Can Rare Events Explain the Equity Premium Puzzle?

Christian Julliard* and Anisha Ghosh◇

*Department of Economics and FMG
London School of Economics, and CEPR

◇Department of Economics
London School of Economics

Federal Reserve Bank of New York, February 4th 2008

Equity Premium Puzzle and Rare Events

The Premium: in the historical data, the U.S. stock market excess return over a risk free asset has been over 7.4% a year

The Puzzle: time separable CRRA utility with a RRA of 10 implies a risk premium of less than 1% a year (e.g. Mehra and Prescott (1985))

- higher RRA is unrealistic: risk-free puzzle; certainty equivalent paradox; micro evidence.

The Rare Events Explanation: (Rietz (1988))

- Equity owners demand high return to compensate for extreme losses they may incur during **unlikely, but severe, economic downturns and market crashes**.
- If returns have been high with too few of these events, equity owners have been compensated for events that did not occur.

⇒ If in a given period these events occur with a frequency smaller than their true probability, investors will appear irrational and economists will misestimate their preferences.

Equity Premium Puzzle and Rare Events

The Premium: in the historical data, the U.S. stock market excess return over a risk free asset has been over 7.4% a year

The Puzzle: time separable CRRA utility with a RRA of 10 implies a risk premium of less than 1% a year (e.g. Mehra and Prescott (1985))

- higher RRA is unrealistic: risk-free puzzle; certainty equivalent paradox; micro evidence.

The Rare Events Explanation: (Rietz (1988))

- Equity owners demand high return to compensate for extreme losses they may incur during **unlikely, but severe, economic downturns and market crashes**.
- If returns have been high with too few of these events, equity owners have been compensated for events that did not occur.

⇒ If in a given period these events occur with a frequency smaller than their true probability, investors will appear irrational and economists will misestimate their preferences.

Equity Premium Puzzle and Rare Events

The Premium: in the historical data, the U.S. stock market excess return over a risk free asset has been over 7.4% a year

The Puzzle: time separable CRRA utility with a RRA of 10 implies a risk premium of less than 1% a year (e.g. Mehra and Prescott (1985))

- higher RRA is unrealistic: risk-free puzzle; certainty equivalent paradox; micro evidence.

The Rare Events Explanation: (Rietz (1988))

- Equity owners demand high return to compensate for extreme losses they may incur during **unlikely, but severe, economic downturns and market crashes**.
- If returns have been high with too few of these events, equity owners have been compensated for events that did not occur.

⇒ If in a given period these events occur with a frequency smaller than their true probability, investors will appear irrational and economists will misestimate their preferences.

Equity Premium Puzzle and Rare Events

The Premium: in the historical data, the U.S. stock market excess return over a risk free asset has been over 7.4% a year

The Puzzle: time separable CRRA utility with a RRA of 10 implies a risk premium of less than 1% a year (e.g. Mehra and Prescott (1985))

- higher RRA is unrealistic: risk-free puzzle; certainty equivalent paradox; micro evidence.

The Rare Events Explanation: (Rietz (1988))

- Equity owners demand high return to compensate for extreme losses they may incur during **unlikely, but severe, economic downturns and market crashes**.
- If returns have been high with too few of these events, equity owners have been compensated for events that did not occur.

⇒ If in a given period these events occur with a frequency smaller than their true probability, investors will appear irrational and economists will misestimate their preferences.

Equity Premium Puzzle and Rare Events

The Premium: in the historical data, the U.S. stock market excess return over a risk free asset has been over 7.4% a year

The Puzzle: time separable CRRA utility with a RRA of 10 implies a risk premium of less than 1% a year (e.g. Mehra and Prescott (1985))

- higher RRA is unrealistic: risk-free puzzle; certainty equivalent paradox; micro evidence.

The Rare Events Explanation: (Rietz (1988))

- Equity owners demand high return to compensate for extreme losses they may incur during **unlikely, but severe, economic downturns and market crashes**.
- If returns have been high with too few of these events, equity owners have been compensated for events that did not occur.

⇒ If in a given period these events occur with a frequency smaller than their true probability, investors will appear irrational and economists will misestimate their preferences.

Outline

Outline

Rare Events – Related Literature

"A throw of dice will never abolish chance." Mallarmé (1897)

- **Stock markets don't like the CLT:** Mandelbrot (1962, 1963), Mandelbrot-Taylor (1967) ...

$\Rightarrow$ Jump and Lévy price processes, min-max, extreme value theory and tail-related risk measures

- **Rare Events and the EPP:** Rietz (1988), Barro (2005), Danthine-Donaldson (1999), Copeland-Zhu (2006), Gabaix (2007) $\Rightarrow$ all calibration exercises

"Perhaps just as puzzling as the high equity premium is why Rietz's framework has not been taken more seriously." Barro (2005)

- **RE and GMM:** Saikkonen-Ripatti (2000).
- **RE and Learning:** Sandroni (1998), Veronesi (2004), Liu et al. (2005), Weitzman (2007).
- **RE, Term Structure and more:** Lewis(1990), Bekaert et al.(2001), Gourinchas-Tornell(2004), Lopes-Michaelides(2005), Gabaix-Fahri(2007)

Rare Events – Related Literature

"A throw of dice will never abolish chance." Mallarmé (1897)

- **Stock markets don't like the CLT:** Mandelbrot (1962, 1963), Mandelbrot-Taylor (1967) ...

⇒ Jump and Lévy price processes, min-max, extreme value theory and tail-related risk measures

- **Rare Events and the EPP:** Rietz (1988), Barro (2005), Danthine-Donaldson (1999), Copeland-Zhu (2006), Gabaix (2007) ⇒ all calibration exercises

"Perhaps just as puzzling as the high equity premium is why Rietz's framework has not been taken more seriously." Barro (2005)

- **RE and GMM:** Saikkonen-Ripatti (2000).
- **RE and Learning:** Sandroni (1998), Veronesi (2004), Liu et al. (2005), Weitzman (2007).
- **RE, Term Structure and more:** Lewis(1990), Bekaert et al.(2001), Gourinchas-Tornell(2004), Lopes-Michaelides(2005), Gabaix-Fahri(2007)

Rare Events – Related Literature

"A throw of dice will never abolish chance." Mallarmé (1897)

- **Stock markets don't like the CLT:** Mandelbrot (1962, 1963), Mandelbrot-Taylor (1967) ...

$\Rightarrow$ Jump and Lévy price processes, min-max, extreme value theory and tail-related risk measures

- **Rare Events and the EPP:** Rietz (1988), Barro (2005), Danthine-Donaldson (1999), Copeland-Zhu (2006), Gabaix (2007) $\Rightarrow$ all calibration exercises

"Perhaps just as puzzling as the high equity premium is why Rietz's framework has not been taken more seriously." Barro (2005)

- **RE and GMM:** Saikkonen-Ripatti (2000).
- **RE and Learning:** Sandroni (1998), Veronesi (2004), Liu et al. (2005), Weitzman (2007).
- **RE, Term Structure and more:** Lewis(1990), Bekaert et al.(2001), Gourinchas-Tornell(2004), Lopes-Michaelides(2005), Gabaix-Fahri(2007)

Rare Events – Related Literature

"A throw of dice will never abolish chance." Mallarmé (1897)

- **Stock markets don't like the CLT:** Mandelbrot (1962, 1963), Mandelbrot-Taylor (1967) ...

$\Rightarrow$ Jump and Lévy price processes, min-max, extreme value theory and tail-related risk measures

- **Rare Events and the EPP:** Rietz (1988), Barro (2005), Danthine-Donaldson (1999), Copeland-Zhu (2006), Gabaix (2007) $\Rightarrow$ all calibration exercises

"Perhaps just as puzzling as the high equity premium is why Rietz's framework has not been taken more seriously." Barro (2005)

- **RE and GMM:** Saikkonen-Ripatti (2000).
- **RE and Learning:** Sandroni (1998), Veronesi (2004), Liu et al. (2005), Weitzman (2007).
- **RE, Term Structure and more:** Lewis(1990), Bekaert et al.(2001), Gourinchas-Tornell(2004), Lopes-Michaelides(2005), Gabaix-Fahri(2007)

Rare Events – Related Literature

"A throw of dice will never abolish chance." Mallarmé (1897)

- **Stock markets don't like the CLT:** Mandelbrot (1962, 1963), Mandelbrot-Taylor (1967) ...

$\Rightarrow$ Jump and Lévy price processes, min-max, extreme value theory and tail-related risk measures

- **Rare Events and the EPP:** Rietz (1988), Barro (2005), Danthine-Donaldson (1999), Copeland-Zhu (2006), Gabaix (2007) $\Rightarrow$ all calibration exercises

"Perhaps just as puzzling as the high equity premium is why Rietz's framework has not been taken more seriously." Barro (2005)

- **RE and GMM:** Saikkonen-Ripatti (2000).
- **RE and Learning:** Sandroni (1998), Veronesi (2004), Liu et al. (2005), Weitzman (2007).
- **RE, Term Structure and more:** Lewis(1990), Bekaert et al.(2001), Gourinchas-Tornell(2004), Lopes-Michaelides(2005), Gabaix-Fahri(2007)

Outline

Sample Analogs and Rare Events

- The CCAPM of Rubinstein (1976) and Breeden (1979) implies

$$0 = E\left[m_t(\gamma_0)R^e_{i,t}\right] \equiv \int m_t(\gamma_0)R^e_{i,t}dF \qquad (1)$$

where $m_t = (C_t/C_{t-1})^{-\gamma}$ is the pricing kernel, γ is the RRA parameter, $R^e_{i,t}$ is the return on the risk asset i in excess of the risk-free rate, and F is the true distribution of the data.

- The standard approach is to estimate γ_0 as

$$\hat{\gamma} := \arg\min g\left(E^T\left[m_t(\gamma)\right], E^T\left[R^e_{i,t}\right], E^T\left[m_t(\gamma), R^e_{i,t}\right]\right)$$

for some function $g(.)$, where $E^T[x_t] = \frac{1}{T}\sum_{t=1}^T x_t$, and then judge whether $\hat{\gamma}$ (or some function of it) is "reasonable"

- $E^T[.]$ justified by WLLN+CLT $\longrightarrow$ problem with rare events

$\Rightarrow$ if in a given sample extreme events happened to occur with a frequency smaller than their true probability, preferences might be misestimated.

Sample Analogs and Rare Events

- The CCAPM of Rubinstein (1976) and Breeden (1979) implies

$$0 = E\left[m_t(\gamma_0) R^e_{i,t}\right] \equiv \int m_t(\gamma_0) R^e_{i,t} dF \qquad (1)$$

where $m_t = (C_t / C_{t-1})^{-\gamma}$ is the pricing kernel, γ is the RRA parameter, $R^e_{i,t}$ is the return on the risk asset i in excess of the risk-free rate, and F is the true distribution of the data.

- The standard approach is to estimate γ_0 as

$$\hat{\gamma} := \arg\min g\left(E^T\left[m_t(\gamma)\right], E^T\left[R^e_{i,t}\right], E^T\left[m_t(\gamma), R^e_{i,t}\right]\right)$$

for some function $g(.)$, where $E^T[x_t] = \frac{1}{T}\sum_{t=1}^T x_t$, and then judge whether $\hat{\gamma}$ (or some function of it) is "reasonable"

- $E^T[.]$ justified by WLLN+CLT $\rightarrow$ problem with rare events

$\Rightarrow$ if in a given sample extreme events happened to occur with a frequency smaller than their true probability, preferences might be misestimated.

