v^2 \right) + \frac{1}{2} G_{\mu v} V \left( \sigma_{\mu c}^2 + \sigma_{\mu v}^2 + \sigma_v^2 \right) + \sigma_{xc} G_{ex} V - \sigma_{xc} G_{ex} V - \sigma_{xc} G_{ex} V - G_{xv} V \left( \sigma_{xc} \sigma_{x\mu} + \sigma_{xv} \sigma_v \right) + G_{xv} V \left( \sigma_{xc} \sigma_{xv} + \sigma_{x\mu} \sigma_v + \sigma_{xv} \sigma_{\mu} \right) + G_{\mu v} V \left( \sigma_{\mu c} \sigma_{xv} + \sigma_{\mu v} \sigma_{\mu} \right) + \lambda E_t \left[ G(c - \bar{\eta}_c, x - \bar{\eta}_x, \mu + \bar{\eta}_\mu, V + \bar{\eta}_v) - G(c, \mu, V) \right]. \tag{108}
\]
Collecting terms linear in $c_t, x_t, \mu_t, V_t$, and independent of $(c_t, x_t, \mu_t, V_t)$, and defining $\tau \equiv (T - t)$, we identify the five equations

$$A'_{x,\beta}(\tau) = -\kappa_x A_{x,\beta}$$

$$A'_{c,\beta}(\tau) = \kappa_x A_{c,\beta}$$

$$A'_{\mu,\beta}(\tau) = A_{c,\beta} - \kappa_\mu A_{\mu,\beta} + \phi_\mu A_{x,\beta}$$

$$A'_{V,\beta}(\tau) = -\frac{1}{2} A_{e,\beta} + A_{e,\beta} \alpha_x \lambda_V + \phi_V A_{x,\beta} + A_{e,\beta} \alpha_x \bar{\eta}_V \lambda_V + A_{x,\beta} \alpha_x \bar{\eta}_V \lambda_V - A_{\mu,\beta} \alpha_\mu \bar{\eta}_V \lambda_V$$

$$-\kappa_V A_{V,\beta} - A_{V,\beta} \bar{\eta}_V \lambda_V + \frac{1}{2} A_{e,\beta}^2 + \frac{1}{2} A_{x,\beta}^2 \left( \sigma_{x,e}^2 + \sigma_{x,\mu}^2 + \sigma_{x,V}^2 \right) + \frac{1}{2} A_{\mu,\beta}^2 \left( \sigma_{\mu,e}^2 + \sigma_{\mu,\mu}^2 + \sigma_{\mu,V}^2 \right) + A_{c,\beta} A_{x,\beta} \sigma_{x,e} - A_{c,\beta} A_{\mu,\beta} \sigma_{\mu,e}$$

$$-A_{c,\beta} A_{V,\beta} \sigma_{V,e} - A_{x,\beta} A_{\mu,\beta} \left( \sigma_{x,e} \sigma_{x,\mu} + \sigma_{x,\mu} \sigma_{x,V} \right) - A_{x,\beta} A_{V,\beta} \left( \sigma_{x,e} \sigma_{x,V} + \sigma_{x,\mu} \sigma_{x,V} + \sigma_{x,V} \sigma_V \right)$$

$$+ A_{x,\beta} A_{V,\beta} \left( \sigma_{x,e} \sigma_{V,e} + \sigma_{x,\mu} \sigma_{V,e} \right) + \lambda_V \left[ \frac{1}{\bar{\eta}_V} \left( \frac{1}{\bar{\eta}_V} + \alpha_c A_{c,\beta} + \alpha_x A_{x,\beta} - \alpha_{\mu} A_{\mu,\beta} - A_{V,\beta} \right)^{-1} - 1 \right]$$

$$A'_{0,\beta}(\tau) = A_{e,\beta} \alpha_x \lambda_0 + A_{e,\beta} \alpha_x \bar{\eta}_V \lambda_0 - \phi_V A_{x,\beta} \bar{V} + A_{x,\beta} \kappa_x \phi_x + A_{x,\beta} \alpha_x \bar{\eta}_V \lambda_0 + \kappa_\mu \bar{\eta} A_{\mu,\beta}$$

