Appendix A. Long-Run Risk

Appendix A.1. Benchmark Model

We here lay out the technology and preference specifications of our model featuring LRR but not disasters. The simple baseline model of section 4.1 is clearly a special case.

Appendix A.1.1. Consumption Process

We take as a benchmark the Long Run Risk (LRR) model:

$$\begin{aligned}
\log g_{t+1} &= G_0 + x_t + v_t^{0.5} w_{g,t+1} \\
\log g_{d,t+1} &= G_{d,0} + \phi x_t + \varphi_d v_t^{0.5} w_{d,t+1} + \tau_d v_t^{0.5} w_{g,t+1} \\
x_{t+1} &= \rho x_t + \varphi_x v_t^{0.5} w_{x,t+1} \\
v_{t+1} &= (1-\varphi_v)\,\bar{v} + \varphi_v v_t + \sigma_v w_{v,t+1}
\end{aligned}$$

where $g_t \equiv \frac{C_t}{C_{t-1}}$ and $g_{d,t} \equiv \frac{D_t}{D_{t-1}}$ are consumption and dividend growth, respectively, and $w_{i,t+1}$ for $i \in \{g, d, x, v\}$ are standard Normal iid innovations.

Appendix A.1.2. Preferences

There is an algebraic equivalence between the recursion defining the utility of a particular Epstein-Zin agent and the recursively defined indirect utility function of a robust agent. Hence we start with an Epstein-Zin recursive definition of utility and will draw out the equivalence in the following analysis. For a more formal treatment we refer the reader to Barillas, Hansen, and Sargent (2009).

Consider an Epstein-Zin utility recursion

$$\begin{aligned}
U_t &= \left[(1-\beta)\, C_t^{\rho} + \beta \mu_t\,(U_{t+1})^{\rho}\right]^{\frac{1}{\rho}} \\
\mu_t\,(U_{t+1}) &= E_t\left[U_{t+1}^{\alpha}\right]^{\frac{1}{\alpha}}
\end{aligned}$$

where ρ controls time preference (with $\frac{1}{1-\rho}$ being the IES, so $\rho \to 0$ implies $IES \to 1$) and α controls the 'risk aversion' inherent in the certainty equivalent function, μ_t (with $CRRA = 1-\alpha$).[22]

[22] If we take $\rho \to 0$ (unity intertemporal elasticity of substitution) we obtain

$$U_t = C_t^{1-\beta} \mu_t\,(U_{t+1})^{\beta}$$

Since we will be working within a non-stationary environment we scale by consumption as follows, where $u_t \equiv \frac{U_t}{C_t}$

$$\begin{aligned} u_t &= \left[(1-\beta) + \beta E_t \left[\frac{U_{t+1}^{\alpha}}{C_t^{\alpha}}\right]^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \\ &= \left[(1-\beta) + \beta E_t \left[\left(\frac{C_{t+1}}{C_t}\right)^{\alpha}\left(\frac{U_{t+1}}{C_{t+1}}\right)^{\alpha}\right]^{\frac{\rho}{\alpha}}\right]^{\frac{1}{\rho}} \end{aligned}$$

which implies a stochastic discount factor, Λ_{t+1}, of the form

$$\begin{aligned} \Lambda_{t+1} &= \Lambda_{t+1}^{r}\Lambda_{t+1}^{u} \\ \Lambda_{t+1}^{r} &\equiv \beta g_{t+1}^{\rho-1} \\ \Lambda_{t+1}^{u} &\equiv \left(\frac{g_{t+1}u_{t+1}}{\mu_t\left(g_{t+1}u_{t+1}\right)}\right)^{\alpha-\rho} \\ &= \left(\frac{g_{t+1}u_{t+1}}{E_t\left[g_{t+1}^{\alpha}u_{t+1}^{\alpha}\right]^{\frac{1}{\alpha}}}\right)^{\alpha-\rho} \end{aligned}$$

In the unit elasticity of substitution (EIS) case, where $\rho \to 0$, we have

$$\log U_t = (1-\beta)\log C_t + \frac{\beta}{\alpha}\log E_t\left[\exp\left\{\alpha \log U_{t+1}\right\}\right] \tag{A.1}$$

or

$$W_t = (1-\beta)\log C_t - \beta\theta \log E_t\left[\exp\left\{-\frac{W_{t+1}}{\theta}\right\}\right] \tag{A.2}$$

$$W_t \equiv \log U_t \tag{A.3}$$

$$\theta \equiv -\frac{1}{\alpha} \tag{A.4}$$

so that we can obtain, in equation (A.2), a recursion of the form of equation (7) in the text, interpretable as that of a robust agent with $h(s_t) = (1-\beta)\log C_t$. The robustness parameter θ can be mapped to α and *vice versa*.

Re-scaling by consumption yields a recursion in terms of stationary variables:

$$W_t - \log C_t = -\beta \log E_t \left[\exp \left\{ -\frac{W_{t+1} - \log C_{t+1} + g_{t+1}}{\theta} \log U_{t+1} \right\} \right] \tag{A.5}$$

$$or - \beta \log E_t \left[\exp \left\{ -\frac{w_{t+1} + g_{t+1}}{\theta} \log U_{t+1} \right\} \right] \tag{A.6}$$

$$w_t \equiv W_t - \log C_t \tag{A.7}$$

$$\equiv \log u_t \tag{A.8}$$

Returning to the Epstein-Zin representation:

$$\begin{aligned}
\log u_t &= -\beta \log C_t + \frac{\beta}{\alpha} \log E_t \left[\exp \left\{ \alpha \log U_{t+1} \right\} \right] \\
&= \frac{\beta}{\alpha} \log E_t \left[\exp \left\{ \alpha \left(\log U_{t+1} - \log C_{t+1} \right) + \alpha \log C_{t+1} - \alpha \log C_t \right\} \right] \\
&= 0 + \beta \log E_t \left[\exp \left\{ \alpha \left(\log g_{t+1} + \log u_{t+1} \right) \right\} \right]^{\frac{1}{\alpha}}
\end{aligned}$$

and

$$\begin{aligned}
\Lambda^r_{t+1} &\equiv \beta g^{-1}_{t+1} \\
\Lambda^u_{t+1} &\equiv \left(\frac{g_{t+1} u_{t+1}}{\mu_t \left(g_{t+1} u_{t+1} \right)} \right)^{\alpha}
\end{aligned} \tag{A.9}$$

So we observe that, as noted by Backus, Chernov, and Zin (2014), the log scaled value of the agent's problem is affine in the log of the certainty equivalent of the scaled value of the problem in the next period, where the intercept of this affine relationship is, in fact, 0 and the slope is β. It is useful to derive expressions in terms of $\alpha \equiv -\frac{1}{\theta}$ although we retain the robustness, rather than EZ, interpretations of the objects.

