Habit, Long Run Risk, Prospect?
A Statistical Inquiry

Eric Aldrich
Duke University
Department of Economics
Durham NC 27708-0097 USA

A. Ronald Gallant
Duke University
Fuqua School of Business
Durham NC 27708-0120 USA

Robert E. McCulloch
University of Texas
Graduate School of Business
Austin TX 78712-1175 USA

Outline

- Overview
  - Goal of this work
  - Related literature
  - Data
  - Summary of results
- Models considered
- Bayesian inference for general scientific models
- Results
- Sensitivity analysis

Goal

- Systematic comparison of several macro/finance models.

- Adhering to the principles of statistical science.

- Using Bayesian methods because data are sparse.
    - ▷ Prior information augments the data.

Statistical Literature – Frequentist

- Bansal, Ravi, A. Ronald Gallant, and George Tauchen (2007), "Rational Pessimism, Rational Exuberance, and Asset Pricing Models," *Review of Economic Studies* 74, 1005–1033.

- Concerns
  - ▷ Modified proposer's models – imposed co-integration
  - ▷ Used a general purpose solution method.
  - ▷ A statistical comparison was defeated by sparse data.
  - ▷ Models compared by performance on macro "puzzles"

Statistical Literature – Bayesian

- Gallant, A. Ronald, and Robert E. McCulloch (2009), "On the Determination of General Scientific Models with Application to Asset Pricing," *Journal of the American Statistical Association* 104, 117–131.
  - ▷ Related: Dejong, Ingram, and Whiteman (2000), Del Negro and Schorfheide (2004), etc.

- Advantages
  - ▷ Can be used when no likelihood is available.
  - ▷ Augments sparse data with prior information.
  - ▷ Permits latent variables

Macro/Finance Literature

- Current practice
  - ▷ List some puzzles – i.e. list some sample moments
  - ▷ Propose a model
  - ▷ Check it against the list of puzzles

- Concerns
  - ▷ Chaotic – lists vary
  - ▷ Few organized head-to-head comparisons
  - ▷ In the hands of the proposers

Models Considered

- **Habit**
  Campbell, J. Y., and J. Cochrane. (1999). "By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107, 205–251.

- **Long run risk**
  Bansal, R., and A. Yaron. (2004). "Risks For the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance* 59, 1481–1509.

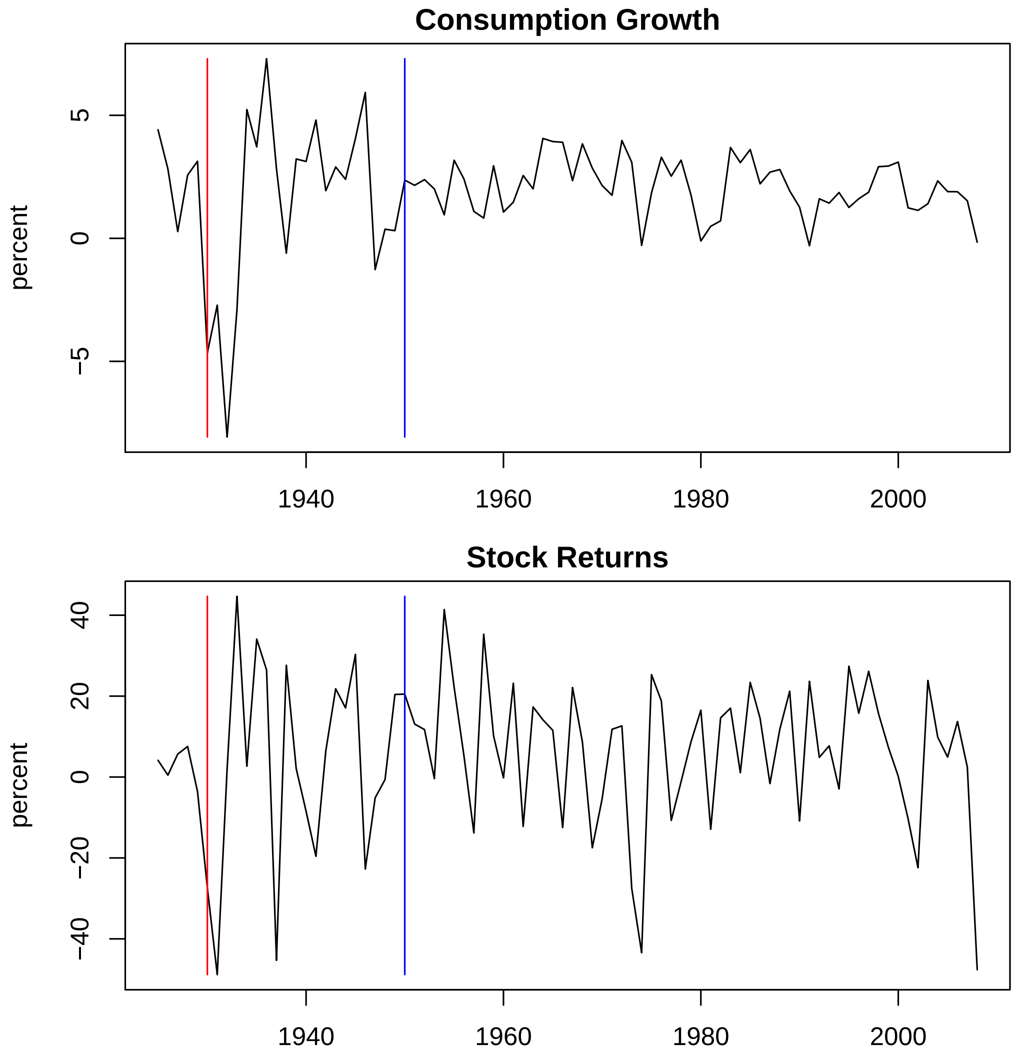
- **Prospect theory**
  Barberis, N, M.Huang, and T. Santos, "Prospect Theory and Asset Prices," *The Quarterly Journal of Economics* 116, 1–53.

Fairness

- Use the proposer's solution method.

- Use the same prior across all models.
  - ▷ Augmented by weak prior information from the proposer's calibration to guarantee that the prior is proper.

Fig 1. Data, 1925–2008



The red line is at 1930 and the blue at 1950.

Findings

- Models roughly equivalent when fit to equity returns in a relative comparison.
- Long run risk preferred for joint returns and consumption growth over 1930–2008 in a relative comparison.
- Habit persistence preferred for joint returns and consumption growth over 1950–2008 in a relative comparison.
- None of the models fit returns and consumption growth over 1930–2008 in an absolute assessment.
- All models add counter factuals.
- Counter factuals cannot be removed by a prior.
- Incomplete: Findings may change.

Outline

- Overview
- Models considered
  - Habit persistence
  - Long run risk
  - Prospect theory
- Bayesian inference for general scientific models
- Results
- Sensitivity analysis

Habit Persistence Asset Pricing Model

Driving Processes

$$\text{Consumption: } c_t - c_{t-1} = g + v_t$$

$$\text{Dividends: } d_t - d_{t-1} = g + w_t$$

$$\text{Random Shocks: } \begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim \text{NID}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma\sigma_w \\ \rho\sigma\sigma_w & \sigma_w^2 \end{pmatrix}\right]$$

The time increment is one month. Lower case denotes logarithms of upper case quantities; i.e. $c_t = \log(C_t)$, $d_t = \log(D_t)$. From Campbell and Cochrane (1999).

Habit Persistence Asset Pricing Model

Utility function

$$\mathcal{E}_0\left(\sum_{t=0}^{\infty}\delta^t\frac{(S_tC_t)^{1-\gamma}-1}{1-\gamma}\right),$$

Habit persistence

$$\text{Surplus ratio: } s_t-\bar{s}=\phi\,(s_{t-1}-\bar{s})+\lambda(s_{t-1})v_{t-1}$$

$$\text{Sensitivity function: } \lambda(s)=\begin{cases}\frac{1}{\bar{S}}\sqrt{1-2(s-\bar{s})}-1 & s_t\le s_{\max}\\ 0 & s_t>s_{\max}\end{cases}$$

$\mathcal{E}_t$ is conditional expectation with respect to S_t, S_{t-1}, ... Lower case denotes logarithms of upper case quantities: $s_t=\log(S_t)$. $\bar{S}$ and $s_{\max}$ can be computed from model parameters $\theta=(g,\sigma,\rho,\sigma_w,\phi,\delta,\gamma)$ as $\bar{S}=\sigma\sqrt{\gamma/(1-\phi)}$, $s_{\max}=\bar{s}+(1-\bar{S}^2)/2$. From Campbell and Cochrane (1999).

Habit Persistence Asset Pricing Model

Utility function

$$\mathcal{E}_0\left(\sum_{t=0}^{\infty} \delta^t \frac{(S_t C_t)^{1-\gamma} - 1}{1-\gamma}\right)$$

Surplus ratio

$$S_t/\bar{S} = e^{\phi}\,(S_{t-1}/\bar{S})e^{\lambda_{t-1}v_{t-1}}$$

Substitute

$$(e^{\phi})^{1-\gamma}\mathcal{E}_0\left(\sum_{t=0}^{\infty} \delta^t \frac{[(S_{t-1}/\bar{S})e^{\lambda_{t-1}v_{t-1}}C_t]^{1-\gamma} - 1}{1-\gamma}\right)$$

Conclusion

Product of an decreasing function of ϕ and and increasing function of δ.