$$-A_{\mu,\beta} \alpha_x \bar{\eta}_V \lambda_0 + \kappa_V \bar{V} A_{V,\beta} - A_{V,\beta} \bar{\eta}_V \lambda_0$$

$$+ \lambda_0 \left[ \frac{1}{\bar{\eta}_V} \left( \frac{1}{\bar{\eta}_V} + \alpha_c A_{c,\beta} + \alpha_x A_{x,\beta} - \alpha_{\mu} A_{\mu,\beta} - A_{V,\beta} \right)^{-1} - 1 \right]$$

subject to the “initial” conditions $A_{x,\beta}(\tau = 0) = \beta$, and all other initial conditions $A_{c,\beta}(\tau = 0) = 0$. The functions with analytic solutions are

$$A_{x,\beta} = \beta e^{-\kappa_x \tau}.$$  

$$A_{c,\beta} = \beta \left(1 - e^{-\kappa_x \tau} \right).$$

$$A_{\mu,\beta} = \left( \frac{\beta}{\kappa_\mu} \right) \left[1 - e^{-\kappa_\mu \tau}\right] + \left( \phi_\mu - 1 \right) \left( \frac{\beta}{\kappa_\mu - \kappa_x} \right) \left[ e^{-\kappa_x \tau} - e^{-\kappa_\mu \tau} \right].$$
quickly.

We use these functions for $\beta = (1, 2)$ to identify the term structure of net payout volatilities:

$$\sigma_{x, \tau} = \sqrt{\left(\frac{1}{\tau}\right) \log \left[ \frac{E_0 \left[e^{\mu(x_\tau - x_0)}\right]}{(E_0 \left[e^{x_\tau - x_0}\right])^2} \right]} \tag{114}$$

### A.7 Pricing Net Payout Strips

The date-$t$ price $P_T(t, c_t, x_t, \mu_t, V_t)$ of the claim to $e^{x_\tau}$ satisfies the equation

$$\Lambda_t P_T(t, c_t, x_t, \mu_t, V_t) = E_t [\Lambda_T e^{x_\tau}]. \tag{115}$$

This implies that $\Lambda_T P_T(t, c_t, x_t, \mu_t, V_t)$ is a $P$-martingale, implying that its expected change is zero:

$$0 = \frac{1}{d\Lambda} E_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dP}{P} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dP}{P} \right) \right]. \tag{116}$$

Given that state vector dynamics are affine, the solution takes the form:

$$P_T(t, c_t, x_t, \mu_t, V_t) = e^{F_0(T-t) + F_c(T-t)c_t + F_x(T-t)x_t + F_\mu(T-t)\mu_t - F_V(T-t)V_t}. \tag{117}$$

Now, from Itô’s lemma, we have

$$\frac{dP}{P} = \mu_{(P,T)} dt + \sqrt{V_t} dz_c \left( F_c + F_x \sigma_{xc} - F_\mu \sigma_{\mu c} + F_V \sigma_{Vc} \right) + \sqrt{V_t} dz_\mu \left( F_x \sigma_{x\mu} - F_\mu \sigma_{\mu \mu} + F_V \sigma_{V\mu} \right) + \sqrt{V_t} dz_V \left( F_c \sigma_{cV} + F_V \sigma_{Vc} \right) + \sqrt{V_t} dz_x \left( F_x \sigma_{x} \right) + dq \left[ e^{-F_c \delta_c - F_x \delta_x + F_\mu \delta_\mu - F_V \delta_V - 1} \right] - \lambda \ dt \ E_t \left[ e^{-F_c \delta_c - F_x \delta_x + F_\mu \delta_\mu - F_V \delta_V - 1} \right]. \tag{118}$$
where we have defined

