Complementary Bayesian Method of Moments Strategies

A. Ronald Gallant
Penn State University
Department of Economics
University Park PA 16802 USA

Paper: http://www.aronaldg.org/papers/cb.pdf
Slides: http://www.aronaldg.org/papers/cbclr.pdf

Bayesian Method of Moments Strategies

- Moment Constrained Bayes
  - Estimates both likelihood parameters $\theta_{(1)}$ and moment function parameters $\theta_{(2)}$
  - Computationally challenging because parameter space is singular
  - Can get approximate MCMC chain for $\theta = (\theta_{(1)}, \theta_{(2)})$ using a penalty function approach: λ-prior method.

- Moment Induced Bayes
  - Estimates moment function parameters $\theta_{(2)}$ parameters only.
  - Must assume a distribution Ψ for a semi-pivotal derived from moment conditions.

- Plan is to use the λ-prior approximation to moment constrained Bayes to infer Ψ for moment induced Bayes for a representative macro-finance application.

Example – Primitives

C_t consumption endowment

P_{ct} price of an asset that pays the consumption endowment

$R_{ct} = (P_{ct} + C_t)/P_{c,t-1}$ gross return on consumption endowment

D_{st} any cash flow S

P_{st} price of cash flow S

$R_{st} = (P_{st} + D_{st})/P_{s,t-1}$ gross return on cash flow S

Example – Epstein-Zin-Weil Utility

Marginal Rate of Substitution

$$M_{t,t+1} = \delta^{\beta} \left(\frac{C_{t+1}}{C_t}\right)^{-(\beta/\psi)} \left(R_{c,t+1}\right)^{(\beta-1)}, \quad \beta = \frac{1-\gamma}{1-1/\psi}$$

δ time preference parameter

γ coefficient of risk aversion

ψ elasticity of intertemporal substitution

Euler equation

$$1 = \mathcal{E}_t \left(\text{MRS}_{t,t+1} R_{c,t+1}\right)$$

$$R_{ct} = \frac{\frac{C_{t-1}}{C_{t-2}} \sum_{j=1}^{\infty} \mathcal{E}_t \prod_{k=1}^{j} \left(\frac{C_{t+k}}{C_{t+k-1}} \text{MRS}_{t+k-1,t+k}\right) + \frac{C_t}{C_{t-1}} \frac{C_{t-1}}{C_{t-2}}}{\sum_{j=1}^{\infty} \mathcal{E}_{t-1} \prod_{k=1}^{j} \left(\frac{C_{t+k-1}}{C_{t+k-2}} \text{MRS}_{t+k-2,t+k-1}\right)}$$

Example – Data

Annual for years 1930 to 2015, construction discussed later

- s_t = log real gross stock return (value weighted NYSE/AMEX/NASDAQ).
- b_t = log real gross bond return (30 day T-bill return).
- c_t = log real per capita consumption growth (nondurables and services).
- w_t = log real gross wealth return, $w_t = \log(R_{ct})$
- $\text{mrs}_{t-1,t}$ = log marginal rate of substitution, $\text{mrs}_{t-1,t} = \log(\text{MRS}_{t-1,t})$

x denotes an array of extent n whose columns are $x_t = (s_t, b_t, c_t, w_t)'$.

Example – Moment Conditions

Given parameters $\theta_{(2)} = (\gamma, \psi, \delta)$ and data x, define the pricing errors

$$\begin{aligned} e_{1,t,t-1} &= 1 - \exp(\mathrm{mrs}_{t-1,t} + s_t) \\ e_{2,t,t-1} &= 1 - \exp(\mathrm{mrs}_{t-1,t} + b_t) \end{aligned}$$

and moment conditions

$$\begin{aligned} m_1(x_t, x_{t-1}, \theta_{(2)}) &= e_{1,t,t-1} \\ m_2(x_t, x_{t-1}, \theta_{(2)}) &= e_{2,t,t-1} \\ m_3(x_t, x_{t-1}, \theta_{(2)}) &= e_{1,t,t-1} \times s_{t-1} \\ m_4(x_t, x_{t-1}, \theta_{(2)}) &= e_{1,t,t-1} \times b_{t-1} \\ m_5(x_t, x_{t-1}, \theta_{(2)}) &= e_{1,t,t-1} \times c_{t-1} \\ m_6(x_t, x_{t-1}, \theta_{(2)}) &= e_{1,t,t-1} \times w_{t-1} \\ m_7(x_t, x_{t-1}, \theta_{(2)}) &= e_{2,t,t-1} \times s_{t-1} \\ m_8(x_t, x_{t-1}, \theta_{(2)}) &= e_{2,t,t-1} \times b_{t-1} \\ m_9(x_t, x_{t-1}, \theta_{(2)}) &= e_{2,t,t-1} \times c_{t-1} \\ m_{10}(x_t, x_{t-1}, \theta_{(2)}) &= e_{2,t,t-1} \times w_{t-1} \end{aligned}$$

Example – Moment Equation

$$\bar{m}(x, \theta_{(2)}) = \frac{1}{n} \sum_{t=2}^{n} m(x_t, x_{t-1}, \theta_{(2)}). \tag{1}$$

$$m(x_t, x_{t-1}, \theta_{(2)}) = \begin{pmatrix} m_1(x_t, x_{t-1}, \theta_{(2)}) \\ m_2(x_t, x_{t-1}, \theta_{(2)}) \\ \vdots \\ m_{10}(x_t, x_{t-1}, \theta_{(2)}) \end{pmatrix} \tag{2}$$