$\Rightarrow$ ϕ and δ are not separately identified.

$\Rightarrow$ Identification comes from the prior.

Habit Persistence Asset Pricing Model

Return on dividends

$$V(S_t) = \mathcal{E}_t \left\{ \delta \left(\frac{S_{t+1} C_{t+1}}{S_t C_t} \right)^{-\gamma} \left(\frac{D_{t+1}}{D_t} \right) \left[1 + V(S_{t+1}) \right] \right\}$$

$$r_{dt} = \log \left[\frac{1 + V(S_t)}{V(S_{t-1})} \left(\frac{D_t}{D_{t-1}} \right) \right]$$

$V(\cdot)$ is defined as the solution of the Euler condition above. It is the price dividend ratio; i.e. $P_{dt}/D_t = V(S_t)$, where P_{dt} is the price of the asset that pays the dividend stream. r_{dt} is the logarithmic real return, i.e. $r_{dt} = \log(P_{dt} + D_t) - \log(P_{d,t-1})$, where P_{dt} and D_t are measured in real (inflation adjusted) dollars. Dividend error can be integrated out analytically. Consumption error integrated by quadrature. From Campbell and Cochrane (1999).

Habit Persistence Asset Pricing Model

Solution Method

Approximate the log policy function

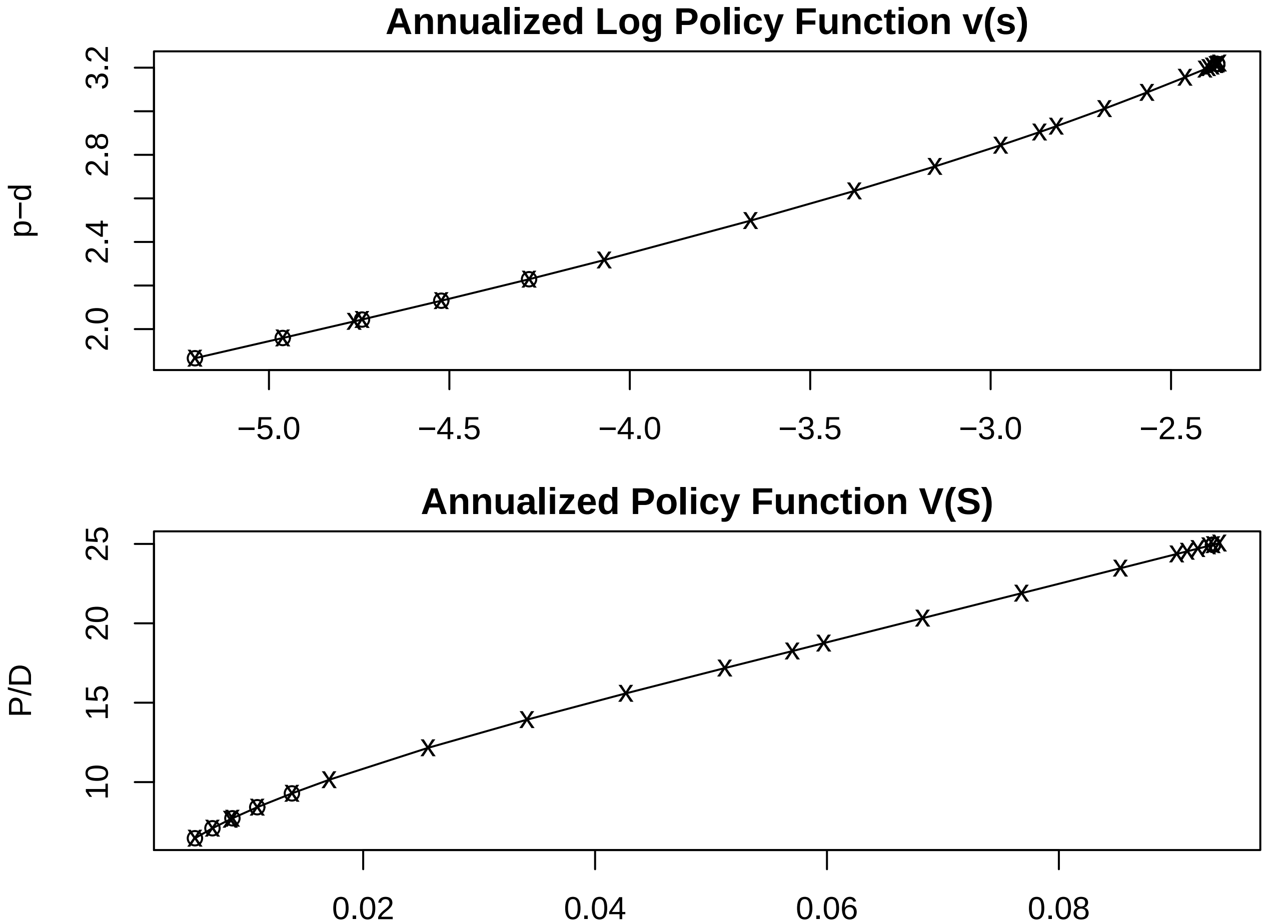
$$v(s_t) = \log V(e^{s_t})$$

by a piecewise linear function and use policy function iteration.

Campbell and Cochrane used Gauss's intquad1 and set join points at $\bar{s}$, s_{max}, $s_{\text{max}}-0.01$, $s_{\text{max}}-0.02$, $s_{\text{max}}-0.03$, $s_{\text{max}}-0.04$, and $\log[iS/(m+1)]$ for $i = 1, \ldots, m = 10$. We used Gauss-Hermite quadrature; we added the abscissae of the Gauss-Hermite quadrature formula at the maximum and minimum of the above join points; we deleted all points less than 0.001 apart.

Figure 2, next slide, plots the approximation at the Campbell and Cochrane parameter values.

Fig 2. Piecewise Linear Approximation



x's mark Campbell and Cochrane join points; o's mark extra join points from the quadrature rule.

Habit Persistence Asset Pricing Model

Risk Free Rate

$$r_{ft} = -\log\left\{\mathcal{E}_t\left[\delta\left(\frac{S_{t+1}C_{t+1}}{S_t C_t}\right)^{-\gamma}\right]\right\}$$

r_{ft} is the logarithmic return on an asset that pays one real dollar one month hence with certainty. From Campbell and Cochrane (1999).

Solution method is similar to the foregoing.

Habit Persistence Asset Pricing Model

Large Model Output

Given model parameters

$$\theta = (g, \sigma, \rho, \sigma_w, \phi, \delta, \gamma)$$

simulate monthly and aggregate to annual:

$$C_t^a = \sum_{k=0}^{11} C_{12t-k}$$

$$c_t^a = \log(C_t^a)$$

$$r_{dt}^a = \sum_{k=0}^{11} r_{d,12t-k}$$

$$r_{ft}^a = \sum_{k=0}^{11} r_{f,12t-k}$$

Habit Persistence Asset Pricing Model

Prior Distribution

$$p(\theta) = \mathsf{N}\left[r_f \,|\, 0.89, \left(\frac{1}{1.96}\right)^2\right] \prod_{i=1}^{p} \mathsf{N}\left[\theta_i \,|\, \theta_i^*, \left(\frac{0.1\theta_i^*}{1.96}\right)^2\right]$$

where the θ_i^* are the calibrated values from Campbell and Cochrane (1999).

The scale factor on ψ and δ is 0.001 rather than 0.1.

This is not an independence prior (next slide).

The first term of the prior dominates (slide after next).

Table 1. Correlation Matrix of the Habit Model Prior

	g	σ	ρ	σ_w	ϕ	δ	γ
g	1.00	0.05	0.05	-0.02	-0.02	0.18	0.06
σ	0.05	1.00	0.04	-0.07	0.03	0.05	0.06
ρ	0.05	0.04	1.00	0.03	0.10	0.04	0.08
σ_w	-0.02	-0.07	0.03	1.00	-0.02	0.02	-0.01
ϕ	-0.02	0.03	0.10	-0.02	1.00	0.47	0.32
δ	0.18	0.05	0.04	0.02	0.47	1.00	-0.26
γ	0.06	0.06	0.08	-0.01	0.32	-0.26	1.00