Appendix A.1.3. Value Function Approximation

We will guess and verify an exponentially affine solution for the scaled utility function, u_t

$$u_t = \exp \left\{ F_0 + F_1 x_t + F_2 v_t \right\}$$

which implies

$$\begin{aligned}
g_{t+1} u_{t+1} &= \eta_t \cdot \exp \left\{ F_1 \varphi_x v_t^{0.5} w_{x,t+1} + F_2 \sigma_v w_{v,t+1} + v_t^{0.5} w_{g,t+1} \right\} \\
\eta_t &\equiv \exp \left\{ F_0 + G_0 + \left(1 + F_1 \rho \right) x_t + F_2 \left(1 - \varphi_v \right) \bar{v} + F_2 \varphi_v v_t \right\}
\end{aligned}$$

where we have conveniently split the exponent into components known in t and those revealed in $t+1$. We then have

$$\mu_t\left(g_{t+1}u_{t+1}\right) = \eta_t \cdot \mu_t\left(e^{F_1\varphi_x v_t^{0.5} w_{x,t+1}}\right)\mu_t\left(e^{F_2\sigma_v w_{v,t+1}}\right)\mu_t\left(e^{v_t^{0.5} w_{g,t+1}}\right)$$

and we can make use of the fact that

$$\begin{aligned} \log \mu_t\left(\exp\left\{\delta_0 + \delta_1 Y_{t+1}\right\}\right) &= \delta_0 + \frac{k_t\left(\delta_1\alpha\right)}{\alpha} \\ k_t\left(s\right) &\equiv \log E_t\left[\exp\left\{sY_{t+1}\right\}\right] \end{aligned}$$

where $k_t\left(s\right)$ is the conditional cumulant generating function (CGF) for a random variable, Y_{t+1}, and δ_0 and δ_1 are known in t. Consequently, since the CGF of a standard Normal variable is,

$$k_t\left(s\right) = \frac{s^2}{2}$$

we have

$$\begin{aligned} \log \mu_t\left(g_{t+1}u_{t+1}\right) &= F_0 + G_0 + F_2\left(1-\varphi_v\right)\bar{v} + \frac{\alpha F_2^2\sigma_v^2}{2} \\ &\quad + \left(1 + F_1\rho\right)x_t \\ &\quad + \left(F_2\varphi_v + \frac{\alpha\left(1 + F_1^2\varphi_x^2\right)}{2}\right)v_t \end{aligned}$$

and (recalling our guess), matching coefficients yields

$$\begin{aligned} F_0 &= \beta\left(F_0 + G_0 + F_2\left(1-\varphi_v\right)\bar{v} + \frac{\alpha F_2^2\sigma_v^2}{2}\right) \\ F_1 &= \beta\left(1 + F_1\rho\right) \\ F_2 &= \beta\left(F_2\varphi_v + \frac{\alpha\left(1 + F_1^2\varphi_x^2\right)}{2}\right) \end{aligned}$$

from which we can obtain

$$\begin{aligned} F_1 &\equiv \frac{\beta}{1-\beta\rho} \\ F_2 &= \frac{\beta}{1-\beta\varphi_v}\frac{\alpha\left(1 + F_1^2\varphi_x^2\right)}{2} \\ F_0 &= \frac{\beta}{1-\beta}\left(G_0 + F_2\left(1-\varphi_v\right)\bar{v} + \frac{\alpha F_2^2\sigma_v^2}{2}\right) \end{aligned}$$

Appendix A.1.4. Stochastic Discount Factor Approximation

We will need $\log(g_{t+1}u_{t+1}) - \log\mu_t(g_{t+1}u_{t+1})$ for deriving the stochastic discount factor (SDF)

$$\begin{aligned}\log\left(\frac{g_{t+1}u_{t+1}}{\mu_t(g_{t+1}u_{t+1})}\right) &= F_1\varphi_x v_t^{0.5} w_{x,t+1} + F_2\sigma_v w_{v,t+1} + v_t^{0.5} w_{g,t+1} \\ &\quad -\alpha\left(\frac{1+F_1^2\varphi_x^2}{2}\right)v_t - \frac{\alpha F_2^2\sigma_v^2}{2}\end{aligned}$$

Given the above derivations we can obtain two expressions for the log of the stochastic discount factor (SDF), which are each more convenient than the other in particular contexts. The first expression makes clear the distinction between properties of the SDF that reflect the standard, expected utility elements of the agent's preferences, and those that reflect her preference for robustness arising from an aversion to perceived model uncertainty:

$$\log\Lambda_{t+1} = \log\Lambda^r_{t+1} + \log\Lambda^u_{t+1} \tag{A.10}$$

$$\log\Lambda^r_{t+1} = k_t + \lambda^r_{g,t} w_{g,t+1} \tag{A.11}$$

$$k_t \equiv \log\beta - G_0 - x_t \tag{A.12}$$

$$\lambda^r_{g,t} \equiv -v_t^{0.5} \tag{A.13}$$

$$\log\Lambda^u_{t+1} \equiv \xi_t + \lambda^u_{g,t} w_{g,t+1} + \lambda^u_{x,t} w_{x,t+1} + \lambda^u_v w_{v,t+1} \tag{A.14}$$

$$\xi_t \equiv -\alpha^2\left(\frac{F_2^2\sigma_v^2 + \left(1+F_1^2\varphi_x^2\right)v_t}{2}\right) \tag{A.15}$$

$$\lambda^u_{g,t} \equiv \alpha v_t^{0.5} \tag{A.16}$$

$$\lambda^u_{x,t} \equiv \alpha F_1\varphi_x v_t^{0.5} \tag{A.17}$$

$$\lambda^u_v \equiv \alpha F_2\sigma_v \tag{A.18}$$

A second representation is useful when deriving other endogenous variables, such as asset prices:[23]

$$
\begin{aligned}
\log \Lambda_{t+1} &= m_0 + m_1 x_t + m_2 v_t \\
&\quad + m_3 v_t^{0.5} w_{d,t+1} + m_4 v_t^{0.5} w_{g,t+1} \\
&\quad + m_5 v_t^{0.5} w_{x,t+1} + m_6 w_{v,t+1} \\
m_0 &\equiv \log \beta - G_0 - \frac{\alpha^2 F_2^2 \sigma_v^2}{2} \\
m_1 &\equiv -1 \\
m_2 &\equiv -\frac{\alpha^2 \left(1 + F_1^2 \varphi_x^2\right)}{2} \\
m_3 &\equiv 0 \\
m_4 &\equiv \alpha - 1 \\
m_5 &\equiv \alpha F_1 \varphi_x \\
m_6 &\equiv \alpha F_2 \sigma_v
\end{aligned}
$$

Appendix A.2. Worst Case Model

The worst case model is a distribution over sequences and can implicitly capture a broad class of possible misspecifications in the benchmark. It is convenient to represent this distribution recursively by constructing state dependent worst case conditional (on t) distributions of the innovations, $w_{i,t+1}$ and combining these with the deterministic elements of the law of motion under the benchmark.

Appendix A.2.1. Benchmark and Worst Case Innovation Distributions

For $i \in \{d, g, x, v\}$ the innovation $w_{i,t+1}$ is $N(0,1)$ under the benchmark model and therefore has conditional moment generating function

$$
\begin{aligned}
h_t^i(u) &= E_t\left[\exp\left\{u w_{i,t+1}\right\}\right] \\
&= \exp\left\{\frac{u^2}{2}\right\}
\end{aligned}
$$

To obtain the moment generating function of $w_{i,t+1}$ under the worst case, denoted $\tilde{h}_t^i$, we apply the worst case change in measure captured in Λ_{t+1}^u. Note that given the linearity of $\log \Lambda_{t+1}^u$ in the innovations and their independence under the benchmark, they are also conditionally independent under the worst case. We abstract from $w_{d,t+1}$ as it is undistorted under the worst case, being

[23] $m_3 \equiv 0$ since the innovation to the dividend is simply noise and not priced, but the term is included nevertheless as it is notationally convenient.

welfare irrelevant.

