Disastrous disappointments: Asset-pricing with disaster risk and disappointment aversion

Presentation for CEF 2013, Vancouver

Jim Dolmas

Federal Reserve Bank of Dallas
jim@jimdolmas.net

07/12/13

What does this paper do?

- Look at asset-pricing implications of rare disasters (RD) together with disappointment aversion (DA)
- The model is in the class of representative agent exchange economies
- Two environments in paper:
    - Similar to Rietz (1988 *JME*), 'normal times' given by Mehra-Prescott Markov chain, disaster probability is constant.
    - Similar to Gourio (2008 *FRL*), 'normal times' are lognormal consumption growth, disaster probability is time-varying.
- Disappointment aversion as defined by Gul (1991 *Ecta*), used by Routledge and Zin (2010 *JF*) or Campanale, Castro and Clementi (2010 *RED*).

Why would we want to do this?

- In nearly 30 years since Mehra-Prescott, we've gone from having no good explanations for asset returns and consumption to having lots.
- For example. . .
  - Rare disasters
  - Long-run risk
  - Campbell-Cochrane habits
  - Disappointment aversion
  - State-dependent preferences
- And those are just the representative agent resolutions!

Why would we want to do this?

- All these resolutions add features to the basic CRRA/lognormal model that are hard to quantify.
- Just think about DA and RD:
    - Hard to discriminate between something that happens a couple times every hundred years or once every couple hundred years.
    - We have some experimental evidence that people over-weight bad outcomes, but not really definitive parameter estimates.

Two rationales for combining DA and RD

Rationale #1

We think disaster risk and disappointment aversion are plausible mechanisms, but we're skeptical about the magnitudes.

It would be nice to know they're complementary—i.e., you only need a little of each when the other is present.

Rationale #2

We're convinced (independent of data on asset returns) that disappointment aversion and disaster risk are important features of taste and technology.

It would be nice to know they can rationalize asset market data when they both are present in a model.

What do I find?

- RD and DA are complementary in sense that a little of each, together, works as well as a lot of either, separately.
- In context of Rietz-like Markov chain model, can basically match means and standard deviations of asset returns in a model with:
  - Disaster probabilities and sizes one-half or one-fourth the average sizes in Barro's *QJE* paper;
  - Strength of DA less that one-fourth size as in Routledge and Zin;
  - And plausible discounting, relative risk aversion, and IES parameters.

Roadmap of the talk

1. Introduction and motivation
2. Background
3. Model
4. Numerical experiments

Background: Rare disasters

- One of the earliest responses to Mehra-Prescott (Rietz 1988 *JME*), but lost favor.
  - Mehra-Prescott's response (1988 *JME*): disasters are too big; risk aversion, discount factor too high; risk-free asset not likely risk-free in a crisis.
  - Merits just a footnote mention in Kocherlakota's 1996 *JEL* survey.
- Interest revived by empirical work of Barro (2006 *QJE*), theoretical contributions by Gabaix, Gourio, others.
- Disaster risk provides potential resolution to equity premium puzzle by providing a channel for holding down risk-free rate.

More

Background: Disappointment aversion

- As with Epstein-Zin preferences, uncertain future utility reduced to a certainty equivalent value. Disappointment aversion enters through the form of the certainty equivalent operator.
    - Disappointing outcomes get extra weight.
- But what outcomes are disappointing?
- Gul (1991 *Ecta*), certainty equivalent obeys a consistency requirement—threshold for disappointment is certainty equivalent itself.
    - So certainty equivalent only defined implicitly.

Disappointment aversion: In words

- Very roughly paraphrased: disappointment averse agents take expectations over-weighting outcomes that would fall below expectations.
- Used by Routledge and Zin (2010 *JF*) in a Mehra-Prescott framework; Campanale, Castro, Clementi (2010 *RED*) in a model with capital accumulation.
    - Routledge and Zin use a 'generalized' form of DA that allows them some control over which states are disappointing.

Disappointment aversion: Formally

- More formally, the agent's intertemporal preferences have the Epstein-Zin form

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

  for $\rho \geq 0$, $\rho \neq 1$.
- The agent's IES is $1/\rho$.
- Conditionality in μ_t will come from Markov structure of consumption growth; for now, just ignore it.

Disappointment aversion: Formally

- 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

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

- $I(\,\cdot\,)$ is an indicator function, taking the value one when $z_i < \mu_t(z)$ (z_i disappoints), and zero otherwise.
- $\alpha \geq 0$, $\alpha \neq 1$ and $\theta \geq 0$ are parameters.

How might DA interact with RD?

- Suppose state $i = 1$ is the only disappointing state. Then:

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

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

$$\hat{p}_1 = \frac{p_1 + \theta p_1}{1+\theta p_1}$$
$$\hat{p}_j = \frac{p_j}{1+\theta p_1} \quad (j > 1)$$

- Think of state 1 as a rare disaster. Then, $p_1 = .01$ and $\theta = 1$ gives same c.e. as $p_1 = .02$ and $\theta = 0$.

