Online Supplement to

Dissecting the Equity Premium

Tyler Beason David Schreindorfer

Arizona State University

November 30, 2021

Abstract

This appendix provides details about our dataset and estimation methodology, and presents robustness tests. We also discuss connections between $EP(x)$ and the pricing kernel, and between $EP(x)$ and the term structure of equity risk premia. Lastly, we detail the approach for computing $EP(x)$ in equilibrium models.

Contents

A Data description and estimation details

This section shows summary statistics for the options data and presents goodness of fit statistics for the interpolant used to estimate risk-neutral densities.

The 1990-2019 CBOE dataset contains 18.4 million end-of-day quotes. Our estimation of risk-neutral densities is based on 4.7 million quotes that remain after excluding in-the-money and long-dated options, and applying the filters described in the paper. To illustrate how quote availability evolved over time, we plot the median number of daily quotes for each calendar year in the top panel of Figure 1. The number of available contracts (after filters) increased steadily from 44 per day in 1990 to 3,297 per day in 2019.

Precise estimates of risk-neutral densities rely crucially on quotes that span a wide strike region. The middle panel of Figure 1 reports the minimum and maximum observed strike K by calendar year, expressed in standardized moneyness units as $\frac{\log(K/F_t)}{VIX_t/100\times\sqrt{\tau}}$. To focus on typical values, we record the minimum and maximum observed strike for each trading day and then plot the median values across days for each year. The observed strike range is represented by the shaded region. In the early part of the sample, observed quotes typically extended from 2.5 standard deviations below the money to 1 standard deviation above the money. This moneyness range extended considerably towards the end of the sample, where it typically spanned [-8,2] standard deviations. To put these numbers into perspective, we also show the 1st and 99th percentiles of the estimated risk-neutral 30-day return distribution in the same standardized moneyness units. The plot shows that, while observed quotes typically need to be extrapolated beyond the observed strike range in the early sample, available quotes span the vast majority of the density's support from 1998 onwards.

As explained in the paper, we estimate risk-neutral densities by mapping price quotes to implied volatility units, fitting an interpolant to them, mapping interpolated implied volatilities back to prices, and finally differentiating option prices via

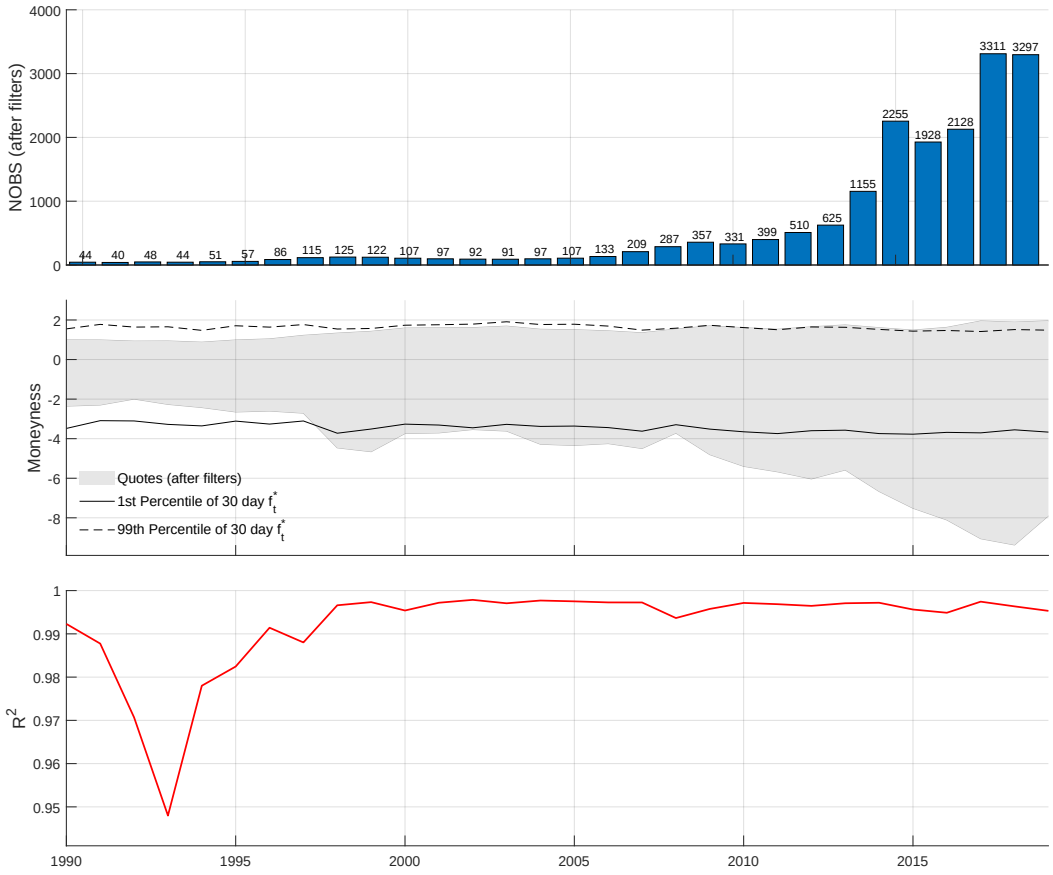

Figure 1: Summary statistics

finite differences. For an accurate representation of the options data, it is important that the interpolant provides a good fit to observed implied volatilities. We assess the fit via the R^2 it implies. The bottom panel of Figure 1 shows the median R^2 across trading days in each calendar year. The SVI interpolant fits implied volatilities very well, with typical R^2 above 94% for all years and well above 99% from 1998 onwards. The mean (median) R^2 across all days equals 98.8% (99.6%).

To illustrate the interpolant's fit more directly, we plot it for 3 select days in Figure 2. The left column shows the first day of the sample (1990/01/02), the middle column the day of Lehman Brothers' default (2008/09/15), and the right column the last day of the sample (2019/12/31). In terms of fit and quote availability, these

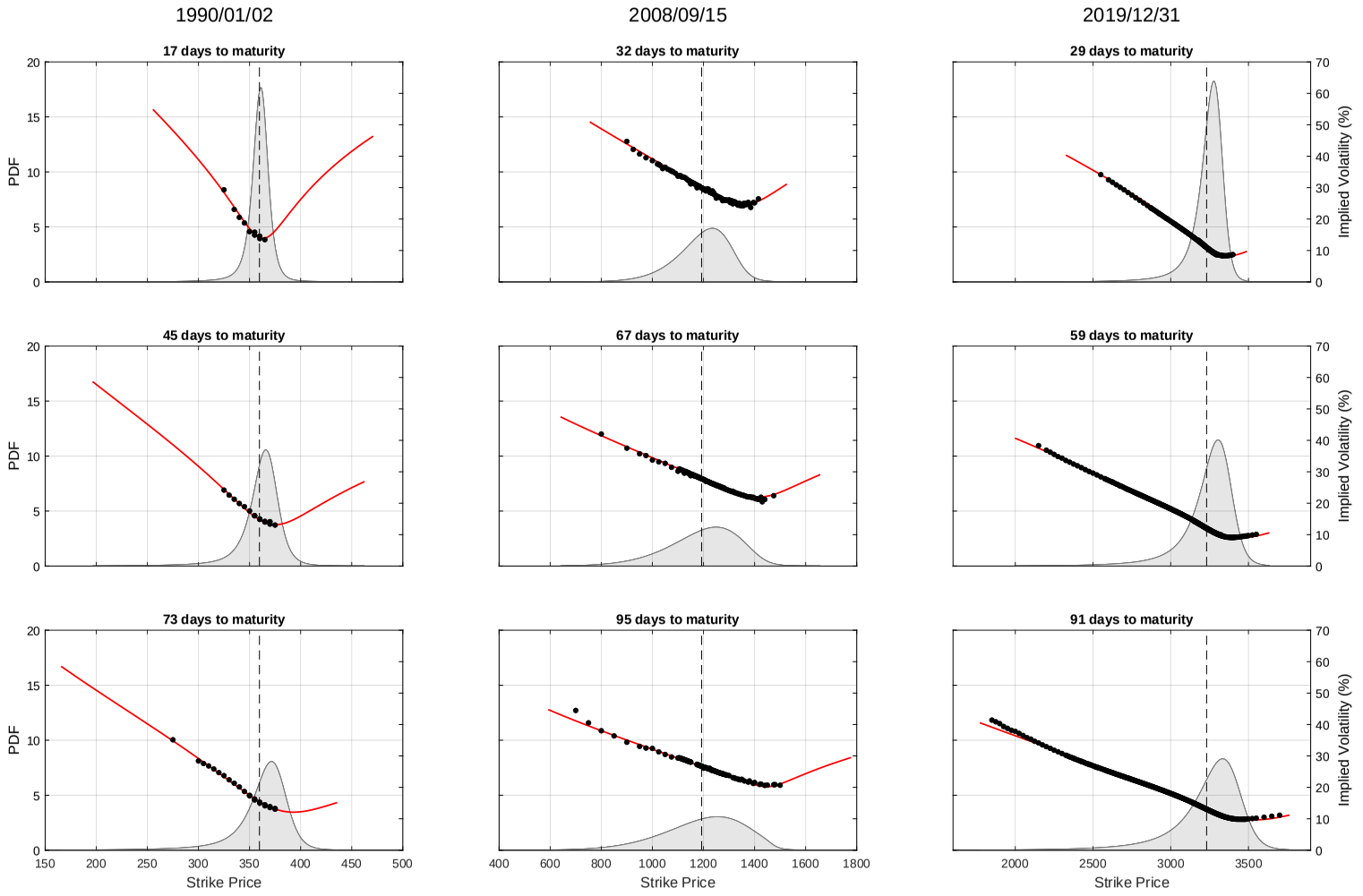

Figure 2: SVI estimation for select days

days are fairly representative of typical days in the early, middle, and late part of the sample, respectively. Recall from the paper that our SVI interpolant fits implied volatilities across all strikes and maturities simultaneously. For illustration purposes, Figure 2 focuses on the observed maturities closest to 1, 2, and 3 months for each day. In line with Figure 1, we see that (i) the interpolant provides an excellent fit to observed option prices for all 3 days and all 3 maturities and (ii) extrapolation only plays an important role for the early part of the sample. Importantly, the figure illustrates that the SVI method provides a very sensible way to extrapolate implied volatilities, especially in the left tail that matters most for our main result.[1]

To ensure that our findings are not unduely affected by the pre-1998 sample, which coincides with fewer observed strikes, a somewhat lower R^2, and the need to extrapolate more frequently, we next illustrate them for different subsamples.