$$\begin{aligned}
\tilde{h}_t^i(u) &= E_t\left[\exp\left\{uw_{i,t+1}\right\}\Lambda_{t+1}^u\right] \\
&= E_t\left[\exp\left\{uw_{g,t+1}+\xi_t+\lambda_{g,t}^u w_{g,t+1}+\lambda_{x,t}^u w_{x,t+1}+\lambda_v^u w_{v,t+1}\right\}\right] \\
&= \exp\left\{\xi_t\right\}E_t\left[\exp\left\{\left(u+\lambda_{g,t}^u\right)w_{g,t+1}\right\}\right]E_t\left[\exp\left\{\lambda_{x,t}^u w_{x,t+1}\right\}\right]E_t\left[\exp\left\{\lambda_v w_{v,t+1}\right\}\right] \\
&= h_t^g\left(u+\lambda_{g,t}^u\right)h_t^x\left(\lambda_{x,t}\right)h_t^v\left(\lambda_v\right)\exp\left\{\xi_t\right\}
\end{aligned}$$

or, in terms of cumulant generating functions ($k_t^i(u) \equiv \log h_t^i(u) = \log E_t\left[u \cdot w_{i,t+1}\right]$, $\tilde{k}_t$ is defined similarly and tiu denotes terms independent of u)

$$\begin{aligned}
\tilde{k}_t^d(u) &= k_t^d(u)+tiu \\
\tilde{k}_t^g(u) &= k_t^g\left(u+\lambda_{g,t}^u\right)+tiu \\
\tilde{k}_t^x(u) &= k_t^x\left(u+\lambda_{x,t}^u\right)+tiu \\
\tilde{k}_t^v(u) &= k_t^v\left(u+\lambda_v^u\right)+tiu
\end{aligned}$$

In our case we have $k_t^d = k_t^g = k_t^x = k_t^v = k$ where $k(u) = \frac{u^2}{2}$. Now, the j^{th} derivative of the CGF, evaluated at 0 yields the j^{th} cumulant. The first two cumulants are the mean and variance, respectively. In our case we have, for a variable b_t known in t

$$\begin{aligned}
\tilde{k}_t^1(u) &= k^1\left(u+b_t\right)=u+b_t \\
\tilde{k}_t^2(u) &= k^2\left(u+b_t\right)=1 \\
\tilde{k}_t^j(u) &= k^j\left(u+b_t\right)=0 \; for \; j>2
\end{aligned}$$

where $k^{(j)}$ means the j^{th} derivative of k. We therefore see that the only effects on the innovation distributions under the worst case are mean shifts. The properties of unit standard deviation and Normality (as indicated by higher order cumulants being zero) are preserved under the worst case. Consequently, the worst case distribution over sequences can be implicitly defined in terms of a Gaussian system of the same form as the benchmark model but with different innovation distributions. In the case of $w_{g,t+1}$, $w_{x,t+1}$ and $w_{v,t+1}$ the means are pessimistically shifted and, in the case of $w_{g,t+1}$ and $w_{x,t+1}$, these shifts are state dependent, through reliance of $\lambda_{g,t}^u$ and $\lambda_{x,t}^u$ on v_t. Since the innovation $w_{d,t+1}$ simply captures noise and is welfare-irrelevant, the 'innovations representation' of the worst case distribution over sequences makes no use of distortions in its moments.

$$\begin{aligned}
\tilde{E}\left[w_{d,t+1}\right] &= E\left[w_{d,t+1}\right]=0 \\
\tilde{E}_t\left[w_{g,t+1}\right] &= \lambda_{g,t}^u=\alpha v_t^{0.5} \\
\tilde{E}_t\left[w_{x,t+1}\right] &= \lambda_{x,t}^u=\alpha F_1 \varphi_x v_t^{0.5} \\
\tilde{E}\left[w_{v,t+1}\right] &= \lambda_v^u=\alpha F_2 \sigma_v
\end{aligned}$$

Appendix A.2.2. Alternative Representations of the Worst Case

In the absence of jump components we can derive a Gaussian representation of the worst case where innovations are again taken to be standard Normal, by simply incorporating the mean shifts described above into the conditional mean dynamics:

$$\begin{aligned}
\log g_{t+1} &= G_0+x_t+v_t^{0.5} w_{g,t+1} \\
&= G_0+x_t+\alpha v_t+v_t^{0.5} \varepsilon_{g,t+1} \\
\log g_{d,t+1} &= G_{d,0}+\phi x_t+\varphi_d v_t^{0.5} w_{d,t+1}+\tau_d v_t^{0.5} w_{g,t+1} \\
&= G_{d,0}+\phi x_t+\tau_d \alpha v_t+\varphi_d v_t^{0.5} \varepsilon_{d,t+1}+\tau_d v_t^{0.5} \varepsilon_{g,t+1} \\
x_{t+1} &= \rho x_t+\varphi_x v_t^{0.5} w_{x,t+1} \\
&= \rho x_t+\varphi_x^2 \alpha F_1 v_t+\varphi_x v_t^{0.5} \varepsilon_{x,t+1} \\
v_{t+1} &= \left(1-\varphi_v\right) \bar{v}+\varphi_v v_t+\sigma_v w_{v,t+1} \\
&= \left(1-\varphi_v\right) \bar{v}+\sigma_v^2 \alpha F_2+\varphi_v v_t+\sigma_v \varepsilon_{v,t+1} \\
&= \left(1-\varphi_v\right) \tilde{v}+\varphi_v v_t+\sigma_v \varepsilon_{v,t+1} \\
\varepsilon_{i,t+1} &\backsim N(0,1) \;\; for \; i \in\{g, x, v, d\}
\end{aligned}$$

One can usefully represent this system in terms of loadings on an 'economic state', $\hat{s}_{t+1} \equiv \left[1, x_t, v_t, \varepsilon_{d,t+1}, \varepsilon_{g,t+1}, \varepsilon_{x,t+1}, \varepsilon_{v,t+1}\right]$ and note that the benchmark can also be represented in this way, with appropriate zeroing out / redefinition of coefficients.

Appendix A.3. Asset Pricing

We here lay out various derivations of asset pricing objects that we use or which may be of interest.

Appendix A.3.1. Market Price of 'Risk'

We have that

$$\begin{aligned}
E_t\left[\Lambda_{t+1}\right] &= \exp\left\{m_0 + m_1 x_t + m_2 v_t\right\} E_t\left[\hat{w}_{t+1}\right] \\
\sigma_t\left(\Lambda_{t+1}\right) &= \exp\left\{m_0 + m_1 x_t + m_2 v_t\right\} \sigma_t\left(\hat{w}_{t+1}\right) \\
\hat{w}_{t+1} &\equiv \exp\left\{\left(m_3 w_{d,t+1} + m_4 w_{g,t+1} + m_5 w_{x,t+1}\right) v_t^{0.5} + m_6 w_{v,t+1}\right\}
\end{aligned}$$

and, thus, the conditional market price of risk (MPR) is then given by

$$MPR_t \equiv \frac{\sigma_t\left(\Lambda_{t+1}\right)}{E_t\left[\Lambda_{t+1}\right]} = \frac{\sigma_t\left(\hat{w}_{t+1}\right)}{E_t\left[\hat{w}_{t+1}\right]}$$

Let us first consider the expectation in the denominator

$$\begin{aligned}
E_t\left[\hat{w}_{t+1}\right] &= \exp\left\{\Omega_0 + \Omega_v v_t\right\} \\
\Omega_0 &\equiv \frac{m_6^2}{2} \\
\Omega_v &\equiv \frac{m_3^2 + m_4^2 + m_5^2}{2}
\end{aligned}$$

and the numerator is given by

$$\sigma_t^2\left(\hat{w}_{t+1}\right) = \exp\left\{2\left(\Omega_0 + \Omega_v v_t\right)\right\}\left(\exp\left\{2\left(\Omega_0 + \Omega_v v_t\right)\right\} - 1\right)$$

so we have

$$\begin{aligned}
\frac{\sigma_t\left(\hat{w}_{t+1}\right)}{E_t\left[\hat{w}_{t+1}\right]} &= \sqrt{\exp\left\{2\left(\Omega_0 + \Omega_v v_t\right)\right\} - 1} \\
&= \sqrt{E_t\left[\hat{w}_{t+1}\right]^2 - 1}
\end{aligned}$$

