Disasters and the Lucas Orchard

Essays in Finance and Macroeconomics

A dissertation presented

by

Ian William Richard Martin

to

The Department of Economics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Economics

Harvard University
Cambridge, Massachusetts

February 2008

Dissertation Advisor: Professor John Campbell Ian William Richard Martin

Disasters and the Lucas Orchard:
Essays in Finance and Macroeconomics

ABSTRACT

This dissertation consists of three chapters linked by a common thread, namely the impact of disasters on financial markets. In Chapter 1, 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—or, equivalently, cumulants—of consumption growth is encoded in the *cumulant-generating function* (CGF). 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 require calibration of a consumption process. First, I use properties of the CGF to derive restrictions on the time-preference rate and elasticity of intertemporal substitution in terms of the equity premium, riskless rate, and consumption-wealth ratio. 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.

In Chapter 2, I investigate the properties of a continuous-time endowment economy in which a representative agent with power utility consumes the dividends of

multiple assets. The assets are Lucas trees; a collection of Lucas trees is a Lucas orchard. Prices, expected returns, and interest rates are determined endogenously on the basis of exogenous dividends. The model replicates various features of the data. Assets with independent dividends exhibit comovement in returns. Jumps spread across assets. Assets with high price-dividend ratios have low risk premia. Small assets exhibit momentum. High yield spreads forecast high excess returns on long term bonds and on the market. Special attention is paid to the behavior of very small assets which, in the limit, may comove endogenously and hence earn positive risk premia even if their dividends are independent of the rest of the economy.

In Chapter 3, I explore the long-run implications of the fundamental equation of asset pricing, which states that the expected time- and risk-adjusted cumulative return on any asset equals one at all horizons. I arrive, via a theorem of Kakutani, at an apparently paradoxical result: for a typical asset, the realized time- and risk-adjusted cumulative return tends to zero with probability one. As a special case, this result strengthens the familiar fact that the growth-optimal portfolio outperforms other assets at long horizons. The apparent paradox is resolved by a further result, which shows that the long-run value of a non-growth-optimal asset is driven by the possibility of extremely good news at the level of the individual asset or extremely bad news at the aggregate level.

TABLE OF CONTENTS

Abstract iii

Acknowledgments vii

1. Consumption-Based Asset Pricing with Higher Cumulants 1
1.1 Asset-pricing fundamentals and the CGF 5
1.1.1 The Gordon growth model 13
1.1.2 The asymptotic lognormality of consumption 15
1.2 The continuous-time case 16
1.2.1 Calculations 19
1.2.2 A concrete example: disasters 20
1.3 Restrictions on preference parameters 26
1.3.1 Hansen-Jagannathan and good-deal bounds 29
1.4 The cost of consumption fluctuations 33
1.4.1 The elimination of all uncertainty 35
1.4.2 A reduction in the variance of consumption growth 41
1.5 The multivariate case and heterogeneity 43
1.5.1 Heterogeneity in the presence of disasters 46
1.6 Conclusion 52

2. The Lucas Orchard 54
2.1 Setup 60
2.1.1 The stochastic discount factor 60
2.1.2 Dividend processes 61
2.1.3 Closing the model 63
2.2 A simple example 63
2.3 General solution in the two-asset case 65
2.3.1 Prices 65
2.3.2 Returns 69
2.3.3 Interest rates 69
2.4 The Brownian motion case 71
2.5 Two calibrations 73
2.5.1 Dividends follow geometric Brownian motions 73
2.5.2 Dividends are subject to occasional disasters 81
2.6 Equilibrium pricing of small assets 89
2.7 N assets 95

2.7.1 The robustness of contagion and flight-to-quality 98
2.8 Conclusion 100

3. Asset Returns in the Long Run 101
3.1 The main result 103
3.2 The growth-optimal portfolio 105
3.2.1 Examples 107
3.2.2 Relationship with previous results 109
3.3 Where's the value in a Type 2 asset? 110
3.4 An example 112
3.5 Conclusion 115

Appendix 116

A. Appendix to Consumption-Based Asset Pricing with Higher Cumulants . . 117
A.1 Cumulants and cumulant-generating functions 117
A.1.1 Definition and standard properties 117
A.2 Calculations with Epstein-Zin preferences 119
A.3 Derivation of results in continuous time 122
A.3.1 Asset pricing calculations 123
A.4 The relationship between cumulants and moments 123
A.4.1 The univariate case 124
A.4.2 The bivariate case 124

B. Appendix to The Lucas Orchard 125
B.1 General solution in the two-asset case 125
B.1.1 Preliminary mathematical results 125
B.1.2 Prices 129
B.1.3 Returns 131
B.1.4 Real interest rates 132
B.2 The ridge property 133
B.3 The Brownian motion case 138
B.3.1 Simple special cases with symmetric Brownian motions 145
B.4 Small asset asymptotics 147
B.5 The N-asset case 159
B.5.1 The Fourier transform $\mathscr{F}_\gamma^N(\boldsymbol{v})$ 159
B.5.2 The expectation 162
B.5.3 Prices 164
B.5.4 Returns 165
B.5.5 Interest rates 166
B.5.6 A final change of variables 167
B.6 Some results from complex analysis 168

Bibliography 172

ACKNOWLEDGMENTS

I cannot imagine having had a better advisor than John Campbell. His influence on me is apparent throughout this dissertation. I am enormously grateful to him.

I was also fortunate to receive the advice and support of Robert Barro and Jeremy Stein. The thread that links my three chapters—the impact of disasters on financial markets—is a consequence of Robert's recent interest in the subject. I thank them both.

I would like to acknowledge financial support from Harvard University, the Center for Retirement Research at Boston College, and the Institute for Humane Studies; and to express my gratitude to the Federal Reserve Bank of New York for hosting me during Summer 2007.

I have Thomas Baranga and Anthony Niblett to thank for a fun working environment over the last few years. I will miss G9!

This thesis is dedicated to my family, and to Hannah, with love.

1. CONSUMPTION-BASED ASSET PRICING WITH HIGHER CUMULANTS

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.[1] 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 this chapter 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[2] 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,

---

[1] I thank Dave Backus, Robert Barro, Emmanuel Farhi, Xavier Gabaix, Simon Gilchrist, Francois Gourio, Greg Mankiw, Anthony Niblett, Steve Ross, Jeremy Stein, Adrien Verdelhan, Martin Weitzman and, in particular, John Campbell for their comments.

[2] 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.

riskless rate, consumption-wealth ratio[3] 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 chapter.

The first few cumulants of consumption growth can in principle be estimated from consumption data, though this approach is not taken here 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.[4] The i.i.d. growth assumption is replaced by its continuous-time analogue: 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

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

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

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.[5] 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 1.4.) These restrictions depend only on the Epstein-Zin-i.i.d. assump-

[5] I illustrate this point with more reasonable numbers in section 1.2.2 below, in which I consider the effect of perturbing parameters in a continuous-time disaster model.

tions 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 (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 perceived by the representative agent.

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 view 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%.

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)). Lustig, Van Nieuwerburgh and Verdelhan (2008) present estimates of the wealth-consumption ratio. 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 approach taken here.

1.1 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.,[6] and the moment-generating function of G (defined below) exists on the interval $[-\gamma, 1]$.[7]

Assumption A1 allows risk aversion γ to be disentangled from the elasticity of intertemporal substitution ψ. To keep things simple, those calculations that appear

[6] An alternative, weaker, assumption is that the representative agent *perceives* himself as having i.i.d. consumption growth and prices assets accordingly; the results of the chapter then go through without modification. For example, the cost of uncertainty, discussed in Section 1.4, depends on the probability distribution perceived by the representative agent.

[7] 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).

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

$$\mathbb{E}\sum_{t=0}^{\infty} e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} \quad \text{if } \gamma \neq 1\,, \quad \text{or} \quad \mathbb{E}\sum_{t=0}^{\infty} e^{-\rho t}\log C_t \quad \text{if } \gamma = 1\,. \tag{1.1}$$

Results for the more general Epstein-Zin case are reported and discussed in the main text, but calculations and proofs are relegated to Appendix A.2.

Assumption A2 is strong, and it is essential for the calculations of this chapter. 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

$$\mathbb{E}\sum_{t=0}^{\infty}\left|e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma}\right| < \infty \quad \text{if } \gamma \neq 1. \tag{1.2}$$

I discuss this requirement further below.

The Euler equation relates the price of an asset this period to the payoff next period. Expectations are calculated with respect to the measure perceived by the representative agent:

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

Iterating forward, we get

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

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

$$P(D) = \mathbb{E}\left(\sum_{t=1}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma} D_t\right). \tag{1.3}$$

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 (1.3),

$$\begin{aligned} P_\lambda &= \mathbb{E}\left(\sum_{t=1}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma}(C_t)^\lambda\right) \\ &= (C_0)^\lambda \sum_{t=1}^{\infty} e^{-\rho t}\mathbb{E}\left(\left(\frac{C_t}{C_0}\right)^{\lambda-\gamma}\right) \\ &= D_\lambda \sum_{t=1}^{\infty} e^{-\rho t}\mathbb{E}\left(e^{(\lambda-\gamma)G_t}\right) \\ &= D_\lambda \sum_{t=1}^{\infty} e^{-\rho t}\left(\mathbb{E}\left(e^{(\lambda-\gamma)G}\right)\right)^t. \end{aligned} \tag{1.4}$$

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 1.1. *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*

$$\boldsymbol{m}(\theta) \equiv \mathbb{E}\exp(\theta G) \tag{1.5}$$

$$\boldsymbol{c}(\theta) \equiv \log \boldsymbol{m}(\theta), \tag{1.6}$$

for all θ for which the expectation in (1.5) is finite.

In the particular application of this chapter, 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.1; 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, (1.4) 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 $d_\lambda/p_\lambda \equiv \log(1 + D_\lambda/P_\lambda)$.[8] Then,

$$d_\lambda/p_\lambda = \rho - \boldsymbol{c}(\lambda - \gamma) \tag{1.7}$$

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 (1.2) 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)

$$\begin{aligned} 1 + R_{t+1} &= \frac{D_{t+1} + P_{t+1}}{P_t} \\ &= \frac{P_{t+1}}{P_t}\left(1 + \frac{D_{t+1}}{P_{t+1}}\right) \\ &= \frac{D_{t+1}}{D_t}\left(e^{\rho - \boldsymbol{c}(\lambda-\gamma)}\right) \end{aligned} \tag{1.8}$$

and thus the expected gross return is

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

[8] 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 $\log D/P$.

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 1.1 (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*

$$r_f = \rho - \boldsymbol{c}(-\gamma) \tag{1.9}$$

$$c/w = \rho - \boldsymbol{c}(1-\gamma) \tag{1.10}$$

$$rp = \boldsymbol{c}(1) + \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\,. \tag{1.11}$$

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

$$r_f = \rho - \sum_{n=1}^{\infty} \frac{\kappa_n(-\gamma)^n}{n!} \tag{1.12}$$

$$c/w = \rho - \sum_{n=1}^{\infty} \frac{\kappa_n(1-\gamma)^n}{n!} \tag{1.13}$$

$$rp = \sum_{n=2}^{\infty} \frac{\kappa_n}{n!} \cdot \Big\{1 + (-\gamma)^n - (1-\gamma)^n\Big\}. \tag{1.14}$$

Writing the first few terms of the series out more explicitly, (1.12) 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

$$r_f = \rho + \mu\gamma - \frac{1}{2}\sigma^2\gamma^2 + \frac{\text{skewness}}{3!}\sigma^3\gamma^3 - \frac{\text{excess kurtosis}}{4!}\sigma^4\gamma^4 + \text{higher order terms}\,. \tag{1.15}$$

In the lognormal case, the skewness, excess kurtosis and all higher cumulants are zero, so (1.15) 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 (1.13) can be rewritten as

$$\begin{aligned} c/w &= \rho + \mu(\gamma-1) - \frac{1}{2}\sigma^2(\gamma-1)^2 + \frac{\text{skewness}}{3!}\sigma^3(\gamma-1)^3 - \\ &\quad - \frac{\text{excess kurtosis}}{4!}\sigma^4(\gamma-1)^4 + \text{higher order terms}\,. \end{aligned} \tag{1.16}$$

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 (1.14) becomes

$$\begin{aligned} rp &= \gamma\sigma^2 + \frac{\text{skewness}}{3!}\sigma^3\left(1-\gamma^3-(1-\gamma)^3\right) + \\ &\quad + \frac{\text{excess kurtosis}}{4!}\sigma^4\left(1+\gamma^4-(1-\gamma)^4\right) + \text{higher order terms}\,. \end{aligned} \tag{1.17}$$

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 1.1 to allow for Epstein-Zin preferences.

Proposition 1.2 (Fundamental quantities, Epstein-Zin case)**.** *Defining* $\vartheta \equiv (1-\gamma)/(1-1/\psi)$*, we have*

$$r_f = \rho - \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\left(\frac{1}{\vartheta}-1\right) \tag{1.18}$$

$$c/w = \rho - \boldsymbol{c}(1-\gamma)/\vartheta \tag{1.19}$$

$$rp = \boldsymbol{c}(1) + \boldsymbol{c}(-\gamma) - \boldsymbol{c}(1-\gamma)\,, \tag{1.20}$$

and the obvious counterparts of (1.12)–(1.14) *which result on expanding the CGFs in* (1.18)–(1.20) *as power series.*

Proof. See Appendix A.2. □

Equation (1.20) 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 1.1 and 1.2 is that expressions (1.12)–(1.14), 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.

1.1.1 The Gordon growth model

From equations (1.18)–(1.20), we see that

$$c/w = rp + r_f - \boldsymbol{c}(1) \tag{1.21}$$

or, more generally, that

$$d_\lambda/p_\lambda = er_\lambda - \boldsymbol{c}(\lambda)\,. \tag{1.22}$$

This is a version of the traditional Gordon growth model. (For example, the last term of (1.21), $\boldsymbol{c}(1) = \log \mathbb{E} C_{t+1}/C_t$, 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 (1.8) 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

$$D_t/P_t = (R-\Gamma)/(1+\Gamma) \tag{1.23}$$

or, in classic Gordon growth model terms,

$$\frac{\mathbb{E}_t D_{t+1}}{P_t} = R - \Gamma. \tag{1.24}$$

To recover (1.22) from (1.23), 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 (1.21).

This observation, in conjunction with equations (1.18)–(1.20), 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 (1.21), 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 (1.20), given knowledge of $\boldsymbol{c}(\cdot)$. On the other hand, knowledge of the riskless rate leaves ρ and ψ indeterminate in equation (1.18), 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 2, the use of Epstein-Zin preferences aids the interpretation of results.

1.1.2 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:[9] 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

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

[9] 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.

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

$$\begin{aligned} r_f &= \rho - \sum_{n=1}^{\infty} \frac{\kappa_n (-\gamma)^n}{n!} \\ &= \rho - \sum_{n=1}^{\infty} \frac{SFC_n \sigma^n (-\gamma)^n}{n!} \end{aligned} \tag{1.25}$$

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.

1.2 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[10]

$$\mathbb{E}\int_{t=0}^{\infty} e^{-\rho t}\frac{C_t^{1-\gamma}}{1-\gamma} \quad \text{if } \gamma \neq 1\,, \quad \text{or} \quad \mathbb{E}\int_{t=0}^{\infty} e^{-\rho t}\log C_t \quad \text{if } \gamma = 1 \tag{1.26}$$

A2c The log consumption path, G_t, of the representative agent follows a Lévy process (defined in Appendix A.3), and $\boldsymbol{m}(\theta)$ exists for θ in $[-\gamma, 1]$.

As before, we need a condition that ensures finiteness of (1.26); 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

$$\mathbb{E}e^{\theta G_t} = \left(\mathbb{E}e^{\theta G}\right)^t \tag{1.27}$$

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, (1.27) 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

[10] 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)).

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 $\omega > 0$.

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.[11]

[11] 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

1.2.1 Calculations

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

$$\begin{aligned} P_\lambda &= \mathbb{E}_0\left(\int_{t=0}^{\infty} e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma}(C_t)^\lambda\, dt\right) \\ &= \frac{D_\lambda}{\rho - \boldsymbol{c}(\lambda-\gamma)} \end{aligned} \tag{1.28}$$

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 1.3 (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*

---

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.

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 A.3. □

1.2.2 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

$$G_t = \widetilde{\mu} t + \sigma_B B_t + \sum_{i=1}^{N(t)} Y_i \tag{1.29}$$

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 (1.29).[12] 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 $t = 1$:

$$\begin{aligned}\mathbb{E}\exp\left\{\theta \sum_{i=1}^{N(1)} Y_i\right\} &= \sum_{0}^{\infty} \frac{e^{-\omega}\omega^n}{n!} \mathbb{E}\exp\left\{\theta \sum_{1}^{n} Y_i\right\} \\ &= \sum_{0}^{\infty} \frac{e^{-\omega}\omega^n}{n!} \left[\mathbb{E}\exp\{\theta Y_1\}\right]^n \\ &= \exp\left\{\omega\left(\mathbb{E}e^{\theta Y_1} - 1\right)\right\} \\ &= \exp\left\{\omega\left(\boldsymbol{m}_{Y_1}(\theta) - 1\right)\right\},\end{aligned}$$

Finally, we have

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

[12] 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 (1.29); I include them out of deference to the previous literature.

and so

$$\boldsymbol{c}(\theta) = \widetilde{\mu}\theta + \sigma_B^2\theta^2/2 + \omega\left(\boldsymbol{m}_{Y_1}(\theta) - 1\right) . \tag{1.30}$$

The cumulants can be read off from the CGF (1.30):

$$\begin{aligned} \kappa_n(G) &= \boldsymbol{c}^{(n)}(0) \\ &= \begin{cases} \widetilde{\mu} + \omega\,\mathbb{E}Y & n = 1 \\ \sigma_B^2 + \omega\,\mathbb{E}Y^2 & n = 2 \\ \omega\,\mathbb{E}Y^n & n \geq 3 \end{cases} \end{aligned} \tag{1.31}$$

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

$$\boldsymbol{c}(\theta) = \widetilde{\mu}\theta + \frac{1}{2}\sigma_B^2\theta^2 + \omega\left(e^{-\theta b + \frac{1}{2}\theta^2 s^2} - 1\right) . \tag{1.32}$$

Figure 1.1 shows the CGF of (1.32) 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

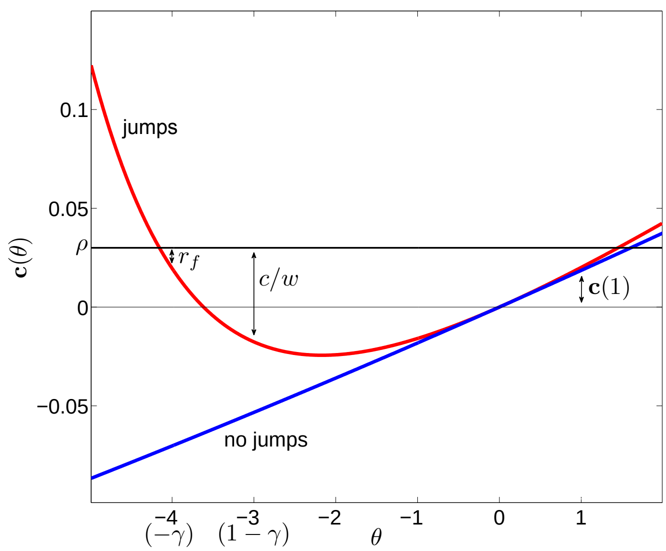


Figure 1.1: The CGF in equation (1.32) shown with and without ($\omega = 0$) jumps. The figure assumes that $\gamma = 4$.

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 1.2.

The standard lognormal model predicts a counterfactually high riskless rate; in Figure 1.1, 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 1.1, 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 1.1, we obtain Figure 1.3, 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

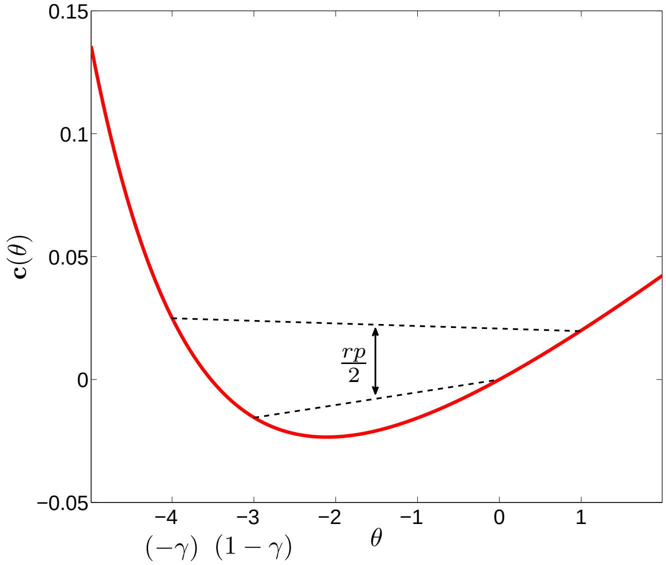


Figure 1.2: The risk premium. The figure assumes that $\gamma = 4$.

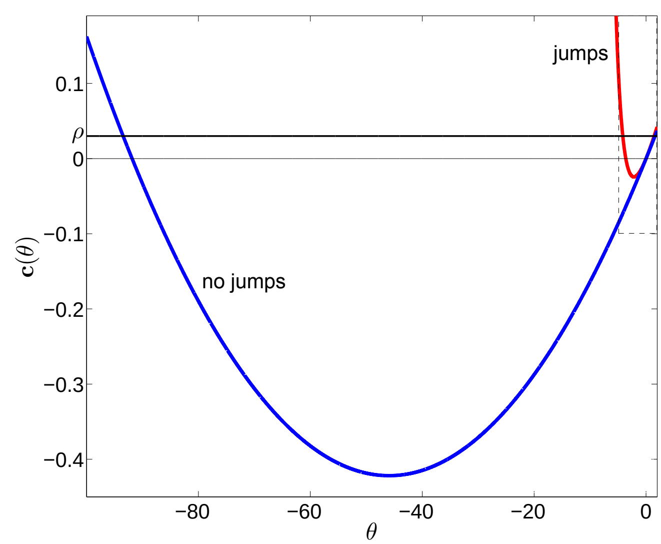


Figure 1.3: Zooming out to see the equity premium and riskless rate puzzles. The dashed box in the upper right-hand corner is the boundary of the region plotted in Figure 1.1.

visible for θ greater than about -5. The jump-free lognormal CGF has incredibly low curvature. For a realistic riskless rate and equity premium, the model requires a risk aversion above 80.

With the explicit expression (1.32) for the CGF in hand, it is easy to investigate the sensitivity of a disaster model's predictions to the parameter values assumed. Table 1.1 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.

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

As before, the CGF can also be thought of as a power series in θ. Table 1.2 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.

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 1.2: The impact of approximating the disaster model by truncating at the nth cumulant. All parameters as in baseline case of Table 1.1.

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 (1.21). 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 1.3 summarizes these assumptions.

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 1.3: Assumed values of the observables.

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 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 1.1. *CGFs are convex.*

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

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

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 1.1.

Proposition 1.4. *In the power utility case, we have*

$$\boldsymbol{c}(1) - rp \leq \frac{c/w - \rho}{\gamma - 1} \leq \boldsymbol{c}(1) \tag{1.33}$$

In the Epstein-Zin case, we have

$$\boldsymbol{c}(1) - rp \leq \frac{c/w - \rho}{1/\psi - 1} \leq \boldsymbol{c}(1) \tag{1.34}$$

Proof. From equation (1.19) 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

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

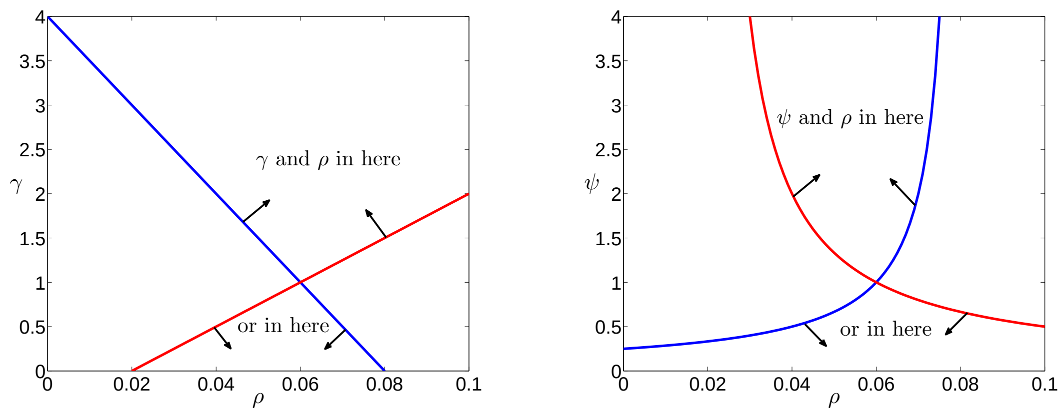
After some rearrangement of the left-hand inequality using (1.18) and (1.19), this gives (1.34). Equation (1.33) follows since $\gamma = 1/\psi$ in the power utility case. □

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

$$4 - \frac{1}{\psi} \leq 50\rho \leq 1 + \frac{2}{\psi} \quad \text{if} \quad \psi \leq 1$$
$$1 + \frac{2}{\psi} \leq 50\rho \leq 4 - \frac{1}{\psi} \quad \text{if} \quad \psi \geq 1\,.$$

Figures 1.4a and 1.4b 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) Power utility case: γ and ρ

(b) Epstein-Zin case: ψ and ρ

Figure 1.4: Parameter restrictions for i.i.d. models with $rp = 6\%$, $r_f = 2\%$ and log expected consumption growth of 2%.

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.

1.3.1 Hansen-Jagannathan and good-deal bounds

The restrictions in Proposition 1.4 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:

$$SR \leq \frac{\sigma(M)}{\mathbb{E}M}\,. \tag{1.35}$$

In the Epstein-Zin-i.i.d. setting, the right-hand side of (1.35) becomes

$$\begin{aligned}\frac{\sigma(M)}{\mathbb{E}M} &= \sqrt{\frac{\mathbb{E}M^2}{(\mathbb{E}M)^2}-1} \\ &= \sqrt{e^{\boldsymbol{c}(-2\gamma)-2\boldsymbol{c}(-\gamma)}-1}\,;\end{aligned} \tag{1.36}$$

combining (1.35) and (1.36), we obtain a Hansen-Jagannathan bound translated into CGF notation:

$$\log\left(1+SR^2\right) \leq \boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma)\,. \tag{1.37}$$

Cochrane and Saá-Requejo (2000) observe that inequality (1.35) suggests a natural way to restrict asset-pricing models. Suppose that $\sigma(M)/\mathbb{E}M \leq h$; then (1.35) 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

$$\boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma) \leq \log\left(1+h^2\right) \tag{1.38}$$

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 1.2.2. 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 (1.38) 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 (1.37) improves on a conclusion which follows from the convexity of the CGF, namely, that

$$0 \leq \boldsymbol{c}(-2\gamma) - 2\boldsymbol{c}(-\gamma)\,. \tag{1.39}$$

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

$$\frac{\boldsymbol{c}(-\gamma) - \boldsymbol{c}(-2\gamma)}{\gamma} \leq \frac{\boldsymbol{c}(0) - \boldsymbol{c}(-\gamma)}{\gamma} \tag{1.40}$$

from which (1.39) follows immediately, given that $\boldsymbol{c}(0) = 0$. Combining (1.38) and (1.39), we obtain the somewhat underwhelming result that

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

However, we can sharpen (1.39) 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

$$\frac{\sigma(M)}{\mathbb{E}M} \geq \sqrt{e^{(\gamma-1)(c/w - r_f) + \vartheta(c/w-\rho)} - 1}\,. \tag{1.41}$$

The good deal bound therefore implies that

$$(\gamma - 1)(c/w - r_f) + \vartheta(c/w - \rho) \leq \log\left(1 + h^2\right)\,. \tag{1.42}$$

Working with the power utility case for simplicity ($\vartheta = 1$) and setting $c/w = 0.06, r_f = 0.02$, Figure 1.5 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 (1.42) then implies the simple restriction

$$\gamma \leq 1 + \frac{\log\left(1+h^2\right)}{c/w - r_f}\,. \tag{1.43}$$

(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 (1.42) and (1.43) is that they do not require

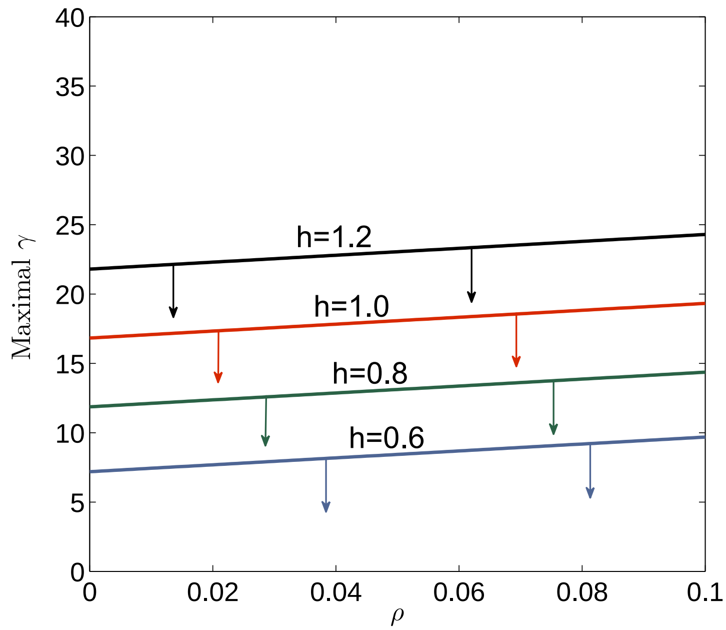


Figure 1.5: Restrictions on γ and ρ implied by good-deal bounds in the power utility case with $c/w = 0.06, r_f = 0.02$.

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.

1.4 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.[13]

A starting point is the close correspondence between expected utility and the

[13] Calculations in the Epstein-Zin case are in Appendix A.2.

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

$$U(\gamma) = \frac{C_0^{1-\gamma}}{1-\gamma} \cdot \left(1 + \frac{W_0}{C_0}\right)\,. \tag{1.44}$$

This correspondence between expected utility and the consumption-wealth ratio, and hence (1.44), 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:

$$U(\gamma) = \frac{C_0^{1-\gamma}}{1-\gamma}\left(1 + \frac{1}{e^{\rho - \mathbf{c}(1-\gamma)} - 1}\right)\,, \qquad \gamma \neq 1\,. \tag{1.45}$$

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.[14]

Equation (1.44) gives expected utility under the status quo; expression (1.45) 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,[15]

[14] 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.

[15] Since the consumption growth process is unchanged, the consumption-wealth ratio remains

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.

1.4.1 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 (1.44) and (1.45), ϕ solves the equation

$$\frac{[(1+\phi)C_0]^{1-\gamma}}{1-\gamma} \cdot \left(1 + \frac{W_0}{C_0}\right) = \frac{C_0^{1-\gamma}}{1-\gamma} \cdot \frac{e^{\rho - \boldsymbol{c}(1)\cdot(1-\gamma)}}{e^{\rho - \boldsymbol{c}(1)\cdot(1-\gamma)} - 1} . \tag{1.46}$$

Simplifying, we have

$$\phi = \left(1 + \frac{W_0}{C_0}\right)^{\frac{1}{\gamma-1}} \left\{1 - e^{-\rho}\left[\mathbb{E}\left(\frac{C_1}{C_0}\right)\right]^{1-\gamma}\right\}^{\frac{1}{\gamma-1}} - 1 . \tag{1.47}$$

What assumptions are required to derive (1.47)? The left-hand side of (1.46) relies on the correspondence between expected utility and the consumption-wealth

constant. The increase in initial consumption therefore corresponds to an increase in initial *wealth* by proportion ϕ.

ratio that was noted at the beginning of section 1.4. 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 (1.47) 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 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 (1.47) 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 1.5. *In the Epstein-Zin case with elasticity of intertemporal substitution ψ, the cost of uncertainty, ϕ, satisfies*

$$\phi = \left(1 + \frac{W_0}{C_0}\right)^{\frac{1}{1/\psi - 1}} \left\{1 - e^{-\rho}\left[\mathbb{E}\left(\frac{C_1}{C_0}\right)\right]^{1-\frac{1}{\psi}}\right\}^{\frac{1}{1/\psi-1}} - 1\,. \tag{1.48}$$

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

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

Proof. See appendix A.2 for the Epstein-Zin calculations. □

Proposition 1.5 shows that if the mean consumption growth rate in levels, consumption-wealth ratio[16] 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).[17] The important feature is that I treat the consumption-wealth 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.

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 (1.47) 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

[16] 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.

[17] 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.

in this chapter. 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 (1.47), 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 1.4: The cost of consumption fluctuations with power utility.

Table 1.4 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 1.1 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 1.4, 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 1.1. 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 1.6. 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.

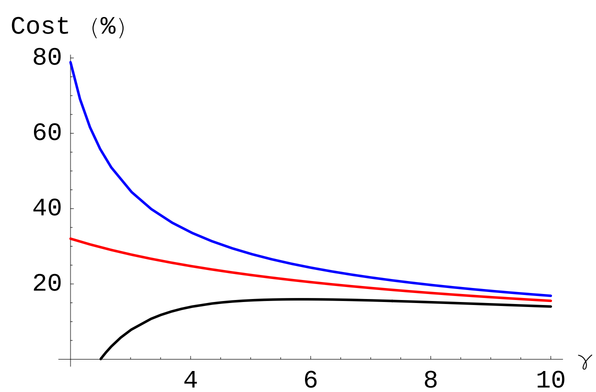


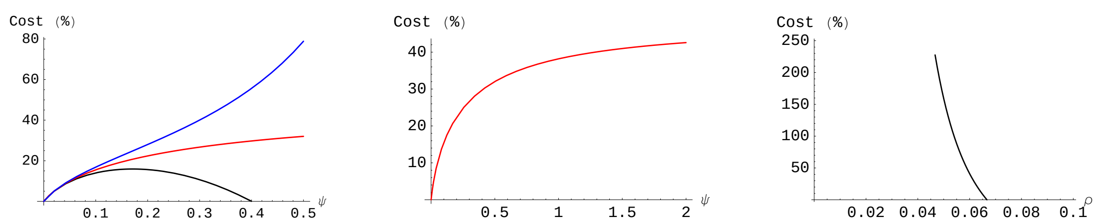
Figure 1.6: The cost of consumption uncertainty plotted against risk aversion, γ, when $\rho = 0.03$ (bottom), $\rho = 0.06$ (middle) and $\rho = 0.09$ (top). The cost of uncertainty ultimately declines as γ increases: for very high values of γ, c/w can only equal 0.06 if there is relatively little risk in consumption growth.

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 1.6, 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.

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.

The cost calculations made in the previous section can be mapped directly into the Epstein-Zin case if $\psi = 0.25$. Figure 1.7a, which is the dual of Figure 1.6 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 $\rho = 0.03$ (bottom), $\rho = 0.06$ (middle) and $\rho = 0.09$ (top).