B Robustness check 1: Subsamples

First, we compute $EP(x)$ based on post-1997 data to investigate whether our results are robust to the exclusion of lower quality data from the early sample. The dotted line in Figure 3 shows that the post-1997 $EP(x)$ curve attributes the entire equity premium to returns below -10%, and over 8/10 to the interval $[-30\%, -10\%]$. Our main finding – the large importance of intermediate left tail events for the equity premium – is therefore even stronger post-1997.

A second potential concern about our sample is that it is not representative of the population distribution of economic states. For example, recessions may be under- or over-represented relative to their ergodic frequency. To the extent that the importance of left tail states changes in recessions, the $EP(x)$ curve in our sample could be a biased estimate of the corresponding population curve. A similar concern was recently raised Bansal et al. (2020) about estimates of the term structure of

[1] To enable readers to retrace our calculations, we include data for these 3 days in our replication code package.

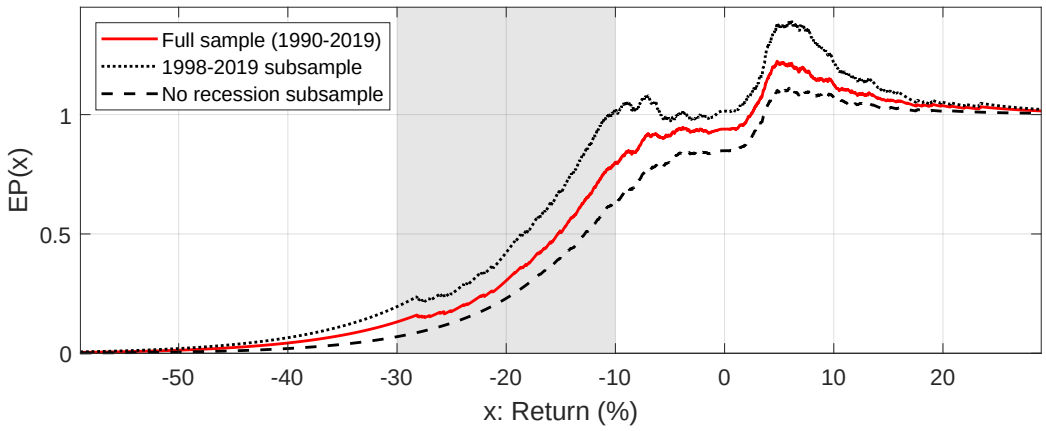

Figure 3: Equity premium decomposition for different subsamples.

the equity premium, which, like our estimates, are most commonly derived from derivative data with short historical samples. The dashed line in Figure 3 shows $EP(x)$ for the part of our sample that excludes the NBER recessions priods 1990/07-1991/03, 2001/03-2001/11, and 2007/12-2009/06. In this subsample, returns below -10% account for 63/100 of the equity premium, whereas returns in the interval $[-30\%, -10\%]$ account for 56/100. As intuition suggests, tail states therefore indeed play a more prominent role for risk premia during recessions. The difference is slight, however, and intermediate left tail events continue to account for over half of the equity premium even outside of recessions. We return to the model of Bansal et al. (2020) in section G of this supplement, where we evaluate its implications for $EP(x)$.

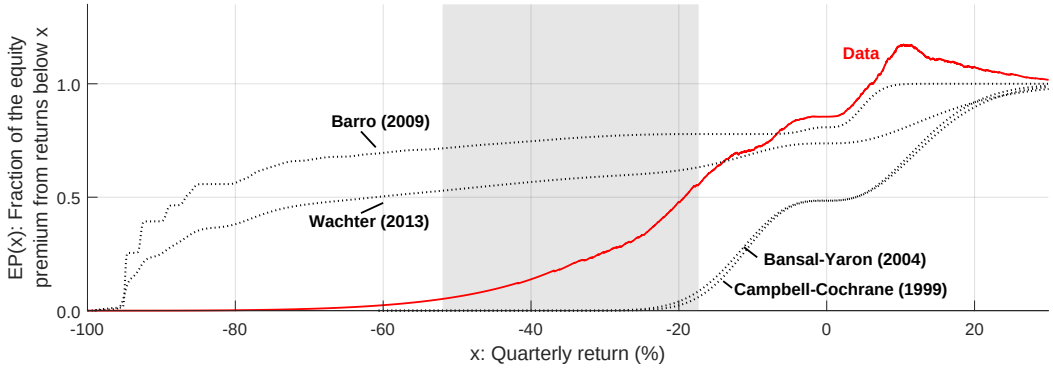

Figure 4: $EP(x)$ **for quarterly returns**

C Robustness check 2: Longer return horizons

The paper focuses on a return horizon of one month. To understand the economic mechanism of existing models, and to relate our results to literature on the term structure of asset prices, it is similarly useful to decompose the equity premium for longer horizons. The fact that option prices and returns can only be observed jointly for 30 years, however, makes it difficult to compute the necessary estimate of the physical return distribution with reasonable precision at very long horizons.

With this caveat in mind, we examine the importance of different return states at horizons of a few months. Figure 4 replicates our main analysis at the quarterly (90 calendar day) horizon. To account for the larger volatility of quarterly returns, we scale the left tail region of interest by a factor of $\sqrt{3}$. In the data, quarterly returns in the interval $\sqrt{3} \times [-30\%, -10\%]$ account for 50.4/100 of the equity premium, while they account for 10.3/100 or less in the four depicted models. At the quarterly horizon, the models therefore perform similarly poorly in capturing sources of the equity premium. The habit and long-run risks models continue to attribute risk premia to less extreme and the disaster model to more extreme left tail states than the data. Of course, it is possible (but hard to evaluate) that the models align more closely with the data at very long horizons. Because long horizons make it challenging to distinguish disasters (jumps) from large Gaussian shocks, however,

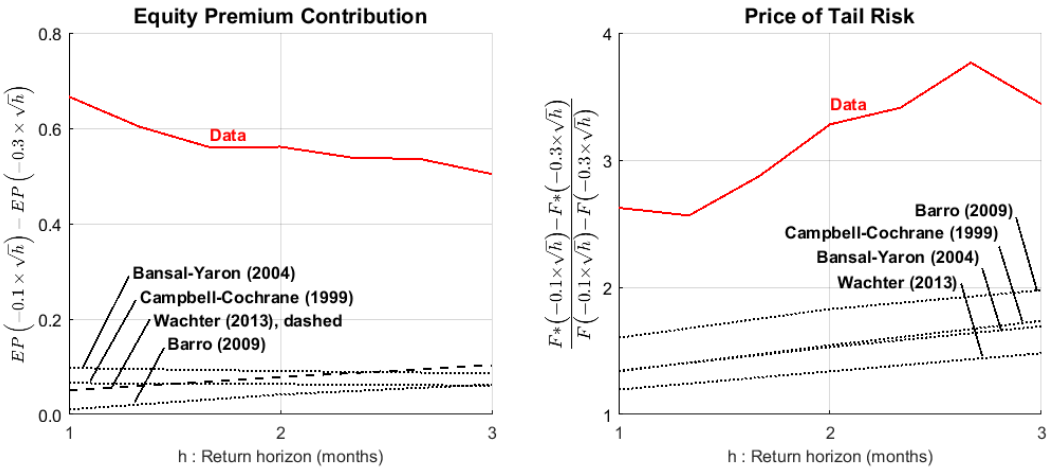

Figure 5: The importance of left tail events at different return horizons

they are less informative about economic mechanisms.

Figure 5 provides a closer look at the term structure dimension by plotting the equity premium contribution (left panel) and price of risk (right panel) for returns in $\sqrt{h} \times [-30\%, -10\%]$ as a function of the return horizon, h. The left panel shows that, in the data, left tail events contribute slightly less to the equity premium at longer horizons. In contrast, they contribute equally much (habits and long-run risks) or more (disasters) in the models. The figure suggests that the gap between models and data remains significant for horizons well beyond three months.