Now, using our expressions for m_i above and approximating $(\alpha - 1)^2$ with α^2 (which will be a reasonable approximation for our heteroscedastic white noise calibrations, given the 'large' magnitudes of α) then we obtain the maximal Sharpe ratio (taking the benchmark measure as the true DGP) under the expected utility and robustness, respectively, as

$$\begin{aligned}
\sqrt{\exp\left(v_t\right) - 1} &\approx \sqrt{v_t} \\
\sqrt{\exp\left(\alpha^2\left(v_t + F_1^2\varphi_x^2 + F_2^2\sigma_v^2\right)\right) - 1} &\approx \alpha\sqrt{v_t + F_1^2\varphi_x^2 + F_2^2\sigma_v^2}
\end{aligned}$$

In fact, the latter is equivalent to the 'market price of uncertainty', $\frac{\sigma_t\left(\Lambda_{t+1}^U\right)}{E_t\left[\Lambda_{t+1}^U\right]} \equiv \sigma_t\left(\Lambda_{t+1}^U\right)$. This again emphasizes that, for our models and calibrations, the *uncertainty* component of the stochastic factor will drive the behavior of what is typically termed the market price of *risk*.

For the unconditional counterparts, temporal dependence among the states prevents analytic expressions, and we obtain unconditional means and standard deviations of the relevant components of the stochastic discount factor via simulation over long horizons.

Appendix A.3.2. Risk Free Rate

Since the 1-period risk free rate is the reciprocal of the conditional expectation of the stochastic discount factor we have

$$\begin{aligned}
R^f_{t,t+1} &= \frac{1}{E_t\left[\Lambda_{t+1}\right]} \\
&= \exp\left\{\chi_0 + \chi_1 x_t + \chi_2 v_t\right\} \\
\chi_0 &\equiv -\left(m_0 + \frac{m_6^2}{2}\right) \\
\chi_1 &\equiv -m_1 \\
\chi_2 &\equiv -\left(m_2 + \frac{m_3^2 + m_4^2 + m_5^2}{2}\right)
\end{aligned}$$

where we have used CGFs / conditional log-Normality as we did in deriving the MPR.

Appendix A.3.3. Term Structures

Consider the 'Euler' equation

$$\begin{aligned}
x_{n,t} &= E_t\left[\Omega_{t+1} x_{n-1,t}\right] \\
x_{0,t} &\equiv 0 \\
x_{n,t} &= \exp\left\{A_n + B_n x_t + C_n v_t\right\} \\
\log \Omega_{t+1} &= h_0 + h_1 x_t + h_2 v_t \\
&\quad + h_3 v_t^{0.5} w_{d,t+1} + h_4 v_t^{0.5} w_{g,t+1} \\
&\quad + h_5 v_t^{0.5} w_{x,t+1} + h_6 w_{v,t+1}
\end{aligned}$$

We show how to obtain recursions for the coefficients A_n, B_n and C_n beginning from the initial condition of $A_0 = B_0 = C_0 = 0$, and then specialize this setup to particular assets. Formally, we have

$$
\begin{aligned}
x_{n,t} &= \exp\left\{h_0 + A_{n-1} + C_{n-1}\left(1-\varphi_v\right)\bar{v} + \left(h_1 + B_{n-1}\rho\right)x_t + \left(h_2 + C_{n-1}\varphi_v\right)v_t\right\} \\
&\quad \times \\
&\quad E_t\left[\exp\left\{h_3 v_t^{0.5} w_{d,t+1} + h_4 v_t^{0.5} w_{g,t+1} + \left(h_5 + B_{n-1}\varphi_x\right)v_t^{0.5} w_{x,t+1} + \left(h_6 + C_{n-1}\sigma_v\right)w_{v,t+1}\right\}\right]
\end{aligned}
$$

and (using cumulant generating functions) we can match coefficients to obtain

$$
\begin{aligned}
A_n &= A_{n-1} + h_0 + C_{n-1}\left(1-\varphi_v\right)\bar{v} + \frac{\left(h_6 + C_{n-1}\sigma_v\right)^2}{2} \\
B_n &= h_1 + B_{n-1}\rho \\
C_n &= C_{n-1}\varphi_v + h_2 + \frac{h_3^2 + h_4^2 + \left(h_5 + B_{n-1}\varphi_x\right)^2}{2}
\end{aligned}
$$

If we wish to derive the real term structure of riskless zero coupon bonds, we would let $\Omega_{t+1} = \Lambda_{t+1}$ (the SDF, so $h_i = m_i$) and in that case $x_{n,t}$ would be the price in t of a riskless claim to a unit of consumption in $t+n$.

Appendix A.3.4. Return on Equity

We begin with the Euler equation

$$
P_t = E_t\left[\Lambda_{t+1}\left(P_{t+1} + D_{t+1}\right)\right]
$$

or, expressed in terms of returns,

$$
\begin{aligned}
1 &= E_t\left[\Lambda_{t+1}\frac{D_{t+1}}{D_t}\frac{\frac{P_{t+1}}{D_{t+1}} + 1}{\frac{P_t}{D_t}}\right] \\
&= E_t\left[\Lambda_{t+1}R_{t+1}\right] \\
R_{t+1} &\equiv g_{d,t+1}\frac{\frac{P_{t+1}}{D_{t+1}} + 1}{\frac{P_t}{D_t}}
\end{aligned}
$$

Then, defining $pd_t \equiv \log\left(\frac{P_t}{D_t}\right)$ and we have $r_{t+1} \equiv \log R_{t+1} = \log g_{d,t+1} - pd_t + \log\left(1 + \frac{P_t}{D_t}\right)$ which, using the Campbell-Shiller approach can be approximated by

$$
r_{t+1} \approx b_0 + b_1 pd_{t+1} - pd_t + \log g_{d,t+1}
$$

where b_0 is a constant and b_1 is a linearization parameter slightly less than 1.[24] Now, we conjecture that pd_t can be expressed as follows

$$pd_t = A_0 + A_1 x_t + A_2 v_t$$

which implies that we can express r_{t+1} as

$$\begin{aligned}
r_{t+1} &= \mu_0 + \mu_1 x_t + \mu_2 v_t \\
&\quad +\mu_3 v_t^{0.5} w_{d,t+1} + \mu_4 v_t^{0.5} w_{g,t+1} \\
&\quad +\mu_5 v_t^{0.5} w_{x,t+1} + \mu_6 w_{v,t+1} \\
\mu_0 &\equiv b_0 + (b_1 - 1) A_0 + G_{d,0} + b_1 A_2 (1 - \varphi_v) \bar{v} \\
\mu_1 &\equiv (b_1 \rho - 1) A_1 + \phi \\
\mu_2 &\equiv (b_1 \varphi_v - 1) A_2 \\
\mu_3 &\equiv \varphi_d \\
\mu_4 &\equiv \tau_d \\
\mu_5 &\equiv b_1 A_1 \varphi_x \\
\mu_6 &\equiv b_1 A_2 \sigma_v
\end{aligned}$$

