DALLASFED Working Paper

Disastrous Disappointments: Asset-Pricing with Disaster Risk And Disappointment Aversion

Jim Dolmas

Federal Reserve Bank of Dallas
Research Department
Working Paper 1309
October 2013

Disastrous Disappointments: Asset-Pricing with Disaster Risk and Disappointment Aversion

Jim Dolmas*

October 16, 2013

Abstract

In this paper, I combine disappointment aversion, as employed by Routledge and Zin [28] and Campanale, Castro and Clementi [9], with rare disasters in the spirit of Rietz [27], Barro [4], Gourio [16], Gabaix [15] and others. I find that, when the model's representative agent is endowed with an empirically plausible degree of disappointment aversion, a rare disaster model can produce moments of asset returns that match the data reasonably well, using disaster probabilities and disaster sizes much smaller than have been employed previously in the literature.

This is good news. Quantifying the disaster risk faced by any one country is inherently difficult with limited time series data. And, it is open to debate whether the disaster risk relevant to, say, US investors is well-approximated by the sizable risks found by Barro [4] and co-authors [6, 7, 26] in cross-country data. On the other hand, we have evidence (see [30], [8], or [11]) that individuals tend to over-weight bad or disappointing outcomes, relative to the outcomes' weights under expected utility. Recognizing aversion to disappointment means that disaster risks need not be nearly as large as suggested by the cross-country evidence for a rare disaster model to produce average equity premia and risk-free rates that match the data.

I illustrate the interaction between disaster risk and disappointment aversion both analytically and in the context of a simple Rietz-like model of asset-pricing with rare disasters. I then analyze a richer model, in the spirit of Barro [4], with a distribution of disaster sizes, Epstein-Zin preferences, and partial default (in the event of a disaster) on the economy's 'risk-free' asset. For small elasticities of intertemporal substitution, the model is able

*Federal Reserve Bank of Dallas, 2200 North Pearl Street, Dallas, TX 75201. E-mail: jim@jimdolmas.net. URL: http://www.jimdolmas.net/economics. I would like to thank, without implicating, Karen Lewis, and participants at the 2013 Conference on Computing in Economics and Finance, Vancouver, and the 2013 Econometric Society European Meeting, Göteborg, for comments on earlier versions of this paper. Typeset in LaTeX, using TeXstudio 2.6

to match almost exactly the means and standard deviations of the equity return and risk-free rate, for disaster risks one-half or one-fourth the estimated sizes from Barro [4]. For larger elasticities of intertemporal substitution, the model's fit is less satisfactory, though it fails in a direction not often viewed as problematic—it under-predicts the volatility of the risk-free rate. Even so, apart from that failing, the results are broadly similar to those obtained by Gourio [16], but with disaster risks one-half or one-fourth as large.

Keywords: Rare disasters, disappointment aversion, asset pricing
JEL: E43, E44, G12

1 Introduction

In the nearly 30 years since Mehra and Prescott [23] and others[1] first highlighted the problematic relationship between the data on aggregate consumption growth and asset returns, the literature on equilibrium asset-pricing has gone from having essentially no good models to rationalize those data to having several more or less successful candidates to choose from. Within the class of representative agent models, responses to the equity premium (and related) puzzles have tended to fall into two broad categories: models that modify the agent's preferences—including habit formation [10], disappointment aversion [28, 9], or state-dependent preferences [25]—or models that modify the aggregate consumption process—moving from *i.i.d.* or near-*i.i.d.* log-normal models of consumption growth to models that feature disaster risk [27, 4, 16, 15] or persistent movements in the conditional mean of consumption growth [3].[2]

In this paper, I focus on—and combine—two of these candidates: disappointment aversion and disaster risk. Routledge and Zin [28] show that preferences displaying disappointment aversion—generalized to allow some flexibility over which outcomes are disappointing—are capable of closely matching both the first and second moments of the equity return and the risk-free rate in a model using Mehra and Prescott's aggregate consumption process. On disaster risk, Gabaix [15] argues that rare disasters are capable of resolving a number of asset-pricing anomalies. Gourio [16], for example, writes down a model with Epstein-Zin preferences and disaster risk calibrated according to Barro's [4] estimates of the size and frequency of disasters. When he allows for time-variation in the probability of disaster, his model can match reasonably well first and second moments of returns data, as well as some aspects of time-series variation in returns and price-dividend ratios.

Why combine disappointment aversion and disaster risk? Disaster risk is inherently difficult to quantify with limited time series data—even a hundred years' worth of observations is insufficient to say with any precision whether

[1] As Cochrane [12] notes, the essence of Mehra and Prescott's equity premium puzzle was also present in Shiller [29], Grossman and Shiller [17], and Hansen and Singleton [20].

[2] Another class of responses dispense with the representative agent assumption altogether; see, for example [14], [19], or [31].

disasters occur a couple times every hundred years or once every couple hundred years. The use of cross-country evidence—pioneered by Barro in [4] and refined by Barro and co-authors in various papers—provides more observations, but is applicable only if we accept that all countries draw disasters from a common distribution. It is open to debate whether the commonly used estimates from Barro's cross-country sample—a sizable 1.7% per annum chance of an average 29% real output decline—are a good representation of the disaster risk faced by an economy such as the US. If the relevant disaster risk is only a half or a fourth as large, do rare disasters then lose their potential to explain US data on consumption and asset returns? This is where disappointment aversion figures in.

There is abundant experimental and field evidence indicating that individuals are averse to loss or disappointment, or at the very least over-weight adverse outcomes relative to what expected utility would suggest (see, for example, the data analyzed by Camerer and Ho [8], the experimental results in Choi, Fisman, Gale and Kariv [11], or the evidence surveyed by Starmer [30, pp. 357–366]). To be sure, the calibration of preferences from laboratory experiments or field data is fraught with its own set of measurement issues—for example, whether to treat laboratory choices in isolation from a subject's wealth or background risk—but it nevertheless seems plausible to endow our agents with preferences embodying some degree of disappointment aversion.

Once we allow agents to have some aversion to disappointment, it's possible for the rare disaster model to generate realistic asset returns—returns with first and second moments that come close to matching what we see in the data—with disasters that are much rarer or much smaller than those typically assumed in the literature.

An overview of the organization of the paper—and a synopsis of the main results—is as follows. The next section, 2, offers some background on asset pricing with rare disasters and disappointment aversion, and establishes some notation. Given that models of disaster risk are more common in the literature, I devote relatively more time in section 2 to discussing disappointment aversion—including the intuition for its impact on the utility value of rare events, considerations as to what degrees of disappointment aversion could be considered reasonable, and the form of the stochastic discount factor under disappointment aversion.

Section 3 then combines disappointment aversion and disaster risk in a simple variant of Rietz's model—a model with a single disaster state, a completely risk-free bond, and preferences that (in the absence of disappointment aversion) would collapse to the standard CRRA time-separable form. The exercize is mainly intended to highlight the trade-off between the probability of disaster (or its size) and the strength of aversion to disappointment. In that model, I show that one can easily substitute a reasonable amount of disappointment aversion for a reduction in disaster risk, while still maintaining the model's asset return implications. The first and second moments of returns implied, for example, by a 2.5% chance (per year) of a 25% consumption decline are roughly the same as those implied by a probability of disaster one-fourth as large (0.625%) if the representative agent's preferences embody an amount of disappointment aversion well within the bounds of the experimental evidence.

While the simple model of section 3 illustrates the mechanisms at work, the match between the model-implied moments and the data is less than ideal. In section 4, which contains the paper's main results, I examine a much richer model. To begin with, I relax the link between the agent's elasticity of intertemporal substitution and coefficient of relative risk aversion, taking full advantage of the Epstein-Zin aggregator that describes the evolution of the agent's lifetime utility. Following Barro [4], I also assume a distribution of possible disasters, and calibrate their relative sizes and frequencies using Barro's data. As in Barro [4] and Gourio [16], I also allow for partial default on the model's 'risk-free' asset in the event of disaster, addressing one of Mehra and Prescott's [24] criticisms of Rietz's proposed resolution to their puzzle.

The quality of the richer model's fit depends importantly on the value assumed for the agent's elasticity of intertemporal substitution. If we assume a small value for the intertemporal substitution elasticity, the model is capable of matching almost exactly the means and standard deviations of the risk-free rate and equity return. This is true for disaster probabilities and average disaster sizes one-half or one-fourth the values from Barro's [4] sample. The model accomplishes this without assuming either an excessive degree of disappointment aversion, an extreme value for the agent's relative risk aversion (curvature) parameter, or a utility discount factor greater than one.

For larger values of the elasticity of intertemporal substitution—I look at values of 1 and 1.5—the results are less satisfactory. However, the main failing of the model—the dimension along which our results depart most significantly from the data—is a failing not normally viewed as problematic: the model produces a risk-free rate that is not sufficiently volatile.[3]

Also, with elasticities of intertemporal substitution greater than one, reductions in disaster risk—in particular reductions in average disaster size—can only be compensated for by degrees of disappointment aversion at the upper end of the range consistent with the experimental evidence.

Nevertheless, the model can still match the first moments of returns quite well, and apart from the too-low volatility of the risk-free rate, the model moments—across varying degrees of disaster risk—are broadly similar to those obtained by Gourio [16] using the disaster probability and average (effective) disaster size consistent with the estimates of Barro [4].

I offer some conclusions and directions for further work in section 5.

Some details of the solution method are provided in an Appendix.

[3]*Cf.* Campbell and Cochrane [10], who set out to design a model with a constant risk-free rate. Also, any rare disaster model in which normal times are described by *i.i.d.* lognormal consumption growth, and in which the disaster probability is constant, will produce a constant risk-free rate in non-disaster samples.

2 Background

In this section, I briefly review the two strands of the literature that come together in this paper. I devote a bit more time to disappointment aversion, as it is somewhat less well known and more difficult to convey without at least writing down some pieces of a model. In addition to providing background and establishing notation, I also illustrate the potential for disappointment aversion to substitute for disaster risk (section 2.3.1) and address the question of what constitutes a plausible amount of disappointment aversion (section 2.3.2).

2.1 Disaster risk

Rare consumption disasters were among the earliest responses to Mehra and Prescott's puzzle. Rietz [27] added a disaster state to Mehra and Prescott's Markov chain, and showed calibrations of the model (varying the disaster size and probability, and the representative agent's risk aversion and discounting parameters) that came close to matching the mean returns on equity and a risk-free asset, as calculated by Mehra and Prescott.