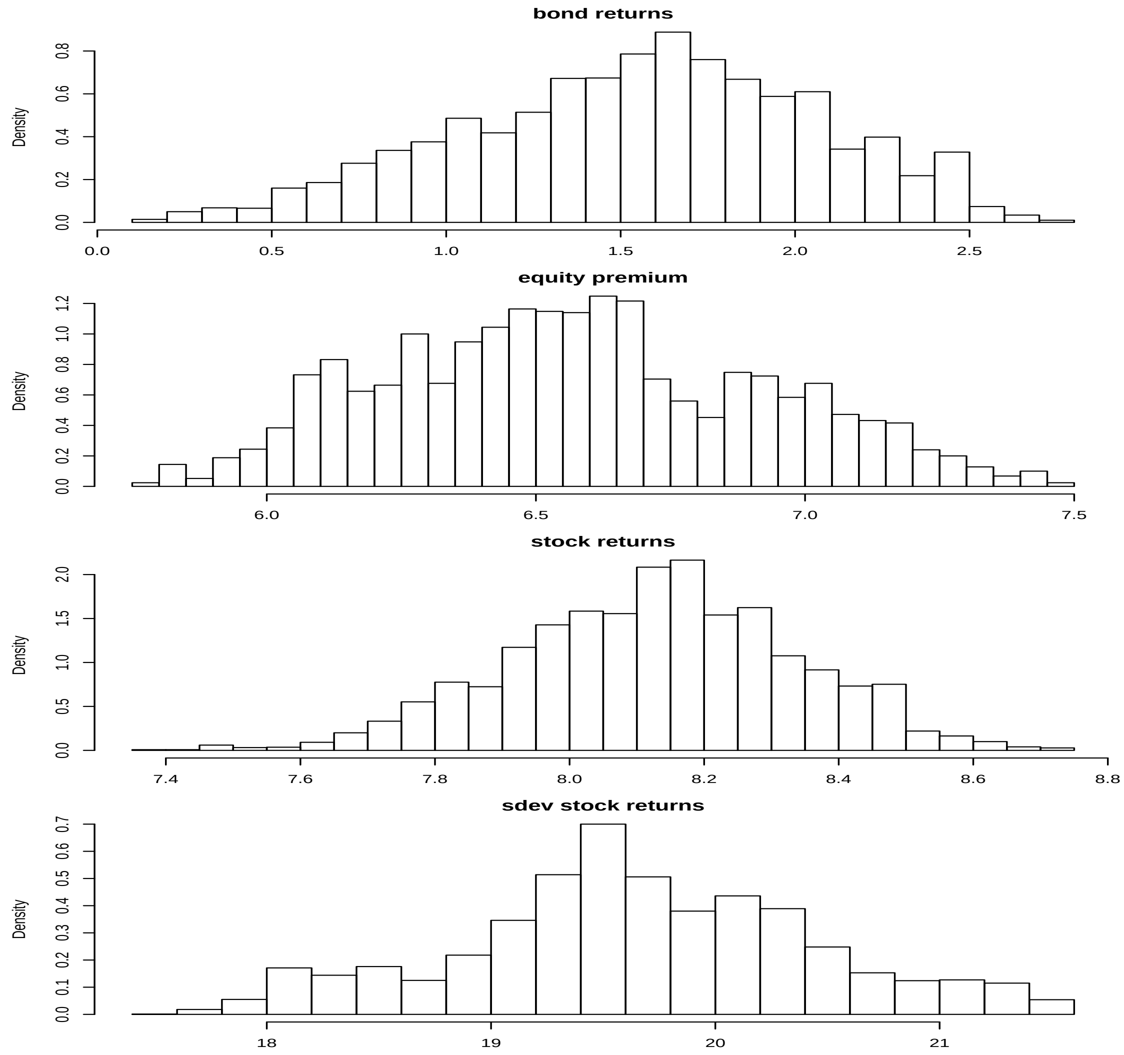
Table 2. Habit Model

	Prior		Posterior	
Parameter	Mode	Std.Dev.	Mode	Std.Dev.
g	0.00157928	0.00008886	0.00164032	0.00007627
σ	0.00430298	0.00021240	0.00503540	0.00017226
ρ	0.19970703	0.00999614	0.19287109	0.00901699
σ_w	0.03240967	0.00162020	0.03240967	0.00126099
ϕ	0.98482513	0.00039484	0.98622894	0.00028433
δ	0.98744965	0.00070869	0.98824310	0.00033228
γ	1.99414062	0.09276708	2.04101562	0.07825575
r_f	0.77806800	0.55238138	1.34504400	0.50401723
$r_d - r_f$	7.78410000	0.38838322	6.83274000	0.34762939
σ_{r_d}	21.46457547	1.04018247	20.14800238	0.79315876

Parameter values are for the monthly frequency. Returns are annualized. Mode is the mode of the multivariate density. It actually occurs in the MCMC chain whereas other measures of central tendency may not even satisfy support conditions. In the data, $r_d - r_f = 5.59 - 0.89 = 5.5$ and $\sigma_{r_d} = 19.72$. The auxiliary model is GARCH with normal errors. The data are annual stock returns and consumption growth 1930–2008.

Not updated to 0.001 scale factor for ϕ and δ.

Fig 3. Posterior Returns



Long Run Risk Asset Pricing Model

Driving Processes

$$
\begin{aligned}
\text{Consumption:}\quad & c_{t+1} - c_t = \mu_c + x_t + \sigma_t \eta_{t+1} \\
\text{Long Run Risk:}\quad & x_{t+1} = \rho x_t + \phi_e \sigma_t e_{t+1} \\
\text{Stochastic Volatility:}\quad & \sigma_{t+1}^2 = \bar{\sigma}^2 + \nu(\sigma_t^2 - \bar{\sigma}^2) + \sigma_w w_{t+1} \\
\text{Dividends:}\quad & d_{t+1} - d_t = \mu_d + \phi_d x_t + \pi_d \sigma_t \eta_{t+1} + \phi_u \sigma_t u_{t+1} \\
\text{Random Shocks:}\quad & \begin{pmatrix} \eta_t \\ e_t \\ w_t \\ u_t \end{pmatrix} \sim \text{NID}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right]
\end{aligned}
$$

The time increment is one month. Lower case denotes logarithms of upper case quantities; i.e. $c_t = \log(C_t)$, $d_t = \log(D_t)$. From Bansal, Kiku, and Yaron (2007).

Long Run Risk Asset Pricing Model

Epstein-Zin utility function

$$U_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(\mathcal{E}_t U_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

where

γ	is the coefficient of risk aversion
ψ	is the elasticity of inter temporal substitution

and

$$\theta = \frac{1-\gamma}{1-1/\psi}$$

$\mathcal{E}_t$ is conditional expectation with respect to x_t, σ_t.

Long Run Risk Asset Pricing Model

Return on consumption

$$
\begin{aligned}
\text{mrs}_{t+1} &= \delta^{\theta} \exp[-(\theta/\psi)(c_{t+1} - c_t) + (\theta - 1) r_{c,t+1}] \\
V_C(x_t, \sigma_t) &= \mathcal{E}_t \left\{ \text{mrs}_{t+1} \left(\frac{C_{t+1}}{C_t} \right) \left[1 + V_C(x_{t+1}, \sigma_{t+1}) \right] \right\} \\
r_{ct} &= \log \left[\frac{1 + V_C(x_t, \sigma_t)}{V_C(x_{t-1}, \sigma_{t-1})} \left(\frac{C_t}{C_{t-1}} \right) \right]
\end{aligned}
$$

$V_C(\cdot)$ is defined as the solution of the Euler condition above. It is the price consumption ratio; i.e. $P_{ct}/C_t = V_C(x_t, \sigma_t)$, where P_{ct} is the price of the asset that pays the consumption stream. r_{ct} is the logarithmic real return, i.e. $r_{ct} = \log(P_{ct} + C_t) - \log(P_{c,t-1})$, where P_{ct} and C_t are measured in real (inflation adjusted) dollars.

Long Run Risk Asset Pricing Model

Solution Method

Use the log linear approximation

$$\begin{aligned} r_{c,t+1} &\doteq \kappa_0 + \kappa_1 z_{t+1} + \Delta c_{t+1} - z_t \\ \kappa_1 &= [\exp(\bar{z})]/[1 + \exp(\bar{z})] \\ k_0 &= \log[1 + \exp(\bar{z})] - \kappa_1 \bar{z} \end{aligned}$$

where $z_t = \log(P_{c,t}/C_t)$ and $\bar{z}$ is its endogenous mean.

To compute $\bar{z}$, use the approximation

$$\begin{aligned} z_t &\doteq A_0(\bar{z}) + A_1(\bar{z})\, x_t + A_2(\bar{z})\, \sigma_t^2 \\ A_i(\bar{z}) &= \text{tedious expressions in model parameters and } \bar{z} \end{aligned}$$

and solve the fixed point problem

$$\bar{z} = A_0(\bar{z}) + A_1(\bar{z})\, x_t + A_2(\bar{z})\, \sigma_t^2$$

Long Run Risk Asset Pricing Model

Return on dividends

$$\begin{aligned}
\text{mrs}_{t+1} &= \delta^{\theta} \exp[-(\theta/\psi)(c_{t+1} - c_t) + (\theta - 1) r_{c,t+1}] \\
V_D(x_t, \sigma_t) &= \mathcal{E}_t \left\{ \text{mrs}_{t+1} \left(\frac{D_{t+1}}{D_t} \right) \left[1 + V_D(x_{t+1}, \sigma_{t+1}) \right] \right\} \\
r_{dt} &= \log \left[\frac{1 + V_D(x_t, \sigma_t)}{V_D(x_{t-1}, \sigma_{t-1})} \left(\frac{D_t}{D_{t-1}} \right) \right]
\end{aligned}$$

$V_D(\cdot)$ is defined as the solution of the Euler condition above. It is the price dividend ratio; i.e. $P_{dt}/D_t = V_D(x_t, \sigma_t)$, where P_{ct} is the price of the asset that pays the dividend stream. r_{dt} is the logarithmic real return, i.e. $r_{dt} = \log(P_{dt} + D_t) - \log(P_{d,t-1})$, where P_{dt} and D_t are measured in real (inflation adjusted) dollars.