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

(c) Against ρ, with $\psi = 1.5$.

Figure 1.7: 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 1.7a, 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 1.7b 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 1.7c 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.

1.4.2 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.[18]

Under the new reduced-volatility process, the CGF is

$$\widetilde{\boldsymbol{c}}(\theta) = \boldsymbol{c}(\theta) + \alpha^2\theta/2 - \alpha^2\theta^2/2\,. \tag{1.49}$$

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

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

[18] 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.

Substituting in from (1.49), and replacing $\rho - \boldsymbol{c}(1-\gamma)$ with the observable $c/w = \log(1 + C/W)$, we obtain after some simplification

$$\phi_\alpha = \left\{1 + \frac{W_0}{C_0}\left[1 - e^{-\frac{1}{2}\alpha^2\gamma(\gamma-1)}\right]\right\}^{1/(\gamma-1)} - 1\,. \tag{1.50}$$

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

Proposition 1.6. *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*

$$\phi_\alpha = \left\{1 + \frac{W_0}{C_0}\left[1 - e^{-\frac{1}{2}\alpha^2\gamma(\frac{1}{\psi}-1)}\right]\right\}^{\frac{1}{1/\psi-1}} - 1\,. \tag{1.51}$$

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

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

$$\phi_\alpha \approx \frac{W_0}{C_0}\frac{\gamma\alpha^2}{2}\,. \tag{1.52}$$

Proof. See appendix A.2 for the Epstein-Zin calculations. □

Obstfeld (1994) observes that (1.52) holds in the power utility case with i.i.d. log-normal 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 (1.52) 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.

1.5 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 1.2. *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

$$\boldsymbol{m}_{G_t,H_t}(\boldsymbol{\theta}) = \left(\boldsymbol{m}_{G,H}(\boldsymbol{\theta})\right)^t \tag{1.53}$$

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 A.4, 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

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

as usual.

Proposition 1.7 (Multivariate results in continuous time)**.**

$$D/P = -\boldsymbol{c}_{G,H}(-1,1) \tag{1.54}$$

$$R_f = -\boldsymbol{c}_{G,H}(-1,0) \tag{1.55}$$

$$RP = \boldsymbol{c}_{G,H}(0,1)+\boldsymbol{c}_{G,H}(-1,0)-\boldsymbol{c}_{G,H}(-1,1) \tag{1.56}$$

The discrete-time case is very similar, and Proposition 1.7 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 1.7 can therefore be replaced with R, defined to be the logarithm of the asset's one-period return.

Proposition 1.8 (Multivariate results in discrete time)**.** *Defining R to be the logarithm of the asset in question's one period gross return, we have*

$$d/p = -\boldsymbol{c}_{G,R}(-1,1) \tag{1.57}$$

$$r_f = -\boldsymbol{c}_{G,R}(-1,0) \tag{1.58}$$

$$rp = \boldsymbol{c}_{G,R}(0,1) + \boldsymbol{c}_{G,R}(-1,0) - \boldsymbol{c}_{G,R}(-1,1) \tag{1.59}$$

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.

1.5.1 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 1.2.2 above. All agents have power utility with relative risk aversion γ.

Aggregate consumption, written C_t, is as in (1.29), so

$$\log \frac{C_t}{C_0} = \mu t + \sigma B_t + \sum_{j=1}^{N_t} Y_j \tag{1.60}$$

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 (1.60). 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

$$\log \frac{C_{i,t}}{C_{i,0}} = \log \frac{C_t}{C_0} + \underbrace{\sigma_1 B_{it} - \frac{1}{2}\sigma_1^2 t}_{\text{type (i)}} + \underbrace{\sum_{j=1}^{N_{i,t}} X_{i,j}}_{\text{type (ii)}} + \underbrace{\sum_{k=1}^{N_t} Y_{i,k}}_{\text{type (iii)}} \tag{1.61}$$

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, (1.61) is consistent with the evolution of aggregate consumption in (1.60) under the maintained assumption that for all i and k,

$$\mathbb{E}e^{X_{i,k}} = \mathbb{E}e^{Y_{i,k}} = 1\,. \tag{1.62}$$

(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 $\boldsymbol{m}(\theta) \equiv \mathbb{E}e^{\theta Y_j}$.[19] 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 $\boldsymbol{m}_3(\theta) \equiv \mathbb{E}e^{\theta Y_{i,j}}$.

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

$$P = \mathbb{E}\int_0^\infty e^{-\rho t}\left(\frac{C_{i,t}}{C_{i,0}}\right)^{-\gamma} \cdot C_t\, dt \tag{1.63}$$

as usual. Heterogeneity matters: dropping the is in (1.63) 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

$$G_{i,t} \equiv -\log\frac{M_{i,t}}{M_{i,0}} = \rho + \gamma\cdot\log\frac{C_{i,t}}{C_{i,0}} \tag{1.64}$$

$$H_t \equiv \log\frac{C_t}{C_0}\,. \tag{1.65}$$

We can apply the results of Proposition 1.7 directly; I retain the i subscript in $G_{i,t}$ as a reminder that individuals, not aggregates, price assets.

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

By Proposition 1.7, we have

$$D/P = -\boldsymbol{c}_{G_i,H}(-1,1) \tag{1.66}$$

$$R_f = -\boldsymbol{c}_{G_i,H}(-1,0) \tag{1.67}$$

$$RP = \boldsymbol{c}_{G_i,H}(0,1) + \boldsymbol{c}_{G_i,H}(-1,0) - \boldsymbol{c}_{G_i,H}(-1,1) \tag{1.68}$$

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

$$\begin{aligned}\boldsymbol{c}_{G_i,H}(\theta_1,\theta_2) &= \rho\theta_1 + \mu\left(\gamma\theta_1+\theta_2\right) + \frac{1}{2}\sigma^2\left(\gamma\theta_1+\theta_2\right)^2 + \frac{1}{2}\sigma_1^2\gamma\theta_1(\gamma\theta_1-1) + \\ &\quad + \omega\left[\boldsymbol{m}\left(\gamma\theta_1+\theta_2\right)\boldsymbol{m}_3\left(\gamma\theta_1\right)-1\right] + \omega_2\left[\boldsymbol{m}_2(\gamma\theta_1)-1\right]\end{aligned} \tag{1.69}$$

The correct consumption-wealth ratio, riskless rate and risk premium on the consumption claim can be obtained from (1.66)–(1.69). 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 (1.69). The discrepancies between true fundamentals and incorrect predictions based on aggregate quantities (denoted by bars) are given by

$$C/W - \overline{C/W} = -\sigma_1^2\gamma(\gamma+1)/2 - \omega_2\left[\boldsymbol{m}_2(-\gamma)-1\right] - \omega\,\boldsymbol{m}(1-\gamma)\left[\boldsymbol{m}_3(-\gamma)-1\right] \tag{1.70}$$

$$R_f - \overline{R_f} = -\sigma_1^2\gamma(\gamma+1)/2 - \omega_2\left[\boldsymbol{m}_2(-\gamma)-1\right] - \omega\,\boldsymbol{m}(-\gamma)\left[\boldsymbol{m}_3(-\gamma)-1\right] \tag{1.71}$$

$$RP - \overline{RP} = \omega\left[\boldsymbol{m}(-\gamma)-\boldsymbol{m}(1-\gamma)\right]\left[\boldsymbol{m}_3(-\gamma)-1\right] \tag{1.72}$$

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 (1.70) and (1.71) are negative and that (1.72) is positive.

Proposition 1.9 (Asset pricing implications of heterogeneity)**.** *Heterogeneity drives down the consumption-wealth ratio and riskless rate and increases the risk premium:*

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

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 (1.70)–(1.72).

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

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

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 $\boldsymbol{\psi}(1-\gamma) \le 0$; then we're done, since $\boldsymbol{\psi}(-\gamma) > 0 \ge \boldsymbol{\psi}(1-\gamma)$. 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 (1.72), this increases the equity premium by

$$\omega\left(e^{b\gamma+s^2\gamma^2/2} - e^{b(\gamma-1)+s^2(\gamma-1)^2/2}\right)\left(e^{s_i^2\gamma(\gamma+1)/2} - 1\right) = \Delta RP(\gamma, s_i)\,. \qquad (1.73)$$

Using the now familiar parameter values $\omega = 0.017, b = 0.39, s = 0.25$, Figure 1.8 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.

1.6 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 the heterogeneous agent model of section 1.5.1. In more complicated settings—as in Chapter 2—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

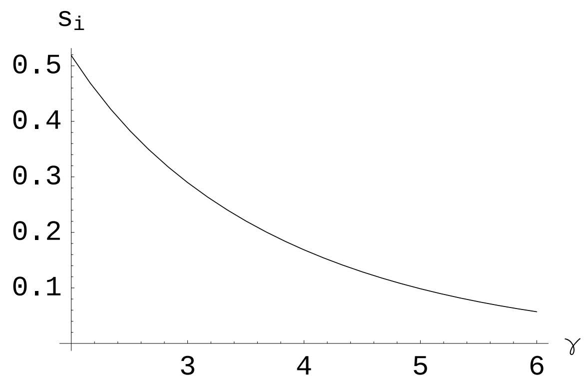


Figure 1.8: The value of s_i needed to give $RP - \overline{RP} = 2\%$, against γ.

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 1.9 in section 1.5.1 would have been considerably more complicated. Convexity arguments were also employed in section 1.3, 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 chapter, 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 1.4 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.

2. THE LUCAS ORCHARD

This chapter investigates the properties of asset prices, risk premia, and the term structure of interest rates in a continuous-time economy in which a representative agent with power utility consumes the sum of the dividends of N assets.[1] The assets can be thought of as Lucas trees, so I call the collection of assets a Lucas orchard. In applications, the assets may represent claims to the dividends of industries, asset classes, or countries.

Each of the assets is assumed to have i.i.d. dividend growth over time, though there may be correlation between the dividend growth rates of different assets. Formally, the vector of log dividends follows a Lévy process. This framework allows for the case in which dividends follow geometric Brownian motions, but also allows for a rich structure of jumps in dividends. Standard lognormal models make poor predictions for key asset-pricing quantities such as the equity premium and riskless rate (Mehra and Prescott (1985)), and recently there has been increased interest in models which allow for the possibility of disasters (Rietz (1988), Barro (2006a), Gabaix (2008)). By allowing for jumps in dividends, I avoid these puzzles without relying on implausible levels of risk aversion or dividend growth volatility.

Despite its simple structure, the model exhibits surprisingly rich asset price be-

[1] I am grateful to Tobias Adrian, Malcolm Baker, Thomas Baranga, Robert Barro, John Cochrane, George Constantinides, Josh Coval, John Cox, Emmanuel Farhi, Xavier Freixas, , Xavier Gabaix, Lars Hansen, Jakub Jurek, David Laibson, Robert Lucas, Greg Mankiw, Emi Nakamura, Martin Oehmke, Lubos Pastor, Roberto Rigobon, David Skeie, Jeremy Stein, Jon Steinsson, Aleh Tsyvinski, Pietro Veronesi, Luis Viceira, James Vickery, Jiang Wang and, in particular, John Campbell and Chris Rogers for their comments on this chapter.

havior, including several phenomena that have been documented in the empirical literature; it illustrates "the importance of explicit recognition of the essential interdependences of markets in theoretical and empirical specifications of financial models" (Brainard and Tobin (1968)).

Similarly, although the assumptions underlying the model are simple and natural, the interaction between multiplicative structure (induced by i.i.d. growth in log dividends) and additive structure (consumption is the sum of dividends) makes the model hard to solve. I use techniques from complex analysis to solve for prices, returns, and interest rates in terms of integral formulas that can be evaluated numerically. These integral formulas are valid for arbitrary i.i.d. dividend growth processes, subject to conditions that ensure finiteness of the representative agent's expected utility (and hence of asset prices). When there are two assets whose dividends follow geometric Brownian motions, the integrals can be solved in closed form. It is hoped that these techniques may find application in other areas of economics.

In the general case considered here, dividends—and hence prices, expected returns, and interest rates—can *jump*, and neither the conditional consumption-CAPM (Breeden (1979)) nor the ICAPM (Merton (1973)) hold. In the special case in which dividends follow geometric Brownian motions, asset prices follow diffusions, so the ICAPM and conditional consumption-CAPM do hold.[2] Here, though, price processes are not taken as given but are determined endogenously based on exogenous fundamentals, in the spirit of Cox, Ingersoll and Ross (1985).

The tractability of the model in the general i.i.d. case is due in part to the use of cumulant-generating functions (CGFs). Chapter 1 expressed the riskless rate, risk premium, and consumption-wealth ratio in terms of the CGF in the case $N = 1$, and the expressions found there are echoed in the more complicated scenario considered

[2] The conditional CAPM itself holds only if dividends follow geometric Brownian motions and the representative agent has log utility, as in Cochrane, Longstaff and Santa-Clara (2008).

here. In effect, working with CGFs makes the mathematics no harder than when working with lognormal models; the advantage of doing so is that one then "gets jumps for free". In fact, the use of CGFs may even make things *simpler* because one can follow the CGF's progress through the algebra: the mathematical equivalent of a barium meal! Furthermore, CGFs have useful properties that I use in various proofs.

For simplicity, I introduce the model in the case $N = 2$. I present two calibrations, each intended to highlight different features of the model. In the first, dividends follow geometric Brownian motions. In the second, I use a calibration based on Barro (2006a) to highlight the impact of rare disasters in a multi-asset framework.

A central feature of both calibrations is that assets whose dividends make up a large proportion of consumption are riskier, all else equal, than assets that make up a small proportion of consumption. Large assets have low price-dividend ratios; small assets have high price-dividend ratios. (For simplicity, this discussion assumes that assets are independent and have fundamentally the same prospects—the same mean dividend growth rate, dividend volatility, susceptibility to disasters, and so on.)

Various properties of the model spring from this fact. Since dividend growth is i.i.d., it is not forecastable. High price-dividend ratios therefore *cannot* forecast high dividend growth, and are instead associated with low expected returns. In calibrations, I also show that assets with high price-dividend ratios also have low expected *excess* returns: the value-growth effect of Fama and French (1993). Moreover, the expected excess return on a value-minus-growth strategy is time-varying and moves with the value spread (the difference in dividend yields between value and growth assets), as has been found in the data by Cohen, Polk and Vuolteenaho (2003).

The model generates price comovement even between assets whose dividends are independent. To see why this happens, suppose that one asset's price increases as a result of a positive shock to dividends. The other asset now contributes a smaller proportion of overall consumption, and therefore typically has a lower required return and hence a higher price.[3] Such comovement is a feature of the data. Shiller (1989) demonstrates that stock prices in the US and UK move together more closely than do fundamentals; Forbes and Rigobon (2002) allow for heteroskedasticity in returns and find consistently high levels of interdependence between markets.

In the model, high market price-dividend ratios forecast low market expected returns. It follows that the market displays "excess" volatility, as documented by Shiller (1981) and many others, in the sense that its returns are more volatile than its dividends. It should be acknowledged, however, that in the calibrations presented here the model does not generate as much excess volatility as is observed in the data.

The riskless rate varies over time, so the term structure of interest rates is not flat. The term structure can be upward-sloping, downward-sloping or hump-shaped (with medium-term bonds earning higher yields than short- and long-term bonds). When the term structure slopes up—the more usual case in the scenarios I consider—long-term bonds earn positive risk premia. High yield spreads forecast high excess returns on the market and high excess returns on long-term bonds, replicating a finding of the empirical literature (for example, Fama and French (1989)).

I decompose realized returns into dividend-driven returns and valuation-driven returns. The latter are returns due to changes in price-dividend ratios—for example, when one asset comoves with another that has received good news, it earns a positive valuation-driven return. Most of the variance in asset returns, particularly for large assets, is dividend-driven. For small assets, however, valuation-driven returns are more important. Small assets also exhibit momentum, in the sense that their

[3] In some circumstances, discussed further below, movements in the riskless rate may partially offset or reverse this effect.

dividend-driven returns and valuation-driven returns are negatively correlated.

In the second calibration, occasional disasters afflict the two assets. The phenomena described above are present, and there are now some new features. First, the introduction of disasters enables the calibration, like that of Barro (2006a), to avoid the equity premium and riskless rate puzzles. Second, jumps are transmitted across assets. When a large asset experiences a disaster, the price of the other (small) asset also jumps downwards. This corresponds to the "typical" case of comovement described above. When, on the other hand, a very small asset suffers a disaster, interest rates drop and the other (large) asset's price jumps *up*. I label these phenomena "contagion" and "flight-to-quality".

Contagion effects provide a new channel through which disasters can contribute to high risk premia. For example, suppose that asset 1 has perfectly stable dividends, but that asset 2 is subject to occasional disastrous declines in dividends. Contagion leads to declines in the price of asset 1 at times when asset 2 experiences a disaster. These occasional price drops may induce a substantial risk premium in asset 1, an ostensibly perfectly safe asset.

I next consider the limit in which one of the two assets is negligibly small by comparison with the other. This case is of special interest because it represents the most extreme departure from simple models in which price-dividend ratios are constant. Closed-form solutions are available without any restrictions on the dividend growth process, and an unexpected phenomenon emerges.

To illustrate this, suppose that the two assets have independent dividend streams. Intuition suggests that a small idiosyncratic asset earns no risk premium, that its expected return is therefore equal to the riskless rate and that it can be valued using a Gordon growth formula; in other words, its dividend yield should equal the riskless rate minus expected dividend growth. I show that this intuition is correct whenever the result of the calculation is meaningful, which is to say positive. What happens if the riskless rate (determined by the characteristics of the large asset) is less than the mean dividend growth of the small asset? I show that the negligibly small asset then has a well-defined price-consumption ratio that, as one would expect, tends to zero in the limit. It has, however, an extremely high valuation in the sense that its price-dividend ratio is infinite in the limit. This valuation effect is reminiscent of, and complementary to, that present in the papers of Pástor and Veronesi (2003, 2006). Despite its independent fundamentals and negligible size, such an asset co-moves endogenously, and hence earns a positive risk premium. In the general case, I provide a precise characterization of when the Gordon growth model does and does not work, and solve for limiting expected returns and price-dividend ratios in closed form.

Various authors have investigated related models. Cole and Obstfeld (1991) consider a similar framework, but focus on the welfare gains from international risk sharing rather than the implications for asset prices, and they do not present any analytical results in the case considered here, in which the the dividends of the two assets are perfect substitutes. Brainard and Tobin (1992, section 8) investigate a framework that is almost identical to the one presented here, differing only in that the dividends of the two assets are very good, rather than perfect, substitutes, and in that per-period endowments are specified by a Markov chain with a small number of states. They present limited numerical results, and—after noting that their "model is simple and abstract; nevertheless it is not easy to analyze"—no analytical results. Menzly, Santos and Veronesi (2004) and Santos and Veronesi (2006) present models in which the dividend shares of assets are assumed to follow mean-reverting processes. By picking convenient functional forms for these processes, closed-form pricing formulas are available. Pavlova and Rigobon (2007) investigate the consequences of demand shocks in an international asset pricing model, but impose log-linear preferences, so price-dividend ratios are constant.

A closely related paper is that of Cochrane, Longstaff and Santa-Clara (2008), who solve a model in which a representative investor with log utility consumes the dividends of two assets whose dividend processes follow geometric Brownian motions. My solution technique is entirely different, and permits me to allow for power utility, for jumps in dividends, and for $N \geq 2$ assets. I also solve for bond yields, and hence expand the set of predictions made by the model.

Section 2.1 sets up the model in the two-asset case. Section 2.2 explains why it is hard to solve and introduces a suggestive special case that is easily solved. Section 2.3 presents integral formulas for prices, expected returns, and real interest rates. Section 2.4 provides closed-form solutions in the Brownian motion case. Section 2.5 explores two calibrations. Section 2.6 investigates the small asset case. Section 2.7 provides integral formulas in the N-asset case. Section 2.8 concludes. Proofs are collected in the appendices.

2.1 Setup

For the time being, I restrict to the two-asset case for clarity. General results in the N-asset case are presented in Section 2.7.

Setting the model up amounts to making *technological* assumptions about dividend processes; making assumptions about the *preferences* of the representative investor that, together with consumption, pin down the stochastic discount factor; and closing the model by specifying that the representative investor's consumption is equal to the sum of the two assets' dividends.

2.1.1 The stochastic discount factor

Time is continuous, and runs from 0 (the present) to infinity. I assume that there is a representative agent with power utility over consumption C_t, with coefficient of relative risk aversion γ and time preference rate ρ. The Euler equation, derived by Lucas (1978) and applied in the two-country context by Lucas (1982), states that the price of an asset with dividend stream $\{X_t\}$ is

$$P_X = \mathbb{E} \int_0^\infty e^{-\rho t} \left(\frac{C_t}{C_0} \right)^{-\gamma} \cdot X_t \, dt \,. \tag{2.1}$$

2.1.2 Dividend processes

The two assets, indexed $i = 1, 2$, throw off random dividend streams D_{it}. Dividends are positive, which makes it natural to work with log dividends, $y_{it} \equiv \log D_{it}$. At time 0, the dividends (y_{10}, y_{20}) of the two assets are arbitrary. The vector $\widetilde{\boldsymbol{y}}_t \equiv \boldsymbol{y}_t - \boldsymbol{y}_0 \equiv (y_{1t} - y_{10}, y_{2t} - y_{20})$ is assumed to follow a Lévy process.

Definition 2.1. *A stochastic process $(L_t)_{t \geq 0}$ taking values in $\mathbb{R}^d$ is a* Lévy process *if*

(i) $L_0 = 0$

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

(iii) For any $n \in \mathbb{N}$ and $0 \leq t_0 < t_1 < \ldots < t_n$, the random variables $L_{t_0}, L_{t_1} - L_{t_0}, L_{t_2} - L_{t_1}, \ldots, L_{t_n} - L_{t_{n-1}}$ are independent.

(iv) The probability distribution of $L_{t+h} - L_t$ does not depend on t.

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

This is the continuous-time analogue of the familiar discrete-time assumption that dividend growth is i.i.d. It is helpful to keep in mind the special case in which $\widetilde{\boldsymbol{y}}$ is a jump-diffusion, in which case we can write

$$\boldsymbol{y}_t = \boldsymbol{y}_0 + \boldsymbol{\mu} t + \boldsymbol{A} \boldsymbol{Z}_t + \sum_{k=1}^{N(t)} \boldsymbol{J}^k \,. \tag{2.2}$$

Here $\boldsymbol{\mu}$ is a two-dimensional vector of "drifts", $\boldsymbol{A}$ a 2×2 matrix of factor loadings, $\boldsymbol{Z}_t$ a 2-dimensional Brownian motion, $N(t)$ a Poisson process with arrival rate ω that represents the number of jumps that have taken place by time t, and $\boldsymbol{J}^k$ are two-dimensional random variables which are distributed like the random variable $\boldsymbol{J}$, and which are assumed to be i.i.d. across time. The covariance matrix of the diffusion components of the two dividend processes is $\boldsymbol{\Sigma} \equiv \boldsymbol{A}\boldsymbol{A}'$, whose elements I write as σ_{ij}.

As in Chapter 1, it is convenient to define the cumulant-generating function, an object that turns out to capture all relevant information about the stochastic processes driving dividend growth.

Definition 2.2. *The* cumulant-generating function $\boldsymbol{c}(\boldsymbol{\theta})$ *of the Lévy process* $\widetilde{\boldsymbol{y}}_t$ *is defined by*

$$\boldsymbol{c}(\boldsymbol{\theta}) \equiv \log \mathbb{E} \exp \boldsymbol{\theta}' (\widetilde{\boldsymbol{y}}_{t+1} - \widetilde{\boldsymbol{y}}_t) \,. \tag{2.3}$$

By properties (i) and (iv) of the definition of a Lévy process, I could equivalently have defined $\boldsymbol{c}(\boldsymbol{\theta}) = \log \mathbb{E} \exp \boldsymbol{\theta}' \widetilde{\boldsymbol{y}}_1$, but the expression (2.3) emphasizes the fact that the cumulant-generating function (CGF) summarizes information about dividend *growth*. Specifically, the CGF summarizes information about the higher moments of $\widetilde{\boldsymbol{y}}$; see also Chapter 1 for more discussion of the role of CGFs in the standard consumption-based framework with one asset.

Some conditions on the Lévy process $\widetilde{\boldsymbol{y}}$ are required to ensure that asset prices are finite; these are discussed further below. In particular, they will ensure that the CGF exists in an appropriate open set containing the origin.

If log dividends follow Brownian motions, the CGF takes the simple form

$$\boldsymbol{c}(\boldsymbol{\theta}) = \boldsymbol{\theta}' \boldsymbol{\mu} + \frac{1}{2} \boldsymbol{\theta}' \boldsymbol{\Sigma} \boldsymbol{\theta} \,.$$

If log dividends follow a jump-diffusion as in (2.2), then

$$\boldsymbol{c}(\boldsymbol{\theta}) = \boldsymbol{\theta}'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\theta}'\boldsymbol{\Sigma}\boldsymbol{\theta} + \omega\left(\mathbb{E}e^{\boldsymbol{\theta}'\boldsymbol{J}} - 1\right) .$$

If the jump sizes are Normally distributed, $\boldsymbol{J} \sim N(\boldsymbol{\mu}_J, \boldsymbol{\Sigma}_J)$, then the CGF becomes

$$\boldsymbol{c}(\boldsymbol{\theta}) = \boldsymbol{\theta}'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\theta}'\boldsymbol{\Sigma}\boldsymbol{\theta} + \omega\left(\exp\left\{\boldsymbol{\theta}'\boldsymbol{\mu}_J + \frac{1}{2}\boldsymbol{\theta}'\boldsymbol{\Sigma}_J\boldsymbol{\theta}\right\} - 1\right) .$$

2.1.3 Closing the model

Dividends are not storable, and the representative investor must hold the market, so the model is closed by stipulating that the representative agent's consumption equals the sum of the two dividends: $C_t = D_{1t} + D_{2t}$.

2.2 A simple example

Consider the problem of pricing the claim to asset 1's output in the simplest case $\gamma = 1$: log utility. We have

$$\begin{aligned} P_1 &= \mathbb{E}\int_0^\infty e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-1} \cdot D_{1t}\,dt \\ &= \mathbb{E}\int_0^\infty e^{-\rho t}\frac{D_{10} + D_{20}}{D_{1t} + D_{2t}} \cdot D_{1t}\,dt \\ &= (D_{10} + D_{20})\int_0^\infty e^{-\rho t}\mathbb{E}\left(\frac{1}{1 + D_{2t}/D_{1t}}\right)dt\,; \end{aligned}$$

and, unfortunately, the expectation is not easily calculated. If, say, the D_{it} are geometric Brownian motions, then we have to compute the expected value of the reciprocal of one plus a lognormal random variable. This, essentially, is the major analytical challenge confronted by Cochrane, Longstaff and Santa-Clara (2008).

Here, though, is an instructive case in which the expectation simplifies considerably. Suppose that $D_{2t} < D_{1t}$ at all times t. Perhaps, for example, D_{1t} is constant and initially larger than D_{2t}, which is subject to downward jumps at random times.[4] (The jumps may be random in size, but they must always be downwards.) Then $D_{2t}/D_{1t} < 1$ and so we can expand the expectation as a geometric sum. To make things simple, set $D_{1t} \equiv 1$: then,

$$\begin{aligned} \mathbb{E}\left(\frac{1}{1+D_{2t}}\right) &= \mathbb{E}\left[1 - D_{2t} + D_{2t}^2 - \ldots\right] \\ &= \sum_{n=0}^{\infty}(-1)^n D_{20}^n \mathbb{E}\left[(D_{2t}/D_{20})^n\right] \\ &= \sum_{n=0}^{\infty}(-1)^n D_{20}^n e^{\boldsymbol{c}(0,n)t} \,. \end{aligned} \tag{2.4}$$

Substituting back, we find that

$$\begin{aligned} P_1 &= (1+D_{20}) \int_{t=0}^{\infty} e^{-\rho t} \sum_{n=0}^{\infty}(-1)^n D_{20}^n e^{\boldsymbol{c}(0,n)t}\, dt \\ &= (1+D_{20}) \sum_{n=0}^{\infty}(-1)^n D_{20}^n \int_{t=0}^{\infty} e^{-[\rho-\boldsymbol{c}(0,n)]t}\, dt \\ &= (1+D_{20}) \sum_{n=0}^{\infty} \frac{(-1)^n D_{20}^n}{\rho - \boldsymbol{c}(0,n)} \end{aligned}$$

If we define $s \equiv D_{10}/(D_{10}+D_{20})$ to be the share of asset 1 in global output—a definition which is maintained throughout—we can rewrite this in a form that is more directly comparable with subsequent results:

$$P/D_1 = \frac{1}{\sqrt{s(1-s)}} \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{1-s}{s}\right)^{n+1/2}}{\rho - \boldsymbol{c}(0,n)} \tag{2.5}$$

P/D_1 is the price-dividend ratio of asset 1 at time 0. When time subscripts are dropped, here and elsewhere, the relevant time is time 0.

[4] This approach fails in the Brownian motion case, since if either D_{1t} or D_{2t} has a Brownian component we cannot say that $D_{2t} < D_{1t}$ with probability one.

The expression (2.5) is not in closed form, but it is easy to evaluate numerically, once the process driving the dividends of asset 2—and hence $\boldsymbol{c}(0,n)$—is specified. For example, if asset 2's log dividend is subject to downward jumps of constant size $-b$ which occur at intervals dictated by a Poisson process with arrival rate ω, then $\boldsymbol{c}(0,n) = \omega(e^{-bn} - 1)$, so $\rho - \boldsymbol{c}(0,n) \to \rho + \omega$ as $n \to \infty$. Meanwhile, $(1-s)/s < 1$ so the terms in the numerator of the summand decline at geometric rate. A numerical summation will therefore converge fast.

The extremely special structure of this example made it legitimate to write $1/(1 + D_{2t})$ as a geometric sum $1 - D_{2t} + D_{2t}^2 - \ldots$ in (2.4). In the general case, it will turn out to be possible to make an analogous move, writing the equivalent of $1/(1 + D_{2t})$ as a Fourier integral before computing the expectation.

2.3 General solution in the two-asset case

It is convenient to work with a generic asset with dividend stream $D_{\boldsymbol{\alpha},t} \equiv D_{1t}^{\alpha_1} D_{2t}^{\alpha_2}$, where $\boldsymbol{\alpha} \equiv (\alpha_1, \alpha_2) \in \{(1,0), (0,1), (0,0)\}$. The three alternatives represent asset 1, asset 2, and a riskless perpetuity respectively.

2.3.1 Prices

Asset prices turn out to depend on the value of a single state variable $s \in [0,1]$, the share of aggregate consumption contributed by the dividend of asset 1:

$$s = \frac{D_{10}}{D_{10} + D_{20}} .$$

The following Proposition supplies an integral formula for the price-dividend ratio on the $\boldsymbol{\alpha}$-asset. The formula is perfectly suited for numerical implementation but also permits further analytical results to be derived.

Proposition 2.1 (The general pricing formula)**.** *The price-dividend ratio on a* *generic asset which pays dividend stream $D_{\boldsymbol{\alpha},t} \equiv D_{1t}^{\alpha_1} D_{2t}^{\alpha_2}$ is given by the expression*[5]

$$\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}}(s) = \frac{1}{\sqrt{s^{\gamma}(1-s)^{\gamma}}} \int_{-\infty}^{\infty} \frac{\mathscr{F}_{\gamma}(v)\left(\frac{1-s}{s}\right)^{iv}}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)}\, dv\,, \tag{2.6}$$

where $\mathscr{F}_{\gamma}(v)$ is defined by

$$\mathscr{F}_{\gamma}(v) \equiv \frac{1}{2\pi} \cdot \frac{\Gamma(\gamma/2 + iv)\Gamma(\gamma/2 - iv)}{\Gamma(\gamma)}\,. \tag{2.7}$$

Proof. See Appendix B.1. □

The gamma function $\Gamma(z)$ that appears in (2.7) is defined for complex numbers z with positive real part by

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t}\, dt\,.$$

For real v and integer $\gamma > 0$, $\mathscr{F}_{\gamma}(v)$ is a strictly positive function which is symmetric about $v = 0$, where it attains its maximum, and decays exponentially fast towards zero as v tends to plus or minus infinity.

In its present form, the pricing formula (2.6) appears rather complicated, but it is worth emphasizing that it allows for different assets ($\boldsymbol{\alpha}$) and for the stochastic process governing log outputs to be *any* Lévy process that leads to finite asset prices—a class which includes, for example, constant deterministic growth, drifting Brownian motion, compound Poisson processes, variance gamma processes, Normal inverse Gaussian processes, and a host of others, including linear combinations of the processes mentioned.

[5] Wherever it appears, i is the complex number $\sqrt{-1}$.

We do, however, require that expected utility and asset prices are finite. I show in Appendix B.1.2 that finiteness of the prices of the two assets—which implies that expected utility is finite—is assured by the *finiteness condition* that

$$\rho - \boldsymbol{c}(1-\gamma/2, -\gamma/2) > 0 \qquad \text{and} \qquad \rho - \boldsymbol{c}(-\gamma/2, 1-\gamma/2) > 0\,. \tag{2.8}$$

Given that perpetuities in zero net supply plausibly also have finite prices, we may also want to impose a requirement that ensures that this is the case,

$$\rho - \boldsymbol{c}(-\gamma/2, -\gamma/2) > 0\,.$$

This restriction is not necessary from a mathematical point of view; I impose it because it seems empirically plausible that real perpetuities in zero net supply have finite prices. (If either of the assets in positive net supply *is* a perpetuity, then this restriction is implied by (2.8).)

These assumptions ensure that aggregate wealth is finite for all $s \in (0,1)$. I impose one final restriction, that aggregate wealth is finite at the one-tree limit points, $s = 0$ and $s = 1$. Asset 1's price-dividend ratio is finite as $s \to 1$ if and only if $\rho - \boldsymbol{c}(1-\gamma, 0) > 0$; asset 2's price-dividend ratio is finite as $s \to 0$ if and only if $\rho - \boldsymbol{c}(0, 1-\gamma) > 0$. These assumptions are summarized in Table 2.1.

Restriction	Reason
$\rho - \boldsymbol{c}(1-\gamma/2, -\gamma/2) > 0$	finite price of asset 1
$\rho - \boldsymbol{c}(-\gamma/2, 1-\gamma/2) > 0$	finite price of asset 2
$\rho - \boldsymbol{c}(-\gamma/2, -\gamma/2) > 0$	finite perpetuity price
$\rho - \boldsymbol{c}(1-\gamma, 0) > 0$	finite aggregate wealth in limit $s \to 1$
$\rho - \boldsymbol{c}(0, 1-\gamma) > 0$	finite aggregate wealth in limit $s \to 0$

Table 2.1: The restrictions imposed on the model.