The right panel shows that the ratio of average risk-neutral to physical probabilities – our measure of risk prices – increases at longer horizons in the data. While the models capture this increasing pattern, they continue to substantially undershoot the magnitude of risk prices at longer horizons. The horizon over which tail events are measured therefore does not appear to be essential for our finding.

D Robustness check 3: International evidence

We repeat the empirical analysis for the British FTSE 100 and German DAX to investigate whether our findings continue to hold internationally. Options data for both indices was purchased from Refinitiv, a subsidiary of the London Stock Exchange Group. Refinitiv's "Tick History Data" spans 2003-2019 and contains intra-daily option quotes and trades, from which we extract end-of-day prices to mimic the analysis for the S&P 500. As for the S&P 500, we rely on end-of-day midquotes for the DAX. Because quotes are very sparse for the first few years of the FTSE sample, however, we instead rely on closing prices. The end-of-day FTSE data consists of 5.1 million (0.72 million) observations before (after) filters, while the DAX data contains 4.1 million (0.91 million) observations. For the FTSE, the median number of daily observations after filters increased from 152 in 2003 to 265 in 2019, while it increased from 83 in 2003 to 280 in 2019 for the DAX.

As for the S&P 500, we use LIBOR rates for the corresponding market to map option prices back and forth between British Pound (or Euro) and IV units, and we interpolate implied volatilities across the moneyness-maturity spectrum based on the SVI method. Moneyness is measured in standard deviation units as $\kappa \equiv \frac{\log(K/F_{t,\tau})}{V_t/100\times\sqrt{\tau}}$, where V_t is the VIX-equivalent for the British (VFTSE index) and German (VDAX index) market, respectively. The observed moneyness range for both indices is relatively constant across years and extends from about -7 to +2 standard deviations for the FTSE and from -7 to +2.5 standard deviations for the DAX. Similar to what Figure 1 shows for the S&P 500 over 2003-2019, these ranges cover the vast majority of the estimated risk-neutral return distributions' support in each calendar year. The mean (median) R^2 of the SVI fit to implied volatilities equals 0.987 (0.992) across all days in the sample for the FTSE, and 0.967 (0.986) for the DAX.

Figure 6 shows $EP(x)$ curves for both markets and compares them to the U.S. curve for the equivalent sample period. Returns in $[-30\%, -10\%]$ contribute somewhat less to the equity premium over 2003-2019 for the S&P 500 and DAX, likely

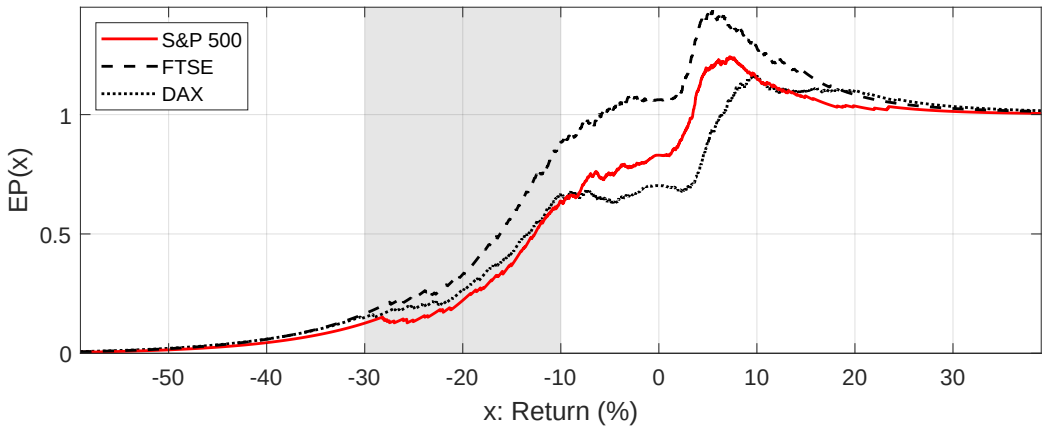

Figure 6: $EP(x)$ **over 2003-2019 in international and U.S. data**

because this subsample excludes one of the two large recessions in our sample. In contrast, the British market attributes 72/100 of the equity premium to such states. The likely reason for this difference is that fact that BREXIT created large uncertainty in the U.K. and, while not resulting in a recession, was associated with large insurance premia on protective put positions.

The international evidence shows that, in each of the markets we examine, the majority of the equity premium is attributable to shocks that coincide with market returns between -30% and -10%. In itself, this observation does not reveal the type of fundamental shocks that command large risk premia. The fact that the three countries have highly correlated stock markets[2], however, suggests that the equity premium reflects compensation for shocks that are common across countries.

[2]30-day returns on the British and U.S. indices have a correlation of 0.834 over our sample period, while the German and U.S. indices have a correlation of 0.814. Similarly, the VIX has a correlation of 0.952 with its British equivalent and a correlation of 0.871 with its German equivalent.

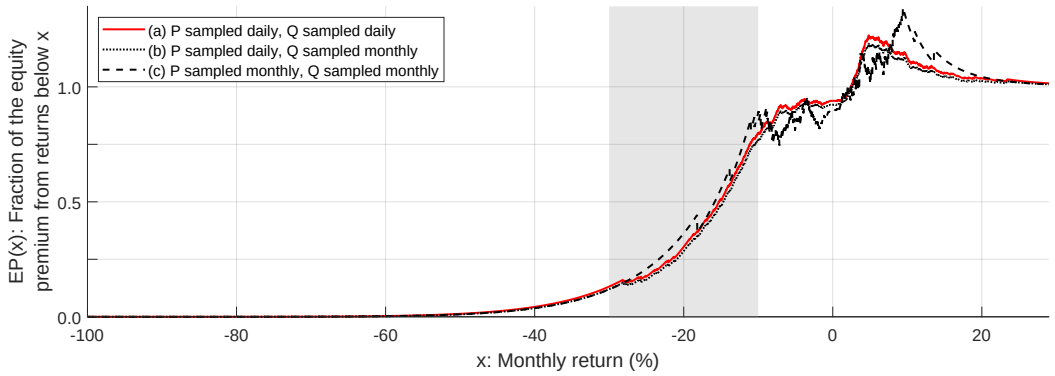

Figure 7: $EP(x)$ **estimates for alternative sampling schemes**

E Robustness check 4: Monthly sampling

Our main analysis is based on a daily sample of overlapping 30 calendar day returns and corresponding estimates of the conditional risk-neutral distribution, f_t^*. The main reason for sampling daily is that it allows us to obtain a more precise estimate of the unconditional return distribution f than would be the case with monthly sampling. A disadvantage of daily sampling is that it induces autocorrelation into our estimates of f_t^*, both due to the overlapping nature of the data and potentially due to stale option price quotes. As a result, statistical inference requires a block bootstrap. In the following, we illustrate that our main result based on daily sampling (sampling scheme "a") is robust to two alternative sampling schemes:

b f^* equals the average of end-of-month estimates of the conditional 30-day risk-neutral density f_t^*, and f is estimated as in the main text.

c f^* equals the average of end-of-month estimates of the conditional 30-day risk-neutral density f_t^*, and f equals the empirical distribution of 30-day returns, sampled at the end of every calendar month.

Figure 7 compares estimates of $EP(x)$ across the three sampling schemes. The estimate under sampling scheme "b" is nearly identical to the benchmark case, with returns in the interval $[-30\%, -10\%]$ contributing 65/100 to the equity premium

(compared to 67/100 in the benchmark). In contrast, sampling scheme "c" results in a much noisier estimate of $EP(x)$ because f is estimated based on just 360 monthly returns, as opposed to $> 7,000$ returns under sampling scheme "a". Sampling scheme "c" implies that $R \in [-30\%, -10\%]$ contribute 77/100 to the equity premium, i.e., ten percentage points more than the benchmark estimates based on daily sampling. The estimates we report in the paper should therefore be considered conservative.

F Relating $EP(x)$ to the pricing kernel

Asset pricing theories are commonly characterized in terms of a pricing kernel. Projected onto returns, the pricing kernel equals

$$\mathbb{E}_t[M_{t+1}|R_{t+1}] = \mathbb{E}_t[M_{t+1}]\frac{f_t^*(R_{t+1})}{f_t(R_{t+1})}, \tag{1}$$

i.e. it reflects the ratio of (conditional) risk-neutral to physical return probabilities. Unfortunately, $\mathbb{E}_t[M_{t+1}|R_{t+1}]$ is challenging to use for state-by-state comparisons between models and data because $f_t(R_{t+1})$ is difficult to measure empirically. An important insight of our paper is that state-by-state comparisons are possible based on unconditional probabilities, which are straightforward to estimate. In particular, we quantify risk prices based on the ratio of *average* risk-neutral ($f^*(R) = \mathbb{E}[f_t^*(R)]$) and physical ($f(R) = \mathbb{E}[f_t(R)]$) probabilities.