Solution method is similar to the foregoing.

Long Run Risk Asset Pricing Model

Risk Free Rate

$$r_{ft} = -\log \mathcal{E}_t \left\{ \delta^{\theta} \exp \left[-(\theta/\psi)(c_{t+1} - c_t) + (\theta - 1) r_{c,t+1} \right] \right\}$$

r_{ft} is the logarithmic return on an asset that pays one real dollar one month hence with certainty.

Solution method is similar to the foregoing.

Long Run Risk Asset Pricing Model

Large Model Output

Given model parameters

$$\theta = (\delta, \gamma, \psi, \mu_c, \rho, \phi_e, \bar{\sigma}^2, \eta, \sigma_w, \mu_d, \phi_u)$$

simulate monthly and aggregate to annual:

$$C_t^a = \sum_{k=0}^{11} C_{12t-k}$$

$$c_t^a = \log(C_t^a)$$

$$r_{dt}^a = \sum_{k=0}^{11} r_{d,12t-k}$$

$$r_{ft}^a = \sum_{k=0}^{11} r_{f,12t-k}$$

Long Run Risk Asset Pricing Model

Prior Distribution

$$p(\theta) = \mathsf{N}\left[r_f \,|\, 0.89, \left(\frac{1}{1.96}\right)^2\right] \prod_{i=1}^{p} \mathsf{N}\left[\theta_i \,|\, \theta_i^*, \left(\frac{0.1\theta_i^*}{1.96}\right)^2\right]$$

where the θ_i^* are the calibrated values from Kiku (2006).

The standard deviation on ρ and ν is 0.01 rather than 0.1.

This is not an independence prior (next slide).

The first term dominates (slide after next).

Table 3. Correlation Matrix of the Long Run Risk Model Prior

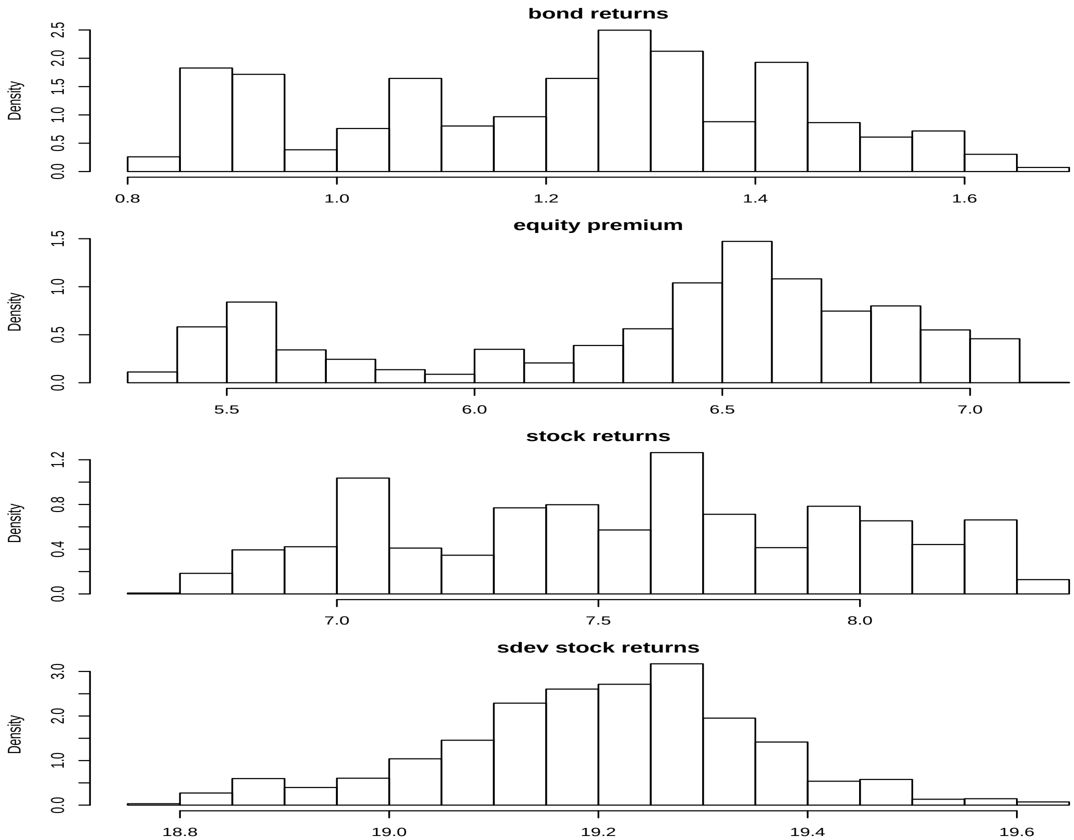
	δ	γ	ψ	μ_c	ρ	ϕ_e	$\bar{\sigma}^2$	ν	σ_w	μ_d	ϕ_d	π_d	ϕ_u
δ	1.00	-0.44	-0.16	-0.08	0.20	-0.31	0.19	-0.12	-0.12	0.18	-0.12	0.31	0.05
γ	-0.44	1.00	-0.04	-0.15	-0.36	0.11	0.07	0.18	0.18	-0.05	-0.17	-0.27	0.00
ψ	-0.16	-0.04	1.00	0.06	-0.07	-0.05	-0.08	-0.04	0.11	0.13	-0.15	0.17	0.05
μ_c	-0.08	-0.15	0.06	1.00	0.22	0.19	-0.12	0.07	0.02	0.05	0.22	0.14	-0.04
ρ	0.20	-0.36	-0.07	0.22	1.00	0.20	-0.42	0.05	-0.03	0.08	0.17	0.16	-0.26
ϕ_e	-0.31	0.11	-0.05	0.19	0.20	1.00	-0.26	0.19	-0.02	-0.21	0.29	-0.22	-0.17
$\bar{\sigma}^2$	0.19	0.07	-0.08	-0.12	-0.42	-0.26	1.00	-0.40	0.09	-0.07	-0.26	-0.02	0.33
ν	-0.12	0.18	-0.04	0.07	0.05	0.19	-0.40	1.00	-0.14	-0.05	0.36	-0.03	-0.04
σ_w	-0.12	0.18	0.11	0.02	-0.03	-0.02	0.09	-0.14	1.00	0.05	-0.31	-0.18	0.03
μ_d	0.18	-0.05	0.13	0.05	0.08	-0.21	-0.07	-0.05	0.05	1.00	-0.05	0.30	0.10
ϕ_d	-0.12	-0.17	-0.15	0.22	0.17	0.29	-0.26	0.36	-0.31	-0.05	1.00	-0.04	-0.31
π_d	0.31	-0.27	0.17	0.14	0.16	-0.22	-0.02	-0.03	-0.18	0.30	-0.04	1.00	0.02
ϕ_u	0.05	0.00	0.05	-0.04	-0.26	-0.17	0.33	-0.04	0.03	0.10	-0.31	0.02	1.00

Table 4. Long Run Risk Model

	Prior		Posterior	
Parameter	Mode	Std.Dev.	Mode	Std.Dev.
δ	0.99961472	0.00040805	0.99938583	0.00008310
γ	9.84765625	0.49029970	10.44921875	0.09925565
ψ	1.49609375	0.07646445	1.49609375	0.06036612
μ_c	0.00148392	0.00006882	0.00152969	0.00006995
ρ	0.98410034	0.01087285	0.98672485	0.00081755
ϕ_e	0.03201294	0.00170142	0.03305054	0.00054844
$\bar{\sigma}^2$	0.00004050	0.00000208	0.00004074	0.00000055
ν	0.98800659	0.01697010	0.99166870	0.00092517
σ_w	0.00000168	0.00000009	0.00000172	0.00000006
μ_d	0.00120926	0.00006146	0.00115585	0.00007563
ϕ_d	2.80859375	0.14518166	2.79296875	0.05577070
π_d	4.06640625	0.20073309	4.06640625	0.09732309
ϕ_u	6.12890625	0.29974344	6.35546875	0.06256562
r_f	0.94983600	0.48203234	1.02063600	0.21291457
$r_d - r_f$	4.20937200	3.12923103	6.06766800	0.48360307
σ_{r_d}	18.30042622	1.30394579	19.25797497	0.14919467

Parameter values are for the monthly frequency. Returns are annualized. Mode is the mode of the multivariate density. It actually occurs in the MCMC chain whereas other measures of central tendency may not even satisfy support conditions. In the data, $r_d - r_f = 5.59 - 0.89 = 5.5$ and $\sigma_{r_d} = 19.72$. The auxiliary model is GARCH with normal errors. The data are annual stock returns and consumption growth 1930–2008.