Consequently, we have (using the earlier expression for the SDF)

$$\begin{aligned}
0 &= \log E_t \left[\exp \left\{ \begin{array}{c} \Omega_0 + \Omega_1 x_t + \Omega_2 v_t \\ +\Omega_3 v_t^{0.5} w_{d,t+1} + \Omega_4 v_t^{0.5} w_{g,t+1} \\ +\Omega_5 v_t^{0.5} w_{x,t+1} + \Omega_6 w_{v,t+1} \end{array} \right\} \right] \\
&= \Omega_0 + \Omega_1 x_t + \left(\Omega_2 + \frac{\Omega_3^2 + \Omega_4^2 + \Omega_5^2}{2} \right) v_t + \frac{\Omega_6^2}{2} \\
\Omega_i &\equiv m_i + \mu_i
\end{aligned}$$

[24] $b_0 = \log(1 + \exp(pd)) - pd \frac{\exp(pd)}{1+\exp(pd)}$ and $b_1 = \frac{\exp(pd)}{1+\exp(pd)}$ where pd is the point of approximation for the price dividend ratio around which $\log\left(1 + \frac{P_t}{D_t}\right)$ is linearized.

which implies

$$
\begin{aligned}
0 &= \Omega_0 + \frac{\Omega_6^2}{2} \\
0 &= \Omega_1 \\
0 &= \Omega_2 + \frac{\Omega_3^2 + \Omega_4^2 + \Omega_5^2}{2}
\end{aligned}
$$

Thus, $\Omega_1 = 0$ implies

$$A_1 = \frac{\phi + m_1}{1 - b_1 \rho}$$

and we then consider $0 = \Omega_2 + \frac{\Omega_3^2+\Omega_4^2+\Omega_5^2}{2}$ which implies

$$A_2 = \frac{m_2 + \frac{\Omega_3^2+\Omega_4^2+\Omega_5^2}{2}}{1 - b_1 \varphi_v}$$

Finally, we consider $0 = \Omega_0 + \frac{\Omega_6^2}{2}$, which implies

$$A_0 = \frac{1}{1 - b_1} \left(b_0 + G_{d,0} + b_1 A_2 \left(1 - \varphi_v\right) \bar{v} + m_0 + \frac{\Omega_6^2}{2} \right)$$

where Ω_6 depends on A_2, for which we have already solved.

Appendix A.4. Detection Error Probabilities

We will calculate detection error probabilities (DEPs) as a way of assessing how easy it is to distinguish the benchmark and worst case models. We simulate n_{loop} samples of length T from both the benchmark and worst case models and each sample of length T is subject to a likelihood ratio test between the two models. The detection error probability is the equally weighted average rate of a detection error (the likelihood ratio test favors the wrong model) under the benchmark and worst case simulations.[25]

In order to calculate the likelihood ratios we must make a decision on how to initialize the simulations and how to treat the $t = 0$ period in the likelihood evaluations. We will draw the initial state for each T-long simulation from the relevant unconditional distribution (or an approximation thereof, if necessary). We choose not to incorporate the unconditional likelihood of the time zero observation into the likelihoods used for the DEP calculations, as is consistent with Hansen and

[25] Clearly this number is subject to sampling variability so one must use many draws of sequences of length T to obtain a reliable measure of detectability.

Sargent (2008), Chapter 9. If we were to allow the agent to use the time zero observation then it would implicitly be as if we were allowing the agent far longer hypothetical samples than length T to distinguish the two models.

Recall that we have

$$\begin{aligned}
\log \Lambda_{t+1}^{u} &\equiv \xi_t + \lambda_{g,t}^{u} w_{g,t+1} + \lambda_{x,t}^{u} w_{x,t+1} + \lambda_{v}^{u} w_{v,t+1} \\
\xi_t &\equiv -\alpha^2 \left(\frac{F_2^2 \sigma_v^2 + \left(1 + F_1^2 \varphi_x^2\right) v_t}{2} \right) \\
\lambda_{g,t}^{u} &\equiv \alpha v_t^{0.5} \\
\lambda_{x,t}^{u} &\equiv \alpha F_1 \varphi_x v_t^{0.5} \\
\lambda_{v}^{u} &\equiv \alpha F_2 \sigma_v
\end{aligned}$$

Under the worst case the conditional likelihood for observation, $\{w_{g,t}, w_{x,t}, w_{v,t}\}$ in t given $t-1$ is

$$\begin{aligned}
L_{t|t-1}^{wc} &= \pi\left(w_{g,t}\right) \pi\left(w_{x,t}\right) \pi\left(w_{v,t}\right) \exp\left\{\xi_{t-1} + \lambda_{g,t-1}^{u} w_{g,t} + \lambda_{x,t-1}^{u} w_{x,t} + \lambda_{v}^{u} w_{v,t}\right\} \\
&= L_{t|t-1}^{bench} \exp\left\{\xi_{t-1} + \lambda_{g,t-1}^{u} w_{g,t} + \lambda_{x,t-1}^{u} w_{x,t} + \lambda_{v}^{u} w_{v,t}\right\}
\end{aligned}$$

where $L_{t|t-1}^{bench}$ is the corresponding likelihood under the benchmark and π is the probability density function of a standard Normal random variable. So the difference in log likelihoods between the worst case and benchmark is

$$\begin{aligned}
\Delta LL_{t|t-1} &\equiv \log \frac{L_{t|t-1}^{wc}}{L_{t|t-1}^{bench}} \\
&= \delta_0 + \delta_1 v_{t-1} + \delta_2 v_{t-1}^{0.5} w_{g,t} + \delta_3 v_{t-1}^{0.5} w_{x,t} + \delta_4 w_{v,t} \\
\delta_0 &\equiv -\frac{\alpha^2 F_2^2 \sigma_v^2}{2} \\
\delta_1 &\equiv -\frac{\alpha^2 \left(1 + F_1^2 \varphi_x^2\right)}{2} \\
\delta_2 &\equiv \alpha \\
\delta_3 &\equiv \alpha F_1 \varphi_x \\
\delta_4 &\equiv \alpha F_2 \sigma_v
\end{aligned}$$

We can therefore construct the conditional (on v_0 and, redundantly, x_0) difference in log likelihoods for a sample for $t = 1 : T$ as

$$\begin{aligned}\Delta LL|v_0, x_0 &\equiv \sum_{t=1}^{T} \Delta LL_{t|t-1} \\ &= \delta_0 T + \delta_1 \sum_{t=1}^{T} v_{t-1} + \delta_2 \sum_{t=1}^{T} v_{t-1}^{0.5} w_{g,t} + \delta_3 \sum_{t=1}^{T} v_{t-1}^{0.5} w_{x,t} + \delta_4 \sum_{t=1}^{T} w_{v,t}\end{aligned}$$

and the unconditional difference in log likelihoods as

$$\Delta LL \equiv \Delta LL|v_0, x_0 + \log \frac{\pi_0^{wc}(v_0, x_0)}{\pi_0^{bench}(v_0, x_0)}$$

where π_0^{wc} and π_0^{bench} are the probability density functions of the unconditional distributions of v_t and x_t under the worst case and benchmark, respectively. We will approximate these with Gaussian distributions matching the first and second moment.