Disaster risk provides a potential resolution to Mehra and Prescott's puzzle by providing a channel for holding down the risk-free rate as risk aversion is increased. For example, in an environment where consumption growth is log-normal and a representative agent has standard, additively separable, power utility preferences, we know that the log risk-free rate obeys

$$R^f = -\log(\beta) + \alpha\gamma - \frac{\alpha^2\sigma^2}{2} \tag{1}$$

where β is the utility discount factor, α is the coefficient of relative risk aversion, and γ and σ are the mean and standard deviation of log consumption growth. As is well-known, for typical estimates of γ and σ, the first-order term $\alpha\gamma$ in (1) dominates as α increases, up to extremely large values of α. A high value of α, necessary to get a non-negligible equity risk premium, thus produces a counterfactually high risk-free rate.

Let a disaster add $\log(1-b) < 0$ to $\log(c_{t+1}/c_t)$, and occur with probability p (independently of the log-normal component). Then, the log risk-free rate becomes

$$R^f = -\log(\beta) + \alpha\gamma - \frac{\alpha^2\sigma^2}{2} - \log(1 - p + p(1-b)^{-\alpha}). \tag{2}$$

For typical values of γ and σ, and for small values of p, (2) can generate low values of R^f even for large values of α.

Mehra and Prescott [24] responded to Rietz's resolution with a number of criticisms—concerning Rietz's calibrations of the size and frequency of disasters; the values of the risk aversion parameter and the utility discount factor needed to resolve the puzzle; and the assumption that the model's risk-free asset would remain risk-free in the event of a disaster—and for a while work on disaster risk as a potential resolution to asset-pricing puzzles was scant.[4]

Following the work of Barro [4], however, there has been renewed interest in disasters. Barro looked over time and across countries to provide new evidence on the size and frequency of disasters. Barro [4, 5], Gabaix [15], and Gourio [16] also made some modeling advances, relative to Rietz, including the use of Epstein-Zin preferences, and the possibility that the model's 'risk-free' asset is subject to default risk in disaster states.[5]

Estimates of the frequency and size of disasters were further refined by Barro and co-authors: Nakamura, Steinsson, Barro, and Ursúa [26], Barro and Ursúa [7], and Barro and Jin [6].

Barro [4] identified disasters with declines in per capita real GDP of at least 15 percent; the frequency of these events—in a sample of 35 countries over 100 years—was 1.7 percent per year. The average size of a disaster in Barro's sample is a 29 percent decline, though as Barro notes, "[b]ecause of diminishing marginal utility of consumption, larger contractions count more than smaller ones; hence, the effective average value of b [the disaster size] exceeds 0.29." In the simulations of his model, Barro uses the frequency distribution of disasters (in contrast to the single disaster state in Rietz's original formulation). Gourio [16] replaces Barro's distribution of disasters with a single disaster state that occurs with 1.7 percent probability and entails a 43 percent consumption decline.[6]

As noted in the introduction—and readily conceded by Barro and co-authors [26, p. 37]—looking at the experience of multiple countries provides more data with which to estimate the parameters governing disaster risk, but also assumes that all countries draw disaster events from a common distribution.

2.2 Disappointment aversion

Similar to the standard Epstein-Zin formulation, preferences embodying disappointment aversion reduce random lifetime utility from tomorrow onward to a certainty equivalent value and combine that certainty equivalent with current consumption to obtain lifetime utility as of today. Aversion to disappointment is embodied in the certainty equivalent operator, where disappointing outcomes are given greater weight relative to non-disappointing outcomes.

Which outcomes are disappointing? In Gul's [18] original formulation, the certainty equivalent obeys a consistency requirement—the threshold for dis-

[4] Kocherlakota's [21] survey on the equity premium puzzle in the *Journal of Economic Literature*, for example, published 11 years after Mehra and Prescott's original article, mentions disasters only passingly, in a single footnote.

[5] Gourio also added time-variation in the probability of disaster. Using average size and frequency numbers from Barro, he gives calibrations of the model that come close not only to first and second moments of asset returns, but also to regression coefficients—estimated by Cochrane [13]—relating asset returns to price-dividend ratios. More recently, Wachter [32] also examines a model with time-varying disaster risk.

[6] This is also roughly the value that would be needed in Gabaix's [15] model if his distribution of disaster sizes were collapsed to a single state.

appointment is the certainty equivalent value itself (so the certainty equivalent is defined only implicitly).[7] In a very rough sense, a disappointment-averse agent takes expectations overweighting outcomes that would fall below expectations.

Routledge and Zin [28] showed how to adapt Gul's notion of disappointment aversion to an asset-pricing framework, while at the same time generalizing it to allow some control over which outcomes are disappointing. Depending on a parameter, only sufficiently bad outcomes may disappoint, and the set of disappointing outcomes may vary with the state of the economy. They dubbed their framework 'generalized disappointment aversion.' Their model—like all those I will consider here—is a representative agent exchange economy in the sense of Lucas [22].

Campanale, Castro, and Clementi [9] hew more closely to Gul's original conception, but also incorporate a form of production and physical capital accumulation.

The preferences I will employ in this paper are similar to those used by Campanale *et al.* While the performance of the model here could no doubt be marginally improved by allowing for generalized disappointment aversion *à la* Rouledge and Zin, the model moments I will report in section 4 are already quite close to the data. Generalized disappointment aversion would simply complicate the exposition without much offsetting gain in terms of the model's performance.

2.3 Some formalism

Our representative agent's intertemporal preferences have the Epstein-Zin form

$$U_t = [(1-\beta)c_t^{1-\rho} + \beta\mu_t(U_{t+1})^{1-\rho}]^{1/(1-\rho)} \tag{3}$$

for $\rho \geq 0$, $\rho \neq 1$.[8] The agent's elasticity of intertemporal substitution (over deterministic consumption paths) is $1/\rho$, and $\beta \in (0,1)$ is the utility discount factor.

Disappointment aversion is captured in the conditional certainty equivalent operator $\mu_t(\,\cdot\,)$. For the moment, set aside the source of the conditionality in μ_t; eventually, this will come from the model's Markov chain structure for consumption growth. For now, though, I will drop the subscript t, to avoid confusion.

Let z be a discrete random variable taking the value z_i with probability p_i (given the state at t). The certainty equivalent of z is defined implicitly by

$$f(\mu(z)) = \sum_i p_i f(z_i) - \theta \sum_i p_i [f(\mu(z)) - f(z_i)] I\left(z_i < \mu(z)\right). \tag{4}$$

In (4), $I(\,\cdot\,)$ is an indicator function, taking the value one when $z_i < \mu(z)$—that is, if z_i is disappointing—and zero otherwise. $\theta \geq 0$ is a parameter. The

[7] Thus, these risk preferences are in the Chew-Dekel class; see Backus, Routledge, and Zin [2].

[8] For the case of $\rho = 1$, we set $U_t = c_t^{1-\beta}\mu_t(U_{t+1})^{\beta}$.

function f has the form

$$f(x) = \frac{x^{1-\alpha}}{1-\alpha} \tag{5}$$

for $\alpha \geq 0$, $\alpha \neq 1$.[9]

With f taking the form in (5), the certainty equivalent can also be written (still only implicitly) as

$$\mu(z)^{1-\alpha} = \frac{\sum_i p_i z_i^{1-\alpha}[1 + \theta I\,(z_i < \mu(z))]}{1 + \theta \sum_i p_i I\,(z_i < \mu(z))} \tag{6}$$

Note that the certainty equivalent operator is positively linearly homogeneous: $\mu(kz) = k\mu(z)$ for any random variable z and any real number $k > 0$. This feature will be useful for the transformation of lifetime utility I employ in solving the model.[10]

The parameter θ most directly influences the strength of disappointment aversion—higher values of θ increase the weight that μ places on disappointing outcomes.[11] When $\theta = 0$, there is no disappointment aversion, and preferences revert to the standard Epstein-Zin form. If also $\alpha = \rho$, time-additive expected utility obtains.[12]

2.3.1 Disappointment aversion and low-probability events

To see how disappointment aversion might interact with rare disasters, imagine that in some state today, the only disappointing state tomorrow is the disaster state. For the moment, I am going to be somewhat vague about what constitutes a disaster; it's enough to think of a disaster as a low-probability state that corresponds to the lowest realization of z in $\mu(z)$.

Let $i = 1$ be the disaster state. Then, from (6) we have

$$\mu(z)^{1-\alpha} = \frac{1+\theta}{1+\theta p_1} p_1 z_1^{1-\alpha} + \frac{1}{1+\theta p_1} \sum_{i \neq 1} p_i z_i^{1-\alpha}$$

[9]For $\alpha = 1$, we take $f(x)$ to be $\log(x)$.

[10]See equation (9) below.

[11]Note, though, that changing θ can change the set of disappointing outcomes as well as the strength of disappointment aversion.

[12]Routledge and Zin's generalized disappointment aversion introduces another parameter, δ, in the implicit definition of the certainty equivalent operator. For $\delta \leq 1$, the expression (6) becomes:

$$\mu(z)^{1-\alpha} = \frac{\sum_i p_i z_i^{1-\alpha}[1 + \theta I\,(z_i < \delta\mu(z))]}{1 + \theta\delta^{1-\alpha} \sum_i p_i I\,(z_i < \delta\mu(z))}.$$

Varying δ allows some control over which states are disappointing. Consider the case of two outcomes, $z_1 < z_2$. If $\delta = 1$, it is always the case that $z_1 < \mu(z) < z_2$—the worse of the two outcomes is always disappointing. For $\delta < 1$—assuming $\alpha \geq 0$—it's possible for neither outcome to be disappointing. Routledge and Zin exploit this property within a two-state, Mehra-Prescott environment, calibrating preferences so that no outcome is disappointing if the current state is the high-growth state, while the low-growth state disappoints if today's state is the low-growth state.

This is precisely the form of a standard, Epstein-Zin certainty equivalent, with distorted probabilities $\hat{p}$,

$$\hat{p}_1 = \frac{p_1 + \theta p_1}{1 + \theta p_1} \tag{7}$$

$$\hat{p}_j = \frac{p_j}{1 + \theta p_1} \quad (j > 1) \tag{8}$$

Assuming state 1 remains the only disappointing state as we vary p_1 and θ—at least within some range of values—there is a clear trade-off between the strength of aversion to disappointment and the probability of disaster: we will obtain the same certainty equivalent value with, for example, a one percent disaster probability ($p_1 = .01$) and some disappointment aversion ($\theta = 1$) as we would with a two percent probability ($p_1 = .02$) and no disappointment aversion ($\theta = 0$).