\[
\mu_{(e \ell)} \equiv -F''_0 - c F'_e - x F'_x - \mu F'_\mu + V F'_V + F_e \left[ \mu - \frac{1}{2} V + \alpha_s \lambda_e + \alpha_s \bar{\eta}_V \lambda_e \right] + F_\mu \left[ \kappa_\mu (\bar{\mu} - \mu) - \alpha_\mu \bar{\eta}_V \lambda_e \right] \\
+ F_x \left[ \kappa_x (c_e + \phi_x - x) + \phi_\mu (\mu - \bar{\mu}) + \phi_V (V - \bar{V}) + \alpha_s \bar{\eta}_V \lambda_e \right] - F_V \left[ \kappa_V (V - \bar{V}) - \bar{\eta}_V \lambda_e \right] \\
+ \frac{1}{2} F'^2 e V + \frac{1}{2} F'^2 x V \left( \sigma_{x e}^2 + \sigma_{x \mu}^2 + \sigma_{x V}^2 \right) + \frac{1}{2} F'^2 \left( \sigma_{\mu e}^2 + \sigma_{\mu V}^2 \right) + \frac{1}{2} F'^2 V \left( \sigma_{e e}^2 + \sigma_{e \mu}^2 + \sigma_{e V}^2 \right) \\
+ \sigma_{x e} F_e F_x V - \sigma_{x \mu} F_e F_\mu V + \sigma_{x V} F_e F_V V - F_x F_\mu V \left( \sigma_{x e} \sigma_{x \mu} + \sigma_{x V} \sigma_{x V} \right) - F_e F_\mu V \left( \sigma_{e \mu} \sigma_{e e} + \sigma_{e \mu} \sigma_{e V} \right) \\
+ \lambda \varepsilon_t \left[ e^{-F_e \bar{\eta}_e - F_x \bar{\eta}_x + F_\mu \bar{\eta}_\mu - F_V \bar{\eta}_V} - 1 \right]. \quad (119)
\]

This implies that:

\[
\frac{dP}{dA} \frac{dA}{\Lambda} = -\theta_e V dt \left[ F_e + F_x \sigma_{e x} - F_\mu \sigma_{e \mu} + F_V \sigma_{e V} \right] - \theta_\mu V dt \left[ F_\mu \sigma_{x \mu} - F_\mu \sigma_{e \mu} + F_V \sigma_{x V} \right] \\
- \theta_V V dt \left[ F_x \sigma_{x V} + F_V \sigma_{V V} \right] + dq \left[ (\Gamma_\Lambda - 1) \left( e^{-F_e \bar{\eta}_e - F_x \bar{\eta}_x + F_\mu \bar{\eta}_\mu - F_V \bar{\eta}_V} - 1 \right) \right]. \quad (120)
\]

Note that

\[
E_t \left[ e^{-F_e \bar{\eta}_e - F_x \bar{\eta}_x + F_\mu \bar{\eta}_\mu - F_V \bar{\eta}_V} \right] = \left( \frac{1}{\bar{\eta}_V} \right) \left[ \frac{1}{\bar{\eta}_V} + F_e \alpha_e + F_x \alpha_x + F_\mu \alpha_\mu + F_V \right]^{-1}
\]

\[
E_t \left[ \Gamma_\Lambda e^{-F_e \bar{\eta}_e - F_x \bar{\eta}_x + F_\mu \bar{\eta}_\mu - F_V \bar{\eta}_V} \right] = \left( \frac{w_e}{\bar{\eta}_V} \right) \left[ \frac{1}{\bar{\eta}_V} + (F_e - \gamma_L) \alpha_e + F_x \alpha_x + F_\mu (\rho - \gamma_L) \alpha_\mu + F_V (\rho - \gamma_L) \alpha_V \right]^{-1}
\]

\[
+ \left( \frac{w_\mu}{\bar{\eta}_V} \right) \left[ \frac{1}{\bar{\eta}_V} + (F_e - \gamma_H) \alpha_e + F_x \alpha_x + F_\mu (\rho - \gamma_H) \alpha_\mu + F_V (\rho - \gamma_H) \alpha_V \right]^{-1}.
\]