Sample Analogs and Rare Events

- The CCAPM of Rubinstein (1976) and Breeden (1979) implies

$$0 = E\left[m_t(\gamma_0) R^e_{i,t}\right] \equiv \int m_t(\gamma_0) R^e_{i,t} dF \qquad (1)$$

where $m_t = (C_t/C_{t-1})^{-\gamma}$ is the pricing kernel, γ is the RRA parameter, $R^e_{i,t}$ is the return on the risk asset i in excess of the risk-free rate, and F is the true distribution of the data.

- The standard approach is to estimate γ_0 as

$$\hat{\gamma} := \arg\min g\left(E^T[m_t(\gamma)], E^T\left[R^e_{i,t}\right], E^T\left[m_t(\gamma), R^e_{i,t}\right]\right)$$

for some function $g(.)$, where $E^T[x_t] = \frac{1}{T}\sum_{t=1}^T x_t$, and then judge whether $\hat{\gamma}$ (or some function of it) is "reasonable"

- $E^T[.]$ justified by WLLN+CLT $\longrightarrow$ problem with rare events

$\Rightarrow$ if in a given sample extreme events happened to occur with a frequency smaller than their true probability, preferences might be misestimated.

Sample Analogs and Rare Events

- The CCAPM of Rubinstein (1976) and Breeden (1979) implies

$$0 = E\left[m_t(\gamma_0) R^e_{i,t}\right] \equiv \int m_t(\gamma_0) R^e_{i,t} dF \qquad (1)$$

where $m_t = (C_t / C_{t-1})^{-\gamma}$ is the pricing kernel, γ is the RRA parameter, $R^e_{i,t}$ is the return on the risk asset i in excess of the risk-free rate, and F is the true distribution of the data.

- The standard approach is to estimate γ_0 as

$$\hat{\gamma} := \arg\min g\left(E^T[m_t(\gamma)], E^T\left[R^e_{i,t}\right], E^T\left[m_t(\gamma), R^e_{i,t}\right]\right)$$

for some function $g(.)$, where $E^T[x_t] = \frac{1}{T}\sum_{t=1}^{T} x_t$, and then judge whether $\hat{\gamma}$ (or some function of it) is "reasonable"

- $E^T[.]$ justified by WLLN+CLT $\longrightarrow$ problem with rare events

$\Rightarrow$ if in a given sample extreme events happened to occur with a frequency smaller than their true probability, preferences might be misestimated.

Information-Theoretic Alternatives: Empirical Likelihood

- Consider the model

$$E\left[f(z_t;\theta_0)\right] \equiv \int f(z_t;\theta_0)d\mu = \underline{0}, \quad \theta \in \Theta \subset \mathbb{R}^s \qquad (2)$$

where f is a known $\mathbb{R}^q$-valued function, $z_t \in \mathbb{R}^k$, $q \geqslant s$.

- We observe draws of $\{z_t\}_{t=1}^T$, from the unknown measure μ.
- Let $\Delta := \left\{(p_1, ..., p_T) : \sum_{t=1}^T p_t = 1, \, p_t \geq 0, \, t = 1, ..., T\right\}$, the nonparametric log likelihood at $(p_1, ..., p_T)$ is

$$\ell_{NP}(p_1, p_2, ..., p_T) = \sum_{t=1}^T \log(p_t), \; (p_1, ..., p_T) \in \Delta$$

- The EL estimator (Owen (1988)), $\left(\widehat{\theta}_{EL}, \widehat{p}_1^{EL}, ..., \widehat{p}_T^{EL}\right)$, solves

$$\max_{\{\theta, p_1, ..., p_T\} \in \Theta \times \Delta} \ell_{NP} = \sum_{t=1}^T \log(p_t) \quad \text{subject to} \quad \sum_{t=1}^T f(z_t;\theta)p_t = 0$$

- The NPMLE of μ is $\hat{\mu}_{EL} = \sum_{t=1}^T \widehat{p}_t^{EL} \delta_{z_t}$ ($\delta_z = 1$ at z).

Information-Theoretic Alternatives: Empirical Likelihood

- Consider the model

$$E\left[f(z_t;\theta_0)\right] \equiv \int f(z_t;\theta_0)d\mu = \underline{0}, \quad \theta \in \Theta \subset \mathbb{R}^s \qquad (2)$$

where f is a known $\mathbb{R}^q$-valued function, $z_t \in \mathbb{R}^k$, $q \geqslant s$.

- We observe draws of $\{z_t\}_{t=1}^{T}$, from the unknown measure μ.
- Let $\Delta := \left\{(p_1, ..., p_T) : \sum_{t=1}^{T} p_t = 1, \, p_t \geq 0, \, t = 1, ..., T\right\}$, the nonparametric log likelihood at $(p_1, ..., p_T)$ is

$$\ell_{NP}(p_1, p_2, ..., p_T) = \sum_{t=1}^{T} \log(p_t), \; (p_1, ..., p_T) \in \Delta$$

- The EL estimator (Owen (1988)), $\left(\widehat{\theta}_{EL}, \widehat{p}_1^{EL}, ..., \widehat{p}_T^{EL}\right)$, solves

$$\max_{\{\theta, p_1, ..., p_T\} \in \Theta \times \Delta} \ell_{NP} = \sum_{t=1}^{T} \log(p_t) \quad \text{subject to} \quad \sum_{t=1}^{T} f(z_t;\theta)p_t = 0$$

- The NPMLE of μ is $\hat{\mu}_{EL} = \sum_{t=1}^{T} \widehat{p}_t^{EL} \delta_{z_t}$ ($\delta_z = 1$ at z).

Information-Theoretic Alternatives: Empirical Likelihood

- Consider the model

$$E\left[f(z_t;\theta_0)\right] \equiv \int f(z_t;\theta_0)d\mu = \underline{0}, \quad \theta \in \Theta \subset \mathbb{R}^s \qquad (2)$$

where f is a known $\mathbb{R}^q$-valued function, $z_t \in \mathbb{R}^k$, $q \geqslant s$.

- We observe draws of $\{z_t\}_{t=1}^T$, from the unknown measure μ.
- Let $\Delta := \left\{(p_1, ..., p_T) : \ \sum_{t=1}^T p_t = 1,\ p_t \geq 0,\ t = 1, ..., T\right\}$,

the nonparametric log likelihood at $(p_1, ..., p_T)$ is

$$\ell_{NP}(p_1, p_2, ..., p_T) = \sum_{t=1}^T \log(p_t),\ (p_1, ..., p_T) \in \Delta$$

- The EL estimator (Owen (1988)), $\left(\widehat{\theta}_{EL}, \widehat{p}_1^{EL}, ..., \widehat{p}_T^{EL}\right)$, solves

$$\max_{\{\theta, p_1, ..., p_T\} \in \Theta \times \Delta} \ell_{NP} = \sum_{t=1}^T \log(p_t) \quad \text{subject to} \sum_{t=1}^T f(z_t;\theta)p_t = 0$$

- The NPMLE of μ is $\hat{\mu}_{EL} = \sum_{t=1}^T \widehat{p}_t^{EL}\delta_{z_t}$ ($\delta_z = 1$ at z).

Information-Theoretic Alternatives: Empirical Likelihood

- Consider the model

$$E\left[f(z_t;\theta_0)\right] \equiv \int f(z_t;\theta_0)d\mu = \underline{0}, \quad \theta \in \Theta \subset \mathbb{R}^s \qquad (2)$$

where f is a known $\mathbb{R}^q$-valued function, $z_t \in \mathbb{R}^k$, $q \geqslant s$.

- We observe draws of $\{z_t\}_{t=1}^T$, from the unknown measure μ.
- Let $\Delta := \left\{(p_1, ..., p_T) : \ \sum_{t=1}^T p_t = 1, \ p_t \geq 0, \ t = 1, ..., T\right\}$, the nonparametric log likelihood at $(p_1, ..., p_T)$ is

$$\ell_{NP}(p_1, p_2, ..., p_T) = \sum_{t=1}^T \log(p_t), \ (p_1, ..., p_T) \in \Delta$$

- The EL estimator (Owen (1988)), $\left(\widehat{\theta}_{EL}, \widehat{p}_1^{EL}, ..., \widehat{p}_T^{EL}\right)$, solves

$$\max_{\{\theta, p_1, ..., p_T\} \in \Theta \times \Delta} \ell_{NP} = \sum_{t=1}^T \log(p_t) \quad \text{subject to} \quad \sum_{t=1}^T f(z_t;\theta)p_t = 0$$

- The NPMLE of μ is $\hat{\mu}_{EL} = \sum_{t=1}^T \widehat{p}_t^{EL} \delta_{z_t}$ ($\delta_z = 1$ at z).

Information-Theoretic Alternatives: Empirical Likelihood

- Consider the model

$$E\left[f(z_t;\theta_0)\right] \equiv \int f(z_t;\theta_0)d\mu = \underline{0}, \quad \theta \in \Theta \subset \mathbb{R}^s \qquad (2)$$

where f is a known $\mathbb{R}^q$-valued function, $z_t \in \mathbb{R}^k$, $q \geqslant s$.

- We observe draws of $\{z_t\}_{t=1}^T$, from the unknown measure μ.
- Let $\Delta := \left\{(p_1, ..., p_T) : \ \sum_{t=1}^T p_t = 1, \ p_t \geq 0, \ t = 1, ..., T\right\}$, the nonparametric log likelihood at $(p_1, ..., p_T)$ is

$$\ell_{NP}(p_1, p_2, ..., p_T) = \sum_{t=1}^T \log(p_t), \ (p_1, ..., p_T) \in \Delta$$

- The EL estimator (Owen (1988)), $\left(\widehat{\theta}_{EL}, \widehat{p}_1^{EL}, ..., \widehat{p}_T^{EL}\right)$, solves

$$\max_{\{\theta, p_1, ..., p_T\} \in \Theta \times \Delta} \ell_{NP} = \sum_{t=1}^T \log(p_t) \quad \text{subject to} \quad \sum_{t=1}^T f(z_t;\theta)p_t = 0$$

- The NPMLE of μ is $\hat{\mu}_{EL} = \sum_{t=1}^T \widehat{p}_t^{EL} \delta_{z_t}$ ($\delta_z = 1$ at z).