The unconditional distribution of v_t is Normal under the worst case and benchmark, with them only differing in the mean $v_t \backsim N\left(\bar{v}, \frac{\sigma_v^2}{1-\varphi_v^2}\right)$ under the benchmark and $v_t \backsim N\left(\tilde{v}, \frac{\sigma_v^2}{1-\varphi_v^2}\right)$ under the worst case where $\tilde{v} \equiv \bar{v} + \frac{\sigma_v^2 \alpha F_2}{1-\varphi_v}$. The unconditional distribution of x_t is not Normal (due to the presence of stochastic volatility) but we will approximate it with a Normal distribution that matches the unconditional mean and variance of the true distributions, along with the unconditional correlation with x_t.[26]

Appendix B. Rare Disasters Model

Appendix B.1. Benchmark Model

We here lay out the technology and preference specifications of our model featuring LRR but not disasters. The simple baseline model of section 4.1 is clearly a special case.

[26] We could also simulate and estimate t distributions or use a Monte Carlo approximation but the difference should be *de minimis.*

Appendix B.1.1. Consumption Process

The rare disasters setup entails using Gaussian and Poisson-Normal mixture components for consumption growth and dividends (levered consumption):

$$\begin{aligned}
\log g_{t+1} &= G_0 + w_{z,t+1} + \bar{v}^{0.5} w_{g,t+1} \\
\log g_{d,t+1} &= \phi \log g_{t+1} \\
w_{z,t+1} | j_{t+1} &\backsim N\left(j_{t+1}\theta, j_{t+1}\delta^2\right) \\
j_{t+1} | h_t &\backsim \frac{e^{-h_t} h_t^{j_{t+1}}}{j_{t+1}!} \\
h_t &\backsim ARG\left(c_h, \varphi_h, \delta_h\right)
\end{aligned}$$

The conditional cumulant generating functions associated with $w_{g,t+1}$, $w_{z,t+1}$, h_{t+1} and $w_{h,t+1}$ are

$$\begin{aligned}
k_{w_g}(s) &= \frac{s^2}{2} \\
k_{t,w_z}(s) &= h_t \left(e^{s\theta + \frac{(s\delta)^2}{2}} - 1 \right) \\
k_{t,h}(s) &= \varphi_h \frac{s}{1 - sc_h} h_t - \delta_h \log\left(1 - sc_h\right) \\
k_{t,w_h}(s) &= s\varphi_h \left(\frac{1}{1 - sc_h} - 1 \right) h_t - \delta_h \left(sc_h + \log\left(1 - sc_h\right) \right)
\end{aligned}$$

Note that if we set $c_h = \frac{\sigma_h^2}{2}$ and $\delta_h = \left(1 - \varphi_h\right) \frac{\bar{h}}{c_h}$ (which defines $\bar{h}$ and σ_h) then we have $h_t = \left(1 - \varphi_h\right) \bar{h} + \varphi_h h_{t-1} + w_{h,t}$.

Appendix B.1.2. Value Function Approximation

We seek an exponentially affine solution for u_t

$$\begin{aligned}
u_t &= \exp\left\{F_0 + F_h h_t\right\} \\
F_0 &= \beta \left(G_0 + F_0 + \alpha \frac{\bar{v}}{2} - \frac{\delta_h}{\alpha} \log\left(1 - \alpha F_h c_h\right) \right) \\
F_h &= \frac{\beta}{\alpha} \left(e^{\alpha\theta + \frac{(\alpha\delta)^2}{2}} - 1 + \frac{\varphi_h \alpha F_h}{1 - \alpha F_h c_h} \right)
\end{aligned}$$

In the case of F_h we must solve a quadratic

$$\begin{aligned}
0 &= aF_h^2 + bF_h + c \\
a &= \alpha^2 c_h \\
b &= \alpha\left(\beta\left(\varphi_h - c_h\left(e^{\alpha\theta+\frac{(\alpha\delta)^2}{2}} - 1\right)\right) - 1\right) \\
c &= \beta\left(e^{\alpha\theta+\frac{(\alpha\delta)^2}{2}} - 1\right)
\end{aligned}$$

We take the root that tends to the solution in the constant intensity case (in which case $u = \exp\left\{F_0 + F_h\bar{h}\right\}$) as we take $c_h \to 0$, adjusting δ_h to ensure $E\left[h_t\right] = \bar{h}$. Given this solution, we can then obtain F_0.

Appendix B.1.3. Stochastic Discount Factor Approximation

Our framework implies a stochastic discount factor

$$\begin{aligned}
\log\Lambda_{t+1} &= \log\Lambda_{t+1}^r + \log\Lambda_{t+1}^u \\
\log\Lambda_{t+1}^r &= k + \lambda_g^r w_{g,t+1} + \lambda_z^r w_{z,t+1} \\
\log\Lambda_{t+1}^u &= \zeta_t + \lambda_g^u w_{g,t+1} + \lambda_z^u w_{z,t+1} + \lambda_h^u w_{h,t+1} \\
k &= \log\beta - G_0 \\
\zeta_t &= \delta_h \log\left(1 - \alpha F_h c_h\right) + h_t\left(\alpha F_h\varphi_h - \left(e^{\alpha\theta+\frac{(\alpha\delta)^2}{2}} - 1 + \frac{\varphi_h\alpha F_h}{1-\alpha F_h c_h}\right)\right) - \alpha^2\frac{\bar{v}}{2} + \alpha F_h\delta_h c_h \\
\lambda_g^r &= -\bar{v}^{0.5} \\
\lambda_z^r &= -1 \\
\lambda_g^u &= \alpha\bar{v}^{0.5} \\
\lambda_z^u &= \alpha \\
\lambda_h^u &= \alpha F_h
\end{aligned}$$

or, alternatively expressed,

$$\log\Lambda_{t+1} = m_0 + m_h h_t + m_2 w_{g,t+1} + m_3 w_{z,t+1} + m_4 w_{h,t+1}$$

Appendix B.2. Worst Case Distribution

The worst case distribution of $w_{g,t+1}$ as can be surmised from the LRR case in Appendix A is simply $N\left(\alpha\bar{v}^{0.5}, 1\right)$. Here we describe the more complicated cases of $w_{z,t+1}$ and $w_{h,t+1}$.

Appendix B.3. Worst Case Distribution of $w_{z,t+1}$

Under the benchmark we have (suppressing subscripts, λ denotes λ_z^u and w_{t+1} denotes $w_{z,t+1}$)

$$\begin{aligned} k(s) &\equiv \log E_t\left[e^{sw_{t+1}}\right] \\ &= h_t\left(e^{s\theta+\frac{(s\delta)^2}{2}}-1\right) \\ &= h_t\left(f(s)-1\right) \\ f(s) &\equiv e^{g(s)} \\ g(s) &\equiv s\theta+\frac{(s\delta)^2}{2} \end{aligned}$$

whereas under the worst case we have

$$\begin{aligned} \tilde{k}(s) &= \log E_t\left[e^{(s+\lambda)w_{t+1}}\right] \\ &= h_t\left(e^{(s+\lambda)\theta+\frac{((s+\lambda)\delta)^2}{2}}-1\right) \\ &= h_t e^{\lambda\theta+\frac{(\lambda\delta)^2}{2}}\left(e^{s\left(\theta+\lambda\delta^2\right)+\frac{(s\delta)^2}{2}}-e^{-\left(\lambda\theta+\frac{(\lambda\delta)^2}{2}\right)}\right) \\ &= \hat{h}_t\left(e^{s\tilde{\theta}+\frac{(s\delta)^2}{2}}-1\right)+\hat{h}_t\left(1-e^{-\left(\lambda\theta+\frac{(\lambda\delta)^2}{2}\right)}\right) \\ &= \hat{h}_t\left(e^{s\tilde{\theta}+\frac{(s\delta)^2}{2}}-1\right)+tis \\ \hat{h}_t &\equiv h_t e^{\lambda\theta+\frac{(\lambda\delta)^2}{2}} \\ \tilde{\theta} &\equiv \theta+\lambda\delta^2 \end{aligned}$$