Collecting terms linear in \( x_t, \ c_t, \ \mu_t, \ V_t, \) and terms independent of state vector, we find
equations as follows,

\[ F_s(\tau)' = -\kappa_s F_s \]  \hspace{1cm} (122)

\[ F_e(\tau)' = \kappa_s F_s \]  \hspace{1cm} (123)

\[ F_\mu(\tau)' = -\rho + F_c - \kappa_\mu F_\mu + \phi_\mu F_s \]  \hspace{1cm} (124)

\[ F_V(\tau)' = -r_V + \left( \frac{1}{2} - \alpha_s \lambda_V - \alpha_s \bar{\eta}_V \lambda_V \right) F_e - \phi_v F_s - \alpha_s \bar{\eta}_V \lambda_V F_z + \alpha_\mu \bar{\eta}_V \lambda_V F_\mu \]

\[- (\kappa_V + \bar{\eta}_V \lambda_V) F_V - \frac{1}{2} F_c^2 - \frac{1}{2} \left( \sigma_{x e}^2 + \sigma_{x \mu}^2 + \sigma_{s V}^2 + \sigma_{s}^2 \right) F_x^2 - \frac{1}{2} \left( \sigma_{\mu c}^2 + \sigma_{\mu}^2 \right) F_\mu^2 \]

\[- \frac{1}{2} \left( \sigma_{V c}^2 + \sigma_{V \mu}^2 + \sigma_{V}^2 \right) F_V^2 - \sigma_{x c} F_x F_z + \sigma_{\mu} F_\mu - \sigma_{V c} F_c F_v + (\sigma_{x \mu} \sigma_{c} + \sigma_{x} \sigma_{\mu}) F_s F_\mu \]

\[- (\sigma_{x c} \sigma_{V c} + \sigma_{x \mu} \sigma_{V \mu} + \sigma_{x} \sigma_{V}) F_z F_V + (\sigma_{\mu} \sigma_{V c} + \sigma_{\mu} \sigma_{V \mu}) F_\mu F_V \]

\[ + \theta_v \left( F_c + \sigma_{x c} F_x - \sigma_{\mu} F_\mu + \sigma_{V c} F_v \right) + \theta_\mu \left( \sigma_{x \mu} F_x - \sigma_{\mu} F_\mu + \sigma_{V \mu} F_v \right) \]

\[ + \theta_V \left( \sigma_{x V} F_z + F_v \sigma_v \right) - \lambda_V E_t \left[ \Gamma_{\lambda} e^{-F_c \tilde{\eta}_c - F_z \tilde{\eta}_z + F_\mu \tilde{\eta}_\mu - F_V \tilde{\eta}_V} \right] + \lambda_V E_t \left[ \Gamma_{\lambda} \right] \]  \hspace{1cm} (125)

\[ F_0(\tau)' = -r_0 + (\alpha_\lambda \lambda_0 + \alpha_s \bar{\eta}_V \lambda_0) F_c + (\kappa_s \phi_s - \bar{\mu} \phi_\mu - \bar{\nu} \phi_v + \alpha_s \bar{\eta}_V \lambda_0) F_x + (\kappa_\mu \bar{\mu} - \alpha_\mu \bar{\eta}_V \lambda_0) F_\mu \]

\[- (\kappa_V \bar{V} - \bar{\eta}_V \lambda_0) F_V + \lambda_0 E_t \left[ \Gamma_{\lambda} e^{-F_c \tilde{\eta}_c - F_z \tilde{\eta}_z + F_\mu \tilde{\eta}_\mu - F_V \tilde{\eta}_V} \right] - \lambda_0 E_t \left[ \Gamma_{\lambda} \right] \]  \hspace{1cm} (126)

subject to the “initial” conditions \( F_s(\tau = 0) = 1 \), and all other initial conditions \( F(\tau = 0) = 0 \).

