Consumption-Based Asset Pricing with Higher Cumulants

Ian Martin*

6 November, 2007

Abstract

I extend the Epstein-Zin-lognormal consumption-based asset-pricing model to allow for general i.i.d. consumption growth processes. Information about the higher moments—equivalently, cumulants—of consumption growth is encoded in the *cumulant-generating function* (CGF). I express four observable quantities (the equity premium, riskless rate, consumption-wealth ratio and mean consumption growth) and the Hansen-Jagannathan bound in terms of the CGF, and present applications. Models in which consumption is subject to occasional disasters can be handled easily and flexibly within the framework. The importance of higher cumulants is a double-edged sword: those model parameters which are most important for asset prices, such as disaster parameters, are also the hardest to calibrate. It is therefore desirable to make statements which do not depend on a particular calibrated consumption process. First, I use properties of the CGF to derive restrictions on the time-preference rate and elasticity of intertemporal substitution that must hold in any Epstein-Zin-i.i.d. model which is consistent with the observable quantities. Second, I show that "good deal" bounds on the maximal Sharpe ratio can be used to derive restrictions on preference parameters without calibrating the consumption process. Third, given preference parameters, I calculate the welfare cost of uncertainty directly from mean consumption growth and the consumption-wealth ratio without having to estimate the amount of risk in the economy. Fourth, I analyze heterogeneous-agent models with jumps.

*iwmartin@fas.harvard.edu; http://www.people.fas.harvard.edu/∼iwmartin/. First draft: 20 August, 2006. I thank Robert Barro, Emmanuel Farhi, Xavier Gabaix, Simon Gilchrist, Francois Gourio, Greg Mankiw, Anthony Niblett, Adrien Verdelhan, Martin Weitzman and, in particular, John Campbell for their comments.

The combination of power utility and i.i.d. lognormal consumption growth makes for a tractable benchmark model in which asset prices and expected returns can be found in closed form. Introducing the consumption-based model, Cochrane (2005, p. 12) writes, "The combination of lognormal distributions and power utility is one of the basic tricks to getting analytical solutions in this kind of model." A message of the present paper is that the lognormality assumption can be relaxed without sacrificing tractability.

Following Barro's (2006a) rehabilitation of Rietz (1988), the ability to generalize beyond the lognormal assumption is evidently desirable. Working under two assumptions—that there is a representative agent with Epstein-Zin preferences[1] and that consumption growth is i.i.d.—I introduce, in Section 1, a mathematical object (the cumulant-generating function) in terms of which four fundamental quantities which are at the heart of consumption-based asset pricing can be simply expressed. Those fundamental quantities, or fundamentals for short, are the equity premium, riskless rate, consumption-wealth ratio[2] and mean consumption growth.

The expressions derived relate the fundamentals directly to the cumulants (equivalently, moments) of consumption growth, and show that familiar concepts such as precautionary saving can be generalized in the presence of higher cumulants. The lognormal assumption is equivalent to the assumption that all cumulants above the second are zero; hence the title of the paper.

The first few cumulants of consumption growth can in principle be estimated from consumption data, though this approach is not taken in the present paper because, given the sizes of the relevant samples in practice, estimates of higher cumulants (or moments) have large standard errors. This is especially troubling because the higher cumulants which are hardest to estimate are extremely influential for asset prices.

In Section 2, I show that these results carry over to a continuous-time setting. If one is in the business of making up stochastic processes, many suggest themselves most naturally in continuous time. Although there is an obvious discrete-time analogue of Brownian motion—a random walk with Normally distributed increments—it is less natural to map Poisson processes, say, into discrete time, and therefore harder to deal with the possibility of jumps in consumption.[3] The i.i.d. growth assumption is replaced by its continuous-time analogue:

[1] Epstein-Zin preferences nest the power utility case. Kocherlakota (1990) demonstrates that when consumption growth is i.i.d., Epstein-Zin preferences and power utility are observationally equivalent. For the sake of intuition, though, it is helpful to use Epstein-Zin preferences in order to distinguish clearly between the effects of risk aversion, intertemporal elasticity of substitution, and time discount rate.

[2] Or, depending on one's preferred interpretation, the dividend-price ratio on the Lucas tree.

[3] According to Kingman (1993), "In the theory of random processes there are two that are fundamental,

log consumption is a Lévy process. I specialize to power utility for simplicity.

I illustrate the CGF framework by investigating a continuous-time model featuring rare disasters in the style of Rietz (1988) or Barro (2006a). By working in continuous time, simple expressions are obtained without the need for Taylor series approximations. The model's predictions are sensitively dependent on the calibration assumed.

As a stark illustration, take a consumption-based model in which the representative agent has relative risk aversion equal to 4. Now imagine adding to the model a certain type of disaster which strikes, on average, once every 100,000 years. When the disaster strikes, it destroys 90 per cent of wealth. (Barro (2006a) documents that Germany and Greece each suffered a 64 per cent fall in per capita real GDP in the course of the Second World War, so such a disaster is not beyond the bounds of possibility.) The introduction of the very rare, very severe disaster will drive the riskless rate down by 10 percentage points—1000 basis points—and will increase the equity premium by 9 per cent.[4] Very rare, very severe events exert an extraordinary influence on the benchmark model, and we do not expect to estimate their frequency and intensity directly from the data.

We can, however, detect the influence of disaster events *indirectly*, by observing asset prices. I argue, therefore, that the standard approach—calibrating a particular model and trying to fit the fundamental quantities—is not the way to go. By turning things round—viewing the fundamental quantities as observable and seeing what they imply—it becomes possible to make statements which are robust to the details of the consumption growth process.

My first application, presented in Section 3, exploits the fact that cumulant-generating functions are convex. I derive robust restrictions on preference parameters which are valid in *any* Epstein-Zin-i.i.d. model which is consistent with the observed fundamentals. My results restrict the time-preference rate, ρ, and elasticity of intertemporal substitution, ψ, to lie in a certain subset of the positive quadrant. (See Figure 4.) These restrictions depend only on the Epstein-Zin-i.i.d. assumptions and on observed values of the fundamentals. They are complementary to econometric or experimental estimates of ψ and ρ, and are of particular interest because there is little agreement about the value of ψ. (Campbell

and occur over and over again, often in surprising ways. There is a real sense in which the deepest results are concerned with their interplay. One, the Bachelier-Wiener model of Brownian motion, has been the subject of many books. The other, the Poisson process, seems at first sight humbler and less worthy of study in its own right .... This comparative neglect is ill judged, and stems from a lack of perception of the real importance of the Poisson process."

[4]I illustrate this point with more reasonable numbers in section 2.2 below, in which I consider the effect of perturbing parameters in a continuous-time disaster model.

(2003) summarizes the conflicting evidence.) I also show how good-deal bounds (Cochrane and Saá-Requejo (2000)) can be used to provide upper bounds on risk aversion without calibrating a consumption process.

The theme of making inferences from observable fundamentals recurs in Section 4, which takes up the question, surveyed by Lucas (2003), of the cost of consumption risk. This cost turns out to depend on ρ and ψ and on two observables: mean consumption growth and the consumption-wealth ratio. The cost does not depend on risk aversion other than through the consumption-wealth ratio, which summarizes all relevant information about the attitude to risk of the representative agent and the amount of risk in the economy, as captured by the cumulants.

In the power utility subcase of Epstein-Zin, the welfare calculations apply more generally to *any* consumption growth process, i.i.d. or not. These results therefore generalize Lucas (1987), Obstfeld (1994) and Barro (2006b). Unlike these authors, I use the consumption-wealth ratio as an observable. Using Barro's preferred preference parameters, I find that the cost of consumption fluctuations is about 14 per cent. I also calculate the welfare gains from a reduction in the variance of consumption growth, and show that the representative agent would sacrifice on the order of one per cent of initial consumption to reduce the standard deviation of consumption growth from 2% to 1%.

Finally, in Section 5, I exhibit the convenience of the CGF approach in a heterogeneous agent model with jumps. The model is intended to resolve the tension between the results of Grossman and Shiller (1982), who show that heterogeneity is irrelevant in continuous time if consumption processes follow diffusions, and those of Constantinides and Duffie (1996), who show that heterogeneity is important in discrete time. I show that in continuous-time i.i.d. models, heterogeneity matters to the extent that it is present at times of aggregate jumps. Jumps lend a discrete-time flavor to the model, which in a sense occupies a position intermediate between Grossman-Shiller and Constantinides-Duffie.

There is a large body of literature that applies Lévy processes to derivative pricing (Carr and Madan (1998), Cont and Tankov (2004)) and, more recently, portfolio choice (Kallsen (2000), Cvitanić, Polimenis and Zapatero (2005), Aït-Sahalia, Cacho-Diaz and Hurd (2006)). Backus, Foresi and Telmer (2001), Shaliastovich and Tauchen (2005), and Lentzas (2007) derive expressions that relate cumulants to risk premia, though the philosophy of these papers is very different from the calibration-free approach taken here.

1 Asset-pricing fundamentals and the CGF

Define $G_t \equiv \log C_t/C_0$ and write $G \equiv G_1$. I make two assumptions.

A1 There is a representative agent whose Epstein-Zin preferences have relative risk aversion γ and elasticity of intertemporal substitution ψ.

A2 The consumption growth, $\log C_t/C_{t-1}$, of the representative agent is i.i.d., and the moment-generating function of G (defined below) exists on the interval $[-\gamma, 1]$.[5]

Assumption A1 allows risk aversion γ to be disentangled from the elasticity of intertemporal substitution ψ. To keep things simple, those calculations that appear in the main text restrict to the power utility case in which ψ is constrained to equal $1/\gamma$; in this case, the representative agent maximizes

$$\mathbb{E}\sum_{t=0}^{\infty} e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} \quad \text{if } \gamma \neq 1\,, \quad \text{or} \quad \mathbb{E}\sum_{t=0}^{\infty} e^{-\rho t}\log C_t \quad \text{if } \gamma = 1\,. \tag{1}$$

Results for the more general Epstein-Zin case are reported and discussed in the main text, but calculations and proofs are relegated to Appendix B.

Assumption A2 is strong, and it is essential for the calculations of this paper. Cogley (1990) and Barro (2006b) present evidence in support of A2 in the form of variance-ratio statistics close to one, on average, across nine (Cogley) or 19 (Barro) countries.

For the time being, I restrict to power utility. We need expected utility to be well defined in that

$$\mathbb{E}\sum_{t=0}^{\infty}\left| e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma}\right| < \infty \quad \text{if } \gamma \neq 1. \tag{2}$$

I discuss this requirement further below.

The Euler equation of asset pricing relates the price of an asset this period to the payoff next period:

$$P_0 = \mathbb{E}_0\left(e^{-\rho}\left(\frac{C_1}{C_0}\right)^{-\gamma}(D_1 + P_1)\right)\,.$$

Iterating forward, we get

$$P_0 = \mathbb{E}\left(\sum_{t=1}^{T} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma} D_t\right) + \mathbb{E}_0 e^{-\rho T}\left(\frac{C_T}{C_0}\right)^{-\gamma} P_T\,.$$

[5] If this is not so, the consumption-based asset-pricing approach is invalid. This assumption ensures that all moments of G are finite. See Billingsley (1995, Section 21).

Finally, allowing $T \to \infty$ (and imposing the no-bubble condition that the second term in the above expression tends to zero in the limit) leads to the familiar equation

$$P(D) = \mathbb{E}\left(\sum_{t=1}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma} D_t\right). \tag{3}$$

I start by considering an asset which pays dividend stream $D_t \equiv (C_t)^\lambda$ for some constant λ (the λ-asset). The central cases of interest will later be $\lambda = 0$ (the *riskless bond*) and $\lambda = 1$ (the *wealth portfolio* which pays consumption as its dividend), but, as in Campbell (1986) and Abel (1999), it is possible to view values $\lambda > 1$ as a tractable way of approximating levered equity claims. I write P_λ for the price of this asset at time 0, and D_λ for the dividend at time 0.

From (3),

$$\begin{aligned}
P_\lambda &= \mathbb{E}\left(\sum_{t=1}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma} (C_t)^\lambda\right) \\
&= (C_0)^\lambda \sum_{t=1}^{\infty} e^{-\rho t}\mathbb{E}\left(\left(\frac{C_t}{C_0}\right)^{\lambda-\gamma}\right) \\
&= D_\lambda \sum_{t=1}^{\infty} e^{-\rho t}\mathbb{E}\left(e^{(\lambda-\gamma)G_t}\right) \\
&= D_\lambda \sum_{t=1}^{\infty} e^{-\rho t}\left(\mathbb{E}\left(e^{(\lambda-\gamma)G}\right)\right)^t. \qquad (4)
\end{aligned}$$

The last equality follows from the assumption that log consumption growth is i.i.d. To make further progress, I now introduce a pair of definitions.

Definition 1. *Given some arbitrary random variable, G, the* moment-generating function $\boldsymbol{m}(\theta)$ *and* cumulant-generating function *or* CGF $\boldsymbol{c}(\theta)$ *are defined by*

$$\begin{aligned}
\boldsymbol{m}(\theta) &\equiv \mathbb{E}\exp(\theta G) \qquad (5)\\
\boldsymbol{c}(\theta) &\equiv \log \boldsymbol{m}(\theta), \qquad (6)
\end{aligned}$$

for all θ for which the expectation in (5) is finite.

In the particular application of this paper, G is, of course, to be viewed as an annual increment of log consumption, $G = \log C_{t+1} - \log C_t$.

Notice that $\boldsymbol{c}(0) = 0$ for any growth process and that $\boldsymbol{c}(1)$ is equal to log mean gross consumption growth—so in practice we will want to ensure that $\boldsymbol{c}(1) \approx 2\%$.

I expand further on the CGF in Appendix A; for now, it can be thought of as capturing information about all moments of G. More precisely, we can expand $\boldsymbol{c}(\theta)$ as a power series in θ,

$$\boldsymbol{c}(\theta) = \sum_{n=1}^{\infty} \frac{\kappa_n \theta^n}{n!},$$

and define κ_n to be the nth *cumulant* of log consumption growth. A small amount of algebra confirms that, for example, $\kappa_1 \equiv \mu$ is the mean, $\kappa_2 \equiv \sigma^2$ the variance, κ_3/σ^3 the skewness and κ_4/σ^4 the kurtosis of log consumption growth. Knowledge of the cumulants of a random variable implies knowledge of the moments, and vice versa.

With this definition, (4) becomes

$$\begin{aligned} P_\lambda &= D_\lambda \sum_{t=1}^{\infty} e^{-[\rho - \boldsymbol{c}(\lambda-\gamma)]t} \\ &= D_\lambda \cdot \frac{e^{-[\rho-\boldsymbol{c}(\lambda-\gamma)]}}{1 - e^{-[\rho-\boldsymbol{c}(\lambda-\gamma)]}}, \end{aligned}$$

or,

$$\frac{D_\lambda}{P_\lambda} = e^{\rho - \boldsymbol{c}(\lambda-\gamma)} - 1$$

It is convenient to define the log dividend yield $d_\lambda/p_\lambda \equiv \log(1 + D_\lambda/P_\lambda)$.[6] Then,

$$d_\lambda/p_\lambda = \rho - \boldsymbol{c}(\lambda - \gamma) \tag{7}$$

Two special cases are of particular interest. The first is $\lambda = 0$, in which case the asset in question is the riskless bond, whose dividend yield is the riskless rate. The second is $\lambda = 1$, in which case the asset pays consumption as its dividend, and can therefore be interpreted as aggregate wealth. The dividend yield is then the consumption-wealth ratio.

This calculation also shows that the necessary restriction on consumption growth for the expected utility to be well defined in (2) is that $\rho > \boldsymbol{c}(1-\gamma)$, or equivalently that the consumption-wealth ratio is positive. When the condition fails, the standard consumption-based asset pricing approach is no longer valid.