To make our metric easier to relate to existing work, we compare it to the average pricing kernel implied by different models. In IID models such as Barro (2009), conditional distributions do not vary and $f^*(R)/f(R)$ is exactly equal to $\mathbb{E}_t[M_{t+1}|R_{t+1}]/\mathbb{E}_t[M_{t+1}]$. In that case, our metric corresponds to the traditional metric for the price of risk. In non-IID models, $f^*(R)/f(R) = \mathbb{E}[f_t^*(R)]/\mathbb{E}[f_t(R)]$ generally differs from $\mathbb{E}[f_t^*(R)/f_t(R)]$ and hence from the average pricing kernel $\mathbb{E}\left[\mathbb{E}_t[M_{t+1}|R_{t+1}]/\mathbb{E}_t[M_{t+1}]\right]$. Figure 8 compares both metrics for the four asset pricing models in Figure 1 of the paper, and shows that they imply a very similar values. Our finding that $f^*(R)/f(R)$ is counterfactually low for left tail returns in the models therefore appears directly related to the shape of the (projected) pricing kernel they imply. Note that the model-based price of risk in figure 8 can be directly compared to the empirical $f^*(R)$-to-$f(R)$ probability ratio in figure II in the paper, which is considerably steeper in the left tail.

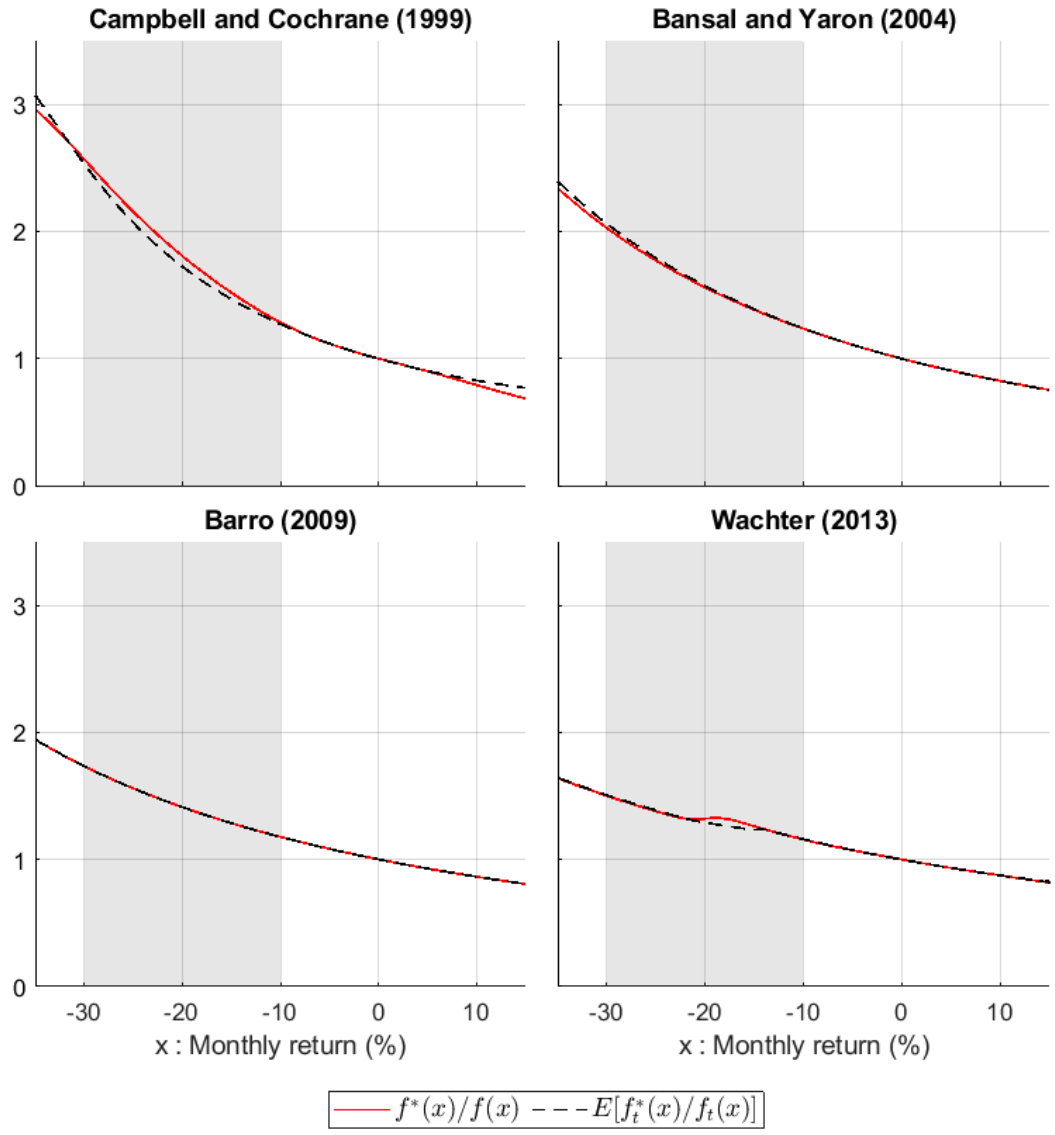

Figure 8: Alternative measures of the price of risk.

G Relating $EP(x)$ to the term structure of the equity premium

van Binsbergen et al. (2012) extract prices of dividend strips with different maturities from index options, and show that claims to near-term dividends have higher average returns and Sharpe ratios. In contrast, they find that habit and long-run risks models imply an upward-sloping term structure of dividend risk premia. Our analysis resembles theirs in that we use options data to characterize properties of the equity premium. It is therefore interesting to ask whether the two findings are related. We do so in two ways.

First, we implement $EP(x)$ for two modified long-run risks models that were explicitly designed to address the term structure finding. Belo et al. (2015) show that the downward-sloping term structure can be explained by a modified dividend process that respects stationary financial leverage ratios. Intuitively, capital structure rebalancing shifts risk from long- to short-horizon dividend strips by forcing shareholders to divest (invest) when leverage is low (high).[3] Bansal et al. (2020) argue that the downward-sloping term structure reflects an oversampling of recessions in van Binsbergen et al.'s sample. Their regime-switching long-run risks model reproduces a negative slope in small samples that overweight recessions as in the data, but implies a positive slope unconditionally. Following Bansal et al.'s argument, we compute $EP(x)$ based on a simulation of their model's state that reflects an (oversampled) recession frequency of 8%.[4] We find that the two modified long-run risks models closely resemble the model of Bansal and Yaron (2004) in their implications

[3] Belo et al. (2015) rely on Bansal et al.'s (2012) calibration of the long-run risks model. We follow their log-linear solution approach and compute ex-dividend returns as $R^{ex}_{t+1} = e^{z_{y,t+1}-z_{y,t}+\Delta y_{t+1}}\left(\frac{1-e^{l_{t+1}}}{1-e^{l_t}}\right)$, which follows from the price-dividend ratio, $\exp(z_{d,t}) = \exp(z_{y,t} + y_t - d_t)(1-\exp(l_t))$, in the model. Option prices are computed via Monte Carlo simulation.

[4] We follow the paper's log-linear solution approach and are able to reproduce the moments reported in section C.10. Option prices and the conditional return distribution f_t can be computed analytically based on standard results for normal random variables.

for $EP(x)$. Returns between -30% and -10% contribute 0.093 and 0.080 to the equity premium, respectively (versus 0.095 in the original model), they occur with a probability of 0.011 and 0.013 (versus 0.013 in the original model), and feature of price of risk of 1.237 and 1.174 (versus 1.280 in the original model).

Second, to understand the conceptual difference between our finding and that of van Binsbergen et al. (2012), it is helpful to write the market return as a weighted average of the returns on individual dividend strips, i.e.[5]

$$R_{t+1}^{cum} = \sum_{n=1}^{\infty} w_{t,n} R_{t+1,n}, \tag{2}$$

where $w_{t,n}$ is the price of a divided strip with maturity n divided by the index level, $R_{t+1,n}$ is the spot return on that strip, and $\sum_{n=1}^{\infty} w_{t,n} = 1$. van Binsbergen et al.'s finding implies that transient shocks (which only affect near-term dividends) command larger risk premia than persistent shocks (which affect dividends far into the future). Market returns reflect both transient and persistent shocks. In order for the market to drop by 10% or more, however, shocks have to be fairly persistent. To see this, note that the first 2 years of dividends contribute only about 3-4% to the total index value.[6] Even if the value of such dividends would fall to zero (a return of -100%), the market would only fall by 3-4%. The market returns we show to contribute the bulk of the equity premium, which fall between -30% and -10%, therefore likely reflect shocks that affect dividends over considerably longer horizons. Both theoretical results from existing models and conceptual considerations therefore suggest that $EP(x)$ reflects different economic forces than the term structure of risk premia.

[5] See Appendix A of van Binsbergen and Koijen (2017) for a derivation.

[6] See figure 2 in van Binsbergen et al. (2012).

H Bootstrap for $EP(x)$

We detail the block bootstrap used to compute the empirical sampling distribution for $EP(x)$ in Figure III of the main text. Our empirical sample consists of joint observations for $\{R_{t:t+30}, f_t^*(R_{t:t+30})\}$, $t = 1, ..., T$, where $f_t^*(R_{t:t+30})$ is the conditional risk-neutral PDF of a 30 calendar day return at time t, and $R_{t:t+30}$ is the realized return over the subsequent 30 calendar days. Using this sample, $EP(x)$ is computed as

$$EP(x) = \frac{\frac{1}{T}\sum_{t=1}^{T} R_{t:t+30}\mathbf{1}\{R_{t:t+30} \leq x\} - \int_{-1}^{x} R\left[\frac{1}{T}\sum_{t=1}^{T} f_t^*(R)\right] dR}{\frac{1}{T}\sum_{t=1}^{T} R_{t:t+30} - \int_{-1}^{\infty} R\left[\frac{1}{T}\sum_{t=1}^{T} f_t^*(R)\right] dR}.$$

To compute the joint sampling distribution of $EP(-30\%)$ and $EP(-10\%) - EP(-30\%)$, we use a block bootstrap with 10 million bootstrap samples and a block length of 21 trading days (the average number of trading days over periods of 30 calendar days). Specifically, the j^{th} pair $\{EP(-30\%), EP(-10\%)\}$ is created as follows.