The actual workings of the model(s) below are more complex than this simple example; in particular, the disaster states are not always the only disappointing states, though they are always disappointing. Nevertheless, disappointment aversion succeeds in assigning more weight to the disastrous outcomes (both in the certainty equivalent and the resulting stochastic discount factor) than would obtain in the case with no disappointment aversion.

2.3.2 How much disappointment aversion is reasonable?

I have claimed above—and will show below—that one can get reasonable moments of asset returns from models incorporating disaster probabilities (and disaster sizes) much smaller than commonly assumed in the literature when agents are endowed with a reasonable amount of disappointment aversion. How much disappointment aversion is reasonable? I use experimental data to discipline the range of θ values I consider.

Camerer and Ho [8] and Choi *et al.* [11] both give estimates of θ based on experimental results. Camerer and Ho synthesize the data from nine separate experimental studies of choice under uncertainty. Treating the observations as generated by the choices of a representative agent, they estimate θ to be around 3.[13] Choi *et al.* perform an experiment in which subjects make a large number of portfolio choices at different state prices, and estimate θ's for each individual. They find considerable heterogeneity in the individual θ's. The estimated values have a mean of 0.315 across subjects, with a standard deviation of 0.493.

Based on these estimates, I restrict θ to lie between 0 and 3 in all the numerical simulations below. Only for the model of section 4, when we assume a high elasticity of intertemporal substitution (section 4.2.3), does the constraint $\theta \leq 3$ prove binding. More common values for θ—which I choose based on model fit, subject to $\theta \in [0, 3]$—are between 1 and 2.

[13] They also estimate θ's separately for each study; those estimates are mostly in the range of 1 to 10.

How does the assumption of $\theta \in [0, 3]$ compare with calibrations of this parameter in other applications of disappointment aversion to asset-pricing? The range I am allowing encompasses larger values than those used by Campanale *et al.* [9], though much smaller values than those considered by Routledge and Zin [28].

Basing their calibration only on the results in Choi *et al.*, Campanale *et al.* consider values of θ between 0 and 0.4. Routledge and Zin use θ values ranging from 9 (the most frequent choice) to 24 (in one case) [28, Tables I, II]. Note, though, comparisons to Routledge and Zin are complicated by the fact that they are assuming a generalized version of disappointment aversion, with an additional parameter controlling the set of disappointing states. What the experimental evidence implies for a plausible choice of θ in that framework is unclear.

One can get some sense of where, effectively, my allowance of any $\theta \in [0, 3]$ lies relative to the choices in these other papers by considering again the 'distorted probabilities' example of section 2.3.1. A 1% probability of the rare state and $\theta = 3$ lead to a distorted probability—from equation (7)—of 3.9%. The same 1% probability with $\theta = 0.4$ implies a distorted probability of the rare event equal to 1.4%. Thus, even though my upper bound on θ is 7.5 times the upper bound in Campanale *et al.*, for a very low probability event, the effect is not quite 3 times as large. For $\theta = 9$, the distorted probability of the rare event is 9.2%. Compared with Routledge and Zin's $\theta = 9$, then, the impact of setting $\theta = 3$ is about 4/10 as large.

2.3.3 Disappointment aversion and asset-pricing

To see the potential channels through which disappointment aversion affects asset-pricing, it's enough to consider the stochastic discount factor implied by preferences of the form given by (3) and (4) or (6). This section does that; readers familiar with either [9] or [28] may wish to skip ahead to section 3.

Imagine an environment in which aggregate log consumption growth,

$$\log(c_{t+1}/c_t) \equiv x_{t+1},$$

follows a Markov chain $\{x_1, x_2, \ldots, x_n; P\}$, with $P_{i,j} = \Pr\{x_{t+1} = x_j : x_t = x_i\}$. Let ϕ_t denote a representative agent's lifetime utility from date t onward, scaled by consumption c_t—*i.e.*, $\phi_t = v_t/c_t$, where v_t is the (equilibrium) value of lifetime utility. ϕ_t, which will enter into the model's stochastic discount factor (SDF), obeys the Bellman-like equation

$$\phi_t = [1 - \beta + \beta\mu_t(\phi_{t+1}e^{x_{t+1}})^{1-\rho}]^{1/(1-\rho)}, \tag{9}$$

which we obtain from evaluating the utility process (3) at the equilibrium consumption stream and dividing both sides by the level of current consumption.

I'll focus on environments in which assets are priced by a representative agent's stochastic discount factor. If q_t is the price of a claim that pays a sure F

units of consumption at date $t+1$, then

$$q_t = \mathsf{E}_t(m_{t+1}F)$$

where m_{t+1} is the stochastic discount factor. If p_t^c denotes the price of an equity claim to the aggregate consumption stream, then

$$p_t^c = \mathsf{E}_t(m_{t+1}(p_{t+1}^c + c_{t+1}))$$

and so forth, in similar fashion, for other types of claims.

Let $\mu_i(v)$ denote the certainty equivalent of v, conditional on state i, and let

$$\begin{aligned} I_{i,j} &\equiv I[v_j < \mu_i(v)] \\ &= I[\phi_j e^{x_j} < \mu_i(\phi e^x)] \end{aligned} \tag{10}$$

The stochastic discount factor under disappointment aversion can then be written as

$$m_{i,j} = \beta e^{-\rho x_j} \left(\frac{\phi_j e^{x_j}}{\mu_i(\phi e^x)} \right)^{\rho-\alpha} \frac{1+\theta I_{i,j}}{1+\theta \sum_j P_{i,j} I_{i,j}}, \tag{11}$$

which is similar to the version given by Campanale *et al.* [9].[14] If $\theta = 0$, we're back to a standard Epstein-Zin SDF,

$$m_{i,j} = \beta e^{-\rho x_j} \left(\frac{\phi_j e^{x_j}}{\mu_i(\phi e^x)} \right)^{\rho-\alpha}.$$

If also $\alpha = \rho$—*i.e.*, expected utility—we obtain Mehra and Prescott's (and Rietz's) SDF,

$$m_j = \beta e^{-\rho x_j},$$

which is independent of the current state.

With this background in mind, I now proceed to flesh out one version of a model incorporating both disappointment aversion and disasters, similar to the model of Rietz. This simple setting will prove a useful laboratory for illustrating the trade-off—in terms of asset return implications—between risk of disaster and aversion to disappointment.

3 Illustrating the mechanics: A Rietz-like example

In this section I consider a simple three-state Markov chain model in the spirit of Rietz. For the calculations here, I assume $\alpha = \rho$, so that, absent disappointment aversion (if $\theta = 0$), the representative agent's preferences are of the standard time- and state-separable CRRA form.

[14]Routledge and Zin follow the equivalent method of defining the stochastic discount factor in terms of the return on the representative agent's wealth.

As in Rietz, my starting point is Mehra and Prescott's two-state Markov chain for consumption growth. Recall that Mehra and Prescott specified gross consumption growth $c_{t+1}/c_t = e^{x_{t+1}} \equiv y_{t+1}$ as

$$y_{t+1} \in \begin{bmatrix} y_L \\ y_H \end{bmatrix} = \begin{bmatrix} 0.982 \\ 1.054 \end{bmatrix} \tag{12}$$

with the transition matrix

$$Q = \begin{bmatrix} Q_{LL} & Q_{LH} \\ Q_{HL} & Q_{HH} \end{bmatrix} = \begin{bmatrix} 0.43 & 0.57 \\ 0.57 & 0.43 \end{bmatrix}. \tag{13}$$

Gross consumption growth has an unconditional mean of 1.018 and a standard deviation of 0.036. Mehra and Prescott's transition matrix implies a slight negative autocorrelation in consumption (-0.14).

To this, I add a 'disaster' outcome y_D and modify the transition matrix as follows. As in Rietz, I assume that if the state today is either L or H, then the D state occurs with probability p. If today's state is the disaster state, growth returns next period to $\{y_L, y_H\}$ according to the long-run probabilities associated with Q—*i.e.*, with probabilities $\{1/2, 1/2\}$. There is zero probability of remaining in the disaster state.

The modified Markov chain is thus given by the set of states

$$y_{t+1} \in \begin{bmatrix} y_D \\ y_L \\ y_H \end{bmatrix} = \begin{bmatrix} y_D \\ 0.982 \\ 1.054 \end{bmatrix} \tag{14}$$

and the transition matrix

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ p & (1-p)Q_{LL} & (1-p)Q_{LH} \\ p & (1-p)Q_{HL} & (1-p)Q_{HH} \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ p & (1-p)0.43 & (1-p)0.57 \\ p & (1-p)0.57 & (1-p)0.43 \end{bmatrix}. \tag{15}$$

The long-run (invariant) probabilities associated with P are

$$\pi^* = \left[\frac{p}{1+p}, \frac{0.5}{1+p}, \frac{0.5}{1+p} \right]. \tag{16}$$

Note that I take the Mehra-Prescott Markov chain as a description of consumption growth in 'normal,' non-disaster states. This is in contrast to Rietz's approach. Recall that, for each given disaster size and probability, Rietz recalibrated y_L, y_H, and Q so that the resulting three-state Markov chain matched Mehra and Prescott's empirical estimates of the mean, standard deviation, and autocorrelation of consumption growth.

In the calculations I make below, I focus on moments of asset returns conditional on being in the non-disaster states. Keeping the consumption process in non-disaster states constant as I vary the probability of disaster makes for a cleaner interpretation of the results.

