Measuring the “Dark Matter” in Asset Pricing Models

Hui Chen\textsuperscript{1}  Winston Dou\textsuperscript{2}  Leonid Kogan\textsuperscript{1}

\textsuperscript{1}MIT and NBER
\textsuperscript{2}UPenn

MFM Summer Session

June 20, 2017
Outline

Overview

Fragility Measures

Finite-sample Interpretation

Applications

Conclusion
Outline

Overview

Fragility Measures

Finite-sample Interpretation

Applications
  Disaster Risk
  Long-Run Risk

Conclusion
“Dark matter”-based asset pricing

- Some asset pricing models rely heavily on “dark matter”:
  
  1. parameters are difficult to measure directly in the data;
  2. results are sensitive to how the “dark matter” is specified.

- Common defense:
  
  - Assumed parameter values/model structure are acceptable if not rejected by the data;
  
  - The fact that they help “explain” asset prices is viewed as evidence supporting such choices.
“Dark matter”-based asset pricing

- Some asset pricing models rely heavily on “dark matter”:
  (1) parameters are difficult to measure directly in the data;
  (2) results are sensitive to how the “dark matter” is specified.

- Common defense:
  - Assumed parameter values/model structure are acceptable if not rejected by the data;
  - The fact that they help “explain” asset prices is viewed as evidence supporting such choices.

- However, (1) + (2) imply model fragility:
  - Implicit degrees of freedom is large;
  - The model tends to over-fit the data in sample;
  - Poor out-of-sample performance.
A traditional approach: sensitivity analysis

- A model is robust if its results are “not sensitive to small perturbations” in parameter values.
A traditional approach: sensitivity analysis

- A model is robust if its results are "not sensitive to small perturbations" in parameter values.

- How to formalize:
  1. How small is "small" — how should we choose the size of perturbation?
  2. What does it mean for the results to be "not sensitive" to such perturbations?
  3. Difficult to extend sensitivity analysis to high-dimensional parameter vectors.
A traditional approach: sensitivity analysis

- A model is robust if its results are "not sensitive to small perturbations" in parameter values.

- How to formalize:
  1. How small is "small" — how should we choose the size of perturbation?
  2. What does it mean for the results to be "not sensitive" to such perturbations?
  3. Difficult to extend sensitivity analysis to high-dimensional parameter vectors.

- We formally define a measure of model robustness and establish a convenient way to compute it.
  - multivariate sensitivity analysis
  - formal connection to over-fitting tendency
Model fragility: informational approach

- Prices $P_t$ are linked to dividends $x_t$ through a pricing model:

  \[ P_t = f(x_t, \theta, \varepsilon_t) \]

  $\theta$: governs the dynamics of $x_t$ and affects stock prices

- We compare information about $\theta$ in
  
  (I) History of $x_t$
  
  (II) History of $\{x_t, P_t\}$ with cross-equation restrictions
Prices are sensitive to $\theta \Leftrightarrow$ prices are informative about $\theta$
(relative to fundamentals) $\Rightarrow$ sign of fragility
Outline

Overview

Fragility Measures

Finite-sample Interpretation

Applications
  Disaster Risk
  Long-Run Risk

Conclusion
Baseline model: specifies dynamics for $x_t$, with distribution $\mathcal{P}$

- indexed by baseline parameters $\theta$: $D_\Theta \times 1$
- true baseline parameter value is $\theta_0$
A Generic Structural Model

- **Baseline model**: specifies dynamics for $x_t$, with distribution $P$
  - indexed by baseline parameters $\theta$: $D_\Theta \times 1$
  - true baseline parameter value is $\theta_0$

- **Full model**: joint dynamics of $(x_t, y_t)$, with distribution $Q$
  - indexed by $\theta$ and nuisance parameter $\psi$: $D_\psi \times 1$
  - true nuisance parameter value is $\psi_0$
A Generic Structural Model

- **Baseline model:** specifies dynamics for $x_t$, with distribution $P$
  - indexed by baseline parameters $\theta$: $D_\Theta \times 1$
  - true baseline parameter value is $\theta_0$