1. Randomly draw an integer i between 1 and $T - 20$.
2. Add observations $i, ..., i + 20$ to the bootstrap sample.
3. Repeat steps 1 and 2 until the bootstrap sample contains at least T observations. In practice, there will be slightly more than T observations if $T/21$ does not equal an integer.
4. Use the bootstrap sample created in steps 1 through 3 to compute $EP(-30\%)$ and $EP(-10\%)$ based on the above formula for $EP(x)$.

The sampling distribution is created by repeating steps 1 through 4 ten million times.

I Model Solutions

This section discusses our solution approach for each model and points out which results from the original studies were replicated for validation. The calculation of $EP(x)$ relies on option prices, which, among the papers we consider, only Backus et al. (2011) and Schreindorfer (2020) solve for. We therefore derive our own solutions for the remaining models. Matlab programs are provided in our replication code package.

Regardless of the model, we can write the put-to-spot price ratio for a strike price K and a 1-period maturity as

$$\begin{aligned}\mathcal{P}(X,\xi_t) =& \frac{1}{S_t}\mathbb{E}_t\left[M(\Delta c_{t+1},\xi_t,\xi_{t+1})\max\{0,K-S_{t+1}\}\right]\\ =& \mathbb{E}_t\left[M(\Delta c_{t+1},\xi_t,\xi_{t+1})\left(X-\frac{PD(\xi_{t+1})}{PD(\xi_t)}e^{\Delta d_{t+1}}\right)\mathbf{1}_{\Delta d_{t+1}\leq\log\left(X\frac{PD(\xi_t)}{PD(\xi_{t+1})}\right)}\right],\end{aligned} \tag{3}$$

where $X = K/S_t$ equals moneyness, the pricing kernel M is a function of log consumption growth Δc_{t+1}, today's state ξ_t, and tomorrow's state ξ_{t+1}, and the price-dividend ratio PD is a function of the contemporaneous state. In what follows, we detail how the expectation on the RHS of (3) is evaluated in each model, and we defer details on the calculation of $EP(x)$ to Section J. Throughout, we rely on the same notation as the original studies and refer interested readers there for definitions and additional details.

I.1 Campbell and Cochrane (1999)

The external habits model does not admit analytical solutions. Wachter (2005) shows that the numerical solution in the original study is inaccurate and proposes the "series method" as a more precise alternative. We follow her solution approach. Specifically, the model is solved on a grid of 1001 unequally-spaced points for the model's state S_{t+1} (the surplus consumption ratio) – "Grid 3" in Wachter (2005). We employ Gauss-Chebyshev quadrature with 500 points covering ± 7 standard

deviations when computing expectations. Cubic splines are used to interpolate the price-dividend ratio to off-grid values resulting from the quadrature. Our results closely match Table 2 in Wachter (2005).

The Campbell-Cochrane model features a single shock to each consumption and dividend growth, but no separate shock to the state. We evaluate the expectation in (3) in two steps. First, we condition on Δc_{t+1} and evaluate the expectation over Δd_{t+1} based on standard results for truncated normal random variables, which yields[7]

$$\mathcal{P}(X,\xi_t) = \mathbb{E}_t\Bigg[M(\Delta c_{t+1},\xi_t)\Bigg(X\Phi\left(\nu(\Delta c_{t+1},\xi_t)\right) - e^{\mathbb{E}[\Delta d_{t+1}|\Delta c_{t+1}]+\sigma[\Delta d_{t+1}|\Delta c_{t+1}]^2/2}\frac{PD(\Delta c_{t+1},\xi_t)}{PD(\xi_t)}\Phi\left(\nu(\Delta c_{t+1},\xi_t)-\sigma[\Delta d_{t+1}|\Delta c_{t+1}]\right)\Bigg)\Bigg]$$

where

- The functional form for the pricing kernel M is shown in Equation 5 of Campbell and Cochrane (1999)
- $\nu(\Delta c_{t+1},\xi_t) = \frac{\log\left(X\frac{PD(\xi_t)}{PD(\Delta c_{t+1},\xi_t)}\right)-E[\Delta d_{t+1}|\Delta c_{t+1}]}{\sigma[\Delta d_{t+1}|\Delta c_{t+1}]}$
- $E[\Delta d_{t+1}|\Delta c_{t+1}] = g + \rho\sigma_w v_{t+1}$
- $\sigma[\Delta d_{t+1}|\Delta c_{t+1}] = \sqrt{1-\rho^2}\sigma_w$

and $\Phi(\cdot)$ denotes the CDF of a standard normal. The remaining expectation over Δc_{t+1} is evaluated based on the aforementioned quadrature method.

I.2 Bekaert and Engstrom (2017)

To solve the habit model of Bekaert and Engstrom (2017), we closely follow the replication code on the publisher's website,

`https://doi.org/10.1086/691450`

[7] The relevant results for truncated normals are summarized by Lemma 1 in Schreindorfer (2020).

The model features two shocks from a gamma distribution that affect both consumption and dividend growth, but no separate shock to the model's state. We use a grid of 101 equally-spaced points for n_t (the time-varying shape parameter of the first gamma shock) and 100 equally-spaced points for s_t (the log surplus consumption ratio), both for the model solution and to compute option prices. Expectation in (3) are evaluated via Monte Carlo integration, using 5×10^8 draws and an identical seed for each grid value of the model's bivariate state.

I.3 Bansal and Yaron (2004)

We solve the long-run risks model numerically rather than relying on the log-linear approximation in the original paper. The model's bivariate state $\xi_t = [x_t, \sigma_t^2]$ is discretized with grids of size $N_x = 200$ and $N_\sigma = 100$, respectively. The grids are linearly-spaced, centered around the unconditional mean of each process, and extend 5 standard deviations in each direction. For this purpose, we set the standard deviation of the conditional mean process to the one implied by the largest grid value for volatility. We use Gauss-Chebyshev quadrature with 50 nodes covering ± 7 standard deviations for shocks to the conditional mean and variance processes when computing expectations. We solve for the value-consumption and price-dividend ratios by iterating on the the Euler equations for both ratios until convergence, using bivariate cubic splines to interpolate to off-grid values when necessary. Our solution replicates Table 2 in Beeler and Campbell (2012), who provide an examination of long-run risk models and show more moments that the original paper.

The model features four IID normal shocks, one to consumption growth, one to dividend growth, and one to each of the model's two states. We apply standard results for truncated normal random variables to integrate over the shocks to consumption and dividend growth, which yields

$$\mathcal{P}(X, \xi_t) = \mathbb{E}_t \left[M_1(\xi_t, \xi_{t+1}) X \Phi\left(\nu(\xi_t, \xi_{t+1})\right) - M_2(\xi_t, \xi_{t+1}) \frac{PD(\xi_{t+1})}{PD(\xi_t)} \Phi\left(\nu(\xi_t, \xi_{t+1}) - \varphi_d \sigma_t\right) \right]$$

where, based on the Hansen et al. (2008) formulation of the Epstein and Zin (1989)

pricing kernel with VC being the utility-consumption ratio and $\mathcal{R}$ its certainty equivalent,

- $\nu(\xi_t, \xi_{t+1}) = \frac{\log\left(X \frac{PD(\xi_t)}{PD(\xi_{t+1})}\right) - \mu - \phi x_t}{\varphi_d \sigma_t}$
- $M_1(\xi_t, \xi_{t+1}) = \beta \left(\frac{VC(\xi_{t+1})}{\mathcal{R}(\xi_t)}\right)^{\frac{1}{\psi}-\gamma} e^{-\gamma\mu - \gamma x_t + \gamma^2\sigma^2/2}$
- $M_2(\xi_t, \xi_{t+1}) = \beta \left(\frac{VC(\xi_{t+1})}{\mathcal{R}(\xi_t)}\right)^{\frac{1}{\psi}-\gamma} e^{(1-\gamma)\mu + (\phi-\gamma)x_t + (\gamma^2+\varphi_d^2)\sigma^2/2}$

and $\Phi(\cdot)$ denotes the CDF of a standard normal. The remaining expectation over shocks to the model's future state is evaluated based on the aforementioned quadrature method.

I.4 Drechsler and Yaron (2011)

Drechsler and Yaron (2011) solve their model analytically based on the Campbell-Shiller log-linearization. While Pohl et al. (2018) do not explicitly solve the Drechsler-Yaron model, they show that log-linearization induces large numerical errors in many similar long-run risks models. Because we would like to ensure that our results are economically meaningful, we rely on a global, numerical, non-linear solution approach instead.