Fig 4. Posterior Returns



Prospect Theory Asset Pricing Model

Driving Processes

$$\text{Aggregate Consumption: } \bar{c}_{t+1} - \bar{c}_t = g_C + \sigma_C \eta_{t+1}$$

$$\text{Dividends: } d_{t+1} - d_t = g_D + \sigma_D \epsilon_{t+1}$$

$$\text{Random Shocks: } \begin{pmatrix} \eta_t \\ \epsilon_t \end{pmatrix} \sim \text{NID} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix} \right]$$

$\bar{C}_t$ is aggregate, per capita consumption which is exogenous to the agent. The time increment is one year. Lower case denotes logarithms of upper case quantities; i.e. $\bar{c}_t = \log(\bar{C}_t)$, $d_t = \log(D_t)$. All variables are real. From Barberis, Huang, Santos (2001).

Prospect Theory Asset Pricing Model

Other Model Variables

- Gross Stock Return: R_t
- Gross Risk Free Rate: $R_f = \rho^{-1} \exp\left(\gamma g_C - \gamma^2 \sigma_C^2 / 2\right)$
- Allocation to Risky Asset: S_t
- Gain or Loss: $X_{t+1} = S_t(R_{t+1} - R_f)$
- Benchmark Level (State Variable): $z_{t+1} = \eta \left(z_t \frac{\bar{R}}{R_{t+1}}\right) + (1 - \eta)$
- Choose $\bar{R}$ to make Median $\{z_t\} = 1$
- The Agent's Consumption: C_t

Prospect Theory Asset Pricing Model

Utility function

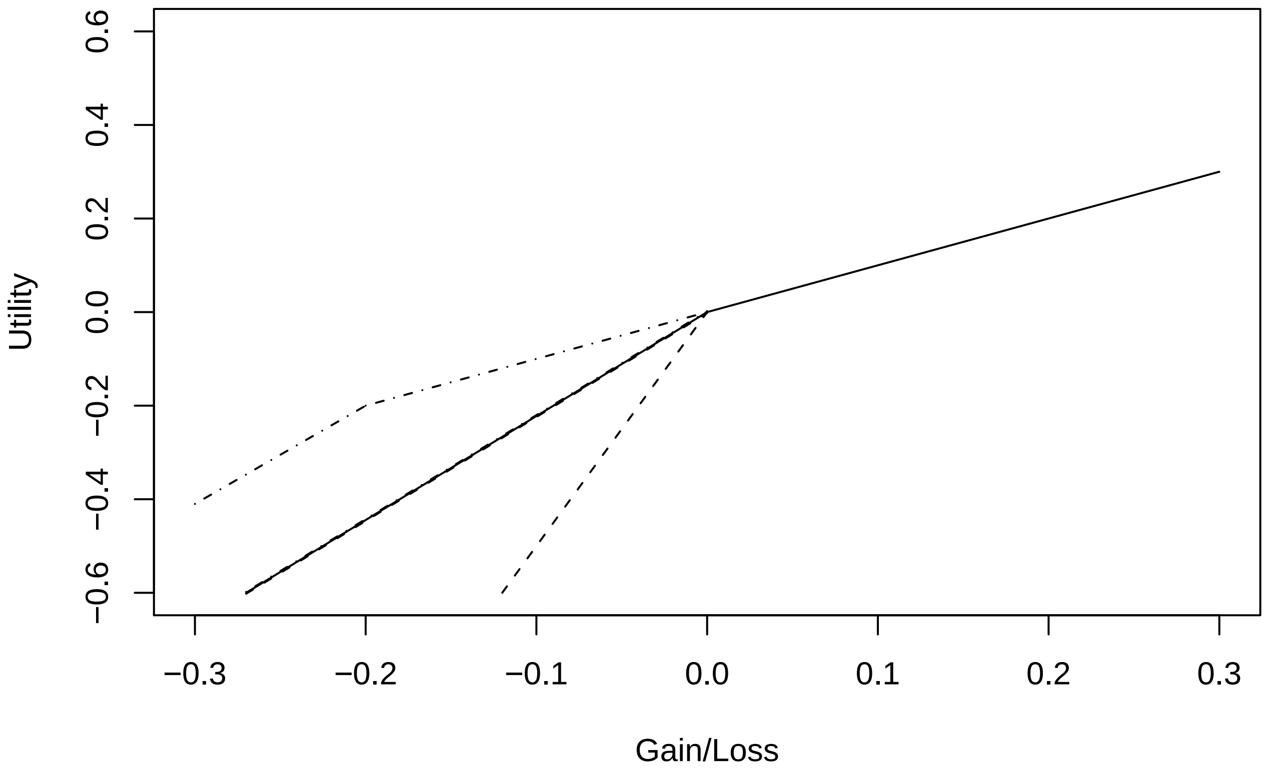
$$\mathcal{E}_0\left[\sum_{t=0}^{\infty}\left(\rho^t \frac{C_t^{1-\gamma}-1}{1-\gamma} + b_0 \bar{C}_t^{-\gamma}\rho^{t+1}\left[S_t\,\hat{v}(R_{t+1}, z_t)\right]\right)\right]$$

Utility from Gains and Losses: $\left[S_t\,\hat{v}(R_{t+1}, z_t)\right]$

$$\begin{aligned}
&\hat{v}(R_{t+1}, z_t) \\
&\quad = R_{t+1} - R_f && z_t \le 1, R_{t+1} \ge z_t R_f \\
&\quad = (z_t R_f - R_f) + \lambda(R_{t+1} - z_t R_f) && z_t \le 1, R_{t+1} < z_t R_f \\
&\quad = R_{t+1} - R_f && z_t > 1, R_{t+1} \ge R_f \\
&\quad = \lambda(z_t)(R_{t+1} - R_f) && z_t > 1, R_{t+1} < R_f
\end{aligned}$$

$$\lambda(z_t) = \lambda + k(z_t - 1)$$

Fig 5. Utility of Gains and Losses



The dot-dash line represents the case where the investor has prior gains ($z < 1$), the dashed line the case of prior losses ($z > 1$), and the solid line the case where the investor has neither prior gains nor losses ($z = 1$).

Prospect Theory Asset Pricing Model

Return on dividends

$$
\begin{aligned}
1 &= \rho \exp\left(g_D - \gamma g_C + \gamma^2 \sigma_C^2 (1-\omega^2)/2\right) \\
&\quad \times \mathcal{E}_t \left[\frac{1 + f(z_{t+1})}{f(z_t)} \exp[(\sigma_D - \gamma \omega \sigma_C)\epsilon_{t+1}] \right] \\
&+ \quad b_0 \rho \, \mathcal{E}_t \left[\hat{v} \left(\frac{1 + f(z_{t+1})}{f(z_t)} \exp(g_D + \sigma_D \epsilon_{t+1}), z_t \right) \right]
\end{aligned}
$$

$$
r_{dt} = \log \left[\frac{1 + f(z_t)}{f(z_{t-1})} \exp(g_D + \sigma_D \epsilon_t) \right]
$$

$f(\cdot)$ is defined as the solution of the Euler condition above. It is the price dividend ratio; i.e. $P_{dt}/D_t = f(z_t)$, where P_{ct} is the price of the asset that pays the dividend stream. r_{dt} is the logarithmic real return, i.e. $r_{dt} = \log(P_{dt} + D_t) - \log(P_{d,t-1})$, where P_{dt} and D_t are measured in real (inflation adjusted) dollars.

Prospect Theory Asset Pricing Model

Self Referential Equations

$$
\begin{aligned}
z_{t+1} &= \eta \left(z_t \frac{\bar{R}}{R_{t+1}} \right) + (1-\eta) \\
R_{t+1} &= \frac{1 + f(z_{t+1})}{f(z_t)} \exp(g_D + \sigma_D \epsilon_{t+1}) \\
1 &= \text{Median}\{z_t\}
\end{aligned}
$$

Prospect Theory Asset Pricing Model

Solution Method

Approximate f by a piecewise linear function $f^{(0)}(z)$.