- **Full model:** joint dynamics of $(x_t, y_t)$, with distribution $Q$
  - indexed by $\theta$ and nuisance parameter $\psi$: $D_\Psi \times 1$
  - true nuisance parameter value is $\psi_0$

- Model performance is summarized by a set of moment conditions (why not MLE?)
  - baseline sub-model: $\mathbb{E}[g_P(\theta_0; x_t)] = 0$
  - full model: $\mathbb{E}[g_Q(\theta_0, \psi_0; x_t, y_t)] = 0$
  - $g_P(\theta_0; x_t)$ is a sub-vector of $g_Q(\theta_0, \psi_0; x_t, y_t)$
Quick Review of GMM

- Empirical moment conditions for the full model:
  \[ \hat{g}_{Q,n}(\theta, \psi) \equiv \frac{1}{n} \sum_{t=1}^{n} g_{Q}(\theta, \psi; x_t, y_t) \]

- Efficient GMM estimator \((\hat{\theta}^Q, \hat{\psi}^Q)\) minimizes:
  \[ \hat{J}_{n,S_Q}(\theta, \psi) \equiv n\hat{g}_{Q,n}(\theta, \psi)^T S^{-1}_Q \hat{g}_{Q,n}(\theta, \psi) \]

- Spectral density matrix:
  \[ S_Q = \sum_{\ell=-\infty}^{+\infty} \mathbb{E}_Q \left[ g(\theta_0, \psi_0; x_t, y_t) g(\theta_0, \psi_0; x_{t-\ell}, y_{t-\ell})^T \right] \]

- Counterparts for the baseline model: \(\hat{g}_{P,n}(\theta), \hat{J}_{n,S_P}(\theta), S_P, \hat{\theta}^P\)
GMM Fisher Information Matrix

- **Baseline model:**

\[ I_P(\theta) \equiv G_P(\theta)^T S_P^{-1} G_P(\theta) \]

where \( G_P(\theta) \equiv \mathbb{E} [\nabla g_P(\theta; x_t)] \)

- **Full model:**

\[ I_Q(\theta, \psi) \equiv G_Q(\theta, \psi)^T S_Q^{-1} G_Q(\theta, \psi) \]

where \( G_Q(\theta, \psi) \equiv \mathbb{E} [\nabla g_Q(\theta, \psi; x_t, y_t)] \)
**Marginal Fisher Information Matrix**

- Efficient asymptotic variances of $\theta$ in the baseline model:

$$\nabla \mathcal{P} (\theta) = \mathbf{I}_\mathcal{P} (\theta)^{-1}$$
**Marginal Fisher Information Matrix**

- Efficient asymptotic variances of $\theta$ in the baseline model:

  $$\nabla_{P}(\theta) = I_{P}(\theta)^{-1}$$

- Efficient asymptotic variances of $\theta$ in the full model:

  $$\nabla_{Q}(\theta|\psi) = I_{Q}(\theta|\psi)^{-1}$$

- Marginal GMM Fisher information matrix:

  $$I_{Q}(\theta|\psi) \equiv I_{Q}^{(1,1)}(\theta, \psi) - \underbrace{I_{Q}^{(1,2)}(\theta, \psi)I_{Q}^{(2,2)}(\theta, \psi)I_{Q}^{(1,2)}(\theta, \psi)^{T}}_{\text{conditional information}} - \underbrace{I_{Q}^{(2,1)}(\theta, \psi)}_{\text{information loss due to uncertainty in } \psi}$$

  $$I_{Q}(\theta, \psi) = \begin{bmatrix} I_{Q}^{(1,1)}(\theta, \psi), & I_{Q}^{(1,2)}(\theta, \psi) \\ I_{Q}^{(2,1)}(\theta, \psi), & I_{Q}^{(2,2)}(\theta, \psi) \end{bmatrix}$$
Fisher Fragility Measure

Definition

\[ \varrho(\theta_0|\psi_0) = \text{tr} \left[ \mathbb{V}_Q(\theta_0|\psi_0)^{-1} \mathbb{V}_P(\theta_0) \right] \]
Fisher Fragility Measure

Definition

\[ \varrho(\theta_0|\psi_0) = \text{tr} \left[ \sqrt{Q}(\theta_0|\psi_0)^{-1}P(\theta_0) \right] \]