For many practical purposes this is, in a sense, the end of the story, since the integral formula is very well behaved and can be calculated effectively instantly in *Mathematica* or *Maple*. After providing similar integral formulas for expected returns, the riskless rate, and bond yields, I take this simple and direct route in section 2.5. Nonetheless, it is possible to push the pen-and-paper approach further in the case in which log dividends follow drifting Brownian motions: the integral (2.6) is then soluble in closed form. See section 2.4.

It is sometimes more convenient to work with the state variable u, a monotonic transformation of s which is defined by

$$u = \log\left(\frac{1-s}{s}\right) = y_{20} - y_{10}$$

While s ranges between 0 and 1, u takes values between $-\infty$ and $+\infty$. As asset 1 becomes small, u tends to infinity; as asset 1 becomes large, u tends to minus infinity.

Proposition 2.2 (The general pricing formula, alternative version)**.** *In terms of the state variable u, the price-dividend ratio on a generic asset which pays dividend stream $D_{\boldsymbol{\alpha},t} \equiv D_{1t}^{\alpha_1} D_{2t}^{\alpha_2}$ is given by the expression*

$$\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}}(u) = [2\cosh(u/2)]^{\gamma} \cdot \int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_{\gamma}(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)}\, dv \tag{2.9}$$

or equivalently by

$$\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}}(u) = \frac{\displaystyle\int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_{\gamma}(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)}\, dv}{\displaystyle\int_{-\infty}^{\infty} e^{iuv}\mathscr{F}_{\gamma}(v)\, dv} \tag{2.10}$$

Proof. See Appendix B.1. □

2.3.2 Returns

An expression for the expected return on a general asset paying dividend stream $D_{\boldsymbol{\alpha},t}$ can be found in terms of integrals very similar to those that appear in the general price-dividend formula. The instantaneous expected return is defined by

$$R_{\boldsymbol{\alpha}}dt \equiv \underbrace{\frac{\mathbb{E}dP_{\boldsymbol{\alpha}}}{P_{\boldsymbol{\alpha}}}}_{\text{capital gains}} + \underbrace{\frac{D_{\boldsymbol{\alpha}}}{P_{\boldsymbol{\alpha}}}dt}_{\text{dividend yield}} \tag{2.11}$$

Proposition 2.3 (Expected returns)**.** *$R_{\boldsymbol{\alpha}}$, the instantaneous expected return on an asset which pays dividend stream $D_{1t}^{\alpha_1}D_{2t}^{\alpha_2}$, is given by*

$$R_{\boldsymbol{\alpha}}(u) = \frac{\sum_{m=0}^{\gamma}\binom{\gamma}{m}e^{-mu}\int_{-\infty}^{\infty}h(v)e^{iuv}\cdot\boldsymbol{c}(\boldsymbol{w}_m(v))\,dv}{\sum_{m=0}^{\gamma}\binom{\gamma}{m}e^{-mu}\int_{-\infty}^{\infty}h(v)e^{iuv}\,dv} + \frac{D_{\boldsymbol{\alpha}}}{P_{\boldsymbol{\alpha}}}(u)\,. \tag{2.12}$$

where

$$h(v) \equiv \frac{\mathscr{F}_{\gamma}(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)}\,, \tag{2.13}$$

and

$$\boldsymbol{w}_m(v) \equiv (\alpha_1 - \gamma/2 + m - iv, \alpha_2 + \gamma/2 - m + iv)\ .$$

An analogous formula written in terms of the state variable s can be obtained by setting $u = \log[(1-s)/s]$ throughout (2.12)*.*

Proof. Appendix B.1 contains the details of the capital gains calculation. The dividend yield component is given by the reciprocal of (2.9). □

2.3.3 Interest rates

The calculations of sections 2.3.1 and 2.3.2 deal with assets which pay a constant stream of dividends. This section calculates zero coupon bond prices and yields.

First, some notation. I write B_T for the time-0 price of a zero-coupon bond which pays one unit of the consumption good at time T. The (zero-coupon) yield to time $T > 0$, $\mathscr{Y}(T)$, is defined by

$$B_T = e^{-\mathscr{Y}(T)\cdot T}\,.$$

Interest rates are not constant in this economy unless the two assets have identical, perfectly correlated, output processes. For example, the prices of perpetuities and zero coupon bonds fluctuate over time. Define, therefore, the instantaneous riskless rate, r, by

$$r \equiv \lim_{T\downarrow 0} \mathscr{Y}(T).$$

The following Proposition summarizes the behavior of real interest rates, in terms of the state variable u. Depending on the particular stochastic process driving dividends, the model can generate upward- or downward-sloping curves and humped curves with a local maximum.

Proposition 2.4 (Real interest rates)**.** *The yield to time T is*

$$\mathscr{Y}(T) = \rho - \frac{1}{T}\log\left\{[2\cosh(u/2)]^{\gamma}\int_{-\infty}^{\infty}\mathscr{F}_{\gamma}(v)e^{iuv}\cdot e^{\boldsymbol{c}(-\gamma/2-iv,-\gamma/2+iv)T}\,dv\right\}\,. \quad (2.14)$$

The instantaneous riskless rate is

$$r = [2\cosh(u/2)]^{\gamma}\int_{-\infty}^{\infty}\mathscr{F}_{\gamma}(v)e^{iuv}\cdot[\rho - \boldsymbol{c}(-\gamma/2-iv,-\gamma/2+iv)]\,dv\,. \qquad (2.15)$$

As before, we can set $u = \log[(1-s)/s]$ in (2.14) *and* (2.15) *to express yields and the riskless rate in terms of the output share s.*

Proof. See Appendix B.1. □

2.4 The Brownian motion case

When dividends follow geometric Brownian motions,[6] closed-form solutions can be obtained for asset prices. Suppose, then, that log dividend processes are driven by a pair of Brownian motions,

$$dy_i = \mu_i \, dt + \sqrt{\sigma_{ii}} \, dz_i \,, \tag{2.16}$$

where dz_1 and dz_2 may be correlated: $\sqrt{\sigma_{11}\sigma_{22}} \, dz_1 \, dz_2 = \sigma_{12} \, dt$.

The following result expresses the price-dividend ratio in terms of the hypergeometric function $F(a, b; c; z)$, which is defined in the region $|z| < 1$ by the power series

$$F(a, b; c; z) = 1 + \frac{a \cdot b}{1! \cdot c} z + \frac{a(a+1) \cdot b(b+1)}{2! \cdot c(c+1)} z^2 + \frac{a(a+1)(a+2) \cdot b(b+1)(b+2)}{3! \cdot c(c+1)(c+2)} z^3 + \cdots \,, \tag{2.17}$$

and in the region $|z| \geq 1$ by the integral representation

$$F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 w^{b-1}(1-w)^{c-b-1}(1-wz)^{-a} \, dw \quad \text{if Re } (c) > \text{Re } (b) > 0 \,.$$

Proposition 2.5 (The Brownian motion case)**.** *When log dividends are determined by equation* (2.16)*, the price-dividend ratio of the* $\boldsymbol{\alpha}$*-asset is given by*

$$\begin{aligned} P/D_1(s) \quad = \quad & \frac{1}{B(\lambda_1 - \lambda_2)} \left[\frac{1}{(\gamma/2 + \lambda_1) \, s^\gamma} \, F\left(\gamma, \gamma/2 + \lambda_1; 1 + \gamma/2 + \lambda_1; \frac{s-1}{s}\right) + \right. \\ & \left. + \frac{1}{(\gamma/2 - \lambda_2) \, (1-s)^\gamma} \, F\left(\gamma, \gamma/2 - \lambda_2; 1 + \gamma/2 - \lambda_2; \frac{s}{s-1}\right) \right] \end{aligned} \tag{2.18}$$

As before, $F(a, b; c; z)$ *is Gauss's hypergeometric function.*

---

[6] Under the Lévy process assumption, this is the unique case in which dividends are not subject to jumps. See Rogers and Williams (2000, pp. 76–77) for a proof.

The variables λ_1, λ_2, and B are given by

$$\begin{aligned} B &\equiv \frac{1}{2}X^2 \\ \lambda_1 &\equiv \frac{\sqrt{Y^2 + X^2Z^2} - Y}{X^2} \\ \lambda_2 &\equiv -\frac{\sqrt{Y^2 + X^2Z^2} + Y}{X^2}, \end{aligned}$$

where

$$\begin{aligned} X^2 &\equiv \sigma_{11} - 2\sigma_{12} + \sigma_{22} \\ Y &\equiv \mu_1 - \mu_2 + \alpha_1(\sigma_{11} - \sigma_{12}) - \alpha_2(\sigma_{22} - \sigma_{12}) - \frac{\gamma}{2}(\sigma_{11} - \sigma_{22}) \\ Z^2 &\equiv 2(\rho - \alpha_1\mu_1 - \alpha_2\mu_2) - (\alpha_1^2\sigma_{11} + 2\alpha_1\alpha_2\sigma_{12} + \alpha_2^2\sigma_{22}) + \\ &\quad + \gamma\left[\mu_1 + \mu_2 + \alpha_1\sigma_{11} + (\alpha_1 + \alpha_2)\sigma_{12} + \alpha_2\sigma_{22}\right] - \frac{\gamma^2}{4}(\sigma_{11} + 2\sigma_{12} + \sigma_{22}). \end{aligned}$$

and as the notation suggests, X^2 and Z^2 are strictly positive.

The instantaneous riskless rate is given by

$$\begin{aligned} r &= \rho + \gamma\left[s\left(\mu_1 + \frac{\sigma_{11}}{2}\right) + (1-s)\left(\mu_2 + \frac{\sigma_{22}}{2}\right)\right] - \\ &\quad - \frac{\gamma(\gamma+1)}{2}\left[s^2\sigma_{11} + 2s(1-s)\sigma_{12} + (1-s)^2\sigma_{22}\right]. \end{aligned} \tag{2.19}$$

Proof. See Appendix B.3 for the price-dividend ratio calculation. In the Brownian motion case, the riskless rate r is given by $r\,dt = -\mathbb{E}(dM/M)$, where $M_t \equiv e^{-\rho t}C_t^{-\gamma}$; (2.19) follows by Itô's lemma. □

Since (2.18) is not obviously more informative than the more general (2.9), which applies equally well to non-Brownian dividend processes, I do not supply a formula for the expected return although, given the above result, this can be calculated, after some algebra, along the same lines as the analogous calculation in Cochrane, Longstaff and Santa-Clara (2008).

Equation (2.18) generalizes the result of Cochrane, Longstaff and Santa-Clara (2008) (equation (50) in their paper) beyond the log utility special case. Intriguingly, these authors show that in some circumstances the price-dividend ratio takes on a simpler form. If, say, log dividends follow independent and symmetric Brownian motions with volatility σ, the time discount rate of the representative agent happens to equal σ^2, and we are in the log utility case, then the price-dividend ratio of asset 1 is

$$P/D_1 = \frac{1}{2\rho s}\left[1 + \left(\frac{1-s}{s}\right)\log(1-s) - \left(\frac{s}{1-s}\right)\log s\right] .$$

I show in Appendix B.3.1 how—and why—such simple expressions can be found in the Brownian motion case when parameters are chosen judiciously.

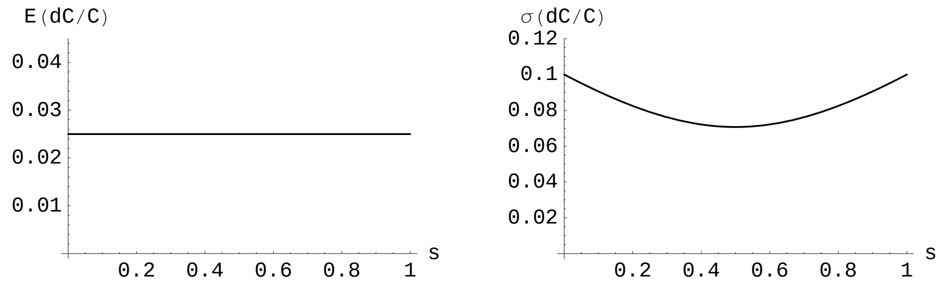
2.5 Two calibrations

I now present two simple calibrations. In each, the representative agent has time discount rate $\rho = 0.03$ and relative risk aversion $\gamma = 4$.

2.5.1 Dividends follow geometric Brownian motions

To explore the distinctive features of the model in a setting that is as simple as possible, consider a calibration in which the two assets are independent and have dividends which follow geometric Brownian motions. Each has mean log dividend growth of 2% and dividend volatility of 10%. In the notation of equation (2.2), $\mu_1 = \mu_2 = 0.02$, $\sigma_{11} = \sigma_{22} = 0.1^2$, and $\sigma_{12} = 0$.

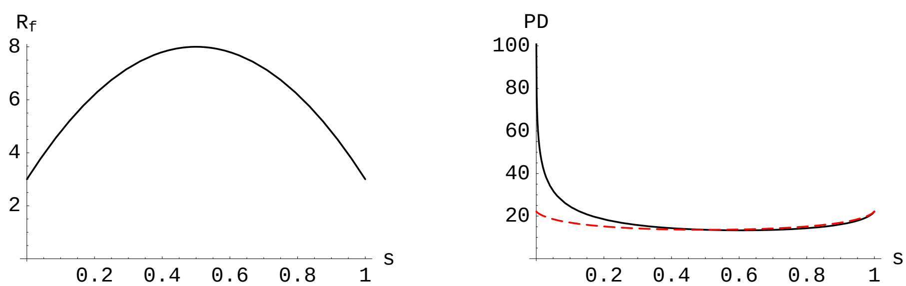
Although the dividend processes for the individual assets are i.i.d., consumption is *not* i.i.d., as documented in Figure 2.1. In this calibration, both assets have the same mean dividend growth, so mean consumption growth does not vary with s. But the standard deviation of consumption growth does vary: it is lower "in the



(a) $\mathbb{E}\left(\frac{dC}{C}\right)$ (b) $\sigma\left(\frac{dC}{C}\right)$

Figure 2.1: Left: Mean consumption growth, $\mathbb{E}(dC/C)$, against asset 1's dividend share, s. Right: The standard deviation of consumption growth, $\sigma(dC/C)$, against s.

middle", where there is most diversification. At the edges, where s is close to 0 or to 1, one of the two assets dominates the economy, and consumption growth is more volatile: the representative agent's eggs are all in one basket.



(a) Riskless rate (b) P/D

Figure 2.2: Left: The riskless rate against s. Right: The price-dividend ratio of asset 1 (solid) and of the market (dashed) against s.

Time-varying consumption growth volatility leads to a time-varying riskless rate. Figure 2.2a plots the riskless rate against asset 1's share of output s. Riskless rates are high for intermediate values of s because consumption volatility is low, which diminishes the motive for precautionary saving.

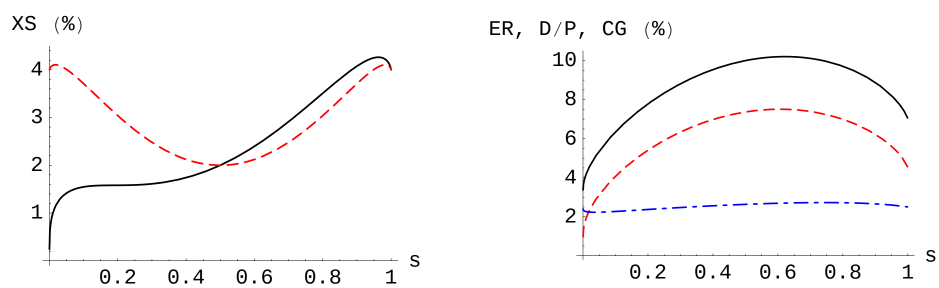
The right-hand graph, Figure 2.2b, shows the price-dividend ratio of asset 1 (solid) and of the market (dashed).[7] When asset 1 is a small part of the market,

[7] The market price-dividend ratio is calculated by observing that

$$\frac{P_1 + P_2}{D_1 + D_2} = s \cdot \frac{P_1}{D_1} + (1 - s) \cdot \frac{P_2}{D_2}.$$

it has a high valuation—P/D shoots up to the left of the figure—because it has very little systematic risk. As asset 1's share increases from $s = 0$, its discount rate increases both because the riskless rate increases and because its risk premium increases, as discussed further below.

Another notable feature of figure 2.2b is that the model predicts the existence of extreme growth assets (at the left of the figure) but not of extreme value assets.[8] This extreme growth case, which occurs as an asset's dividend share approaches zero, is of particular interest because it represents the most radical departure from a constant discount rate framework in which price-dividend ratios are constant; it is explored in more detail in section 2.6.



(a) Excess return on asset 1 (b) Expected return decomposition

Figure 2.3: Left: The excess return on asset 1 (solid) and on the market (dashed), against s. Right: Decomposition of expected returns (solid) into dividend yield (dashed) and expected capital gains (dot-dashed).

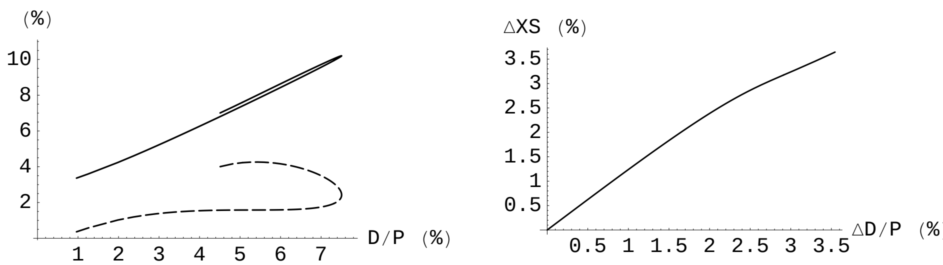
Figure 2.3a shows how the risk premium on asset 1 and on the market depends on the state variable s. Due to the diversification effect discussed above, the market risk premium is smallest when the two assets are of equal size. The risk premium on asset 1 increases as asset 1's dividend share increases. In the limit as s tends to zero, the risk premium on asset 1 tends to zero. The figure shows, however, that in this calibration even very small assets earn economically significant risk premia. In

[8] I use the term "growth" to refer to assets with high price-dividend ratios, and "value" to refer to assets with low price-dividend ratios.

other calibrations, idiosyncratic assets can earn strictly positive risk premia even in the limit in which they become negligible.

A comparison of figures 2.2b and 2.3a reveals that there is a value-growth effect: assets with high valuations earn low excess returns.[9]

Figure 2.3b decomposes expected returns into dividend yield plus expected capital gain. In this calibration, almost all cross-sectional variation in expected returns can be attributed to cross-sectional differences in dividend yield.



(a) Expected returns and expected excess returns on asset 1 against D/P.

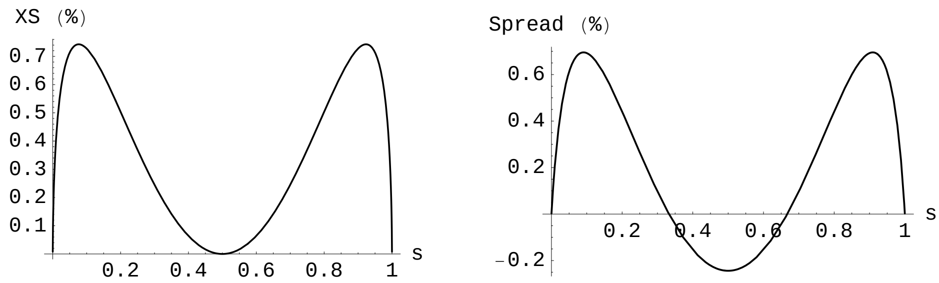
(b) Expected excess returns on the value-minus-growth strategy, plotted against the value spread.

Figure 2.4: Left: Expected returns (solid) and expected excess returns (dashed) on asset 1 against its dividend yield. Right: Expected excess return on the value-minus-growth strategy against the value spread.

Figure 2.4a makes this point in a different way, by plotting expected returns and risk premia against dividend yield. Figure 2.4b demonstrates that the excess return on a zero-cost investment in a value-minus-growth portfolio is increasing in the value spread (that is, the difference in dividend yield between the value and the growth asset). This echoes the empirical finding of Cohen, Polk and Vuolteenaho (2003) that "the expected return on value-minus-growth strategies is atypically high at times when their spread in book-to-market ratios is wide."

It is also of interest to consider the behavior of assets in zero net supply, such

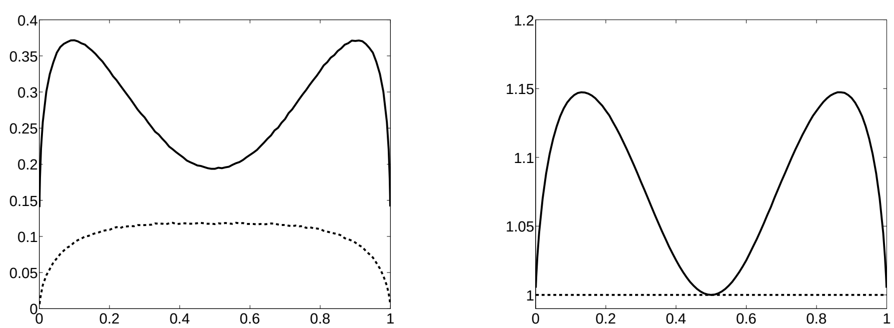
[9] Of course, in this simple example, there are only two assets in the cross-section. But the results of Section 2.7 confirm that price-dividend ratios are state-dependent with N assets. Since dividend growth is i.i.d., high price-dividend ratios forecast low expected returns in the general case. See Cochrane (2005), p. 399.



(a) Excess returns on a perpetuity. (b) The yield spread.

Figure 2.5: A high yield spread, $\mathscr{Y}(30)-\mathscr{Y}(0)$, signals high expected excess returns on a perpetuity.

as perpetuities and zero coupon bonds. Figure 2.5a plots the risk premium on a real perpetuity which pays one unit of consumption good per unit time. Figure 2.5b shows how the spread in yields between a 30-year zero-coupon bond and the instantaneous riskless rate varies with s. A high yield spread forecasts high excess returns on long-term bonds. Looking back at figure 2.3a, we see that a high yield spread also forecasts high excess returns on the market.



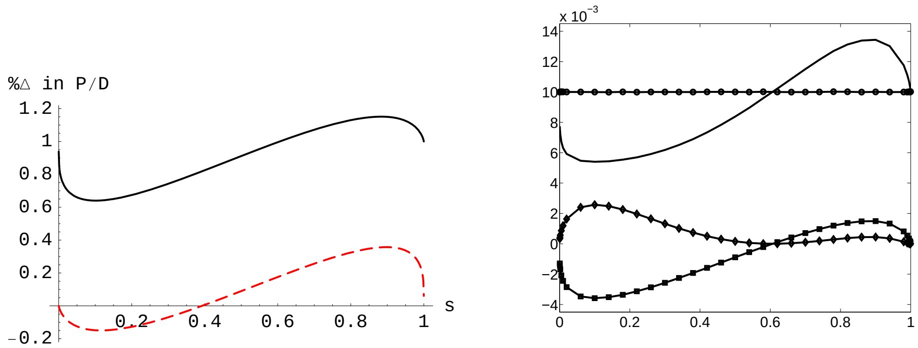
(a) Correlation between asset returns. (b) Excess volatility on the market.

Figure 2.6: Left: The correlation between the returns of asset 1 and asset 2 against s. Right: The ratio of market return volatility to dividend volatility against s. Solid lines, $\gamma = 4$; dashed lines, $\gamma = 1$.

Figure 2.6a demonstrates that the model generates significant comovement between the returns of the two assets, even though the two assets have independent fundamentals.[10] There is considerably more comovement when $\gamma = 4$ than in the

[10] These figures, unlike the preceding ones, are calculated by Monte Carlo methods, as follows.

log utility case. Figure 2.6b shows that the model generates excess volatility in the aggregate market when $\gamma > 1$. (When $\gamma = 1$—the log utility case, indicated with a dashed line—there is no excess volatility because the price-dividend ratio of the aggregate market is constant. For the same reason, there is no excess volatility in the $\gamma = 4$ case when $s = 1/2$: the market price-dividend ratio is locally flat, as a function of s, at this point.)



(a) Response to cashflow shock to asset 1.

(b) A variance decomposition.

Figure 2.7: Left: The response of asset 1 (solid) and asset 2 (dashed) to a +1% increase in the dividend of asset 1. Right: Decomposition of the variance of returns (solid) into three parts: the variance of dividend-driven returns (circles), the variance of valuation-driven returns (diamonds) and the covariance between the two types of returns (squares).

What drives asset 1's returns? In the two-asset case, two types of shock move an asset's price: a shock to its dividends, or a shock to the other asset's dividends, which changes the asset's price by changing its price-dividend ratio. In the terminology

For each of 109 different starting values of $s \in [0, 1]$, I generate 4000 sample paths of log dividends. (The 109 different values are the points 0.01, 0.02, ..., 0.99, five points between 0 and 0.01, and five points between 0.99 and 1.) Each sample path simulates a drifting Brownian motion over a very short time horizon: 3×10^{-5} years, slightly less than 16 minutes. Over this time horizon, each drifting Brownian motion is simulated by dividing the interval into 600 time steps; Normal random variables determine the evolution of log dividends between these time steps. Given a particular sample path for dividends, prices can be calculated, given the price-dividend functions; and hence also total returns, and the covariance matrix of realized returns on the two assets. Finally, I estimate variances and covariance between the two assets, at each value of s, by averaging over the covariance matrices estimated for each of the 4000 sample paths.

of Campbell (1991), the first type of shock corresponds to the arrival of "cashflow news" and the second to the arrival of "discount-rate news". Figure 2.7a plots the percentage price response of asset 1 (solid) and asset 2 (dashed) to a 1% increase in asset 1's dividends. When asset 1 is small, it underreacts to good news about its own cashflow shock: the price response is considerably less than 1%. At the same time, asset 2 moves in the opposite direction. When asset 1 is large, it overreacts to good news about its own cashflow shock, and asset 2 moves in the same direction. Note also that asset 2's price moves considerably more, in response to dividend news for asset 1, when asset 1 is large than when asset 1 is small.

A better understanding of these effects can be gained by exploiting a simple identity that breaks realized returns on any asset into two pieces:

$$
\begin{aligned}
R_{t+1} &= \frac{P_{t+1} + D_{t+1}}{P_t} \\
&= \underbrace{\frac{D_{t+1}}{D_t}\left(1 + \frac{D_t}{P_t}\right)}_{\text{dividend-driven}} + \underbrace{\frac{D_{t+1}}{D_t}\frac{D_t}{P_t}\left(\frac{P_{t+1}}{D_{t+1}} - \frac{P_t}{D_t}\right)}_{\text{valuation-driven}} \\
&\equiv R_{D,t+1} + R_{V,t+1} \, . \qquad (2.20)
\end{aligned}
$$

The last line defines the *dividend-driven return* $R_{D,t+1}$ and the *valuation-driven return* $R_{V,t+1}$. In an economy in which price-dividend ratios are constant—for example, one with a single Lucas tree with i.i.d. dividend growth and a representative agent with power utility or Epstein-Zin preferences—the valuation-driven component disappears, and returns are exclusively dividend-driven. The above identity holds exactly, with no log-linearizations needed. It is similar to the decomposition of Campbell (1991), but changes in price-dividend ratio appear on the right-hand side of (2.20), as opposed to the changes in future returns that appear in Campbell's decomposition. It has the advantage that the two components can be estimated directly from historical data.

Using the decomposition in (2.20), we have

$$\operatorname{var}_t R_{t+1} = \operatorname{var}_t R_{D,t+1} + \operatorname{var}_t R_{V,t+1} + 2\operatorname{cov}_t(R_{D,t+1}, R_{V,t+1}), \tag{2.21}$$

an equation that provides another way to think about the sources of variation in expected returns.

In the continuous-time case relevant for my purposes here, the above equations are modified slightly: we have

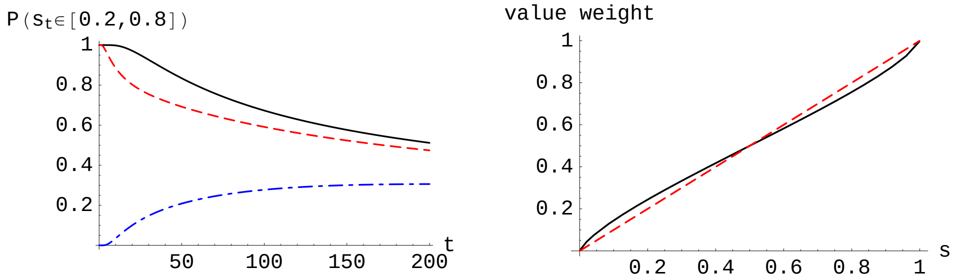
$$R_{t+dt} = \frac{D_{t+dt}}{D_t}\left(1 + \frac{D_t}{P_t}dt\right) + \frac{D_{t+dt}}{D_t}\frac{D_t}{P_t}\left(\frac{P_{t+dt}}{D_{t+dt}} - \frac{P_t}{D_t}\right),$$

and again the first term on the right-hand side can be thought of as the dividend-driven return $R_{D,t+dt}$ and the second as the valuation-driven return $R_{V,t+dt}$.

I estimate the three components, $\operatorname{var}_t R_{D,t+dt}$, $\operatorname{var}_t R_{V,t+dt}$, and $\operatorname{cov}_t(R_{D,t+dt}, R_{V,t+dt})$ by simulating the underlying Brownian processes as described in Footnote 10. The results are shown in figure 2.7b. The figure shows that (i) most of the variance in asset returns is driven by cash-flow news, (ii) dividend-driven returns and valuation-driven returns are negatively correlated for small assets and positively correlated for large assets,[11] (iii) for large assets, a far higher proportion of variation in expected returns is due to cashflow news than to discount rate news, while (iv) for small assets, valuation-driven returns are much more important: the variance of dividend-driven returns is only about four times higher than the variance of valuation-driven returns.

Figure 2.8 plots the probability that the dividend share at time t, s_t, remains in the region $[0.2, 0.8]$ for t between 0 and 200 years, and for starting shares $s_0 = 0.1, 0.3, 0.5$. (The cases $s_0 = 0.7, 0.9$ can also be read off the graph, because the

[11] In the language of Campbell (1991), cashflow news and discount-rate news are *positively* correlated for small assets and negatively correlated for large assets.



(a) $\mathbb{P}(s_t \in [0.2, 0.8])$

(b) The value weight (solid) and the 45 degree line (dashed)

Figure 2.8: Left: In calibration 1, the probability that s_t lies between 0.2 and 0.8, plotted against time t, measured in years, assuming starting shares $s_0 = 0.5$ (solid), $s_0 = 0.3$ (dashed), and $s_0 = 0.1$ (dot-dashed). Right: In calibration 1, the value weight of asset 1 (solid), and the 45 degree line (dashed), against s.

world is symmetric.) It also plots the value weight of asset 1 in the aggregate market against s.

2.5.2 *Dividends are subject to occasional disasters*

The second calibration is intended to highlight the effect of disasters. Again, the two assets are symmetric for simplicity. In the notation of equation (2.2), the drifts are $\mu_1 = \mu_2 = 0.02$. The two Brownian motions driving dividends are independent and each has volatility of 2%, so $\sigma_{11} = \sigma_{22} = 0.02^2$ and $\sigma_{12} = 0$.

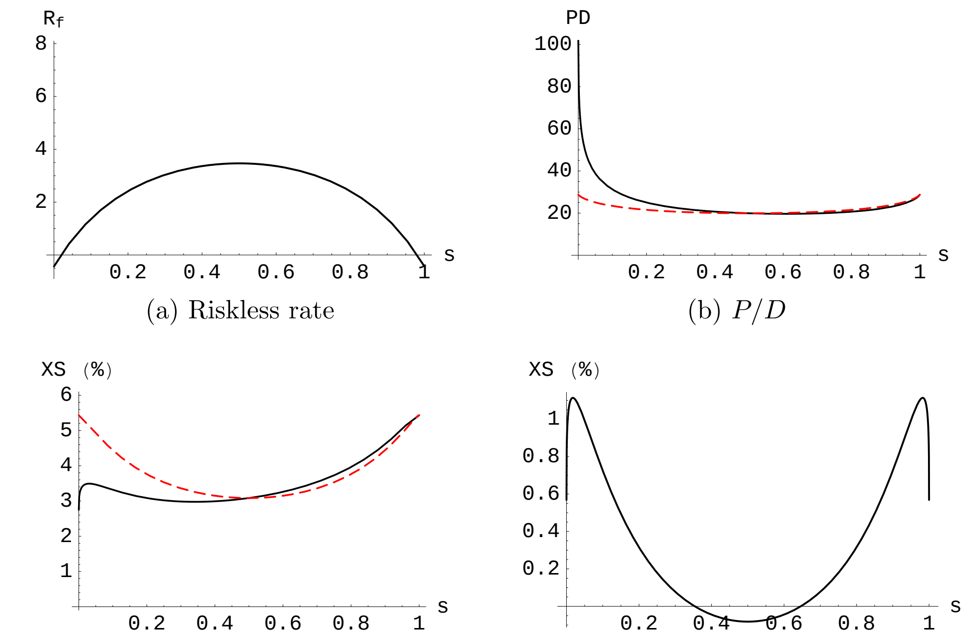
There are also jumps in dividends, caused by the arrival of disasters, of which there are three types. One type affects only asset 1: it arrives at times dictated by a Poisson process with rate 0.017/2. When the disaster strikes, it shocks log dividends by a Normal random variable with mean −0.38 and standard deviation 0.25. The second is exactly the same, except that it affects only asset 2. The third type arrives at rate 0.017/2 and shocks the log dividends of *both* assets by the same amount,[12] which is, again, a random variable with mean −0.38 and standard deviation of 0.25.

[12] These disasters are therefore simultaneous and of perfectly correlated—in fact, identical—sizes; the framework also easily handles the case in which disasters are simultaneous but uncorrelated or imperfectly correlated.

If the two assets are thought of as claims to a country's output, then the first two types are examples of local disasters while the third is a global disaster.

From the perspective of either asset, then, disasters occur at rate $0.017/2 + 0.017/2 = 0.017$: on average, about once every 60 years. There is a 50-50 chance that any given disaster is local or global. These disaster arrival rates—and the mean and standard deviation of the disaster sizes—are chosen to match exactly the empirical disaster frequency estimated by Barro (2006a), and to match approximately the disaster size distribution documented in the same paper.

Taking everything into account, these parameter values imply an unconditional mean dividend growth rate (in levels, not logs) of 1.6%. Conditional on disasters not occurring, the mean dividend growth rate is 2.0%.



(a) Riskless rate

(b) P/D

(c) Excess returns on asset 1 and the market

(d) Excess returns on a perpetuity

Figure 2.9: The riskless rate; price-dividend ratio on asset 1 (solid) and on the market (dashed); excess returns on asset 1 (solid) and on the market (dashed); and excess returns on a perpetuity.