I consider claims to a one-period riskless asset, paying one unit of consumption next period, and two forms of 'equity.' One is the standard claim to aggregate consumption and the other, meant to represent a dividend claim, is a claim to a process whose log growth rate is a multiple of log consumption growth. Letting y^d_{t+1} denote gross dividend growth from t to $t+1$,

$$y^d_{t+1} = (y_{t+1})^{\lambda}, \tag{17}$$

where $\lambda \geq 1$ is the "leverage" parameter. This formulation is by now quite standard.[15]

If w^c and w^d denote the price-dividend ratios for the two equity-like claims, then the key pricing equations are

$$w^c_t = \mathsf{E}_t \left[m_{t+1} y_{t+1} \left(1 + w^c_{t+1} \right) \right] \tag{18a}$$

$$w^d_t = \mathsf{E}_t \left[m_{t+1} y^d_{t+1} \left(1 + w^d_{t+1} \right) \right], \tag{18b}$$

where m_{t+1} is the model's stochastic discount factor. The price of the one-period riskless claim to one unit of consumption, q_t, of course obeys

$$q_t = \mathsf{E}_t \left[m_{t+1} \right]. \tag{19}$$

Gross asset returns are defined in the standard way. For Markov states i today and j tomorrow:

$$R^f_i = \frac{1}{q_i} \qquad \text{(Risk-free rate)}$$

$$R^c_{i,j} = \frac{y_j (w^c_j + 1)}{w^c_i} \qquad \text{(Consumption claim return)}$$

$$R^d_{i,j} = \frac{y^d_j (w^d_j + 1)}{w^d_i} \qquad \text{(Dividend claim return)}$$

Of the consumption and dividend claims, in what follows I will mostly focus on the dividend claim, and I will refer to its return as the 'equity return.'

For a given stochastic discount factor, the model can be solved in standard fashion. Equations (18a), (18b), and (19) are solved for the price-dividend ratios w^c and w^d, and the price q. Under the Markov chain assumption, these are vectors in R^3, and solving for the price-dividend ratios involves a simple matrix inversion. Deriving the model's stochastic discount factor is complicated only slightly by the presence of disappointment aversion—evaluating certainty equivalents entails an additional fixed point calculation. I give details of the solution method in the Appendix.

[15] The formulation originates with Abel [1] and is employed, notably, in Bansal and Yaron [3], as well as in Gourio [16].

3.1 The trade-off between disaster probability and strength of disappointment aversion

In this section, I'm primarily interested in the trade-off between the strength of disappointment aversion and the probability of disaster. In the computations that follow, I will vary both p, the probability of disaster, and θ, the strength of aversion to disappointment.

I fix the size of the disaster at a 25% drop in consumption, or $y_D = 0.75$. The leverage parameter λ is set to 3 throughout, a typical value.[16] In setting the other preference parameters, I focus on the case of $\alpha = \rho$—absent disappointment aversion, these preferences collapse to the constant relative risk aversion form employed by Rietz. I set $\alpha = 5$ (the upper end of the range Mehra and Prescott considered admissible), and I set the utility discount factor $\beta = 0.97$, another standard choice. Table 1 summarizes these basic parameter settings.

α	ρ^{-1}	β	λ	y_D
5	0.20	0.97	3	0.75

Table 1: Basic parameter settings for Rietz-like model. These parameter values are used throughout this section. Other parameters of the consumption process are as given in equations (14) and (15).

As a reference point, Table 2 shows the first and second moments of asset returns from the model without disappointment aversion, for disaster probabilities of 0 (no disasters) and 2.5%. These are moments conditional on being outside the disaster state—*i.e.*, the average behavior of the model in the non-disaster states. The data in the last row, for comparison, are from Gourio [16], based on the same 1926–2004 sample in Cochrane [13]. The results in the table simply confirm what we know from Mehra-Prescott and Rietz—the shortcomings of the time-additive, expected utility specification, in particular its too-high risk-free rate, can be remedied by introducing a small probability of a severe decline in consumption.

Analogous results for the effect of disappointment aversion alone—setting $p = 0$—are shown in Table 3. Given the other parameters, the value of $\theta = 2.44$ reported in the table is simply the value that minimizes the (weighted) Euclidean distance between the four model moments and the data.[17] For that value of θ, the disappointing states are the disaster and low growth states, regardless of the state today.

[16] This is the value used, for example, by Bansal and Yaron [3]. For technical reasons, Gourio [16] uses a value of 2.5, though he describes 3 as the 'standard' value.

[17] That is to say, the measure of fit is a weighted sum of squared errors. Any choice of weights is somewhat arbitrary, but generally speaking, being off by a percentage point on the average equity return is obviously very different from being off by the same amount on the average risk-free rate, given their very different magnitudes. To at least account for the different relative magnitudes of the four moments, I use weights that are inversely proportional to the squares of the estimates from the data. In effect, the loss function minimizes the average squared percentage deviations of the model moments from their values in the data.

	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
$p = 0.025$	1.72	10.67	2.23	12.74
$p = 0.0$	10.69	12.90	2.71	13.03
Data	1.03	8.91	4.36	15.04

Table 2: Asset returns with and without disasters; no disappointment aversion. Disappointment aversion strength $\theta = 0$. Other parameters are as given in Table 1. Model moments are for a sample with no disasters. Expected returns are in percent, and their standard deviations are in percentage points. Data are from Gourio [16].

	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
$\theta = 2.44$	1.06	9.02	1.59	12.28
$\theta = 0.0$	10.69	12.90	2.71	13.03
Data	1.03	8.91	4.36	15.04

Table 3: Asset returns with and without disappointment aversion; no disasters. Disaster probability $p = 0$. Other parameters are as given in Table 1. Model moments are for a sample with no disasters. Expected returns are in percent, and their standard deviations are in percentage points. Data are from Gourio [16].

I now illustrate the trade-off between p, the probability of disaster, and θ, the strength of disappointment aversion. Table 4 shows the results. I begin with $p = 0.025$ and consider, in turn, probabilities that are $1/2$, $1/4$, and $1/8$ as large. The first and last rows recapitulate cases we've already seen—disaster risk without disappointment aversion, and disappointment aversion without disaster risk.

As the first two rows of Table 4 show, when $p = 0.025$, adding a small degree of disappointment aversion lowers the average risk-free rate with only a negligible impact on the other moments. In all the calibrations—as was the case as well in Tables 2 and 3—the model combining disappointment aversion and disaster risk tends to under-predict the volatility of the riskless rate, but otherwise performs quite well.[18] With a small disaster probability (just 0.625%) and a plausible amount of disappointment aversion ($\theta = 1.22$), the model first moments and the second moment of the equity return all come reasonably close to the data.

It's worth pointing out that—at least for the simple model of this section—both disappointment aversion and disaster risk are essential to the reasonable fit in the middle rows of Table 4. The two mechanisms are, in a sense, complements—a small amount of either factor, on its own, makes little im-

[18]Is the too-low volatility of the riskless rate a serious problem? Too much volatility—as, for example, characterized early habit models—has generally been viewed as a serious problem. Too little volatility, on the other hand, seems to raise few concerns, as discussed in footnote 3 above.

		$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
$p = 0.025$	$\theta = 0.00$	1.72	10.67	2.23	12.74
$p = 0.025$	$\theta = 0.06$	1.04	10.42	2.17	12.70
$p = 0.0125$	$\theta = 0.63$	1.04	9.87	1.94	12.56
$p = 0.00625$	$\theta = 1.22$	1.05	9.50	1.78	12.44
$p = 0.003125$	$\theta = 1.70$	1.05	9.28	1.69	12.37
$p = 0.0$	$\theta = 2.44$	1.06	9.02	1.59	12.28
Data	–	1.03	8.91	4.36	15.04

Table 4: The trade-off between disaster probability and strength of disappointment aversion. The (non-zero) θ values minimize the weighted distance between the model moments and the data, for each value of p. Disaster size is 25% ($y_D = 0.75$). Other parameters are as given in Table 1. Model moments are for a sample with no disasters. Expected returns are in percent, and their standard deviations are in percentage points. Data are from Gourio [16].

provement in the model's performance, but does improve the model's fit in the presence of the other. The combination of $p = 0.00625$ and $\theta = 1.22$ in the fourth row of Table 4 illustrates this point well. Table 5 contrasts the performance of the model with $p = 0.00625$ and no disappointment aversion; $\theta = 1.22$ and no disasters; and the combination of the two together. Clearly, neither alone is sufficient.

		$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
$p = 0.00625$	$\theta = 1.22$	1.05	9.50	1.78	12.44
$p = 0.00625$	$\theta = 0.0$	8.30	12.32	2.58	12.95
$p = 0.0$	$\theta = 1.22$	3.89	10.26	2.05	12.58

Table 5: Complementarity between disaster probability and strength of disappointment aversion. Disaster size is 25% ($y_D = 0.75$). Other parameters are as given in Table 1. Model moments are for a sample with no disasters. Expected returns are in percent, and their standard deviations are in percentage points.

3.2 The trade-off between disaster size and aversion to disappointment

There is also a trade-off between the strength of disappointment aversion and the disaster size, for a given disaster probability. To illustrate this, I fix $p = 0.01$, and let y_D vary from 0.64 to 0.98—*i.e.*, disaster sizes from 36% down to 2%. The latter value, of course, is no real disaster at all, and is just a slightly worse outcome than the low-growth state in 'normal times.' I keep the other parameters the same—the remaining parameters of the Markov chain are given in (14) and (15), $\beta = 0.97$, $\alpha = \rho = 5$, and leverage $\lambda = 3$.

As in the construction of Table 4, for each y_D I choose the value of θ to minimize the weighted distance between the model first and second moments and the data. These θ values range from $\theta = 0.04$ when the disaster size is 36% to $\theta = 2.4$ when the disaster size is 2%.

In the top two panels of Figure 1 I've plotted the model's average risk-free rate (Panel A) and average equity return (Panel B) against the size of disaster—note that the disaster sizes on the horizontal axes are in *descending* order.