- Comparing estimation precision under full model and baseline model: \( \sqrt{Q}(\theta_0|\psi_0) \) vs. \( P(\theta_0) \)
- Informativeness of cross-equation restrictions
- \( \varrho(\theta_0|\psi_0) = \lambda_1 + \lambda_2 + \cdots + \lambda_{D_\theta} \)

\( \lambda_1 \geq \cdots \geq \lambda_{D_\theta} \) are eigenvalues of \( \sqrt{Q}(\theta_0|\psi_0)^{-1/2}P(\theta_0)\sqrt{Q}(\theta_0|\psi_0)^{-1/2} \)
THE “WORST-CASE” DIRECTIONS

Definition
For any $1 \leq D \leq D_\Theta$, 

$$\varrho^D(\theta_0|\psi_0) = \max_{v \in \mathbb{R}^{D \times D_\Theta}, \text{Rank}(v) = D} \text{tr} \left[ (v \vee_Q (\theta_0|\psi_0)v^T)^{-1} (v \vee_P (\theta_0)v^T) \right]$$
THE “WORST-CASE” DIRECTIONS

Definition
For any $1 \leq D \leq D_\Theta$, 

$$\varrho^D(\theta_0|\psi_0) = \max_{v \in \mathbb{R}^{D \times D_\Theta}, \text{Rank}(v)=D} \text{tr} \left[ (v \vee Q(\theta_0|\psi_0)v^T)^{-1} (v \vee P(\theta_0)v^T) \right]$$

- Search all directions $v$ for the largest discrepancy between $\vee Q(\theta_0|\psi_0)$, $\vee P(\theta_0)$
- “Effective sample size”: extra sample needed to match the variance of $v\theta$
- $\varrho^D(\theta_0|\psi_0) = \lambda_1 + \cdots + \lambda_D$
**WHY IS ϱ A MEASURE OF MODEL FRAGILITY?**

Over-fitting tendency of model Q relative to P:

\[
\varrho_o(\theta_0|\psi_0, x^n, y^n) \equiv \int d_{S_Q}\{\theta; x^n, y^n\} \pi_P(\theta|x^n) d\theta,
\]

- \(d_{S_Q}\{\theta; x^n, y^n\}\) measures the over-fitting for a particular model \(\theta\), based on \(J\)-distance (GMM analog of the log likelihood ratio)

\[
d_{S_Q}\{\theta; x^n, y^n\} = \hat{J}_{n,S_Q}(\theta, \hat{\psi}^Q) - \hat{J}_{n,S_Q}(\hat{\theta}^Q, \hat{\psi}^Q)
\]

J-distance gap \(\geq 0\)

\[
\rightarrow (\hat{\theta}^Q, \hat{\psi}^Q) \text{ is the GMM estimator}
\]

\[
\rightarrow (\theta, \hat{\psi}^Q) \text{ is the constrained GMM estimator with fixed } \theta
\]
Why is $\varrho$ a measure of model fragility?

Over-fitting tendency of model $Q$ relative to $P$:

$$\varrho_o(\theta_0|\psi_0, x^n, y^n) \equiv \int d_{SQ}\{\theta; x^n, y^n\} \pi_P(\theta|x^n) d\theta,$$

- $d_{SQ}\{\theta; x^n, y^n\}$ measures the over-fitting for a particular model $\theta$, based on $J$-distance (GMM analog of the log likelihood ratio)

$$d_{SQ}\{\theta; x^n, y^n\} = \widehat{J}_{n,S_Q}(\theta, \psi_Q^Q) - \widehat{J}_{n,S_Q}(\hat{\theta}^Q, \hat{\psi}_Q^Q)$$

\[ \text{J-distance gap} \geq 0 \]

- $(\hat{\theta}_Q^Q, \hat{\psi}_Q^Q)$ is the GMM estimator

- $(\theta, \psi_Q^Q)$ is the constrained GMM estimator with fixed $\theta$

- Averaged over $\pi_P(\theta|x^n)$, the posterior of $\theta$ in the baseline model.