The model's state is 3-dimensional. Both the utility-consumption ratio V/C and price-dividend ratio P/D are solved on a grid with 20 equally-spaced points for x_t between -0.012 and 0.01, 20 equally-spaced points for $\bar{\sigma}_t$ between 1e-16 and 4, and 30 equally-spaced points for σ_t between 1e-16 and 35. The endpoints are chosen to ensure that the grids contain the vast majority of the models' state. The model features 9 IID shocks: 5 normal shocks, two Poisson shocks, and 2 exponential shocks. To evaluate expectations, we analytically integrate over Gaussian shocks to consumption and dividend growth and Poisson shocks to the model's state, rely on Gaussian quadrature for normal shocks to the model's state, and on Gauss-Laguerre quadrature for exponential shocks to the model's state. The quadrature routine uses 5 nodes in each dimension. Cubic splines are used to interpolate V/C and P/D to

Table 1: Drechsler and Yaron (2011) – Asset prices for two solution approaches

	Log-linear			Global		
	5	50	95	5	50	95
$\mathbb{E}[r_m]$	3.94	6.92	10.08	3.43	6.42	9.57
$\sigma[r_m]$	14.50	17.44	22.09	14.02	16.83	21.00
$skew[r_m]$	-0.79	0.03	0.89	-0.76	0.03	0.87
$skew[r_m]$(M)	-1.26	-0.42	0.08	-1.06	-0.34	0.12
$AC1[r_m]$	-0.26	-0.06	0.14	-0.09	-0.01	0.06
$\mathbb{E}[r_f]$	0.23	0.94	1.43	0.28	0.98	1.48
$\sigma[r_f]$	0.83	1.49	2.71	0.83	1.47	2.65
$skew[r_f]$	-4.43	-2.30	-0.93	-4.38	-2.26	-0.91
$AC1[r_f]$	0.23	0.47	0.68	0.23	0.47	0.68
$\mathbb{E}[p-d]$	2.92	3.01	3.08	3.02	3.11	3.18
$\sigma[p-d]$	0.13	0.17	0.22	0.12	0.16	0.22
$skew[p-d]$	0.34	0.56	0.73	0.31	0.55	0.73
$AC1[p-d]$	-1.23	-0.30	0.42	-1.48	-0.46	0.24

This table shows percentiles 5, 50, and 95 across 10,000 finite sample simulations of length 73 years for each indicated statistic. Results are shown both for the log-linearized approximate solution from the original paper and for a numerical, global, non-linear solution. All moments are time-aggregated to an annual frequency, expect when labeled M (monthly).

any resulting off-grid values that result from the quadrature. In the model, the probability of jumps is a function of σ_t. Our algorithm truncates the number of jumps at the largest (state-specific) value whose probability exceeds 1 basis point when evaluating expectations. As such, no jumps are allowed at the lowest grid value of σ_t (where the jump intensity is very low), but up to 11 jumps are allowed at the highest grid value (where the jump intensity is very high). This aspect is a crucial feature of the numerical solution: results change significantly when jumps are truncated too aggressively. Both V/C and P/D are initialized at the log-linear solution from the original paper, and we iterate on the Euler equations of both functions until their absolute change falls below 10^{-6} for all grid values.

To compare results from our non-linear solution to those from the log-linear approximation, we simulate 10,000 samples of 77 years (the length of Drechsler

and Yaron's sample) and compute finite sample moments of time-aggregated data. To ensure comparability, we use the same seed for the simulation of both solution approaches. As Table 1 shows, results for basic asset prices are fairly similar across both solution approaches. Log-linearization therefore results in much smaller errors than in the long-run risks models considered by Pohl et al. (2018). We nevertheless base all of our results for the model on the global solution approach.

To compute option prices, we evaluate the expectation in (3) via Monte Carlo integration, using 5×10^8 draws and an identical seed for each grid value of the model's trivariate state.

I.5 Barro (2009)

The rare disaster model is cast in discrete time, but Barro (2009) solves it based on continuous time approximations. Because such approximations are not available for option prices, we instead rely on exact, analytical, discrete time solutions for the price-consumption ratio and risk-free rate. We find, however, that these solutions differ only marginally from the values implied by the approximations in Equations 5, 7, and 12 of the original paper.

Barro (2009) provides only a back-of-the-envelope calculation for the (levered) equity premium, i.e. he does not formally define a dividend process. Because we require the price of an equity claim in order to price equity index options, we follow Abel (1999) in defining dividends as $D_t = C_t^{\phi}$, where ϕ captures leverage. We set leverage to a value of 2.6, the same number used by Wachter (2013). All of our results for the Barro model are based on this equity claim.

The model relies on a mixture of Gaussian shocks and disasters, where disaster sizes are drawn from the empirical distribution of international disaster occurrences. We follow Barro (2009) in assuming that the probability of multiple disasters in a short time period (in our case, a month) is negligible. To compute option prices, we condition on the disaster realization, evaluate expectations over Gaussian shocks

analytically, and then average over disasters based on their empirical distribution:

$$\mathcal{P}(X,\xi_t) =$$
$$(1-p)\left(Xe^{-\gamma g+\gamma^2\sigma^2/2}\Phi\left(\nu+\gamma\sigma\right)-e^{(\phi-\gamma)g+(\phi-\gamma)^2\sigma^2/2}\Phi\left(\nu-(\phi-\gamma)\sigma\right)\right)$$
$$+\frac{p}{N}\sum_{i=1}^{N}\left(Xe^{-\gamma(g+Z_i)+\gamma^2\sigma^2/2}\Phi\left(\nu-\frac{Z_i}{\sigma}+\gamma\sigma\right)-e^{(\phi-\gamma)(g+Z_i)+(\phi-\gamma)^2\sigma^2/2}\Phi\left(\nu-\frac{Z_i}{\sigma}-(\phi-\gamma)\sigma\right)\right)$$

where

- $\nu=\frac{\log(X)/\phi-g}{\sigma}$
- N denotes the number of empirical disaster realizations
- Z_i is the log consumption drop observed during disaster i

and $\Phi(\cdot)$ denotes the CDF of a standard normal.

I.6 Wachter (2013)

We solve a discrete time version of Wachter (2013) in order to simplify the calculation of option prices. Log consumption growth Δc and the disaster probability λ follow

$$\begin{aligned}\Delta c_{t+1} &= \mu\, dt+\sigma\sqrt{dt}\varepsilon^c_{t+1}+Z_{t+1}\nu_{t+1}\\ \lambda_{t+1} &= \lambda_t+\kappa(\bar{\lambda}-\lambda_t)dt+\sigma_l\sqrt{dt}\sqrt{\lambda_t}\varepsilon^l_{t+1},\end{aligned} \tag{4}$$

where $\varepsilon^c_t,\varepsilon^l_t\sim N(0,1)$, $N_t\sim$Berloulli$(\lambda_t dt)$, and the disaster size Z_t is drawn from the empirical disaster size distribution.[8] All shocks are IID. The discretization assumes that either zero or one disaster occurs every period, with a disaster probability of $\lambda_t dt$ per period. Dividend growth is given by $\Delta d_t=\phi\times\Delta c_t$, and the agent has Epstein and Zin (1991) utility.

The calibration is identical to that in Wachter (2013). We solve the model at a monthly frequency ($dt=1/12$), using an equally-spaced grid for λ_t with 500 points

[8]The distribution of $1-e^{Z_t}$ is shown in Panel A of Figure 7 in Wachter (2013). We rely on the same data.

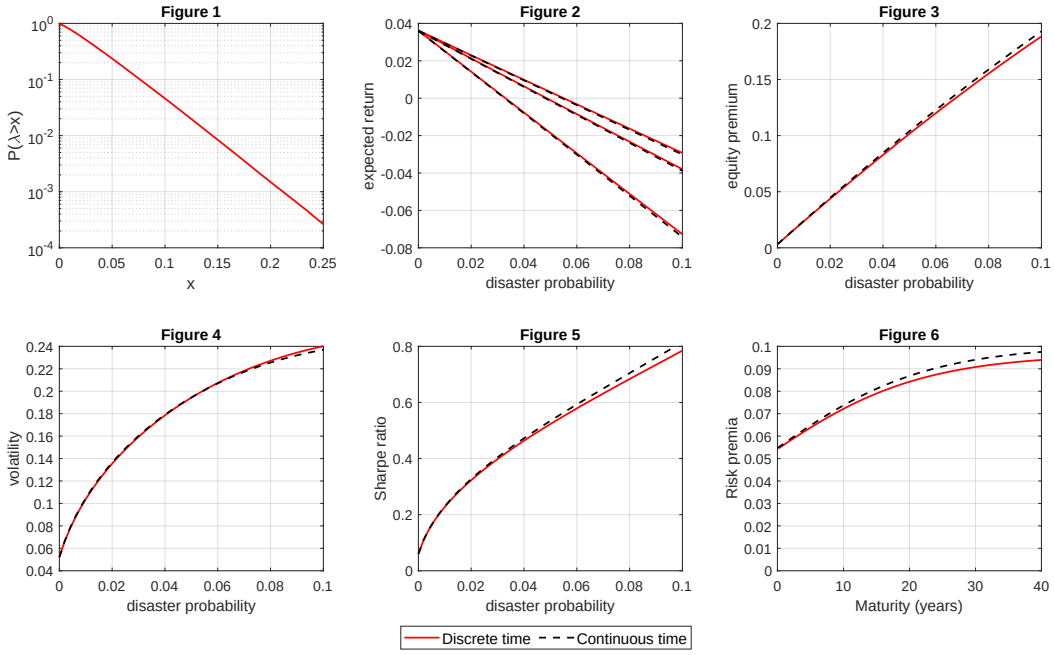

Figure 9: Replication of Wachter (2013). We replicate figures 1 through 6 from Wachter (2013) for the original continuous time model, and compare it to our discrete time version of the model.

between 1e-16 and 1. Expectations over ε^c_{t+1}, Z_{t+1}, and N_{t+1} are evaluated analytically, while expectations over ε^l_{t+1} are evaluated via Gauss-Chebyshev quadrature with 500 points between ± 7 standard deviations. We solve the value-consumption ratio and the price-dividend ratio by iterating on the system until convergence. Both ratios are interpolated with cubic splines to any off-grid values of λ_{t+1} that result from the quadrature. Figure 9 replicates figures 1 through 6 from Wachter (2013), and compares them to the results from our discrete time version of the model. The plots show that the discretized model provides a good approximation to its continuous time counterpart.