DA and asset-pricing

- What does the SDF look like?
- Assume log consumption growth, x_{t+1}, follows a Markov chain, $\{x_i, P_{i,j}\}$.
- The SDF has the form:

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

  which is similar to the version given by Campanale *et al.*
- ϕ is lifetime utility scaled by consumption (v_t/c_t), and $I_{i,j} = 1$ if state j is disappointing conditional on state i today.

Roadmap of the talk

1. Introduction and motivation
2. Background
3. Model
4. Numerical experiments

Consumption process

- Starting point is Mehra and Prescott's Markov chain—two states, with gross consumption growth y_L and y_H, and symmetric transition matrix Q.
- Add a 'disaster' outcome y_D and modify the transition matrix as follows:
  - 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 $\{y_L, y_H\}$ with probabilities $\{1/2, 1/2\}$.
  - There is zero probability of remaining in the disaster state.
- Unlike Rietz, I assume Mehra-Prescott Markov chain describes "normal" (non-disaster) times.
  - Keep y_L, y_H, Q fixed as we vary p, y_D.

Markov chain

Assets

- I consider claims to a one-period 'riskless' asset, paying one unit of consumption next period (if no disaster), and two forms of 'equity.'
- Riskless asset is 'riskless' because I allow partial default in the disaster state. ▸ Default
- Two equity claims:
  - One is standard claim to aggregate consumption
  - The other (a 'dividend' claim) is a claim to a leveraged version of aggregate consumption. ▸ Leverage
- Returns are defined in the usual way. ▸ Returns

Roadmap of the talk

1. Introduction and motivation
2. Background
3. Model
4. Numerical experiments

Structure of the experiments

- Vary the environment by varying p, disaster probability, and $1 - y_D$, the disaster size. Find parameters of certainty equivalent—θ, strength of DA, and α, curvature—that get us close to first and second moments of returns data (Cochrane 2008 *RFS*).
    - I use discrete grids for θ and α.
    - $\theta \in \{0, 0.1, \dots, 2\}$, $\alpha \in \{0, 0.1, \dots 5\}$.
    - For comparison, Routledge and Zin mostly use $\theta = 9$. Choosing θ
- Model moments I compare to data are conditional on being outside the disaster state ('normal times').
- Focus on 'risk-free' return and dividend (leveraged consumption) return.

Structure of the experiments

- Starting point is average p and $1 - y_D$ from Barro's *QJE* paper:

$$\begin{aligned} p &= 0.017 \\ 1 - y_D &= 0.43 \end{aligned}$$

- Consider in turn (and in combinations) disaster probabilities and disaster sizes that are 1/2 or 1/4 the averages from Barro's data. I.e., p and $1 - y_D$ such that

$$\begin{aligned} p &\in \{0.017, 0.0085, 0.00425\} \\ 1 - y_D &\in \{0.43, 0.215, 0.1075\}. \end{aligned}$$

Fixing other parameters

- Some parameters fixed throughout experiments—discounting, leverage, default fraction, IES.
- In all the environments, data prefer a very low IES, which I fix at 0.1 (close to optimal).
- I follow Gourio (2008 *FRL*), in setting fraction of 'risk-free' asset defaulted on in disaster to be 40% of the disaster size.

Basic parameters

ρ^{-1}	β	λ	η
0.10	0.97	3	$0.4(1-y_D)$

Some results: Varying p

- Keep disaster size at 43%, consider p's of 1.7%, 0.85%, and 0.425%.

Best fits, varying disaster probability

p	θ	α	$\mathrm{E}(R^f)$	$\mathrm{E}(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
1.7%	0.2	2.1	1.19	7.30	5.41	19.15
0.85%	0.8	2.3	1.13	8.29	4.51	17.01
0.425%	1.3	2.1	1.11	8.69	4.01	15.75
Data	–	–	1.03	8.91	4.36	15.04

- Data are from Cochrane (2008 *RFS*).
- The $p = 0.425\%$ line is actually the best overall fit across all the experiments, based on a weighted Euclidean distance criterion (weights inversely proportional to the data).

Some results: Varying $1 - y_D$

- Keep disaster probability constant at 1.7%, consider $1 - y_D$'s of 43%, 21.5%, and 10.75%.

Best fits, varying disaster size

$1 - y_D$	θ	α	$E(R^f)$	$E(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
43%	0.2	2.1	1.19	7.30	5.41	19.15
21.5%	0.8	3.7	1.02	8.55	4.19	16.74
10.75%	1.4	3.9	1.09	8.88	3.67	15.25
Data	–	–	1.03	8.91	4.36	15.04

Some results: Varying both disaster size and probability

- Cut both disaster size and disaster probability by a half and a fourth.