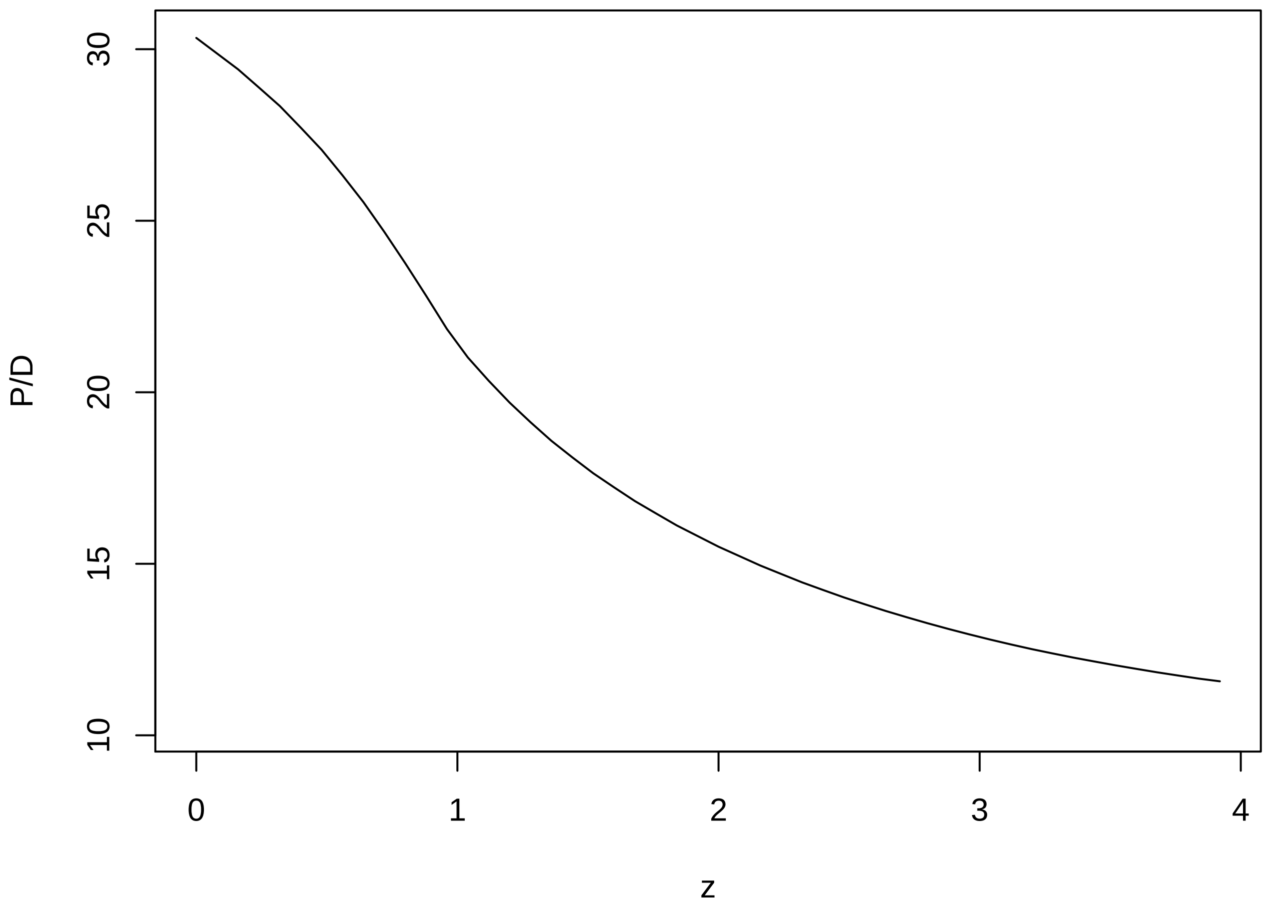
Approximate $\bar{R}$ by $(1+f(1))\exp(g_D)/f(1)$, which is a departure from Barberis, Huang, and Santos (2001).

Define $h^{(0)}$ such that $z_{t+1} = h^{(0)}(z_t, \epsilon_{t+1})$ solves the self referential equations that define z_{t+1} and R_{t+1} on previous slide. A root finding problem. We use Brent's method.

Substitute $h^{(0)}(z_t, \epsilon_{t+1})$ into the Euler equation. Use Gauss-Hermite quadrature to integrate out ϵ_{t+1}. Solve for $f^{(1)}(z)$. A root finding problem at each join point of $f^{(1)}$.

Repeat $h^{(i)} \rightarrow f^{(i+1)}$ until convergence.

Fig 6. Piecewise Linear Approximation



Prospect Theory Asset Pricing Model

Risk Free Rate

$$r_f = \log\left[\rho^{-1}\exp\Big(\gamma g_C - \gamma^2\sigma_C^2/2\Big)\right]$$

r_{ft} is the logarithmic return on an asset that pays one real dollar one year hence with certainty.

Prospect Theory Asset Pricing Model

Large Model Output

Given model parameters

$$\theta = (g_C, g_D, \sigma_C, \sigma_D, \omega, \gamma, \rho, \lambda, k, b_0, \eta)$$

simulate annually and set

$$c_t^a = \log(C_t)$$

$$r_{dt}^a = r_{dt}$$

$$r_{ft}^a = r_f$$

Prospect Asset Pricing Model

Prior Distribution

$$p(\theta) = \mathsf{N}\left[r_f \,|\, 0.89, \left(\frac{1}{1.96}\right)^2\right] \prod_{i=1}^{p} \mathsf{N}\left[\theta_i \,|\, \theta_i^*, \left(\frac{0.1\theta_i^*}{1.96}\right)^2\right]$$

where the θ_i^* are the calibrated values from Barberis, Huang, Santos (2001).

This is not an independence prior (seen next slide).

The first term dominates (slide after next).

Table 5. Correlation Matrix of the Prospect Theory Model Prior

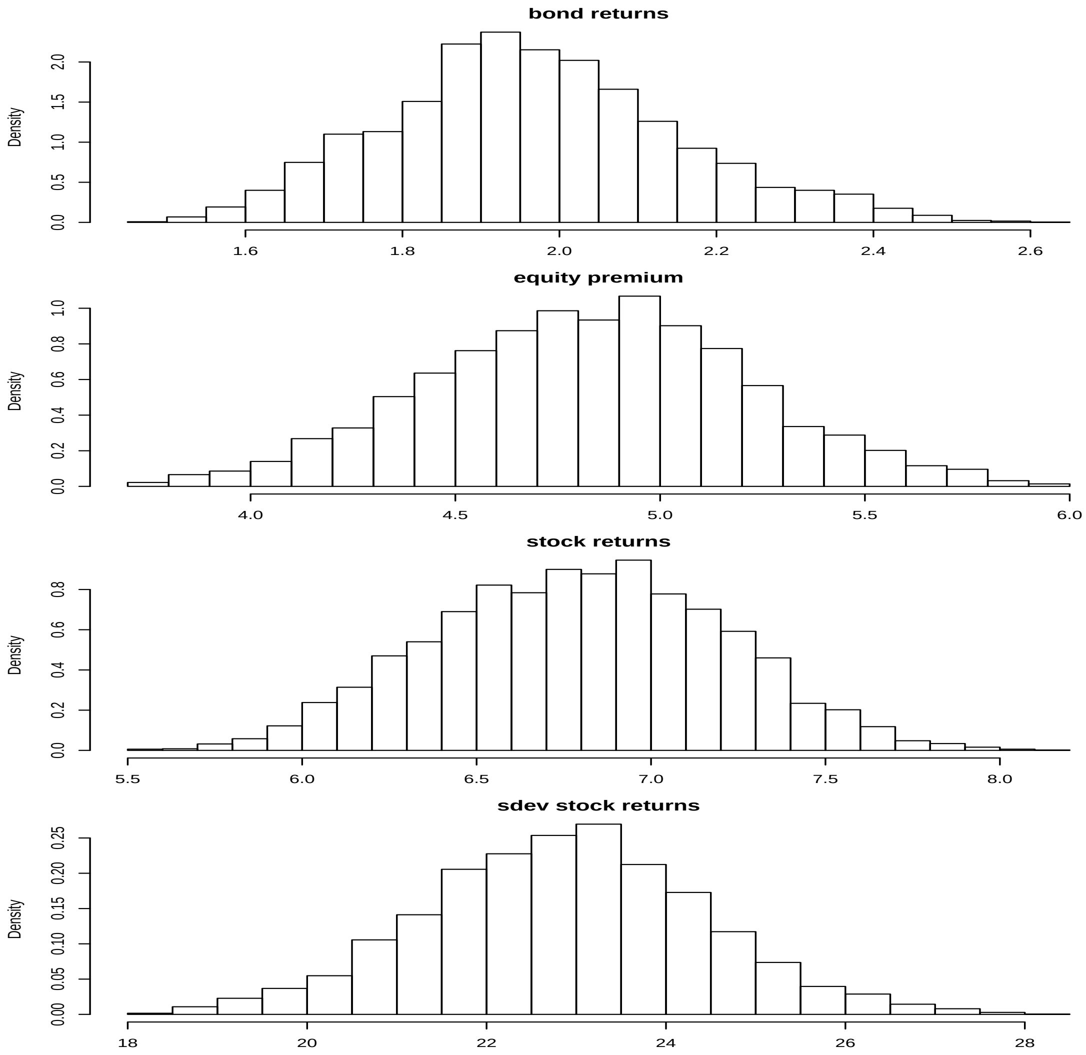
	g_C	g_D	σ_C	σ_D	ω	γ	ρ	λ	k	b_0	η
g_C	1.00	-0.19	-0.14	-0.03	-0.09	0.09	-0.05	0.11	-0.04	-0.08	0.31
g_D	-0.19	1.00	-0.03	0.06	0.03	0.15	0.11	-0.05	0.12	-0.36	-0.28
σ_C	-0.14	-0.03	1.00	0.14	-0.06	-0.21	-0.24	0.18	-0.20	-0.06	-0.15
σ_D	-0.03	0.06	0.14	1.00	0.12	-0.23	-0.12	0.13	0.10	-0.01	-0.06
ω	-0.09	0.03	-0.06	0.12	1.00	-0.17	0.16	-0.01	-0.06	0.04	-0.25
γ	0.09	0.15	-0.21	-0.23	-0.17	1.00	0.06	0.03	-0.11	-0.07	0.05
ρ	-0.05	0.11	-0.24	-0.12	0.16	0.06	1.00	-0.10	0.15	0.26	0.07
λ	0.11	-0.05	0.18	0.13	-0.01	0.03	-0.10	1.00	-0.07	-0.08	-0.02
k	-0.04	0.12	-0.20	0.10	-0.06	-0.11	0.15	-0.07	1.00	-0.25	0.07
b_0	-0.08	-0.36	-0.06	-0.01	0.04	-0.07	0.26	-0.08	-0.25	1.00	0.24
η	0.31	-0.28	-0.15	-0.06	-0.25	0.05	0.07	-0.02	0.07	0.24	1.00