Information-Theoretic Alternatives: Empirical Likelihood

- Consider the model

$$E\left[f(z_t;\theta_0)\right] \equiv \int f(z_t;\theta_0)d\mu = \underline{0}, \quad \theta \in \Theta \subset \mathbb{R}^s \qquad (2)$$

  where f is a known $\mathbb{R}^q$-valued function, $z_t \in \mathbb{R}^k$, $q \geqslant s$.
- We observe draws of $\{z_t\}_{t=1}^T$, from the unknown measure μ.
- Let $\Delta := \left\{(p_1, ..., p_T) : \ \sum_{t=1}^T p_t = 1,\ p_t \geq 0,\ t = 1, ..., T\right\}$, the nonparametric log likelihood at $(p_1, ..., p_T)$ is

$$\ell_{NP}(p_1, p_2, ..., p_T) = \sum_{t=1}^T \log(p_t),\ \ (p_1, ..., p_T) \in \Delta$$

- The EL estimator (Owen (1988)), $\left(\widehat{\theta}_{EL}, \widehat{p}_1^{EL}, ..., \widehat{p}_T^{EL}\right)$, solves

$$\max_{\{\theta,p_1,...,p_T\}\in\Theta\times\Delta} \ell_{NP} = \sum_{t=1}^T \log(p_t) \quad \text{subject to} \ \sum_{t=1}^T f(z_t;\theta)p_t = 0$$

- The NPMLE of μ is $\hat{\mu}_{EL} = \sum_{t=1}^T \widehat{p}_t^{EL}\delta_{z_t}$ ($\delta_z = 1$ at z).

- The EL estimator is first, higher-order, and Large Deviation efficient, and has good small sample properties.
- For a function $a(z;\theta_0)$, $\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})\widehat{p}_t^{EL}$ is a more efficient estimator of $E\left[a(z;\theta_0)\right]$ than $\frac{1}{T}\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})$.
- Most importantly, the EL estimator solves the problem

$$\inf_{\theta\in\Theta}\ \inf_{p\in P(\theta)} \int \log\left(\frac{d\mu}{dp}\right) d\mu = \inf_{\theta\in\Theta}\ \inf_{p\in P(\theta)} K(\mu,p)$$

where $K(Q,Q')$ is Kullback-Leibler Information Criterion (KLIC) "distance" between probability measures Q and Q', $P(\theta) := \left\{p \in M : \int f(z;\theta)dp = 0\right\}$ and M is the set of all probability measures on $\mathbb{R}^k$ (absolutely continuous w.r.t. μ)

$\Rightarrow$ EL minimizes the distance – in the information sense – between the estimated prob. measure and the unknown one.

- Moreover, it endogenously re-weights rare events to fit the data (WLLN for rare events, Brown and Smith (1986); KLIC is very sensitive to deviations between measures, Robinson (1991))

- The EL estimator is first, higher-order, and Large Deviation efficient, and has good small sample properties.
- For a function $a(z;\theta_0)$, $\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})\widehat{p}_t^{EL}$ is a more efficient estimator of $E\left[a(z;\theta_0)\right]$ than $\frac{1}{T}\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})$.
- Most importantly, the EL estimator solves the problem

$$\inf_{\theta\in\Theta} \inf_{p\in P(\theta)} \int \log\left(\frac{d\mu}{dp}\right) d\mu = \inf_{\theta\in\Theta} \inf_{p\in P(\theta)} K(\mu,p)$$

where $K(Q,Q')$ is Kullback-Leibler Information Criterion (KLIC) "distance" between probability measures Q and Q', $P(\theta) := \left\{p \in M : \int f(z;\theta)dp = 0\right\}$ and M is the set of all probability measures on $\mathbb{R}^k$ (absolutely continuous w.r.t. μ)

⇒ EL minimizes the distance – in the information sense – between the estimated prob. measure and the unknown one.

- Moreover, it endogenously re-weights rare events to fit the data (WLLN for rare events, Brown and Smith (1986); KLIC is very sensitive to deviations between measures, Robinson (1991))

- The EL estimator is first, higher-order, and Large Deviation efficient, and has good small sample properties.
- For a function $a(z;\theta_0)$, $\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})\widehat{p}_t^{EL}$ is a more efficient estimator of $E\left[a(z;\theta_0)\right]$ than $\frac{1}{T}\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})$.
- Most importantly, the EL estimator solves the problem

$$\inf_{\theta\in\Theta} \inf_{p\in P(\theta)} \int \log\left(\frac{d\mu}{dp}\right) d\mu = \inf_{\theta\in\Theta} \inf_{p\in P(\theta)} K(\mu,p)$$

where $K(Q,Q')$ is Kullback-Leibler Information Criterion (KLIC) "distance" between probability measures Q and Q', $P(\theta) := \left\{p \in M : \int f(z;\theta)dp = 0\right\}$ and M is the set of all probability measures on $\mathbb{R}^k$ (absolutely continuous w.r.t. μ)

$\Rightarrow$ EL minimizes the distance – in the information sense – between the estimated prob. measure and the unknown one.

- Moreover, it endogenously re-weights rare events to fit the data (WLLN for rare events, Brown and Smith (1986); KLIC is very sensitive to deviations between measures, Robinson (1991))

- The EL estimator is first, higher-order, and Large Deviation efficient, and has good small sample properties.
- For a function $a(z;\theta_0)$, $\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})\widehat{p}_t^{EL}$ is a more efficient estimator of $E\left[a(z;\theta_0)\right]$ than $\frac{1}{T}\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})$.
- Most importantly, the EL estimator solves the problem

$$\inf_{\theta\in\Theta}\ \inf_{p\in P(\theta)} \int \log\left(\frac{d\mu}{dp}\right) d\mu = \inf_{\theta\in\Theta}\ \inf_{p\in P(\theta)} K(\mu, p)$$

where $K(Q, Q')$ is Kullback-Leibler Information Criterion (KLIC) "distance" between probability measures Q and Q', $P(\theta) := \left\{p \in M : \int f(z;\theta)dp = 0\right\}$ and M is the set of all probability measures on $\mathbb{R}^k$ (absolutely continuous w.r.t. μ)

⇒ EL minimizes the distance – in the information sense – between the estimated prob. measure and the unknown one.

- Moreover, it endogenously re-weights rare events to fit the data (WLLN for rare events, Brown and Smith (1986); KLIC is very sensitive to deviations between measures, Robinson (1991))

- The EL estimator is first, higher-order, and Large Deviation efficient, and has good small sample properties.
- For a function $a(z;\theta_0)$, $\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})\widehat{p}_t^{EL}$ is a more efficient estimator of $E\left[a(z;\theta_0)\right]$ than $\frac{1}{T}\sum_{t=1}^{T} a(z_t;\widehat{\theta}_{EL})$.
- Most importantly, the EL estimator solves the problem

$$\inf_{\theta\in\Theta} \inf_{p\in P(\theta)} \int \log\left(\frac{d\mu}{dp}\right) d\mu = \inf_{\theta\in\Theta} \inf_{p\in P(\theta)} K(\mu,p)$$

where $K(Q,Q')$ is Kullback-Leibler Information Criterion (KLIC) "distance" between probability measures Q and Q', $P(\theta) := \left\{p\in M : \int f(z;\theta)dp = 0\right\}$ and M is the set of all probability measures on $\mathbb{R}^k$ (absolutely continuous w.r.t. μ)

$\Rightarrow$ EL minimizes the distance – in the information sense – between the estimated prob. measure and the unknown one.

- Moreover, it endogenously re-weights rare events to fit the data (WLLN for rare events, Brown and Smith (1986); KLIC is very sensitive to deviations between measures, Robinson (1991))

Exponential Tilting and Bayesian Interpretations

- Since the KLIC divergence is not symmetric, we can also define the Exponential Tilting, ET, estimator (e.g. Kitamura and Stutzer (1997)), $\left(\widehat{\theta}_{ET}, \widehat{p}_1^{ET}, \ldots, \widehat{p}_T^{ET}\right)$, as

$$\inf_{\theta\in\Theta} \inf_{p\in P(\theta)} \int \log\left(\frac{dp}{d\mu}\right) dp = \inf_{\theta\in\Theta} \inf_{p\in P(\theta)} K(p,\mu)$$

- Given a prior $\pi(\theta)$, Lazar (2003) shows that Bayesian EL (BEL) posterior inference can be accurately based on

$$p\left(\theta | \{z_t\}_{t=1}^{T}\right) \propto \pi(\theta) \times \prod_{t=1}^{T} \hat{p}_t^{EL}$$

- Also, under a diffuse prior for $\{p_t\}_{t=1}^{T}$, a proper posterior can be obtained from $\left\{\hat{p}^{ET}\right\}_{t=1}^{T}$ (BETEL, Schennach (2005))

Estimation results
Data Description

Exponential Tilting and Bayesian Interpretations

- Since the KLIC divergence is not symmetric, we can also define the Exponential Tilting, ET, estimator (e.g. Kitamura and Stutzer (1997)), $\left(\widehat{\theta}_{ET}, \widehat{p}_1^{ET}, \dots, \widehat{p}_T^{ET}\right)$, as

$$\inf_{\theta\in\Theta} \inf_{p\in P(\theta)} \int \log\left(\frac{dp}{d\mu}\right) dp = \inf_{\theta\in\Theta} \inf_{p\in P(\theta)} K(p,\mu)$$

- Given a prior $\pi(\theta)$, Lazar (2003) shows that Bayesian EL (BEL) posterior inference can be accurately based on

$$p\left(\theta | \{z_t\}_{t=1}^{T}\right) \propto \pi(\theta) \times \prod_{t=1}^{T} \hat{p}_t^{EL}$$

- Also, under a diffuse prior for $\{p_t\}_{t=1}^{T}$, a proper posterior can be obtained from $\{\hat{p}^{ET}\}_{t=1}^{T}$ (BETEL, Schennach (2005))

Estimation results
Data Description

Exponential Tilting and Bayesian Interpretations

- Since the KLIC divergence is not symmetric, we can also define the Exponential Tilting, ET, estimator (e.g. Kitamura and Stutzer (1997)), $\left(\widehat{\theta}_{ET}, \widehat{p}_1^{ET}, ..., \widehat{p}_T^{ET}\right)$, as

$$\inf_{\theta\in\Theta} \inf_{p\in P(\theta)} \int \log\left(\frac{dp}{d\mu}\right) dp = \inf_{\theta\in\Theta} \inf_{p\in P(\theta)} K(p,\mu)$$

- Given a prior $\pi(\theta)$, Lazar (2003) shows that Bayesian EL (BEL) posterior inference can be accurately based on

$$p\left(\theta | \{z_t\}_{t=1}^{T}\right) \propto \pi(\theta) \times \prod_{t=1}^{T} \hat{p}_t^{EL}$$

- Also, under a diffuse prior for $\{p_t\}_{t=1}^{T}$, a proper posterior can be obtained from $\left\{\hat{p}^{ET}\right\}_{t=1}^{T}$ (BETEL, Schennach (2005))

Estimation results
Data Description

Estimation

"Really, the most natural thing to do with the consumption-based model is to estimate it and test it, as one would do for any economic model." Cochrane (2005).

- Given their properties, EL, ET, BEL and BETEL are the ideal device for the estimation of the consumption Euler equation (1) if we are concerned about rare events

Note: the GMM estimator does not focus on the distance between measures, but only on the inability of the parameters to satisfy the sample analog of the moment condition

Remark: inference based on BEL and BETEL satisfies the "likelihood priciple" $\longrightarrow$ it depends only on the data

$\Rightarrow$ we estimate and test the Euler equation (1) for a representative agent and the market return using the EL, ET, BEL and ETEL.

Estimation results

Estimation

"Really, the most natural thing to do with the consumption-based model is to estimate it and test it, as one would do for any economic model." Cochrane (2005).