The gross return on the λ-asset is (dropping λ subscripts for clarity)

$$\begin{aligned} 1 + R_{t+1} &= \frac{D_{t+1} + P_{t+1}}{P_t} \\ &= \frac{P_{t+1}}{P_t}\left(1 + \frac{D_{t+1}}{P_{t+1}}\right) \\ &= \frac{D_{t+1}}{D_t}\left(e^{\rho - \boldsymbol{c}(\lambda-\gamma)}\right) \end{aligned} \tag{8}$$

[6]It is worth emphasizing that log dividend yield, as I have defined it, is a number close to D/P, since $\log(1+x) \approx x$ for small x. d/p is *not* the same as $d - p$ as used elsewhere in the literature to mean $\log D/P$.

and thus the expected gross return is

$$\begin{aligned} 1+\mathbb{E}R_{t+1} &= \mathbb{E}\left(\left(\frac{C_{t+1}}{C_t}\right)^{\lambda}\right)\cdot e^{\rho-\boldsymbol{c}(\lambda-\gamma)} \\ &= \mathbb{E}\left(e^{G\lambda}\right)\cdot e^{\rho-\boldsymbol{c}(\lambda-\gamma)} \\ &= e^{\rho-\boldsymbol{c}(\lambda-\gamma)+\boldsymbol{c}(\lambda)} \end{aligned}$$

Once again, it turns out to be more convenient to work with log expected gross return, $er_\lambda \equiv \log(1+\mathbb{E}R_{t+1}) = \rho + \boldsymbol{c}(\lambda) - \boldsymbol{c}(\lambda-\gamma)$.

The above calculations are summarized in

Proposition 1 (Fundamental quantities, power utility case)**.** *The riskless rate, $r_f \equiv \log(1+R_f)$, consumption-wealth ratio, $c/w \equiv \log(1+C/W)$, and risk premium on aggregate wealth, $rp \equiv er_1 - r_f$, are given by*

$$r_f = \rho - \boldsymbol{c}(-\gamma) \tag{9}$$

$$c/w = \rho - \boldsymbol{c}(1-\gamma) \tag{10}$$

$$rp = \boldsymbol{c}(1) + \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\,. \tag{11}$$

Writing these quantities explicitly in terms of the underlying cumulants by expanding $\boldsymbol{c}(\theta)$ in power series form, we obtain

$$r_f = \rho - \sum_{n=1}^{\infty} \frac{\kappa_n(-\gamma)^n}{n!} \tag{12}$$

$$c/w = \rho - \sum_{n=1}^{\infty} \frac{\kappa_n(1-\gamma)^n}{n!} \tag{13}$$

$$rp = \sum_{n=2}^{\infty} \frac{\kappa_n}{n!}\cdot\Big\{1+(-\gamma)^n-(1-\gamma)^n\Big\}. \tag{14}$$

Writing the first few terms of the series out more explicitly, (12) implies that

$$r_f = \rho + \kappa_1\gamma - \frac{\kappa_2}{2}\gamma^2 + \frac{\kappa_3}{3!}\gamma^3 - \frac{\kappa_4}{4!}\gamma^4 + \text{higher order terms}\,.$$

By definition of the first four cumulants, this can be rewritten as

$$r_f = \rho + \mu\gamma - \frac{1}{2}\sigma^2\gamma^2 + \frac{\text{skewness}}{3!}\sigma^3\gamma^3 - \frac{\text{excess kurtosis}}{4!}\sigma^4\gamma^4 + \text{higher order terms}\,. \tag{15}$$

In the lognormal case, the skewness, excess kurtosis and all higher cumulants are zero, so (15) reduces to the familiar $r_f = \rho + \mu\gamma - \sigma^2\gamma^2/2$. More generally, the riskless rate is low if mean log consumption growth μ is low (an intertemporal substitution effect); if the variance of log consumption growth σ^2 is high (a precautionary savings effect); if there is negative skewness; or if there is a high degree of kurtosis.

Similarly, the consumption-wealth ratio (13) can be rewritten as

$$\begin{aligned} c/w &= \rho + \mu(\gamma-1) - \frac{1}{2}\sigma^2(\gamma-1)^2 + \frac{\text{skewness}}{3!}\sigma^3(\gamma-1)^3 - \\ &\quad - \frac{\text{excess kurtosis}}{4!}\sigma^4(\gamma-1)^4 + \text{higher order terms}\,. \end{aligned} \tag{16}$$

The log utility case, $\gamma = 1$, is evidently a special case, in which the consumption-wealth ratio is determined only by the rate of time preference: $c/w = \rho$. If $\gamma \neq 1$, the consumption-wealth ratio is low when cumulants of even order are large (high variance, high kurtosis, and so on). The importance of cumulants of odd order depends on whether γ is greater or less than 1. In the empirically more plausible case $\gamma > 1$, the consumption-wealth ratio is low when odd cumulants are low: when mean log consumption growth is low, or when there is negative skewness, for example. If the representative agent is more risk-tolerant than log, the reverse is true: the consumption-wealth ratio is high when mean log consumption growth is low, or when there is negative skewness.

The risk premium (14) becomes

$$\begin{aligned} rp &= \gamma\sigma^2 + \frac{\text{skewness}}{3!}\sigma^3\left(1-\gamma^3-(1-\gamma)^3\right) + \\ &\quad + \frac{\text{excess kurtosis}}{4!}\sigma^4\left(1+\gamma^4-(1-\gamma)^4\right) + \text{higher order terms}\,. \end{aligned} \tag{17}$$

In the lognormal case, this is just $rp = \gamma\sigma^2$. Since $1 + \gamma^n - (1-\gamma)^n > 0$ for even n, the risk premium is increasing in variance, excess kurtosis and higher cumulants of even order. The effect of skewness and higher cumulants of odd order depends on γ. For odd n, $1 - \gamma^n - (1-\gamma)^n$ is positive if $\gamma < 1$, zero if $\gamma = 1$, and negative if $\gamma > 1$. If $\gamma = 1$, skewness and higher odd-order cumulants have no effect on the risk premium. Otherwise, the risk premium is decreasing in skewness and higher odd cumulants if $\gamma > 1$ and increasing if $\gamma < 1$.

The following result generalizes Proposition 1 to allow for Epstein-Zin preferences.

Proposition 2 (Fundamental quantities, Epstein-Zin case)**.** *Defining* $\vartheta \equiv (1-\gamma)/(1-1/\psi)$, *we have*

$$r_f = \rho - \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\left(\frac{1}{\vartheta} - 1\right) \tag{18}$$

$$c/w = \rho - \boldsymbol{c}(1-\gamma)/\vartheta \tag{19}$$

$$rp = \boldsymbol{c}(1) + \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\,, \tag{20}$$

and the obvious counterparts of (12)–(14) *which result on expanding the CGFs in* (18)–(20) *as power series.*

Proof. See Appendix B. □

Equation (20) shows as expected that when the CGF is linear—that is, when consumption growth is deterministic—there is no risk premium. Roughly speaking, the CGF of the driving consumption process must have a significant amount of convexity over the range $[-\gamma, 1]$ to generate an empirically reasonable risk premium. It also confirms that risk aversion alone influences the risk premium: the elasticity of intertemporal substitution is not a factor.

An interesting feature of Propositions 1 and 2 is that expressions (12)–(14), and their analogues in the Epstein-Zin case, can in principle be estimated directly by estimating the cumulants of log consumption, given a sufficiently long data sample, without imposing any further structure on the model. If, say, the high equity premium results from the occasional occurrence of severe disasters, this will show up in the cumulants. No particular assumption—beyond (A1) and (A2)—need be made about the arrival rate or distribution of disasters, nor of any other feature of the consumption process. In practice, of course, we cannot estimate infinitely many cumulants from a finite data set. One solution to this is to impose some particular distribution on log consumption growth, and then to estimate the parameters of the distribution.

An alternative approach, more in the spirit of model-independence, is to approximate the equations by truncating after the first N cumulants, N being determined by the amount of data available. (In this context it is worth noting that the assumption that consumption growth is lognormal is equivalent to truncating at $N = 2$, since, as noted above, when log consumption growth is Normal all cumulants above the variance are equal to zero—that is, $\kappa_n = 0$ for n greater than 2.) Nonetheless, for the reasons stated in the Introduction, I do not follow this route.

1.1 The Gordon growth model

From equations (18)–(20), we see that

$$c/w = rp + r_f - \boldsymbol{c}(1) \tag{21}$$

or, more generally, that

$$d_\lambda/p_\lambda = er_\lambda - \boldsymbol{c}(\lambda)\,. \tag{22}$$

This is a version of the traditional Gordon growth model. (For example, the last term of (21), $\boldsymbol{c}(1) = \log \mathbb{E} C_{t+1}/C_t$, measures mean consumption growth.)

The connection is even more explicit in *levels* rather than logs. To see this, note that $\mathbb{E}_t R_{t+1} \equiv R$ is constant, and write $1+\Gamma$ for the gross growth rate of consumption, $\mathbb{E}_t C_{t+1}/C_t$. Taking expectations of (8) and imposing a no-bubbles condition, we get

$$
\begin{aligned}
P_t &= \mathbb{E}_t \left(\frac{D_{t+1} + P_{t+1}}{1+R} \right) \\
&= \mathbb{E}_t \sum_{k=1}^{\infty} \frac{D_{t+k}}{(1+R)^k} \\
&= D_t \cdot \sum_{k=1}^{\infty} \left(\frac{1+\Gamma}{1+R} \right)^k \\
&= \frac{D_t(1+\Gamma)}{R-\Gamma}
\end{aligned}
$$

This can be expressed as

$$D_t/P_t = (R-\Gamma)/(1+\Gamma) \tag{23}$$

or, in classic Gordon growth model terms,

$$\frac{\mathbb{E}_t D_{t+1}}{P_t} = R - \Gamma. \tag{24}$$

To recover (22) from (23), apply $\log(1+\cdot)$ to both sides, and note that $\log(1+\Gamma) = \boldsymbol{c}(\lambda)$.

Since the Gordon growth model holds in this framework, only three of the riskless rate, risk premium, consumption-wealth ratio and mean consumption growth can be independently specified: the fourth is then mechanically determined by (21).

This observation, in conjunction with equations (18)–(20), provides another way to look at Kocherlakota's (1990) point. In principle, given sufficient asset price and consumption data, we could determine the riskless rate, the risk premium, and CGF $\boldsymbol{c}(\cdot)$ to any desired level of accuracy. (In view of (21), the consumption-wealth ratio would contain no extra information.) Since γ is the only preference parameter that determines the risk premium, it could be calculated from (20), given knowledge of $\boldsymbol{c}(\cdot)$. On the other hand, knowledge of the riskless rate leaves ρ and ψ indeterminate in equation (18), even given knowledge of γ and $\boldsymbol{c}(\cdot)$. That is, the time discount rate and elasticity of intertemporal substitution cannot be disentangled. On the other hand, as noted in footnote 1, the use of Epstein-Zin preferences aids the interpretation of results.

1.2 The asymptotic lognormality of consumption

If G has mean μ and (finite) variance σ^2, the central limit theorem shows that consumption is asymptotically lognormal:[7] as $t \to \infty$

$$\frac{G_t - \mu t}{\sqrt{t}} \xrightarrow{d} N(0, \sigma^2).$$

It therefore appears that if one measures over very long periods, only the first two cumulants will be needed to capture information about consumption growth. Why, then, does the representative agent care about cumulants of log consumption growth other than mean and variance? To answer this question, it is helpful to define the scale-free cumulants

$$SFC_n \equiv \frac{\kappa_n}{\sigma^n}$$

For example, SFC_3 is skewness and SFC_4 is kurtosis. These scale-free cumulants are normalized to be invariant if the underlying random variable is scaled by some constant factor. Since the (unscaled) cumulants of G_t are linear in t, the nth scale-free cumulant of G_t is proportional to $t \cdot t^{-n/2} = t^{(2-n)/2}$ and so tends to zero for n greater than 2. The asymptotic Normality of $(G_t - \mu t)/\sqrt{t}$ is reflected in the fact that its scale-free cumulants of orders greater than two tend to zero as t tends to infinity. But in terms of the scale-free cumulants, the riskless rate (for example) can be expressed as

$$\begin{aligned} r_f &= \rho - \sum_{n=1}^{\infty} \frac{\kappa_n (-\gamma)^n}{n!} \\ &= \rho - \sum_{n=1}^{\infty} \frac{SFC_n \sigma^n (-\gamma)^n}{n!} \end{aligned} \tag{25}$$

Thus, even though skewness, kurtosis and higher scale-free cumulants tend to zero as the period length is allowed to increase, the relevant asset-pricing equation scales these variables by σ—and this tends to infinity as period length increases, in such a way that higher cumulants remain relevant.

[7] Informally, $G_t - \mu t$ is typically $O(\sqrt{t})$, so for positive α, $\mathbb{P}(G_t - \mu t \geq \alpha t) \to 0$ as $t \to \infty$, or equivalently, $\mathbb{P}(C_t \geq C_0 e^{(\mu+\alpha)t}) \to 0$. The Cramér-Chernoff theorem tells us how fast this probability decays to zero, and provides an opportunity to mention another context in which the CGF arises. It implies that

$$\frac{1}{t} \log \mathbb{P}\left(C_t \geq C_0 e^{\alpha t}\right) \longrightarrow \inf_{\theta \geq 0} \boldsymbol{c}(\theta) - \alpha\theta$$

and

$$\frac{1}{t} \log \mathbb{P}\left(C_t \leq C_0 e^{\alpha t}\right) \longrightarrow \inf_{\theta \leq 0} \boldsymbol{c}(\theta) - \alpha\theta.$$

Van der Vaart (1998) has a proof.

2 The continuous-time case

For the purposes of constructing concrete examples, it is convenient to confirm that the simplicity of the above framework carries over to the continuous-time case.

Assumptions A1 and A2 are modified slightly. They become

A1c There is a representative agent with constant relative risk aversion γ, who therefore maximizes[8]

$$\mathbb{E} \int_{t=0}^{\infty} e^{-\rho t} \frac{C_t^{1-\gamma}}{1-\gamma} \quad \text{if } \gamma \neq 1\,, \quad \text{or} \quad \mathbb{E} \int_{t=0}^{\infty} e^{-\rho t} \log C_t \quad \text{if } \gamma = 1 \tag{26}$$

A2c The log consumption path, G_t, of the representative agent follows a Lévy process (defined in Appendix C), and $\boldsymbol{m}(\theta)$ exists for θ in $[-\gamma, 1]$.

As before, we need a condition that ensures finiteness of (26); as before, the pricing calculation, below, yields the required condition.

The analysis is almost identical to that in the discrete-time case; all that is needed is that an equality of the form

$$\mathbb{E} e^{\theta G_t} = \left(\mathbb{E} e^{\theta G} \right)^t \tag{27}$$

holds, where G_t is now a continuous-time process. In the discrete time case, this was an obvious consequence of the facts that

$$G_t = \log C_1/C_0 + \log C_2/C_1 + \cdots + \log C_t/C_{t-1}$$

and that each of the terms $\log C_i/C_{i-1}$ was assumed i.i.d. with the same law as G. In continuous time, (27) follows from Assumption A2c; see Sato (1999) for a proof.

The assumption that $\boldsymbol{m}(\theta)$ exists over the appropriate interval has bite, for example, in the case of Mandelbrot's stable processes: since stable processes (other than Brownian motion) do not have well-defined moments, I am excluding them from consideration.

Example 1 Brownian motion with drift, $L_t = ct + \sigma_B B_t$. These are the only continuous Lévy processes.

Example 2 The Poisson counting process, $L_t = N_t$: N_t counts the number of jumps that have taken place by time t and is distributed according to a Poisson distribution with parameter ωt for some $\omega > 0$.

[8]For simplicity, I restrict to the power utility case, although it should be clear that the analysis can be easily generalized to allow for the continuous-time analogue of Epstein-Zin preferences (Duffie and Epstein (1992)).

Example 3 A compound Poisson process, $L_t = \sum_{i=1}^{N_t} Y_i$, where the random variables Y_i are i.i.d. Example 2 is the special case in which $Y_1 \equiv 1$.