We can solve analytically for \( F_s(\tau) \), \( F_e(\tau) \), \( F_\mu(\tau) \):

\[ F_s(\tau) = e^{-\kappa_s \tau}. \]  \hspace{1cm} (127)

\[ F_e(\tau) = 1 - e^{-\kappa_s \tau}. \]

\[ F_\mu(\tau) = \left( \frac{1 - \rho}{\kappa_\mu} \right) (1 - e^{-\kappa_\mu \tau}) + \left( \frac{\phi_\mu - 1}{\kappa_\mu - \kappa_\gamma} \right) (e^{-\kappa_\gamma \tau} - e^{-\kappa_\mu \tau}). \]  \hspace{1cm} (128)

\( F_0(\tau) \) and \( F_V(\tau) \) must be solved numerically.
A.8 Risk Free Bond Prices

The date-$t$ price $B^T(t, \mu_t, V_t)$ of a risk free bond that pays one unit of consumption at date-$T$ satisfies the equation

$$\Lambda_t B^T(t, \mu_t, V_t) = E_t [\Lambda_T 1].$$

(129)

This implies that $\Lambda_t B^T(t, \mu_t, V_t)$ is a P-martingale, implying that its expected change is zero:

$$0 = \frac{1}{dt} E_t \left[ \frac{d\Lambda}{\Lambda} + \frac{dB}{B} + \left( \frac{d\Lambda}{\Lambda} \right) \left( \frac{dB}{B} \right) \right].$$

(130)

Given that state vector dynamics are affine, the solution takes the form:

$$B^T(t, \mu_t, V_t) = e^{G_0(T-t)-G_u(T-t)\mu_t+G_V(T-t)V_t}.$$  

(131)

Now, from Ito’s lemma, we have

$$\frac{dB}{B} = \mu_B dt + \sqrt{V_t} dz_e (G_\mu \sigma_e - G_V \sigma_V) + \sqrt{V_t} dz_\mu (G_\mu \sigma_\mu - G_V \sigma_V) + \sqrt{V_t} dz_V (-G_V \sigma_V)$$

$$+ dq \left[ e^{-G_\mu \tilde{\eta}_u + G_V \tilde{\eta}_v} - 1 \right] - \lambda_t dt E_t \left[ e^{-G_\mu \tilde{\eta}_u + G_V \tilde{\eta}_v} - 1 \right],$$

(132)

where we have defined

$$\mu_B \equiv -G'_0 + \mu G'_\mu - VG'_V - G_\mu \left[ \kappa_\mu (\overline{\mu} - \mu) - \alpha_\mu \overline{\nu} \lambda_1 \right] + G_V \left[ \kappa_V (\overline{V} - V) - \overline{\nu} \lambda_1 \right]$$

$$+ \frac{1}{2} G^2 \sigma_e^2 + \frac{1}{2} G^2 \sigma_V^2 + G_\mu G_V (\sigma_\mu \sigma_e + \sigma_\mu \sigma_V)$$

$$+ \lambda E_t \left[ e^{-G_\mu \tilde{\eta}_u + G_V \tilde{\eta}_v} - 1 \right].$$

(133)
This implies that:

\[
\frac{1}{dt} E_i \left[ \frac{dB \, d\lambda}{B \, \lambda} \right] = -\theta_v V \left[ G_\mu \sigma_{\mu c} - G_v \sigma_{v c} \right] - \theta_v V \left[ G_\mu \sigma_{\mu} - G_v \sigma_{v \mu} \right] + \theta_v V \left[ G_v \sigma_v \right] + \lambda E_i \left[ (\Gamma_\lambda - 1) \left( e^{-G_\mu \tilde{\eta}_\mu + G_v \tilde{\eta}_v} - 1 \right) \right].
\]

(134)

Note that

\[
E_i \left[ e^{-G_\mu \tilde{\eta}_\mu + G_v \tilde{\eta}_v} \right] = \left( \frac{1}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + G_\mu \alpha_\mu - G_v \right]^{-1}
\]

\[
E_i \left[ \Gamma_\lambda e^{-G_\mu \tilde{\eta}_\mu + G_v \tilde{\eta}_v} \right] = \left( \frac{\omega_L}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + G_\mu \alpha_\mu - G_v - \alpha_v \gamma_L - (\rho - \gamma_L)u_\mu \alpha_\mu + (\rho - \gamma_L)u_v \right]^{-1}
\]

\[
\left( \frac{\omega_H}{\tilde{\eta}_v} \right) \left[ \frac{1}{\tilde{\eta}_v} + G_\mu \alpha_\mu - G_v - \alpha_v \gamma_H - (\rho - \gamma_H)u_\mu \alpha_\mu + (\rho - \gamma_H)u_v \right]^{-1}.
\]