Figure 2.9 exhibits the central features of asset prices and returns in this calibration. In broad outline, the pictures are very similar to those presented previously—

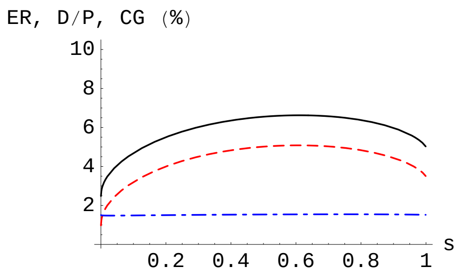
and for the same reasons—but some new features stand out. The riskless rate is lower across the range of values of s. Also, despite considerably lower Brownian volatility, the presence of jumps induces a higher risk premium, both at the individual asset level and at the market level. As in Rietz (1988) and Barro (2006a), incorporating rare disasters makes it easier to match the observed riskless rate and equity premium. A more unusual feature is that disasters can propagate to apparently safe assets: since the state variable can jump, interest rates can jump, and hence bond prices can jump. Consequently, at times when the current riskless rate is low (for $s \approx 0$ or $s \approx 1$), the risk premium on a perpetuity is significantly higher than previously, despite the fact that disasters do not affect its cashflows. A perpetuity earns a negative risk premium near $s = 1/2$, since in this state long-dated bonds act as a hedge against disasters: when a disaster strikes one of the assets, riskless rates drop and the price of a long-dated bond jumps up.

Figure 2.10 shows an expected return decomposition; expected returns and risk premia against dividend yield; price responses to a 1% dividend shock to asset 1; and the yield spread. The qualitative features are substantially the same as in the previous calibration in each case.

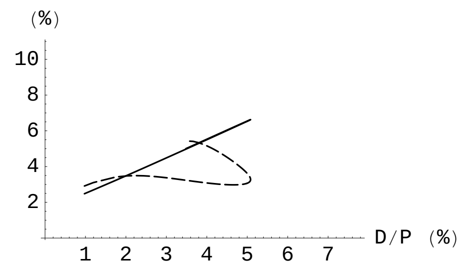
In the presence of jumps, the cross-asset effects present in the previous calibration become more pronounced. Notably, disasters propagate across assets.

This is shown graphically in Figure 2.11, which plots a single sample time series. Time, along the x-axis, runs from 0 to 60 years. The sequence of figures should be read clockwise, starting from the top left. Asset 1 (in red) is the small asset—with an initial dividend share of 10%. Asset 2 is shown in black. From exogenous dividend processes we calculate the dividend share of asset 1, and hence price-dividend ratios. Finally, from dividends and price-dividend ratios, we calculate prices.

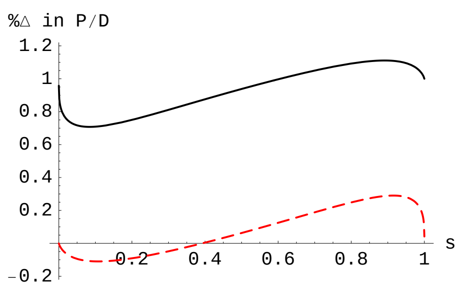
In the particular realization shown here, each asset suffers one negative shock to fundamentals; there is no "global" shock. When the large asset suffers its disaster,



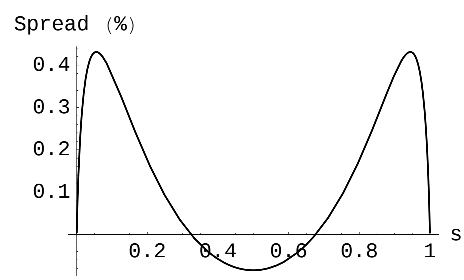
(a) Expected return decomposition



(b) Expected returns (solid) and expected excess returns (dashed) against dividend yield



(c) Responses to a 1% dividend shock to asset 1



(d) Yield spread $\mathscr{Y}(30) - \mathscr{Y}(0)$

Figure 2.10: Figure 2.10a decomposes expected returns (solid) into dividend yield (dashed) and expected capital gains (dot-dashed). Figure 2.10b plots expected returns (solid) and expected excess returns (dashed) on asset 1 against its dividend yield. Figure 2.10c has the response of asset 1 (solid) and asset 2 (dashed) to a 1% dividend shock to asset 1.

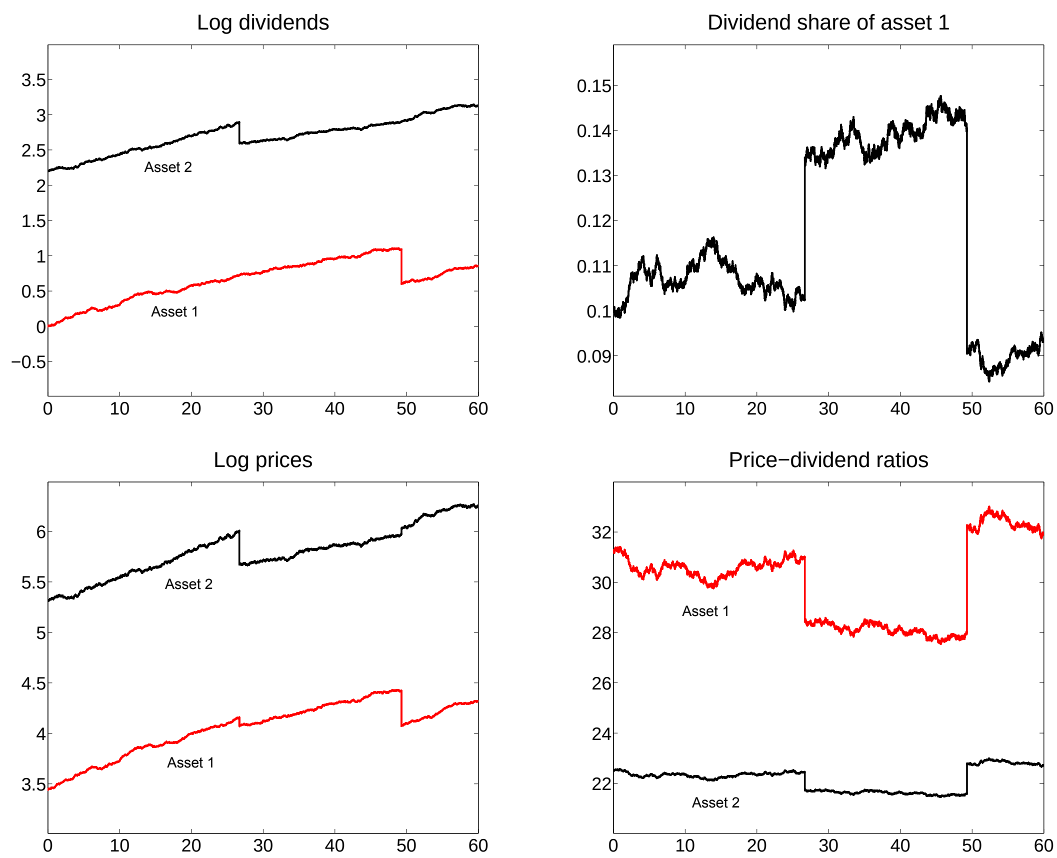


Figure 2.11: Dividends, dividend share, prices, and price-dividend ratios against time.

after about 26 years, its dividend drops by 25% and its price drops by 28%. Two forces act on the small asset. A disaster to the large asset makes the economy more balanced, so riskless rates jump up; at the same time, the risk premium on the small asset jumps up because it is a larger part of the economy. These effects act in the same direction, and the small asset experiences a downward price jump of 8.2%: contagion.

When the small asset suffers its disaster, after about 49 years, its dividend drops by 39% and its price drops by 30%. Now, two *opposing* forces act on the large asset. On one hand, its risk premium rises as it is a larger share of the market. On the other, the riskless rate declines in response to the increasingly unbalanced world. The riskless rate effect dominates, and the large asset experiences an upward price jump of 5.7%: flight-to-quality.

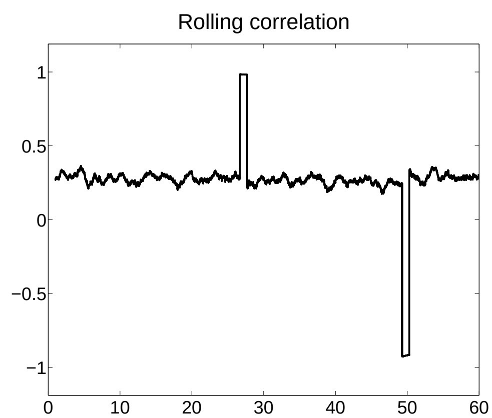
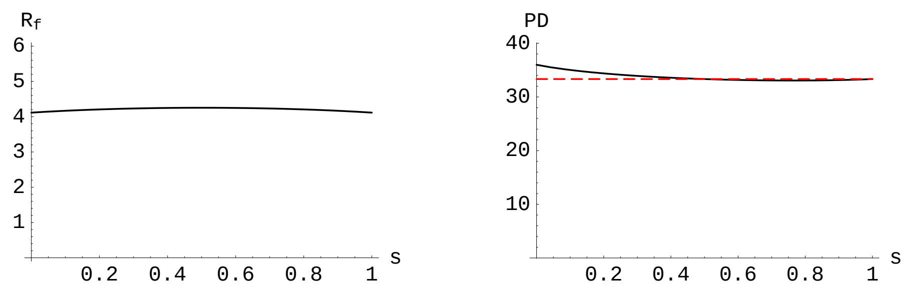


Figure 2.12: The one-year rolling correlation between assets 1 and 2, calculated along the sample path of Figure 2.11.

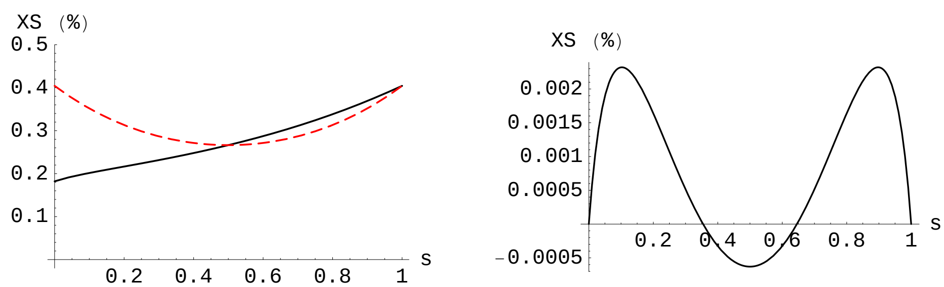
We can also calculate rolling 1-year realized return correlations along this sample path, as shown in Figure 2.12. During normal times, the correlation hovers around 0.3, despite the fact that, conditional on no jumps, the two assets have independent dividend streams. When the first disaster ("contagion") takes place, the measured correlation spikes up almost as far as +1 due to the spectacular outlying return.

When the second disaster ("flight-to-quality") takes place, the measured correlation spikes down almost as far as -1. Despite the fact that naively calculated correlations display occasional spikes, the correlation between the two assets, conditional on some given s, is constant over time—and is economically significant even if one conditions on jumps not taking place. These results are therefore reminiscent of the findings of Forbes and Rigobon (2002), who demonstrate that although naively calculated correlations spike at times of crisis, once one corrects for the heteroskedasticity induced by high market volatility at times of crisis, it can be seen that markets have a high level of "interdependence" in all states of the world.



(a) Riskless rate (b) P/D

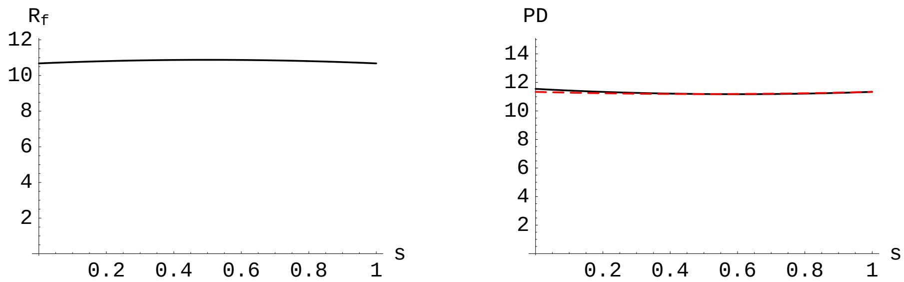
Figure 2.13: Left: The riskless rate against s. Right: The price-dividend ratio of asset 1 (solid) and of the market (dashed), against s. Log utility with jumps.



(a) Excess return on asset 1 (b) Excess return on a perpetuity

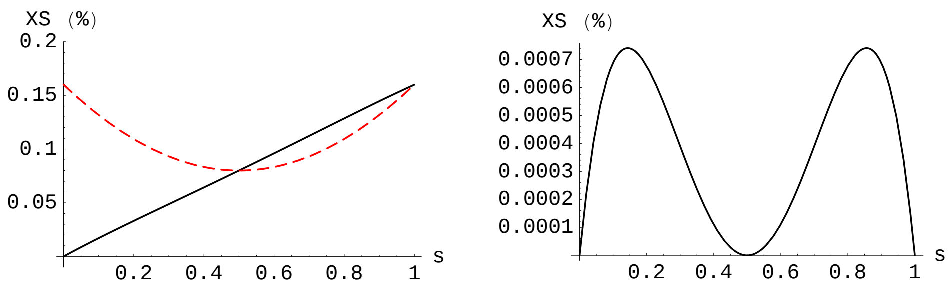
Figure 2.14: Left: The excess return on asset 1 (solid) and on the market (dashed) against s. Right: The excess return on a perpetuity against s. Log utility with jumps.

Figures 2.13–2.18 explore the consequences of using log utility or removing jumps from the second calibration. In Figures 2.13–2.14, $\gamma = 1$ and there are jumps; in



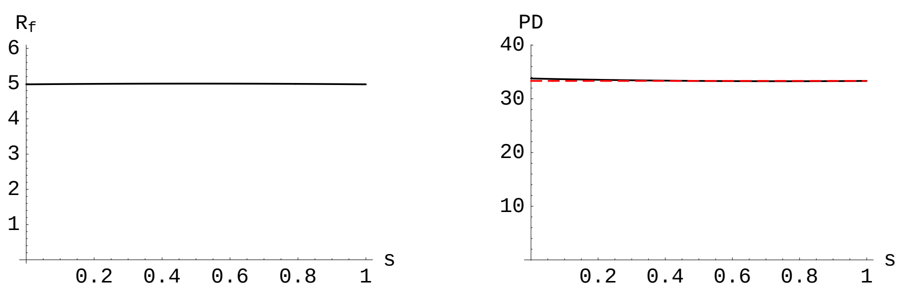
(a) Riskless rate (b) P/D

Figure 2.15: Left: The riskless rate against s. Right: The price-dividend ratio on asset 1 (solid) and on the market (dashed), against s. $\gamma = 4$, no jumps.



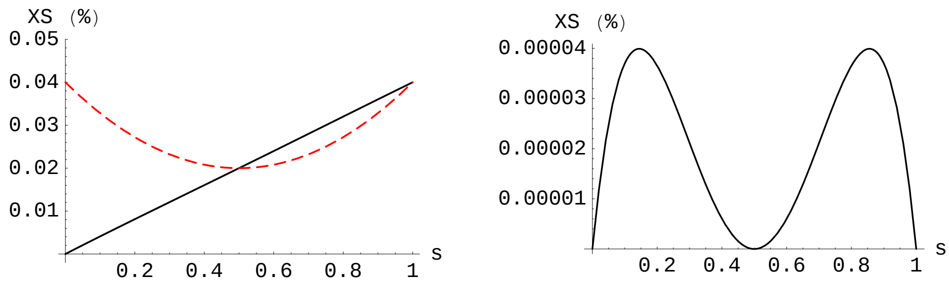
(a) Excess return on asset 1 (b) Excess return on a perpetuity

Figure 2.16: Left: The excess return on asset 1 (solid) and on the market (dashed), against s. Right: The excess return on a perpetuity against s. $\gamma = 4$, no jumps.



(a) Riskless rate (b) P/D

Figure 2.17: Left: The riskless rate against s. Right: The price-dividend ratio of asset 1 (solid) and of the market (dashed), against s. Log utility, no jumps.



(a) Excess return on asset 1 (b) Excess return on a perpetuity

Figure 2.18: Left: The excess return on asset 1 (solid) and on the market (dashed) against s. Right: The excess return on a perpetuity against s. Log utility, no jumps.

Figures 2.15–2.16, $\gamma = 4$ but there are no jumps; in Figures 2.17–2.18, $\gamma = 1$ and there are no jumps.

In all cases, the results are quantitatively uninteresting: both high risk aversion ($\gamma \approx 4$) and occasional disasters are needed to generate interesting predictions using parameter values normally considered reasonable in the consumption-based asset pricing literature.

2.6 Equilibrium pricing of small assets

A distinctive qualitative prediction of the model is that there should exist extreme growth assets, but not extreme value assets. (Look back at the left-hand side of Figure 2.2b.) The extreme growth case also represents the starkest departure from simple models in which price-dividend ratios are constant (as, for example, in a one-tree model with power utility and i.i.d. dividend growth). Furthermore, it is important to understand whether the complicated dynamics exhibited above are relevant for small assets.[13] These considerations lead me to investigate the price behavior of asset 1 in the limit $s \to 0$ in which it becomes tiny relative to the rest of the market (that is, asset 2).

To preview the results, consider the problem of pricing a negligibly small asset,

[13] This is analogous to the "small-country case" in international finance.

whose fundamentals are independent of all other assets, in an environment in which the (real) riskless rate is 6%. If the small asset has mean dividend growth rate of 4%, the following logic seems plausible. Since the asset is negligibly small, it need not earn a risk premium, so the appropriate discount rate is the riskless rate. Next, since dividends are supposed throughout to be i.i.d., it seems sensible to apply the Gordon growth model to conclude that for this small asset,

$$\begin{aligned}\text{dividend yield} &= \text{riskless rate} - \text{mean dividend growth} \\ &= 6\% - 4\% \\ &= 2\%\,.\end{aligned}$$

It turns out that this argument can be made formal; I do so below.

Now, consider the (more realistic) situation in which the riskless real rate is 2%. If the asset does not earn a risk premium, Gordon growth logic seems to suggest that the dividend yield should be $2\% - 4\% = -2\%$, an obviously nonsensical result. I show below how to value assets in situations such as these, in which the Gordon growth model breaks down. In the limit, such an asset has a price-consumption ratio of zero, as one would expect. More surprisingly, though, it has an infinite price-dividend ratio—reminiscent of Pástor and Veronesi (2003, 2006)—and a strictly positive risk premium. Moreover, since the dividend yield is zero, expected returns on the asset are entirely attributable to expected capital gains.

I now return to the general setup in which the assets may have correlated dividend growth and make a pair of definitions.

Definition 2.3. *If the inequality*

$$\rho - \boldsymbol{c}(1, -\gamma) > 0 \tag{2.22}$$

holds then we are in the subcritical *case.*

If the reverse inequality

$$\rho - \boldsymbol{c}(1, -\gamma) < 0 \tag{2.23}$$

holds then we are in the supercritical *case.*[14]

In the supercritical case, define θ^* to be the unique $\theta \in (0,1)$ which satisfies

$$\rho - \boldsymbol{c}(1-\theta, \theta-\gamma) = 0\,. \tag{2.24}$$

In the supercritical case we have $\theta^* \in (0,1)$ because equation (2.24) is negative at $\theta = 0$ by (2.23) and positive at $\theta = 1$ by the finiteness assumptions in Table 2.1. In the Brownian motion case, (2.24) is simply a quadratic equation in θ. More generally, the fact that the solution is unique follows from the fact, proved in Appendix B.4, that $\rho - \boldsymbol{c}(1-\theta, \theta-\gamma)$ is a concave function of θ.

The next two Propositions supply various asymptotics. To highlight the link with the traditional Gordon growth formula, I write $G_1 \equiv \boldsymbol{c}(1,0) = \log \mathbb{E} D_{11}/D_{10}$ and $G_2 \equiv \boldsymbol{c}(0,1) = \log \mathbb{E} D_{21}/D_{20}$ for (log) mean dividend growth, and R_1 and R_2 for the expected instantaneous returns on assets 1 and 2.

Proposition 2.6. *In the subcritical case, in the limit as $s \downarrow 0$, we have*

$$R_f = \rho - \boldsymbol{c}(0, -\gamma) \tag{2.25}$$

$$R_1 = \rho - \boldsymbol{c}(1, -\gamma) + \boldsymbol{c}(1,0) \tag{2.26}$$

$$D/P_1 = R_1 - G_1 \tag{2.27}$$

[14] There is also a third case, the *critical* case in which $\rho - \boldsymbol{c}(1, -\gamma) = 0$; I omit it for the sake of brevity. Briefly, price-dividend ratios are asymptotically infinite and excess returns asymptotically zero, assuming independent dividend growth. The simple example presented in Section 1 of Cochrane, Longstaff and Santa-Clara (2008) is precisely critical. This is no coincidence: the condition that implies criticality also ensures that the expression for the price-dividend ratio is relatively simple. See Appendix B.3.1.

If the two assets are independent, then in the limit $R_f = R_1 < R_2$.

Proof. See Appendix B.4. □

The results of Proposition 2.6 correspond to the first example above. A small idiosyncratic asset with i.i.d. dividend growth earns no risk premium, and can be valued with the Gordon growth model (2.27). The next result shows that this is not the whole story: more intriguing behavior may emerge.

Proposition 2.7. *In the supercritical case, in the limit as* $s \downarrow 0$*, we have*

$$R_f = \rho - \boldsymbol{c}(0, -\gamma) \tag{2.28}$$

$$R_1 = \boldsymbol{c}(1 - \theta^*, \theta^*) \tag{2.29}$$

$$D/P_1 = 0 \tag{2.30}$$

If the two assets are independent, then in the limit $R_f < R_1 < R_2$*. If* $G_1 \geq G_2$*, then in the limit* $R_1 < G_1$*, whether or not the assets are independent.*

Proof. See Appendix B.4. □

These results are much more surprising. To understand what is going on, consider the case in which dividend growth is independent across assets, so that the risk in question is both small and idiosyncratic. Proposition 2.7 demonstrates that in the supercritical regime, such an asset has an enormous valuation ratio and earns a strictly positive risk premium.

A naive attempt to apply the Gordon growth model breaks down in the supercritical case because (2.23) holds, so the riskless rate minus dividend growth is *negative.* Nonetheless, the asymptotically small asset still has a well-defined dividend-price ratio and expected return, as demonstrated in Proposition 2.7. What happens to the price in the asymptotic limit?

The first point is that this is not quite the right question. Suppose that we are in the supercritical scenario, and imagine holding the dividend of asset 1 fixed while allowing the dividend of asset 2 (and hence total consumption) to increase without limit. Since s then tends to zero, this is one way asset 1 can become "small". Because D_1 is held constant, the price of asset 1—measured, as always, in units of consumption—is unbounded in this limit. A more informative question is to ask for the asymptotic behavior of the price-consumption ratio.

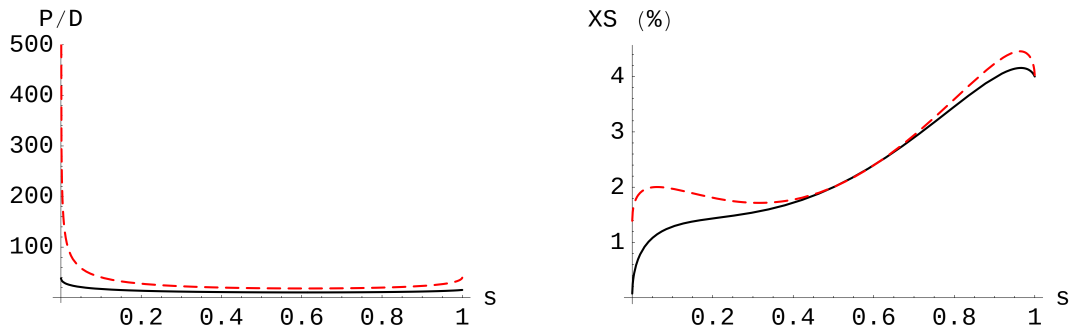
Alternatively, imagine holding the dividend of asset 2 fixed while the dividend of asset 1, and hence s, tends to zero. The price-dividend ratio goes to infinity, but the dividend goes to zero: what happens to the price? The answer is that since consumption remains finite in this example, the price is zero, finite or infinite in the limit depending on whether the price-consumption ratio is zero, finite or infinite in the limit.

In short, it is useful to focus on the price-consumption ratio, $P/C = s \cdot P/D$. Appendix B.4 shows that the fact that the price-consumption ratio is zero in the limit follows from the fact that $\theta^* < 1$.

Examination of the subcritical condition (2.22) and supercritical condition (2.23) reveals that the supercritical regime occurs whenever ρ is sufficiently small. More generally, the supercritical regime is relevant in environments in which the riskless rate is low.

I now exhibit these phenomena in the simple Brownian motion example considered earlier in the chapter. This will make it clear that, first, the supercritical case is neither pathological nor dependent on extreme parameter values and, second, the size of the strictly positive excess return earned on the small asset in the supercritical case is economically meaningful. To recap, the world is symmetric, and the two assets are independent with 2% mean dividend growth and 10% dividend volatility.

As usual, $\gamma = 4$. If $\rho = 0.05$, then we are in the subcritical case.[15] If on the other hand $\rho = 0.01$, we are in the supercritical case.



(a) Asset 1's price-dividend ratio

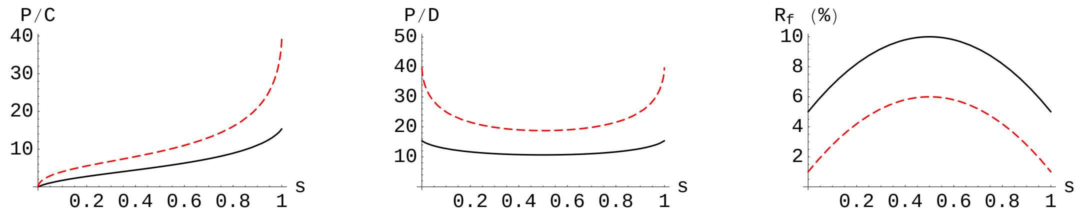
(b) Asset 1's excess return

Figure 2.19: Left: Price-dividend ratio of asset 1 against s. Right: Excess return of asset 1 against s. Supercritical case is dashed, subcritical case is solid.

Figure 2.19 shows the price-dividend ratio and excess return of asset 1 against s. The asymptotic limits are to the left of the graph, as $s \downarrow 0$. In the subcritical case, the price-dividend ratio remains below 40 for all s and the excess return tends to zero. In the supercritical case, the price-dividend ratio explodes and the excess return tends to roughly 1.3 per cent. (Notice also that for intermediate values of the state variable, the risk premium on asset 1 is not sensitive to the value of ρ, as would be the case in a standard one-tree model.) Asymptotically, the dividend yield is zero, so all of the expected return of the small asset can be attributed to expected capital gains.

Finally, to allay suspicions that something strange is going on in the background, Figure 2.20 demonstrates that asset 1's price-consumption ratio, the market price-dividend ratio and the riskless rate are all well-behaved in the limit.

[15] In the calibration presented earlier, I set $\rho = 0.03$. This case is also subcritical. I have chosen to use $\rho = 0.05$ here in order to make the distinction between the two cases clearer in the figures.



(a) Asset 1's price-consumption ratio (b) Market price-dividend ratio (c) Riskless rate

Figure 2.20: The price-consumption ratio of asset 1, market price-dividend ratio and riskless rate plotted against asset 1's share of output, s. Supercritical case is dashed, subcritical case is solid.

2.7 N assets

The general results presented in Section 2.3 can be generalized to the case in which the representative agent's consumption stream is provided by the output of N assets, $C_t = D_{1t} + D_{2t} + \cdots + D_{Nt}$.

With this modification, equations (2.1)–(2.3) are unchanged, except that boldface vectors are now understood to have N entries, as opposed to just two. The fundamental ideas underlying the calculation are also the same. The main technical difficulty lies in calculating $\mathscr{F}_\gamma^N(\boldsymbol{v}) \equiv \mathscr{F}_\gamma^N(v_1, \ldots, v_{N-1})$, the generalization of $\mathscr{F}_\gamma(v)$ to the N-asset case. It turns out that we have

$$\mathscr{F}_\gamma^N(\boldsymbol{v}) = \frac{\Gamma\left(\gamma/N + iv_1 + iv_2 + \ldots + iv_{N-1}\right)}{(2\pi)^{N-1}\Gamma(\gamma)} \cdot \prod_{k=1}^{N-1} \Gamma\left(\gamma/N - iv_k\right). \tag{2.31}$$

Before stating the main result, it will be useful to recall some old, and to define some new, notation. Let $\boldsymbol{e}_j$ be an N-vector with a one at the jth entry and zeros elsewhere, and define the N-vectors $\boldsymbol{y}_0 \equiv (y_{10}, \ldots, y_{N0})'$ and $\boldsymbol{\gamma} \equiv (\gamma, \ldots, \gamma)'$, and the $(N-1)\times N$ matrix $\boldsymbol{U}$ and the $(N-1)$-vector $\boldsymbol{u}$ by

$$\boldsymbol{U} \equiv \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{u} \equiv \begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ u_N \end{pmatrix} \equiv \boldsymbol{U}\boldsymbol{y}_0 \,. \tag{2.32}$$

In the two-asset case, there was one state variable. We worked with s, the dividend share of asset one, or with $u = \log(1-s)/s = y_{20} - y_{10}$. With N assets, there are $N-1$ state variables. One natural set of state variables is $\{s_i\}$, $i = 1, \ldots, N-1$, where

$$s_i = \frac{D_{i0}}{D_{10} + \cdots + D_{N0}}$$

is the dividend share of asset i; in fact, though, it turns out to be more convenient to work with the $(N-1)$-dimensional state vector $\boldsymbol{u}$. The first entry of $\boldsymbol{u}$ is $u_2 = y_{20} - y_{10}$, which corresponds to the state variable u of previous sections. More generally, $u_k = y_{k0} - y_{10}$ is a measure of the size of asset k relative to asset 1. Consistent with this notation, I will also write $u_1 \equiv y_{10} - y_{10} = 0$ and define the N-vector $\boldsymbol{u}_+ \equiv (u_1, u_2, \ldots, u_N)' = (0, u_2, \ldots, u_N)'$ to make subsequent formulas easier to read.

The following Proposition generalizes earlier integral formulas to the N-asset case. All integrals are over $\mathbb{R}^{N-1}$: $\boldsymbol{v}$ is an $(N-1)$-vector. Again, they can be evaluated on the computer. The condition that ensures finiteness of the price of asset j is that

$$\rho - \boldsymbol{c}(\boldsymbol{e}_j - \boldsymbol{\gamma}/N) > 0 \,.$$

I assume that this condition holds for all assets j.

Proposition 2.8 (Integral formulas in the N-asset case)**.** *The price-dividend ratio* *on asset j is*

$$P/D = e^{-\boldsymbol{\gamma}'\boldsymbol{u}_+/N}\left(e^{u_1}+\cdots+e^{u_N}\right)^{\gamma}\int \frac{\mathscr{F}_{\gamma}^{N}(\boldsymbol{v})e^{i\boldsymbol{u}'\boldsymbol{v}}}{\rho - \boldsymbol{c}(\boldsymbol{e}_j - \boldsymbol{\gamma}/N + i\boldsymbol{U}'\boldsymbol{v})}\,d\boldsymbol{v}\,. \tag{2.33}$$

Defining the expected return by $ER\,dt \equiv \mathbb{E}(dP + D\,dt)/P$, we have

$$ER = \frac{\Phi}{P/D} + D/P\,, \tag{2.34}$$

where

$$\Phi = \sum_{\boldsymbol{m}} \binom{\gamma}{\boldsymbol{m}} e^{(\boldsymbol{m}-\boldsymbol{\gamma}/N)'\boldsymbol{u}_+}\int \frac{\mathscr{F}_{\gamma}^{N}(\boldsymbol{v})e^{i\boldsymbol{u}'\boldsymbol{v}}\boldsymbol{c}(\boldsymbol{e}_j + \boldsymbol{m} - \boldsymbol{\gamma}/N + i\boldsymbol{U}'\boldsymbol{v})}{\rho - \boldsymbol{c}(\boldsymbol{e}_j - \boldsymbol{\gamma}/N + i\boldsymbol{U}'\boldsymbol{v})}\,d\boldsymbol{v}\,.$$

The summation is over all vectors $\boldsymbol{m} = (m_1, \ldots, m_N)'$ whose entries are non-negative and add up to γ. I have made use of the multinomial coefficient

$$\binom{\gamma}{\boldsymbol{m}} = \frac{\gamma!}{m_1!\cdots m_N!}\,.$$

The zero-coupon yield to time T is

$$\mathscr{Y}(T) = \rho - \frac{1}{T}\log\left[e^{-\boldsymbol{\gamma}'\boldsymbol{u}_+/N}\left(e^{u_1}+\cdots+e^{u_N}\right)^{\gamma}\int \mathscr{F}_{\gamma}^{N}(\boldsymbol{v})e^{i\boldsymbol{u}'\boldsymbol{v}}e^{\boldsymbol{c}(-\boldsymbol{\gamma}/N+i\boldsymbol{U}'\boldsymbol{v})T}\,d\boldsymbol{v}\right]. \tag{2.35}$$

The riskless rate is

$$r = e^{-\boldsymbol{\gamma}'\boldsymbol{u}_+/N}\left(e^{u_1}+\cdots+e^{u_N}\right)^{\gamma}\int \mathscr{F}_{\gamma}^{N}(\boldsymbol{v})e^{i\boldsymbol{u}'\boldsymbol{v}}\left[\rho - \boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{U}'\boldsymbol{v})\right]d\boldsymbol{v}\,. \tag{2.36}$$

These formulas can be expressed in terms of the dividend shares $\{s_i\}$ by making the substitution $u_k = \log(s_k/s_1)$.

Proof. See Appendix B.5. □

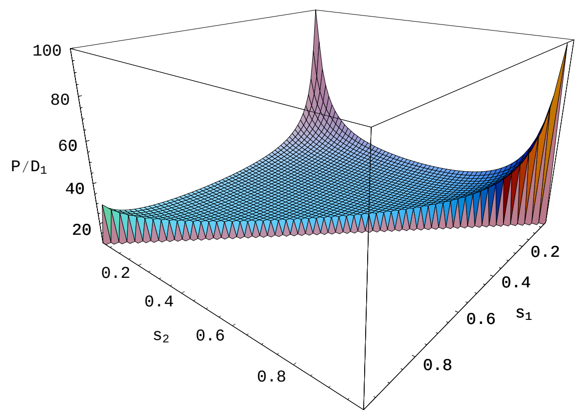


Figure 2.21: The price-dividend ratio of asset 1 against s_1 and s_2, in an example in which the three assets have identical and independent fundamentals.

The integral formula (2.33), for example, is a generalization of (2.9). Figure 2.21 illustrates how the price-dividend ratio of asset 1 depends on the two state variables s_1 and s_2 in a three-asset example.

2.7.1 The robustness of contagion and flight-to-quality

Above, I presented a two-asset calibration in which a small asset experiences a negative shock ("contagion") if a large asset has bad dividend news. On the other hand, a sufficiently large asset experiences a *positive* shock when a sufficiently small asset has bad dividend news; this was labelled "flight-to-quality". This flight-to-quality effect was dependent on a decrease in the riskless rate outweighing the effect of an increase in the risk premium on the large asset.