Example 4 There exist Lévy processes with L_1 distributed according to any of the following distributions (amongst others): the t-distribution, the Cauchy distribution, the Pareto distribution, the F-distribution, the gamma distribution. Only in the last case does the moment-generating function exist for some $\theta > 0$, and thus only in the last case can the techniques of standard consumption-based asset pricing be brought to bear. (See Weitzman (2005).)

Example 5 The α-stable Paretian processes advocated by Mandelbrot (1963, 1967) are Lévy processes with the additional property that for any constant $c > 0$, the law of $\{L_{ct}\}_{t \geq 0}$ is the same as the law of $\{c^{1/\alpha} L_t\}_{t \geq 0}$; $\alpha \in (0, 2]$ is the *index* of the process. Loosely speaking, the sample paths of such a process look similar as one "zooms in" on them. The case $\alpha = 2$ gives Brownian motion; this is the only α-stable process with finite variance.

Example 6 The time-change of one Lévy process with another independent increasing Lévy process; that is, $L_t = P_{Q_t}$ is a Lévy process if P is a Lévy process and Q is an increasing Lévy process. Thus B_{N_t}, for example, is a Lévy process.

Example 7 The sum of two independent Lévy processes is a Lévy process.

Iterating the steps in these last two examples produces a wide variety of Lévy processes.[9] Appendix E provides some examples of models in which consumption can be thought of as following a Lévy process.

2.1 Calculations

In continuous time, the price of a claim to the dividend stream $\{D_t\} \equiv \{(C_t)^\lambda\}$ is

$$\begin{aligned} P_\lambda &= \mathbb{E}_0 \left(\int_{t=0}^{\infty} e^{-\rho t} \left(\frac{C_t}{C_0} \right)^{-\gamma} (C_t)^\lambda \, dt \right) \\ &= \frac{D_\lambda}{\rho - \boldsymbol{c}(\lambda - \gamma)} \end{aligned} \tag{28}$$

[9]It is tempting to think that given some arbitrary random variable X, a Lévy process L_t can be defined such that $L_1 = X$; this would be the continuous time analogue of an i.i.d. sequence whose increments are distributed like X. This intuition is incorrect: for example, if X has bounded support, such a Lévy process will never exist. (Sato (1999, section 24) has a proof.) This means, for example, that there is no continuous time equivalent of the discrete time process whose increments are uniformly distributed on $[-1, 1]$. This apparent defect is a flaw only if one believes that there is a particular distinguishing feature of certain identifiable points in time that makes the discrete-time approach valid; otherwise, it should be viewed as a desirable discipline imposed by the continuous framework.

Once again, the condition that ensures finiteness of expected utility is that $\rho > \boldsymbol{c}(1-\gamma)$; if $\rho > 0$, this condition is satisfied for γ in some neighborhood of 1.

The instantaneous return, R_λ, and instantaneous expected return, ER_λ, are given by

$$\begin{aligned} R_\lambda \, dt &\equiv \frac{dP_\lambda}{P_\lambda} + \frac{D_\lambda}{P_\lambda} \, dt \\ &= \frac{dD_\lambda}{D_\lambda} + \frac{D_\lambda}{P_\lambda} \, dt \\ ER_\lambda \, dt &\equiv \mathbb{E}\left(\frac{dD_\lambda}{D_\lambda}\right) + \frac{D_\lambda}{P_\lambda} \, dt \end{aligned}$$

The following proposition shows that the discrete-time results go through almost unchanged, except that the equations that previously held for log dividend yields, the log riskless rate and the log risk premium now apply to the instantaneous dividend yield, the instantaneous riskless rate and the instantaneous risk premium.

Proposition 3 (Reprise of earlier results)**.** *The riskless rate, R_f, consumption-wealth ratio, C/W, and risk premium on aggregate wealth, $RP \equiv ER_1 - R_f$, are given by*

$$\begin{aligned} R_f &= \rho - \boldsymbol{c}(-\gamma) \\ C/W &= \rho - \boldsymbol{c}(1-\gamma) \\ RP &= \boldsymbol{c}(1) + \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\,. \end{aligned}$$

The Gordon growth model holds:

$$D_\lambda / P_\lambda = ER_\lambda - \boldsymbol{c}(\lambda).$$

Proof. See Appendix C. □

2.2 A concrete example: disasters

To aid intuition, it is helpful to demonstrate the above results in the context of a particular model. In this section, I show how to derive a convenient continuous-time version of Barro (2006a). I use the model to show that i.i.d. disaster models make predictions for the fundamentals that are sensitively dependent on the parameter values assumed. In particular, making disasters more frequent or more severe drives the riskless rate down sharply.

Suppose that log consumption follows the jump-diffusion process

$$G_t = \widetilde{\mu} t + \sigma_B B_t + \sum_{i=1}^{N(t)} Y_i \tag{29}$$

where B_t is a standard Brownian motion, $N(t)$ is a Poisson counting process with parameter ω and Y_i are i.i.d. random variables with some arbitrary distribution. The significance of this example is that any Lévy process can be approximated arbitrarily accurately by a process of the form (29).[10] I will assume that all moments of the disaster size Y_1 are finite, from which it follows that all moments of G are finite.

The CGF is $\boldsymbol{c}(\theta) = \log \boldsymbol{m}(\theta)$, where

$$\begin{aligned} \boldsymbol{m}(\theta) &= \mathbb{E}e^{\theta G_1} \\ &= e^{\widetilde{\mu}\theta} \cdot \mathbb{E}e^{\sigma_B \theta B_1} \cdot \mathbb{E}e^{\theta \sum_{i=1}^{N(1)} Y_i} ; \end{aligned}$$

separating the expectation into two separate products is legitimate since the Poisson jumps and Y_i are independent of the Brownian component B_t. The middle term is the expectation of a lognormal random variable: $\mathbb{E}e^{\theta\sigma_B B_1} = e^{\sigma_B^2\theta^2/2}$. The final term is slightly more complicated, but can be evaluated by conditioning on the number of Poisson jumps that take place before $t = 1$:

$$\begin{aligned} \mathbb{E}\exp\left\{\theta \sum_{i=1}^{N(1)} Y_i\right\} &= \sum_0^\infty \frac{e^{-\omega}\omega^n}{n!} \mathbb{E}\exp\left\{\theta \sum_1^n Y_i\right\} \\ &= \sum_0^\infty \frac{e^{-\omega}\omega^n}{n!} \left[\mathbb{E}\exp\left\{\theta Y_1\right\}\right]^n \\ &= \exp\left\{\omega\left(\mathbb{E}e^{\theta Y_1} - 1\right)\right\} \\ &= \exp\left\{\omega\left(\boldsymbol{m}_{Y_1}(\theta) - 1\right)\right\} , \end{aligned}$$

Finally, we have

$$\boldsymbol{m}(\theta) = \exp\left\{\widetilde{\mu}\theta + \sigma_B^2\theta^2/2 + \omega\left(\boldsymbol{m}_{Y_1}(\theta) - 1\right)\right\}$$

and so

$$\boldsymbol{c}(\theta) = \widetilde{\mu}\theta + \sigma_B^2\theta^2/2 + \omega\left(\boldsymbol{m}_{Y_1}(\theta) - 1\right) . \tag{30}$$

[10] In fact a stronger result holds: for any Lévy process L_t, there exists a sequence of compound Poisson processes $\{L_t^n\}_{n=1}^\infty$ such that

$$\mathbb{P}\left[\lim_{n\to\infty} \sup_{t\le u} |L_t - L_t^n| = 0, \forall u \ge 0\right] = 1.$$

See Sato (1999, Chapter 9) for a proof. In view of this, I could leave the drift term $\widetilde{\mu}t$ and Brownian term $\sigma_B B_t$ out of (29); I include them out of deference to the previous literature.

The cumulants can be read off from the CGF (30):

$$\begin{aligned}\kappa_n(G) &= \boldsymbol{c}^{(n)}(0) \\ &= \begin{cases} \widetilde{\mu} + \omega\,\mathbb{E}Y & n=1 \\ \sigma_B^2 + \omega\,\mathbb{E}Y^2 & n=2 \\ \omega\,\mathbb{E}Y^n & n\geq 3 \end{cases}\end{aligned} \tag{31}$$

Turning off the Brownian motion component of consumption growth ($\sigma_B = 0$) affects only the second cumulant (variance). Turning off jumps, on the other hand, corresponds to setting $\omega = 0$, which alters *all* the cumulants and in particular sets $\kappa_n = 0$ for $n \geq 3$. This illustrates how introducing jumps can significantly alter a model's asset-pricing implications.

Take the case in which $Y \sim N(-b, s^2)$; b is assumed to be greater than zero, so the jumps represent disasters. The CGF is then

$$\boldsymbol{c}(\theta) = \widetilde{\mu}\theta + \frac{1}{2}\sigma_B^2\theta^2 + \omega\left(e^{-\theta b + \frac{1}{2}\theta^2 s^2} - 1\right). \tag{32}$$

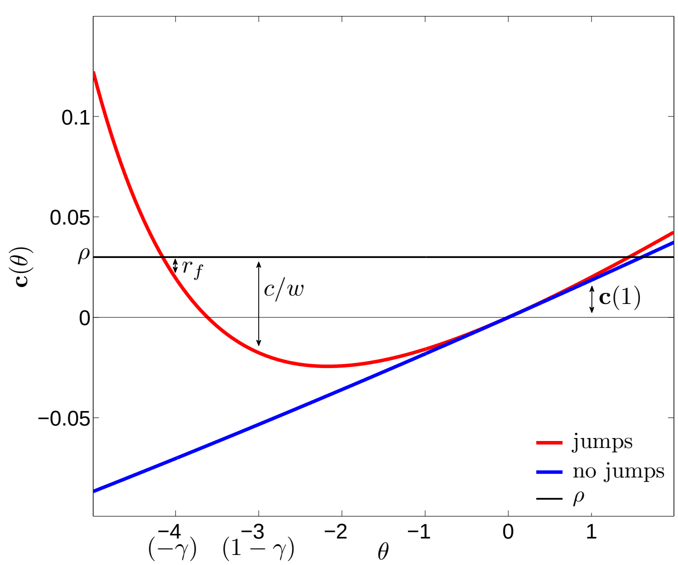


Figure 1: The CGF in equation (32) shown with and without ($\omega = 0$) jumps. The figure assumes that $\gamma = 4$.

Figure 1 shows the CGF of (32) plotted against θ. I set parameters which correspond to Barro's (2006a) baseline calibration—$\gamma = 4, \sigma_B = 0.02, \rho = 0.03, \widetilde{\mu} = 0.025, \omega = 0.017$—and choose $b = 0.39$ and $s = 0.25$ to match the mean and variance of the distribution of jumps used in the same paper. I also plot the CGF that results in the absence of jumps ($\omega = 0$).

In the latter case, I adjust the drift of consumption growth to keep mean log consumption growth constant.

The riskless rate, consumption-wealth ratio and mean consumption growth can be read directly off the graph, as indicated by the arrows. The risk premium can be calculated from these three via the Gordon growth formula ($rp = c/w + \boldsymbol{c}(1) - r_f$), or read directly off the graph as follows. Draw one line from $(-\gamma, \boldsymbol{c}(-\gamma))$ to $(1, \boldsymbol{c}(1))$ and another from $(1-\gamma, \boldsymbol{c}(1-\gamma))$ to $(0,0)$. The midpoint of the first line lies above the midpoint of the second by convexity of the CGF. The risk premium is twice the distance from one midpoint to the other. This procedure is illustrated in Figure 2.

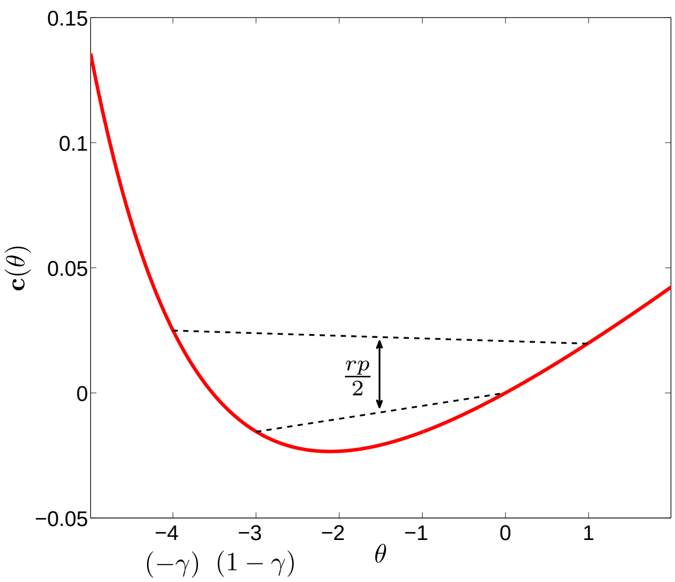


Figure 2: The risk premium. The figure assumes that $\gamma = 4$.

The standard lognormal model predicts a counterfactually high riskless rate; in Figure 1, this is reflected in the fact that the no-jumps CGF lies well below ρ for reasonable values of θ. Similarly, the standard lognormal model predicts a counterfactually low equity premium. In Figure 1, this manifests itself in a no-jump CGF which is practically linear over the relevant range and which is upward-sloping between $-\gamma$ and $1-\gamma$. Conversely, the disaster CGF has a shape which allows it to match observed fundamentals closely.

Zooming out on Figure 1, we obtain Figure 3, which further illustrates the equity premium and riskless rate puzzles. With jumps, the CGF is visible at the right-hand side of the figure; the CGF explodes so quickly as θ declines that it is only visible for θ greater than about -5. The jump-free lognormal CGF has incredibly low curvature. For a realistic

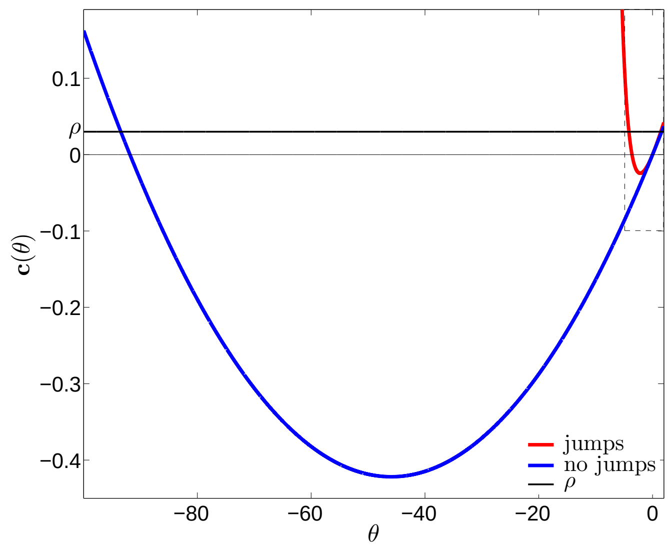


Figure 3: Zooming out to see the equity premium and riskless rate puzzles. The dashed box in the upper right-hand corner is the boundary of the region plotted in Figure 1.

riskless rate and equity premium, the model requires a risk aversion above 80.

With the explicit expression (32) for the CGF in hand, it is easy to investigate the sensitivity of a disaster model's predictions to the parameter values assumed. Table 1 shows how changes in the calibration of the distribution of disasters affect the relevant fundamentals and the cost of consumption uncertainty, ϕ. As is evident from the table, the predictions of the disaster model are sensitively dependent on the precise calibration. In particular, small changes in any of the parameters ω, b or s have large effects on the riskless rate and equity premium. For example, increasing s (the standard deviation of disaster sizes) from 0.25 to 0.30 drives the riskless rate down by more than three per cent. Given that these parameters are hard to estimate—disasters happen very rarely—this is a significant difficulty.

As before, the CGF can also be thought of as a power series in θ. Table 2 investigates the consequences of truncating this power series at the term of order θ^n. When $n = 2$, this is equivalent to making a lognormality assumption, as noted above. With $n = 3$, it can be thought of as an approximation which accounts for the influence of skewness; $n = 4$ also allows for kurtosis. As is clear from the table, however, much of the action is due to cumulants of fifth order and higher. This suggests that one should not expect calculations based on third- or fourth-order approximations to capture fully the influence of disasters.