(135)

Collecting terms linear in \( \mu_v \), \( V_v \), and terms independent of state vector, we find equations as follows,

\[
G'_\mu(\tau) = \rho - \kappa_\mu G_\mu
\]

\[
G'_v(\tau) = r_v + \alpha_v \tilde{\eta}_v \lambda_v G_\mu - (\kappa_v + \tilde{\eta}_v \lambda_v) G_v + \frac{1}{2} \left( \sigma^2_{\mu c} + \sigma^2_\mu \right) G_\mu^2 + \frac{1}{2} \left( \sigma^2_{v c} + \sigma^2_v \right) G_v^2 - \left( \sigma_\mu \sigma_{v c} + \sigma_\mu \sigma_{v \mu} \right) G_\mu G_v - \theta_v \left( \sigma_\mu G_\mu - \sigma_v G_v \right) - \theta_v \left( \sigma_\mu G_\mu - G_v \sigma_v \right)
\]

\[
+ \theta_v G_v \sigma_v + \lambda E_i \left[ (\Gamma_\lambda) \left( e^{-G_\mu \tilde{\eta}_\mu + G_v \tilde{\eta}_v} - 1 \right) \right]
\]

\[
G'_0(\tau) = -r_0 - (\kappa_\mu \tilde{\mu} - \alpha_\mu \tilde{\eta}_v \lambda_0) G_\mu + (\kappa_v \tilde{V} - \tilde{\eta}_v \lambda_0) G_v + \lambda_0 E_i \left[ (\Gamma_\lambda) \left( e^{-G_\mu \tilde{\eta}_\mu + G_v \tilde{\eta}_v} - 1 \right) \right]
\]

subject to the “initial” conditions \( G_\mu(\tau = 0) = 0 \), \( G_0(\tau = 0) = 0 \), \( G_v(\tau = 0) = 0 \).

We can solve analytically for \( G_\mu(\tau) \):

\[
G_\mu(\tau) = \left( \frac{\rho}{\kappa_\mu} \right) \left[ 1 - e^{-\kappa_\mu \tau} \right].
\]

(137)
$G_0(\tau)$ and $G_v(\tau)$ must be solved numerically.

We define the yield to maturity for a time horizon $\tau$ into the future as

$$B^T(t, \mu_t, V_t) = e^{-\tau y(\tau, \mu_t, V_t)}.$$  \hspace{1cm} (138)

It therefore follows that

$$y(\tau, \mu_t, V_t) = y_0 + y_\mu \mu_t - y_v V_t,$$  \hspace{1cm} (139)

where

$$y_0 = \frac{-G_0}{\tau},$$

$$y_\mu = \frac{G_\mu}{\tau},$$

$$y_v = \frac{G_v}{\tau}. \hspace{1cm} (140)$$

The instantaneous volatility of long term yield can be identified by noting that

$$dy_t|_{\text{stochastic}} = \sqrt{V}dz_v (y_v \sigma_{v_v} - y_\mu \sigma_{v_\mu}) + \sqrt{V}dz_\mu (y_v \sigma_{v_\mu} - y_\mu \sigma_\mu) + \sqrt{V}dz_v (y_v \sigma_v) + dq (y_\mu \tilde{n}_\mu - y_v \tilde{n}_v),$$

and hence

$$\frac{1}{dt}E_t[dy^2] = V_t \left[ (y_v \sigma_{v_v} - y_\mu \sigma_{v_\mu})^2 + (y_v \sigma_{v_\mu} - y_\mu \sigma_\mu)^2 + (y_v \sigma_v)^2 \right] + 2\lambda t \tilde{\eta}_v^2 (y_\mu \alpha_\mu - y_v)^2. \hspace{1cm} (141)$$