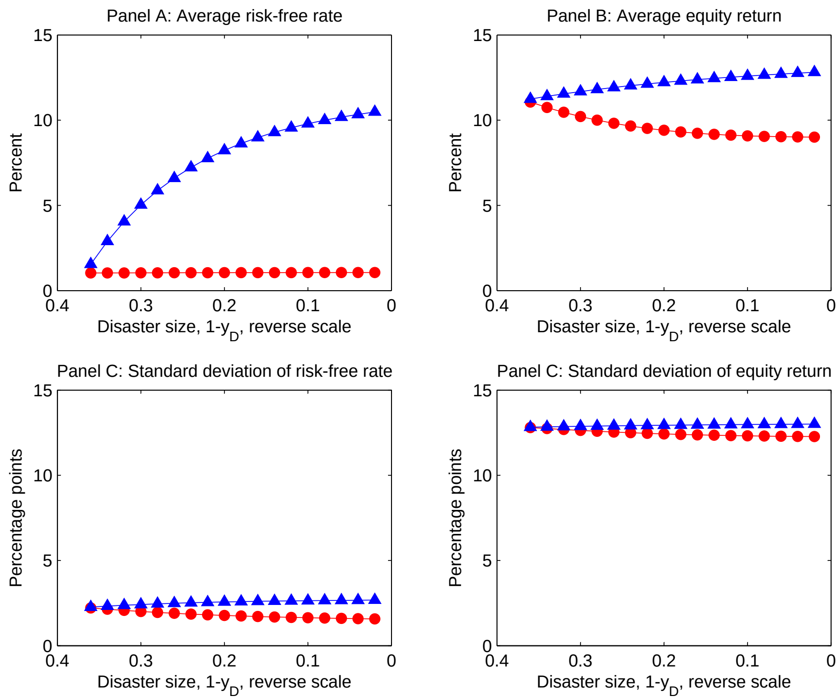

Figure 1: The trade-off between disaster size and disappointment aversion. Panel A: The average risk-free rate as the disaster size decreases from 36% to 2%, for a fixed probability of disaster $p = 0.01$. Under disappointment aversion (red circles), θ is set at each y_D to minimize the distance between model first and second moments and the data. Without disappointment aversion (blue triangles), θ is held fixed at zero as y_D varies. Panels B–D: Analogous plots for the average equity return, $\mathsf{E}(R^d)$, and the standard deviations of the risk-free rate, $\sigma(R^f)$, and equity return, $\sigma(R^d)$.

With the strength of disappointment aversion adjusted as the disaster size shrinks, the model produces a low risk-free rate across all disaster sizes. For comparison, in Panel A of Figure 1 I've also plotted the model's average risk-free rate absent disappointment aversion—*i.e.*, assuming $\theta = 0$. In that case, the risk-free rate quickly rises to unreasonable levels as the size the disaster falls.

As the disaster size decreases and θ is adjusted in response, the model's average equity return—the red circles in Panel B—declines slightly, but is essentially flat for disaster sizes below about 25%. With θ fixed at zero (the blue triangles), the average return increases by about 2.5 percentage points.

Varying the disaster size has less pronounced effects on the standard deviations of returns, shown in Panels C and D, and the effects are similar regardless of whether disappointment aversion is present or not.

4 Main results: A richer model with multiple disaster states

In this section, which contains the main results of the paper, I work with a richer model along the lines of Barro [5]—disentangling intertemporal substitution from risk aversion, incorporating a richer set of disaster states, and allowing for default, in disaster states, of the otherwise-risk-free asset. I discuss each of these modifications in turn, before moving to the model calibration and results.

The calculations in the previous section did not take full advantage of the Epstein-Zin aggregator (3); rather, they imposed $\rho = \alpha$, so that in the special case of no disappointment aversion, the representative agent's preferences collapse to the standard form used in Mehra and Prescott [23] and Rietz [27]. As we'll see, divorcing intertemporal substitution from risk aversion—allowing $\rho \neq \alpha$—makes possible a much closer fit of the model's moments to the data.

A second important difference here from the model of section 3 is that, following Barro [4] and Gabaix [15], the consumption process in this section features multiple disaster states. I will use the Barro's distribution [4, Figure I.A, p. 832] to calibrate the relative frequencies of those states and the relative sizes of the associated disasters.

I will continue to let consumption growth in normal times be given by the Mehra-Prescott Markov chain from (12) and (13). As before, the probability of transitioning from states L or H today to states L or H tomorrow is given by $(1-p)Q$, while with probability p tomorrow's state is one of N disaster states. The probability of disaster state $i \in \{1, 2, \ldots, N\}$ is $p_i = pf_i$, where f_i is the relative frequency (among disaster states) of state i (so $f_i \geq 0$, $\sum_i f_i = 1$). I will write gross consumption growth in disaster state i as

$$y_{D,i} = 1 - bz_i,$$

where the relative disaster sizes z_i obey $\sum_i f_i z_i = 1$; thus, the average disaster size will be b.

As in the model of the previous section, if the economy is in one of the disaster states today, then tomorrow's state is either L or H according to their long-run probabilities from Q—*i.e.,* $1/2$ and $1/2$. The probability of staying in today's disaster state or transitioning to one of the other $N-1$ disaster states is assumed to be zero.

Under those assumptions, the Markov chain for the model of this section can be written as follows. The set of gross consumption growth rates is:

$$y_{t+1} \in \begin{bmatrix} \mathbf{1} - b\mathbf{z} \\ y_L \\ y_H \end{bmatrix} \tag{20}$$

where $\mathbf{1}$ is an $N \times 1$ vector of ones, and $\mathbf{z}$ is the $N \times 1$ vector of relative disaster sizes z_i. The state transition matrix is:

$$P = \begin{bmatrix} \mathbf{0} & \frac{1}{2}[\mathbf{1}, \mathbf{1}] \\ p[\mathbf{f}, \mathbf{f}]^\top & (1-p)Q \end{bmatrix} \tag{21}$$

where $\mathbf{0}$ is an $N \times N$ matrix of zeros, and $\mathbf{f}$ is the $N \times 1$ vector consisting of the relative frequencies f_i.

A last important difference from the model of section 3 is that, following Barro [4], Gourio [16] and others, I allow for partial default on the 'riskless' in the event of a disaster. Rather than adding one more layer of uncertainty to the model, though, I assume for simplicity—and in contrast to the implementations by Barro and Gourio—that partial default occurs with probability one conditional on a disaster occurring. In particular, I assume there is a one-period zero-coupon bond (in zero net supply) whose price at t is again q_t and whose payoff at $t+1$ is F_{t+1}, where

$$F_{t+1} = \begin{cases} 1 - \eta_i & \text{if } y_{t+1} = y_{D,i} \\ 1 & \text{if } y_{t+1} \in \{y_L, y_H\} \end{cases} \tag{22}$$

The bond is priced according to

$$q_t = \mathsf{E}_t(m_{t+1} F_{t+1}).$$

I will still use the notation R^f to denote the return on this asset, but in order to avoid confusion with a truly risk-free asset, I will henceforward refer to that return as the model's bill rate.

Dividend growth continues to be represented by a leveraged version of consumption growth, as given in (17), and the key pricing equations for the equity claims are again the price-dividend relationships (18a) and (18b).

4.1 Model calibration

In the numerical experiments that will follow, I will vary the disaster probability p and average disaster size b, holding fixed the relative frequencies f_i and relative sizes z_i. I will calibrate those relative frequencies and sizes according to Barro's data; those parameters (along with Barro's data) are summarized in Table 6.

Disaster size	0.17	0.22	0.27	0.32	0.37	0.42	0.47	0.52	0.57	0.62
Frequency	20	13	3	9	5	0	2	3	3	2
z_i	0.59	0.77	0.94	1.11	1.29	1.46	1.63	1.81	1.98	2.16
f_i	0.33	0.22	0.05	0.15	0.08	0.00	0.03	0.05	0.05	0.03

Table 6: Distribution of disaster sizes. The data in the first two lines are from Barro's histogram of 60 episodes [4, Figure I.A, p. 832], taking the sizes to be the midpoints of Barro's bins. The z_i's and f_i's are the implied relative sizes and frequencies, as described in the text.

In the previous section, the probability of disaster and the disaster size—variations on a 2% probability of a 25% consumption decline—were rather arbitrary choices. In the calculations below, I focus on the frequency of disaster and average disaster size from Barro's [4] study: a 1.7% per year chance of a 29% output decline.[19] In terms of the model outlined above, setting $p = 0.017$ and $b = 0.29$ corresponds to the case considered by Barro.[20]

The calculations will examine—separately and in combination—disaster probabilities and sizes that are either 1, $\frac{1}{2}$, or $\frac{1}{4}$ times as large as those averages. That is, I will consider values of p such that

$$p \in \{0.017, 0.017/2, 0.017/4\}$$

and values of b such that

$$b \in \{0.29, 0.29/2, 0.29/4\}.$$

The various choices of p and b will define the experiments I perform in this section. We'll see that with appropriate choices of the curvature parameter, α, and the strength of disappointment aversion, θ, the model can come quite close to matching exactly the first and second moments of asset returns. The values of α and θ that maximize the model's fit are, moreover, all of plausible size: I restrict $\alpha \leq 5$ and, based on the empirical evidence discussed in section 2.3.2, $\theta \leq 3$.

Other parameter values that will remained fixed throughout the experiments include the leverage parameter, λ, and the discount factor β. I will continue to set $\lambda = 3$ and $\beta = 0.97$.

The new parameters η_i introduced in equation (22)—the default losses on the otherwise riskless asset in disaster states—I take to be proportional to the

[19] In [4], Barro defines a disaster as a decline in real output per capita of 15% or more; such events occur with a frequency of 1.7% per year in his sample (60 episodes out of 3500 country-years). The average size of a decline is 29%.

[20] While 1.7% and 29% may appear to be not too far from the 2% and 25% benchmarks of the previous section, the similarity is superficial: the 29% average disaster size does not capture the long left tail in the distribution of disaster states in the present model, and those tail outcomes matter a great deal in the representative agent's preferences (with or without aversion to disappointment). Gourio [16], for example, proxies Barro's distribution of disasters with a single disaster state involving a 43% drop in consumption.

disaster sizes bz_i. In line with Gourio's [16] specification,[21] I set $\eta_i = 0.4bz_i$.

Finally, given these parameter choices, the data favor very low elasticities of intertemporal substitution. Rather than look for choices of ρ, α, and θ which produce the best fit to the data, I will initially fix ρ^{-1} (the elasticity of intertemporal substitution) at 0.1. Over the range of disaster sizes and probabilities I consider, this is close to an optimal choice, and even with this constraint in place, there are a range of choices of α and θ which, to practical purposes, produce moments that essentially match the data. Section 4.2.1 presents the results for the model with a low elasticity of intertemporal substitution.

We know, however, that if we look beyond the means and standard deviations of asset returns, there are reasons to prefer larger elasticities of intertemporal substitution. Barro [5], for example, notes that elasticities greater than one are necessary if we want greater disaster risk to result in lower equity prices. Gourio [16] also shows that an elasticity of intertemporal substitution greater than one is necessary for replicating certain facts about the time series predictability of returns—namely that dividend yields forecast equity returns and excess equity returns with regression coefficients of identical (positive) sign.[22]

Hence, I will also examine the standard case of $\rho^{-1} = 1$ and the (increasingly standard) case of $\rho > 1$ (in particular, $\rho^{-1} = 1.5$). Those results are presented in section 4.2.3.

All of these basic parameter choices are summarized in Table 7.