How robust is this effect? Intuition suggests that when more assets are introduced, the riskless rate effect will be muted, while the risk premium effect will continue to matter for individual assets. This section evaluates that intuition.

In the two-asset case, an asset is subject to contagion when its price-dividend ratio is decreasing in its dividend share, and to flight-to-quality when its price-dividend ratio is increasing in its dividend share. In the calibration of Section 2.5.1, the share, s^*, at which the transition takes place occurs at the minimum point of the price-dividend curve shown in Figure 2.2b: that is, at $s^* \approx 0.61$.

In the N-asset case, whether an asset experiences contagion or flight-to-quality depends on the $(N-1)$-dimensional state vector and also on which other asset is assumed to experience a shock. Suppose, for example, that there are $N-1$ equally sized small assets and an Nth large asset, and that all assets have independent and identically distributed dividend processes, following geometric Brownian motions with $\mu = 0.02$ and $\sigma = 0.1$. As in the two-asset case, we can calculate the critical dividend share, s^*, above which the Nth asset exhibits flight-to-quality, and below which the Nth asset exhibits contagion, following a negative dividend shock to any one of the $N-1$ small assets.

N	Critical share, s^*	Relative share
2	0.61	1.56
3	0.48	1.83
4	0.41	2.11
5	0.37	2.38
6	0.35	2.66

Table 2.2: Above the critical share the large asset experiences flight-to-quality; below, it experiences contagion. Relative share is the ratio of the large asset's dividend to the dividend of one of the small assets, at this critical share.

Table 2.2 demonstrates that s^* is decreasing in N. An alternative measure of the large asset's relative size at this critical point is the ratio of the its dividend to the dividend of any one of the $N-1$ small assets. This quantity is reported as "Relative share" in Table 2.2. The relative share is increasing in N: when $N = 6$, an asset that has dividends two and a half times as large as any other asset will still experience contagion rather than flight-to-quality, whereas such an asset experiences flight-to-quality if $N \leq 5$. On the basis of this evidence, it appears that when there are several assets of broadly similar size, contagion, not flight-to-quality, is the norm.

2.8 Conclusion

It seems worthwhile to summarize the solution method for readers who are not inclined to look through the appendices. By means of a change of measure followed by a Fourier transform, the Lucas asset-pricing equation (2.1) is converted into the integral formulas (2.6) and (2.33) which can be evaluated numerically.

In the two-asset case, the integral formula (2.6) can be simplified further if dividends follow geometric Brownian motions. Techniques from complex analysis enable the integral to be expressed as an infinite sum of residues that can be evaluated in closed form, leading to the expression (2.18). Closed forms are also available in the limit as an asset becomes negligibly small, because only one residue is then relevant: the tractable expressions (2.27) and (2.30) are valid for general dividend processes.

Complicated, interesting, and empirically relevant phenomena emerge from simple assumptions. In various regions of the parameter space, the model exhibits momentum, mean-reversion, contagion, flight to quality, the value-growth effect, and excess volatility. Notably, the model demonstrates that comovement is a robust feature of the neoclassical model. As an extreme example, even a negligibly small asset whose fundamentals are independent of the rest of the economy may comove endogenously and hence earn a risk premium.

3. ASSET RETURNS IN THE LONG RUN

Working in a rather general setting, this chapter explores the long run implications of the fundamental asset-pricing equation,

$$\mathbb{E}_t M_{t+1} R_{t+1} = 1\,.$$

I have introduced a stochastic discount factor, M_{t+1}, that prices payoffs at time $t+1$ from the perspective of time t. R_{t+1} is the gross return, from time t to $t+1$, on some arbitrary asset.[1]

The objects of interest are the martingale $X_t \equiv M_1 R_1 \cdots M_t R_t$, and the random variable $X_\infty \equiv \lim_{t\to\infty} X_t$. The asset-pricing equation states that $\mathbb{E}X_t = 1$ for all finite t, so it is natural to expect that $\mathbb{E}X_\infty = 1$, too. In Section 3.1, I show that this may or may not be true; typically, in fact, it is not, and when it is not, $X_\infty = 0$.[2] This dichotomy, together with a diagnostic that determines which of the two cases applies, is the main result of the chapter.

The result applies to any valid stochastic discount factor, but to understand it better I consider, in Section 3.2, a particular stochastic discount factor: the reciprocal of the return on the growth-optimal portfolio. In this special case, the main result provides conditions under which the following statement can be made precise: in the long run, the growth-optimal portfolio outperforms, and continues

[1] I thank Brandon Bates, John Campbell, John Cochrane, and Lars Hansen for their comments.

[2] To be precise, $X_\infty = 0$ with probability one, or $X_\infty = 0$ almost surely. Throughout the chapter, I drop such qualifications in the interests of readability.

to outperform, any other asset by an arbitrarily large factor. The prior result can then be interpreted as saying that $X_\infty = 0$ for any asset that is not growth-optimal.

Under this interpretation, related results have been obtained by a variety of authors. The assumption of this chapter is that random variables are independent, but not necessarily identically distributed, across periods. Latané (1959) and Samuelson (1971) make a stronger assumption—the world is i.i.d.—and prove a weaker result. Breiman (1960) does not make an independence assumption, but requires that gross returns are bounded away from 0 and from infinity, so his analysis does not apply to situations in which returns are, say, lognormal. Markowitz (1976) uses the Strong Law of Large Numbers to find the result presented here but, again, requires the i.i.d. assumption. I discuss the distinctions between the various different results at greater length below.

In Section 3.3, I return to the more general setting in which M_t is any valid stochastic discount factor. I interpret X_t as a measure of the realized value of an asset. How are we to square the fact that $\mathbb{E}X_t = 1$ with the fact that non-growth-optimal assets have $X_\infty = 0$? I show that such assets derive their value—their $\mathbb{E}X_t = 1$—from low-probability events in which the realized value of X_t is enormous. When such an event happens, either $M_1 \cdots M_t$ is large or $R_1 \cdots R_t$ is large, or both. The former possibility, which is driven by extreme *left*-tail events, can be thought of as representing the importance of rare disasters; this interpretation becomes particularly clear when considering the riskless strategy in which $R_1 \cdots R_t$ is deterministic. The latter possibility represents the importance of the extreme *right* tail of the distribution of returns on the asset in question, and is particularly clear in a risk-neutral world in which $M_1 \cdots M_t$ is deterministic.

Section 3.4 conducts simulations that illustrate the preceding results in the context of a simple economy featuring two assets—one riskless, the other i.i.d. lognormal. Section 3.5 concludes.

3.1 The main result

Today is time 0. I make three assumptions:

(i) There is no arbitrage.

(ii) The asset of interest has limited liability.

(iii) All random variables are independent across periods.

Assumption (i) implies that for all $t \geq 1$ we can define M_t to be a stochastic discount factor which prices payoffs at time t from the perspective of time $t-1$, and then we have

$$M_t > 0 \quad \text{and } \mathbb{E}\left(M_1 R_1 \cdot M_2 R_2 \cdot \ldots \cdot M_t R_t\right) = 1\,. \tag{3.1}$$

Assumption (ii) implies that for any $t \geq 1$

$$R_t \geq 0\,, \tag{3.2}$$

where R_t is the gross realized return from time $t-1$ to time t on some arbitrary asset or investment strategy. M_t and R_t are random variables that only become known at time t.

To simplify notation, define the random variables X_t, $t = 1, 2, \ldots$, by

$$X_t \equiv M_1 R_1 \cdot M_2 R_2 \cdot \ldots \cdot M_t R_t\,.$$

X_t is a non-negative martingale, because

$$\begin{aligned}
\mathbb{E}_{t-1} X_t &= \mathbb{E}_{t-1}\left(M_1 R_1 \cdots M_t R_t\right) \\
&= M_1 R_1 \cdots M_{t-1} R_{t-1}\, \mathbb{E}_{t-1}\left(M_t R_t\right) \\
&= M_1 R_1 \cdots M_{t-1} R_{t-1} \\
&= X_{t-1}\,.
\end{aligned}$$

Next, define

$$X_\infty \equiv \lim_{t\to\infty} X_t = \lim_{t\to\infty} M_1 R_1 \cdot M_2 R_2 \cdot \ldots \cdot M_t R_t\,.$$

Since X_t is a non-negative martingale, the random variable X_∞ exists almost surely.

Define also the constants $\{a_t\}$, for $t = 1, 2, 3, \ldots$, by

$$a_t \equiv \mathbb{E}\sqrt{M_t R_t}\,.$$

By Jensen's inequality,

$$a_t = \mathbb{E}\sqrt{M_t R_t} \le \sqrt{\mathbb{E} M_t R_t} = \sqrt{1} = 1\,,$$

so $0 < a_t \le 1$. Whenever $M_t R_t$ is non-constant, the inequality is strict.

The main result of the chapter is the following dichotomy.

Proposition 3.1. *Under assumptions (i)–(iii), either*

$$\prod_{t=1}^{\infty} a_t > 0 \quad \textit{and} \quad \mathbb{E} X_\infty = 1 \tag{3.3}$$

or

$$\prod_{t=1}^{\infty} a_t = 0 \quad \textit{and} \quad X_\infty = 0\,. \tag{3.4}$$

Proof. In view of assumptions (i)–(iii), the result follows directly from Kakutani's (1948) product martingale theorem, as presented in Williams (1995, pp. 144–5). □

The quantity $\prod a_t$ is to be viewed as a diagnostic that tells us which of the two cases applies. I will refer to assets for which $\prod a_t > 0$ as Type 1 assets, and assets for which $\prod a_t = 0$ as Type 2 assets.

At first glance, one might have expected (3.3) always to apply since, for any finite t,

$$\mathbb{E}X_t = 1,$$

and so it is tempting—but wrong—to conclude that in general

$$\mathbb{E}X_\infty = \mathbb{E} \lim_{t\to\infty} X_t \stackrel{?}{=} \lim_{t\to\infty} \mathbb{E}X_t = \lim_{t\to\infty} 1 = 1.$$

The interchange of expectation and limit is the weak link in this chain. Proposition 3.1 shows that this interchange is legitimate only for Type 1 assets. I show below that the case $X_\infty = 0$ is relevant in many asset-pricing examples.

3.2 The growth-optimal portfolio

The return on the growth-optimal portfolio (G-OP) between $t-1$ and t, which I will write as R_t^*, solves the problem

$$\max_{R_t} \mathbb{E}_{t-1} \log R_t \quad \text{s.t.} \quad \mathbb{E}_{t-1} M_t R_t = 1\,.$$

If the market is complete, it follows (on "taking first-order conditions state-by-state") that the return on the G-OP satisfies $R_t^* = 1/M_t$. Since t was arbitrary, this holds for all t.

In the general case in which there are N assets with returns $R_t^{(i)}$, $i = 1, \dots, N$, the G-OP is obtained by picking $\alpha_i, i = 1, \dots, N$ to solve

$$\max_{\{\alpha_i\}} \mathbb{E} \log \sum \alpha_i R_t^{(i)} \quad \text{s.t.} \quad \sum \alpha_i = 1 \,.$$

The first-order conditions are that, for each i,

$$\mathbb{E} \, \frac{R_t^{(i)}}{\sum \alpha_j R_t^{(j)}} = \lambda \,.$$

Multiplying both sides of this equation by α_i and summing over i, we find $\lambda = 1$, so

$$\mathbb{E} \, \frac{R_t^{(i)}}{\sum \alpha_j R_t^{(j)}} = 1 \quad \text{for all } i \,,$$

which exhibits $1/\sum \alpha_j R_t^{(j)} = 1/R_t^*$ as a valid stochastic discount factor.

If markets are complete, the stochastic discount factor is unique and, necessarily, $M_t = 1/R_t^*$. In the incomplete market case, the main result applies to any valid stochastic discount factor, but we can get further insight into Proposition 3.1 by choosing to focus on the stochastic discount factor $1/R_t^*$. Doing so, we have

$$X_T = \frac{R_1 \cdot R_2 \cdot \ldots \cdot R_T}{R_1^* \cdot R_2^* \cdot \ldots \cdot R_T^*} \,,$$

so in this case X_T has a simple interpretation as the *relative performance* of the asset in question by comparison with the G-OP.

We can conclude the section by rephrasing Proposition 3.1 as follows.

Proposition 3.2. *Suppose that assumptions (i)–(iii) hold, and define* $a_t = \mathbb{E}\sqrt{R_t/R_t^*}$. *Either*

$$\prod_{t=1}^{\infty} a_t > 0 \quad \textit{and} \quad \mathbb{E}\left[\lim_{T \to \infty} \frac{R_1 \cdot R_2 \cdot \ldots \cdot R_T}{R_1^* \cdot R_2^* \cdot \ldots \cdot R_T^*}\right] = 1 \qquad (3.5)$$

or

$$\prod_{t=1}^{\infty} a_t = 0 \quad and \quad \lim_{T\to\infty} \frac{R_1 \cdot R_2 \cdot \ldots \cdot R_T}{R_1^* \cdot R_2^* \cdot \ldots \cdot R_T^*} = 0\,. \tag{3.6}$$

Proof. Follows from Proposition 3.1, after setting $M_t = 1/R_t^*$. □

3.2.1 Examples

Example 1: the G-OP. Suppose the asset in question *is* the G-OP. Then R_t/R_t^* is trivially equal to 1, so

$$a_t = 1 \quad \text{and hence} \quad \prod_{t=1}^{\infty} a_t = 1\,.$$

The growth-optimal portfolio is the archetypal Type 1 asset with $\mathbb{E}X_\infty = 1$.

Example 2: an i.i.d. world. Fix any asset other than the G-OP. Since the world is i.i.d., a_t equals a, some constant. Since the asset in question is not the G-OP and the world is nondeterministic, the mean-1 random variable $M_t R_t$ is nonconstant, so a Jensen's inequality argument delivers the *strict* inequality $a < 1$. It follows that

$$\prod_{t=1}^{\infty} a_t = \prod_{t=1}^{\infty} a = 0\,.$$

Any fixed asset which is not the growth-optimal portfolio is of Type 2: $X_\infty = 0$ and equation (3.6) holds.

Example 3: an i.i.d. risk-neutral world. In this case, the stochastic discount factor is deterministic, so the G-OP has deterministic returns and so must be invested in the riskless asset. As a result of the previous example, we can say that any strategy which invests in the same risky asset each period must eventually have returns which satisfy

$$\frac{R_1 \cdot R_2 \cdot \cdots \cdot R_T}{R_{f,1} \cdot R_{f,2} \cdot \cdots \cdot R_{f,T}} < \varepsilon \tag{3.7}$$

where $R_{f,t}$ indicates the riskless rate from time $t-1$ to time t. The realized return on any risky asset is ultimately negligible by comparison with the riskless return.

Example 4: eventually-growth-optimal strategies. Consider the strategy of investing in arbitrary fashion until some fixed finite time T' and then investing in the G-OP. Such a strategy is *eventually-growth-optimal*. Since $a_t = 1$ for t larger than T', we have

$$\prod_{t=1}^{\infty} a_t = \prod_{t=1}^{T'} a_t > 0$$

and so eventually-growth-optimal strategies are of Type 1, and satisfy $\mathbb{E}X_\infty = 1$.

Example 5: fixed strategies that invest in the G-OP infinitely often. A trading strategy which invests in the G-OP infinitely often may nonetheless be of Type 2. Suppose, for example, that the strategy invests in the G-OP during time periods 1, 3, 5, ..., and in some other i.i.d. asset during time periods 2, 4, 6, ...; write a for the value of a_t during these even periods and note that $a < 1$ by Jensen's inequality. We have

$$\prod_{t=1}^{\infty} a_t = \prod_{t=1}^{\infty} a_{2t} = \prod_{t=1}^{\infty} a = 0$$

and so $X_\infty = 0$.

Example 6: strategies that are never growth-optimal but which satisfy $\mathbb{E}X_\infty = 1$. If a trading strategy becomes increasingly similar to the G-OP over time, it may be possible to sustain the case $\mathbb{E}X_\infty = 1$. Suppose for example that we have

$$a_t = 1 - 1/t^2$$

for all t. It follows that $\sum_{t=1}^{\infty}(1-a_t) < \infty$; this condition implies that $\prod_{t=1}^{\infty} a_t > 0$.

3.2.2 *Relationship with previous results*

Various authors have obtained results similar to Proposition 3.2. Latané (1959) and Samuelson (1971) assume that the world is i.i.d., and rely on the Weak Law of Large Numbers and the Central Limit Theorem respectively. They show that

$$\mathbb{P}\left[\frac{R_1 \cdot R_2 \cdot \cdots \cdot R_T}{R_1^* \cdot R_2^* \cdot \cdots \cdot R_T^*} < 1\right] \longrightarrow 1 \text{ as } T \to \infty..$$

This conclusion is weaker than the conclusion presented above for three reasons. First, the result holds only *at* the time horizon T, and gives no guarantee about what happens thereafter. Second, the result shows only that the G-OP outperforms, rather than that it overwhelmingly outperforms. Third, the result holds with probability approaching one, rather than with probability equal to one.

Markowitz (1976) also assumes that the world is i.i.d., and shows that the Strong Law of Large Numbers delivers the conclusion of Proposition 3.2,[3]

$$\lim_{T\to\infty} \frac{R_1 \cdot R_2 \cdot \ldots \cdot R_T}{R_1^* \cdot R_2^* \cdot \ldots \cdot R_T^*} = 0\,.$$

Each result, weak or strong, can be derived by applying the appropriate Law of Large Numbers, Weak or Strong, to the random variables $\log M_t + \log R_t$, which are i.i.d. by assumption and which have mean $\mathbb{E} \log M_t + \log R_t \equiv \mu < 0$ by Jensen's inequality.

Breiman (1960) does not require that random variables are independent across time. But he does assume that returns are bounded away from zero and from infinity, thereby ruling out lognormality of returns or the possibility of bankruptcy,

---

[3] In fact, under the i.i.d. assumption of Markowitz's paper, a stronger result can be obtained, namely that

$$\lim_{T\to\infty} \left(\frac{R_1 \cdot R_2 \cdot \ldots \cdot R_T}{R_1^* \cdot R_2^* \cdot \ldots \cdot R_T^*}\right)^{1/T} < 1\,,$$

where R_t is the return on a fixed non-growth-optimal asset.

for example. Given this boundedness assumption, he shows that Proposition 3.2 holds if a condition equivalent to

$$\prod_{t=1}^{\infty} e^{\mathbb{E}_{t-1} \log(R_t/R_t^*)} = 0$$

holds.

3.3 Where's the value in a Type 2 asset?

Who would buy a Type 2 asset, if a dollar placed in the growth-optimal portfolio will outperform it, in the long run, with probability one? Why aren't such assets *cheaper*?

Fix, for the sake of argument, some particular Type 2 asset. How are we to square the fact that X_t tends to zero with the fact that $\mathbb{E}X_t = 1$ for all finite t? It seems intuitively clear that there must be some unlikely states of the world in which X_t is *very large*, and that the value of the Type 2 asset in question is driven by these unlikely states of the world.

The following Proposition makes this idea formal.

Proposition 3.3. *For Type 1 assets, we have*

$$\mathbb{E} \sup_{t \geq 1} X_t < \infty \,, \tag{3.8}$$

while for Type 2 assets, we have

$$\mathbb{E} \sup_{t \geq 1} X_t = \infty \quad \textit{and} \quad \sup_{t \geq 1} \mathbb{E} \left(X_t \log^+ X_t \right) = \infty \,, \tag{3.9}$$

where $\log^+(x) \equiv \max\{\log x, 0\}$.

Proof. Equation (3.8) is established in the course of the proof of Kakutani's product martingale theorem in Williams (1995, pp. 144–5). (It leads to the conclusion that for Type 1 assets, the family $\{X_t\}$ is uniformly integrable,[4] from which Proposition 3.1 follows.)

Both parts of (3.9) can be established by contradiction. If $\mathbb{E}\,\sup_t X_t < \infty$ then it would follow that the family of random variables $\{X_t\}$, being dominated by the integrable random variable $\sup_t X_t$, would be uniformly integrable, and hence that $\mathbb{E}X_\infty = \mathbb{E}X_1 = 1$. But this contradicts the conclusion of Proposition 3.1.

Similarly, if $\sup_t \mathbb{E}\left(X_t \log^+ X_t\right) < \infty$ it would follow, by Proposition IV-2-10 of Neveu (1975, p. 70), that $\sup_t X_t$ would be integrable, and hence, as in the previous paragraph, that $\{X_t\}$ would be a uniformly integrable family of random variables. Again, this contradicts Proposition 3.1, and the result follows. □

When contemplating (3.8) and (3.9), it is helpful to keep in mind the fact that

$$\sup_{t\geq 0} \mathbb{E}X_t = \sup_{t\geq 1} 1 = 1\,.$$

Although the expected value of X_t is equal to 1 for all t, and the expected value of the supremum of X_t is finite for Type 1 assets, the expected value of the supremum of X_t is infinite in the case of Type 2 assets: there are rare states of the world in which X_t becomes very large indeed.

In such states, we have

$$M_1R_1 \cdot M_2R_2 \cdots M_tR_t \quad \text{very large,}$$

and so we must have some combination of large $M_1 \cdots M_t$ and large $R_1 \cdots R_t$. The

[4] A family $\{X_t\}$ of random variables is *uniformly integrable* if

$$\sup_{t\geq 1} \mathbb{E}\left(|X_t|\,\mathbf{1}[|X_t| > a]\right) \longrightarrow 0 \quad \text{as } a \to \infty.$$

former possibility, large $M_1 \cdots M_t$, corresponds roughly to the realization of a disastrously bad state of the world. In a consumption-based model with time-separable utility, for example, $M_1 \cdots M_t$ is large when marginal utility at time t is high. The latter possibility, large $R_1 \cdots R_t$, corresponds to a particularly favorable return realization for the asset in question.

In some sense, therefore, the value in Type 2 assets derives either from aggregate disasters (large $M_1 \cdots M_t$) or asset-specific triumphs (large $R_1 \cdots R_t$). At a general level, we can say no more. Nonetheless, for the sake of intuition, it is interesting to consider simple special cases that focus attention on each of the two channels separately.

Suppose, first, that we are in a risk-neutral i.i.d. world, as in Example 3 above, and consider a risky Type 2 asset. Since $M_1 \cdots M_t = (1/R_f)^t$ is deterministic, the value of the asset is driven by very occasional asset-specific triumphs—explosions in $R_1 \cdots R_t$—that is, by extreme *right*-tail events.

Conversely, suppose that the world is i.i.d. but not risk-neutral, so that M_t is not constant, and that we are considering the riskless strategy that rolls cash over in the riskless asset. Now, $R_1 \cdots R_t = (R_f)^t$ is deterministic. Again, this is a Type 2 strategy, but now the value is derived from aggregate disasters—states of the world which occur with very low probability, but in which $M_1 \cdots M_t$ is far larger than its expected value. In other words, the value of this strategy, in the long run, is driven by the presence of extreme *left*-tail events. Weitzman (2004) emphasizes the importance of this effect.

3.4 An example

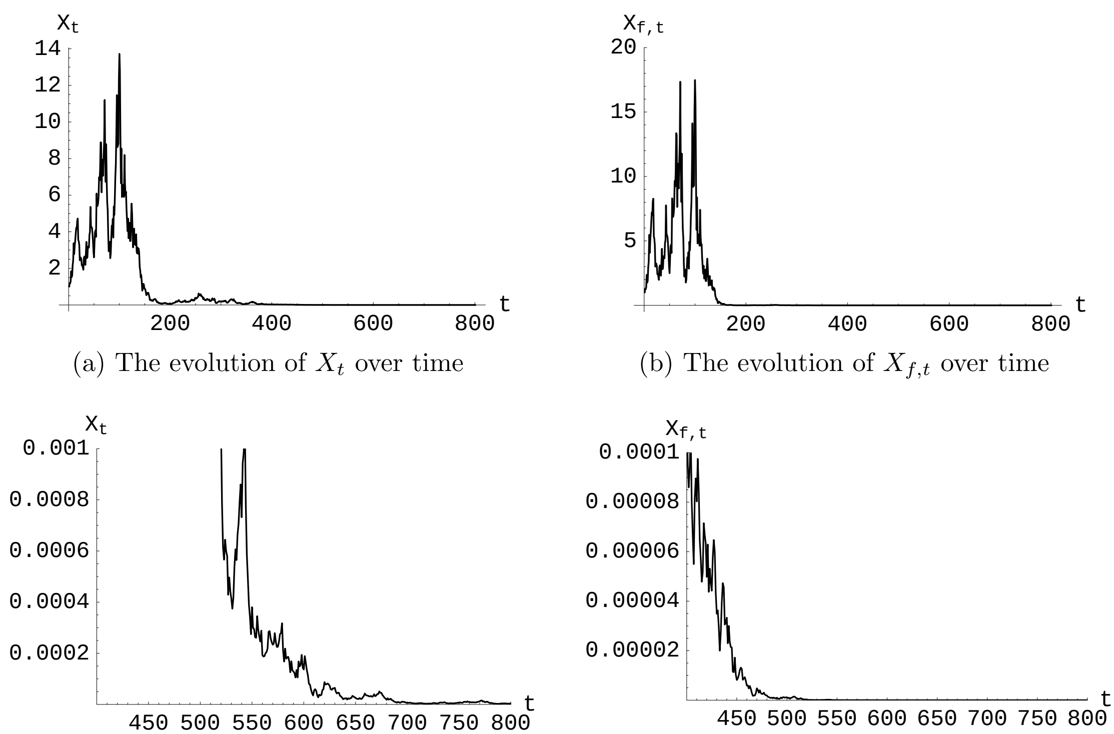
Consider an i.i.d. economy with two assets, a riskless asset which pays the certain return $R_{f,t} \equiv e^{r_f}$ and a risky asset which pays the lognormal return $R_t \equiv e^{\mu - \sigma^2/2 + \sigma Z_t}$, where Z_t is a standard Normal random variable. It is easy to check that the stochastic discount factor $M_t \equiv e^{-r_f - \lambda^2/2 - \lambda Z_t}$ prices the assets, where λ is the Sharpe ratio $(\mu - r_f)/\sigma$. (Notice that M_t so defined is not the reciprocal of the return on the growth-optimal portfolio, since the latter is not lognormally distributed.) Each asset is of Type 2, as is easily checked.

Writing $X_{f,t} \equiv M_1 R_{f,1} \cdots M_t R_{f,t}$ and $X_t \equiv M_1 R_1 \cdots M_t R_t$, we have

$$X_{f,t} = e^{-\lambda(Z_1 + \cdots + Z_t) - \lambda^2 t/2} \tag{3.10}$$

$$X_t = e^{(\sigma-\lambda)(Z_1 + \cdots + Z_t) - (\sigma-\lambda)^2 t/2} \tag{3.11}$$

Notice that in this example with just one kind of shock, realistic values of σ and λ imply that X_t is large when $(Z_1 + \cdots + Z_t)$ is small: in the long run, only disasters matter.



(a) The evolution of X_t over time

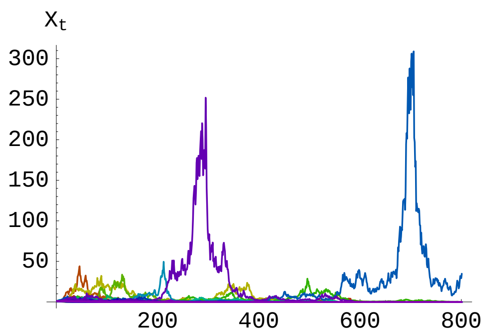
(b) The evolution of $X_{f,t}$ over time

(c) Zooming in on X_t

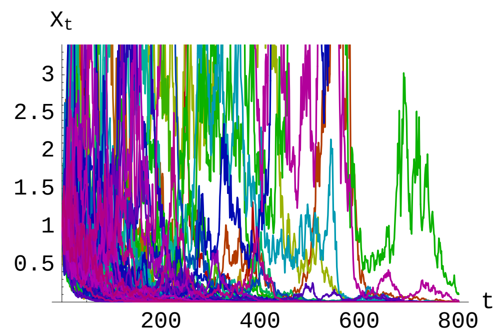
(d) Zooming in on $X_{f,t}$

Figure 3.1: Realizations of X_t and $X_{f,t}$ on one particular sample path.

Figure 3.1 plots a realization of X_t and $X_{f,t}$. Each time period represents one quarter. I have set $\sigma = 0.08$ and $\lambda = 0.25$, which corresponds to an annualized standard deviation of 16% and Sharpe ratio of 50% for the risky asset. Proposition 3.1 states that X_t and $X_{f,t}$ tend to zero as t tends to infinity; along this particular sample path, X_t and $X_{f,t}$ are indistinguishable from zero, even in the zoomed-in graphs, after about 700 quarters. Note also the occasional spikes, which Proposition 3.3 led us to expect.



(a) 100 sample paths of X_t



(b) Zooming in

Figure 3.2: Realizations of X_t on 100 sample paths.

Figure 3.2 shows realizations of X_t along 100 different sample paths. As before, each period represents one quarter. On two sample paths, X_t spikes above 250. These spikes are so large, compared with the values of X_t attained on the vast majority of sample paths, that only about six of the sample paths are visible on the first, unzoomed, diagram. Despite these spikes, after 800 quarters only one sample path remains above 0.5.

In the continuous-time limit, the analogue of X_t would follow a geometric Brownian martingale of the form $e^{\alpha W_t - \alpha^2 t/2}$, where α is some constant and W_t is a Brownian motion. In this special case, we can see directly that $e^{\alpha W_t - \alpha^2 t/2} \to 0$, because $\alpha W_t - \alpha^2 t/2 \to -\infty$ as $t \to \infty$; this follows, in turn, from the fact that $W_t/t \to 0$ as $t \to \infty$ (Karatzas and Shreve (1991, p. 104)).

3.5 Conclusion

Although expected time- and risk-adjusted cumulative returns on any asset equal one at all horizons, *realized* time- and risk-adjusted cumulative returns on Type 2 assets tend to zero with probability one.

This apparent paradox is resolved in Section 3.3, which demonstrates that the value of such an asset is driven by the possibility of two types of rare events: spectacular outperformance of the asset itself, and occasional aggregate disasters. Only the first is relevant for the valuation of risky assets in a risk-neutral economy; only the second is relevant for the valuation of riskless strategies in a risky, risk-averse world.

Just three assumptions underpin these results. Two of these—no arbitrage and limited liability—are uncontroversial. The third—independence across time of the relevant random variables—is less desirable. Ritter (1979) presents a generalization of Kakutani's theorem that relaxes the independence assumption, and it may be that the ideas in that paper can be used to improve Proposition 3.1; in the interests of simplicity, I have not pursued such an extension here.

APPENDIX

A. APPENDIX TO CONSUMPTION-BASED ASSET PRICING WITH HIGHER CUMULANTS

A.1 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

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

Here, and throughout Chapter 1, μ and σ denote the unconditional mean and standard deviation of log consumption growth.

A.1.1 Definition and standard properties

Definition A.1. *The* cumulants *of G are the coefficients κ_n in the power series expansion of the CGF $\boldsymbol{c}(\theta)$:*

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

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

Proposition A.1. *We have the following properties.*

1. $\mathbb{E}G = \kappa_1$; $varG = \kappa_2 \equiv \sigma^2$; $skewness(G) = \kappa_3/\sigma^3$; $excess\ kurtosis(G) = \kappa_4/\sigma^4$.

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* n*th 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 $\theta = 0$. □

Thus, the CGF is a convex[1] 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.

[1] As shown in the main text, Fact 1.1.

A.2 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

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

and hence that

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

and so, finally, that

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

Defining d/p as usual,

$$d/p = \rho\vartheta + \nu(1-\vartheta) - \boldsymbol{c}(\lambda-\gamma)\,. \tag{A.2}$$

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 (A.2), 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

$$a_0^{(1-\gamma)/\vartheta} = 1 - e^{-\rho} + e^{-\rho}\mathbb{E}\left[\left(\frac{C_1}{C_0}\right)^{1-\gamma} a_1^{1-\gamma}\right]^{1/\vartheta}, \tag{A.3}$$

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

I now conjecture that $a_i = a$, some constant, solves (A.3). If so,

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

from which it follows that

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

which confirms the conjecture that a was constant. Hence,

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

The cost of all uncertainty, ϕ, solves the equation

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

from which (1.48) 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 $\widetilde{\boldsymbol{c}}(\theta)$ from equation (1.49), we obtain the expression (1.51).

A.3 Derivation of results in continuous time

Definition A.2. *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* $\varepsilon > 0$*,* $\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \varepsilon) = 0$*.*

A.3.1 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 (1.27) 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,

$$\begin{aligned}
ER_\lambda &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathbb{E} \left(\left(\frac{C_{t+\Delta t}}{C_t} \right)^\lambda - 1 \right) + \rho - \boldsymbol{c}(\lambda - \gamma) \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \mathbb{E} \left(e^{\lambda G_{\Delta t}} - 1 \right) + \rho - \boldsymbol{c}(\lambda - \gamma) \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \cdot \left(e^{\boldsymbol{c}(\lambda)\Delta t} - 1 \right) + \rho - \boldsymbol{c}(\lambda - \gamma) \\
&= \boldsymbol{c}(\lambda) + \rho - \boldsymbol{c}(\lambda - \gamma) \qquad \text{(A.4)}
\end{aligned}$$

A.4 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.

A.4.1 The univariate case

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

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

A.4.2 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.