Table 6. Prospect Model

	Prior		Posterior	
Parameter	Mode	Std.Dev.	Mode	Std.Dev.
g_C	0.01818848	0.00090476	0.01892090	0.00092879
g_D	0.01818848	0.00090163	0.01818848	0.00089555
σ_C	0.03759766	0.00205060	0.03271484	0.00199789
σ_D	0.12011719	0.00597515	0.10839844	0.00537066
ω	0.14941406	0.00812375	0.14550781	0.00817098
γ	1.00781250	0.05078639	0.99218750	0.04934629
ρ	0.99981689	0.00107731	0.99969482	0.00146920
λ	2.28125000	0.11296097	2.15625000	0.11651081
k	9.78125000	0.52518433	9.71875000	0.46313708
b_0	2.10205078	0.02325629	2.05712891	0.02103218
η	0.89636230	0.01482200	0.86242676	0.01774379
r_f	1.77957600	0.16310681	1.85515200	0.18479550
$r_d - r_f$	5.53539600	0.37949302	4.25407200	0.38999789
σ_{r_d}	25.92442092	2.00779757	21.10334571	1.56307799

Parameter values are for the annual frequency. Mode is the mode of the multivariate density. It actually occurs in the MCMC chain whereas other measures of central tendency may not even satisfy support conditions. In the data, $r_d - r_f = 5.59 - 0.89 = 5.5$ and $\sigma_{r_d} = 19.72$. The auxiliary model is GARCH with normal errors. The data are annual stock returns and consumption growth 1930–2008.

Fig 7. Posterior Returns



Outline

- Overview
- Models considered
- Bayesian inference for general scientific models
  - ▷ Genesis in II/EMM
  - ▷ Relative model comparison
  - ▷ Absolute model assessment
  - ▷ Graphical interpretation of methodology
- Results
- Sensitivity analysis

Bayesian Inference for General Scientific Models

- Gallant and McCulloch (2009)

- The ideas for model estimation are not new.
  - ▷ What is new is a computational strategy that works.
  - ▷ Extremely computationally intensive.

- The ideas for absolute model assessment are probably new.
  - ▷ "No attribution is correct." Steve Stigler.

Its Genesis is in II/EMM Notions

- Structural model: $p(y|x,\theta)$

- Auxiliary model: $f(y|x,\eta)$

- Binding function:

  $g(\theta) \mapsto \underset{\eta}{\text{argmin}} \int\int \log\, p(y|x,\theta) - \log\, f(y|x,\eta)\, dP(y,x|\theta)$

  ▷ Use KL because it can be computed without knowledge of $p(y|x,\theta)$ provided simulation from $p(y|x,\theta)$ is possible.

  ▷ $g(\theta) \mapsto \underset{\eta}{\text{argmax}} \sum_{t=1}^{N} \log\, f(\hat{y}_t|\hat{x}_t,\eta)$

- Likelihood: $p(y|x,\theta) = f(y|x,g(\theta))$

Computing the Binding Function

$$g(\theta) \mapsto \underset{\eta}{\text{argmax}} \sum_{t=1}^{N} \log\ f(\hat{y}_t | \hat{x}_t, \eta)$$

1. For each θ of an MCMC chain of length R, generate a simulation $\{\hat{y}_t, \hat{x}_t\}_{t=1}^{N}$ from $p(y|x,\theta)$, $N = 5000$.

2. The start value of η is the mode of an MCMC chain $\{\eta_t\}_{t=1}^{K}$ with likelihood $\sum_{t=1}^{N} \log\ f(\hat{y}_t|\hat{x}_t, \eta)$ and a flat prior, $K = 200$.

   - For use later compute $S_\eta \leftarrow S_\eta + (\eta_{K/2} - \eta_K)(\eta_{K/2} - \eta_K)'$
   - $\Sigma_\eta = \frac{1}{R} S_\eta$

3. Compute $\underset{\eta}{\text{argmax}} \sum_{t=1}^{N} \log\ f(\hat{y}_t|\hat{x}_t, \eta)$ using BFGS.

Computing the Posterior

1. For data $\{y_t, x_t\}_{t=1}^{n}$ use MCMC with prior $\pi(\theta)$ and likelihood $\sum_{t=1}^{n} \log f(y_t|x_t, g(\theta))$.

   - $g(\theta) \mapsto \underset{\eta}{\text{argmax}} \sum_{t=1}^{N} \log\ f(y_t|x_t, \eta)$

   - The prior can depend on functionals of $p(y|x,\theta)$ that can be computed from the simulation $\{\widehat{y}_t, \widehat{x}_t\}_{t=1}^{N}$, e.g. risk free rate.

2. Posterior probabilities can be computed from the MCMC chain for θ.

   - For model comparison use method f_5 of Gamerman and Lopes (2006),

Relative Model Comparison

- Compute posterior probabilities for structural models $p_1(y|x,\theta_1)$, $p_2(y|x,\theta_2)$, $p_3(y|x,\theta_3)$ with priors $\pi(\theta_1)$, $\pi(\theta_2)$, $\pi(\theta_3)$ from their chains using GL f_5.
  - ▷ Use the same auxiliary model $f(y|x,\eta)$ for each model.

- Equivalent to comparing the models $f(y|x,g_1(\theta_1))$, $f(y|x,g_2(\theta_2))$, $f(y|x,g_3(\theta_3))$, with priors $\pi(\theta_1)$, $\pi(\theta_2)$, $\pi(\theta_3)$.
  - ▷ This is an important observation.
  - ▷ Inference is actually being conducted with likelihoods $\prod f(y|x,g_1(\theta_1))$, $\prod f(y|x,g_2(\theta_2))$, $\prod f(y|x,g_3(\theta_3))$, not $\prod p_1(y|x,\theta_1)$, $\prod p_2(y|x,\theta_2)$, $\prod p_3(y|x,\theta_3)$.
  - ▷ If $f(y|x,\eta)$ encompasses $p_1(y|x,\theta_1)$, $p_2(y|x,\theta_2)$, $p_3(y|x,\theta_3)$, then the former and later are the same.

Absolute Model Assessment

- Likelihood: auxiliary model $f(y|x,\eta)$.

- Prior: $\pi_\kappa(\eta) \propto \exp\left(-\frac{1}{2}\min_\theta\,[\eta - g(\theta)]'(\kappa\Sigma_\eta)^{-1}[\eta - g(\theta)]\right)$

- Assign equal prior probability to a sequence of models that differ in their κ priors.

- Compute the posterior probabilities of the sequence using GL f_5.

- High posterior probability for large κ is evidence against the model.

Tinker Toy Example

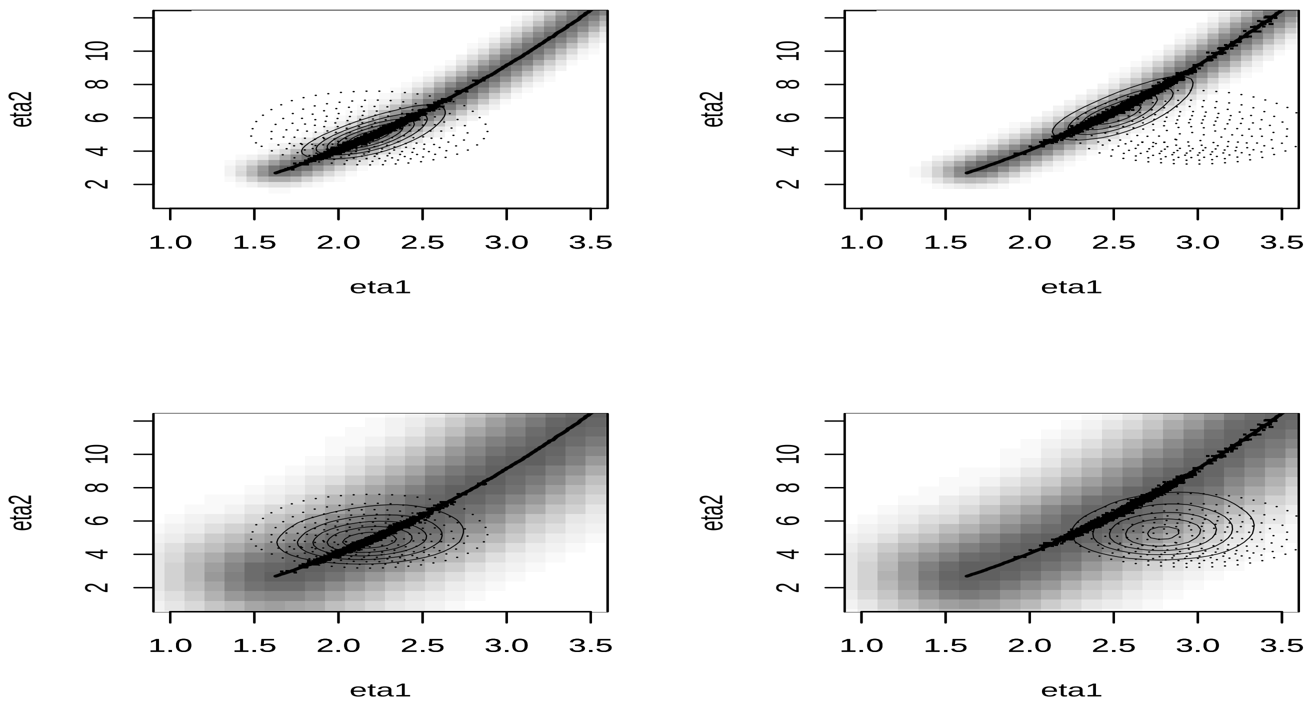
Structural Model: $p(y \,|\, \theta) = n(y \,|\, \theta, \theta^2)$