ρ^{-1}	β	λ	η_i	(f_i, z_i)
$\{0.1, 1.0, 1.5\}$	0.97	3	$0.4bz_i$	Values in Table 6

Table 7: Basic parameter settings for the model. These parameter values are used throughout this section. b is the average disaster size, which will be varied, while the z_i's are the fixed relative disaster sizes. Other parameters of the consumption process (20)–(21) are the values of Q, y_L, and y_H given in expressions (12)–(13).

4.2 Numerical results

The experiments vary the disaster probability p and average disaster size b, holding fixed the relative size distribution of disaster states. Given values of p and b, I ask what values of the certainty equivalent parameters α and θ bring the model's predictions for the moments of asset returns close to the data. In particular, I'm interested in whether modest disaster risks—p and b much less

[21]Gourio assumes a 40% chance of default equal to disaster size; as noted above, I will treat the partial defaults as certain conditional on a disaster occurring.

[22]The same forces—essentially, the need for heightened rate of return risk to reduce asset demands—motivated Bansal and Yaron [3] to consider intertemporal substitution elasticities greater than one.

than 0.017 and 0.29—when faced by an agent with a plausible amount of disappointment aversion—restricting $\theta \leq 3$—can produce return moments that match the data.

As in section 3, my choices of θ and α here attempt to minimize the (weighted) distance between the model moments and the data. Footnote 17 describes the loss function in more detail. Rather than solve for highly precise minimizers—realistically, we can never hope to have evidence that distinguishes between, say, $\theta = 0.5$ and $\theta = 0.51$, or between $\alpha = 2.2$ and $\alpha = 2.24$—I work on small discrete grids:

$$\theta \in \{0, 0.1, 0.2, \ldots, 2.9, 3\}$$

and

$$\alpha \in \{0, 0.1, 0.2, \ldots, 4.9, 5\}.$$

Also as in section 3, the moments I report are what one would obtain for a sample in which there were no disasters—*i.e.,* they summarize the model's average behavior in 'normal' times.

4.2.1 Low elasticity of intertemporal substitution

As noted above, the best fit to the four moments of the data is obtained when the intertemporal elasticity of substitution is low. Table 8 shows results for the model when $\rho^{-1} = 0.1$

Line 1 of the table shows results using Barro's [4] probability of disaster and average disaster size. Lines 2–3 reduce the probability of disaster to one-half and one-fourth the size in Line 1, holding fixed the average disaster size. Lines 4–5 perform a similar exercise on the average disaster size, holding the disaster probability fixed, while Lines 6–7 reduce both the average size and probability. Line 8 gives results for the model with no disasters.

The first thing to note about the results are the generally close matches between the model moments and the data—the biggest misses are on the order of 0.7–0.8 percentage points, mainly in the average equity return and the standard deviation of the bill rate. Of these, only the lowest values of the standard deviation of the bill rate may represent significant departures from the data.[23]

Also, only in two instances—namely, the case with a disaster probability of 0.425% and average disaster size of just 7.25% (Line 7) plus the case with no disaster risk at all (Line 8)—do we come near our upper bound of 3 on the strength of disappointment aversion parameter θ. And, in all cases the values for α, the curvature parameter in the certainty equivalent, are well below their upper bound of 5.

[23] *If* returns were normally distributed in the data (they are not), then a two-standard-error band around the 4.36 estimate of the standard deviation of the bill rate—based on 78 years' worth of data—would be roughly $[3.6, 5.1]$.

Line	p	b	θ	α	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
1	0.01700	0.2900	0.3	2.9	1.02	9.58	4.37	15.17
2	0.00850	0.2900	0.9	2.8	1.01	9.33	3.97	14.81
3	0.00425	0.2900	1.6	2.1	1.04	9.09	3.71	14.55
4	0.01700	0.1450	1.2	3.3	1.04	9.32	3.75	14.64
5	0.01700	0.0725	1.9	1.5	1.06	9.04	3.56	14.42
6	0.00850	0.1450	1.7	2.6	1.04	9.21	3.62	14.51
7	0.00425	0.0725	2.4	0.3	1.03	8.97	3.52	14.37
8	0	0	2.5	0.5	1.04	9.07	3.50	14.36
Data	–	–	–	–	1.03	8.91	4.36	15.04

Table 8: Results for the model with multiple disaster states and $\rho^{-1} = 0.1$. The parameters p and b are the disaster probability and average disaster size. θ and α are selected to minimize the distance between the model moments and the data. Other model parameters are as given in Tables 6 and 7. Model moments are for a sample with no disasters. Expected returns are in percent, and their standard deviations are in percentage points. Data are from Gourio [16].

4.2.2 The trade-off between curvature and strength of disappointment aversion

Comparing the α values in Lines 1, 4, and 5 of Table 8—experiments which hold p at 0.017 while setting b equal to 0.29, 0.145, and 0.0725—presents an apparent puzzle. While θ increases monotonically as the disaster size is reduced, α increases from 2.9 to 3.3, then falls to 1.5.

The non-monotonicity in α across these three cases is a reflection of a close substitutability between α and θ. For most of the cases recorded in Table 8, it's possible to alter the strength of disappointment aversion somewhat, while compensating with a change in the curvature parameter, and achieve results nearly as good as the optima presented in the table. In other words, there's a locus of α–θ pairs on which the loss function being minimized is nearly flat. The loss-minimizing choice of (θ, α) at $(p, b) = (0.017, 0.0725)$ happens to have $\theta = 1.9$ and $\alpha = 1.5$, but the model's fit is almost at good at, for example, $\theta = 1.8$ and $\alpha = 2.3$.

To illustrate this substitutability, Table 9 gives results for the case of $(p, b) = (0.017, 0.0725)$—the environment from Line 5 of Table 8—as θ varies from 1.5 to 2.0. The corresponding α values are best, given the values of θ. The resulting sets of moments are barely distinguishable from one another; any of them might, in fact, might have been optimal if our loss function used slightly different weights.[24]

There are limits to the trade-off: for θ greater than 2.0, we would run into the $\alpha \geq 0$ constraint, while for θ less than 1.5, we would violate our $\alpha \leq 5$

[24] For example, if we placed a bit more weight on the average equity return relative to the average bill rate.

θ	α	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
1.5	4.9	1.05	9.60	3.56	14.57
1.6	4.0	1.06	9.45	3.56	14.52
1.7	3.2	1.02	9.29	3.55	14.48
1.8	2.3	1.06	9.17	3.56	14.45
1.9	1.5	1.06	9.04	3.56	14.42
2.0	0.8	1.01	8.89	3.53	14.36
Data		1.03	8.91	4.36	15.04

Table 9: Trade-off between θ and α. Results for the model with $\rho^{-1} = 0.1$, $p = 0.017$ and $b = 0.0725$. The α values give the best fit at each value of θ. Other aspects of the calculations are as described in the notes to Table 8.

constraint.[25] The limits on the substitutability of θ and α of course depend on the disaster risk specification (p, b). For example, when both the disaster size and probability are one-fourth their original values from Barro—*i.e.*, $(p, b) = (0.00425, 0.0725)$, as in Line 7 of Table 8—the $\alpha \leq 5$ constraint binds for $\theta < 1.9$, the $\alpha \geq 0$ constraint binds for $\theta > 2.4$.

It's important to note that the substitutability between the curvature and disappointment aversion parameters is with respect to the small set of moments we are focusing on in these experiments. Given a rich enough set of securities, $(\theta, \alpha) = (2.0, 0.8)$ and $(\theta, \alpha) = (1.5, 4.9)$ can be expected to produce different pricing implications for at least some of the securities.

To illustrate this possibility, Figure 2 plots the logarithms of normalized conditional state price vectors for the environments corresponding to the first and penultimate lines in Table 9—*i.e.*, for $(\theta, \alpha) = (1.5, 4.9)$ and $(2.0, 0.8)$. The state price vectors are normalized by the state price vector for an agent with the same elasticity of intertemporal substitution, but no aversion to risk or disappointment, $(\theta, \alpha) = (0, 0)$. Panel A shows the log normalized state prices conditional on being in the low-growth (y_L) state today; the prices in Panel B are conditional on being in the high-growth (y_H) state.[26]

The disaster states D_1–D_9 in the Figure are the nine disaster states that have non-zero frequencies in the distribution summarized in Table 6. With an average disaster size of 7.25%, the smallest (D_1) and largest (D_9) consumption declines are 4.3% and 15.7%, respectively.

Relative to the state prices that would obtain with a risk-neutral agent, consumption in any of the disaster states is more valuable under the prices for

[25]Recall that we are are choosing values of θ and α from discrete grids, with points spaced 0.1 apart. Increasing θ from 2.0 to 2.1 results in $\alpha = 0.0$, while lowering θ from 1.5 to 1.4 results in $\alpha = 5.0$.

[26]If $\psi_{i,j}$ denotes the price in state i today of a unit of consumption in state j next period, then—as usual—$\psi_{i,j} = P_{i,j} m_{i,j}$. The normalized state prices plotted in the figure are $\log(\psi_{i,j} / \hat{\psi}_{i,j})$, where $\hat{\psi}_{i,j}$ are state prices for an agent with no (timeless) risk aversion ($\theta = \alpha = 0$). The matrix of $\hat{\psi}$ values is approximately βP. *N.B.*, the similarity in terminology notwithstanding, the state prices $\hat{\psi}$ obtained from a risk-neutral agent are not the risk-neutral prices for these environments.