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

B. APPENDIX TO THE LUCAS ORCHARD

B.1 General solution in the two-asset case

B.1.1 Preliminary mathematical results

An expectation

This section contains a calculation which is used below. It may be helpful to glance ahead to equation (B.13) for motivation. The goal is to evaluate

$$E \equiv \mathbb{E}\left(\frac{e^{\alpha_1 \widetilde{y}_{1t} + \alpha_2 \widetilde{y}_{2t}}}{\left[e^{y_{10}+\widetilde{y}_{1t}} + e^{y_{20}+\widetilde{y}_{2t}}\right]^{\gamma}}\right)$$

for general $\alpha_1, \alpha_2, \gamma > 0$. First, I rewrite the expectation, noting that

$$\begin{aligned} \mathbb{E}\left(\frac{e^{\alpha_1 \widetilde{y}_{1t} + \alpha_2 \widetilde{y}_{2t}}}{\left[e^{y_{10}+\widetilde{y}_{1t}} + e^{y_{20}+\widetilde{y}_{2t}}\right]^{\gamma}}\right) &= e^{-\gamma/2(y_{10}+y_{20})} \times \\ &\quad \mathbb{E}\left(\frac{e^{(\alpha_1-\gamma/2)\widetilde{y}_{1t} + (\alpha_2-\gamma/2)\widetilde{y}_{2t}}}{\left[2\cosh((y_{20}-y_{10}+\widetilde{y}_{2t}-\widetilde{y}_{1t})/2)\right]^{\gamma}}\right) \end{aligned} \quad \text{(B.1)}$$

To take care of the exponential in the numerator inside the expectation, I transform the probability law, defining

$$\widetilde{\mathbb{E}}\left[Y\right] \equiv e^{-t\boldsymbol{c}(\alpha_1-\gamma/2, \alpha_2-\gamma/2)} \cdot \mathbb{E}\left[e^{(\alpha_1-\gamma/2)\widetilde{y}_{1t} + (\alpha_2-\gamma/2)\widetilde{y}_{2t}} \cdot Y\right] \ . \quad \text{(B.2)}$$

This is an Esscher transform of the original law, and it has the property that

$$\widetilde{\boldsymbol{c}}(v_1, v_2) \equiv \log \widetilde{\mathbb{E}}\left[e^{v_1 \widetilde{y}_{11} + v_2 \widetilde{y}_{21}}\right] = \boldsymbol{c}(\alpha_1 - \gamma/2 + v_1, \alpha_2 - \gamma/2 + v_2) - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2)\,. \tag{B.3}$$

In terms of this transformed law, the right hand side of (B.1) equals

$$e^{-\gamma(y_{10}+y_{20})/2 + \boldsymbol{c}(\alpha_1-\gamma/2,\alpha_2-\gamma/2)t}\widetilde{\mathbb{E}}\left(\frac{1}{[2\cosh((y_{20}-y_{10}+\widetilde{y}_{2t}-\widetilde{y}_{1t})/2)]^{\gamma}}\right) \tag{B.4}$$

To make further progress, we can now attack the expectation in (B.4) by exploiting the fact that $1/\left[2\cosh(u/2)\right]^{\gamma}$ has a Fourier transform which can be found in closed form for integer $\gamma > 0$. Define the Fourier transform $\mathscr{F}_\gamma(v)$ by

$$\frac{1}{[2\cosh(u/2)]^{\gamma}} = \int_{-\infty}^{\infty} e^{iuv}\mathscr{F}_\gamma(v)\,dv \tag{B.5}$$

We have, then,

$$\begin{aligned} E &= e^{-\gamma(y_{10}+y_{20})/2 + \boldsymbol{c}(\alpha_1-\gamma/2,\alpha_2-\gamma/2)t}\ \widetilde{\mathbb{E}}\left(\int_{-\infty}^{\infty} e^{iv(y_{20}-y_{10}+\widetilde{y}_{2t}-\widetilde{y}_{1t})}\mathscr{F}_\gamma(v)\,dv\right) \\ &= e^{-\gamma(y_{10}+y_{20})/2 + \boldsymbol{c}(\alpha_1-\gamma/2,\alpha_2-\gamma/2)t}\left(\int_{-\infty}^{\infty} e^{\widetilde{\boldsymbol{c}}(-iv,iv)t}\cdot e^{iv(y_{20}-y_{10})}\mathscr{F}_\gamma(v)\,dv\right) \\ &= e^{-\gamma(y_{10}+y_{20})/2}\int_{-\infty}^{\infty} e^{\boldsymbol{c}(\alpha_1-\gamma/2-iv,\alpha_2-\gamma/2+iv)t}\cdot e^{iv(y_{20}-y_{10})}\mathscr{F}_\gamma(v)\,dv\,. \end{aligned} \tag{B.6}$$

The Fourier transform $\mathscr{F}_\gamma(v)$

By the Fourier inversion theorem, definition (B.5) implies that

$$\begin{aligned} \mathscr{F}_\gamma(v) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-iuv}}{(2\cosh(u/2))^{\gamma}}\,du \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-iuv}}{(e^{u/2}+e^{-u/2})^{\gamma}}\,du\,. \end{aligned} \tag{B.7}$$

Make the change of variable

$$u = \log \frac{t}{1-t}\,. \tag{B.8}$$

It follows that

$$du = \frac{dt}{t(1-t)}$$

so on making this substitution in (B.7), we have

$$\begin{aligned} \mathscr{F}_\gamma(v) &= \frac{1}{2\pi}\int_0^1 \frac{\left(\frac{t}{1-t}\right)^{-iv}}{\left(\sqrt{\frac{t}{1-t}}+\sqrt{\frac{1-t}{t}}\right)^{\gamma}} \frac{dt}{t(1-t)} \\ &= \frac{1}{2\pi}\int_0^1 t^{\gamma/2-iv}(1-t)^{\gamma/2+iv}\frac{dt}{t(1-t)}\,. \end{aligned} \tag{B.9}$$

This is a Dirichlet surface integral. As shown in Andrews, Askey and Roy (1999, p. 34), it can be evaluated in terms of Γ-functions, giving

$$\mathscr{F}_\gamma(v) = \frac{1}{2\pi}\frac{\Gamma(\gamma/2-iv)\Gamma(\gamma/2+iv)}{\Gamma(\gamma)}\,. \tag{B.10}$$

For future reference, it is useful to note an equivalent representation of $\mathscr{F}_\gamma(v)$. Contour integration reveals that $\mathscr{F}_1(v) = \frac{1}{2}\mathrm{sech}\pi v$ and $\mathscr{F}_2(v) = \frac{1}{2}v\,\mathrm{cosech}\pi v$. From these two facts, expression (B.10), and the fact that $\Gamma(x) = (x-1)\Gamma(x-1)$, it follows that for positive integer γ, we have

$$\mathscr{F}_\gamma(v) = \begin{cases} \dfrac{v\,\mathrm{cosech}(\pi v)}{2(\gamma-1)!}\cdot \displaystyle\prod_{n=1}^{\gamma/2-1}\left(v^2+n^2\right) & \text{for even } \gamma\,, \\ \dfrac{\mathrm{sech}(\pi v)}{2(\gamma-1)!}\cdot \displaystyle\prod_{n=1}^{(\gamma-1)/2}\left(v^2+(n-1/2)^2\right) & \text{for odd } \gamma\,. \end{cases} \tag{B.11}$$

An Itô calculation

Given a jump-diffusion $\boldsymbol{y}$, with

$$d\boldsymbol{y} = \boldsymbol{\mu} dt + \boldsymbol{A} d\boldsymbol{Z} + \boldsymbol{J} dN \ ,$$

this section seeks a simple formula for

$$\mathbb{E}d(e^{\boldsymbol{w}'\boldsymbol{y}})$$

where $\boldsymbol{w}$ is a constant vector.

First, define $x \equiv \boldsymbol{w}'\boldsymbol{y}$; then

$$dx = \boldsymbol{w}'\boldsymbol{\mu} dt + \boldsymbol{w}'\boldsymbol{A} d\boldsymbol{Z} + \boldsymbol{w}'\boldsymbol{J} dN$$

We seek $\mathbb{E}d(e^x)$. By Itô 's formula for jump-diffusions, we have

$$d(e^x) = e^x \left[\left(\boldsymbol{w}'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} \right) dt + \boldsymbol{w}'\boldsymbol{A} d\boldsymbol{Z} + \left(e^{\boldsymbol{w}'\boldsymbol{J}} - 1 \right) dN \right]$$

where $\boldsymbol{\Sigma} \equiv \boldsymbol{A}\boldsymbol{A}'$; and so, after taking expectations,

$$\begin{aligned} \mathbb{E}d(e^{\boldsymbol{w}'\boldsymbol{y}}) &= e^{\boldsymbol{w}'\boldsymbol{y}} \cdot \left[\boldsymbol{w}'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} + \omega \left(\mathbb{E}e^{\boldsymbol{w}'\boldsymbol{J}} - 1 \right) \right] dt \\ &= e^{\boldsymbol{w}'\boldsymbol{y}} \cdot \boldsymbol{c}(\boldsymbol{w}) dt \, . \end{aligned} \tag{B.12}$$

In the case in which $\boldsymbol{y}$ is a general Lévy process, (B.12) holds by Proposition 8.20 of Cont and Tankov (2004).

B.1.2 Prices

Proof of Proposition 2.1

The price of the $\boldsymbol{\alpha}$-asset is

$$
\begin{aligned}
P_{\boldsymbol{\alpha}} &= \mathbb{E} \int_0^\infty e^{-\rho t} \left(\frac{C_t}{C_0} \right)^{-\gamma} D_{1t}^{\alpha_1} D_{2t}^{\alpha_2} \, dt \\
&= (C_0)^\gamma \int_0^\infty e^{-\rho t} \, \mathbb{E} \left(\frac{e^{\alpha_1 (y_{10} + \widetilde{y}_{1t}) + \alpha_2 (y_{20} + \widetilde{y}_{2t})}}{\left[e^{y_{10} + \widetilde{y}_{1t}} + e^{y_{20} + \widetilde{y}_{2t}} \right]^\gamma} \right) dt
\end{aligned}
$$

It follows that

$$
\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}} = (e^{y_{10}} + e^{y_{20}})^\gamma \int_{t=0}^\infty e^{-\rho t} \, \mathbb{E} \left(\frac{e^{\alpha_1 \widetilde{y}_{1t} + \alpha_2 \widetilde{y}_{2t}}}{\left[e^{y_{10} + \widetilde{y}_{1t}} + e^{y_{20} + \widetilde{y}_{2t}} \right]^\gamma} \right) dt \tag{B.13}
$$

The expectation inside the integral was calculated above in Appendix B.1.1. Substituting (B.6) into (B.13), interchanging the order of integration,[1] and writing u for $y_{20} - y_{10}$, we get

$$
\begin{aligned}
\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}} &= [2\cosh(u/2)]^\gamma \int_{v=-\infty}^\infty \int_{t=0}^\infty e^{-\rho t} e^{\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)t} \cdot e^{iuv} \mathscr{F}_\gamma(v) \, dt \, dv \\
&\overset{(a)}{=} [2\cosh(u/2)]^\gamma \int_{v=-\infty}^\infty \frac{e^{iuv} \mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)} \, dv
\end{aligned} \tag{B.14}
$$

For equality (a) to hold, I have assumed that

$$
\rho - \mathrm{Re}[\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)] > 0 \quad \text{for all } v \in \mathbb{R}.
$$

I show in Appendix B.2 that this follows from the apparently weaker assumption that the inequality holds at $v = 0$:

$$
\rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) > 0 \tag{B.15}
$$

[1] Since the integrand is absolutely integrable, this is a legitimate application of Fubini's theorem.

In particular, for the problem under consideration to be well-defined, we must impose a requirement that expected utility is finite. Finiteness of expected utility is guaranteed by the finiteness of the prices of the two assets. Therefore I refer to the two inequalities generated by substituting $(\alpha_1, \alpha_2) = (1, 0)$ and $(0, 1)$ into (B.15) as the *finiteness condition*. It is assumed throughout that this condition holds. (See equation (2.8) and Table 2.1.)

In terms of the state variable s, the price-dividend ratio is therefore

$$\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}} = \frac{1}{\sqrt{s^\gamma (1-s)^\gamma}} \cdot \int_{-\infty}^{\infty} \frac{\left(\frac{1-s}{s}\right)^{iv} \mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)} \, dv \tag{B.16}$$

where I have defined $s \equiv D_{10}/(D_{10} + D_{20})$.

Proof of Proposition 2.2

Since $u = \log[(1-s)/s]$, we have

$$\frac{1-s}{s} = e^u$$

and

$$\frac{1}{\sqrt{s^\gamma (1-s)^\gamma}} = [2\cosh(u/2)]^\gamma \ .$$

Furthermore, $\mathscr{F}_\gamma(v)$ was defined by

$$\frac{1}{[2\cosh(u/2)]^\gamma} = \int_{-\infty}^{\infty} e^{iuv} \mathscr{F}_\gamma(v) \, dv \, .$$

Substituting these observations into the pricing formula (2.6), we find the expressions of Proposition 2.2.

B.1.3 Returns

Expected returns contain a dividend yield component and a capital gain component:

$$R_{\boldsymbol{\alpha}} dt = \frac{D_{\boldsymbol{\alpha}}}{P_{\boldsymbol{\alpha}}} dt + \frac{\mathbb{E} dP_{\boldsymbol{\alpha}}}{P_{\boldsymbol{\alpha}}}$$

The first term is supplied by the pricing formula derived in the previous section. This section therefore focusses on calculating $\mathbb{E} dP_{\boldsymbol{\alpha}}/P_{\boldsymbol{\alpha}}$ in the case in which γ is an integer.

We have

$$P_{\boldsymbol{\alpha}} = (D_{10} + D_{20})^{\gamma}\, e^{(\alpha_1 - \gamma/2) y_{10} + (\alpha_2 - \gamma/2) y_{20}} \int_{-\infty}^{\infty} \frac{e^{iv(y_{20} - y_{10})} \mathscr{F}_{\gamma}(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)}\, dv \tag{B.17}$$

For convenience, I write

$$h(v) \equiv \frac{\mathscr{F}_{\gamma}(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)} \quad \text{and} \quad \binom{n}{m} \equiv \frac{n!}{m!(n-m)!}$$

throughout this section.

Introducing this notation,

$$\begin{aligned} P_{\boldsymbol{\alpha}} &= \int_{-\infty}^{\infty} h(v) \cdot (e^{y_{10}} + e^{y_{20}})^{\gamma}\, e^{(\alpha_1 - \gamma/2 - iv) y_{10} + (\alpha_2 - \gamma/2 + iv) y_{20}}\, dv \\ &= \int_{-\infty}^{\infty} h(v) \cdot \sum_{m=0}^{\gamma} \left[\binom{\gamma}{m} e^{m y_{10}} \cdot e^{(\gamma - m) y_{20}} \right] e^{(\alpha_1 - \gamma/2 - iv) y_{10} + (\alpha_2 - \gamma/2 + iv) y_{20}}\, dv \\ &= \sum_{m=0}^{\gamma} \binom{\gamma}{m} \int_{-\infty}^{\infty} h(v) \cdot e^{(\alpha_1 - \gamma/2 + m - iv) y_{10} + (\alpha_2 + \gamma/2 - m + iv) y_{20}}\, dv \\ &\equiv \sum_{m=0}^{\gamma} \binom{\gamma}{m} \int_{-\infty}^{\infty} h(v) \cdot e^{\boldsymbol{w}_m(v) \cdot \boldsymbol{y}}\, dv\,, \end{aligned} \tag{B.18}$$

where

$$\boldsymbol{w}_m(v) \equiv (\alpha_1 - \gamma/2 + m - iv, \alpha_2 + \gamma/2 - m + iv)'$$

The calculation of Appendix B.1.1, above, can now be used in (B.18) to show that

$$\mathbb{E}(dP_{\boldsymbol{\alpha}}) = \left\{ \sum_{m=0}^{\gamma} \binom{\gamma}{m} \int_{-\infty}^{\infty} h(v) \cdot e^{\boldsymbol{w}_m(v) \cdot \boldsymbol{y}} \boldsymbol{c} \left[\boldsymbol{w}_m(v) \right] \, dv \right\} \cdot dt \tag{B.19}$$

Dividing (B.19) by (B.18) and rearranging, the expected capital gain is given by the formula

$$\frac{\mathbb{E} dP_{\boldsymbol{\alpha}}}{P_{\boldsymbol{\alpha}}} = \frac{\displaystyle\sum_{m=0}^{\gamma} \binom{\gamma}{m} e^{-mu} \int_{-\infty}^{\infty} h(v) e^{iuv} \cdot \boldsymbol{c}(\boldsymbol{w}_m(v)) \, dv}{\displaystyle\sum_{m=0}^{\gamma} \binom{\gamma}{m} e^{-mu} \int_{-\infty}^{\infty} h(v) e^{iuv} \, dv} \cdot dt \tag{B.20}$$

B.1.4 Real interest rates

From the Euler equation, we have

$$\begin{aligned} B_T &= \mathbb{E} \left[e^{-\rho T} \left(\frac{C_T}{C_0} \right)^{-\gamma} \right] \\ &= e^{-\rho T} C_0^{\gamma} \mathbb{E} \left[\frac{1}{(D_{1T} + D_{2T})^{\gamma}} \right] \end{aligned}$$

Using the result of Appendix B.1.1, we find that

$$\begin{aligned} B_T &= e^{-\rho T} \left(e^{y_{10}} + e^{y_{20}} \right)^{\gamma} e^{-\gamma (y_{10} + y_{20})/2} \int_{-\infty}^{\infty} e^{iv(y_{20} - y_{10})} \mathscr{F}_{\gamma}(v) e^{\mathbf{c}(-\gamma/2 - iv, -\gamma/2 + iv) T} \, dv \\ &= e^{-\rho T} \left[2 \cosh(u/2) \right]^{\gamma} \int_{-\infty}^{\infty} \mathscr{F}_{\gamma}(v) e^{iuv} \cdot e^{\mathbf{c}(-\gamma/2 - iv, -\gamma/2 + iv) T} \, dv \, , \end{aligned}$$

as claimed. The yield, $\mathscr{Y}(T)$, follows directly from this expression:

$$\mathscr{Y}(T) = \rho - \frac{1}{T} \log \left\{ \left[2 \cosh(u/2) \right]^{\gamma} \int_{-\infty}^{\infty} \mathscr{F}_{\gamma}(v) e^{iuv} \cdot e^{\mathbf{c}(-\gamma/2 - iv, -\gamma/2 + iv) T} \, dv \right\} . \tag{B.21}$$

The riskless rate is found by taking the limit as $T \downarrow 0$ in (B.21). To calculate this limit, first use the fact that

$$[2\cosh(u/2)]^{\gamma}\int_{-\infty}^{\infty}\mathscr{F}_{\gamma}(v)e^{iuv}\,dv = 1$$

to rewrite equation (B.21) as

$$\mathscr{Y}(T) = \rho - \frac{1}{T}\log\left\{1 + [2\cosh(u/2)]^{\gamma}\int_{-\infty}^{\infty}\mathscr{F}_{\gamma}(v)e^{iuv}\left[e^{\boldsymbol{c}(-\gamma/2-iv,-\gamma/2+iv)T} - 1\right]\,dv\right\}.$$

It follows, after applying L'Hôpital's rule, that

$$\begin{aligned} r &= \rho - [2\cosh(u/2)]^{\gamma}\int_{-\infty}^{\infty}\mathscr{F}_{\gamma}(v)e^{iuv}\boldsymbol{c}(-\gamma/2 - iv, -\gamma/2 + iv)\,dv \\ &= [2\cosh(u/2)]^{\gamma}\int_{-\infty}^{\infty}\mathscr{F}_{\gamma}(v)e^{iuv}\cdot[\rho - \boldsymbol{c}(-\gamma/2 - iv, -\gamma/2 + iv)]\,dv \end{aligned}$$

as required.

B.2 The ridge property

This section expands on two closely related issues. First, as mentioned in Appendix B.1.2, the required assumption that

$$\rho - \mathrm{Re}[\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)] > 0 \quad \text{for all } v \in \mathbb{R}$$

follows from the apparently weaker assumption that the inequality holds at $v = 0$:

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) > 0\,.$$

Second, when considering the small-asset asymptotic (see Section 2.6 and Appendix B.4), it is of interest to find the zero of

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)$$

in the upper half-plane which is closest to the real axis (the *minimal* zero, in the terminology of Appendix B.4).

In either case, we are led to explore the properties of $\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)$, considered as a function of v. Recalling the change of measure of Appendix B.1.1, we can exploit the fact that

$$\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv) = \widetilde{\boldsymbol{c}}(-iv, iv) + \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2)$$

where $\widetilde{\boldsymbol{c}}(v_1, v_2)$ is the cumulant-generating function under the changed measure. Next, note that

$$\widetilde{\boldsymbol{c}}(-iv, iv) = \log \widetilde{\mathbb{E}}\, e^{iv(\widetilde{y}_{21} - \widetilde{y}_{11})} \equiv \log \psi(v)$$

which defines $\psi(v)$ as the characteristic function of the random variable $\widetilde{y}_{21} - \widetilde{y}_{11}$.

The characteristic function ψ has the *ridge property*. That is, for real v and w, we have

$$|\psi(v + iw)| \leq |\psi(iw)| \ .$$

This follows because (writing X for $\widetilde{y}_{21} - \widetilde{y}_{11}$)

$$|\psi(v + iw)| = \left| \widetilde{\mathbb{E}}\, e^{iX(v+iw)} \right| \leq \widetilde{\mathbb{E}} \left| e^{iX(v+iw)} \right| = \widetilde{\mathbb{E}}\, e^{-wX} = \psi(iw) = |\psi(iw)| \ .$$

Figure B.1 illustrates the ridge property using the calibration of section 2.5.2.

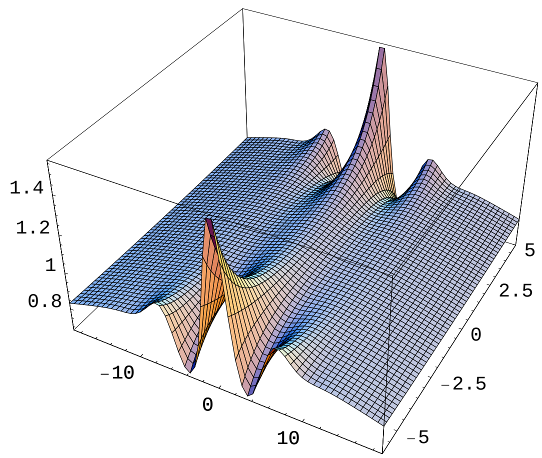


Figure B.1: The ridge property. The figure plots $|\psi(v)|$ (on the z-axis) over a portion of the complex plane around the origin. A ridge runs up the imaginary axis.

Proposition B.1. *The assumption that*

$$\rho - Re[\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)] > 0 \quad \textit{for all } v \in \mathbb{R}.$$

follows from the apparently weaker assumption that the inequality holds at $v = 0$*:*

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) > 0\,. \tag{B.22}$$

Proof. Suppose the apparently weaker inequality holds. In terms of the characteristic function ψ, we have

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv) = \rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) - \log\psi(v)\,; \tag{B.23}$$

note that the middle term on the right is real. So, for $v \in \mathbb{R}$, we have

$$\begin{aligned}\rho - \mathrm{Re}[\boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)] &= \rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) - \mathrm{Re}\log\psi(v) \\ &= \rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) - \log|\psi(v)| \\ &\geq \rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) - \log|\psi(0)| \\ &= \rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) \\ &> 0 \quad \text{by assumption,}\end{aligned}$$

which establishes the claim. The first inequality in this chain follows by the ridge property. □

Under assumption (B.22) this proposition implies, for example, that there are no zeros of $\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)$ *on* the real axis. The following proposition documents an important property of the closest zero above the real axis.

Proposition B.2. *Consider*

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv) \tag{B.24}$$

as a function of $v \in \mathbb{C}$*, and suppose that the condition*

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) > 0 \tag{B.25}$$

holds. Then the zero of (B.24) *in the upper half-plane which is closest to the real axis lies on the imaginary axis.*

Proof. Using equation (B.23) above, any zero, z, satisfies

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2, \alpha_2 - \gamma/2) = \log\psi(z)\,.$$

Writing the left-hand side as $\widehat{\rho} \in \mathbb{R}$ for convenience, any zero z must satisfy $\psi(z) = \exp \widehat{\rho}$. The fact that $\widehat{\rho} > 0$ follows from (B.25).

Let z^* be the zero in the upper half-plane with smallest imaginary part, and suppose (for a contradiction) that Re $z^* \neq 0$. Let $\widetilde{z} = (\text{Im } z^*)i$ be the projection of z^* onto the imaginary axis. By the ridge property, we have $\psi(\widetilde{z}) > |\psi(z^*)| = \exp \widehat{\rho}$. So, $\psi(\widetilde{z}) > \exp \widehat{\rho} > 1 = \psi(0)$. By continuity of ψ, there must be a purely imaginary $\widehat{z}$ which lies between 0 and $\widetilde{z}$ and satisfies $\psi(\widehat{z}) = \exp \widehat{\rho}$—but this contradicts the assumption that z^* had smallest imaginary part. Therefore the zero with smallest imaginary part must, in fact, lie on the imaginary axis. □

Condition (B.25) holds when $(\alpha_1, \alpha_2) = (1, 0)$ or $(0, 1)$ by the finiteness condition.

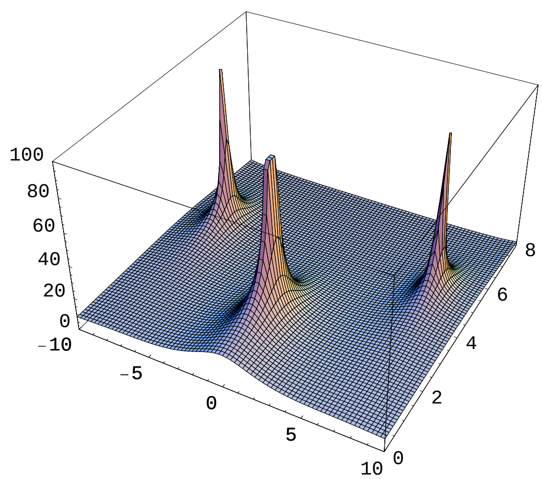


Figure B.2: $1/|\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)|$ plotted for v in a region of the complex plane close to the origin. Zeros of $\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)$ occur at the spikes. The pole nearest the real axis lies on the imaginary axis, at roughly $3i$ in this example.

Figure B.2 illustrates Proposition B.2 (using the calibration of section 2.5.2) by plotting the real-valued function $1/\left|\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)\right|$ over a region of the complex plane close to the origin. When $\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)$ has a zero, this function explodes. Proposition B.2 says that the spike which is closest to the real axis should lie on the imaginary axis—and of course it does.

B.3 The Brownian motion case

From (2.9), the price-dividend ratio on the $\boldsymbol{\alpha}$-asset is

$$P/D(u) = [2\cosh(u/2)]^{\gamma} \cdot \int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_{\gamma}(v)}{\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv)}\, dv\,. \tag{B.26}$$

In this Brownian motion case,

$$\boldsymbol{c}(\theta_1, \theta_2) = \mu_1\theta_1 + \mu_2\theta_2 + \frac{1}{2}\sigma_{11}\theta_1^2 + \sigma_{12}\theta_1\theta_2 + \frac{1}{2}\sigma_{22}\theta_2^2\,.$$

There are two solutions to the equation $\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv) = 0$, each of which lies on the imaginary axis. One—call it $\lambda_1 i$—lies in the upper half-plane; the other—call it $\lambda_2 i$—lies in the lower half-plane. We can rewrite

$$\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv) = B(v - \lambda_1 i)(v - \lambda_2 i)$$

for some $B > 0, \lambda_1 > 0, \lambda_2 < 0$. (I establish the claims made in this paragraph in Step 5, below.)

The aim, then, is to evaluate

$$I \equiv \int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_{\gamma}(v)}{B(v - \lambda_1 i)(v - \lambda_2 i)}\, dv\,, \tag{B.27}$$

in terms of which the price-dividend ratio is

$$P/D = [2\cosh(u/2)]^{\gamma} \cdot I\,. \tag{B.28}$$

The proof of Proposition 2.5, which amounts to evaluating the integral (B.27), is somewhat involved, so I have divided it into several steps. Step 1 starts from the assumption that the state variable u is positive—an assumption that will later be relaxed—and demonstrates that the integral (B.26) can be calculated via the Residue Theorem. (Appendix B.6 defines residues and provides a statement of this theorem.) Steps 2 and 3 carry out these calculations and simplify. Step 4 demonstrates that the resulting expression is also valid for negative u. Step 5 calculates B, λ_1, and λ_2 in terms of fundamental parameters, which concludes the proof.

Step 1. Let $u > 0$. Consider the case in which γ is even. Let $R_n \equiv n + 1/2$, where n is an integer. Define the large semicircle Ω_n to be the semicircle whose base lies along the real axis from $-R_n$ to R_n and which has a semicircular arc (ω_n) passing through the upper half-plane from R_n through $R_n i$ and back to $-R_n$. I will first show that

$$I = \lim_{n\to\infty} \int_{\Omega_n} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{B(v-\lambda_1 i)(v-\lambda_2 i)}\, dv\,. \tag{B.29}$$

Then, from the residue theorem, it will follow that

$$I = 2\pi i \cdot \sum \operatorname{Res}\left\{ \frac{e^{iuv}\mathscr{F}_\gamma(v)}{B(v-\lambda_1 i)(v-\lambda_2 i)}; v_p \right\}\,, \tag{B.30}$$

where the sum is taken over all poles v_p in the upper half-plane.

The first step is to establish that (B.29) holds. The right-hand side is equal to

$$\lim_{n\to\infty} \quad \underbrace{\int_{-R_n}^{R_n} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{B(v-\lambda_1 i)(v-\lambda_2 i)}\, dv}_{I_n} + \underbrace{\int_{\omega_n} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{B(v-\lambda_1 i)(v-\lambda_2 i)}\, dv}_{J_n}$$

The integral I_n tends to I as n tends to infinity. The aim, then, is to establish that the second term J_n tends to zero as n tends to infinity. Along the arc ω_n, we have $v = R_n e^{i\theta}$ where θ varies between 0 and π.

At this point of the argument it is convenient to work with the representation of $\mathscr{F}_\gamma(v)$ of equation (B.11). Substituting from (B.11), we have

$$J_n = \int_0^\pi \frac{e^{iuR_n \cos\theta - uR_n \sin\theta} P(R_n e^{i\theta})}{Q(R_n e^{i\theta}) \left(e^{\pi R_n(\cos\theta + i\sin\theta)} - e^{-\pi R_n(\cos\theta + i\sin\theta)}\right)} \cdot R_n i e^{i\theta}\, d\theta$$

with $P(\cdot)$ and $Q(\cdot)$ polynomials.

To show that J_n tends to zero as n tends to infinity, I separate the range of integration $[0, \pi]$ into two parts: $[\pi/2 - \delta, \pi/2 + \delta]$ and its complement in $[0, \pi]$. Here δ will be chosen to be extremely small.

First, consider

$$\begin{aligned} J_n^{(1)} &\equiv \left| \int_{\pi/2-\delta}^{\pi/2+\delta} \frac{P(R_n e^{i\theta}) e^{iuR_n \cos\theta - uR_n \sin\theta} R_n i e^{i\theta}}{Q(R_n e^{i\theta}) \left(e^{\pi R_n(\cos\theta + i\sin\theta)} - e^{-\pi R_n(\cos\theta + i\sin\theta)}\right)}\, d\theta \right| \\ &\leq \int_{\pi/2-\delta}^{\pi/2+\delta} \left| \frac{P(R_n e^{i\theta})}{Q(R_n e^{i\theta})} \right| \frac{e^{-uR_n \sin\theta} R_n}{\left|e^{\pi R_n(\cos\theta + i\sin\theta)} - e^{-\pi R_n(\cos\theta + i\sin\theta)}\right|}\, d\theta \end{aligned} \tag{B.31}$$

Pick δ sufficiently small that

$$\left|e^{\pi R_n(\cos\theta + i\sin\theta)} - e^{-\pi R_n(\cos\theta + i\sin\theta)}\right| \geq 2 - \varepsilon$$

for all $\theta \in [\pi/2 - \delta, \pi/2 + \delta]$; ε is some very small number close to but greater than zero. This is possible because the left-hand side is continuous and equal to 2 when $\theta = \pi/2$. Then,

$$J_n^{(1)} \leq \int_{\pi/2-\delta}^{\pi/2+\delta} \left| \frac{P(R_n e^{i\theta})}{Q(R_n e^{i\theta})} \right| \frac{e^{-uR_n \sin\theta} R_n}{2 - \varepsilon}\, d\theta \tag{B.32}$$

Since

(i) we can also ensure that δ is small enough that $\sin\theta \geq \varepsilon$ for θ in the range of integration,

(ii) $|P(R_n e^{i\theta})| \leq P_2(R_n)$, where P_2 is the polynomial obtained by taking absolute values of the coefficients in P,

(iii) $Q(R_n e^{i\theta})$ tends to infinity as R_n becomes large, and

(iv) decaying exponentials decay faster than polynomials grow, in the sense that for any positive k and λ, $x^k e^{-\lambda x} \to 0$ as $x \to \infty$, $x \in \mathbb{R}$,

we see, finally, that the right-hand side of (B.32), and hence $J_n^{(1)}$, tends to zero as n tends to infinity,

It remains to be shown that

$$J_n^{(2)} \equiv \left| \int_{[0,\pi/2-\delta]\cup[\pi/2+\delta,\pi]} \frac{P(R_n e^{i\theta}) e^{iuR_n\cos\theta - uR_n\sin\theta} R_n i e^{i\theta}}{Q(R_n e^{i\theta})\left(e^{\pi R_n(\cos\theta+i\sin\theta)} - e^{-\pi R_n(\cos\theta+i\sin\theta)}\right)} \, d\theta \right|$$

is zero in the limit. Since $\delta > 0$, for all θ in the range of integration we have that $|\cos\theta| \geq \zeta > 0$, for some small ζ. We have

$$J_n^{(2)} \leq \int_{[0,\pi/2-\delta]\cup[\pi/2+\delta,\pi]} \left| \frac{P(R_n e^{i\theta})}{Q(R_n e^{i\theta})} \right| \frac{e^{-uR_n\sin\theta} R_n}{\left|e^{\pi R_n(\cos\theta+i\sin\theta)} - e^{-\pi R_n(\cos\theta+i\sin\theta)}\right|} \, d\theta \, .$$

Now,

$$\begin{aligned} & \left|e^{\pi R_n(\cos\theta+i\sin\theta)} - e^{-\pi R_n(\cos\theta+i\sin\theta)}\right| \\ \geq & \left|\left|e^{\pi R_n(\cos\theta+i\sin\theta)}\right| - \left|e^{-\pi R_n(\cos\theta+i\sin\theta)}\right|\right| \\ = & \; e^{\pi R_n|\cos\theta|} - e^{-\pi R_n|\cos\theta|} \\ \geq & \; e^{\pi R_n\zeta} - e^{-\pi R_n\zeta} \end{aligned}$$

for all θ in the range of integration. So,

$$\begin{aligned} J_n^{(2)} & \leq \int_{[0,\pi/2-\delta]\cup[\pi/2+\delta,\pi]} \left| \frac{P(R_n e^{i\theta})}{Q(R_n e^{i\theta})} \right| \frac{e^{-uR_n\sin\theta} R_n}{e^{\pi R_n\zeta} - e^{-\pi R_n\zeta}} \, d\theta \\ & \leq \int_{[0,\pi/2-\delta]\cup[\pi/2+\delta,\pi]} \left| \frac{P(R_n e^{i\theta})}{Q(R_n e^{i\theta})} \right| \frac{R_n}{e^{\pi R_n\zeta} - e^{-\pi R_n\zeta}} \, d\theta \end{aligned}$$

which tends to zero as n tends to infinity.

The case of γ odd is almost identical. The only important difference is that we take $R_n = n$ (as opposed to $n + 1/2$) before allowing n to go to infinity. The reason for doing so is that we must take care to avoid the poles of $\mathscr{F}_\gamma(v)$ on the imaginary axis.