- This measure extends the DIC measure for over-fitting tendency in statistics (Spiegelhalter et al. 2002)
**Equivalence Results**

**Theorem:** Under standard regularity conditions,

\[
\lim_{n \to \infty} \varrho_o(\theta_0|\psi_0, x^n, y^n) = \varrho(\theta_0|\psi_0) + \sum_{i=1}^{D_\Theta} \left[ \lambda_i - 1 \right] \chi^2_{i,1}
\]
Theorem: Under standard regularity conditions,

\[
\lim_{n \to \infty} \varrho_o(\theta_0|\psi_0, x^n, y^n) = \varrho(\theta_0|\psi_0) + \sum_{i=1}^{D_\Theta} [\lambda_i - 1] \chi_{i,1}^2
\]

- "wlim": a variable with the limiting distribution (in the sense of weak convergence)
- \(\chi_{1,i}^2\): i.i.d. chi-squared random variables with 1 degree of freedom
AVERAGE TENDENCY OF OVER-FITTING

\[ \mathbb{E} \left[ \operatorname{wlim}_{n \to \infty} \varrho_o(\theta_0|\psi_0, x^n, y^n) \right] = 2\varrho(\theta_0|\psi_0) - D_{\Theta} \]

Interpretation:

- \( \varrho(\theta_0|\psi_0) \) is the implicit degrees of freedom, while \( D_{\Theta} \) is the standard explicit degrees of freedom.

- \( \varrho(\theta_0|\psi_0) \) captures the average tendency of over-fitting when sample size is large.

- When bootstrapping based on a large sample \((x^n, y^n)\) to estimate the average tendency of over-fitting, the result should be close to \( 2\varrho(\theta_0|\psi_0) - D_{\Theta} \).
Tail distribution for $\varrho$

Tail probability of the limiting variable converges to zero at an exponential rate:

$$\lim_{x \to \infty} \frac{1}{x} \ln P \left\{ \lim_{n \to \infty} \varrho_o(\theta_0|\psi_0, x^n, y^n) > x \right\} = -\frac{1}{2(\lambda_1 - 1)}$$

Interpretation:

- The thickness of the tail distribution is fully characterized by the largest eigenvalue $\lambda_1$, which is the worst-case 1-D Fisher fragility measure $\varrho^1(\theta_0|\psi_0)$.

- Given the average level $\varrho(\theta_0|\psi_0)$, the more the distribution of eigenvalues is concentrated on $\lambda_1$, the higher the probability of extreme over-fitting.
Outline

Overview

Fragility Measures

Finite-sample Interpretation

Applications
  Disaster Risk
  Long-Run Risk

Conclusion
INFORMATIVENESS OF ECONOMIC RESTRICTIONS

- In finite sample, we can compare posterior beliefs about $\theta$
  - Compare the entire posterior distributions
  - Quantify the additional information in cross-equation restrictions

- Prior and Posteriors
  - $\pi(\theta)$: the common prior about $\theta$
  - $\pi_P(\theta|x^n)$: posterior based on baseline model and $x^n$
  - $\pi_Q(\theta|x^n, y^n)$: posterior based on full model and $(x^n, y^n)$

- Use relative entropy to measure the statistical discrepancy between $\pi_P(\theta|x^n)$ and $\pi_Q(\theta|x^n, y^n)$:

$$D_{KL}(\pi_Q(\theta|x^n, y^n) \| \pi_P(\theta|x^n)) = \int \ln \left( \frac{\pi_Q(\theta|x^n, y^n)}{\pi_P(\theta|x^n)} \right) \pi_Q(\theta|x^n, y^n) \, d\theta$$
**Effective Sample Size**

- Relative entropy is difficult to interpret. We normalize it by asking:
  - How much extra data \((m^*)\) is needed on average to generate the same amount of additional information about \(\theta\) based only on the baseline model.