- Given their properties, EL, ET, BEL and BETEL are the ideal device for the estimation of the consumption Euler equation (1) if we are concerned about rare events

Note: the GMM estimator does not focus on the distance between measures, but only on the inability of the parameters to satisfy the sample analog of the moment condition

Remark: inference based on BEL and BETEL satisfies the "likelihood priciple" $\longrightarrow$ it depends only on the data

$\Rightarrow$ we estimate and test the Euler equation (1) for a representative agent and the market return using the EL, ET, BEL and ETEL.

Estimation results

Estimation

"Really, the most natural thing to do with the consumption-based model is to estimate it and test it, as one would do for any economic model." Cochrane (2005).

- Given their properties, EL, ET, BEL and BETEL are the ideal device for the estimation of the consumption Euler equation (1) if we are concerned about rare events

Note: the GMM estimator does not focus on the distance between measures, but only on the inability of the parameters to satisfy the sample analog of the moment condition

Remark: inference based on BEL and BETEL satisfies the "likelihood priciple" $\rightarrow$ it depends only on the data

$\Rightarrow$ we estimate and test the Euler equation (1) for a representative agent and the market return using the EL, ET, BEL and ETEL.

Estimation results

Estimation

"Really, the most natural thing to do with the consumption-based model is to estimate it and test it, as one would do for any economic model." Cochrane (2005).

- Given their properties, EL, ET, BEL and BETEL are the ideal device for the estimation of the consumption Euler equation (1) if we are concerned about rare events

Note: the GMM estimator does not focus on the distance between measures, but only on the inability of the parameters to satisfy the sample analog of the moment condition

Remark: inference based on BEL and BETEL satisfies the "likelihood priciple" → it depends only on the data

⇒ we estimate and test the Euler equation (1) for a representative agent and the market return using the EL, ET, BEL and ETEL.

Estimation results

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

▸ Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value $\longrightarrow$ higher returns) and the value premium (high book values relative to market equity $\longrightarrow$ higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Estimation Results

Table 1: Euler Equation Estimation

	EL	*ET*	*BEL*	*BETEL*
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{\gamma}$	102 (48.0)	146 (32.3)	102 [24.8, 263.1]	90 [19.5, 164.9]
$\chi^2_{(1)}$	9.87 (.002)	10.65 (.001)		
$\Pr(\gamma \leq 10 \mid \text{data})$			.64%	.92%
	Panel B: Annual Data (1929-2006)			
$\hat{\gamma}$	32 (10.5)	32 (10.5)	32 [13.4, 64.1]	32 [13.8, 57.1]
$\chi^2_{(1)}$	5.26 (.022)	5.93 (.015)		
$\Pr(\gamma \leq 10 \mid \text{data})$			1.00%	.84%

Note: similar findings with data starting in 1890.

Data Description

Estimation Results

Table 1: Euler Equation Estimation

	EL	*ET*	*BEL*	*BETEL*
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{\gamma}$	102 (48.0)	146 (32.3)	102 [24.8, 263.1]	90 [19.5, 164.9]
$\chi^2_{(1)}$	9.87 (.002)	10.65 (.001)		
$\Pr(\gamma \leq 10 \mid \text{data})$			.64%	.92%
	Panel B: Annual Data (1929-2006)			
$\hat{\gamma}$	32 (10.5)	32 (10.5)	32 [13.4, 64.1]	32 [13.8, 57.1]
$\chi^2_{(1)}$	5.26 (.022)	5.93 (.015)		
$\Pr(\gamma \leq 10 \mid \text{data})$			1.00%	.84%

Note: similar findings with data starting in 1890.

Data Description

Estimation Results

Table 1: Euler Equation Estimation

	EL	*ET*	*BEL*	*BETEL*
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{\gamma}$	102 (48.0)	146 (32.3)	102 [24.8, 263.1]	90 [19.5, 164.9]
$\chi^2_{(1)}$	9.87 (.002)	10.65 (.001)		
$\Pr(\gamma \leq 10 \mid \text{data})$			.64%	.92%
	Panel B: Annual Data (1929-2006)			
$\hat{\gamma}$	32 (10.5)	32 (10.5)	32 [13.4, 64.1]	32 [13.8, 57.1]
$\chi^2_{(1)}$	5.26 (.022)	5.93 (.015)		
$\Pr(\gamma \leq 10 \mid \text{data})$			1.00%	.84%

Note: similar findings with data starting in 1890.

Data Description

Estimation Results

Table 1: Euler Equation Estimation

	EL	*ET*	*BEL*	*BETEL*
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{\gamma}$	102 (48.0)	146 (32.3)	102 [24.8, 263.1]	90 [19.5, 164.9]
$\chi^2_{(1)}$	9.87 (.002)	10.65 (.001)		
$\Pr(\gamma \leq 10 \vert \text{data})$			.64%	.92%
	Panel B: Annual Data (1929-2006)			
$\hat{\gamma}$	32 (10.5)	32 (10.5)	32 [13.4, 64.1]	32 [13.8, 57.1]
$\chi^2_{(1)}$	5.26 (.022)	5.93 (.015)		
$\Pr(\gamma \leq 10 \vert \text{data})$			1.00%	.84%

Note: similar findings with data starting in 1890.

Data Description

Outline

A world without the Equity Premium Puzzle

- The consumption Euler equation implies that

$$\frac{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]} = \underbrace{E^F\left[R_t^e\right] + \frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:epp^F(\gamma)}$$

  where F is the true, unknown, probability measure
- The right hand side is a measure of the EPP under F
- For any γ, EL and ET estimate F with $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$ such that

$$\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e \hat{p}_t^j(\gamma) = 0 \;\forall\gamma,\; j \in \{EL, ET\}$$

$$\therefore E^{\hat{P}^j(\gamma)}\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right] = 0 \rightarrow epp^j(\gamma) = 0,\; j \in \{EL, ET\}$$

  where $\hat{P}^j(\gamma)$ is the prob. measure defined by $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$

A world without the Equity Premium Puzzle

- The consumption Euler equation implies that

$$\frac{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]} = \underbrace{E^F\left[R_t^e\right] + \frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:epp^F(\gamma)}$$

  where F is the true, unknown, probability measure
- The right hand side is a measure of the EPP under F
- For any γ, EL and ET estimate F with $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$ such that

$$\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e \hat{p}_t^j(\gamma) = 0 \;\forall\gamma,\; j \in \{EL, ET\}$$

$$\therefore E^{\hat{P}^j(\gamma)}\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right] = 0 \rightarrow epp^j(\gamma) = 0,\; j \in \{EL, ET\}$$

  where $\hat{P}^j(\gamma)$ is the prob. measure defined by $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$

A world without the Equity Premium Puzzle

- The consumption Euler equation implies that

$$\frac{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]} = \underbrace{E^F\left[R_t^e\right] + \frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:epp^F(\gamma)}$$

  where F is the true, unknown, probability measure
- The right hand side is a measure of the EPP under F
- For any γ, EL and ET estimate F with $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$ such that

$$\sum_{t=1}^T \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e \hat{p}_t^j(\gamma) = 0 \; \forall \gamma, \; j \in \{EL, ET\}$$

$$\therefore E^{\hat{P}^j(\gamma)}\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right] = 0 \rightarrow epp^j(\gamma) = 0, \; j \in \{EL, ET\}$$

where $\hat{P}^j(\gamma)$ is the prob. measure defined by $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$

A world without the Equity Premium Puzzle

- The consumption Euler equation implies that

$$\frac{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]} = \underbrace{E^F\left[R_t^e\right] + \frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R_t^e\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:epp^F(\gamma)}$$

  where F is the true, unknown, probability measure
- The right hand side is a measure of the EPP under F
- For any γ, EL and ET estimate F with $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$ such that

$$\sum_{t=1}^T \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e \hat{p}_t^j(\gamma) = 0 \ \forall\gamma,\ j \in \{EL, ET\}$$

$$\therefore E^{\hat{P}^j(\gamma)}\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_t^e\right] = 0 \rightarrow epp^j(\gamma) = 0,\ j \in \{EL, ET\}$$

  where $\hat{P}^j(\gamma)$ is the prob. measure defined by $\left\{\hat{p}_t^j(\gamma)\right\}_{t=1}^T$

Constructing the Rare Events Distribution of the Data

- Therefore, we can fix γ and have EL and ET estimate the probability measure needed to solve the EPP with that given level of RRA
- We fix $\gamma = 10$ (the upper bound of the "reasonable" range) but also consider $\gamma \in]0, 10]$
- The estimated $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, will minimize the distance - in the information sense - between the unknown probability measure and the one needed to rationalize the EPP

"Thus, data are used to calibrate the model economy so that it mimics the world as closely as possible along a limited, but clearly specified, number of dimensions." Kydland and Prescott (1996)

Note: if rare events are the explanation of the EPP, $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, should identify their distribution

Moreover: $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ delivers – by construction – the *most likely* rare events explanation of the EPP.

Constructing the Rare Events Distribution of the Data

- Therefore, we can fix γ and have EL and ET estimate the probability measure needed to solve the EPP with that given level of RRA
- We fix $\gamma = 10$ (the upper bound of the "reasonable" range) but also consider $\gamma \in]0, 10]$
- The estimated $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, will minimize the distance - in the information sense - between the unknown probability measure and the one needed to rationalize the EPP

"Thus, data are used to calibrate the model economy so that it mimics the world as closely as possible along a limited, but clearly specified, number of dimensions." Kydland and Prescott (1996)

Note: if rare events are the explanation of the EPP, $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, should identify their distribution

Moreover: $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ delivers – by construction – the *most likely* rare events explanation of the EPP.

Constructing the Rare Events Distribution of the Data

- Therefore, we can fix γ and have EL and ET estimate the probability measure needed to solve the EPP with that given level of RRA
- We fix $\gamma = 10$ (the upper bound of the "reasonable" range) but also consider $\gamma \in]0, 10]$
- The estimated $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, will minimize the distance - in the information sense - between the unknown probability measure and the one needed to rationalize the EPP

"Thus, data are used to calibrate the model economy so that it mimics the world as closely as possible along a limited, but clearly specified, number of dimensions." Kydland and Prescott (1996)

Note: if rare events are the explanation of the EPP, $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, should identify their distribution

Moreover: $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ delivers – by construction – the *most likely* rare events explanation of the EPP.