Step 2. From now on, I revert to the definition of $\mathscr{F}_\gamma(v)$ as

$$\mathscr{F}_\gamma(v) = \frac{1}{2\pi} \frac{\Gamma(\gamma/2 - iv)\Gamma(\gamma/2 + iv)}{\Gamma(\gamma)} .$$

The integrand is

$$\frac{e^{iuv}\Gamma(\gamma/2 - iv)\Gamma(\gamma/2 + iv)}{2\pi \cdot B \cdot \Gamma(\gamma) \cdot (v - \lambda_1 i)(v - \lambda_2 i)} , \tag{B.33}$$

which has poles in the upper half-plane at $\lambda_1 i$ and at points v such that $\gamma/2 + iv = -n$ for integers $n \geq 0$, since the Γ-function has poles at the negative integers and zero. In other words, the integrand has poles at $\lambda_1 i$ and at $(n + \gamma/2)i$ for $n \geq 0$.

We can calculate the residue of (B.33) at $v = \lambda_1 i$ directly, using the fact that if $f(z) = g(z)/h(z)$ has a pole at a, and $g(a) \neq 0$, $h(a) = 0$, and $h'(a) \neq 0$, then

$$\text{Res}\{f(z); a\} = \frac{g(a)}{h'(a)} .$$

The residue at $\lambda_1 i$ is therefore

$$\frac{e^{-\lambda_1 u}\Gamma(\gamma/2 + \lambda_1)\Gamma(\gamma/2 - \lambda_1)}{2\pi i \cdot B \cdot \Gamma(\gamma) \cdot (\lambda_1 - \lambda_2)} . \tag{B.34}$$

$\Gamma(z)$ has residue $(-1)^n/n!$ at $z = -n$. (See, for example, Andrews, Askey and Roy (1999, p. 7).) Using this fact, it follows that the residue of (B.33) at $v = (n + \gamma/2)i$ for integers $n \geq 0$ is

$$\frac{-e^{-u(n+\gamma/2)} \cdot \Gamma(\gamma + n) \cdot \frac{(-1)^n}{n!}}{2\pi i \cdot B \cdot \Gamma(\gamma) \cdot (n + \gamma/2 - \lambda_1)(n + \gamma/2 - \lambda_2)} \tag{B.35}$$

Substituting (B.34) and (B.35) into (B.30), we find

$$I = \frac{e^{-\lambda_1 u}\Gamma(\gamma/2+\lambda_1)\Gamma(\gamma/2-\lambda_1)}{B\cdot\Gamma(\gamma)\cdot(\lambda_1-\lambda_2)} - e^{-\gamma u/2}\sum_{n=0}^{\infty}\frac{(-e^{-u})^n\cdot\Gamma(\gamma+n)\cdot\frac{1}{n!}}{B\cdot\Gamma(\gamma)\cdot(n+\gamma/2-\lambda_1)(n+\gamma/2-\lambda_2)}$$

Since $|-e^{-u}| < 1$ under the assumption that $u > 0$, which for the time being is still maintained, we can use the series definition of Gauss's hypergeometric function given in equation (2.17), together with the fact that $\Gamma(\gamma+n)/\Gamma(\gamma) = \gamma(\gamma+1)\cdots(\gamma+n-1)$, to obtain

$$\begin{aligned} I \;=\; & \frac{e^{-\lambda_1 u}}{B(\lambda_1-\lambda_2)}\frac{\Gamma(\gamma/2-\lambda_1)\Gamma(\gamma/2+\lambda_1)}{\Gamma(\gamma)} + \\ & + \frac{e^{-\gamma u/2}}{B(\lambda_1-\lambda_2)}\left[\frac{1}{\gamma/2-\lambda_2}F\left(\gamma,\gamma/2-\lambda_2;1+\gamma/2-\lambda_2;-e^{-u}\right) - \right. \\ & \left. - \frac{1}{\gamma/2-\lambda_1}F\left(\gamma,\gamma/2-\lambda_1;1+\gamma/2-\lambda_1;-e^{-u}\right)\right] \end{aligned} \quad \text{(B.36)}$$

Step 3. A final simplification follows from the fact that

$$\begin{aligned} e^{-\lambda_1 u}\frac{\Gamma(\gamma/2-\lambda_1)\Gamma(\gamma/2+\lambda_1)}{\Gamma(\gamma)} \;=\; & \frac{e^{\gamma u/2}}{\gamma/2+\lambda_1}F\left(\gamma,\gamma/2+\lambda_1;1+\gamma/2+\lambda_1;-e^{u}\right) + \\ & + \frac{e^{-\gamma u/2}}{\gamma/2-\lambda_1}F\left(\gamma,\gamma/2-\lambda_1;1+\gamma/2-\lambda_1;-e^{-u}\right)\ , \end{aligned}$$

which follows from equation (1.8.1.11) of Slater (1966, pp. 35–36).

Using this observation to substitute out the first term in (B.36), we have

$$\begin{aligned} I \;=\; & \frac{1}{B(\lambda_1-\lambda_2)}\left[\frac{e^{\gamma u/2}}{\gamma/2+\lambda_1}F\left(\gamma,\gamma/2+\lambda_1;1+\gamma/2+\lambda_1;-e^{u}\right) + \right. \\ & \left. + \frac{e^{-\gamma u/2}}{\gamma/2-\lambda_2}F\left(\gamma,\gamma/2-\lambda_2;1+\gamma/2-\lambda_2;-e^{-u}\right)\right]. \end{aligned}$$

Substituting this expression into (B.28) gives the formula

$$P/D_1(u) = \frac{[2\cosh(u/2)]^\gamma}{B(\lambda_1-\lambda_2)}\left[\frac{e^{\gamma u/2}}{\gamma/2+\lambda_1}F\left(\gamma,\gamma/2+\lambda_1;1+\gamma/2+\lambda_1;-e^u\right)+ \right.$$
$$\left. + \frac{e^{-\gamma u/2}}{\gamma/2-\lambda_2}F\left(\gamma,\gamma/2-\lambda_2;1+\gamma/2-\lambda_2;-e^{-u}\right)\right]\;; \tag{B.37}$$

thus far, however, the derivation is valid only under the assumption that $u > 0$.

Step 4. Suppose, now, that $u < 0$. Take the complex conjugate of equation (B.27). (This leaves the left-hand side unaltered because the price-dividend ratio is real.) Doing so is equivalent to reframing the problem with $(u, \lambda_1, \lambda_2)$ replaced by $(-u, -\lambda_2, -\lambda_1)$. Since $-u > 0, -\lambda_2 > 0$, and $-\lambda_1 < 0$, the method of steps 1–4 applies unchanged. Since the formula (B.37) is invariant under $(-u, -\lambda_2, -\lambda_1) \mapsto (u, \lambda_1, \lambda_2)$, we can conclude that equation (B.37) is valid for all u. Substituting $u \mapsto \log(1-s)/s$ delivers (2.18).

Step 5. It only remains to find the values of B, λ_1, and λ_2 in terms of the fundamental parameters. The CGF is given, in the general Brownian motion case, by

$$\boldsymbol{c}(\theta_1,\theta_2) = \mu_1\theta_1+\mu_2\theta_2+\frac{1}{2}\sigma_{11}\theta_1^2+\sigma_{12}\theta_1\theta_2+\frac{1}{2}\sigma_{22}\theta_2^2$$

We can rewrite $\rho - \boldsymbol{c}(\alpha_1-\gamma/2-iv, \alpha_2-\gamma/2+iv)$ in the form

$$\rho - \boldsymbol{c}(\alpha_1-\gamma/2-iv,\alpha_2-\gamma/2+iv) = \frac{1}{2}X^2v^2+iYv+\frac{1}{2}Z^2\,, \tag{B.38}$$

where, as in the main text, I have defined

$$
\begin{aligned}
X^2 &\equiv \sigma_{11} - 2\sigma_{12} + \sigma_{22} \\
Y &\equiv \mu_1 - \mu_2 + \alpha_1(\sigma_{11} - \sigma_{12}) - \alpha_2(\sigma_{22} - \sigma_{12}) - \frac{\gamma}{2}(\sigma_{11} - \sigma_{22}) \\
Z^2 &\equiv 2(\rho - \alpha_1\mu_1 - \alpha_2\mu_2) - (\alpha_1^2\sigma_{11} + 2\alpha_1\alpha_2\sigma_{12} + \alpha_2^2\sigma_{22}) + \\
&\quad + \gamma\left[\mu_1 + \mu_2 + \alpha_1\sigma_{11} + (\alpha_1 + \alpha_2)\sigma_{12} + \alpha_2\sigma_{22}\right] - \frac{\gamma^2}{4}(\sigma_{11} + 2\sigma_{12} + \sigma_{22}) .
\end{aligned}
$$

I have chosen to write X^2 and Z^2 to emphasize that these two quantities are positive. The positivity of X^2 follows because it is the variance of the difference of two random variables $(y_{21} - y_{11})$. The positivity of Z^2, on the other hand, follows from the finiteness conditions, after setting $v = 0$ in (B.38).

From (B.38), we have, finally, that

$$
\rho - \boldsymbol{c}(\alpha_1 - \gamma/2 - iv, \alpha_2 - \gamma/2 + iv) = B(v - \lambda_1 i)(v - \lambda_2 i)
$$

where

$$
\begin{aligned}
B &\equiv \frac{1}{2}X^2 \\
\lambda_1 &\equiv \frac{\sqrt{Y^2 + X^2 Z^2} - Y}{X^2} \\
\lambda_2 &\equiv -\frac{\sqrt{Y^2 + X^2 Z^2} + Y}{X^2} .
\end{aligned}
$$

B.3.1 Simple special cases with symmetric Brownian motions

In some special subcases, it is possible to obtain considerably simpler expressions for the price-dividend ratio. In this section, I consider the special case in which the world is symmetrical and the log dividend processes of each asset follow independent drifting Brownian motions with drift μ and volatility σ. It follows that the CGF is given by

$$\boldsymbol{c}(\theta_1,\theta_2) = \mu(\theta_1+\theta_2) + \frac{1}{2}\sigma^2(\theta_1^2+\theta_2^2) \tag{B.39}$$

Recall the general pricing formula, in the form of (2.9):

$$\frac{P_{\boldsymbol{\alpha}}}{D_{\boldsymbol{\alpha}}} = [2\cosh(u/2)]^\gamma \cdot \int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(\alpha_1-\gamma/2-iv, \alpha_2-\gamma/2+iv)}\, dv$$

I will focus on pricing the claim to asset 1, so $\alpha_1 = 1, \alpha_2 = 0$. Substituting in from (B.39), a little algebra confirms the fact that

$$\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv) = \sigma^2\left[(v+i/2)^2 + A^2\right],$$

where $A^2 \equiv (\rho+\mu(\gamma-1))/\sigma^2 - (\gamma-1)^2/4$. The finiteness condition requires that

$$\rho - \boldsymbol{c}(1-\gamma/2, -\gamma/2) > 0 \quad \text{and} \quad \rho - \boldsymbol{c}(1-\gamma, 0) > 0$$

which amounts to the requirement that $A > (\gamma-1)/2$.

The general pricing formula gives the price-dividend ratio of asset 1, written P/D_1, as

$$P/D_1 = [2\cosh(u/2)]^\gamma \cdot \int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\sigma^2\left[(v+i/2)^2+A^2\right]}\, dv\,. \tag{B.40}$$

The question, as before, is where the poles of the integrand are. In the upper half plane, $\mathscr{F}_\gamma(v)$ has infinitely many regularly spaced poles on the imaginary axis, at $(\gamma/2)i$, $(\gamma/2+1)i$, $(\gamma/2+2)i$, ... The other pole is at the zero, in the upper half-plane, of the denominator $\sigma^2\left[(v+i/2)^2+A^2\right]$—that is, at$(A-1/2)i$. It turns out that the integral takes on a relatively simple form if we ensure that the pole at $(A-1/2)i$ is an integer distance from the poles $(\gamma/2)i$, $(\gamma/2+1)i$, etc. (The simple example presented in Cochrane, Longstaff and Santa-Clara (2008) has $\rho = \sigma^2$, so

$A = 1$.) Thus, we want

$$A \in \{(\gamma+1)/2, (\gamma+3)/2, (\gamma+5)/2, \ldots\}$$

For example, if $\gamma = 2$ and $A = 3/2$, the price-dividend ratio of asset 1 is

$$P/D_1(s) = \frac{2(1-s)^3 \log(1-s) + 2s - 5s^2 + 3s^3 - s^3 \log s}{3(1-s)^2 s^3 \sigma^2}\,.$$

B.4 Small asset asymptotics

I start by establishing the claim made in the text that $\rho - \boldsymbol{c}(1-\theta, \theta-\gamma)$ is a concave function of θ. This fact follows directly from

Proposition B.3 (A convexity property of $\boldsymbol{c}(\cdot,\cdot)$)**.** *For arbitrary real numbers α and β, the function $\boldsymbol{c}(\alpha-\theta, \beta+\theta)$ is a convex function of θ.*

Proof. Define the measure $\widehat{\mathbb{P}}$ by

$$\widehat{\mathbb{E}}(A) \equiv e^{-\boldsymbol{c}(\alpha,\beta)} \mathbb{E}\left(e^{\alpha y_{11} + \beta y_{21}} A\right)\,.$$

It follows that the CGF of $y_{21} - y_{11}$, calculated with respect to $\widehat{\mathbb{P}}$, is

$$\begin{aligned}
\widehat{\boldsymbol{c}}(\theta) &= \log \widehat{\mathbb{E}}\left(e^{\theta y_{21} - \theta y_{11}}\right) \\
&= -\boldsymbol{c}(\alpha,\beta) + \log \mathbb{E}\left(e^{(\alpha-\theta) y_{11} + (\beta+\theta) y_{21}}\right) \\
&= -\boldsymbol{c}(\alpha,\beta) + \boldsymbol{c}(\alpha-\theta, \beta+\theta)\,.
\end{aligned}$$

Therefore, $\boldsymbol{c}(\alpha-\theta, \beta+\theta) = \boldsymbol{c}(\alpha,\beta) + \widehat{\boldsymbol{c}}(\theta)$. (Compare also equations (B.2) and (B.3) of Appendix B.1.1.)

The convexity of $\boldsymbol{c}(\alpha-\theta, \beta+\theta)$ follows immediately, because $\widehat{\boldsymbol{c}}(\theta)$, as a CGF, is convex, as shown in Billingsley (1995, pp. 147–8). □

The price-dividend ratio in the small asset limit is given by (2.10), which I reproduce here for the situation in which asset 1 is small:

$$P/D_1 = \lim_{u\to\infty} \frac{\displaystyle\int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)}\,dv}{\displaystyle\int_{-\infty}^{\infty} e^{iuv}\mathscr{F}_\gamma(v)\,dv}. \tag{B.41}$$

By the Riemann-Lebesgue lemma, both the numerator and denominator on the right-hand side of (B.41) tend to zero in the limit as u tends to infinity. What happens to their ratio? This section shows how to calculate limiting price-dividend ratio, riskless rate and excess returns in the small-asset case. For clarity, I work through the price-dividend ratio in detail; the techniques used also apply to the riskless rate and to expected returns, and are very similar to those that were used to provide the closed-form solution in the Brownian motion case.

The following definition provides a convenient label for the poles that will be of interest when evaluating the relevant integrals in the asymptotic limit. (When reading the definition, note that by the finiteness condition and Proposition B.1 of Appendix B.2, the functions to which the definition will be applied will never have poles *on* the real axis.)

Definition B.1. *Let f be an arbitrary meromorphic function. A pole (resp. zero) of f is* minimal *if it lies in the upper half-plane and no other pole (resp. zero) in the upper half-plane has smaller imaginary part.*

Step 1. Consider the integral which makes up the numerator of (B.41),

$$I \equiv \int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)}\,dv\,.$$

If log dividends are drifting Brownian motions, Appendix B.3 showed that this integral could be approached by summing all residues in the upper half-plane. The aim here is to show that the *asymptotic* behavior of this integral is determined only by the minimal residue. Roughly speaking, this is because poles with larger imaginary parts are rendered asymptotically irrelevant by the term e^{iuv}.

To establish this fact, it is convenient to integrate around a contour which avoids all poles except for the minimal pole. If the minimal pole occurs at the minimal zero of $\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)$ then, by Proposition B.2 of Appendix B.2, this pole occurs on the imaginary axis. Otherwise, the minimal pole occurs at the minimal pole of $\mathscr{F}_\gamma(v)$, so is at $i\gamma/2$—which is also on the imaginary axis. In short, we can assume that the minimal pole occurs at the point mi, where $m > 0$ is a real number.

Let $\Box_N$ denote the rectangle in the complex plane with corners at $-N$, N, $N+(m+\varepsilon)i$ and $-N+(m+\varepsilon)i$, with the understanding that integration will take place in the anticlockwise direction. Since the integrand is meromorphic, all poles are isolated, so $\varepsilon > 0$ can be chosen to be sufficiently small that the rectangle $\Box_N$ only contains the pole at mi.

By the residue theorem, we have

$$\begin{aligned} J &\equiv \int_{\Box_N} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)}\, dv \\ &= 2\pi i \operatorname{Res}\left\{ \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)}; mi \right\} \end{aligned}$$

On the other hand, we can also decompose the integral into four pieces:

$$\begin{aligned} J &= \int_{-N}^{N} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)}\, dv + \int_0^{m+\varepsilon} \frac{e^{iu(N+iv)}\mathscr{F}_\gamma(N+iv)}{\rho - \boldsymbol{c}(\ldots)}\, i\, dv + \\ &\quad + \int_N^{-N} \frac{e^{iu(v+(m+\varepsilon)i)}\mathscr{F}_\gamma(v+(m+\varepsilon)i)}{\rho - \boldsymbol{c}(\ldots)}\, dv + \int_{m+\varepsilon}^{0} \frac{e^{iu(-N+iv)}\mathscr{F}_\gamma(-N+iv)}{\rho - \boldsymbol{c}(\ldots)}\, i\, dv \\ &\equiv J_1 + J_2 + J_3 + J_4 \end{aligned}$$

In brief, the desired result follows on first letting N tend to infinity; then J_2 and J_4 go to zero. Subsequently letting u go to infinity, J_3 becomes asymptotically irrelevant compared to J_1. By the residue theorem, the integral $I = \lim_{N\to\infty} J_1$ is therefore asymptotically equivalent[2] to $2\pi i$ times the residue at mi:

$$I \sim 2\pi i \cdot \operatorname{Res}\left\{\frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2 - iv, -\gamma/2 + iv)}; mi\right\}.$$

The following calculations justify these statements. Consider J_2. Since the range of integration is a closed and bounded interval, the function $|\rho - \boldsymbol{c}(\ldots)|$ attains its maximum and minimum on the range. Since also the function has no zeros on the interval, we can write $|\rho - \boldsymbol{c}(\ldots)| \geq \delta_1 > 0$ for all v in the range of integration. We have

$$\begin{aligned} |J_2| &\leq \int_0^{m+\varepsilon} \left| \frac{e^{iu(N+iv)}\mathscr{F}_\gamma(N+iv)}{\rho - \boldsymbol{c}(\ldots)} i \right| dv \\ &= \int_0^{m+\varepsilon} \frac{e^{-uv}\,|\mathscr{F}_\gamma(N+iv)|}{|\rho - \boldsymbol{c}(\ldots)|} dv \\ &\leq \frac{1}{\delta_1}\int_0^{m+\varepsilon} |\mathscr{F}_\gamma(N+iv)|\, dv \\ &\to 0 \end{aligned}$$

as N tends to infinity because $|\mathscr{F}_\gamma(N+iv)|$ converges to zero uniformly over v in the range of integration. An almost identical argument shows that $|J_4|$ tends to zero as N tends to infinity.

Now consider J_3. Set $\delta_2 = |\rho - \boldsymbol{c}(1-\gamma/2 + m + \varepsilon, -\gamma/2 - m - \varepsilon)| > 0$; then by the ridge property discussed in Appendix B.2, $|\rho - \boldsymbol{c}(\ldots)| \geq \delta_2$ for all v in the range

[2] I write $f(x) \sim g(x)$—"$f(x)$ is asymptotically equivalent to $g(x)$"—to indicate that $\lim_{x\to\infty} f(x)/g(x) = 1$. Below, I also use the "big-O" notation $f(x) = O(g(x))$—"$f(x)$ is asymptotically of the same order as $g(x)$"—to indicate that $\lim_{x\to\infty} f(x)/g(x)$ is finite.

of integration. It follows that

$$\begin{aligned} |J_3| &\le \int_{-N}^{N} \frac{e^{-(m+\varepsilon)u}\,|\mathscr{F}_\gamma(v+(m+\varepsilon)i)|}{|\rho-\boldsymbol{c}(\ldots)|}\,dv \\ &\le e^{-u(m+\varepsilon)}\cdot\frac{1}{\delta_2}\int_{-N}^{N}|\mathscr{F}_\gamma(v+(m+\varepsilon)i)|\,dv \\ &\to e^{-u(m+\varepsilon)}\cdot X/\delta_2 \end{aligned}$$

where X is the (finite) limit of the integral $\int_{-N}^{N}|\mathscr{F}_\gamma(v+(m+\varepsilon)i)|\,dv$ as N tends to infinity. (X is finite because $\mathscr{F}_\gamma(v+(m+\varepsilon)i)$ decays to zero exponentially fast as $v\to\pm\infty$.)

By the residue theorem,

$$J_1+J_2+J_3+J_4 = 2\pi i\times\text{residue at } mi = O(e^{-mu})\,.$$

Let N go to infinity; then J_2 and J_4 go to zero, J_1 tends to I and J_3 tends to $e^{-u(m+\varepsilon)}X$, so

$$I+e^{-u(m+\varepsilon)}X = 2\pi i\times\text{residue at } mi = O(e^{-mu})\,.$$

In the limit as $u\to\infty$, $e^{-u(m+\varepsilon)}X$ is exponentially smaller than e^{-mu}, so

$$I\sim 2\pi i\,\mathrm{Res}\left\{\frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)};mi\right\}$$

as $u\to\infty$. The asymptotic behavior of the integral I is dictated by the residue closest to the real line.

Essentially identical arguments can be made to show that the other relevant integrals are asymptotically equivalent to $2\pi i$ times the minimal residue of the relevant integrand; they are omitted to prevent an already complicated argument becoming totally unreadable.

Step 2. I now apply the logic of step 1 to (i) the price-dividend ratio, (ii) the riskless rate and (iii) expected returns.

(i) In the price-dividend ratio case, we have to evaluate

$$\lim_{u\to\infty} P/D(u) = \lim_{u\to\infty} \frac{\displaystyle\int_{-\infty}^{\infty} \frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)}\, dv}{\displaystyle\int_{-\infty}^{\infty} e^{iuv}\mathscr{F}_\gamma(v)\, dv} \equiv \lim_{u\to\infty} \frac{I_n}{I_d}\,.$$

We have just seen that I_n and I_d are asymptotically equivalent to $2\pi i$ times the residue at the pole (of the relevant integrand) with smallest imaginary part. Here, I take this fact as given and refer to the pole (or zero) with least positive imaginary part as the *minimal* pole (or zero).

Consider, then, the more complicated integral I_n. The integrand has a pole at $i\gamma/2$ due to a singularity in $\mathscr{F}_\gamma(v)$. The question is whether or not there is a zero of $\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)$ for some v with imaginary part smaller than $\gamma/2$. If there is, then this is the minimal pole. If not, then $i\gamma/2$ is the minimal pole.

In Appendix B.2, it was shown that the minimal zero of $\rho - \boldsymbol{c}(1-\gamma/2-iv, -\gamma/2+iv)$ lies on the imaginary axis. Thus the zero in question is of the form z^*i for some positive real z^* satisfying

$$\rho - \boldsymbol{c}(1-\gamma/2+z^*, -\gamma/2-z^*) = 0\,. \tag{B.42}$$

If $z^* < \gamma/2$, we are in the supercritical case; if $z^* > \gamma/2$, we are in the subcritical case. (At the end of the proof, I will define $\theta^* = \gamma/2 - z^*$, simply for notational convenience.)

(a) In the subcritical case, the minimal pole for both integrals is at $i\gamma/2$. We therefore have, asymptotically,

$$
\begin{aligned}
P/D \quad &\longrightarrow \quad \frac{\operatorname{Res}\left\{\dfrac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)};i\gamma/2\right\}}{\operatorname{Res}\left\{e^{iuv}\mathscr{F}_\gamma(v);i\gamma/2\right\}} \\
&= \quad \frac{1}{\rho-\boldsymbol{c}(1,-\gamma)}\cdot\frac{\operatorname{Res}\left\{e^{iuv}\mathscr{F}_\gamma(v);i\gamma/2\right\}}{\operatorname{Res}\left\{e^{iuv}\mathscr{F}_\gamma(v);i\gamma/2\right\}} \\
&= \quad \frac{1}{\rho-\boldsymbol{c}(1,-\gamma)}
\end{aligned}
$$

(b) In the supercritical case, the minimal pole is at iz^* for I_n and at $i\gamma/2$ for I_d. We therefore have

$$
\begin{aligned}
P/D \quad &\longrightarrow \quad \frac{\operatorname{Res}\left\{\dfrac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)};iz^*\right\}}{\operatorname{Res}\left\{e^{iuv}\mathscr{F}_\gamma(v);i\gamma/2\right\}} \\
&= \quad e^{u(\gamma/2-z^*)}\cdot\frac{\operatorname{Res}\left\{\dfrac{\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)};iz^*\right\}}{\operatorname{Res}\left\{\mathscr{F}_\gamma(v);i\gamma/2\right\}} \\
&\longrightarrow \quad \infty
\end{aligned}
$$

as u tends to infinity because $\gamma/2-z^*>0$.

To see that the price-consumption ratio, $P/C=s\cdot P/D$, remains finite in this limit, we must evaluate $\lim_{s\to 0} s\cdot P/D$. Since $s=1/(1+e^u)\sim e^{-u}$, we have, asymptotically, that

$$
P/C \longrightarrow e^{u(\gamma/2-z^*-1)}\cdot\frac{\operatorname{Res}\left\{\dfrac{\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)};iz^*\right\}}{\operatorname{Res}\left\{\mathscr{F}_\gamma(v);i\gamma/2\right\}},
$$

which tends to zero as $u\to\infty$ because $\gamma/2-z^*-1<0$.

(ii) In the riskless rate case, we seek the limit of

$$r = \frac{\displaystyle\int_{-\infty}^{\infty} \mathscr{F}_\gamma(v) e^{iuv} \cdot [\rho - \boldsymbol{c}(-\gamma/2 - iv, -\gamma/2 + iv)]\, dv}{\displaystyle\int_{-\infty}^{\infty} \mathscr{F}_\gamma(v) e^{iuv}\, dv}\,.$$

This is much simpler, because the minimal pole is $i\gamma/2$ for both numerator and denominator. It follows that

$$r \longrightarrow \rho - \boldsymbol{c}(-\gamma/2 - i(i\gamma/2), -\gamma/2 + i(i\gamma/2)) = \rho - \boldsymbol{c}(0, -\gamma)\,.$$

(iii) In the expected return case, we need the limit of the expected capital gain expression which is the first term on the right-hand side of (2.12). This expression is asymptotically equivalent to

$$\frac{\displaystyle\int_{-\infty}^{\infty} \frac{e^{iuv} \mathscr{F}_\gamma(v) \boldsymbol{c}(1 - \gamma/2 - iv, \gamma/2 + iv)}{\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)}\, dv}{\displaystyle\int_{-\infty}^{\infty} \frac{e^{iuv} \mathscr{F}_\gamma(v)}{\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)}\, dv} \equiv \frac{J_n}{J_d}$$

since the higher-order exponential terms e^{-mu} for $m \geq 1$ which appear in (2.12) become irrelevant exponentially fast as u tends to infinity. Again, there are two subcases, depending on whether the minimal zero of $\rho - \boldsymbol{c}(1 - \gamma/2 - iv, -\gamma/2 + iv)$ has imaginary part greater than or less than $\gamma/2$.

(a) In the subcritical case, the minimal pole of each of J_n and J_d occurs at

$i\gamma/2$. Therefore we have

$$\lim_{u\to\infty} \mathbb{E}dP/P = \frac{\operatorname{Res}\left\{\frac{e^{iuv}\mathscr{F}_\gamma(v)\boldsymbol{c}(1-\gamma/2-iv,\gamma/2+iv)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)}; i\gamma/2\right\}}{\operatorname{Res}\left\{\frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)}; i\gamma/2\right\}} = \boldsymbol{c}(1,0)\,.$$

(b) In the supercritical case, the minimal pole of each of J_n and J_d occurs at iz^*. Therefore, we have

$$\lim_{u\to\infty} \mathbb{E}dP/P = \frac{\operatorname{Res}\left\{\frac{e^{iuv}\mathscr{F}_\gamma(v)\boldsymbol{c}(1-\gamma/2-iv,\gamma/2+iv)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)}; iz^*\right\}}{\operatorname{Res}\left\{\frac{e^{iuv}\mathscr{F}_\gamma(v)}{\rho-\boldsymbol{c}(1-\gamma/2-iv,-\gamma/2+iv)}; iz^*\right\}} = \boldsymbol{c}(1-\gamma/2+z^*,\gamma/2-z^*)\,.$$

Since instantaneous expected returns are the sum of expected capital gains and the dividend-price ratio, expected returns in the asymptotic limit are

$$\boldsymbol{c}(1,0)+\rho-\boldsymbol{c}(1,-\gamma)$$

in the subcritical case, and

$$\boldsymbol{c}(1-\gamma/2+z^*,\gamma/2-z^*)$$

in the supercritical case.

Subtracting the riskless rate, we have, finally, that excess returns are

$$\boldsymbol{c}(1,0)+\boldsymbol{c}(0,-\gamma)-\boldsymbol{c}(1,-\gamma)$$

in the subcritical case, and

$$\boldsymbol{c}(1-\gamma/2+z^*, \gamma/2-z^*) - \rho + \boldsymbol{c}(0,-\gamma)$$

in the supercritical case. Recalling that $\rho = \boldsymbol{c}(1-\gamma/2+z^*, -\gamma/2-z^*)$ by the definition of z^*, the excess return in the supercritical case can be rewritten as

$$\boldsymbol{c}(1-\gamma/2+z^*, \gamma/2-z^*) + \boldsymbol{c}(0,-\gamma) - \boldsymbol{c}(1-\gamma/2+z^*, -\gamma/2-z^*)\,.$$

This concludes the derivation of the various asymptotics in the general case.

Step 3. If dividends are also independent across assets then we can decompose

$$\boldsymbol{c}(\theta_1, \theta_2) = \boldsymbol{c}_1(\theta_1) + \boldsymbol{c}_2(\theta_2)$$

where $\boldsymbol{c}_i(\theta) \equiv \log \mathbb{E} \exp \theta y_{i1}$. It follows that in the subcritical case,

$$XS \longrightarrow \boldsymbol{c}(1,0) + \boldsymbol{c}(0,-\gamma) - \boldsymbol{c}(1,-\gamma) = 0$$

and in the supercritical case,

$$\begin{aligned} XS &\longrightarrow \boldsymbol{c}(1-\gamma/2+z^*, \gamma/2-z^*) + \boldsymbol{c}(0,-\gamma) - \boldsymbol{c}(1-\gamma/2+z^*, -\gamma/2-z^*) \\ &= \boldsymbol{c}_2(\gamma/2-z^*) + \boldsymbol{c}_2(-\gamma) - \boldsymbol{c}_2(-\gamma/2-z^*)\,. \end{aligned}$$

Step 4. I now show that this last expression is positive. First, note that because $\boldsymbol{c}_2(x)$—as a CGF—is convex, we have that

$$\frac{\boldsymbol{c}_2(e)-\boldsymbol{c}_2(d)}{e-d} < \frac{\boldsymbol{c}_2(g)-\boldsymbol{c}_2(f)}{g-f} \quad \text{whenever } d<e<f<g\,.$$

Next, observe that in the supercritical case, we have

$$-\gamma < -\gamma/2 - z^* < 0 < \gamma/2 - z^* .$$

It follows that

$$\frac{\boldsymbol{c}_2(-\gamma/2 - z^*) - \boldsymbol{c}_2(-\gamma)}{(-\gamma/2 - z^*) - (-\gamma)} < \frac{\boldsymbol{c}_2(\gamma/2 - z^*) - \boldsymbol{c}_2(0)}{(\gamma/2 - z^*) - 0} ,$$

or equivalently, because $\boldsymbol{c}_2(0) = 0$,

$$\boldsymbol{c}_2(-\gamma/2 - z^*) - \boldsymbol{c}_2(-\gamma) < \boldsymbol{c}_2(\gamma/2 - z^*) ;$$

and so

$$\boldsymbol{c}_2(\gamma/2 - z^*) + \boldsymbol{c}_2(-\gamma) - \boldsymbol{c}_2(-\gamma/2 - z^*) > 0$$

as required.

Step 5. The last step showed that $R_1 = R_f$ in the subcritical case and $R_1 > R_f$ in the supercritical case. It only remains to show that the other bounds on expected returns hold: that (i) $R_1 < R_2$, assuming independence, and that (ii) in the supercritical case $R_1 < G_1$, assuming $G_1 \geq G_2$.

Step 5(i). Proof that $R_1 < R_2$, assuming independence:

In the subcritical case, $R_1 = \rho + \boldsymbol{c}(1,0) - \boldsymbol{c}(1,-\gamma)$ and $R_2 = \rho + \boldsymbol{c}(0,1) - \boldsymbol{c}(0,1-\gamma)$. Since we are assuming independence, it remains to show that

$$-\boldsymbol{c}_2(-\gamma) < \boldsymbol{c}_2(1) - \boldsymbol{c}_2(1-\gamma) ,$$

or equivalently that

$$\boldsymbol{c}_2(1-\gamma) < \boldsymbol{c}_2(1) + \boldsymbol{c}_2(-\gamma) ,$$

which follows immediately by convexity of $\boldsymbol{c}_2(\cdot)$.

In the supercritical case, $R_1 = \boldsymbol{c}(1-\gamma/2+z^*, \gamma/2-z^*)$ and $R_2 = \boldsymbol{c}(1-\gamma/2+z^*, -\gamma/2-z^*) + \boldsymbol{c}(0,1) - \boldsymbol{c}(0,1-\gamma)$ (substituting in for ρ from the definition of z^*). By independence, it remains to show that

$$\boldsymbol{c}_2(\gamma/2-z^*) < \boldsymbol{c}_2(-\gamma/2-z^*) + \boldsymbol{c}_2(1) - \boldsymbol{c}_2(1-\gamma)\,,$$

or equivalently that

$$\boldsymbol{c}_2(1-\gamma) + \boldsymbol{c}_2(\gamma/2-z^*) < \boldsymbol{c}_2(1) + \boldsymbol{c}_2(-\gamma/2-z^*)$$

which also follows directly by convexity of $\boldsymbol{c}_2(\cdot)$, noting that $\gamma/2 - z^* \in (0,1)$.