Since tis are terms independent of s we see that we retain the same distributional structure as under the benchmark but with adjusted arrival probability and mean conditional on arrival. Thus the worst case can be represented as a distorted Poisson-mixture of normals (this is also noted by Backus, Chernov, and Martin (2011))

$$\begin{aligned} w_{z,t+1}|j_{t+1} &\backsim N\left(j\tilde{\theta}, j\delta^2\right) \\ j_{t+1}|t &\backsim \frac{e^{-\hat{h}_t}\hat{h}_t^{j_{t+1}}}{j_{t+1}!} \end{aligned}$$

Appendix B.4. Worst Case Conditional Moments of $w_{h,t+1}$

Under the worst case, the cgf of $w_{h,t+1}$ is the benchmark cgf evaluated at a point translated by λ_h^u

$$\begin{aligned}
\tilde{k}_{t,w_h}(s) &= k_{t,w_h}(s+\lambda_h^u) \\
&= (s+\lambda_h^u)\varphi_h\left(\frac{1}{1-(s+\lambda_h^u)c_h}-1\right)h_t \\
&\quad -\delta_h\left((s+\lambda_h^u)c_h+\log\left(1-(s+\lambda_h^u)c_h\right)\right)
\end{aligned}$$

Consequently, the worst case conditional mean and variance of $w_{h,t+1}$ are

$$\begin{aligned}
\tilde{E}_t[w_{h,t+1}] &= \tilde{k}_{t,w_h}^{(1)}(0) \\
&= a_0 + a_h h_t \\
\tilde{\sigma}_t^2(w_{h,t+1}) &= \tilde{k}_{t,w_h}^{(2)}(0) \\
&= f_0 + f_h h_t
\end{aligned}$$

where the loadings are given by

$$\begin{aligned}
a_0 &\equiv \frac{\delta_h c_h^2 \lambda_h^u}{1-\lambda_h^u c_h} \\
a_h &\equiv \varphi_h \frac{1-(1-\lambda_h^u c_h)^2}{\left(1-\lambda_h^u c_h\right)^2} \\
f_0 &\equiv \frac{\delta_h c_h^2}{\left(1-\lambda_h^u c_h\right)^2} \\
f_1 &\equiv 2\frac{\varphi_h c_h}{\left(1-\lambda_h^u c_h\right)^2}\left(1+c_h^2\frac{\lambda_h^u}{1-\lambda_h^u c_h}\right)
\end{aligned}$$

Under the worst case we can then think of h_{t+1} as following an adjusted process

$$\begin{aligned}
h_{t+1} &= (1-\tilde{\varphi}_h)\tilde{h} + \tilde{\varphi}_h h_t + \tilde{w}_{h,t+1} \\
\tilde{\varphi}_h &\equiv \varphi_h + a_h \\
&= \frac{\varphi_h}{\left(1-\lambda_h^u c_h\right)^2} \\
\tilde{h} &\equiv \frac{\delta_h c_h + a_0}{1-\tilde{\varphi}_h}
\end{aligned}$$

where $\tilde{w}_{h,t+1}$ is a Martingale difference sequence. Then, under the benchmark and worst case, respectively we have

$$
\begin{aligned}
E_t\left[h_{t+j}\right] &= \varphi_h^j h_t + \bar{h}\left(1-\varphi_h^j\right)\\
\tilde{E}_t\left[h_{t+j}\right] &= \tilde{\varphi}_h^h h_t + \tilde{h}\left(1-\tilde{\varphi}_h^j\right)
\end{aligned}
$$

so that revisions in expectations following 'news' on observing $w_{h,t}$ are

$$
\begin{aligned}
\Delta E_t\left[h_{t+j}\right] &= \varphi_h^j w_{h,t}\\
\Delta \tilde{E}_t\left[h_{t+j}\right] &= \tilde{\varphi}_h^j w_{h,t}
\end{aligned}
$$

Since the conditional expectation of $w_{z,t+j+1}$ given h_{t+j} is θh_{t+j} under the benchmark and $\tilde{\theta}$ $e^{\lambda_z^u\theta+\frac{(\lambda_z^u\delta)^2}{2}}h_{t+j}$ under the worst case, we can use the law of iterated expectations to obtain the news for $\log g_{t+j+1}$ as

$$
\begin{aligned}
\Delta E_t\left[\log g_{t+j}\right] &= \varphi_h^j \theta w_{h,t}\\
&\equiv N\left(j+1\right)w_{h,t}\\
\Delta \tilde{E}_t\left[\log g_{t+j}\right] &= \tilde{\varphi}_h^j \tilde{\theta} e^{\lambda_z^u\theta+\frac{(\lambda_z^u\delta)^2}{2}} w_{h,t}\\
&\equiv \tilde{N}\left(j+1\right)w_{h,t}
\end{aligned}
$$

Appendix B.5. Asset Pricing

We here lay out various derivations of asset pricing objects that we use or which may be of interest.

Appendix B.5.1. Pricing An Equity Claim

For desired calibrations of the time varying intensity jump-augmented model the Campbell-Shiller approximation appeared not to provide an accurate solution.[27] We therefore adopt the alternative approach advocated by Lettau and Wachter (2011) of obtaining the term structure of equity and approximating the price dividend ratio of the market using price dividend ratios of zero coupon equity.

We will take P_t^n to denote the price in t of a claim to the value of the dividend stream in period $t+n$ alone. Defining $p_t^n \equiv \frac{P_t^n}{D_t}$ (the price of the $t+n$ equity claim scaled by the current dividend of

[27]We experienced non-existence of a fixed point for the mean of the log price dividend ratio around which to carry out the Taylor approximation.

the *whole* market), we have

$$p_t^n = E_t\left[\Lambda_{t+1}\frac{D_{t+1}}{D_t}p_{t+1}^{n-1}\right]$$

and we approximate the price dividend ratio of the market by

$$\frac{P_t}{D_t} = \sum_{n=1}^{\infty} p_t^n = \sum_{n=1}^{\infty} \exp\left\{A_n + B_n h_t\right\}$$

We make use of the Euler equation

$$\begin{aligned}
p_t^n &= E_t\left[\chi_{t+1} p_{t+1}^{n-1}\right] \\
\log \chi_{t+1} &= l_0 + l_1 h_t + l_2 w_{g,t+1} + l_3 w_{z,t+1} + l_4 w_{h,t+1} \\
\log p_t^n &= A_n + B_n h_t
\end{aligned}$$

where $A_0 = B_0 = 0$. Thus, χ_{t+1} is our stochastic discount factor multiplied by dividend growth[28]

$$\begin{aligned}
\chi_{t+1} &= \Lambda_{t+1}\frac{D_{t+1}}{D_t} \\
l_0 &= m_0 + G_{d,0} \\
l_1 &= m_1 \\
l_2 &= m_2 + \phi \\
l_3 &= m_3 + \phi \\
l_4 &= m_4
\end{aligned}$$

We obtain recursions for the coefficients $\{A_n, B_n\}_{n=1}^{\infty}$ from the Euler equation

$$\begin{aligned}
A_n &= A_{n-1} + l_0 + B_{n-1}\delta_h c_h + \frac{l_2^2}{2} - \delta_h\left(\left(l_4 + B_{n-1}\right)c_h + \log\left(1 - \left(l_4 + B_{n-1}\right)c_h\right)\right) \\
B_n &= l_1 + B_{n-1}\varphi_h + e^{l_3\theta + \frac{(l_3\delta)^2}{2}} - 1 + \left(l_4 + B_{n-1}\right)\varphi_h\left(\frac{1}{1-\left(l_4 + B_{n-1}\right)c_h} - 1\right)
\end{aligned}$$