	ω	b	s	R_f	C/W	RP
Baseline case	**0.017**	**0.39**	**0.25**	**1.0**	**4.8**	**5.7**
High ω	0.022			-2.4	3.1	7.4
Low ω	0.012			4.5	6.4	4.1
High b		0.44		-1.9	3.6	7.5
Low b		0.34		3.5	5.8	4.4
High s			0.30	-2.2	3.8	8.1
Low s			0.20	3.2	5.5	4.2

Table 1: The impact of different assumptions about the distribution of disasters. All parameters other than ω, b and s are as before.

n	R_f	C/W	RP	
1	10.3	8.5	0.0	deterministic
2	7.1	6.7	1.6	lognormal
3	4.7	5.7	3.0	
4	3.0	5.1	4.1	
∞	1.0	4.8	5.7	true model

Table 2: The impact of approximating the disaster model by truncating at the nth cumulant. All parameters as in baseline case of Table 1.

3 Restrictions on preference parameters

Any three of the riskless rate, consumption-wealth ratio, risk premium and expected consumption growth pin down the value of the fourth, via the equation $c/w = r_f + rp - \boldsymbol{c}(1)$ of (21). I now assume that these quantities are *observable*, and suppose for simplicity that the riskless rate and mean consumption growth are specified by $r_f = 0.02$ and $\boldsymbol{c}(1) = 0.02$, and that the risk premium and consumption-wealth are given by $rp = 0.06$ and $c/w = 0.06$. One interpretation is that we are interested only in models which avoid the riskless rate and equity premium puzzles and make a reasonable assumption about mean consumption growth. Table 3 summarizes these assumptions.

We have seen, too, that the riskless rate, risk premium, consumption-wealth ratio and mean consumption growth tell us information about the shape of the CGF. I now show how to exploit this observation to find restrictions on preference parameters, in terms of observable fundamentals, that must hold in *any* Epstein-Zin/i.i.d. model, no matter what

riskless rate	r_f	0.02
risk premium	rp	0.06
consumption-wealth ratio	c/w	0.06
mean consumption growth	$\boldsymbol{c}(1)$	0.02

Table 3: Assumed values of the observables.

pattern of (say) rare disasters we allow ourselves to entertain.

Since for example $r_f = \rho - \boldsymbol{c}(-\gamma)$ in the power utility case, observation of the riskless rate tells us something about ρ and something about the value taken by the CGF at $-\gamma$. Similarly, observation of the consumption-wealth ratio tells us something about ρ and something about the value taken by the CGF at $1-\gamma$. Next, $\boldsymbol{c}(1) = \log \mathbb{E}(C_1/C_0)$ is pinned down by mean consumption growth, and $\boldsymbol{c}(0) = 0$ by definition. How, though, can we get control on the enormous range of possible consumption processes? One approach is to exploit the fact that the CGF of any random variable is convex, a property that is so central in what follows that I record it as

Fact 1. *CGFs are convex.*

Proof. Since $\boldsymbol{c}(\theta) = \log \boldsymbol{m}(\theta)$, we have

$$\begin{aligned} \boldsymbol{c}''(\theta) &= \frac{\boldsymbol{m}(\theta) \cdot \boldsymbol{m}''(\theta) - \boldsymbol{m}'(\theta)^2}{\boldsymbol{m}(\theta)^2} \\ &= \frac{\mathbb{E}e^{\theta G}\mathbb{E}G^2 e^{\theta G} - \left(\mathbb{E}Ge^{\theta G}\right)^2}{\boldsymbol{m}(\theta)^2} . \end{aligned}$$

The numerator of this expression is positive by a version of the Cauchy-Schwartz inequality which states that $\mathbb{E}X^2 \cdot \mathbb{E}Y^2 \geq \mathbb{E}(|XY|)^2$ for any random variables X and Y. In this case, we need to set $X = e^{\theta G/2}$ and $Y = Ge^{\theta G/2}$. (See, for example, Billingsley (1995), for further discussion of CGFs.) □

The convexity of the CGF can be thought of as encoding useful inequalities (those of Jensen and Lyapunov, for example) in a memorable and geometrically intuitive form.

I now state the main result, which takes full advantage of Fact 1.

Proposition 4. *In the power utility case, we have*

$$\boldsymbol{c}(1) - rp \leq \frac{c/w - \rho}{\gamma - 1} \leq \boldsymbol{c}(1) \tag{33}$$

In the Epstein-Zin case, we have

$$\boldsymbol{c}(1) - rp \leq \frac{c/w - \rho}{1/\psi - 1} \leq \boldsymbol{c}(1) \tag{34}$$

Proof. From equation (19) we have, in the Epstein-Zin case,

$$\frac{c/w - \rho}{1/\psi - 1} = \frac{\boldsymbol{c}(1-\gamma)}{1-\gamma}.$$

The convexity of $\boldsymbol{c}(\theta)$ and the fact that $\boldsymbol{c}(0) = 0$ imply that

$$\frac{\boldsymbol{c}(-\gamma)}{-\gamma} \leq \frac{\boldsymbol{c}(1-\gamma)}{1-\gamma} \leq \boldsymbol{c}(1);$$

to see this, note that if $f(\theta)$ is a convex function passing through zero, then $f(\theta)/\theta$ is increasing. Putting the two facts together, we have

$$\frac{\boldsymbol{c}(-\gamma)}{-\gamma} \leq \frac{c/w - \rho}{1/\psi - 1} \leq \boldsymbol{c}(1).$$

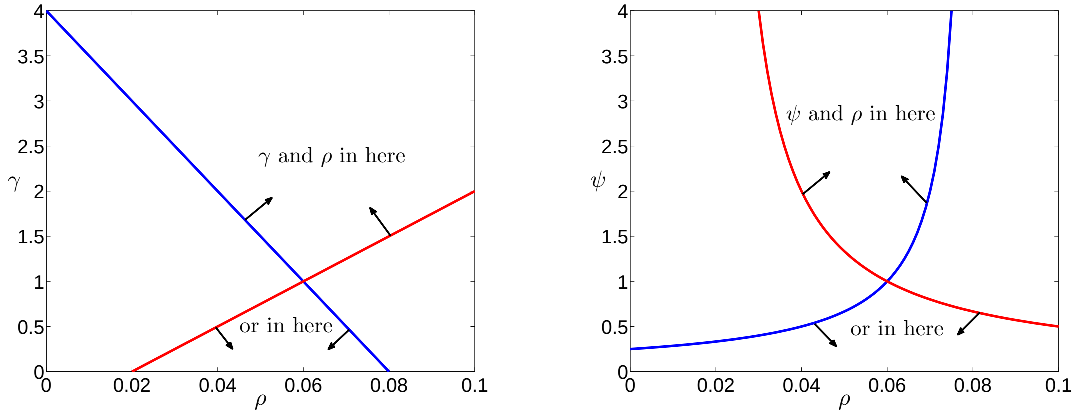
After some rearrangement of the left-hand inequality using (18) and (19), this gives (34). Equation (33) follows since $\gamma = 1/\psi$ in the power utility case. □

The intuition for the result is that as ψ approaches one, the consumption-wealth ratio approaches ρ. Therefore, when the consumption-wealth ratio is far from ρ, ψ must be far from one. Using the empirically reasonable values $rp = 6\%, r_f = 2\%, c/w = 6\%, \boldsymbol{c}(1) = 2\%$, we have the restriction that $-0.04 \leq (0.06 - \rho)/(1/\psi - 1) \leq 0.02$, or equivalently

$$\begin{aligned} 4 - \frac{1}{\psi} \leq 50\rho \leq 1 + \frac{2}{\psi} \quad &\text{if} \quad \psi \leq 1 \\ 1 + \frac{2}{\psi} \leq 50\rho \leq 4 - \frac{1}{\psi} \quad &\text{if} \quad \psi \geq 1. \end{aligned}$$

Figures 4a and 4b illustrate these constraints. Note, for example, that if ψ is greater than one, ρ is constrained to lie between 0.02 and 0.08; if also ψ is less than two, ρ must lie between 0.04 and 0.07.

A pragmatic conclusion that might be drawn from these diagrams is that they can be used to constrain ρ precisely—by setting it equal to the consumption-wealth ratio, c/w—and that following this choice of ρ, ψ (or γ) can be chosen freely.



(a) Power utility case: γ and ρ

(b) Epstein-Zin case: ψ and ρ

Figure 4: Parameter restrictions for i.i.d. models with $rp = 6\%$, $r_f = 2\%$ and log expected consumption growth of 2%.

3.1 Hansen-Jagannathan and good-deal bounds

The restrictions in Proposition 4 are complementary to the bound derived by Hansen and Jagannathan (1991), which relates the standard deviation and mean of the stochastic discount factor, M, to the Sharpe ratio on an arbitrary asset, SR:

$$SR \leq \frac{\sigma(M)}{\mathbb{E}M}\,. \tag{35}$$

In the Epstein-Zin-i.i.d. setting, the right-hand side of (35) becomes

$$\begin{aligned}\frac{\sigma(M)}{\mathbb{E}M} &= \sqrt{\frac{\mathbb{E}M^2}{(\mathbb{E}M)^2}-1} \\ &= \sqrt{e^{\boldsymbol{c}(-2\gamma)-2\boldsymbol{c}(-\gamma)}-1}\,;\end{aligned} \tag{36}$$

combining (35) and (36), we obtain a Hansen-Jagannathan bound translated into CGF notation:

$$\log\left(1+SR^2\right) \leq \boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma)\,. \tag{37}$$

Cochrane and Saá-Requejo (2000) observe that inequality (35) suggests a natural way to restrict asset-pricing models. Suppose that $\sigma(M)/\mathbb{E}M \leq h$; then (35) implies that the maximal Sharpe ratio is less than h. The idea is that assets with higher Sharpe ratios are "good deals"—deals which are in fact too good to be true. In CGF notation, the good-deal bound is that

$$\boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma) \leq \log\left(1+h^2\right) \tag{38}$$

Suppose, for example, that we wish to impose the restriction that Sharpe ratios above 100% are too good a deal to be available. Then the good-deal bound is $\boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma) \leq \log 2$. This expression can be evaluated under particular parametric assumptions about the consumption process. In the case in which consumption growth is lognormal, with volatility of log consumption equal to σ, it supplies an upper bound on risk aversion: $\gamma \leq \sqrt{\log 2}/\sigma$ (which is about 42 if $\sigma = 0.02$). However, this upper bound is rather weak, and in any case the postulated consumption process is inconsistent with observed features of asset markets such as the high equity premium and low riskless rate.

Alternatively, one might model the consumption process as subject to disasters in the sense of Section 2.2. In this case, the good-deal bound implies tighter restrictions on γ, but these restrictions are sensitively dependent on the disaster parameters.

In order to progress from (38) to a bound on γ and ρ which does not require parametrization of the consumption process, we want to relate $\boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma)$ to quantities which can be directly observed. For example, the Hansen-Jagannathan bound (37) improves on a conclusion which follows from the convexity of the CGF, namely, that

$$0 \leq \boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma)\,. \tag{39}$$

This trivial inequality follows by considering the value of the CGF at the three points $\boldsymbol{c}(0)$, $\boldsymbol{c}(-\gamma)$, and $\boldsymbol{c}(-2\gamma)$. Convexity implies that the average slope of the CGF is more negative (or less positive) between -2γ and $-\gamma$ than it is between $-\gamma$ and 0. To be precise, it implies that

$$\frac{\boldsymbol{c}(-\gamma) - \boldsymbol{c}(-2\gamma)}{\gamma} \leq \frac{\boldsymbol{c}(0) - \boldsymbol{c}(-\gamma)}{\gamma} \tag{40}$$

from which (39) follows immediately, given that $\boldsymbol{c}(0) = 0$. Combining (38) and (39), we obtain the somewhat underwhelming result that

$$0 \leq \log\left(1 + h^2\right)\,.$$

However, we can sharpen (39) by comparing the slope of the CGF between -2γ and $-\gamma$ to the slope between $-\gamma$ and $1-\gamma$ (as opposed to that between $-\gamma$ and 0). Making this formal, we have by convexity of the CGF that

$$\frac{\boldsymbol{c}(-\gamma) - \boldsymbol{c}(-2\gamma)}{\gamma} \leq \frac{\boldsymbol{c}(1-\gamma) - \boldsymbol{c}(-\gamma)}{1}\,,$$

from which it follows that

$$\begin{aligned}
\boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma) &\geq (\gamma - 1)\boldsymbol{c}(-\gamma) - \gamma\boldsymbol{c}(1-\gamma) \\
&= (\gamma - 1)(c/w - r_f) + \vartheta(c/w - \rho)
\end{aligned}$$

or equivalently

$$\frac{\sigma(M)}{\mathbb{E}M} \geq \sqrt{e^{(\gamma-1)(c/w-r_f)+\vartheta(c/w-\rho)}-1}\,. \tag{41}$$

The good deal bound therefore implies that

$$(\gamma-1)(c/w-r_f)+\vartheta(c/w-\rho) \leq \log\left(1+h^2\right)\,. \tag{42}$$

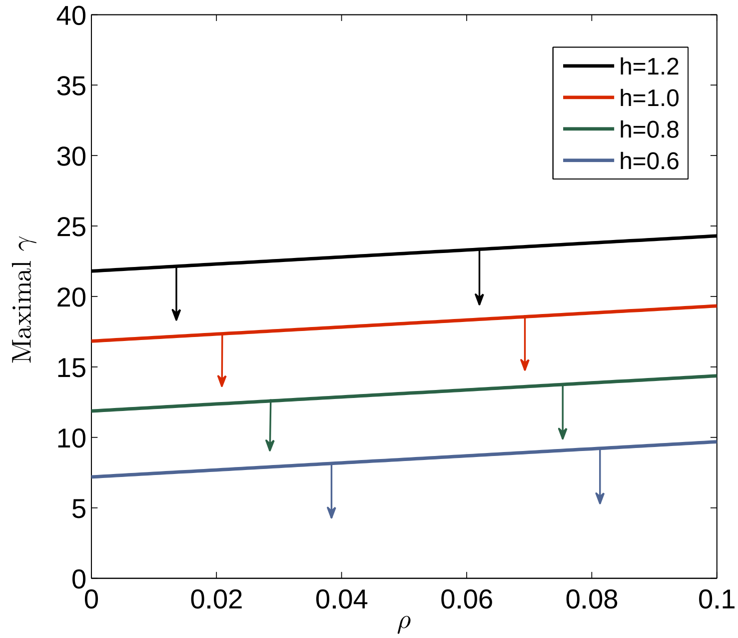


Figure 5: Restrictions on γ and ρ implied by good-deal bounds in the power utility case with $c/w = 0.06, r_f = 0.02$.

Working with the power utility case for simplicity ($\vartheta = 1$) and setting $c/w = 0.06, r_f = 0.02$, Figure 5 shows the upper bounds on γ that result for various different h. Lower values of h imply tighter restrictions. When $h = 1$—ruling out Sharpe ratios above 100%—we have $\gamma \leq 16.8 + 25\rho$. So if $\rho = 0.03$, $\gamma < 17.6$.

Alternatively, we could take the approach suggested at the end of the previous section, by setting $\rho = c/w$. In the general (Epstein-Zin) case, equation (42) then implies the simple restriction

$$\gamma \leq 1 + \frac{\log\left(1+h^2\right)}{c/w-r_f}\,. \tag{43}$$

(To avoid unnecessary complication I have imposed the empirically relevant case $c/w \geq r_f$.) Setting $c/w = 0.06$, $r_f = 0.02$, and $h = 1$, this implies that $\gamma < 18.4$.

The important feature of the bounds (42) and (43) is that they do not require one to take a stand on the details of the higher cumulants of consumption growth. By exploiting the observable consumption-wealth ratio and riskless rate, calibration of the consumption process can be avoided.