Step 5(ii). Next, I show that in the supercritical case, $R_1 \leq G_1$ if $G_1 \geq G_2$. We do not need the independence assumption here. It will be helpful to write $\theta = \gamma/2 - z^* \in (0,1)$. With this notation, the limiting $R_1 = \boldsymbol{c}(1-\theta, \theta)$. The claim is that $\boldsymbol{c}(1-\theta,\theta) \leq \boldsymbol{c}(1,0)$. To show this, we make the same change of measure as was used in the proof of Proposition B.3. We have $R_1 = \boldsymbol{c}(1-\theta,\theta) = \boldsymbol{c}(1,0) + \widehat{\boldsymbol{c}}(-\theta)$. It suffices to show that $\widehat{\boldsymbol{c}}(-\theta) \leq 0$ for all θ in $(0,1)$. We have $\boldsymbol{c}(0,1) = \boldsymbol{c}(1,0) + \widehat{\boldsymbol{c}}(-1)$, and so by assumption $\widehat{\boldsymbol{c}}(-1) \leq 0$. Since $\widehat{\boldsymbol{c}}(0) = 0$, the claim follows by convexity of $\widehat{\boldsymbol{c}}(\cdot)$.

Finally, it is notationally convenient to set $\theta^* = \gamma/2 - z^*$. It follows from (B.42) that the defining property of θ^* in the supercritical case is that

$$\rho - \boldsymbol{c}(1-\theta^*, \theta^* - \gamma) = 0\,.$$

B.5 The N-asset case

B.5.1 The Fourier transform $\mathscr{F}_\gamma^N(\boldsymbol{v})$

To make a start, we seek the integral

$$I_N \equiv \int_{\mathbb{R}^{N-1}} \frac{e^{-ix_1v_1-ix_2v_2-\cdots-ix_{N-1}v_{N-1}}}{\left(e^{x_1/N}+\cdots+e^{x_{N-1}/N}+e^{-(x_1+x_2+\ldots+x_{N-1})/N}\right)^\gamma}\,dx_1\,\ldots\,dx_{N-1}\,. \quad \text{(B.43)}$$

For notational convenience, write $x_N \equiv -x_1-\cdots-x_{N-1}$—so $\sum_1^N x_i = 0$—and, for $i=1,\ldots,N$, define

$$t_i = \frac{e^{x_i/N}}{e^{x_1/N}+\cdots+e^{x_N/N}}\,. \quad \text{(B.44)}$$

Note that the variables t_i range between 0 and 1 (and, by construction, sum to 1) as the variables $\{x_i\}$ range around. Furthermore, we have

$$\begin{aligned} \prod_{k=1}^N t_k &= \frac{e^{(x_1+\cdots+x_N)/N}}{\left(e^{x_1/N}+\cdots+e^{x_N/N}\right)^N} \\ &= \frac{1}{\left(e^{x_1/N}+\cdots+e^{x_N/N}\right)^N} \\ \text{and}\quad t_i^N &= \frac{e^{x_i}}{\left(e^{x_1/N}+\cdots+e^{x_N/N}\right)^N}\,, \end{aligned}$$

so

$$e^{x_i} = \frac{t_i^N}{\prod_{k=1}^N t_k}\,.$$

Of course, because of the linear dependence $\sum_{k=1}^N t_k = 1$, there are only $N-1$ independent variables and $t_N = 1-t_1-\cdots-t_{N-1}$, so we can rewrite

$$x_i = N\log t_i - \sum_{k=1}^{N-1}\log t_k - \log\left(1-\sum_{k=1}^{N-1}t_k\right)\,,\quad i=1,\ldots,N-1\,. \quad \text{(B.45)}$$

To make the change of variables specified in (B.44), we have to calculate the

Jacobian

$$J \equiv \left| \frac{\partial(x_1, \ldots, x_{N-1})}{\partial(t_1, \ldots, t_{N-1})} \right| .$$

From (B.45),

$$\frac{\partial x_i}{\partial t_j} = \frac{1}{t_N} - \frac{1}{t_j} + \frac{N\delta_{ij}}{t_i} ,$$

where δ_{ij} equals one if $i = j$ and zero otherwise, so we can write

$$\begin{aligned}
\frac{\partial(x_1, \ldots, x_{N-1})}{\partial(t_1, \ldots, t_{N-1})} &= \begin{pmatrix} \frac{N}{t_1} & & & \\ & \frac{N}{t_2} & & \\ & & \ddots & \\ & & & \frac{N}{t_{N-1}} \end{pmatrix} + \begin{pmatrix} \frac{1}{t_N} - \frac{1}{t_1} & \frac{1}{t_N} - \frac{1}{t_2} & \cdots & \frac{1}{t_N} - \frac{1}{t_{N-1}} \\ \frac{1}{t_N} - \frac{1}{t_1} & \frac{1}{t_N} - \frac{1}{t_2} & \cdots & \frac{1}{t_N} - \frac{1}{t_{N-1}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{t_N} - \frac{1}{t_1} & \frac{1}{t_N} - \frac{1}{t_2} & \cdots & \frac{1}{t_N} - \frac{1}{t_{N-1}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{N}{t_1} & & & \\ & \frac{N}{t_2} & & \\ & & \ddots & \\ & & & \frac{N}{t_{N-1}} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \frac{1}{t_N} - \frac{1}{t_1} \\ \frac{1}{t_N} - \frac{1}{t_2} \\ \vdots \\ \frac{1}{t_N} - \frac{1}{t_{N-1}} \end{pmatrix}' \\
&\equiv \boldsymbol{A} + \boldsymbol{\alpha}\boldsymbol{\beta}' .
\end{aligned}$$

The last line defines the $(N-1) \times (N-1)$ matrix $\boldsymbol{A}$ and the $(N-1)$-dimensional column vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. $\boldsymbol{A}$ is a diagonal matrix: blanks indicate zeros. The prime symbol ($'$) denotes a transpose.

In order to calculate $J = \det(\boldsymbol{A} + \boldsymbol{\alpha}\boldsymbol{\beta}')$ we can use the following

Fact B.1 (Matrix determinant lemma)**.** *Suppose that* $\boldsymbol{A}$ *is an invertible square matrix and that* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\beta}$ *are column vectors, each of length equal to the dimension of* $\boldsymbol{A}$*. Then*

$$\det(\boldsymbol{A} + \boldsymbol{\alpha}\boldsymbol{\beta}') = \left(1 + \boldsymbol{\beta}'\boldsymbol{A}^{-1}\boldsymbol{\alpha}\right) \det \boldsymbol{A} .$$

This fact is useful in the present case because $\boldsymbol{A}$ is diagonal, so its inverse and determinant are easily calculated. To be specific,

$$\begin{aligned} \det \boldsymbol{A} &= \frac{N^{N-1}}{t_1 \cdots t_{N-1}} \\ \text{and} \quad \boldsymbol{A}^{-1} &= \begin{pmatrix} \frac{t_1}{N} & & & \\ & \frac{t_2}{N} & & \\ & & \ddots & \\ & & & \frac{t_{N-1}}{N} \end{pmatrix}. \end{aligned}$$

It follows that

$$\begin{aligned} J &= \left[1 + \begin{pmatrix} \frac{1}{t_N} - \frac{1}{t_1} \\ \frac{1}{t_N} - \frac{1}{t_2} \\ \vdots \\ \frac{1}{t_N} - \frac{1}{t_{N-1}} \end{pmatrix}' \begin{pmatrix} \frac{t_1}{N} & & & \\ & \frac{t_2}{N} & & \\ & & \ddots & \\ & & & \frac{t_{N-1}}{N} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right] \times \frac{N^{N-1}}{t_1 \cdots t_{N-1}} \\ &= \frac{N^{N-2}}{t_1 \cdots t_N}. \end{aligned}$$

We can now return to the integral I_N. For typographical reasons, I write Π for the product $\prod_{k=1}^{N} t_k$ and suppress the range of integration, which is $[0,1]^{N-1}$. Making the substitution suggested in (B.44),

$$\begin{aligned} I_N &= \int \frac{\left(\frac{t_1^N}{\Pi}\right)^{-iv_1} \left(\frac{t_2^N}{\Pi}\right)^{-iv_2} \cdots \left(\frac{t_{N-1}^N}{\Pi}\right)^{-iv_{N-1}}}{\left(\frac{t_1+t_2+\cdots+t_N}{\Pi^{1/N}}\right)^{\gamma}} \cdot J \, dt_1 \ldots dt_{N-1} \\ &= N^{N-2} \int \Pi^{\gamma/N} \left(\frac{t_1^N}{\Pi}\right)^{-iv_1} \cdots \left(\frac{t_{N-1}^N}{\Pi}\right)^{-iv_{N-1}} \frac{dt_1 \ldots dt_{N-1}}{t_1 \ldots t_{N-1} t_N} \\ &= N^{N-2} \int \Big(t_1^{\gamma/N + iv_1 + \cdots + iv_{N-1} - Niv_1} t_2^{\gamma/N + iv_1 + \cdots + iv_{N-1} - Niv_2} \cdots \\ &\qquad \cdots t_{N-1}^{\gamma/N + iv_1 + \cdots + iv_{N-1} - Niv_{N-1}} \cdot t_N^{\gamma/N + iv_1 + \cdots + iv_{N-1}} \Big) \frac{dt_1 \ldots dt_{N-1}}{t_1 \ldots t_{N-1} t_N}. \end{aligned}$$

As in the two-asset case, this is a Dirichlet surface integral. As shown in Andrews, Askey and Roy (1999, p. 34), it can be evaluated in terms of Γ-functions: we have

$$I_N = \frac{N^{N-2}}{\Gamma(\gamma)} \cdot \Gamma\left(\gamma/N + iv_1 + iv_2 + \ldots + iv_{N-1}\right) \cdot \prod_{k=1}^{N-1} \Gamma\left(\gamma/N + iv_1 + \cdots + iv_{N-1} - Niv_k\right) .$$

Defining $\mathscr{G}_\gamma^N(\boldsymbol{v}) = I_N/(2\pi)^{N-1}$, where $\boldsymbol{v} = (v_1, \ldots, v_{N-1})$, we have

$$\mathscr{G}_\gamma^N(\boldsymbol{v}) = \frac{N^{N-2}}{(2\pi)^{N-1}} \cdot \frac{\Gamma\left(\gamma/N + iv_1 + iv_2 + \ldots + iv_{N-1}\right)}{\Gamma(\gamma)} \cdot \prod_{k=1}^{N-1} \Gamma\left(\gamma/N + iv_1 + \cdots + iv_{N-1} - Niv_k\right) . \tag{B.46}$$

It follows from this definition of $\mathscr{G}_\gamma^N(\boldsymbol{v})$, by the Fourier inversion theorem, that

$$\frac{1}{\left(e^{x_1/N} + e^{x_2/N} + \ldots + e^{-(x_1+x_2+\ldots+x_{N-1})/N}\right)^\gamma} = \int_{\mathbb{R}^{N-1}} \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{x}}\, d\boldsymbol{v} , \tag{B.47}$$

where $\boldsymbol{x} = (x_1, \ldots, x_{N-1})$.

B.5.2 The expectation

We seek the expectation

$$E = \mathbb{E}\left[\frac{e^{\boldsymbol{\alpha}'\widetilde{\boldsymbol{y}}_t}}{\left(e^{y_{10}+\widetilde{y}_{1t}} + \cdots + e^{y_{N0}+\widetilde{y}_{Nt}}\right)^\gamma}\right] ,$$

where $\boldsymbol{\alpha} \equiv (\alpha_1, \ldots, \alpha_N)'$ and $\widetilde{\boldsymbol{y}}_t \equiv (\widetilde{y}_{1t}, \ldots, \widetilde{y}_{Nt})'$.

The calculation is carried out by applying the same three tricks that were useful in the two-asset case: namely, by putting the denominator in a form amenable to a Fourier transform; then changing measure, to take care of the numerator; and finally applying the Fourier transform.

The calculations below also use the vectors $\boldsymbol{y}_0$ and $\boldsymbol{\gamma}$ defined in the main text.

In addition, define the $(N-1)\times N$ matrix $\boldsymbol{Q}$ and vectors $\boldsymbol{q}_i$ by

$$\boldsymbol{Q} \equiv \begin{pmatrix} \boldsymbol{q}_2' \\ \boldsymbol{q}_3' \\ \vdots \\ \boldsymbol{q}_N' \end{pmatrix} \equiv \begin{pmatrix} -1 & N-1 & -1 & \cdots & -1 \\ -1 & -1 & N-1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -1 \\ -1 & -1 & \cdots & -1 & N-1 \end{pmatrix} , \tag{B.48}$$

and let $\boldsymbol{q}_1 \equiv (N-1,\ldots,-1,-1)'$—the "missing" row which does *not* appear as the top row of $\boldsymbol{Q}$. (This definition is only introduced to make certain expressions neater, since $\boldsymbol{q}_1 = -\sum_{k=2}^{N} \boldsymbol{q}_k$.)

We will also need to make a change of measure at one stage, as in the two asset case. Define $\widetilde{\mathbb{E}}$ by

$$\widetilde{\mathbb{E}}\left[Y\right] \equiv e^{-t\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N)} \cdot \mathbb{E}\left[e^{(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N)'\widetilde{\boldsymbol{y}}_t} \cdot Y\right] . \tag{B.49}$$

It follows that

$$\widetilde{\boldsymbol{c}}(\boldsymbol{v}) \equiv \log \widetilde{\mathbb{E}}\left[e^{\widetilde{\boldsymbol{y}}_1'\boldsymbol{v}}\right] = \boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+\boldsymbol{v}) - \boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N) . \tag{B.50}$$

Using the new notation,

$$\begin{aligned} E &= \mathbb{E}\left[\frac{e^{\boldsymbol{\alpha}'\widetilde{\boldsymbol{y}}_t}}{\left(e^{y_{10}+\widetilde{y}_{1t}} + \cdots + e^{y_{N0}+\widetilde{y}_{Nt}}\right)^{\gamma}}\right] \\ &= \mathbb{E}\left[\frac{e^{\boldsymbol{\alpha}'\widetilde{\boldsymbol{y}}_t - \boldsymbol{\gamma}'(\boldsymbol{y}_0+\widetilde{\boldsymbol{y}}_t)/N}}{\left(e^{\boldsymbol{q}_1'(\boldsymbol{y}_0+\widetilde{\boldsymbol{y}}_t)/N} + \cdots + e^{\boldsymbol{q}_N'(\boldsymbol{y}_0+\widetilde{\boldsymbol{y}}_t)/N}\right)^{\gamma}}\right] \\ &= e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N} e^{\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N)t}\, \widetilde{\mathbb{E}}\left[\frac{1}{\left(e^{\boldsymbol{q}_1'(\boldsymbol{y}_0+\widetilde{\boldsymbol{y}}_t)/N} + \cdots + e^{\boldsymbol{q}_N'(\boldsymbol{y}_0+\widetilde{\boldsymbol{y}}_t)/N}\right)^{\gamma}}\right] . \end{aligned}$$

Now, $\boldsymbol{Q}(\boldsymbol{y}_0 + \widetilde{\boldsymbol{y}}_t)$ plays the role of $\boldsymbol{x}$ in expression (B.47). It follows that

$$\begin{aligned} E &= e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N} e^{\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N)t}\, \widetilde{\mathbb{E}}\left[\int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}(\boldsymbol{y}_0+\widetilde{\boldsymbol{y}}_t)}\, d\boldsymbol{v}\right] \\ &= e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N} e^{\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N)t} \int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} e^{\widetilde{\boldsymbol{c}}(i\boldsymbol{Q}'\boldsymbol{v})t}\, d\boldsymbol{v} \\ &= e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N} \int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} e^{\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+i\boldsymbol{Q}'\boldsymbol{v})t}\, d\boldsymbol{v}\,. \end{aligned} \tag{B.51}$$

B.5.3 Prices

Now we proceed along the same lines as in the two-asset case. First, the price of the $\boldsymbol{\alpha}$-asset is given by

$$\begin{aligned} P &= \mathbb{E}\int_0^\infty e^{-\rho t}\left(\frac{C_t}{C_0}\right)^{-\gamma} D_{1t}^{\alpha_1}\cdots D_{Nt}^{\alpha_N}\, dt \\ &= C_0^\gamma \int_0^\infty e^{-\rho t}\mathbb{E}\left[\frac{e^{\alpha_1(y_{10}+\widetilde{y}_{1t})+\cdots+\alpha_N(y_{N0}+\widetilde{y}_{Nt})}}{\left(e^{y_{10}+\widetilde{y}_{1t}}+\cdots+e^{y_{N0}+\widetilde{y}_{Nt}}\right)^\gamma}\right] dt\,. \end{aligned}$$

The price-dividend ratio is therefore equal to

$$P/D = C_0^\gamma \int_0^\infty e^{-\rho t}\mathbb{E}\left[\frac{e^{\alpha_1\widetilde{y}_{1t}+\cdots+\alpha_N\widetilde{y}_{Nt}}}{\left(e^{y_{10}+\widetilde{y}_{1t}}+\cdots+e^{y_{N0}+\widetilde{y}_{Nt}}\right)^\gamma}\right] dt\,,$$

and the expectation was calculated, as E, in the previous section. Substituting in from (B.51),

$$\begin{aligned} P/D &= C_0^\gamma \int_{t=0}^\infty e^{-\rho t}\left(e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N}\int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} e^{\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+i\boldsymbol{Q}'\boldsymbol{v})t}\, d\boldsymbol{v}\right) dt \\ &= C_0^\gamma e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N}\int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0}\left(\int_{t=0}^\infty e^{-[\rho-\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+i\boldsymbol{Q}'\boldsymbol{v})]t}\, dt\right) d\boldsymbol{v} \\ &= C_0^\gamma e^{-\boldsymbol{\gamma}'\boldsymbol{y}_0/N}\int \frac{\mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0}}{\rho-\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v} \qquad \text{(B.52)} \\ &= \left(e^{\boldsymbol{q}_1'\boldsymbol{y}_0/N}+\cdots+e^{\boldsymbol{q}_N'\boldsymbol{y}_0/N}\right)^\gamma \int \frac{\mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0}}{\rho-\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v}\,. \qquad \text{(B.53)} \end{aligned}$$

As in the two-asset case, I assume that $\mathrm{Re}\,[\rho-\boldsymbol{c}(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+i\boldsymbol{Q}'\boldsymbol{v})] > 0$; and as in the two-asset case, this follows from the apparently weaker condition that $\rho - \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N) > 0$. The proof follows exactly the same lines as in the two-asset case, and is therefore omitted.

B.5.4 Returns

From (B.52), the price of the $\boldsymbol{\alpha}$-asset is

$$P = (e^{y_{10}} + \cdots + e^{y_{N0}})^{\gamma}\, e^{(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N)'\boldsymbol{y}_0} \int \frac{\mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0}}{\rho - \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v}\,.$$

Introducing the *multinomial coefficient*,

$$\binom{\gamma}{\boldsymbol{m}} \equiv \frac{\gamma!}{m_1! m_2! \cdots m_N!}\,,$$

we can express the price as

$$P = \sum_{\boldsymbol{m}} \binom{\gamma}{\boldsymbol{m}} \int \frac{\mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+\boldsymbol{m}+i\boldsymbol{Q}'\boldsymbol{v})'\boldsymbol{y}_0}}{\rho - \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v}\,.$$

The sum is taken over all N-dimensional vectors $\boldsymbol{m}$ whose entries are nonnegative integers which add up to γ.

Using the result of Appendix B.1.1, it follows that

$$\mathbb{E}dP = \sum_{\boldsymbol{m}} \binom{\gamma}{\boldsymbol{m}} \int \frac{\mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{(\boldsymbol{\alpha}-\boldsymbol{\gamma}/N+\boldsymbol{m}+i\boldsymbol{Q}'\boldsymbol{v})'\boldsymbol{y}_0} \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + \boldsymbol{m} + i\boldsymbol{Q}'\boldsymbol{v})}{\rho - \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v}\, dt\,,$$

and hence

$$\begin{aligned} \mathbb{E}dP/D &= \sum_{\boldsymbol{m}} \binom{\gamma}{\boldsymbol{m}} \int \frac{\mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{(-\boldsymbol{\gamma}/N+\boldsymbol{m}+i\boldsymbol{Q}'\boldsymbol{v})'\boldsymbol{y}_0} \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + \boldsymbol{m} + i\boldsymbol{Q}'\boldsymbol{v})}{\rho - \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v}\, dt \\ &= \sum_{\boldsymbol{m}} \binom{\gamma}{\boldsymbol{m}} e^{m_1 \boldsymbol{q}_1'\boldsymbol{y}_0/N + \cdots + m_N \boldsymbol{q}_N'\boldsymbol{y}_0/N} \int \frac{\mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + \boldsymbol{m} + i\boldsymbol{Q}'\boldsymbol{v})}{\rho - \boldsymbol{c}(\boldsymbol{\alpha} - \boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})}\, d\boldsymbol{v}\, dt\,. \end{aligned}$$

We then get expected capital gains by dividing through by the price-dividend ratio, calculated above. The other component of expected return is the dividend yield, which is the reciprocal of the price-dividend ratio.

B.5.5 *Interest rates*

The price of a time-T zero-coupon bond is

$$B_T = \mathbb{E} e^{-\rho T} \left(\frac{C_T}{C_0} \right)^{-\gamma} .$$

Using the expectation calculated in section B.5.2, we have

$$\begin{aligned} B_T &= e^{-\rho T} C_0^{\gamma}\, \mathbb{E} \frac{1}{\left(e^{y_{10}+\widetilde{y}_{1T}} + \cdots + e^{y_{N0}+\widetilde{y}_{NT}}\right)^{\gamma}} \\ &= e^{-\rho T} C_0^{\gamma} e^{-\boldsymbol{\gamma}' \boldsymbol{y}_0/N} \int \mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} e^{\boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})T}\, d\boldsymbol{v} \\ &= e^{-\rho T} \left(e^{\boldsymbol{q}_1' \boldsymbol{y}_0/N} + \cdots + e^{\boldsymbol{q}_N' \boldsymbol{y}_0/N} \right)^{\gamma} \int \mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} e^{\boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})T}\, d\boldsymbol{v} . \end{aligned}$$

The yield $\mathscr{Y}(T) = -(\log B_T)/T$. Using the above expression,

$$\mathscr{Y}(T) = \rho - \frac{1}{T} \log \left\{ \left(e^{\boldsymbol{q}_1' \boldsymbol{y}_0/N} + \cdots + e^{\boldsymbol{q}_N' \boldsymbol{y}_0/N} \right)^{\gamma} \int \mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} e^{\boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})T}\, d\boldsymbol{v} \right\} .$$

To calculate the riskless rate, rearrange this expression slightly, using (B.47)—

$$\mathscr{Y}(T) = \rho - \frac{1}{T} \log \left\{ 1 + \left(e^{\boldsymbol{q}_1' \boldsymbol{y}_0/N} + \cdots + e^{\boldsymbol{q}_N' \boldsymbol{y}_0/N} \right)^{\gamma} \int \mathscr{G}_{\gamma}^{N}(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} \left(e^{\boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})T} - 1 \right) d\boldsymbol{v} \right\} .$$

Using L'Hôpital's rule, as in the two-asset case, we have

$$
\begin{aligned}
r &= \lim_{T\downarrow 0} \mathscr{Y}(T) \\
&= \rho - \left(e^{\boldsymbol{q}_1'\boldsymbol{y}_0/N} + \cdots + e^{\boldsymbol{q}_N'\boldsymbol{y}_0/N}\right)^{\gamma} \int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} \boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})\, d\boldsymbol{v} \\
&= \left(e^{\boldsymbol{q}_1'\boldsymbol{y}_0/N} + \cdots + e^{\boldsymbol{q}_N'\boldsymbol{y}_0/N}\right)^{\gamma} \int \mathscr{G}_\gamma^N(\boldsymbol{v}) e^{i\boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0} \left[\rho - \boldsymbol{c}(-\boldsymbol{\gamma}/N + i\boldsymbol{Q}'\boldsymbol{v})\right] d\boldsymbol{v}\,.
\end{aligned}
$$

B.5.6 A final change of variables

The expressions so far obtained can be simplified by a final change of variables. Define $\widehat{\boldsymbol{v}} \equiv \boldsymbol{B}\boldsymbol{v}$, where $\boldsymbol{B}$ is the $(N-1)\times(N-1)$ square matrix

$$
\boldsymbol{B} \equiv \begin{pmatrix} N-1 & -1 & \cdots & -1 \\ -1 & N-1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ -1 & \cdots & -1 & N-1 \end{pmatrix}.
$$

With this definition, we have $\widehat{v}_k = Nv_k - v_1 - \cdots - v_{N-1}$ and $\widehat{v}_1 + \cdots + \widehat{v}_{N-1} = v_1 + \cdots + v_{N-1}$. It is simple to verify that

$$
\boldsymbol{B}^{-1} = \frac{1}{N} \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2 \end{pmatrix}.
$$

Using the matrix determinant lemma (Fact B.1 above) it is easy to calculate the Jacobian: $\det \boldsymbol{B}^{-1} = 1/N^{N-2}$, so—since $\boldsymbol{v} = \boldsymbol{B}^{-1}\widehat{\boldsymbol{v}}$—$d\boldsymbol{v}$ is replaced by $d\widehat{\boldsymbol{v}}/N^{N-2}$. Next, $\widehat{\boldsymbol{v}}$ was defined in such a way that $\mathscr{G}_\gamma^N(\boldsymbol{v})$, defined in equation (B.46), is equal to $N^{N-2}\mathscr{F}_\gamma^N(\widehat{\boldsymbol{v}})$, defined in equation (2.31). Finally, noting that $\boldsymbol{B}^{-1}\boldsymbol{Q} = \boldsymbol{U}$ and

$\boldsymbol{u} \equiv \boldsymbol{U}\boldsymbol{y}_0$, as defined in (2.32), we have

$$\begin{aligned} \boldsymbol{Q}'\boldsymbol{v} &= \boldsymbol{Q}'\boldsymbol{B}^{-1}\widehat{\boldsymbol{v}} \\ &= \boldsymbol{U}'\widehat{\boldsymbol{v}}\,, \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{v}'\boldsymbol{Q}\boldsymbol{y}_0 &= \widehat{\boldsymbol{v}}'\boldsymbol{U}\boldsymbol{y}_0 \\ &= \widehat{\boldsymbol{v}}'\boldsymbol{u} \\ &= \boldsymbol{u}'\widehat{\boldsymbol{v}}\,. \end{aligned}$$

Proposition 2.8 follows after making these substitutions throughout the various integrals and dropping hats on the integration variables $\widehat{\boldsymbol{v}}$.

B.6 Some results from complex analysis

This section provides a brief summary of the definitions and results from complex analysis that were invoked in Appendix B.3. Proofs of the results cited will be found in any introductory complex analysis textbook; I have drawn on Priestley (1995).

A complex-valued function f is said to be *holomorphic* in G, which is some subset of the complex plane, if

$$\lim_{h\to 0} \frac{f(z+h) - f(z)}{h}$$

exists for every point z in some open set containing G. Note that the limit must be the same no matter what direction h approaches 0 from: for example, it may be tend to zero along the imaginary axis or along the real axis. Polynomials, convergent power series, the exponential function, sine, and cosine are holomorphic, as are compositions and finite sums and products of these functions. So, for example, the hyperbolic cosine, $\cosh z \equiv (e^z + e^{-z})/2$ is holomorphic.

Result B.1 (Holomorphic iff analytic)**.** *A function f is holomorphic in the open set $\{z \in \mathbb{C} : |z-a| < r\}$ if and only if it is* analytic*—that is, representable by a power series:*

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \,, \qquad \textit{for } z \textit{ such that} \quad |z-a| < r \,.$$

Proof. See Priestley (1995), pp. 20–21 and 69. □

Complex integrals appear throughout Chapter 2. Real integration takes place on subsets of the real line—for example,

$$\int_a^b f(x)\, dx$$

is an integral "along the path from a to b." Complex integration takes place over paths in the complex plane. Since the complex plane is two-dimensional (as opposed to the one dimension of the real line), these paths can be more complicated. For example, an integral might be "around the unit circle defined by $|z| = 1$," or "along the real line from $-R$ to R, then around a semicircular arc lying in the upper half-plane from R back to $-R$."

The integrals which occur in Chapter 2 (for example, (2.10)) feature integrands which are holomorphic everywhere except for at certain *singularities* at which they explode to infinity. (These singularities do not, of course, occur on the path of integration.) It is an amazing—and powerful—fact that such integrals depend on the behavior of the integrands at singularities elsewhere in the complex plane. I now introduce the mathematical apparatus used in Chapter 2 that relates to this fact.

If f is holomorphic in some punctured disc $\{z \in \mathbb{C} : 0 < |z-a| < r\}$ but not at the point a, then a is an *isolated singularity.* (Keep in mind the example $f(z) = 1/z$, which is holomorphic everywhere except for at an isolated singularity at the origin.) In this case, f can be expanded as a unique power series of the form

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n \qquad \text{for } z \text{ such that} \quad 0 < |z-a| < r\,. \tag{B.54}$$

If $c_n = 0$ for all $n < 0$, the point a is a *removable singularity.* (In other words, it is not "really" a singularity at all. Consider the example $f(z) = (\sin z)/z$, which has a removable singularity at $z = 0$. If the function f is redefined slightly by specifying that $f(0) = 1$, then the singularity has been removed.) If there is some positive m such that $c_{-m} \neq 0$ and $c_k = 0$ for all $k < -m$ then the point a is a *pole of order m.*[3]

These concepts are best illustrated with an example. Take the function

$$\mathscr{F}_2(v) = \frac{v}{2\sinh \pi v}\,.$$

Singularities occur whenever $\sinh \pi v = 0$, in other words at $v = 0, \pm i, \pm 2i, \ldots$.[4] However, it is easy to check that the singularity at the origin is removable. (By L'Hôpital's rule, $f(v)$ tends to $1/2\pi$ as v tends to zero.) In fact, the only non-removable singularities are poles of order 1 at $\pm i, \pm 2i, \pm 3i, \ldots$.

A function f which is holomorphic throughout the complex plane, except at poles, is called *meromorphic.* If a meromorphic function f has a pole at a then the *residue* of f at a, written $\operatorname{Res}\{f(z); a\}$, is defined to be the coefficient on the term $(z-a)^{-1}$ in a power series expansion of the form (B.54). With this final piece of notation, I now state the residue theorem.

Result B.2 (The Residue Theorem)**.** *Let Ω denote a closed path of integration which is to be integrated around in an anticlockwise direction. Suppose f is holomorphic*

[3] If there are arbitrarily large m such that $c_{-m} \neq 0$ then the point a is an *isolated essential singularity*, but this case is not relevant for my purposes.

[4] Remember that when z is real, we have $\sinh(iz) = i \sin z$ and $\cosh(iz) = \cos z$.

inside and on Ω*, except for at a finite number of poles at points* $a_1, \ldots, a_m$ *inside* Ω*. Then*

$$\int_\Omega f(z)\,dz = 2\pi i \sum_{j=1}^{m} Res\{f(z); a_j\}$$

Proof. See Priestley (1995), chapter 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.

Andrews, G. E., Askey, R., and R. Roy (1999), *Special Functions*, 1st edition, Cambridge University Press, UK.

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.

Brainard, W. C. and J. Tobin (1968), "Pitfalls in Financial Model Building," *American Economic Review*, 58:2:99–122.

Brainard, W. C. and J. Tobin (1992), "On the Internationalization of Portfolios," *Oxford Economic Papers*, 44:533–565.

Breeden, D. T. (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7:265–296.

Breiman, L. (1960), "Investment Policies for Expanding Businesses Optimal in a Long-Run Sense," *Naval Research Logistics Quarterly*, 7:647–651.

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

Campbell, J. Y. (1991), "A Variance Decomposition for Stock Returns," *Economic Journal*, 101:157–179.

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., Longstaff, F. A., and P. Santa-Clara (2008), "Two Trees," *Review of Financial Studies*, 21:1:347–385.

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.

Cohen, R. B., Polk, C., and T. Vuolteenaho (2003), "The Value Spread," *Journal of Finance*, 58:2:609–641.

Cole, H. L. and M. Obstfeld (1991), "Commodity Trade and International Risk Sharing," *Journal of Monetary Economics*, 28:3–24.

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.

Cox, J. C., Ingersoll, J. E., and S. A. Ross (1985), "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53:2:363–384.

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.

Fama, E. F. and K. R. French (1989), "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25:23–49.

Fama, E. F. and K. R. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33:131–155.

Forbes, K. J. and R. Rigobon (2002), "No Contagion, Only Interdependence: Measuring Stock Market Comovements," *Journal of Finance*, 57:5:2223–2261.

Gabaix, X. (2008), "Variable Rare Disasters: A Tractable Theory of Ten Puzzles in Macro-Finance," *American Economic Review*, forthcoming.

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.

Kakutani, S. (1948), "On Equivalence of Infinite Product Measures," *Annals of Mathematics*, 49:214–224.

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

Karatzas, I. and S. E. Shreve (1991), *Brownian Motion and Stochastic Calculus*, Springer, New York, NY.

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.

Latané, H. A. (1959), "Criteria for Choice Among Risky Ventures," *Journal of Political Economy*, 67:2:144–155.

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. (1978), "Asset Prices in an Exchange Economy," *Econometrica*, 46:6:1429–1445.

Lucas, R. E. (1982), "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics*, 10:335–359.

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

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

Lustig, H., Van Nieuwerburgh, S. and A. Verdelhan (2008), "The Wealth-Consumption Ratio," working paper.

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

Markowitz, H. M. (1976), "Investment for the Long Run: New Evidence for an Old Rule," *Journal of Finance*, 31:5:1273–1286.

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

Menzly, L., Santos, T., and P. Veronesi (2004), "Understanding Predictability," *Journal of Political Economy*, 112:1:1–47.

Merton, R. C. (1973), "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41:5:867–887.

Neveu, J. (1975), *Discrete-Parameter Martingales*, trans. T. P. Speed, North-Holland, Amsterdam.

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

Pavlova, A. and R. Rigobon (2007), "Asset Prices and Exchange Rates," *Review of Financial Studies*, 20:4:1139–1181.

Priestley, H. A. (1995), *Introduction to Complex Analysis*, Oxford University Press, Oxford, UK.

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

Ritter, G. (1979), "On Kakutani's Dichotomy Theorem for Infinite Products of Not Necessarily Independent Functions," *Mathematische Annalen*, 239:35–53.

Rogers, L. C. G., and D. Williams (2000), *Diffusions, Markov Processes, and Martingales, Volume 1*, Cambridge University Press, Cambridge, UK.

Samuelson, P. A. (1971), "The 'Fallacy' of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling," *Proceedings of the National Academy of Sciences of the United States of America*, 68:10:2493–2496.

Santos, T. and P. Veronesi (2006), "Labor Income and Predictable Stock Returns," *Review of Financial Studies*, 19:1:1–44.

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.

Shiller, R. J. (1981), "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, 71:421–436.

Shiller, R. J. (1989), "Comovements in Stock Prices and Comovements in Dividends," *Journal of Finance*, 44:3:719–729.

Slater, L. J. (1966), *Generalized Hypergeometric Functions*, Cambridge University Press, Cambridge, UK.

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. (2004), "Discounting a Distant Future Whose Technology is Unknown," working paper.

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

Williams, D. (1995), *Probability with Martingales*, Cambridge University Press, Cambridge, UK.