Best fits, varying disaster size and probability

p	$1 - y_D$	θ	α	$E(R^f)$	$E(R^d)$	$\sigma(R^f)$	$\sigma(R^d)$
1.7%	43%	0.2	2.1	1.19	7.30	5.41	19.15
0.85%	21.5%	1.3	3.6	1.01	8.95	3.82	15.62
0.425%	10.75%	2.0	2.8	1.09	9.22	3.55	14.64
Data	–	–	–	1.03	8.91	4.36	15.04

Conclusions, extensions

- DA and RD interact in a complementary way, in terms of asset-pricing implications.
  - Very modest disaster probabilities or sizes can still produce realistic first and second moments of returns, if representative agent has a modest amount of DA.
  - And vice versa—modest amount of DA 'works,' if there's some modest disaster risk.
- If you look at the paper, it also has some exploratory results for a model with time-varying disaster risk (as in Gourio *FRL*).
  - Adds a slow-moving component to the pricing kernel.
  - Needs Routledge-Zin type "generalized" DA to get good fit.
  - Problem is, little or no hope of measuring that time variation.

The current version of the paper and Matlab code for solving the model are available at:
`www.jimdolmas.net/economics/current-work`

RD and the risk-free rate

- Recall that in representative agent, CRRA, lognormal case, risk-free rate obeys:

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

  where α is the CRRA coefficient, γ and σ are mean and standard deviation of log consumption growth.

- Well-known that for typical estimates of γ and σ, the first-order term $\alpha\gamma$ dominates as α increases.
  - A high value of α, necessary to get a non-negligible equity risk premium, thus produces a counterfactually high risk-free rate.

RD and the 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

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

- For typical values of γ and σ, and for small values of p, we can generate low values of R^f even for large values of α, and with $\beta < 1$.

Return

Consumption process: Markov chain

- Mehra and Prescott's two-state Markov chain for gross consumption growth y_{t+1} is:

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

with the transition matrix

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

Consumption process: Markov chain

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

$$y_{t+1} \in \begin{bmatrix} y_D \\ 0.982 \\ 1.054 \end{bmatrix}$$

and the transition matrix

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ p & (1-p)0.43 & (1-p)0.57 \\ p & (1-p)0.57 & (1-p)0.43 \end{bmatrix}$$

Return

The ‘riskless’ asset

- The ‘riskless’ asset here is a one-period zero-coupon bond (in zero net supply) whose price at t is q_t and whose payoff at $t+1$ is F_{t+1}, where

$$F_{t+1} = \begin{cases} 1-\eta & \text{if } y_{t+1} = y_D \\ 1 & \text{if } y_{t+1} \neq y_D \end{cases}$$

- In the experiments, I'll follow Gourio (2010 *FRL*) in making η a fraction of the disaster size.

Return

Dividend claim

- The 'dividend' claim is a claim to a process ('dividends') whose log growth rate is a multiple of log consumption growth.
- Letting y^d_{t+1} denote gross dividend growth from t to $t+1$,

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

where $\lambda \geq 1$ is the "leverage" parameter.

- This dividend formulation is now pretty standard—used, for example, in Routledge and Zin, Gourio, Bansal-Yaron, *etc.*

Return

Asset returns

- Returns are defined in the standard way.
- If q is the price of the 'riskless asset,' and w^c and w^d denote the price-dividend ratios for the two equity claims, then the key pricing equations are

$$
\begin{aligned}
q_t &= \mathrm{E}_t[m_{t+1} F_{t+1}] \\
w_t^c &= \mathrm{E}_t\left[m_{t+1} y_{t+1}\left(1 + w_{t+1}^c\right)\right] \\
w_t^d &= \mathrm{E}_t\left[m_{t+1} y_{t+1}^d\left(1 + w_{t+1}^d\right)\right],
\end{aligned}
$$

Asset returns

- Then, for Markov states i today and j tomorrow:

$$R^f_{i,j} = \frac{F_j}{q_i} \qquad \text{(Risk-free rate)}$$

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

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

where F_j is $1 - \eta$ if $y_j = y_D$, 1 otherwise.

Return

What's a moderate amount of DA?

- Some experimental evidence on size of θ.
  - Choi, Fisman, Gale and Kariv (2007 *AER*): estimate mean around 0.3, with ninety percent of estimates for individual subjects between 0 and around 1.4.
- Routledge and Zin use θ's from 9 (mostly) to 24 (in one case).
  - Also use $\delta \neq 1$, match first and second moments of returns data exactly.
- Campanale *et al.* use $\theta \in [0, 0.4]$, but assume $\beta > 1$.
  - Maybe it's a matter of taste, but I don't like assuming $\beta > 1$.
- The results I obtain are all with $\theta \in [0, 2]$ (and $\beta = 0.97$).

Return