The calculation of option prices is similar to the one for Barro (2009). We condition on the disaster size realization, evaluate integrals over ε^c_{t+1} and ν_{t+1} analytically, integrals over ε^l_{t+1} with the aforementioned quadrature method, and finally evaluate

expectations over disaster realizations based on their empirical distribution.

I.7 Backus et al. (2011)

Backus et al. (2011) present alternative calibrations of the rare disaster model in Barro (2006). We rely on their preferred "small jump" calibration (calibration 4 in their Table II), which is designed to match observed option prices. The model admits analytical solutions for all asset prices, including options. We verified that our solution replicates the equity premium reported in the original study.

I.8 He and Krishnamurthy (2013)

He and Krishnamurthy (2013) propose financial intermediary constraints as a source of risk premia in asset markets. Their study focuses on mortgage-backed securities as an example of an intermediated asset, but the same mechanism has been considered to explain the equity premium in subsequent work (see, e.g. Muir (2017)). In order to apply the He-Krishnamurthy model to the equity market, we re-calibrate the dividend volatility to $\sigma = 15\%$, the time discount rate to $\rho = 0.16$, and the specialist's relative risk aversion to $\gamma = 1.05$. The remaining parameters are kept unchanged relative to the original calibration. Our calibration implies an equity premium of 7.55% per year, a return volatility of 15.05% per year, a risk-free rate of 0.09% per year, and a price-dividend ratio of 17.75. As in the original calibration, the model generates very little excess return volatility and a risk-free rate that is only marginally positive.

We solve the model numerically based on code on Zhiguo He's website,

`https://voices.uchicago.edu/zhiguohe`

Option prices are computed on a grid of 1,000 equally-spaced points for the model's state y_t (the ratio of the households' wealth and the current dividend level). The model is cast in continuous-time. We compute option prices for a maturity of 1

month via Monte Carlo integration by simulating the model at a daily frequency ($dt = 1/252$) for 21 trading days. To do so, the model is time-discretized based on an Euler scheme. We rely on 5×10^8 simulated paths for each value of the state, which is sufficient to ensure that simulation noise has a negligible effect on our results.

I.9 Constantinides and Ghosh (2017)

Constantinides and Ghosh (2017) extend the incomplete markets model of Constantinides and Duffie (1996) with countercyclical left skewness in idiosyncratic risk, and propose a calibration that matches basic quantity and return moments. We follow their solution approach, which approximates the wealth-consumption and price-dividend ratios with log-linear functions and then derives asset prices analytically.

We deviate from the original calibration for two reasons. First, while replicating the paper's main table, Table III, we discovered that it mistakenly calibrates the *quarterly* log price-dividend ratio in the model to the *annual* log price-dividend ratio in the data. The annual log price-dividend ratio in the model therefore exceeds its data counterpart by $\ln(4) = 1.39$. The authors confirmed this mistake via email. Second, the original calibration uses a quarterly frequency, but we require monthly option prices to compute $EP(x)$. We therefore re-calibrate the model at the monthly frequency, as shown in Table 2. Parameter values are chosen to match quarterly moments (and annual moments for the log price-dividend ratio) from the original paper based on simulated model data that is time-aggregated to the corresponding frequency. The quality of our calibration is comparable to that in the original paper and it matches the correct empirical counterpart of the price-dividend ratio.

The model features four IID shocks: standard normals in consumption and dividend growth and Poisson and Gamma shocks to the model's state x_t (household risk). We compute option prices via Monte Carlo integration on grid with 1,000 equally-spaced points for the state. We rely on 10^9 Monte Carlo draws and an identical seed for each value of the state, which is sufficient to ensure that simulation

Table 2: Monthly calibration of the Constantinides and Ghosh (2017) model

Prices									
	$E[r_f]$	$\sigma[r_f]$	$AC1[r_f]$	$E[r_m]$	$\sigma[r_m]$	$AC1[r_m]$	$E[p/d]$	$\sigma[p/d]$	$AC1[p/d]$
Data	0.005	0.005	0.863	0.019	0.087	0.056	3.745	0.411	0.972
Model	0.003	0.023	0.875	0.013	0.094	-0.063	3.728	0.185	0.872

Consumption and Dividends							
	$E[\Delta c]$	$\sigma[\Delta c]$	$E[\Delta d]$	$\sigma[\Delta d]$	$AC1[\Delta d]$	$\mu_2^{1/2}(\Delta c_{CEX})$	$\mu_3(\Delta c_{CEX})$
Data	0.004	0.004	0.005	0.026	0.328	0.379	-0.025
Model	0.004	0.003	0.007	0.016	0.211	0.363	-0.031

Preference Parameters		
γ	ψ	δ
1.18	1.01	0.982

Other Parameters										
μ	σ_a	ν	ξ	ρ	σ	α_d	σ_d	β_d	$\hat{\omega}$	$\hat{\sigma}$
0.001	0.002	0.044	0.00012	0.959	0.010	0.002	0.011	-0.034	0.222	0.272

Data moments are taken from Table III in Constantinides and Ghosh (2017). All series are quarterly. The log price-dividend ratio equals the current price, divided by the sum of dividends over the previous year. The model is solved at the monthly frequency, and simulated model data (for 300,000 months) is time-aggregated to the frequency of the data. Cross-sectional moments of household consumption growth are based on a simulated panel of 100,000 households over the same 300,000 months. Exact parameter values for the calibration are available in our replication code package.

noise has a negligible effect on our results. The option pricing equation (3) requires expressions for the pricing kernel, price-dividend ratio, and dividend growth, which are given in equations 3, 13, and 12 of the original paper, respectively.

I.10 Schreindorfer (2020)

The disappointment aversion model of Schreindorfer (2020) admits analytical solutions for all asset prices, including options. We solve the model based on the replication code on David Schreindorfers's website,

`www.davidschreindorfer.com`

J Computing $EP(x)$ in the models

$EP(x)$ derives from the densities $f(R)$ and $f^*(R)$, which we compute on an equally-spaced grid $R_1, \ldots, R_N$ with $N = 10,000$ points. Table 3 shows the endpoints of the grid, which were chosen to contain the relevant region of the densities' support for each model. Both densities are found by first computing the state-dependent (conditional) densities $f(R, \xi_t)$ and $f^*(R, \xi_t)$, and then averaging them over the model's state ξ_t based on a long simulation.

$f^*(R, \xi_t)$ is computed via Breeden and Litzenberger (1978) based on the put-to-spot price ratio, which we compute from (3) for each model as described in the previous section. Specifically, we find option prices for each point in the grid, as well as the points $R_i \pm \epsilon$ (for a small ϵ), and then compute the conditional density based on the second-order central difference

$$f^*(R_i, \xi_t) = (1 + R^f(\xi_t)) \frac{\mathcal{P}(1 + R_i - \epsilon, \xi_t) + \mathcal{P}(1 + R_i + \epsilon, \xi_t) - 2\mathcal{P}(1 + R_i, \xi_t)}{\epsilon^2}. \quad (5)$$

$f(R, \xi_t)$ is computed with a mixture of analytical, quadrature, and simulation methods. We then simulate each model's state for $L = 10^7$ periods and find the unconditional densities as

$$\begin{aligned} f(R_i) =& \frac{1}{L} \sum_{l=1}^{L} f_t(R_i, \xi_l), \quad i = 1, ..., N \\ f^*(R_i) =& \frac{1}{L} \sum_{l=1}^{L} f_t(R_i, \xi_l), \quad i = 1, ..., N. \end{aligned} \quad (6)$$

Based on the densities, the equity premium decomposition is computed as

$$EP(x) = \frac{\sum_{i=1}^{N} \mathbf{1}\{R_i \leq x\} R_i (f(R_i) - f^*(R_i))\delta}{\sum_{j=1}^{N} R_j (f(R_j) - f^*(R_j))\delta}, \quad (7)$$

where $\delta = R_{i+1} - R_i$ is the bin width and $\mathbf{1}\{\cdot\}$ an indicator function.