Constructing the Rare Events Distribution of the Data

- Therefore, we can fix γ and have EL and ET estimate the probability measure needed to solve the EPP with that given level of RRA
- We fix $\gamma = 10$ (the upper bound of the "reasonable" range) but also consider $\gamma \in]0, 10]$
- The estimated $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, will minimize the distance - in the information sense - between the unknown probability measure and the one needed to rationalize the EPP

"Thus, data are used to calibrate the model economy so that it mimics the world as closely as possible along a limited, but clearly specified, number of dimensions." Kydland and Prescott (1996)

Note: if rare events are the explanation of the EPP, $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, should identify their distribution

Moreover: $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ delivers – by construction – the *most likely* rare events explanation of the EPP

Constructing the Rare Events Distribution of the Data

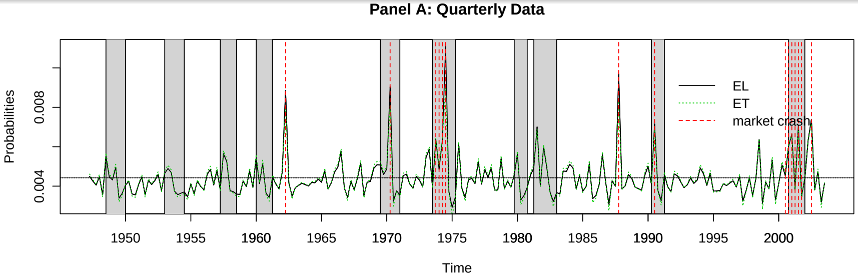
- Therefore, we can fix γ and have EL and ET estimate the probability measure needed to solve the EPP with that given level of RRA
- We fix $\gamma = 10$ (the upper bound of the "reasonable" range) but also consider $\gamma \in]0, 10]$
- The estimated $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, will minimize the distance - in the information sense - between the unknown probability measure and the one needed to rationalize the EPP

"Thus, data are used to calibrate the model economy so that it mimics the world as closely as possible along a limited, but clearly specified, number of dimensions." Kydland and Prescott (1996)

Note: if rare events are the explanation of the EPP, $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, should identify their distribution

Moreover: $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ delivers – by construction – the *most likely* rare events explanation of the EPP.

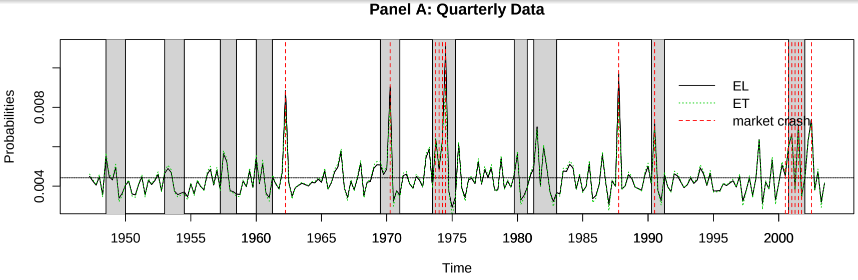
Rare Events Probabilities



Shaded areas are NBER recessions. Vertical dashed lines are the stock market crashes (Mishkin-White (2002)).

- $corr\left(\hat{P}^{EL}(\gamma), \hat{P}^{ET}(\gamma)\right) = .97$
- very few substantial (but small) increases in probability
- Probability of recession: Sample: 19.9%. EL: 21.3%. ET: 20.9%.
- Probability of market crash: Sample: 6.6%. EL: 10.2%. ET: 9.6%.

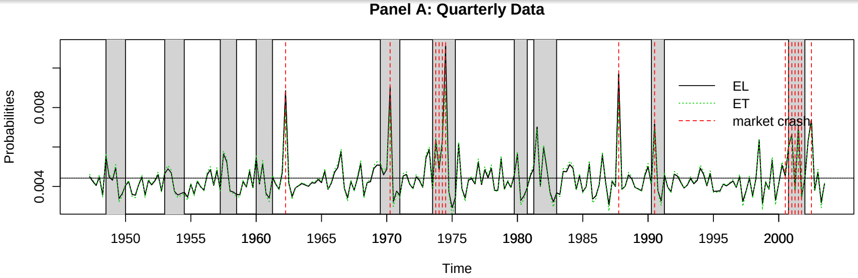
Rare Events Probabilities



Shaded areas are NBER recessions. Vertical dashed lines are the stock market crashes (Mishkin-White (2002)).

- $corr\left(\hat{P}^{EL}(\gamma), \hat{P}^{ET}(\gamma)\right) = .97$
- very few substantial (but small) increases in probability
- Probability of recession: Sample: 19.9%. EL: 21.3%. ET: 20.9%.
- Probability of market crash: Sample: 6.6%. EL: 10.2%. ET: 9.6%.

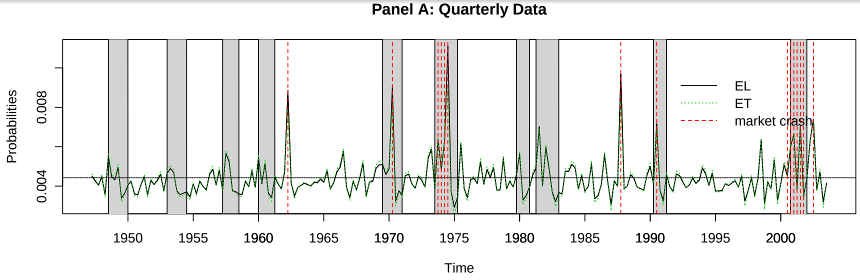
Rare Events Probabilities



Shaded areas are NBER recessions. Vertical dashed lines are the stock market crashes (Mishkin-White (2002)).

- $corr\left(\hat{P}^{EL}(\gamma), \hat{P}^{ET}(\gamma)\right) = .97$
- very few substantial (but small) increases in probability
- Probability of recession: Sample: 19.9%. EL: 21.3%. ET: 20.9%.
- Probability of market crash: Sample: 6.6%. EL: 10.2%. ET: 9.6%.

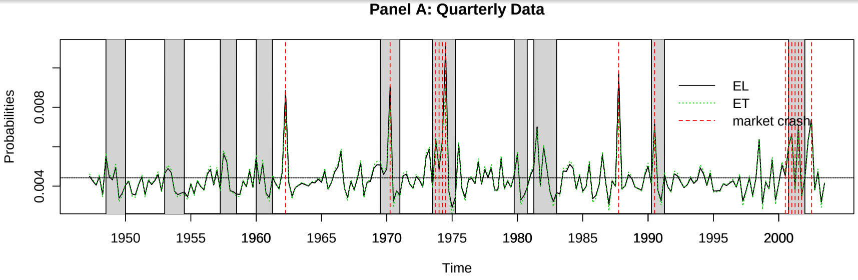
Rare Events Probabilities



Shaded areas are NBER recessions. Vertical dashed lines are the stock market crashes (Mishkin-White (2002)).

- $corr\left(\hat{P}^{EL}(\gamma), \hat{P}^{ET}(\gamma)\right) = .97$
- very few substantial (but small) increases in probability
- Probability of recession: Sample: 19.9%. EL: 21.3%. ET: 20.9%.
- Probability of market crash: Sample: 6.6%. EL: 10.2%. ET: 9.6%.

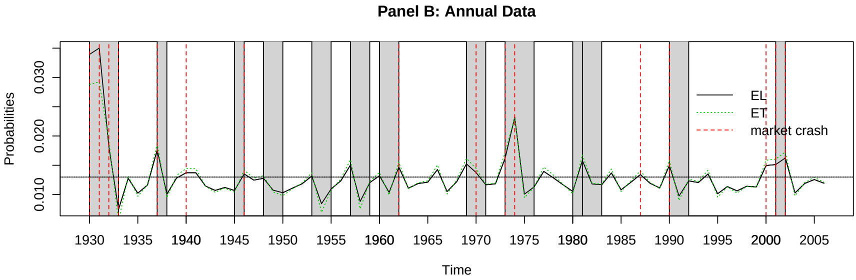
Rare Events Probabilities



Shaded areas are NBER recessions. Vertical dashed lines are the stock market crashes (Mishkin-White (2002)).

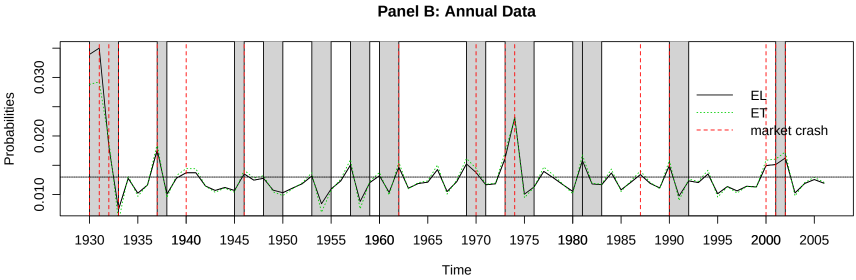
- $corr\left(\hat{P}^{EL}(\gamma), \hat{P}^{ET}(\gamma)\right) = .97$
- very few substantial (but small) increases in probability
- Probability of recession: Sample: 19.9%. EL: 21.3%. ET: 20.9%.
- Probability of market crash: Sample: 6.6%. EL: 10.2%. ET: 9.6%.

Rare Events Probabilities



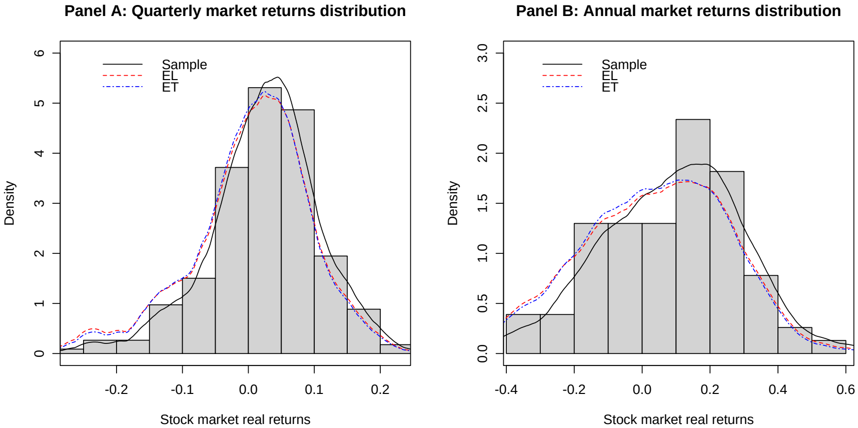
Note: similar findings with data starting in 1890.

Rare Events Probabilities



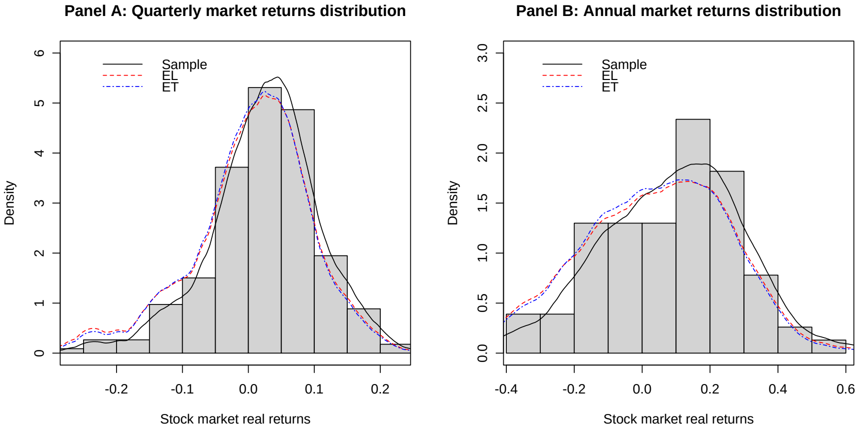
Note: similar findings with data starting in 1890.