4 The cost of consumption fluctuations

Continuing with the theme of extracting information from observable fundamentals, I now explore the implications of the consumption-wealth ratio for estimates of the cost of consumption fluctuations in the style of Lucas (1987), Obstfeld (1994) or Barro (2006b). I work with power utility throughout this section and assume that $\gamma \neq 1$, though results for log utility are stated in the Propositions.[11]

A starting point is the close correspondence between expected utility and the price of the consumption claim (that is, wealth):

$$U(\gamma) \equiv \mathbb{E}\left[\sum_{t=0}^{\infty} e^{-\rho t} \frac{C_t^{1-\gamma}}{1-\gamma}\right] \longleftrightarrow \mathbb{E}\left[\sum_{t=1}^{\infty} e^{-\rho t} \left(\frac{C_t}{C_0}\right)^{1-\gamma}\right] = \frac{W_0}{C_0}\,.$$

In fact we have

$$U(\gamma) = \frac{C_0^{1-\gamma}}{1-\gamma} \cdot \left(1 + \frac{W_0}{C_0}\right)\,. \tag{44}$$

This correspondence between expected utility and the consumption-wealth ratio, and hence (44), does not have a meaningful analogue in the log utility case. In a sense, the consumption-wealth ratio is less informative in the log utility case since it is pinned down by the time discount rate, $C/W = e^{\rho} - 1$.

Expected utility can also be expressed in terms of the CGF:

$$U(\gamma) = \frac{C_0^{1-\gamma}}{1-\gamma} \left(1 + \frac{1}{e^{\rho - \boldsymbol{c}(1-\gamma)} - 1}\right)\,, \qquad \gamma \neq 1\,. \tag{45}$$

When $\gamma < 1$ the representative agent prefers large values of $\boldsymbol{c}(1-\gamma)$ and when $\gamma > 1$ the representative agent prefers small values of $\boldsymbol{c}(1-\gamma)$. When $\gamma > 1$, the representative agent likes positive mean and positive skew and positive cumulants of odd orders but dislikes large values of variance, kurtosis and cumulants of even orders; when $\gamma < 1$ the representative agent likes large means, large variances, large skewness, large kurtosis—large positive values of cumulants of *all* orders.[12]

[11] Calculations in the Epstein-Zin case are in Appendix B.

[12] As always, these cumulants are the cumulants of *log* consumption growth. This explains the result that risk-averse agents with $\gamma < 1$ prefer large variances, which may initially seem counterintuitive.

Equation (44) gives expected utility under the status quo; expression (45) permits the calculation of expected utility under alternative consumption processes with their corresponding CGFs. I compare two quantities: expected utility with initial consumption $(1+\phi)C_0$ and the status quo consumption growth process,[13] and expected utility with initial consumption C_0 and the alternative consumption growth process. The *cost of uncertainty* is the value of ϕ which equates the two. This definition follows the lead of Lucas (1987) and Obstfeld (1994) and Section V of Alvarez and Jermann (2004).

The following sections consider two possible counterfactuals: (i) a scenario in which all uncertainty is eliminated, and (ii) a scenario in which the variance of consumption growth is reduced by α^2 but higher cumulants are unchanged. In each case, mean consumption growth $\mathbb{E}C_{t+1}/C_t$ is held constant.

4.1 The elimination of all uncertainty

Since

$$\mathbb{E}\left(\frac{C_1}{C_0}\right) = \mathbb{E}e^G = e^{\boldsymbol{c}(1)} \ ,$$

keeping mean consumption growth constant is equivalent to holding $\boldsymbol{c}(1) = \log \mathbb{E}(C_1/C_0)$ constant. If all uncertainty is also to be eliminated, log consumption follows the trivial Lévy process $\overline{G}_t$ whose CGF is $\boldsymbol{c}_{\overline{G}}(\theta) = \boldsymbol{c}(1)\cdot\theta$ for all θ.

From (44) and (45), ϕ solves the equation

$$\frac{[(1+\phi)C_0]^{1-\gamma}}{1-\gamma}\cdot\left(1+\frac{W_0}{C_0}\right) = \frac{C_0^{1-\gamma}}{1-\gamma}\cdot\frac{e^{\rho-\boldsymbol{c}(1)\cdot(1-\gamma)}}{e^{\rho-\boldsymbol{c}(1)\cdot(1-\gamma)}-1}\,. \tag{46}$$

Simplifying, we have

$$\phi = \left(1+\frac{W_0}{C_0}\right)^{\frac{1}{\gamma-1}}\left\{1-e^{-\rho}\left[\mathbb{E}\left(\frac{C_1}{C_0}\right)\right]^{1-\gamma}\right\}^{\frac{1}{\gamma-1}} - 1\,. \tag{47}$$

What assumptions are required to derive (47)? The left-hand side of (46) relies on the correspondence between expected utility and the consumption-wealth ratio that was noted at the beginning of section 4. This correspondence follows directly from Lucas's (1978) Euler equation with power utility: the assumption that real-world consumption growth is i.i.d. is not required. The cost of *all* uncertainty given in (47) depends only on the power utility assumption. The counterfactual case of deterministic growth is trivially i.i.d., so it is convenient to work with a CGF, though not necessary. (Below, I calculate the benefit

[13] Since the consumption growth process is unchanged, the consumption-wealth ratio remains constant. The increase in initial consumption therefore corresponds to an increase in initial *wealth* by proportion ϕ.

associated with a reduction in the variance of consumption growth, while higher moments remain constant. In this case, the i.i.d. assumption is required and CGFs are central to my calculations.)

In the Epstein-Zin case it is also necessary to rely on the i.i.d. assumption. It turns out that (47) is misleading in that the γ terms that appear in it are capturing not risk aversion but the elasticity of intertemporal substitution, as the following proposition shows.

Proposition 5. *In the Epstein-Zin case with elasticity of intertemporal substitution ψ, the cost of uncertainty, ϕ, satisfies*

$$\phi = \left(1 + \frac{W_0}{C_0}\right)^{\frac{1}{1/\psi - 1}} \left\{1 - e^{-\rho} \left[\mathbb{E}\left(\frac{C_1}{C_0}\right)\right]^{1-\frac{1}{\psi}}\right\}^{\frac{1}{1/\psi-1}} - 1\,. \tag{48}$$

With power utility and $\gamma \neq 1$, the above equation holds, even in the absence of the i.i.d. assumption, with $1/\psi$ replaced by γ.

With log utility we do require the i.i.d. assumption, and have

$$\begin{aligned} \phi &= \exp\left[(\boldsymbol{c}(1) - \mu) \,/\, (e^{\rho} - 1)\right] - 1 \\ &= \exp\left[(\boldsymbol{c}(1) - \mu)\,\frac{W_0}{C_0}\right] - 1\,. \end{aligned}$$

Proof. See appendix B for the Epstein-Zin calculations. □

Proposition 5 shows that if the mean consumption growth rate in levels, consumption-wealth ratio[14] and preference parameters ρ and ψ can be estimated accurately, then the gains notionally available from eliminating all uncertainty can be estimated without needing to make assumptions about the particular stochastic process followed by consumption. In particular, in the Epstein-Zin case, ϕ is not—directly—dependent on γ, nor on estimates of the variance (and higher cumulants) of consumption growth. The consumption-wealth ratio encodes all relevant information about the amount of risk (that is, the cumulants κ_n, $n \geq 2$) and the representative agent's attitude to risk (γ).

In the power utility case in particular, this result is rather general. It applies to arbitrary consumption processes and so nests results obtained by Lucas (1987, 2003), Obstfeld (1994) and Barro (2006b).[15] The important feature is that I treat the consumption-wealth

[14] If one is prepared to identify the consumption claim with the stock market, as in Mehra and Prescott (1985), then the dividend yield on the market can be used in place of C/W.

[15] There is a slight wrinkle in that Lucas (1987, 2003) assumes that current consumption C_0 is not known in the risky case. I follow Alvarez and Jermann (2004) in assuming that C_0 is known. The distinction turns out not to be quantitatively significant in practice.

ratio as an observable. Lucas, Obstfeld and Barro postulate some particular consumption process and, implicitly or explicitly, calculate the consumption-wealth ratio implied by that consumption process. For these authors, a change in γ is accompanied by a change in C/W; I, on the other hand, hold C/W constant and view it as containing information about the underlying consumption process.

4.1.1 The cost of all uncertainty with power utility

As before, suppose that $c/w = 0.06$ and $\boldsymbol{c}(1) = 0.02$, and that $\rho = 0.03$ and $\gamma = 4$. Substituting these values into (47) gives $\phi \approx 14\%$.

This cost estimate is roughly two orders of magnitude higher than that obtained by Lucas (1987, 2003), even allowing for the higher risk aversion assumed in this paper. Although Lucas's calculations do not make use of the observable consumption-wealth ratio, it is possible to calculate the consumption-wealth ratio implied by his assumptions on the consumption process and my assumptions on ρ and γ; the result is an implied consumption-wealth ratio $c/w = 0.0896$. Substituting this value back into (47), we recover the far lower cost estimate, $\phi \approx 0.14\%$. Once one considers the consumption-wealth ratio as an observable, the cost of uncertainty appears to be considerably higher.

	ρ	γ	$\boldsymbol{c}(1)$	c/w	ϕ
Baseline case	**0.03**	**4**	**0.02**	**0.06**	**14%**
High ρ	0.04				18%
Low ρ	0.02				10%
High γ		5			16%
Low γ		3			7.7%
High growth			0.025		20%
Low growth			0.015		7.5%
High c/w				0.07	8.4%
Low c/w				0.05	21%

Table 4: The cost of consumption fluctuations with power utility.

Table 4 shows how different assumptions on preference parameters and on mean consumption growth and the consumption-wealth ratio affect the estimate of the cost of uncertainty. Apart from the last two lines of the table, the consumption-wealth ratio c/w is held constant in the calculations.

The cost of uncertainty is *higher* when agents are more impatient (high ρ). When ρ is low, the (relatively) high consumption-wealth ratio signals that there is not too much risk in the economy. When ρ is high, the (relatively) low consumption-wealth ratio signals that there is considerable risk in the economy, or that risk aversion is high.

The case in which γ varies is somewhat more complicated. Suppose, first, that ρ is low relative to c/w, as in the above table. If we imagine holding the level of risk constant, then increasing γ from a low level will lead, first, to an increase in c/w because the representative agent is less inclined to substitute consumption intertemporally. Ultimately, however, increasing γ must lead to a decrease in c/w, once the precautionary saving motive starts to dominate. (These statements are most easily understood if one keeps Figure 1 in mind.) Turning the logic around, if γ increases but c/w remains constant, the level of risk in the economy must first be increasing and then declining. It follows that we may expect increases in γ to have ambiguous effects on the cost of uncertainty, holding c/w constant. In table 4, the former effect dominates.

When, on the other hand, ρ is large relative to c/w, the CGF must have significant curvature—look at Figure 1. It follows that there is considerable risk in the economy; in this case, for γ to increase while c/w remains constant, it can only be that the level of risk is declining. Thus we expect to see that for low values of ρ, the cost of uncertainty is first increasing and then decreasing in γ, while for larger values of ρ, the cost is declining in γ.

These observations are borne out by Figure 6. When $\rho = 0.03$, the cost of uncertainty is first increasing and then decreasing in γ. When $\rho = 0.06$ or 0.09, the cost of uncertainty is decreasing in γ.

Finally, when ρ equals 0.03, γ must be at least 2.5 to be consistent with the assumed mean consumption growth and consumption-wealth ratio. In Figure 6, the black line hits zero at $\gamma = 2.5$ because the only possibility consistent with $\rho = 0.03, \gamma = 2.5, \boldsymbol{c}(1) = 0.02, c/w = 0.06$ is that consumption is deterministic.

4.1.2 The cost of all uncertainty with Epstein-Zin preferences

With Epstein-Zin preferences, the intertemporal substitution parameter ψ influences the agent's preference over the timing of resolution of uncertainty. When $\psi > 1/\gamma$, the agent prefers early resolution of uncertainty; when $\psi < 1/\gamma$, the agent prefers late resolution of uncertainty. In this sense, Epstein and Zin (1989) observe that the elasticity of intertemporal substitution, ψ, "seems intertwined with both substitutability and risk aversion." This fact frustrates intuition in the Epstein-Zin case.

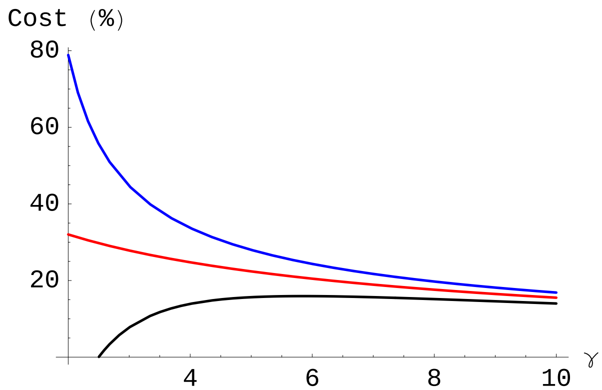
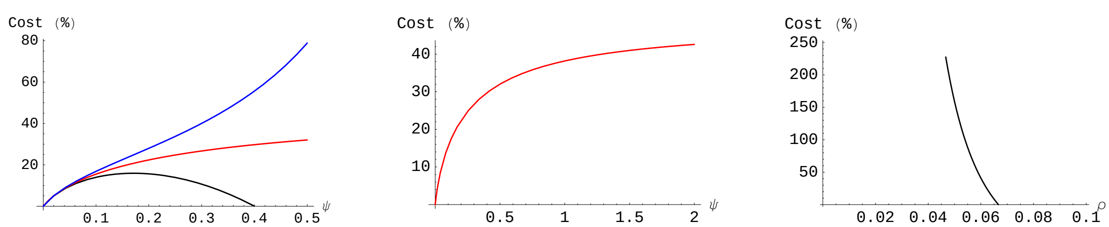


Figure 6: The cost of consumption uncertainty plotted against risk aversion, γ, when $\rho = 0.03$ (in black), $\rho = 0.06$ (in red) and $\rho = 0.09$ (in blue). The cost of uncertainty ultimately declines as γ increases: for very high values of γ, c/w can only equal 0.06 if there is relatively little risk in consumption growth.

The cost calculations made in the previous section can be mapped directly into the Epstein-Zin case if $\psi = 0.25$. Figure 7a, which is the dual of Figure 6 but is more general because it makes no restrictions on γ, illustrates the effects of changes in ρ and ψ. When ρ is high, the cost is high—for the same reasons as above.



(a) Against ψ, with $\rho = 0.03$ (in black), $\rho = 0.06$ (in red) and $\rho = 0.09$ (in blue).

(b) Against ψ, with $\rho = c/w = 0.06$.

(c) Against ρ, with $\psi = 1.5$.

Figure 7: The cost of uncertainty with Epstein-Zin preferences.

As before, it is not possible to set ψ and ρ arbitrarily while retaining consistency with observed values of the consumption-wealth ratio. In Figure 7a, we see that we cannot have ψ between 0.4 and 1 if $\rho = 0.03$. However, if (and *only* if) $\rho = c/w$, then ψ can take any value (specifically, any value around one). Figure 7b therefore sets $\rho = c/w$ and shows that the cost of uncertainty increases in ψ. When ψ is around one, the implied cost of uncertainty is high, at about 40% of current wealth.

Finally, Figure 7c plots the cost of uncertainty against ρ, holding ψ fixed at 1.5. For consistency, ρ must lie between 0.0467 and 0.0667. The cost of uncertainty is extraordinarily sensitively dependent on the relationship between ρ and c/w.

4.2 A reduction in the variance of consumption growth

The preceding section showed that there are significant costs due to uncertainty. This section investigates the utility benefit of a reduction in variance, holding all higher cumulants fixed; it requires the assumption that real-world consumption growth is i.i.d. The counterfactual situation under consideration is one in which the variance of log consumption growth is reduced by α^2 from its current level, which can remain unspecified.[16]

Under the new reduced-volatility process, the CGF is

$$\widetilde{\boldsymbol{c}}(\theta) = \boldsymbol{c}(\theta) + \alpha^2\theta/2 - \alpha^2\theta^2/2\,. \tag{49}$$

The term of order θ^2 decreases the variance of log consumption growth by α^2. The term of order θ adjusts the drift of log consumption growth to hold mean consumption growth constant in levels, that is, to ensure that $\widetilde{\boldsymbol{c}}(1) = \boldsymbol{c}(1)$.