Auxiliary Model: $f(y \,|\, \eta) = n(y \,|\, \eta_1, \eta_2)$

Binding function: $g : \theta \mapsto \eta = (\eta_1, \eta_2) = (\theta, \theta^2)$

The structural model states that excess returns $y_t = r_{dt} - r_{ft}$ are iid normal with a Sharpe ratio of one. The auxiliary model states that excess returns are iid normal.

Fig 8. The Structural and Auxiliary Model



The dotted lines are contours of the likelihood of the auxiliary model $f(y|x, \eta)$ of the tinker toy example. The line is the prior on η determined by the binding function $\eta = g(\theta)$ from the parameters θ of structural model $p(y|x, \theta)$ to the parameters η of the auxiliary model. In the left panels the structural model is true, in the right it is false. The thickness of the line is proportional to the posterior of η. The prior $\pi(\eta)$ can be relaxed as indicated by the shading. The lower panels are more relaxed than the upper. The solid contours show the posterior under the relaxed prior. Relaxation causes the contours to enlarge in all cases. When the structural model is false, the posterior shifts in search of the likelihood.

Outline

- Overview
- Models considered
- Bayesian inference for general scientific models
- Results
  - ▷ Relative comparison
  - ▷ Absolute assessment
- Sensitivity analysis

Model Assessment
Relative Performance in Annual Data
Stock Returns

Posterior Probabilities for Three Models

	1930–2008	1950–2008
Habit Persistence	0.67	0.30
Long Run Risk	0.30	0.54
Prospect Theory	0.03	0.17

Data are annual consumption growth and stock returns. The auxiliary model is GARCH with normal errors.

Model Assessment
Relative Performance in Annual Data
Consumption Growth and Stock Returns

Posterior Probabilities for Three Models

	1930–2008	1950–2008
Habit Persistence	0.00	1.00
Long Run Risk	1.00	0.00
Prospect Theory	0.00	0.00

Data are annual consumption growth and stock returns. The auxiliary model is GARCH with normal errors.

Model Assessment
Absolute Performance in Annual Data
Consumption Growth and Stock Returns

Posterior Probabilities for Long Run Risk

$\kappa = 0.1$	0.03
$\kappa = 1.0$	0.27
$\kappa = 5.0$	0.70

κ is the standard deviation of a prior that imposes the long run risk model upon a model that is regarded as correct.

Data are annual consumption growth and stock returns, 1930–2008 The auxiliary model is GARCH with normal errors.

Outline

- Overview
- Models considered
- Bayesian inference for general scientific models
- Results
- Sensitivity analysis
  - ▷ Role of the auxiliary model
  - ▷ Do results depend on the choice of auxiliary model?

The Auxiliary Model

- Common sense suggests that the auxiliary model $f_1(y|x,\eta)$ that best fits the data should be used, particularly for absolute model assessment.

- Theory dictates that for correct Bayesian inference an auxiliary model $f_5(y|x,\eta)$ that encompasses the structural models under consideration be used.

- How to choose? Particularly in our case because the encompassing model is absurd.

Common Sense Auxiliary Model f_1

- Mean function:
  - ▷ One lag
  - ▷ Linear

  Variance function:
  - ▷ GARCH(1,1)

  Errors:
  - ▷ Normal

Encompassing Auxiliary Model f_5

- Mean function:
  - ▷ Two lags
  - ▷ Nonlinear

  Variance function:
  - ▷ GARCH(1,1)
  - ▷ Leverage

  Errors:
  - ▷ Flexible SNP density

Points of View

- Using f_1 instead if f_5 means that a likelihood that differs from the structural model's likelihood is being used.
  - ▷ Inference cannot be regarded as relating to the structural model.

- Using f_1 instead of f_5 is akin to GMM estimation.
  - ▷ One only asks the structural model to match certain features of the data and allows it to ignore others.

Logically Correct Approach

- Use the encompassing model $f_5(y|x,\eta)$ together with a prior $\pi(\eta)$ that forces equality, i.e., $f_5(y|x,\eta)\pi(\eta) = f_1(y|x,\eta)$.

- Does not work, even for relaxed priors that do not force equality.

  ▷ There do not exist parameter settings for these strucural models and solution methods that will stop them from emitting bizarre simulations.

Sensitivity

- Does the choice of auxiliary model affect results?

- Does the choice of sample period, 1930–2008 or 1950–2008, interact with the choice of auxiliary model?

Table 7. Auxiliary Models

	f_0	f_1	f_2	f_3	f_4	f_5
Mean	1 lag	1 lag	1 lag	1 lag	1 lag	2 lags
	linear	linear	linear	linear	nonlinear	nonlinear
Variance	constant	garch	garch	garch	garch	garch
			leverage	leverage	leverage	leverage
Errors	normal	normal	normal	flexible	flexible	flexible
Parms	3	5	6	10	11	12

$f_0(y_t \,|\, x_{t-1},\, \eta)$ is BIC preferred when fit to the data. $f_5(y_t \,|\, x_{t-1},\, \eta)$ is BIC preferred when fit to model simulations. Variance matrices are of the BEKK form. Data are centered and scaled; lags are attenuated by a spline transform.

Table 8. Posterior Probability, Stock Returns, 1930–2008

Model	f_0	f_1	f_2	f_3	f_4	f_5
Habit	0.29	0.67	0.33	0.25	0.36	0.34
LR Risk	0.62	0.30	0.43	0.25	0.21	0.22
Prospect	0.09	0.03	0.24	0.50	0.43	0.44

The data are annual stock returns 1930–2008. $f_0(y_t \,|\, x_{t-1}, \eta)$ is BIC preferred when fit to the data. $f_5(y_t \,|\, x_{t-1}, \eta)$ is BIC preferred when fit to model simulations. Variance matrices are of the BEKK form. Data are centered and scaled; lags are attenuated by a spline transform.

Table 9. Posterior Probability, Stock Returns, 1950–2008

Model	f_0	f_1	f_2	f_3	f_4	f_5
Habit	0.33	0.30	0.26	0.24	0.30	0.30
LR Risk	0.62	0.54	0.68	0.61	0.58	0.49
Prospect	0.05	0.17	0.06	0.14	0.11	0.20

The data are annual stock returns 1950–2008. $f_0(y_t \,|\, x_{t-1}, \eta)$ is BIC preferred when fit to the data. $f_5(y_t \,|\, x_{t-1}, \eta)$ is BIC preferred when fit to model simulations. Variance matrices are of the BEKK form. Data are centered and scaled; lags are attenuated by a spline transform.

Table 10. Posterior Probability, Consumption Growth and Stock Returns, 1930–2008

Model	f_0	f_1	f_2	f_3	f_4	f_5
Habit	0.00	0.00	0.00	0.08	0.00	0.00
LR Risk	1.00	1.00	1.00	0.92	1.00	1.00
Prospect	0.00	0.00	0.00	0.00	0.00	0.00

The data are annual stock returns and consumption growth 1930–2008. $f_1(y_t \,|\, x_{t-1}, \eta)$ is BIC preferred when fit to the data. $f_5(y_t \,|\, x_{t-1}, \eta)$ is BIC preferred when fit to model simulations. Variance matrices are of the BEKK form. Data are centered and scaled; lags are attenuated by a spline transform.

Table 11. Posterior Probability, Consumption Growth and Stock Returns, 1950–2008

Model	f_0	f_1	f_2	f_3	f_4	f_5
Habit	1.00	1.00	1.00	1.00	1.00	1.00
LR Risk	0.00	0.00	0.00	0.00	0.00	0.00
Prospect	0.00	0.00	0.00	0.00	0.00	0.00

The data are annual stock returns and consumption growth 1950–2008. $f_0(y_t \,|\, x_{t-1},\, \eta)$ is BIC preferred when fit to the data. $f_5(y_t \,|\, x_{t-1},\, \eta)$ is BIC preferred when fit to model simulations. Variance matrices are of the BEKK form. Data are centered and scaled; lags are attenuated by a spline transform.