The Implied Distribution of Returns



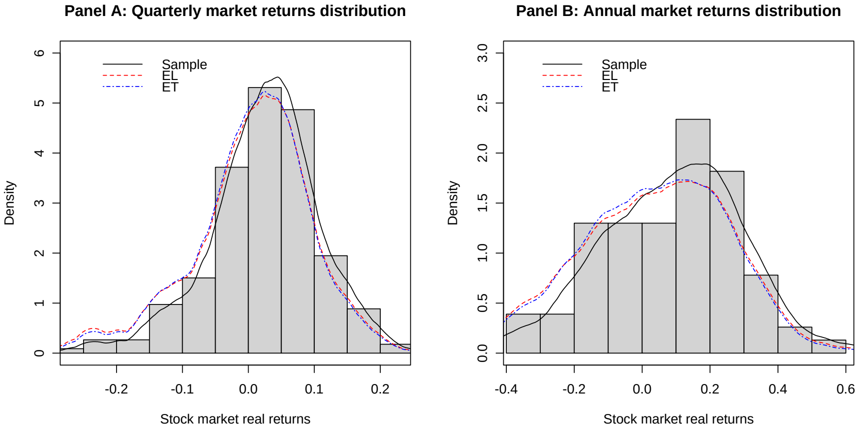
- Ticker left tails, left skewness, median and mean reduction
- Implied median (mean) of return: 4.9%-6.4% (2.1%-5%)
- Barro (2005) calibrated rare events model: 3.7%-8.4%

The Implied Distribution of Returns



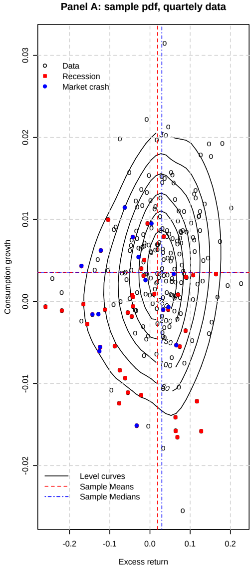
- Ticker left tails, left skewness, median and mean reduction
- Implied median (mean) of return: 4.9%-6.4% (2.1%-5%)
- Barro (2005) calibrated rare events model: 3.7%-8.4%

The Implied Distribution of Returns

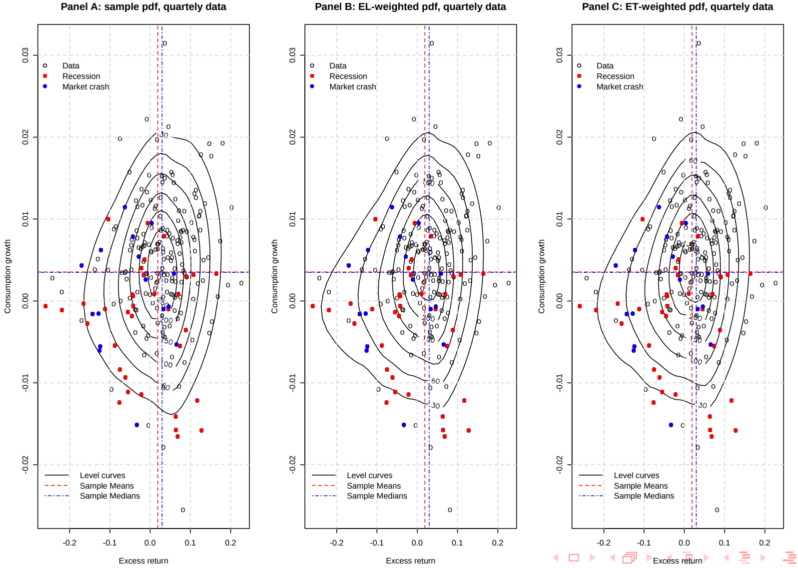


- Ticker left tails, left skewness, median and mean reduction
- Implied median (mean) of return: 4.9%-6.4% (2.1%-5%)
- Barro (2005) calibrated rare events model: 3.7%-8.4%

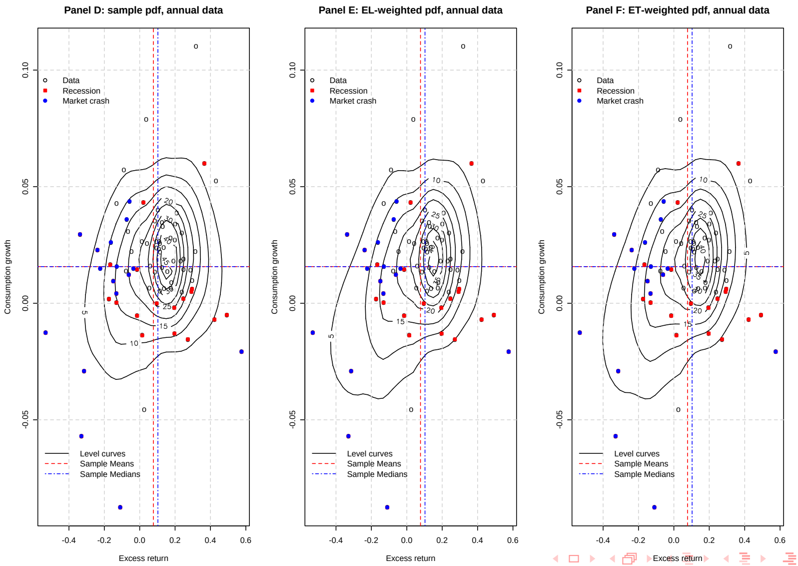
The Distribution of Risk premia and Consumption Growth



The Distribution of Risk premia and Consumption Growth



The Distribution of Risk premia and Consumption Growth



How likely is the Equity Premium Puzzle?

- The $\hat{P}^j(\gamma), \ j \in \{EL, ET\}$, measures provide the most probable (in the likelihood sense) rare events explanation of the EPP

Under the rare events hypothesis, what is the likelihood of having an EPP in a sample of the same size as the historical one?

- To answer this question we perform the following counterfactual exercise:
  1. Using $\hat{P}^j(\gamma), \ j \in \{EL, ET\}$ we generate 100,000 samples of the same size as the historical ones
  2. In each i sample we compute the realized EPP as

$$epp_i^T(\gamma) = E^T\left[R_{i,t}^e\right] + \frac{Cov^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}; R_{i,t}^e\right]}{E^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}\right]}.$$

  3. In each sample we also perform a GMM estimation of γ

How likely is the Equity Premium Puzzle?

- The $\hat{P}^j(\gamma),\ \ j \in \{EL, ET\}$, measures provide the most probable (in the likelihood sense) rare events explanation of the EPP

Under the rare events hypothesis, what is the likelihood of having an EPP in a sample of the same size as the historical one?

- To answer this question we perform the following counterfactual exercise:
  1. Using $\hat{P}^j(\gamma),\ \ j \in \{EL, ET\}$ we generate 100,000 samples of the same size as the historical ones
  2. In each i sample we compute the realized EPP as

$$epp_i^T(\gamma) = E^T\left[R_{i,t}^e\right] + \frac{Cov^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}; R_{i,t}^e\right]}{E^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}\right]}.$$

  3. In each sample we also perform a GMM estimation of γ

How likely is the Equity Premium Puzzle?

- The $\hat{P}^j(\gamma),\ \ j \in \{EL, ET\}$, measures provide the most probable (in the likelihood sense) rare events explanation of the EPP

Under the rare events hypothesis, what is the likelihood of having an EPP in a sample of the same size as the historical one?

- To answer this question we perform the following counterfactual exercise:
  1. Using $\hat{P}^j(\gamma),\ \ j \in \{EL, ET\}$ we generate 100,000 samples of the same size as the historical ones
  2. In each i sample we compute the realized EPP as

$$epp_i^T(\gamma) = E^T\left[R_{i,t}^e\right] + \frac{Cov^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}; R_{i,t}^e\right]}{E^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}\right]}.$$

  3. In each sample we also perform a GMM estimation of γ

How likely is the Equity Premium Puzzle?

- The $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, measures provide the most probable (in the likelihood sense) rare events explanation of the EPP

Under the rare events hypothesis, what is the likelihood of having an EPP in a sample of the same size as the historical one?

- To answer this question we perform the following counterfactual exercise:
  1. Using $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ we generate 100,000 samples of the same size as the historical ones
  2. In each i sample we compute the realized EPP as

$$epp_i^T(\gamma) = E^T\left[R_{i,t}^e\right] + \frac{Cov^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}; R_{i,t}^e\right]}{E^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}\right]}.$$

  3. In each sample we also perform a GMM estimation of γ

How likely is the Equity Premium Puzzle?

- The $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$, measures provide the most probable (in the likelihood sense) rare events explanation of the EPP

Under the rare events hypothesis, what is the likelihood of having an EPP in a sample of the same size as the historical one?

- To answer this question we perform the following counterfactual exercise:
  1. Using $\hat{P}^j(\gamma)$, $j \in \{EL, ET\}$ we generate 100,000 samples of the same size as the historical ones
  2. In each i sample we compute the realized EPP as

$$epp_i^T(\gamma) = E^T\left[R_{i,t}^e\right] + \frac{Cov^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}; R_{i,t}^e\right]}{E^T\left[\left(\frac{C_{i,t}}{C_{i,t-1}}\right)^{-\gamma}\right]}.$$

  3. In each sample we also perform a GMM estimation of γ.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}(\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}(\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}(\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}(\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}(\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}(\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}(\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}(\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}(\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}(\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}(\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}(\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}(\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}(\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}(\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}(\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}(\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}(\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}(\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}(\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}(\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}(\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}(\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}(\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}\,(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}\,(\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}\,(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}\,(\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}\,(\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}\,(\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}\,(\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}\,(\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}\ (\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}\ (\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}\ (\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}\ (\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}\ (\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}\ (\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}\ (\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}\ (\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}\ (\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}\ (\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}\ (\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}\ (\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}\ (\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}\ (\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}\ (\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}\ (\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

How likely is the Equity Premium Puzzle?

Table 2: Counterfactual Equity Premium Puzzle

	epp^T	epp_i^T	$\Pr\left(epp_i^T \geq epp^T\right)$	$\hat{\gamma}_{GMM}$
	Panel A: Quarterly Data (1947:Q1-2003:Q3)			
$\hat{P}^{EL}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.7%]	0.10%	5 [−41, 67]
$\hat{P}^{EL}(\gamma = 10)$	7.3%	0.0% [−4.7%, 4.7%]	0.12%	10 [−36, 69]
$\hat{P}^{ET}(\gamma = 5)$	7.4%	0.0% [−4.6%, 4.5%]	0.10%	5 [−43, 66]
$\hat{P}^{ET}(\gamma = 10)$	7.3%	0.0% [−4.6%, 4.5%]	0.13%	10 [−40, 70]
	Panel B: Annual Data (1929-2006)			
$\hat{P}^{EL}(\gamma = 5)$	7.2%	0.0% [−5.4%, 5.3%]	0.37%	5 [−21, 29]
$\hat{P}^{EL}(\gamma = 10)$	6.5%	0.0% [−5.7%, 5.7%]	1.22%	10 [−12, 32]
$\hat{P}^{ET}(\gamma = 5)$	7.2%	0.0% [−5.1%, 5.1%]	0.33%	5 [−24, 29]
$\hat{P}^{ET}(\gamma = 10)$	6.5%	0.0% [−5.4%, 5.5%]	0.98%	10 [−13, 33]

Note: similar findings with data starting in 1890.