- The additional information provided by “effective sample” is measured by the mutual information between \(\theta\) and \(\tilde{x}^m\) conditional observed data \(x^n\):
  \[
  I(\tilde{x}^m; \theta|x^n) = \mathbb{E}_{\tilde{x}^m|x^n} \left[ D_{KL}(\pi_{\hat{P}}(\theta'|\tilde{x}^m, x^n) || \pi_{\hat{P}}(\theta'|x^n)) \right]
  \]

- Effective sample size ratio:
  \[
  \rho_{KL}(x^n, y^n) = \frac{n + m^*}{n}
  \]
  such that
  \[
  D_{KL}(\pi_Q(\theta|x^n, y^n) || \pi_{\hat{P}}(\theta|x^n)) = I(\tilde{x}^{m*}; \theta|x^n)
  \]
ILLUSTRATION: SOLVING FOR $\varrho_{KL}$

Extra sample: $(n + m)/n$

$D_{KL}(\pi_Q(\theta|x^n, y^n) || \pi_P(\theta|x^n))$

$ightarrow I(\tilde{x}^m; \theta|x^n)$

$ightarrow \varrho_{KL}$
ASYMPTOTIC EQUIVALENCE

Under standard regularity conditions,

$$\mathbb{E}\left[\operatorname{wlim}_{n \to \infty} \ln \rho_{KL}(x^n, y^n)\right] = \ln \rho(\theta_0 | \psi_0)$$

Interpretation:

- $\rho(\theta_0 | \psi_0)$ captures the additional effective sample size under a rigorous information-theoretic framework.

- It further justifies the economic interpretation: informativeness of cross-equation restrictions implied by theories.
Outline

Overview

Fragility Measures

Finite-sample Interpretation

Applications
  Disaster Risk
  Long-Run Risk

Conclusion
Disaster risk model

- **Disaster:** \( z_t \in (0, 1), \ Pr(z_t = 1) = p \)

- **Baseline model** for (log) consumption growth \( g_t \):
  - \( z_t = 0: g_t \sim N(\mu, \sigma^2) \)
  - \( z_t = 1: g_t = -v_t \), where \( v_t \sim 1\{v_t > v\} \xi e^{-\xi(v_t-v)} \)

- **Full model:** adds excess log return on the market portfolio \( r_t \)
  - \( z_t = 0: r_t \) and \( g_t \) are jointly normal,
  \[
  r_t = \eta + \rho \frac{\tau}{\sigma} (g_t - \mu) + \sqrt{1 - \rho^2 \tau} \epsilon_{0,t},
  \]
  - \( z_t = 1: \)
  \[
  r_t = bg_t + \zeta \epsilon_{1,t}
  \]
  - Representative agent with power utility: \( \gamma \)

- **“Dark matter” parameters:** \( p, \xi \)
Cross-equation restriction

- Consumption Euler equation for excess log returns

\[ 1 = \mathbb{E}_t \left[ \frac{m_{t+1}}{\mathbb{E}_t[m_{t+1}]} e^{r_{t+1}} \right] \]

- The log equity premium is then given by

\[ \mathbb{E}[r_t] = (1 - p) \eta - pb \left( v + 1/\xi \right), \]

\[ \eta \approx \gamma \rho \sigma \tau - \frac{\tau^2}{2} + e^{\gamma (\mu - \gamma^2 \sigma^2/2)} \xi \left( \frac{e^{\gamma v}}{\xi - \gamma} - \frac{e^{\xi^2/2 + (\gamma - b)v}}{\xi + b - \gamma} \right) \frac{p}{1 - p} \]

- Key properties:

1. \( \eta \) is unbounded as \( \xi \to \gamma \) (no matter how small \( p \) is)
2. Large disaster size (small \( \xi \)) \( \implies \) high sensitivity of \( \eta \) to \( p \)
Disaster risk model: An Irrefutable Model?

“Dark matter” parameters: \( p, \xi \)
A. $\gamma = 3, p = 3.96\%, \xi = 4.649$

B. $\gamma = 24, p = 1.81\%, \xi = 446.36$

C. $\gamma = 3, p = 0.31\%, \xi = 3.179$

D. $\gamma = 24, p = 0.0699\%, \xi = 28.43$
**Long-Run Risk Model**

- Consumption processes:

  \[
  \Delta c_{t+1} = \mu_c + x_t + \sigma_t \epsilon_{c,t+1}
  \]

  \[
  x_{t+1} = \rho x_t + \phi_x \sigma_t \epsilon_{x,t+1}
  \]

  \[
  \tilde{\sigma}^2_{t+1} = \bar{\sigma}^2 + \nu(\tilde{\sigma}^2_t - \bar{\sigma}^2) + \sigma_w \epsilon_{\sigma,t+1}
  \]

  \[
  \sigma^2_{t+1} = \max(\sigma^2, \tilde{\sigma}^2_{t+1})
  \]