The cost of uncertainty, ϕ_α, solves

$$\frac{[(1+\phi_\alpha)C_0]^{1-\gamma}}{1-\gamma} \cdot \left(1 + \frac{W_0}{C_0}\right) = \frac{C_0^{1-\gamma}}{1-\gamma} \cdot \frac{e^{\rho-\widetilde{\boldsymbol{c}}(1-\gamma)}}{e^{\rho-\widetilde{\boldsymbol{c}}(1-\gamma)} - 1}\,.$$

Substituting in from (49), and replacing $\rho - \boldsymbol{c}(1-\gamma)$ with the observable $c/w = \log(1+C/W)$, we obtain after some simplification

$$\phi_\alpha = \left\{1 + \frac{W_0}{C_0}\left[1 - e^{-\frac{1}{2}\alpha^2\gamma(\gamma-1)}\right]\right\}^{1/(\gamma-1)} - 1\,. \tag{50}$$

Carrying out similar calculations in the Epstein-Zin case, we find

Proposition 6. *In the Epstein-Zin case with elasticity of intertemporal substitution ψ, a reduction in consumption variance of α^2 is equivalent in utility terms to a proportional increase in current consumption of ϕ_α, where*

$$\phi_\alpha = \left\{1 + \frac{W_0}{C_0}\left[1 - e^{-\frac{1}{2}\alpha^2\gamma\left(\frac{1}{\psi}-1\right)}\right]\right\}^{\frac{1}{1/\psi-1}} - 1\,. \tag{51}$$

In the power utility case, the above equation holds with $1/\psi$ replaced by γ.

[16] It is possible to consider such an adjustment in variance alone—leaving higher cumulants unchanged—because the Brownian component of log consumption growth only affects the second cumulant. Conversely, it is not clear how to adjust, say, kurtosis without changing other cumulants.

With log utility, we have

$$\begin{aligned}\phi_\alpha &= \exp\left[\frac{1}{2}\alpha^2/\left(e^\rho - 1\right)\right] - 1 \\ &= \exp\left[\frac{1}{2}\alpha^2 \frac{W_0}{C_0}\right] - 1\,.\end{aligned}$$

In all cases, we have the first-order approximation for small α^2

$$\phi_\alpha \approx \frac{W_0}{C_0}\frac{\gamma\alpha^2}{2}\,. \tag{52}$$

Proof. See appendix B for the Epstein-Zin calculations. □

Obstfeld (1994) observes that (52) holds in the power utility case with i.i.d. lognormal consumption growth, but does not argue that it holds in the Epstein-Zin case or for general i.i.d. consumption processes.

With $\gamma = 4$, and setting $c/w = 0.06$ as usual, it follows from (52) that a reduction in variance of 0.0003—as would be associated with a decline in the standard deviation of log consumption growth from 2% to 1%—is equivalent in welfare terms to an increase in current consumption (or equivalently wealth) of 1.0%. While this is a significant quantity, these calculations suggest that most of the cost of uncertainty can be attributed to higher-order cumulants.

5 The multivariate case and heterogeneity

I now briefly describe how to extend the CGF framework to price assets whose dividends are not a power of the stochastic discount factor. To be concise, I work in continuous time. I assume that there is no arbitrage (in which case there exists a stochastic discount factor to time t for arbitary t, labelled M_t/M_0), and that there is an asset under consideration with well-defined price whose dividend stream is $\{D_t\}$.

Motivated by the analysis above, I define $G_t \equiv -\log M_t/M_0$ and $H_t \equiv \log D_t/D_0$, $G \equiv G_1, H \equiv H_1$, and assume that (G_t, H_t) follows a two-dimensional Lévy process. This assumption allows for the possibility that G_t and H_t are correlated; for example, G_t and H_t may be correlated Brownian motions, or may be subject to correlated jumps. As before, however, the increments of (G_t, H_t) are stationary and independent. We can then define the bivariate CGF.

Definition 2. *Given two random variables G and H, the bivariate moment- and cumulant-generating functions are defined by*

$$\boldsymbol{m}_{G,H}(\boldsymbol{\theta}) \equiv \boldsymbol{m}_{G,H}(\theta_1, \theta_2) \equiv \mathbb{E}\, e^{\theta_1 G + \theta_2 H}$$
$$\boldsymbol{c}_{G,H}(\boldsymbol{\theta}) \equiv \boldsymbol{c}_{G,H}(\theta_1, \theta_2) \equiv \log \boldsymbol{m}_{G,H}(\boldsymbol{\theta})$$

The bivariate cumulants of (G, H), written κ_{rs}, are defined by

$$\boldsymbol{c}_{G,H}(\theta_1, \theta_2) = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \kappa_{rs} \frac{\theta_1^r}{r!} \frac{\theta_2^s}{s!}$$

On the set for which the moment-generating function is defined, we have, as before, that

$$\boldsymbol{m}_{G_t,H_t}(\boldsymbol{\theta}) = (\boldsymbol{m}_{G,H}(\boldsymbol{\theta}))^t \tag{53}$$

Thus,

$$\begin{aligned}
P_0 &\equiv \mathbb{E} \int_0^\infty \frac{M_t}{M_0} D_t dt \\
&= D_0 \int_0^\infty \mathbb{E}\left(e^{-G_t + H_t}\right) dt \\
&= D_0 \int_0^\infty e^{\boldsymbol{c}_{G,H}(-1,1)t} dt \\
&= \frac{D_0}{-\boldsymbol{c}_{G,H}(-1,1)},
\end{aligned}$$

or

$$D/P = -\boldsymbol{c}_{G,H}(-1,1);$$

for the price to be well-defined, we require that $\boldsymbol{c}_{G,H}(-1,1) < 0$. Thus pricing a generic asset in an i.i.d. environment is a matter of analyzing the bivariate cumulants of G and H. In Appendix D, I list the first few bivariate cumulants in terms of the central moments of G and H.

The riskless rate is therefore $R_f = -\boldsymbol{c}_{G,H}(-1,0) = -\boldsymbol{c}_G(-1)$. The instantaneous expected return on a generic asset is

$$\begin{aligned}
ER_t &\equiv D/P + \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathbb{E}\left(\frac{D_{\lambda,t+\Delta t} - D_{\lambda,t}}{D_{\lambda,t}}\right) \\
&= -\boldsymbol{c}_{G,H}(-1,1) + \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \left\{\mathbb{E}\left(e^{H_1}\right)^{\Delta t} - 1\right\} \\
&= -\boldsymbol{c}_{G,H}(-1,1) + \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \left\{e^{\boldsymbol{c}_{G,H}(0,1)\Delta t} - 1\right\} \\
&= -\boldsymbol{c}_{G,H}(-1,1) + \boldsymbol{c}_{G,H}(0,1)
\end{aligned}$$

So,

$$ER = \boldsymbol{c}_{G,H}(0,1) - \boldsymbol{c}_{G,H}(-1,1)$$

The risk premium on a generic asset is

$$RP = \boldsymbol{c}_{G,H}(-1,0) + \boldsymbol{c}_{G,H}(0,1) - \boldsymbol{c}_{G,H}(-1,1)$$

In the lognormal case, this expression becomes

$$\begin{aligned} RP &= \underbrace{-\mathbb{E}G + \frac{1}{2!}\operatorname{var} G}_{\boldsymbol{c}_{G,H}(-1,0)} + \underbrace{\mathbb{E}H + \frac{1}{2!}\operatorname{var} H}_{\boldsymbol{c}_{G,H}(0,1)} - \underbrace{\mathbb{E}(-G+H) - \frac{1}{2!}\operatorname{var}(-G+H)}_{\boldsymbol{c}_{G,H}(-1,1)} \\ &= \operatorname{cov}(G,H) \end{aligned}$$

as usual.

Proposition 7 (Multivariate results in continuous time)**.**

$$\begin{aligned} D/P &= -\boldsymbol{c}_{G,H}(-1,1) && (54) \\ R_f &= -\boldsymbol{c}_{G,H}(-1,0) && (55) \\ RP &= \boldsymbol{c}_{G,H}(0,1) + \boldsymbol{c}_{G,H}(-1,0) - \boldsymbol{c}_{G,H}(-1,1) && (56) \end{aligned}$$

The discrete-time case is very similar, and Proposition 7 holds with D/P, R_f and RP replaced by their lower case counterparts—$d/p \equiv \log(1 + D/P)$, and so on. Since dividend-price ratios are constant, the return on any asset is proportional to its dividend growth. The H terms in Proposition 7 can therefore be replaced with R, defined to be the logarithm of the asset's one-period return.

Proposition 8 (Multivariate results in discrete time)**.** *Defining R to be the logarithm of the asset in question's one period gross return, we have*

$$\begin{aligned} d/p &= -\boldsymbol{c}_{G,R}(-1,1) && (57) \\ r_f &= -\boldsymbol{c}_{G,R}(-1,0) && (58) \\ rp &= \boldsymbol{c}_{G,R}(0,1) + \boldsymbol{c}_{G,R}(-1,0) - \boldsymbol{c}_{G,R}(-1,1) && (59) \end{aligned}$$

We have expressions for dividend-price ratios and risk premia in terms of the bivariate cumulants of the log SDF and log returns. In the presence of jumps, risk premia are not determined by covariances alone, but by co-skewness, co-kurtosis and "co-cumulants" of all orders.

5.1 Heterogeneity in the presence of disasters

This section illustrates the above calculations by presenting a simple model of heterogeneity in the presence of rare disasters. Constantinides and Duffie (1996) have shown that heterogeneity of consumption processes across individuals can have asset pricing implications that appear surprising to an econometrician who uses aggregate data; for example, they show that accounting for heterogeneity may contribute to an understanding of the equity premium puzzle. On the other hand, Grossman and Shiller (1982) have shown that in a continuous-time framework in which the (heterogeneous) consumption processes of different agents follow diffusions, this effect disappears.

I attempt to resolve the tension between these two results by showing that heterogeneity matters to the extent that it is present *at times of aggregate jumps.* The presence of jumps lends a discrete-time flavor to the model which, in a sense, occupies a position intermediate between Constantinides-Duffie and Grossman-Shiller. My starting point is an assumption that agents suffer idiosyncratic shocks to consumption, though I make no serious attempt to explain why agents are unable to insure against these shocks. One story would be that agents have labor income risk which is uninsurable for moral hazard reasons.

The model is set up in such a way that all agents attach the same values to "equity", interpreted as a claim on aggregate consumption, which is subject to jumps as modelled in section 2.2 above. All agents have power utility with relative risk aversion γ.

Aggregate consumption, written C_t, is as in (29), so

$$\log \frac{C_t}{C_0} = \mu t + \sigma B_t + \sum_{j=1}^{N_t} Y_j \tag{60}$$

where, for reference, B_t is a Brownian motion, N_t is the value taken by a Poisson counting process at time t (distributed according to a Poisson distribution with parameter ωt) and Y_j are i.i.d. random variables with distribution currently left unspecified (although I assume that disasters are bad news on average, so $\mathbb{E}e^{Y_j} < 1$). Disasters occur at times when N_t increases.

The log consumption process of an individual agent, i, is determined by layering idiosyncratic shocks on top of the aggregate process specified in (60). I allow for three types of idiosyncratic shocks:

(i) a Brownian motion component, $B_{i,t}$,

(ii) idiosyncratic jumps, $X_{i,k}$, which occur at times determined by an idiosyncratic Poisson process, $N_{i,t}$, and

(iii) idiosyncratic jumps, $Y_{i,k}$, which occur at times determined by the Poisson process N_t, that is, at times of aggregate disaster.

Type (i) shocks are included only in order to demonstrate that they do not affect the risk premium. (They do, however, affect the riskless rate and consumption-wealth ratio.) Type (ii) shocks can be thought of as totally idiosyncratic shocks (to labor income, say). Type (iii) shocks are idiosyncratic in size, but hit all agents at the same time. This allows for the unarguable fact that when a major disaster occurs, some agents are affected more than others. It will turn out that while all three types of shock drive down the riskless rate and consumption-wealth ratio (relative to the homogeneous case), only shocks of type (iii) affect the risk premium.

Formally, I assume that

$$\log \frac{C_{i,t}}{C_{i,0}} = \log \frac{C_t}{C_0} + \underbrace{\sigma_1 B_{it} - \frac{1}{2}\sigma_1^2 t}_{\text{type (i)}} + \underbrace{\sum_{j=1}^{N_{i,t}} X_{i,j}}_{\text{type (ii)}} + \underbrace{\sum_{k=1}^{N_t} Y_{i,k}}_{\text{type (iii)}} \tag{61}$$

where $X_{i,j}$ and $Y_{k,l}$ are i.i.d. across i, j, k and l, and $N_{i,t}$ is a Poisson process, independent across i, with arrival rate ω_2. Finally, σ_1 and ω_2 are constant across all agents i. The upshot of these assumptions is that any two agents attach the same values to any asset whose payoffs are independent of the idiosyncratic components of their consumption processes (in particular, to equity as defined above). As in Constantinides and Duffie (1996), there is, therefore, a no-trade equilibrium with equity in which agents consume $\{C_{i,t}\}$.

Aggregate quantities are computed by summing over agents i; I assume that a law of large numbers holds so that this process is equivalent to taking an expectation over i. With this assumption, (61) is consistent with the evolution of aggregate consumption in (60) under the maintained assumption that for all i and k,

$$\mathbb{E}e^{X_{i,k}} = \mathbb{E}e^{Y_{i,k}} = 1\,. \tag{62}$$

(The drift term $-\sigma_1^2 t/2$ takes care of the type (i) piece.)

For the time being, I leave the distribution of jumps in aggregate log consumption unspecified and, throughout this section, define $\boldsymbol{m}(\theta) \equiv \mathbb{E}e^{\theta Y_j}$.[17] Similarly, the relevant details of the distribution of jumps in idiosyncratic log consumption are summarized by $\boldsymbol{m}_2(\theta) \equiv \mathbb{E}e^{\theta X_{i,k}}$ and $\boldsymbol{m}_3(\theta) \equiv \mathbb{E}e^{\theta Y_{i,j}}$.

[17] As already noted, it is assumed that $\boldsymbol{m}(1) < 1$.

The Euler equation holds for each agent i, so the price of equity, P, must satisfy

$$P = \mathbb{E} \int_0^\infty e^{-\rho t} \left(\frac{C_{i,t}}{C_{i,0}} \right)^{-\gamma} \cdot C_t \, dt \tag{63}$$

as usual. Heterogeneity matters: dropping the is in (63) is not valid.

The analysis of the previous section goes through unchanged. Any agent's consumption process gives rise to a valid stochastic discount factor,

$$\frac{M_{i,t}}{M_{i,0}} = e^{-\rho t} \left(\frac{C_{i,t}}{C_{i,0}} \right)^{-\gamma} ,$$

so I define

$$G_{i,t} \equiv -\log \frac{M_{i,t}}{M_{i,0}} = \rho + \gamma \cdot \log \frac{C_{i,t}}{C_{i,0}} \tag{64}$$

$$H_t \equiv \log \frac{C_t}{C_0} . \tag{65}$$

We can apply the results of Proposition 7 directly; I retain the i subscript in $G_{i,t}$ as a reminder that individuals, not aggregates, price assets.

By Proposition 7, we have

$$D/P = -\boldsymbol{c}_{G_i,H}(-1,1) \tag{66}$$

$$R_f = -\boldsymbol{c}_{G_i,H}(-1,0) \tag{67}$$

$$RP = \boldsymbol{c}_{G_i,H}(0,1) + \boldsymbol{c}_{G_i,H}(-1,0) - \boldsymbol{c}_{G_i,H}(-1,1) \tag{68}$$

where as usual $G_i = G_{i,1}$ and $H = H_1$. By the definition, the dividend on equity is aggregate consumption, and the price of equity is aggregate wealth, so we can also write $D/P = C/W$.