Rare Events and the Cross-Section of Asset Returns

- The consumption Euler equation implies that

$$E^F\left[R^e_{i,t}\right] = \alpha - \underbrace{\frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:\beta_i}\lambda \qquad (3)$$

should hold for any asset i with $\alpha = 0$ and $\lambda = 1$.

- Linearizing the pricing kernel we have that

$$E^F\left[R^e_{i,t}\right] = \alpha + \underbrace{Cov^F\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)}_{=:\beta_i}\lambda \qquad (4)$$

should hold with $\alpha = 0$ and $\lambda > 0$, $\forall i$

- The β_i terms can be interpreted as a measure of the consumption risk that an agent undertakes investing in asset i.

Rare Events and the Cross-Section of Asset Returns

- The consumption Euler equation implies that

$$E^F\left[R^e_{i,t}\right] = \alpha - \underbrace{\frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:\beta_i}\lambda \qquad (3)$$

should hold for any asset i with $\alpha = 0$ and $\lambda = 1$.

- Linearizing the pricing kernel we have that

$$E^F\left[R^e_{i,t}\right] = \alpha + \underbrace{Cov^F\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)}_{=:\beta_i}\lambda \qquad (4)$$

should hold with $\alpha = 0$ and $\lambda > 0$, $\forall i$

- The β_i terms can be interpreted as a measure of the consumption risk that an agent undertakes investing in asset i.

Rare Events and the Cross-Section of Asset Returns

- The consumption Euler equation implies that

$$E^F\left[R^e_{i,t}\right] = \alpha - \underbrace{\frac{Cov^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right]}{E^F\left[\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\right]}}_{=:\beta_i}\lambda \tag{3}$$

  should hold for any asset i with $\alpha = 0$ and $\lambda = 1$.
- Linearizing the pricing kernel we have that

$$E^F\left[R^e_{i,t}\right] = \alpha + \underbrace{Cov^F\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)}_{=:\beta_i}\lambda \tag{4}$$

  should hold with $\alpha = 0$ and $\lambda > 0$, $\forall i$
- The β_i terms can be interpreted as a measure of the consumption risk that an agent undertakes investing in asset i.

- The cross-sectional implications of equations (3) and (4) are generally rejected by the data (e.g. Parker-Julliard (2005))
- Does the rare events rationalization of the EPP help the CCAPM explain the cross-section of asset returns?

1. We focus on the 25 Fama-French (1992) Size and Book-to-market portfolios (1947:Q1-2003:Q3) FF25
2. We adapt the Fama-MacBeth (1973) cross-sectional regression procedure to construct the moments in equations (3) and (4) under the $\hat{P}^{EL}(\gamma)$ and $\hat{P}^{ET}(\gamma)$ measures needed to solve the EPP. $\hat{P}^{j}$-weighted Fama-McBeth
3. We also report the changes in $Var(\beta_i)/Var\left(E\left[R^e_{i,t+1}\right]\right)$,

$$Var\left(corr\left(\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right)\right), \text{ and } Var\left(corr\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)\right)$$

caused by computing the moments under the $\hat{P}^{ET}(\gamma)$ and $\hat{P}^{EL}(\gamma)$ measures instead that as sample analogs.

Counterfactual Cross-Sectional Regressions

- The cross-sectional implications of equations (3) and (4) are generally rejected by the data (e.g. Parker-Julliard (2005))
- Does the rare events rationalization of the EPP help the CCAPM explain the cross-section of asset returns?
  1. We focus on the 25 Fama-French (1992) Size and Book-to-market portfolios (1947:Q1-2003:Q3) FF25
  2. We adapt the Fama-MacBeth (1973) cross-sectional regression procedure to construct the moments in equations (3) and (4) under the $\hat{P}^{EL}(\gamma)$ and $\hat{P}^{ET}(\gamma)$ measures needed to solve the EPP. $\hat{P}^{j}$-weighted Fama-McBeth
  3. We also report the changes in $Var(\beta_i)/Var\left(E\left[R^e_{i,t+1}\right]\right)$, $Var\left(corr\left(\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right)\right)$, and $Var\left(corr\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)\right)$ caused by computing the moments under the $\hat{P}^{ET}(\gamma)$ and $\hat{P}^{EL}(\gamma)$ measures instead that as sample analogs.

Counterfactual Cross-Sectional Regressions

- The cross-sectional implications of equations (3) and (4) are generally rejected by the data (e.g. Parker-Julliard (2005))
- Does the rare events rationalization of the EPP help the CCAPM explain the cross-section of asset returns?
  1. We focus on the 25 Fama-French (1992) Size and Book-to-market portfolios (1947:Q1-2003:Q3) FF25
  2. We adapt the Fama-MacBeth (1973) cross-sectional regression procedure to construct the moments in equations (3) and (4) under the $\hat{P}^{EL}(\gamma)$ and $\hat{P}^{ET}(\gamma)$ measures needed to solve the EPP. $\hat{P}^{j}$-weighted Fama-McBeth
  3. We also report the changes in $Var(\beta_i)/Var\left(E\left[R^e_{i,t+1}\right]\right)$, $Var\left(corr\left(\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right)\right)$, and $Var\left(corr\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)\right)$ caused by computing the moments under the $\hat{P}^{ET}(\gamma)$ and $\hat{P}^{EL}(\gamma)$ measures instead that as sample analogs.

Counterfactual Cross-Sectional Regressions

- The cross-sectional implications of equations (3) and (4) are generally rejected by the data (e.g. Parker-Julliard (2005))
- Does the rare events rationalization of the EPP help the CCAPM explain the cross-section of asset returns?
  1. We focus on the 25 Fama-French (1992) Size and Book-to-market portfolios (1947:Q1-2003:Q3) FF25
  2. We adapt the Fama-MacBeth (1973) cross-sectional regression procedure to construct the moments in equations (3) and (4) under the $\hat{P}^{EL}(\gamma)$ and $\hat{P}^{ET}(\gamma)$ measures needed to solve the EPP. $\hat{P}^{j}$-weighted Fama-McBeth
  3. We also report the changes in $Var(\beta_i)/Var\left(E\left[R^e_{i,t+1}\right]\right)$, $Var\left(corr\left(\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}; R^e_{i,t}\right)\right)$, and $Var\left(corr\left(\ln\frac{C_t}{C_{t-1}}; R^e_{i,t}\right)\right)$ caused by computing the moments under the $\hat{P}^{ET}(\gamma)$ and $\hat{P}^{EL}(\gamma)$ measures instead that as sample analogs.

Counterfactual Cross-Sectional Regressions

Probability Weighted Fama-MacBeth Regressions

I: For each asset i construct the consumption risk β's as

$$\hat{\beta}_i^j := -\frac{\sum_{t=1}^T \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^T \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]\left[\sum_{t=1}^T R_{i,t}^e \hat{p}_t^j\right]}{\left[\sum_{t=1}^T \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]},$$

where $j \in \{EL, ET\}$ and γ is fixed, and as

$$\hat{\beta}_i^j := \sum_{t=1}^T \ln\left(\frac{C_t}{C_{t-1}}\right) R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^T \ln\left(\frac{C_t}{C_{t-1}}\right) \hat{p}_t^j\right]\left[\sum_{t=1}^T R_{i,t}^e \hat{p}_t^j\right]$$

II: For each t, run the cross–sectional regression

$$R_{i,t}^e = \alpha_t + \hat{\beta}_i^j \lambda_t + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ is a mean zero cross-sectional error term, obtaining the sequence of estimates $\left\{\hat{\alpha}_t, \hat{\lambda}_t\right\}_{t=1}^T$.

Cross-Section Table 3 Outline

Probability Weighted Fama-MacBeth Regressions

I: For each asset i construct the consumption risk β's as

$$\hat{\beta}_i^j := -\frac{\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]\left[\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j\right]}{\left[\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]},$$

where $j \in \{EL, ET\}$ and γ is fixed, and as

$$\hat{\beta}_i^j := \sum_{t=1}^{T} \ln\left(\frac{C_t}{C_{t-1}}\right) R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^{T} \ln\left(\frac{C_t}{C_{t-1}}\right) \hat{p}_t^j\right]\left[\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j\right]$$

II: For each t, run the cross–sectional regression

$$R_{i,t}^e = \alpha_t + \hat{\beta}_i^j \lambda_t + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ is a mean zero cross-sectional error term, obtaining the sequence of estimates $\left\{\hat{\alpha}_t, \hat{\lambda}_t\right\}_{t=1}^{T}$.

▸ Cross-Section ▸ Table 3 ▸ Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^{T}$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j,\ \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)},\ E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j,\ \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section Table 3 Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^{T}$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j, \ \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)}, \ E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j, \ \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section Table 3 Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^{T}$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j, \ \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)}, \ E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j, \ \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section Table 3 Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^{T}$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j, \; \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)}, \; E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j, \; \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section Table 3 Outline

Rare Events and the cross-section of asset returns

Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
	Panel A: C-CAPM, $\gamma = 10$				
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
	Panel B: linearized C-CAPM				
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

Data Description
$\hat{P}^j$-weighted Fama-McBeth regressions

Rare Events and the cross-section of asset returns

Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
	Panel A: C-CAPM, $\gamma = 10$				
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
	Panel B: linearized C-CAPM				
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

Data Description
$\hat{P}^j$-weighted Fama-McBeth regressions

Rare Events and the cross-section of asset returns

Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
	Panel A: C-CAPM, $\gamma = 10$				
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
	Panel B: linearized C-CAPM				
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

Data Description

$\hat{P}^j$-weighted Fama-McBeth regressions

Rare Events and the cross-section of asset returns

Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
	Panel A: C-CAPM, $\gamma = 10$				
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
	Panel B: linearized C-CAPM				
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

▸ Data Description
▸ $\hat{P}^j$-weighted Fama-McBeth regressions

Rare Events and the cross-section of asset returns

Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
		Panel A: C-CAPM, $\gamma = 10$			
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
		Panel B: linearized C-CAPM			
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

Data Description
$\hat{P}^j$-weighted Fama-McBeth regressions

Rare Events and the cross-section of asset returns

Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
	Panel A: C-CAPM, $\gamma = 10$				
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
	Panel B: linearized C-CAPM				
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

Data Description $\hat{P}^j$-weighted Fama-McBeth regressions

Rare Events and the cross-section of asset returns

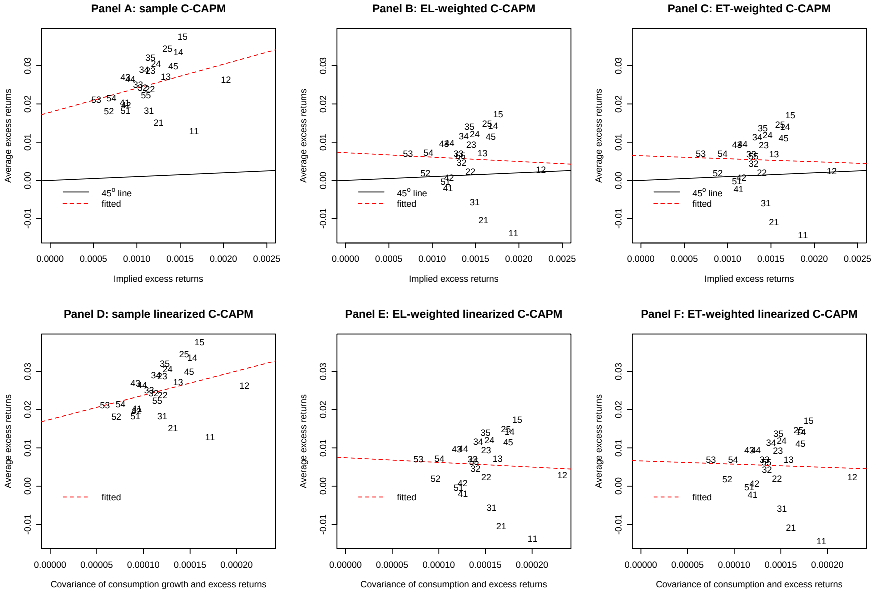
Table 3: Counterfactual Cross-Sectional Regressions