- Dividend process:

  \[
  \Delta d_{t+1} = \mu_d + \phi_d x_t + \phi_{d,c} \sigma_t \epsilon_{c,t+1} + \phi_{d,d} \sigma_t \epsilon_{d,t+1}
  \]

- Representative agent with EZW preference: \( \delta_L, \gamma_L \) and \( \psi_L \)

- From the consumption Euler equation, derive a linear approximation of the stochastic discount factor

  \[
  m_{t+1} = \Gamma_0 + \Gamma_1 x_t + \Gamma_2 \sigma^2_t - \lambda_c \sigma_t \epsilon_{c,t+1} - \lambda_x \phi_x \sigma_t \epsilon_{x,t+1} - \lambda_\sigma \sigma_w \epsilon_{\sigma,t+1}
  \]
In equilibrium, the excess (log) market return can be written as

\[ r_{m,t+1}^e = \mu_{r,t}^e + \beta_\eta \sigma_t \eta_{t+1} + \beta_e \sigma_t e_{t+1} + \beta_w \sigma_w w_{t+1} + \varphi_{d,d} \sigma_t u_{d,t+1} \]

Loadings on shocks

The conditional expected excess (log) return \( \mu_{r,t}^e = \mathbb{E}_t [r_{m,t+1}^e] \) is

\[ \mu_{r,t}^e = \lambda_\eta \beta_\eta \sigma_{p,t}^2 + \lambda_e \beta_e \sigma_{p,t}^2 + \lambda_w \beta_w \sigma_w^2 - \frac{1}{2} \left( \beta_\eta^2 \sigma_t^2 + \beta_e^2 \sigma_t^2 + \beta_w^2 \sigma_w^2 + \varphi_{d,d}^2 \sigma_t^2 \right) \]

\( \beta \)'s and \( \lambda \)'s are determined as fixed points in the equilibrium.
Baseline Model Choice

- Baseline model: $x_t = (\Delta c_{t+1}, x_t, \sigma^2_t, \Delta d_{t+1})$
  - baseline parameters $\theta = (\mu_c, \rho, \phi_x, \overline{\sigma}^2, \nu, \sigma_w, \mu_d, \phi_d, \phi_{d,c}, \phi_{d,d})$

- Full model: asset pricing model with $y_t = r^e_{m,t+1}$
  - nuisance parameters $\psi = (\gamma_L, \psi_L)$
## Model I

- Bansal-Kiku-Yaron (2012) parameters

<table>
<thead>
<tr>
<th></th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences</td>
<td>$\delta_L$</td>
<td>$\gamma_L$</td>
<td>$\psi_L$</td>
</tr>
<tr>
<td>Consumption</td>
<td>$\mu_c$</td>
<td>$\rho$</td>
<td>$\phi_e$</td>
</tr>
<tr>
<td>Dividends</td>
<td>$\mu_d$</td>
<td>$\psi$</td>
<td>$\pi$</td>
</tr>
<tr>
<td>Other</td>
<td>$\sigma$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Does a good job matching asset pricing moments

Simulated and sample moments for the benchmark calibration:

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data Estimate</th>
<th>Model 5%</th>
<th>Model Median</th>
<th>Model 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E}[r_M - r_f]$</td>
<td>7.09</td>
<td>2.33</td>
<td>5.88</td>
<td>10.58</td>
</tr>
<tr>
<td>$\mathbb{E}[r_M]$</td>
<td>7.66</td>
<td>2.91</td>
<td>6.66</td>
<td>11.20</td>
</tr>
<tr>
<td>$\sigma(r_M)$</td>
<td>20.28</td>
<td>12.10</td>
<td>20.99</td>
<td>29.11</td>
</tr>
<tr>
<td>$\mathbb{E}[r_f]$</td>
<td>0.57</td>
<td>-0.20</td>
<td>0.77</td>
<td>1.45</td>
</tr>
<tr>
<td>$\sigma(r_f)$</td>
<td>2.86</td>
<td>0.64</td>
<td>1.07</td>
<td>1.62</td>
</tr>
<tr>
<td>$\mathbb{E}[p - d]$</td>
<td>3.36</td>
<td>2.69</td>
<td>2.99</td>
<td>3.30</td>
</tr>
</tbody>
</table>
## Model II (Alternative Calibration)