Computing the bivariate CGF of G_i and H is a simple exercise, which gives

$$\begin{aligned} \boldsymbol{c}_{G_i,H}(\theta_1,\theta_2) = {} & \rho\theta_1 + \mu\left(\gamma\theta_1 + \theta_2\right) + \frac{1}{2}\sigma^2\left(\gamma\theta_1 + \theta_2\right)^2 + \frac{1}{2}\sigma_1^2\gamma\theta_1(\gamma\theta_1 - 1) + \\ & + \omega\left[\boldsymbol{m}\left(\gamma\theta_1 + \theta_2\right)\boldsymbol{m}_3\left(\gamma\theta_1\right) - 1\right] + \omega_2\left[\boldsymbol{m}_2(\gamma\theta_1) - 1\right] \end{aligned} \tag{69}$$

The correct consumption-wealth ratio, riskless rate and risk premium on the consumption claim can be obtained from (66)–(69). An econometrician who incorrectly uses aggregate consumption in calculations of these fundamentals is implicitly imposing $\sigma_1 = 0$ and $Y_{i,k} \equiv X_{i,j} \equiv 0$, or equivalently $\boldsymbol{m}_2(\theta) = \boldsymbol{m}_3(\theta) = 1$ for all θ, in (69). The discrepancies between true fundamentals and incorrect predictions based on aggregate quantities (denoted by bars) are given by

$$C/W - \overline{C/W} = -\sigma_1^2\gamma(\gamma+1)/2 - \omega_2\left[\boldsymbol{m}_2(-\gamma) - 1\right] - \omega\,\boldsymbol{m}(1-\gamma)\left[\boldsymbol{m}_3(-\gamma) - 1\right] \tag{70}$$

$$R_f - \overline{R_f} = -\sigma_1^2\gamma(\gamma+1)/2 - \omega_2\left[\boldsymbol{m}_2(-\gamma) - 1\right] - \omega\,\boldsymbol{m}(-\gamma)\left[\boldsymbol{m}_3(-\gamma) - 1\right] \tag{71}$$

$$RP - \overline{RP} = \omega\left[\boldsymbol{m}(-\gamma) - \boldsymbol{m}(1-\gamma)\right]\left[\boldsymbol{m}_3(-\gamma) - 1\right] \tag{72}$$

To get some hold on these unwieldy expressions, I now show that whatever the distribution of idiosyncratic jumps, allowing for heterogeneity leads to a lower consumption-wealth ratio and riskless rate and to a higher risk premium than would be predicted by the same naive econometrician. In other words, I show that the expressions (70) and (71) are negative and that (72) is positive.

Proposition 9 (Asset pricing implications of heterogeneity)**.** *Heterogeneity drives down the consumption-wealth ratio and riskless rate and increases the risk premium:*

$$\begin{aligned} C/W &\leq \overline{C/W} \\ R_f &\leq \overline{R_f} \\ RP &\geq \overline{RP}\,. \end{aligned}$$

Proof. I show that (1) $\boldsymbol{m}(\theta) > 0$ for all θ, (2) $\boldsymbol{m}_j(-\gamma) > 1$ for $j = 2, 3$, and (3) $\boldsymbol{m}(-\gamma) > \boldsymbol{m}(1-\gamma)$. The result then follows by inspection of (70)–(72).

The first of these follows simply by observing that since $e^{\theta Y}$ is positive for all Y, the expectation $\boldsymbol{m}(\theta) = \mathbb{E}e^{\theta Y}$ must also be positive. To see the second in the case $j = 3$, note first that the function $f(x) = x^{-\gamma}$, where $\gamma > 0$, is convex on the positive real line, and remember that $\mathbb{E}e^{Y_{i,k}} = 1$. Then, by Jensen's inequality, we have

$$\boldsymbol{m}_3(-\gamma) = \mathbb{E}e^{-\gamma Y_{i,k}} \geq \left[\mathbb{E}e^{Y_{i,k}}\right]^{-\gamma} = 1$$

The case $j = 2$ follows by the same logic, since also $\mathbb{E}e^{X_{i,k}} = 1$.

It remains to be shown that $\boldsymbol{m}(-\gamma) > \boldsymbol{m}(1-\gamma)$. Define $\boldsymbol{\psi}(\theta) \equiv \log \boldsymbol{m}(\theta)$ to be the CGF of the aggregate jump random variable. We want to show that $\boldsymbol{\psi}(-\gamma) > \boldsymbol{\psi}(1-\gamma)$. Since I assume throughout that $\mathbb{E}e^{Y_j} < 1$, we have $\boldsymbol{m}(1) < 1$ and hence $\boldsymbol{\psi}(1) < 0$. We also have $\boldsymbol{\psi}(0) = 0$ as usual for CGFs. Since $\gamma > 0$, convexity of the CGF implies that $\boldsymbol{\psi}(-\gamma) > 0$. Suppose that $\boldsymbol{\psi}(1-\gamma) \leq 0$; then we're done, since $\boldsymbol{\psi}(-\gamma) > 0 \geq \boldsymbol{\psi}(1-\gamma)$. If not, it must be the case that $\boldsymbol{\psi}(1-\gamma) > 0$. The convexity of $\boldsymbol{\psi}(\cdot)$ then implies that $1-\gamma$ must be negative, so $\gamma > 1$. The convexity of $\boldsymbol{\psi}(\cdot)$ also entails that

$$\frac{\boldsymbol{\psi}(-\gamma)}{-\gamma} \leq \frac{\boldsymbol{\psi}(1-\gamma)}{1-\gamma}$$

so

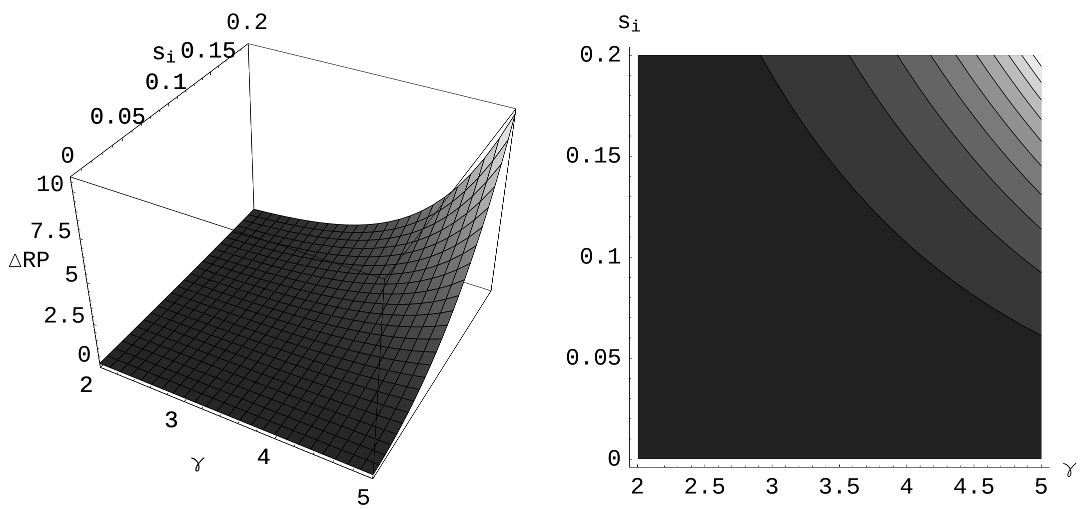
$$\boldsymbol{\psi}(-\gamma) \geq \frac{\gamma}{\gamma-1}\boldsymbol{\psi}(1-\gamma) > \boldsymbol{\psi}(1-\gamma)$$

as required; the last inequality follows from the fact that $\gamma > 1$ and $\boldsymbol{\psi}(1-\gamma) > 0$. $\square$

To get a sense of the quantitative importance of heterogeneity, suppose that aggregate and idiosyncratic—type (iii)—jumps in log consumption are Normally distributed, $Y_j \sim N(-b, s^2)$ and $Y_{i,k} \sim N(-s_i^2/2, s_i^2)$. Then $\boldsymbol{m}(\theta) = e^{-b\theta + s^2\theta^2/2}$ and $\boldsymbol{m}_3(\theta) = e^{s_i^2\theta(\theta-1)/2}$. From (72), this increases the equity premium by

$$\omega\left(e^{b\gamma+s^2\gamma^2/2} - e^{b(\gamma-1)+s^2(\gamma-1)^2/2}\right)\left(e^{s_i^2\gamma(\gamma+1)/2} - 1\right) = \Delta RP(\gamma, s_i)\,. \tag{73}$$

I plot this extra kick to the equity premium, $\Delta RP(\gamma, s_i)$, in Figure 8, using the now familiar parameter values $\omega = 0.017, b = 0.39, s = 0.25$. Heterogeneity has a significant effect for values of γ above about three.



(a) 3d plot

(b) contour plot

Figure 8: The extra kick to the risk premium due to heterogeneity at times of disaster, as a function of γ and s_i.

Using the same parameter values, Figure 9 plots the value of s_i that would boost the equity premium by 2 per cent, relative to the homogeneous case, against γ. For $\gamma \approx 5$, $s_i = 0.1$ is enough; in other words, even if a typical idiosyncratic shock (standard deviation $s_i = 0.1$) is only 40% of the magnitude of a typical aggregate shock (standard deviation $s = 0.25$), heterogeneity is quantitatively important.

6 Conclusion

Cumulant-generating functions make Epstein-Zin- and power utility-i.i.d. models tractable. The mere fact that they simplify notation makes them useful modelling tools, as shown in

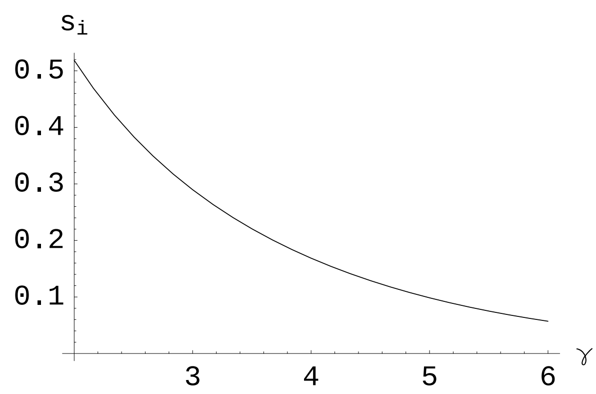


Figure 9: The value of s_i needed to give $RP - \overline{RP} = 2\%$, against γ.

the heterogeneous agent model of section 5.1. In more complicated settings (for example, Martin (2007)), it may even be easier to work with a CGF than to consider a special case such as lognormality, simply because the CGF's progress can be easily tracked through the algebra. In a sense, CGFs make it possible to carry out tractable asset-pricing calculations and nonetheless "get jumps for free".

More fundamentally, however, CGFs have useful mathematical properties. Without appealing to the convexity of a CGF, the *proof*, not just the notation, of Proposition 9 in section 5.1 would have been considerably more complicated. Convexity arguments were also employed in section 3, which derives robust restrictions on preference parameters based on observed values of the riskless rate, equity premium, consumption-wealth ratio and mean consumption growth.

These robust restrictions also exemplify the other theme of this paper, which is that it is desirable, when thinking about disasters, to try to make statements which are not sensitively dependent on the assumed pattern of higher cumulants. Section 4 showed—under assumptions more general than those made by Lucas (1987), Obstfeld (1994) or Barro (2006b)—that it is possible to use the observed consumption-wealth ratio to estimate the welfare cost of uncertainty without specifying a consumption process; and argued also that the cost is high.

7 Bibliography

Aït-Sahalia, Y., Cacho-Diaz, J. and T. R. Hurd (2006), "Portfolio Choice with Jumps: A Closed Form Solution," preprint.

Alvarez, F. and U. J. Jermann (2004), "Using Asset Prices to Measure the Costs of Business Cycles," *Journal of Political Economy*, 112:6:1223–1256.

Backus, D. K., Foresi, S. and C. I. Telmer (2001), "Affine Term Structure Models and the Forward Premium Anomaly," *Journal of Finance*, 56:1:279–304.

Barndorff-Nielsen, O. E. (1997), "Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling," *Scandinavian Journal of Statistics*, 24:1–13.

Barro, R. J. (2006a), "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121:3:823–866.

Barro, R. J. (2006b), "On the Welfare Costs of Consumption Uncertainty," preprint.

Billingsley, P. (1995), *Probability and Measure*, 3rd edition, John Wiley & Sons, New York, NY.

Campbell, J. Y. (1986), "Bond and Stock Returns in a Simple Exchange Model," *Quarterly Journal of Economics*, 101:4:785–804.

Campbell, J. Y. (2003), "Consumption-Based Asset Pricing," Chapter 13 in George Constantinides, Milton Harris and Rene Stulz eds., *Handbook of the Economics of Finance vol. IB*, North-Holland, Amsterdam, 803–887.

Carr, P. and D. Madan (1998), "Option Valuation Using the Fast Fourier Transform," *Journal of Computational Finance*, 2:61–73.

Cochrane, J. H. (2005), *Asset Pricing*, revised edition, Princeton University Press, Princeton, NJ.

Cochrane, J. H. and J. Saá-Requejo (2000), "Beyond Arbitrage: Good Deal Asset Price Bounds in Incomplete Markets," *Journal of Political Economy*, 108:1:79–119.

Cogley, T. (1990), "International Evidence on the Size of the Random Walk in Output," *Journal of Political Economy*, 98:3:501–518.

Constantinides, G. M. and D. Duffie (1996), "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 104:2:219–240.

Cont, R. and P. Tankov (2004), *Financial Modelling with Jump Processes*, Chapman & Hall/CRC, Boca Raton, FL.

Cvitanić, J., Polimenis, V. and F. Zapatero (2005), "Optimal Portfolio Allocation with Higher Moments," preprint.

Duffie, D. and L. J. Epstein (1992), "Asset Pricing with Stochastic Differential Utility," *Review of Financial Studies*, 5:3:411–436.

Duffie, D., Pan, J. and K. Singleton (2000), "Transform Analysis and Asset-Pricing for Affine Jump-Diffusions," *Econometrica*, 68:6:1343–1376.

Grossman, S. J. and R. J. Shiller (1982), "Consumption Correlatedness and Risk Measurement in Economies with Non-Traded Assets and Heterogeneous Information," *Journal of Financial Economics* 10:195–210.

Hansen, L. P. and R. Jagannathan (1991), "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99:2:225–262.

Hansen, L. P. and K. J. Singleton (1983), "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91:249–268.

Kallsen, J. (2000), "Optimal Portfolios for Exponential Lévy Processes," *Mathematical Methods of Operations Research*, 51:357–374.

Kingman, J. F. C. (1993), *Poisson Processes*, Oxford University Press, Oxford, UK.

Kocherlakota, N. R. (1990), "Disentangling the Coefficient of Relative Risk Aversion from the Elasticity of Intertemporal Substitution: An Irrelevance Result," *Journal of Finance*, 45:1:175–190.

Lentzas, G. A. (2007), "The Risk-Premium for the Non-Gaussianity of Asset Returns," working paper, Oxford University.

Lettau, M. and S. Ludvigson (2001), "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56:3:815–849.

Lucas, R. E. (1987), *Models of Business Cycles*, Basil Blackwell, Oxford, UK.

Lucas, R. E. (2003), "Macroeconomic Priorities," *American Economic Review*, 93:1–14.

Madan, D., Carr, P. and E. Chang (1998), "The Variance Gamma Process and Option Pricing," *European Finance Review*, 2:79–105.

Martin, I. W. R. (2007), "The Lucas Orchard," working paper, Harvard University.

Mehra, R. and E. C. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15:145–161.

Obstfeld, M. (1994), "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability," *European Economic Review*, 38:1471–1486.

Rietz, T. A. (1988), "The Equity Premium: A Solution," *Journal of Monetary Economics*, 22:117–131.

Sato, K. (1999), *Lévy Processes and Infinitely Divisible Distributions*, 1st edition, Cambridge University Press, Cambridge, UK.

Shaliastovich, I. and G. Tauchen (2005), "Pricing Implications of Stochastic Volatility, Business Cycle Time Change, and Non-Gaussianity," working paper, Duke University.

Stuart, A. and J. K. Ord (1994), *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory*, Edward Arnold, London, UK.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press, Cambridge, UK.