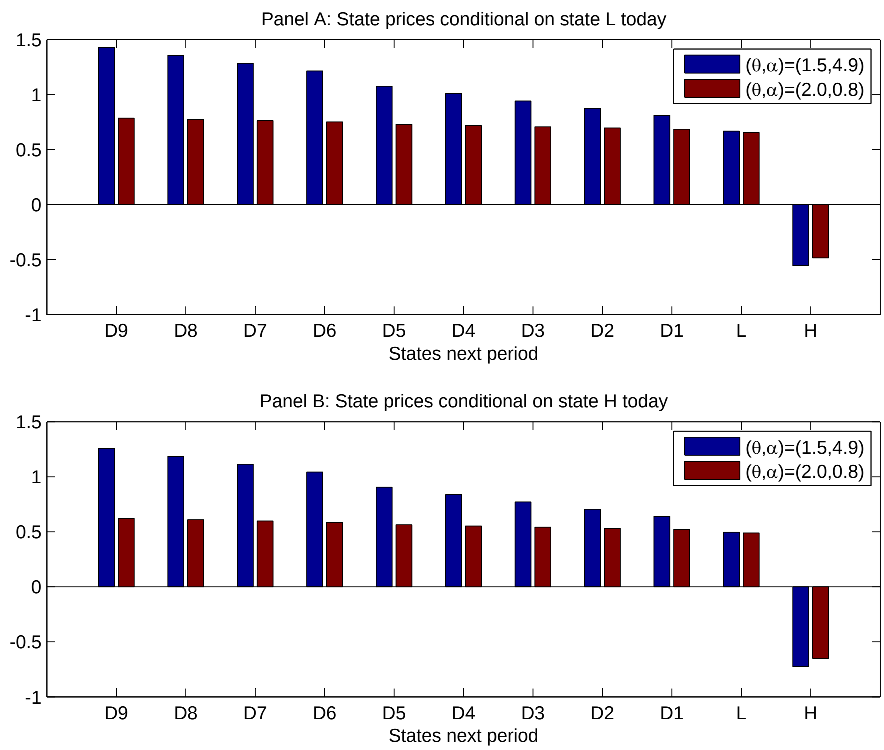

Figure 2: Normalized state price vectors for two different combinations of curvature and disappointment aversion. Log normalized state price vectors for environments defined by $(\theta, \alpha) = (1.5, 4.9)$ and $(2.0, 0.8)$ from Table 9. D_1 through D_9 are disaster states; L and H are 'normal' times. Panel A: State prices conditional on state L today. Panel B: Conditional on state L today. See text, particularly footnote 26, for details.

an agent with either $(\theta, \alpha) = (1.5, 4.9)$ or $(2.0, 0.8)$. In the case with relatively more disappointment aversion and less curvature—that is, $(\theta, \alpha) = (2.0, 0.8)$—the effect is remarkably uniform (in percentage terms) across all the disaster states. With relatively less disappointment aversion and more curvature, $(\theta, \alpha) = (1.5, 4.9)$, the increased value (relative to risk neutrality) is bigger for deeper disasters.

Thus, while the two (θ, α) pairs produce nearly identical first and second moments of asset returns, there are conceivable securities that would be priced quite differently in the two environments—for example, types of size-contingent disaster insurance.

One thing these results do suggest, though, is that attempting to estimate the model's preference parameters using data on consumption growth and a pair of asset returns is likely to run into identification problems. The likelihood function implied by the model will be nearly flat along a locus of (θ, α) pairs.

4.2.3 Results for larger elasticities of intertemporal substitution

While the best fit of the model to the observed means and standard deviations of returns is obtained for a low value of the intertemporal elasticity of substitution, there are other considerations—such as the relationship between the amount of disaster risk and asset values (as in [5]) or the time series properties of returns and asset prices (as in [16])—that point to larger values.

In this section, I will consider larger values of the elasticity of intertemporal substitution, namely $\rho^{-1} = 1$ (which corresponds to a logarithmic Epstein-Zin aggregator) and $\rho^{-1} = 1.5$ (an increasingly common choice, since Bansal and Yaron's analysis in [3]). In the process, the results illustrate why the model favors a low substitution elasticity. In a nutshell, at larger values of the intertemporal elasticity of substitution, the model produces an unrealistically small volatility for the bill rate. The problem is not unique to the combination of disaster risk and disappointment aversion—it's present as well if we assume no aversion to disappointment.

In a model like Barro's [5], with *i.i.d.* 'normal' consumption growth and a constant (independent) disaster probability, the bill rate is constant on non-disaster samples, regardless of the elasticity of intertemporal substitution. Analyses based on such models do not attempt to match the volatility of the bill rate, and hence would not notice this departure from the data. In Gourio [16] and Gabaix [15], time-variation in the probability of disaster—not present in our model—accounts for all the variation (over non-disaster samples) in the bill rate.

Table 10 presents some results from the model with larger values of the intertemporal elasticity of substitution. Results for $\rho^{-1} = 1$ are in the top panel of the table, those for $\rho^{-1} = 1.5$ in the bottom panel. Compared with the results of section 4.2.1, the results in Table 10 are generally worse, with the most notable failing of the model being the uniformly low values for the standard deviation of the bill rate.

Panel A: $\rho^{-1} = 1$

Line	p	b	θ	α	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
1	0.01700	0.2900	0.0	4.0	1.02	10.22	0.41	10.73
2	0.00850	0.2900	0.0	4.8	1.13	9.88	0.41	10.70
3	0.00425	0.2900	0.7	5.0	1.13	10.60	0.29	10.93
4	0.01700	0.1450	3.0	5.0	1.74	11.83	0.19	11.38
5	0.01700	0.0725	3.0	5.0	2.23	10.62	0.24	11.15
6	0.00850	0.1450	3.0	5.0	2.06	10.96	0.23	11.21
7	0.00425	0.0725	3.0	5.0	2.35	10.18	0.26	11.05
Data	–	–	–	–	1.03	8.91	4.36	15.04

Panel B: $\rho^{-1} = 1.5$

Line	p	b	θ	α	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
1	0.01700	0.2900	0.0	4.0	1.11	10.23	0.25	10.58
2	0.00850	0.2900	0.0	4.8	1.07	9.74	0.24	10.52
3	0.00425	0.2900	0.8	5.0	1.14	10.89	0.14	10.86
4	0.01700	0.1450	3.0	5.0	2.14	12.20	0.12	11.33
5	0.01700	0.0725	3.0	5.0	2.48	10.84	0.16	11.08
6	0.00850	0.1450	3.0	5.0	2.35	11.21	0.14	11.14
7	0.00425	0.0725	3.0	5.0	2.55	10.32	0.17	10.97
Data	–	–	–	–	1.03	8.91	4.36	15.04

Table 10: Results for the model with multiple disaster states and larger intertemporal elasticities of substitution. The parameters p and b are the disaster probability and average disaster size. θ and α are selected to minimize the distance between the model moments and the data. Other model parameters are as given in Tables 6 and 7. Model moments are for a sample with no disasters. Expected returns are in percent, and their standard deviations are in percentage points. Data are from Gourio [16].

Apart from the low volatility of the bill rate, though, the results are broadly similar to the results obtained by Gourio [16, Table 3] for similar values of the elasticity of intertemporal substitution and long-run average disaster risk comparable to the estimates from Barro [4]—a probability of 1.7% and a single disaster state of size 43%. The average returns across all our (p, b) specifications are at least as close to the data as they are in Gourio's results, and the misses in the volatility of our equity return are of comparable size to the misses in Gourio, though in the opposite direction.[27]

Moreover, the relatively high values of the bill rate—above 2% in a number of cases—are not especially a cause for concern, since they can be remedied with an appropriate choice of the utility discount factor β. To this point, I've

[27]To be sure, Gourio is trying to fit his model—which incorporates time-varying disaster risk—to a larger set of moments, including time series properties of the returns data.

simply fixed β at 0.97, a somewhat arbitrary choice. Calibrating β to achieve a particular average interest rate is a standard practice in much of macroeconomics; in fact there is little other basis for choosing β.[28] Increasing β lowers the average bill rate and equity return, with little effect on their volatilities. For example, keeping all other parameters the same as described in Table 8, setting $\beta = 0.984$ in the environment of line 5 of Panel B produces the following moments (with the original line 5 shown for comparison):

	p	b	θ	α	$\mathsf{E}(R^f)$	$\mathsf{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
$\beta = 0.970$	0.01700	0.0725	3.0	5.0	2.48	10.84	0.16	11.08
$\beta = 0.984$	0.01700	0.0725	3.0	5.0	1.03	9.25	0.15	10.91
Data	–	–	–	–	1.03	8.91	4.36	15.04

Table 11: Adjusting β to target the bill rate. The line for $\beta = 0.97$ simply replicates line 5 of Panel B from Table 8. The line for $\beta = 0.984$ keeps all parameters the same except for the choice of β. See notes to Table 8 for details of calculations.

In addition to the uniformly low volatilities of the bill rate, another notable feature of the results for larger elasticities of intertemporal substitution is the choice of θ and α. If the average disaster is size 0.29 (Lines 1–3 in either panel), the best fits involve either a small amount of disappointment aversion (when $p = .00425$) or none at all ($p = 0.017$ or 0.0085). In those cases, α is at or just below its upper bound of 5.0. In contrast, for the smaller average disaster sizes (0.145 or 0.0725), the optimal θ and α are both at their upper bounds.[29] For the model with a very small elasticity of intertemporal substitution, the best choices of θ and α varied more smoothly with the amount of disaster risk.

Finally, we note that the choice of $\rho^{-1} = 1.0$ or 1.5 makes little difference for the results, as can be seen in a comparison of Panels A and B of the table. Note, too, that an elasticity of intertemporal substitution equal to one is in no sense a critical value for the performance of the model (in terms of the four moments considered here). Additional simulations (not reported) show that most of the deterioration in the model's fit, relative to the $\rho^{-1} = 0.1$ case, occurs between $\rho^{-1} = 0.1$ and 1.

5 Conclusions and directions for further work

Models featuring rare consumption disasters have recently proven successful at rationalizing many aspects of the data on consumption growth and asset

[28]One might argue introspectively that β should be less than one: If expecting the same consumption next period as this period, would one voluntarily trade a unit of today's consumption for less than a unit next period? I suspect most people would answer no.

[29]Both constraints $\theta \leq 3$ and $\alpha \leq 5$ do bind, but relaxing those constraints—even by large amounts—produced only marginal improvements in the model's fit.

returns. The empirical frequency of disasters, however, is subject to considerable uncertainty. Using the experiences of many countries—as Barro [4] did, effectively jump-starting a literature that had been dormant—provides enough observations to make more precise inferences, but only if those experiences reflect a distribution of disaster risk common to all countries. It's natural to be skeptical of whether the very sizable disaster risks documented by Barro and co-authors using cross-country data are a good characterization of risks faced by investors in any single country, particularly investors in the US or other developed economies.

The message of this paper is that disaster risks do not have to be nearly as great as commonly assumed in the literature, once we allow agents in rare disaster models to display an empirically plausible degree of aversion to disappointment. Experimental evidence indicates that individuals tend to overweight bad outcomes relative to expected utility; disappointment aversion represents a form of risk preferences that have that feature. I impose empirical discipline on the model by restricting the strength of disappointment aversion to be in the range of estimates from the experimental studies of Choi *et al.* [11] and from Camerer and Ho [8].