<table>
<thead>
<tr>
<th>Preferences</th>
<th>$\delta_L$</th>
<th>$\gamma_L$</th>
<th>$\psi_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.9989</td>
<td>27</td>
<td>1.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consumption</th>
<th>$\mu_c$</th>
<th>$\rho$</th>
<th>$\varphi_e$</th>
<th>$\overline{\sigma}$</th>
<th>$\nu$</th>
<th>$\sigma_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0015</td>
<td>0.975</td>
<td>0.038</td>
<td>0.0072</td>
<td>0.98</td>
<td>2.8e−6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dividends</th>
<th>$\mu_d$</th>
<th>$\psi$</th>
<th>$\pi$</th>
<th>$\varphi$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0015</td>
<td>2.5</td>
<td>2.6</td>
<td>5.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Also matches asset pricing moments

Simulated and sample moments for alternative calibration:

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E}[r_M - r_f]$</td>
<td>7.09</td>
<td>3.65</td>
</tr>
<tr>
<td>$\mathbb{E}[r_M]$</td>
<td>7.66</td>
<td>4.42</td>
</tr>
<tr>
<td>$\sigma (r_M)$</td>
<td>20.28</td>
<td>15.01</td>
</tr>
<tr>
<td>$\mathbb{E}[r_f]$</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>$\sigma (r_f)$</td>
<td>2.86</td>
<td>0.73</td>
</tr>
<tr>
<td>$\mathbb{E}[p - d]$</td>
<td>3.36</td>
<td>2.77</td>
</tr>
</tbody>
</table>
But they are very different in terms of fragility

<table>
<thead>
<tr>
<th>Model</th>
<th>$\rho$</th>
<th>$\rho^1$</th>
<th>$\psi^\nu$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_c$</td>
<td>$\rho$</td>
<td>$\varphi_x$</td>
</tr>
<tr>
<td>(M1)</td>
<td>276.3</td>
<td>196.3</td>
<td>1.0</td>
</tr>
<tr>
<td>(M2)</td>
<td>34.0</td>
<td>21.1</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. Nuisance parameter vector $\psi$: $(\gamma_L, \psi_L)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M1)</td>
<td>$3.58 \cdot 10^5$</td>
<td>$3.57 \cdot 10^5$</td>
<td>1.0</td>
</tr>
<tr>
<td>(M2)</td>
<td>323.3</td>
<td>287.7</td>
<td>1.0</td>
</tr>
<tr>
<td>II. Nuisance parameter vector $\psi$: empty</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Diagnosing the sources of fragility

![Graph showing the relationship between Eigenvalues and Rank of 1-D Subspaces. The graph has a downward trend, indicating that as the rank increases, the eigenvalues decrease.]
Outline

Overview

Fragility Measures

Finite-sample Interpretation

Applications
  Disaster Risk
  Long-Run Risk

Conclusion
What does it mean when the cross-equation restrictions in a model appear highly informative?

- Powerful identification if economic restrictions are valid.
- Sign of fragility
  
  Why should we be concerned: prone to over-fitting in sample, poor out-of-sample performance, problematic implications for counter-factual analysis ...
HOW TO USE THE FRAgilITY MEASURE?

■ To improve model selection and inference
  ← Model selection: measure the fragility of a class of models
  ← Structural estimation: guard against mis-specification and false confidence in inference

■ Model diagnostics: To identify aspects of the model that require more data or theory support.
  ← International evidence? Survey data?
  ← If it is about agents’ beliefs, can they be micro-founded, e.g. via preferences, information aggregation, market frictions?
  ← Explicitly analyze agents’ uncertainties inside the model (Hansen-Sargent)

■ Method can be applied to structural models in many fields (structural corporate, IO, macro, ...)