We must ask under what conditions the sum in our approximation to the price dividend ratio

[28]For the real zero coupon riskless term structure we would take P_t^n as the price to a riskless claim in $t+n$ and $\chi_{t+1} = \Lambda_{t+1}$.

will converge. Consider a fixed point of the recursion for B_n and call it B. Then we have

$$B = \frac{1}{1-\varphi_h}\left(l_1 + e^{l_3\theta + \frac{(l_3\delta)^2}{2}} - 1 + \varphi_h c_h \frac{(l_4+B)^2}{(1-c_h l_4) - c_h B}\right)$$

so we must solve the quadratic

$$\begin{aligned}
0 &= aB^2 + bB + c \\
a &= \xi_h + c_h \\
b &= 2\xi_h l_4 - c_h \xi_0 - 1 + c_h l_4 \\
c &= \xi_0 \left(1 - c_h l_4\right) + \xi_h l_4^2 \\
\xi_0 &= \frac{l_1 + e^{l_3\theta + \frac{(l_3\delta)^2}{2}} - 1}{1-\varphi_h} \\
\xi_h &= \frac{\varphi_h}{1-\varphi_h} c_h
\end{aligned}$$

where we choose the root that tends to the value achieved in the constant intensity case as $c_h \to 0$ with δ_h adjusted to ensure $E\left[h_t\right] = \bar{h}$. Clearly, we require $|\varphi_j| < 1$ for convergence and also

$$l_0 + B\delta_h c_h + \frac{l_2^2}{2} - \delta_h \left(\left(l_4 + B\right) c_h + \log\left(1 - \left(l_4 + B\right) c_h\right)\right) < 0$$

so that $A_n \to -\infty$. We check these conditions under our calibrations and, assuming they are satisfied, truncate the sum at which the change in the value is below a threshold of tolerance.

Appendix B.5.2. Defaultable Government Debt

As in Wachter (2013) and Barro (2006) we allow for the possibility that in the case of a 'jump' or 'disaster' the government may default on its debt with some probability. Following Wachter, we set the probability equal to 0.4 and, in the case of a disaster, assume that the reduction in the amount repaid, relative to face, is equal to the proportional decline in consumption growth. Thus, we have the price in t of a claim to Y_{t+1}(the payoff of the defaultable bond) as[29]

$$\begin{aligned} P_t^B &= E_t\left[\Lambda_{t+1} Y_{t+1}\right] \\ Y_{t+1} &= 1 + \Psi_{t+1} w_{z,t+1} \\ \Psi_{t+1} &= 1 \; w/p \; q \\ &= 0 \; w/p \; 1-q \end{aligned}$$

Then we have (using independence of Ψ_{t+1})

$$\begin{aligned} P_t^B &= E_t\left[\Lambda_{t+1} Y_{t+1}\right] \\ &= E_t\left[\Lambda_{t+1}\right] + E_t\left[\Lambda_{t+1} \Psi_{t+1} w_{z,t+1}\right] \\ &= P_t^f + E_t\left[\Psi_{t+1}\right] E_t\left[\Lambda_{t+1} w_{z,t+1}\right] \\ &= P_t^f + q \hat{P}_t \end{aligned}$$

where

$$\begin{aligned} \hat{P}_t &= E_t\left[\Lambda_{t+1} w_{z,t+1}\right] \\ &= \exp\left\{m_0 + m_1 h_t\right\} E_t\left[e^{m_2 w_{g,t+1}}\right] E_t\left[e^{m_4 w_{h,t+1}}\right] E_t\left[w_{z,t+1} e^{m_3 w_{z,t+1}}\right] \\ &= \exp\left\{m_0 + m_1 h_t + \frac{m_2^2}{2} + m_4 \varphi_h \left(\frac{1}{1 - m_4 c_h} - 1\right) h_t - \delta_h \left(m_4 c_h + \log\left(1 - m_4 c_h\right)\right)\right\} \\ &\quad \times E_t\left[w_{z,t+1} \exp\left\{m_3 w_{z,t+1}\right\}\right] \end{aligned}$$

On consulting *Mathematica* we have

$$E_t\left[w_{z,t+1} \exp\left\{m_3 w_{z,t+1}\right\}\right] = h_t\left(\theta + m_3 \delta^2\right) \exp\left\{h_t\left(e^{m_3\theta + \frac{(m_3\delta)^2}{2}} - 1\right) + m_3\theta + \frac{(m_3\delta)^2}{2}\right\}$$

[29]Note that Δ_{t+1} is always being realized but is only economically relevant when $w_{z,t+1} \neq 0$ which will only be the case (almost surely) when a jump has occurred and it is drawn from a non-degenerate, non-zero mean Normal distribution.

so that

$$
\begin{aligned}
\hat{P}_t &= P_t^f \xi(h_t) \\
P_t^f &= \exp\left\{ \begin{array}{c} m_0 + \frac{m_2^2}{2} - \delta_h \left(m_4 c_h + \log\left(1 - m_4 c_h\right)\right) \\ + h_t \left(m_4 \varphi_h \left(\frac{1}{1 - m_4 c_h} - 1 \right) + m_1 + e^{m_3 \theta + \frac{(m_3 \delta)^2}{2}} - 1 \right) \end{array} \right\} \\
\xi(h_t) &= \left(\theta + m_3 \delta^2\right) h_t e^{m_3 \theta + \frac{(m_3 \delta)^2}{2}}
\end{aligned}
$$

and, thus,

$$
P_t^B = P_t^f \left(1 + q \xi(h_t)\right)
$$

We then have the face and expected returns on the defaultable bond as

$$
\begin{aligned}
R_{t+1}^{B,face} &\equiv \frac{1}{P_t^B} \\
E_t\left[R_{t+1}^B\right] &= E_t\left[\frac{Y_{t+1}}{P_t^B}\right] \\
&= R_{t+1}^{B,face} + \frac{q \theta h_t}{P_t^B}
\end{aligned}
$$

Appendix B.6. Detection Error Probabilities

We have

$$
\log \Lambda_{t+1}^u = m_0 + m_1 h_t + m_2 w_{g,t+1} + m_3 w_{z,t+1} + m_4 w_{h,t+1}
$$

implying that the conditional (on h_0) log likelihood difference between the benchmark and worst case models, given $\{h_t, w_{g,t}, w_{z,t}, w_{h,t}\}_{t=1}^T$ is

$$
\Delta LL|_{h_0} = m_0 T + m_1 \sum_{t=1}^T h_{t-1} + m_2 \sum_{t=1}^T w_{g,t} + m_3 \sum_{t=1}^T w_{z,t} + m_4 \sum_{t=1}^T w_{h,t}
$$

so that the unconditional log likelihood difference is

$$
\Delta LL = \Delta LL|_{h_0} + \log \frac{\Pi_h^{wc,erg}(h_0)}{\Pi_h^{bench,erg}(h_0)}
$$

where $\Pi_h^{bench,erg}(h_0)$ is the pdf of a $\Gamma\left(\delta_h, \frac{c_h}{1-\varphi_h}\right)$ random variable (ergodic distribution of h_t under the benchmark) and $\Pi_h^{wc,erg}(h_0)$ is the pdf of a $\Gamma(\eta_1, \eta_2)$ where we obtain η_1 and η_2 by fitting a Gamma distribution to observations from a long simulation of h_t under the worst case.