The main results of the paper are in Tables 8 and 10, which present simulated moments of asset returns for a model incorporating many of the features common to the rare disaster literature—a distribution of possible disaster sizes, partial default (in the event of disaster) on the economy's otherwise riskless asset, and levered equity.

For a low value of the elasticity of intertemporal substitution, the model is able to almost perfectly match the means and standard deviations of the bill rate and equity return, for disaster risks one-half or one-fourth the size implied by the cross-country data. For larger elasticities of intertemporal substitution, the results are less satisfactory, but still—apart from a too-low standard deviation of the model-implied bill rate—broadly similar to other results in the rare disaster literature [16]. These results, too, are for disaster risks much smaller than commonly assumed.

These results are encouraging, especially if one is skeptical about estimates of the magnitude of disaster risk.[30] Still, there are obvious directions for further work.

I have not employed generalized disappointment aversion—the form developed by Routledge and Zin [28]—largely on the strength of the model's results for a low elasticity of intertemporal substitution in section 4.2.1. The model moments in Table 8 are already so close to the data, any improvement in fit that could be gained with generalized disappointment aversion would seem to hardly justify the added complexity or the introduction of one more preference parameter. The results for the larger elasticities of intertemporal substitution, however, are a different story. There is certainly ample room for

[30] A common rejoinder is that one may well be skeptical about the experimental results used to calibrate risk preferences. There is an important difference, though—we can, in principle, run as many experiments as we need to sharpen those parameter estimates; on disaster risk, we're stuck with the limited data we currently have.

improvement in the results reported in Table 10. Whether having additional control over the set of states viewed as disappointing is useful in that context is a question worth exploring.

Incorporating a time-varying disaster probability is also a natural direction for further work, but poses its own particular challenges, not the least of which is calibrating that unobserved process. Moreover, my point in this paper has been to show that a plausible degree of aversion to disappointment reduces the amount of disaster risk necessary for the rare disaster model to produce realistic asset-pricing implications. The numerical experiments therefore took the form of varying the disaster risk environment, summarized by the parameters (p, b). A similar set of experiments in a model with a time-varying disaster probability would require specifying how changes in (p, b)—understood now as parameters of the long-run distribution—affect the other parameters of the disaster probability process, a specification with little empirical guidance.[31] Following the approach of Wachter [32]—who uses the observed persistence of the price-dividend ratio and the volatility of stock returns to back out the persistence and volatility of the disaster probability process—would be one possible way of addressing this ambiguity.

Finally, for simplicity—to keep the exercise as straightforward as possible—I have focused on changes only in the probability of disaster and the average size of disaster, while maintaining the relative frequencies and relative sizes as in Barro's [4] data. Exploring the interaction, if any, between disappointment aversion and changes in the shape of the distribution of disasters also seems a natural avenue for further work.

References

[1] Andrew B. Abel. Risk premia and term premia in general equilibrium. *Journal of Monetary Economics*, 43(1):3–33, 1999.

[2] David K. Backus, Bryan R. Routledge, and Stanley E. Zin. Recursive preferences. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke, 2008.

[3] Ravi Bansal and Amir Yaron. Risks for the long run: A potential resolution of asset-pricing puzzles. *Journal of Finance*, 59(4):1481–1509, 2004.

[4] Robert J. Barro. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3):823–866, 2006.

[5] Robert J. Barro. Rare disasters, asset prices, and welfare costs. *American Economic Review*, 99(1):243–264, 2009.

[31] Consider the model of Gourio [16], for example. There, the time-varying disaster probability is modeled as following a two-state Markov chain. In general, one cannot simply vary the long-run mean of the process without regard to its standard deviation, if the process is required to be strictly positive.

[6] Robert J. Barro and Tao Jin. On the size distribution of macroeconomic disasters. *Econometrica*, 79(5):1567–1589, 2011.

[7] Robert J. Barro and José Ursúa. Macroeconomic crises since 1870. *Brookings Papers on Economic Activity*, Spring:255–335, 2008.

[8] Colin F. Camerer and Teck-Hua Ho. Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2):167–196, 1994.

[9] Claudio Campanale, Rui Castro, and Gian Luca Clementi. Asset pricing in a production economy with Chew-Dekel preferences. *Review of Economic Dynamics*, 13(2):379–402, 2010.

[10] John Y. Campbell and John H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251, 1999.

[11] Syngjoo Choi, Raymond Fisman, Douglas Gale, and Sachar Kariv. Consistency and heterogeneity of individual behavior under uncertainty. *The American economic review*, 97(5):1921–1938, 2007.

[12] John H. Cochrane. Financial markets and the real economy. In Rajnish Mehra, editor, *Handbook of the Equity Premium*. Elsevier, Amsterdam, 2007.

[13] John H. Cochrane. The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, 21(4):1533–1575, 2008.

[14] George M. Constantinides and Darrell Duffie. Asset pricing with heterogeneous consumers. *Journal of Political economy*, 104:219–240, 1996.

[15] Xavier Gabaix. Variable rare disaster: A tractable theory of ten puzzles in macro-finance. *American Economic Review*, 98(2):64–67, 2008.

[16] François Gourio. Time-series predictability in the disaster model. *Finance Research Letters*, 5(4):191–203, 2008.

[17] Sanford Grossman and Robert Shiller. The determinants of the variability of stock market prices. *American Economic Review*, 71:222–227, 1981.

[18] Faruk Gul. A theory of disappointment aversion. *Econometrica*, 59(3):667–686, 1991.

[19] Fatih Guvenen. A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1750, 2009.

[20] Lars P. Hansen and Kenneth Singleton. Stochastic consumption, risk aversion, and the intertemporal behavior of asset returns. *Journal of Political Economy*, 1983.

[21] Narayana R. Kocherlakota. The equity premium: It's still a puzzle. *Journal of Economic Literature*, 34(1):42–71, March 1996.

[22] Robert E. Lucas, Jr. Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445, November 1978.

[23] Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15:145–161, 1985.

[24] Rajnish Mehra and Edward C. Prescott. The equity risk premium: A solution? *Journal of Monetary Economics*, 22:133–136, 1988.

[25] Angelo Melino and Alan X. Yang. State-dependent preferences can explain the equity premium puzzle. *Review of Economic Dynamics*, 6(4):806–830, 2003.

[26] Emi Nakamura, Jón Steinsson, Robert Barro, and José Ursúa. Crises and recoveries in an empirical model of consumption disasters. National Bureau of Economic Research Working Paper 15920, 2010.

[27] Thomas A. Rietz. The equity risk premium: A solution. *Journal of Monetary Economics*, 22(1):117–131, 1988.

[28] Bryan R. Routledge and Stanley E. Zin. Generalized disappointment aversion and asset prices. *Journal of Finance*, 65(4):1303–1332, 2010.

[29] Robert J. Shiller. Consumption, asset returns and economic fluctuations. *Carnegie-Rochester Conference on Public Policy*, 1982.

[30] Chris Starmer. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38:333–382, 2000.

[31] Kjetil Storesletten, Christopher I. Telmer, and Amir Yaron. Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics*, 10(4):519–548, 2007.

[32] Jessica A. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, 68(3):987–1035, 2013.

Appendix: Solution method

Given parameter values (including a Markov process for consumption growth), solving the model involves finding the normalized value function ϕ and certainty equivalent μ that satisfy versions of (9) and either (4) or (6), appropriate to the Markov chain environment.

With n Markov states, both ϕ and μ are n-vectors. The Markov chain version of (9) is

$$\phi_i = [1 - \beta + \beta\mu_i^{1-\rho}]^{1/(1-\rho)} \tag{23}$$

where μ_i satisfies a version of (4)

$$f(\mu_i) = \sum_j P_{i,j} f(\phi_j e^{x_j}) - \theta \sum_j P_{i,j} \left[f(\mu_i) - f(\phi_j e^{x_j}) \right] I_{i,j}, \tag{24}$$

where $f(w) = w^{1-\alpha}/(1-\alpha)$, and I is defined analogously to (10)

$$I_{i,j} = I[\phi_j e^{x_j} < \mu_i] \tag{25}$$

The solution strategy is to iterate on ϕ using (23)–(25). Given an initial ϕ^0, use (24) and (25) to find the vector of μ_i's that enter into the right-hand side of (23), calculate ϕ^1 as the value of the right-hand side of (23), and repeat with ϕ^1 in the role of ϕ^0.

At each iterate ϕ^j, I solve for the certainty equivalent μ using a dichotomy (*i.e.*, bisection) algorithm applied to (24) and (25). More precisely, let $A_i = f(\mu_i)$ and $\xi_j = f\left(\phi_j e^{x_j}\right)$. Given this notation, what we seek is the vector $A = \{A_i\}_{i=1}^n$ such that

$$A_i = \sum_j P_{i,j} \xi_j - \theta \sum_j P_{i,j} \left[A_i - \xi_j \right] I\left(\xi_j < A_i\right) \tag{26}$$

Because $f(w) = w^{1-\alpha}/(1-\alpha)$ is monotone increasing, this is equivalent to (24) and (25). Given an A that satisfies (26), we then get $\mu_i = f^{-1}(A_i)$.

Fix today's state i, and let

$$V(A_i) = A_i - \{\text{the right-hand side of (26)}\}.$$

To justify the dichotomy approach, we will show that $V(A_i)$ is strictly increasing in A_i, that $V(\min\{\xi_k\}) < 0$, and that $V(\max\{\xi_k\}) > 0$. It will then follow that dichotomy beginning on the interval $[\min\{\xi_k\}, \max\{\xi_k\}]$ converges to an A_i such that $V(A_i) = 0$.

That $V(A_i)$ is strictly increasing in A_i is fairly obvious: each non-zero term in

$$\theta \sum_j P_{i,j} \left[A_i - \xi_j \right] I\left(\xi_j < A_i\right)$$

is increasing in A_i and strictly positive. Increasing A_i will either leave the number of terms in the sum the same or add more strictly positive terms.

It also follows that

$$V(A_i) \geq A_i - \sum_j P_{i,j} \xi_j,$$

so $V(\max\{\xi_k\}) > 0$ is immediate.

To show that $V(\min\{\xi_k\}) < 0$, it's enough to note that $I\left(\xi_j < \min\{\xi_k\}\right) = 0$ for all j, so that

$$V(\min\{\xi_k\}) = -\sum_j P_{i,j} \left(\xi_j - \min\{\xi_k\}\right) < 0.$$