Moments:	R^2	$\hat{\alpha}$	$\hat{\lambda}$	$\Delta \frac{Var(\beta_i)}{Var\left(E\left[R^e_{i,t+1}\right]\right)}$	$\Delta Var(\rho_i)$
	Panel A: C-CAPM, $\gamma = 10$				
Sample	0.11	0.017 (0.005)	6.28 (5.04)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−1.15 (5.09)	−35.4%	−18.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−0.78 (5.09)	−38.2%	−12.9%
	Panel B: linearized C-CAPM				
Sample	0.12	0.017 (0.005)	63.35 (49.89)		
$\hat{P}^{EL}(\gamma)$	0.00	0.007 (0.006)	−12.18 (50.31)	−34.9%	−19.4%
$\hat{P}^{ET}(\gamma)$	0.00	0.006 (0.006)	−8.49 (50.37)	−37.8%	−13.7%

Fama-MacBeth (1973) standard errors in parenthesis.

Note: similar results $\forall \gamma \in]0, 10]$ and annual data

Data Description
$\hat{P}^j$-weighted Fama-McBeth regressions


Panel A: sample C-CAPM
Panel B: EL-weighted C-CAPM
Panel C: ET-weighted C-CAPM
Average excess returns
Implied excess returns
45° line
fitted
Panel D: sample linearized C-CAPM
Panel E: EL-weighted linearized C-CAPM
Panel F: ET-weighted linearized C-CAPM
Covariance of consumption growth and excess returns
Covariance of consumption and excess returns

Outline

1. Rare Events – Related Literature
2. Estimation
   - Sample Analogs and Rare Events
   - Information-Theoretic Alternatives
   - Estimation Results
3. Counterfactual Evidence
   - The Rare Events Distribution of the Data
   - How likely is the Equity Premium Puzzle?
   - Rare Events and the Cross-Section of Asset Returns
4. Conclusion

Conclusion

Key findings:

- Rare events are an unlikely explanation of the EPP:
  1. Rare-events-robust estimation approaches still reject the CCAPM and require a very high RRA to rationalize the EPP.
  2. If the data were generated by the rare events distribution needed to rationalize the EPP with a low RRA, the historically observed EPP would be very unlikely to arise.
  3. Rare-events substantially worsen the CCAPM ability of explaining the cross-section of asset returns, since they reduce the cross-sectional dispersion of consumption risk.

Methodological contribution:

- A data-driven, information-theoretic approach for the calibration of structural models.
- Can also be used for "dynamic" model simulation.

(Fairly) straightforward applications:

- Exchange rates, term structures, VaR, DSGE, non-nested model comparison, etc.

Conclusion

Key findings:

- Rare events are an unlikely explanation of the EPP:
  1. Rare-events-robust estimation approaches still reject the CCAPM and require a very high RRA to rationalize the EPP.
  2. If the data were generated by the rare events distribution needed to rationalize the EPP with a low RRA, the historically observed EPP would be very unlikely to arise.
  3. Rare-events substantially worsen the CCAPM ability of explaining the cross-section of asset returns, since they reduce the cross-sectional dispersion of consumption risk.

Methodological contribution:

- A data-driven, information-theoretic approach for the calibration of structural models.
- Can also be used for "dynamic" model simulation.

(Fairly) straightforward applications:

- Exchange rates, term structures, VaR, DSGE, non-nested model comparison, etc.

Conclusion

Key findings:

- Rare events are an unlikely explanation of the EPP:
  1. Rare-events-robust estimation approaches still reject the CCAPM and require a very high RRA to rationalize the EPP.
  2. If the data were generated by the rare events distribution needed to rationalize the EPP with a low RRA, the historically observed EPP would be very unlikely to arise.
  3. Rare-events substantially worsen the CCAPM ability of explaining the cross-section of asset returns, since they reduce the cross-sectional dispersion of consumption risk.

Methodological contribution:

- A data-driven, information-theoretic approach for the calibration of structural models.
- Can also be used for "dynamic" model simulation.

(Fairly) straightforward applications:

- Exchange rates, term structures, VaR, DSGE, non-nested model comparison, etc.

Conclusion

Key findings:

- Rare events are an unlikely explanation of the EPP:
  1. Rare-events-robust estimation approaches still reject the CCAPM and require a very high RRA to rationalize the EPP.
  2. If the data were generated by the rare events distribution needed to rationalize the EPP with a low RRA, the historically observed EPP would be very unlikely to arise.
  3. Rare-events substantially worsen the CCAPM ability of explaining the cross-section of asset returns, since they reduce the cross-sectional dispersion of consumption risk.

Methodological contribution:

- A data-driven, information-theoretic approach for the calibration of structural models.
- Can also be used for "dynamic" model simulation.

(Fairly) straightforward applications:

- Exchange rates, term structures, VaR, DSGE, non-nested model comparison, etc.

Conclusion

Key findings:

- Rare events are an unlikely explanation of the EPP:
    1. Rare-events-robust estimation approaches still reject the CCAPM and require a very high RRA to rationalize the EPP.
    2. If the data were generated by the rare events distribution needed to rationalize the EPP with a low RRA, the historically observed EPP would be very unlikely to arise.
    3. Rare-events substantially worsen the CCAPM ability of explaining the cross-section of asset returns, since they reduce the cross-sectional dispersion of consumption risk.

Methodological contribution:

- A data-driven, information-theoretic approach for the calibration of structural models.
- Can also be used for "dynamic" model simulation.

(Fairly) straightforward applications:

- Exchange rates, term structures, VaR, DSGE, non-nested model comparison, etc.

Outline

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Table 1 Cross-Section Table 3 Outline

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Table 1 Cross-Section Table 3 Outline

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value $\rightarrow$ higher returns) and the value premium (high book values relative to market equity $\rightarrow$ higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Table 1 Cross-Section Table 3 Outline

Data Description

- Market return proxy: CRSP value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ.
- Risk-free rate proxy: one-month Treasury Bill rate
- Consumption: NIPA per capita personal consumption expenditures on nondurable goods

Samples: Quarterly: 1947:Q1-2003:Q3. Annual: 1929-2006.

Estimation results

- Cross-sectional analysis: quarterly returns on the 25 Fama-French (1992) portfolios.
- Designed to focus on the size effect (small market value → higher returns) and the value premium (high book values relative to market equity → higher returns).
- Intersections of 5 portfolios formed on size and 5 portfolios formed on the book equity to market equity ratio.

Table 1 | Cross-Section | Table 3 | Outline

Probability Weighted Fama-MacBeth Regressions

I: For each asset i construct the consumption risk β's as

$$\hat{\beta}_i^j := -\frac{\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]\left[\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j\right]}{\left[\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]},$$

where $j \in \{EL, ET\}$ and γ is fixed, and as

$$\hat{\beta}_i^j := \sum_{t=1}^{T} \ln\left(\frac{C_t}{C_{t-1}}\right) R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^{T} \ln\left(\frac{C_t}{C_{t-1}}\right) \hat{p}_t^j\right]\left[\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j\right]$$

II: For each t, run the cross–sectional regression

$$R_{i,t}^e = \alpha_t + \hat{\beta}_i^j \lambda_t + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ is a mean zero cross-sectional error term, obtaining the sequence of estimates $\left\{\hat{\alpha}_t, \hat{\lambda}_t\right\}_{t=1}^{T}$.

Cross-Section Table 3 Outline

Probability Weighted Fama-MacBeth Regressions

I: For each asset i construct the consumption risk β's as

$$\hat{\beta}_i^j := -\frac{\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]\left[\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j\right]}{\left[\sum_{t=1}^{T}\left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \hat{p}_t^j\right]},$$

where $j \in \{EL, ET\}$ and γ is fixed, and as

$$\hat{\beta}_i^j := \sum_{t=1}^{T} \ln\left(\frac{C_t}{C_{t-1}}\right) R_{i,t}^e \hat{p}_t^j - \left[\sum_{t=1}^{T} \ln\left(\frac{C_t}{C_{t-1}}\right) \hat{p}_t^j\right]\left[\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j\right]$$

II: For each t, run the cross–sectional regression

$$R_{i,t}^e = \alpha_t + \hat{\beta}_i^j \lambda_t + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ is a mean zero cross-sectional error term, obtaining the sequence of estimates $\left\{\hat{\alpha}_t, \hat{\lambda}_t\right\}_{t=1}^{T}$.

Cross-Section | Table 3 | Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^T$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T}\sum_{t=1}^{T}(\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j,\ \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T}\sum_{t=1}^{T}\left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)},\ E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j,\ \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section Table 3 Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^T$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j, \ \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)}, \ E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j, \ \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section | Table 3 | Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^T$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j, \ \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{p}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{p}^j(\gamma)}\left[R_{i,t}^e\right]\right)}, \ E^{\hat{p}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j, \ \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section | Table 3 | Outline

III: Construct point estimates for α and λ as

$$\hat{\alpha} := \sum_{t=1}^{T} \hat{\alpha}_t \hat{p}_t^j \text{ and } \hat{\lambda} := \sum_{t=1}^{T} \hat{\lambda}_t \hat{p}_t^j.$$

Note: $\hat{\alpha}$ and $\hat{\lambda}$ are equivalent to the ones we would obtain from the cross-sectional regression

$$\sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j = \alpha + \hat{\beta}_i^j \lambda + \varepsilon_i$$

IV: Use the weighted sampling variation of $\{\alpha_t, \lambda_t\}_{t=1}^{T}$ to construct the standard deviations of the estimators

$$\sigma^2(\hat{\alpha}) := \frac{1}{T} \sum_{t=1}^{T} (\hat{\alpha}_t - \hat{\alpha})^2 \hat{p}_t^j, \; \sigma^2\left(\hat{\lambda}\right) := \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\lambda}_t - \hat{\lambda}\right)^2 \hat{p}_t^j.$$

V: The cross-sectional R^2 for these regressions is constructed as

$$R^2 := 1 - \frac{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] - \hat{R}_{i,t}^e\right)}{Var\left(E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right]\right)}, \; E^{\hat{P}^j(\gamma)}\left[R_{i,t}^e\right] := \sum_{t=1}^{T} R_{i,t}^e \hat{p}_t^j, \; \hat{R}_{i,t}^e := \hat{\alpha} + \hat{\beta}_i^j \hat{\lambda}.$$

Cross-Section Table 3 Outline