Table 3: Endpoints of the return grid for f and f^*

Model	R_1: lower bound	R_N: upper bound
Campbell-Cochrane (1999)	-0.35	0.35
Bekaert-Engstrom (2017)	-0.80	0.50
Bansal-Yaron (2004)	-0.35	0.35
Drechsler-Yaron (2011)	-0.80	1.00
Barro (2009)	-0.99	$3\sqrt{1/12}$
Wachter (2013)	-0.99	$3\sqrt{1/12}$
Backus-Chernov-Martin (2011)	-0.99	$3\sqrt{1/12}$
Schreindorfer (2020)	-0.60	0.50
Constantinides-Ghosh (2017)	-0.99	1.50
He-Krishnamurthy (2013)	-0.50	0.35

K Small Sample Model Assessment

In Table I of the paper, we compare empirical moments of stock market tail events to the corresponding population moments in asset pricing models. Here, we test the same models by assessing the likelihood with which they replicate the empirical moments in finite samples. We simulate 10 million samples of length 360 months (the length of our data) from each model, compute the equity premium contribution, probability, and price of risk for returns in the interval $[-30\%, -10\%]$ in each sample, and then report percentiles of the resulting sampling distribution.[9]

We initially follow the common approach of computing univariate confidence intervals for each statistic of interest, as for example in Beeler and Campbell (2012). Table 4 shows that the models of Campbell and Cochrane (1999), Barro (2009), and Wachter (2013) are rejected at the 5% significance level based on $EP(-.1) - EP(-.3)$, whereas the model of Bansal and Yaron (2004) is rejected at the 10% level. In contrast, the extensions of Backus et al. (2011), Drechsler and Yaron (2011),

[9]Note that the monthly decision interval of the models prevents us from replicating the daily sub-sampling approach that was used empirically. Model-based confidence intervals are therefore wider than their empirical counterparts.

and Bekaert and Engstrom (2017) cannot be rejected at any common significance level based on $EP(-.1) - EP(-.3)$, and only Backus et al. (2011) and Bekaert and Engstrom (2017) can be rejected based on price of risk $\frac{\int_{-.3}^{-.1} f^*(R)dR}{\int_{-.3}^{-.1} f(R)dR}$ at the 10% level. Based on this evidence alone, models with non-normal shocks therefore appear capable of generating finite samples that mimic our data.

However, an important shortcoming of this assessment is that a sequence of univariate tests is not as powerful as a single, multivariate test. For example, it is possible that a model never comes close to matching all three statistics of interest in a single finite sample, despite frequently generating each of them in different samples. This is problematic, as we have argued that a realistic account of $EP(x)$ requires high risk prices for tail events. Multivariate test of structural models are commonly based on formal estimations, either moment- or likelihood-based, and asymptotic inference procedures. Unfortunately, similar testing procedures do not appear to exist for assessing structural models multivariately in finite samples. We therefore rely on a simple, but admittedly ad-hoc, bivariate extension of the commonly-used univariate test. Specifically, we compute the probability with which a given model simultaneously produces an equity premium contribution $EP(-.1) - EP(-.3)$ and price of risk $\frac{\int_{-.3}^{-.1} f^*(R)dR}{\int_{-.3}^{-.1} f(R)dR}$ in excess of the statistics' respective empirical point estimates. The last column in Table 4 reports the p-value of this bivariate test. We find that, apart from Schreindorfer (2020), none of the models produce such samples with a probability above 5%. It is therefore unlikely for them to have generated the data.

Table 4: Small sample model statistics

Paper	$EP(-.1)-EP(-.3)$	$\int_{-.3}^{-.1} f(R)dR$	$\frac{\int_{-.3}^{-.1} f^*(R)dR}{\int_{-.3}^{-.1} f(R)dR}$	Bivariate p-value
Data, 1990-2019	0.666	0.016	2.627	
Campbell-Cochrane (1999)	0.028	0.003	1.382	0.003
	(-0.11, 0.24)	(0, 0.014)	(0.65, 3.70)	
	[-0.20, 0.33]	[0, 0.017]	[0.57, 4.16]	
Bekaert-Engstrom (2017)	0.388	0.039	1.440	0.013
	(-0.08, 0.79)	(0.020, 0.061)	(0.96, 2.56)	
	[-0.23, 0.93]	[0.017, 0.067]	[0.89, 2.96]	
Bansal-Yaron (2004)	0.072	0.014	1.205	0.005
	(-0.43, 0.49)	(0.003, 0.029)	(0.64, 3.39)	
	[-0.82, 0.84]	[0.003, 0.032]	[0.58, 4.34]	
Drechsler-Yaron (2011)	0.358	0.019	1.683	0.035
	(0.09, 0.93)	(0.006, 0.039)	(1.04, 4.09)	
	[0.02, 1.25]	[0.003, 0.044]	[0.97, 5.46]	
Barro (2009)	0.026	0.000	0.198	0.000
	(-0.13, 0.05)	(0, 0.003)	(0.10, 0.20)	
	[-0.17, 0.06]	[0, 0.003]	[0.10, 0.20]	
Wachter (2013)	0.052	0.008	1.192	0.001
	(-0.15, 0.22)	(0, 0.039)	(0.6, 3.05)	
	[-0.23, 0.27]	[0, 0.045]	[0.50, 3.76]	
Backus-Chernov-Martin (2011)	0.320	0.025	1.384	0.012
	(-0.38, 1.11)	(0.014, 0.039)	(0.89, 2.49)	
	[-0.98, 1.75]	[0.011, 0.044]	[0.78, 3.11]	
Schreindorfer (2020)	0.808	0.014	3.569	0.612
	(0.53, 1.58)	(0.006, 0.025)	(1.98, 8.92)	
	[0.50, 1.96]	[0.003, 0.028]	[1.78, 17.84]	
Constantinides-Ghosh (2017)	0.305	0.011	1.424	0.040
	(-0.03, 0.78)	(0, 0.053)	(0.92, 3.61)	
	[-0.09, 0.91]	[0, 0.067]	[0.86, 4.29]	
He-Krishnamurthy (2013)	0.104	0.006	1.831	0.000
	(-0.04, 0.28)	(0, 0.017)	(0.81, 4.75)	
	[-0.07, 0.34]	[0, 0.017]	[0.73, 5.33]	

This table presents finite sample statistics for the moments in Table I of the paper. We simulate 10 million samples from each model with the same length as the empirical sample and report the median across samples. Intervals in square brackets correspond to percentiles 2.5 and 97.5, and intervals in parentheses to percentiles 5 and 95. The last column shows a p-value for the bivariate test $H_0 : EP(-.1) - EP(-.3) \geq 0.666$ and $\frac{\int_{-.3}^{-.1} f^*(R)dR}{\int_{-.3}^{-.1} f(R)dR} \geq 2.627$, which assesses whether the models are able to simultaneously generate equity premium contributions and risk prices that are at least as large as in the data.

References

Abel, A. B. (1999): "Risk premia and term premia in general equilibrium," *Journal of Monetary Economics*, 43, 3–33.

Backus, D., M. Chernov, and I. Martin (2011): "Disasters Implied by Equity Index Options," *Journal of Finance*, 66, 1969–2012.

Bansal, R., D. Kiku, and A. Yaron (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

Bansal, R., S. Miller, D. Song, and A. Yaron (2020): "The Term Structure of Equity Risk Premia," *Journal of Financial Economics*, forthcoming.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 57, 1481–1509.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *Quarterly Journal of Economics*, 121, 823–866.

——— (2009): "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99, 243–264.

Beeler, J. and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bekaert, G. and E. Engstrom (2017): "Asset Return Dynamics under Habits and Bad Environment–Good Environment Fundamentals," *Journal of Political Economy*, 60, 713–760.

Belo, F., P. Collin-Dufresne, and R. S. Goldstein (2015): "Dividend Dynamics and the Term Structure of Dividend Strips," *Journal of Finance*, 70, 1115–1160.

Breeden, D. and R. H. Litzenberger (1978): "State Contingent Prices Implicit in option Prices," *Journal of Business*, 51, 3–24.

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

Constantinides, G. M. and D. Duffie (1996): "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 104, 219–240.

Constantinides, G. M. and A. Ghosh (2017): "Asset Pricing with Countercyclical Household Consumption Risk," *Journal of Finance*, 72, 415–460.

Drechsler, I. and A. Yaron (2011): "What's Vol Got to Do with It," *Review of Financial Studies*, 24, 1–45.

Epstein, L. G. and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

——— (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, 263–286.

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

He, Z. and A. Krishnamurthy (2013): "Intermediary Asset Pricing," *American Economic Review*, 103, 732–770.

Muir, T. (2017): "Financial Crises and Risk Premia," *Quarterly Journal of Economics*, 132, 765–809.

Pohl, W., K. Schmedders, and O. Wilms (2018): "Higher-Order Effects in Asset Pricing Models with Long-Run Risks," *Journal of Finance*, 73, 1061–1111.

Schreindorfer, D. (2020): "Macroeconomic Tail Risks and Asset Prices," *Review of Financial Studies*, 33, 3541–3582.

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

van Binsbergen, J. H. and R. S. Koijen (2017): "The term structure of returns: Facts and theory," *Journal of Financial Economics*, 124, 1–21.

Wachter, J. (2005): "Solving models with external habit," *Finance Research Letters*, 2, 210–226.

——— (2013): "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 68, 987–1035.