Weitzman, M. L. (2005), "Risk, Uncertainty and Asset-Pricing 'Puzzles'," preprint.

A Cumulants and cumulant-generating functions

This section lays out some important properties of cumulant-generating functions. It turns out that $\boldsymbol{c}(\theta)$ can be thought of as a power series in θ that encodes the cumulants (equivalently, moments) of consumption growth. To preview the main result, we have

$$\boldsymbol{c}(\theta) = \mu \cdot \frac{\theta}{1!} + \frac{\sigma^2\theta^2}{2!} + \text{skewness} \cdot \frac{\sigma^3\theta^3}{3!} + \text{kurtosis} \cdot \frac{\sigma^4\theta^4}{4!} + \ldots$$

Here and throughout the paper, μ and σ denote the unconditional mean and standard deviation of log consumption growth.

A.1 Definition and standard properties

Definition 3. *The* cumulants *of G are the coefficients κ_n in the power series expansion of the CGF $\boldsymbol{c}(\theta)$:*

$$\boldsymbol{c}(\theta) = \sum_{n=1}^{\infty} \frac{\kappa_n(G)\theta^n}{n!}. \tag{74}$$

It turns out that cumulants have many appealing properties, which I collect in a theorem.

Proposition 10. *We have the following properties.*

1. $\mathbb{E}G = \kappa_1$; $\operatorname{var} G = \kappa_2 \equiv \sigma^2$; *skewness* $(G) = \kappa_3/\sigma^3$; *excess kurtosis* $(G) = \kappa_4/\sigma^4$.
2. *For any two independent random variables G and H, $\kappa_n(G+H) = \kappa_n(G) + \kappa_n(H)$ and $\boldsymbol{c}_{G+H}(\theta) = \boldsymbol{c}_G(\theta) + \boldsymbol{c}_H(\theta)$.*
3. $\kappa_1(G) = \boldsymbol{c}'_G(0)$; $\kappa_2(G) = \boldsymbol{c}''_G(0)$; $\kappa_n(G) = \boldsymbol{c}^{(n)}_G(0)$.
4. *κ_n is a polynomial in the first n moments of G (and the nth moment of G is a polynomial in the first n cumulants of G).*

Proof. I only provide the outlines of proofs; for more details, see Billingsley (1995, section 9). Property 2 follows from the definitions of moment- and cumulant-generating functions, and the fact that when G and H are independent, $\mathbb{E}e^{\theta(G+H)} = \mathbb{E}e^{\theta G}\mathbb{E}e^{\theta H}$. Property 3 follows from the definition of cumulants. Properties 1 and 4 follow by noting that

$$\boldsymbol{m}(\theta) = 1 + \sum_{n=1}^{\infty} \frac{\theta^n}{n!}\mathbb{E}G^n$$

and hence that $\boldsymbol{c}(\theta)$ can be expanded as a power series in θ

$$\begin{aligned}\boldsymbol{c}(\theta) &= \log \boldsymbol{m}(\theta) \\ &= \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \left(\sum_{n=1}^{\infty} \frac{\theta^n}{n!} \mathbb{E} G^n \right)^j ,\end{aligned}$$

and then differentiating the requisite number of times with respect to θ and setting $\theta = 0$. $\square$

Thus, the CGF is a convex[18] function which passes through the origin, at which point it has slope equal to mean log consumption growth and second derivative equal to the variance of log consumption growth.

B Calculations with Epstein-Zin preferences

The Epstein-Zin first-order condition leads to the pricing formula

$$P = \mathbb{E} \sum_{1}^{\infty} e^{-\rho\vartheta t} \left(\frac{C_t}{C_0} \right)^{-\vartheta/\psi} (1 + R_{m,0\to t})^{\vartheta - 1} (C_t)^{\lambda} ,$$

where $\vartheta = (1-\gamma)/(1-1/\psi)$ and $R_{m,0\to t}$ is the cumulative return on the wealth portfolio from period 0 to period t. I assume that $\psi \neq 1$ for convenience.

Now,

$$\begin{aligned}1 + R_{m,s-1\to s} &= \frac{C_s + W_s}{W_{s-1}} \\ &= \frac{C_s}{C_{s-1}} \left(\frac{C_{s-1}}{W_{s-1}} + \frac{W_s}{C_s} \frac{C_{s-1}}{W_{s-1}} \right) \\ &= \frac{C_s}{C_{s-1}} e^{\nu} ,\end{aligned}$$

where the last equality follows by making the assumption—provisional for the time being, but subsequently shown to be correct—that the consumption-wealth ratio is constant. I have defined $1 + C/W \equiv e^{\nu}$. It follows, then, that

$$1 + R_{m,0\to t} = \frac{C_t}{C_0} e^{\nu t} ,$$

[18] As shown in the main text, Fact 1.

and hence that

$$
\begin{aligned}
P &= (C_0)^{\lambda} \cdot \mathbb{E} \sum_1^{\infty} e^{-\rho\vartheta t} \left(\frac{C_t}{C_0}\right)^{\lambda-\vartheta/\psi} \left(\frac{C_t}{C_0}\right)^{\vartheta-1} e^{\nu(\vartheta-1)t} \\
&= (C_0)^{\lambda} \cdot \sum_1^{\infty} e^{-[\rho\vartheta+\nu(1-\vartheta)-\boldsymbol{c}(\lambda-\gamma)]t} \\
&= \frac{(C_0)^{\lambda}}{e^{\rho\vartheta+\nu(1-\vartheta)-\boldsymbol{c}(\lambda-\gamma)}-1},
\end{aligned}
$$

and so, finally, that

$$
\frac{D}{P} = e^{\rho\vartheta+\nu(1-\vartheta)-\boldsymbol{c}(\lambda-\gamma)} - 1 .
$$

Defining d/p as usual,

$$
d/p = \rho\vartheta + \nu(1-\vartheta) - \boldsymbol{c}(\lambda-\gamma) . \tag{75}
$$

Setting $\lambda = 1$, we get an expression for $c/w \equiv \nu$ which can be solved for ν:

$$
\nu = c/w = \rho\vartheta + \nu(1-\vartheta) - \boldsymbol{c}(1-\gamma) ,
$$

from which it follows that

$$
\nu = \rho - \boldsymbol{c}(1-\gamma) \cdot \frac{1-\psi}{\psi(\gamma-1)} .
$$

Note that this exercise confirms the provisional assumption made above that ν is constant.

Substituting back into (75), we have

$$
dp = \rho - \frac{1-\psi\gamma}{\psi(\gamma-1)} \boldsymbol{c}(1-\gamma) - \boldsymbol{c}(\lambda-\gamma) .
$$

We also have, as before, that

$$
1 + R_{t+1} = \frac{D_{t+1}}{D_t} \left(e^{\rho\vartheta+\nu(1-\vartheta)-\boldsymbol{c}(\lambda-\gamma)}\right) ,
$$

so

$$
er = \rho\vartheta + \nu(1-\vartheta) + \boldsymbol{c}(\lambda) - \boldsymbol{c}(\lambda-\gamma) .
$$

To summarize, we have

$$
\begin{aligned}
r_f &= \rho - \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\left(\frac{1}{\vartheta} - 1\right) \\
c/w &= \rho - \boldsymbol{c}(1-\gamma)/\vartheta \\
rp &= \boldsymbol{c}(1) + \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma) .
\end{aligned}
$$

The objective function at time 0 satisfies

$$
(U_0)^{(1-\gamma)/\vartheta} = \left(1-e^{-\rho}\right)(C_0)^{(1-\gamma)/\vartheta} + e^{-\rho}\left(\mathbb{E}(U_1)^{1-\gamma}\right)^{1/\vartheta}
$$

or

$$a_0^{(1-\gamma)/\vartheta} = 1 - e^{-\rho} + e^{-\rho}\mathbb{E}\left[\left(\frac{C_1}{C_0}\right)^{1-\gamma} a_1^{1-\gamma}\right]^{1/\vartheta}, \tag{76}$$

where I have defined $a_i \equiv U_i/C_i$.

I now conjecture that $a_i = a$, some constant, solves (76). If so,

$$a^{(1-\gamma)/\vartheta} = 1 - e^{-\rho} + e^{-\rho}a^{(1-\gamma)/\vartheta}e^{\mathbf{c}(1-\gamma)/\vartheta},$$

from which it follows that

$$a = \left(\frac{1-e^{-\rho}}{1-e^{-\rho+\mathbf{c}(1-\gamma)/\vartheta}}\right)^{\vartheta/(1-\gamma)},$$

which confirms the conjecture that a was constant. Hence,

$$U_0 = C_0 \cdot \left(\frac{e^{\rho}-1}{e^{\rho}-e^{\mathbf{c}(1-\gamma)/\vartheta}}\right)^{\vartheta/(1-\gamma)}.$$

The cost of all uncertainty, ϕ, solves the equation

$$(1+\phi)\, C_0 \cdot \left(\frac{e^{\rho}-1}{e^{\rho}-e^{\mathbf{c}(1-\gamma)/\vartheta}}\right)^{\vartheta/(1-\gamma)} = C_0 \left(\frac{e^{\rho}-1}{e^{\rho}-e^{\mathbf{c}(1)\cdot(1-\gamma)/\vartheta}}\right)^{\vartheta/(1-\gamma)},$$

from which (48) follows.

Similarly, ϕ_α solves

$$(1+\phi_\alpha)C_0 \cdot \left(\frac{e^{\rho}-1}{e^{\rho}-e^{\mathbf{c}(1-\gamma)/\vartheta}}\right)^{\vartheta/(1-\gamma)} = C_0 \cdot \left(\frac{e^{\rho}-1}{e^{\rho}-e^{\widetilde{\mathbf{c}}(1-\gamma)/\vartheta}}\right)^{\vartheta/(1-\gamma)},$$

and after substituting in for $\widetilde{\mathbf{c}}(\theta)$ from equation (49), we obtain the expression (51).

C Derivation of results in continuous time

Definition 4. *A real-valued stochastic process $(L_t)_{t\geq 0}$ with $L_0 = 0$ is a* Lévy process *if*

1. *With probability one, L_t is right continuous on $[0,\infty)$, with left limits on $(0,\infty)$.*

2. *For any $n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \ldots < t_n$, the random variables $L_{t_j} - L_{t_{j-1}}$ are independent for $j = 1, \ldots, n$.*

3. *The probability distribution of $L_{t+h} - L_t$ does not depend on t.*

4. *For all $t \geq 0$ and $\varepsilon > 0$, $\lim_{s\to t}\mathbb{P}(|X_s - X_t| > \varepsilon) = 0$.*

C.1 Asset pricing calculations

The price of a claim to the dividend stream $\{D_t\} \equiv \{(C_t)^\lambda\}$ is

$$\begin{aligned}
P_\lambda &= \mathbb{E}_0\left(\int_{t=0}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma}(C_t)^\lambda\,dt\right) \\
&= D_\lambda\,\mathbb{E}_0\left(\int_{t=0}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-(\gamma-\lambda)}dt\right) \\
&= D_\lambda\int_{t=0}^{\infty} e^{-\rho t}\boldsymbol{m}_{G_t}(\lambda-\gamma)\,dt \\
&\overset{(a)}{=} D_\lambda\int_{t=0}^{\infty} e^{-\rho t}\Big(\boldsymbol{m}(\lambda-\gamma)\Big)^t\,dt \\
&= D_\lambda\int_{t=0}^{\infty} e^{-\{\rho-\boldsymbol{c}(\lambda-\gamma)\}t}\,dt \\
&= \frac{D_\lambda}{\rho-\boldsymbol{c}(\lambda-\gamma)}
\end{aligned}$$

The critical property (27) satisfied by Lévy processes manifests itself in equality (a). The riskless rate and consumption-wealth ratio can be calculated by substituting $\lambda = 0$ and $\lambda = 1$ respectively.

From the definition in the main text,

$$\begin{aligned}
ER_\lambda &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\cdot\mathbb{E}\left(\left(\frac{C_{t+\Delta t}}{C_t}\right)^\lambda-1\right)+\rho-\boldsymbol{c}(\lambda-\gamma) \\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\cdot\mathbb{E}\left(e^{\lambda G_{\Delta t}}-1\right)+\rho-\boldsymbol{c}(\lambda-\gamma) \\
&= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\cdot\left(e^{\boldsymbol{c}(\lambda)\Delta t}-1\right)+\rho-\boldsymbol{c}(\lambda-\gamma) \\
&= \boldsymbol{c}(\lambda)+\rho-\boldsymbol{c}(\lambda-\gamma)
\end{aligned} \tag{77}$$

D The relationship between cumulants and moments

Stuart and Ord (1994) list the univariate and bivariate cumulants in terms of central moments; I reproduce the first few of each below.

D.1 The univariate case

Define $\mu_i \equiv \mathbb{E}\left[(G - \mathbb{E}G)^i\right]$.

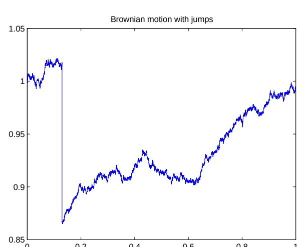
$$
\begin{aligned}
\kappa_2 &= \mu_2 \\
\kappa_3 &= \mu_3 \\
\kappa_4 &= \mu_4 - 3(\mu_2)^2 \\
\kappa_5 &= \mu_5 - 10\mu_3\mu_2 \\
\kappa_6 &= \mu_6 - 15\mu_4\mu_2 - 10(\mu_3)^2 + 30(\mu_2)^3
\end{aligned}
$$

D.2 The bivariate case

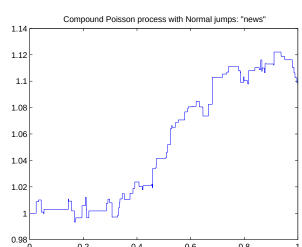
Define $\mu_{ij} \equiv \mathbb{E}\left[(G - \mathbb{E}G)^i\,(H - \mathbb{E}H)^j\right]$. When i or j is equal to zero, the bivariate cumulant reduces to a univariate cumulant.

$$
\begin{aligned}
\kappa_{00} &= 0 \\
\kappa_{11} &= \mu_{11} \\
\kappa_{21} &= \mu_{21} \\
\kappa_{31} &= \mu_{31} - 3\mu_{20}\mu_{11}\,; \;\; \kappa_{22} = \mu_{22} - \mu_{20}\mu_{02} - 2(\mu_{11})^2 \\
\kappa_{41} &= \mu_{41} - 4\mu_{30}\mu_{11} - 6\mu_{21}\mu_{20}\,; \;\; \kappa_{32} = \mu_{32} - \mu_{30}\mu_{02} - 6\mu_{21}\mu_{11} - 3\mu_{20}\mu_{12}
\end{aligned}
$$

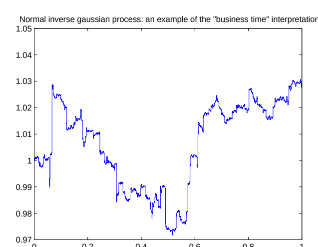
E Examples of Lévy processes



(a) Disasters.



(b) News.



(c) Business time.

Figure 10: Different models of consumption, with interpretations.

Consumption might be thought of as following a Lévy process in the following possible models of the world.

1. *Disasters.* G_t is a geometric Brownian motion with infrequent jumps: occasionally, there is a war or a meteorite strike which causes consumption to be multiplied by some random variable (perhaps with mean much less than one). See Figure 10a.

2. *News.* G_t is a compound Poisson process: consumption adjusts in response to information arrivals which are distributed independently over time, and so follow a Poisson process (perhaps with a fast arrival rate). See Figure 10b.

3. *Business time; inattention or inability to adjust consumption.* G_t is a time-changed process: in the terminology of example 6, Q_t may capture inattention or inability on the part of the consumer to adjust consumption. Suppose, for example, that Q_t is a jump process. Between jumps, Q_t is constant: consumption is "stuck." Figure 10c shows an example in which log consumption is a Brownian motion in terms of business time, but in terms of real time is a Normal inverse Gaussian process (see Barndorff-Nielsen (1997) for more details).