Example – HAC Matrix

$$W\left(x,\theta_{(2)}\right) = \sum_{\tau=-[n^{1/5}]}^{[n^{1/5}]} w\left(\frac{\tau}{[n^{1/5}]}\right)\bar{W}_\tau \tag{3}$$

where

$$w(u) = \begin{cases} 1 - 6|u|^2 + 6|u|^3 & \text{if } 0 < u < \frac{1}{2} \\ 2(1-|u|)^3 & \text{if } \frac{1}{2} \leq u < 1 \end{cases}$$

$$\bar{W}_\tau = \begin{cases} \frac{1}{n}\sum_{t=2+\tau}^{n}\left(m_t - \bar{m}\right)\left(m_{t-\tau} - \bar{m}\right)' & \tau \geq 0 \\ \tilde{W}'_{n,-\tau} & \tau < 0 \end{cases} \tag{4}$$

Example – Semi-Pivotal

$$Z(x, \theta_{(2)}) = \sqrt{n}\left[W(x, \theta_{(2)})\right]^{-\frac{1}{2}}\left[\bar{m}(x, \theta_{(2)})\right] \tag{5}$$

$\left[W(x, \theta_{(2)})\right]^{-\frac{1}{2}}$ is the inverse of the Cholesky factorization of $W(x, \theta_{(2)})$

The semi-pivotal condition is that $\left\{x : Z(x, \theta_{(2)}) = z\right\}$ not be empty for any choice of $(z, \theta_{(2)})$ in parameter space $\Theta_{(2)}$ of $\theta_{(2)}$ and range space $\mathcal{Z}$ of Z.

Moment Constrained Bayes – Defined

- Assert a likelihood $f(x \,|\, \theta_{(1)})$, which is often a sieve.

- Expectation of the moment conditions generates parametric restrictions

$$0 = \rho\left(\theta_{(1)}, \theta_{(2)}\right) = \int m(x_t, x_{t-1}, \theta_{(2)})\, f(x|\theta_{(1)})\, dx \tag{6}$$

- For any non-random positive definite matrix $W\left(\theta_{(1)}, \theta_{(2)}\right)$, an equivalent expression is

$$0 = \rho'\left(\theta_{(1)}, \theta_{(2)}\right) W^{-1}\left(\theta_{(1)}, \theta_{(2)}\right) \rho\left(\theta_{(1)}, \theta_{(2)}\right). \tag{7}$$

- See Bornn, Shephard, and Solgi (2016) and Shin (2015).

Moment Constrained Bayes – Properties

Advantage:

- Provides estimates of likelihood parameters $\theta_{(1)}$ and structural parameters $\theta_{(2)}$

Disadvantage:

- The parameter space
$$\Theta = \left\{\theta \in \mathbb{R}^{\dim(\theta)} \,|\, \theta = \left(\theta_{(1)}, \theta_{(2)}\right), 0 = \rho\left(\theta_{(1)}, \theta_{(2)}\right)\right\} \tag{8}$$
has measure zero.

- Makes estimation of the posterior distribution of θ subject to a prior $p(\theta)$ and constraint (6) by Markov Chain Monte Carlo (MCMC) problematic

Moment Induced Bayes – Defined

- Assumes that the semi-pivotal $z = Z\big(x, \theta_{(2)}\big)$ follows a distribution $\Psi(z)$ with density $\psi(z)$.

- From this distribution, one can infer a probability space upon which Bayesian inference can be conducted (Gallant, 2016a).

- On this space $p(x \mid \theta_{(2)}) = \psi\Big[Z\big(x, \theta_{(2)}\big)\Big]$ is the likelihood.

- One proceeds directly to Bayesian inference using likelihood $p(x \mid \theta_{(2)})$ and a prior $p(\theta_{(2)})$

- Often the normal distribution Φ with density ϕ is used.
  - Sweeting (1986) provides if and only if conditions for convergence of ψ to ϕ uniformly in $\theta_{(2)}$ and x.

Moment Induced Bayes – Properties

Advantage:

- Computationally attractive, e.g., Chernozhukov and Hong (2003)

Disadvantages:

- Only provides estimates of structural parameters $\theta_{(2)}$
- The consequences of using ϕ instead of ψ in finite samples for relevant applications are unknown.

The legitimacy of using ϕ instead of ψ is the subject of this talk.

λ-Prior Method

- Use MCMC with sieve $f\left(x \,|\, \theta_{(1)}\right)$ and prior

$$p_\lambda(\theta) = p(\theta) \times \exp\left[-\lambda\, n\, \rho'\left(\theta_{(1)}, \theta_{(2)}\right) W^{-1}\left(\theta_{(1)}, \theta_{(2)}\right) \rho\left(\theta_{(1)}, \theta_{(2)}\right)\right]. \qquad (9)$$

- Compute expectations in $\rho\left(\theta_{(1)}, \theta_{(2)}\right)$ and $W\left(\theta_{(1)}, \theta_{(2)}\right)$ by simulation from $f\left(x \,|\, \theta_{(1)}\right)$

- For large λ, MCMC θ draws become concentrated near the parameter space Θ thereby providing approximate moment constrained Bayes draws

- For λ sufficiently large, MCMC must fail because Θ is singular.

- Find the largest λ such that the θ-chain mixes and use those draws as the approximation to the posterior on Θ.

- Get draws z from Ψ by simulating x of size n for each MCMC draw θ and evaluating $z = Z(x, \theta_{(2)})$

Data

From BEA and CRSP

- s_t = log real gross stock return (value weighted NYSE/AMEX/NASDAQ).
- b_t = log real gross bond return (30 day T-bill return).
- c_t = log real per capita consumption growth (nondurables and services).

More trouble

- $\text{mrs}_{t-1,t}$ = log marginal rate of substitution, $\text{mrs}_{t-1,t} = \log(\text{MRS}_{t-1,t})$
- w_t = log real gross wealth return, $w_t = \log(R_{ct})$

Ex Post MRS Computation

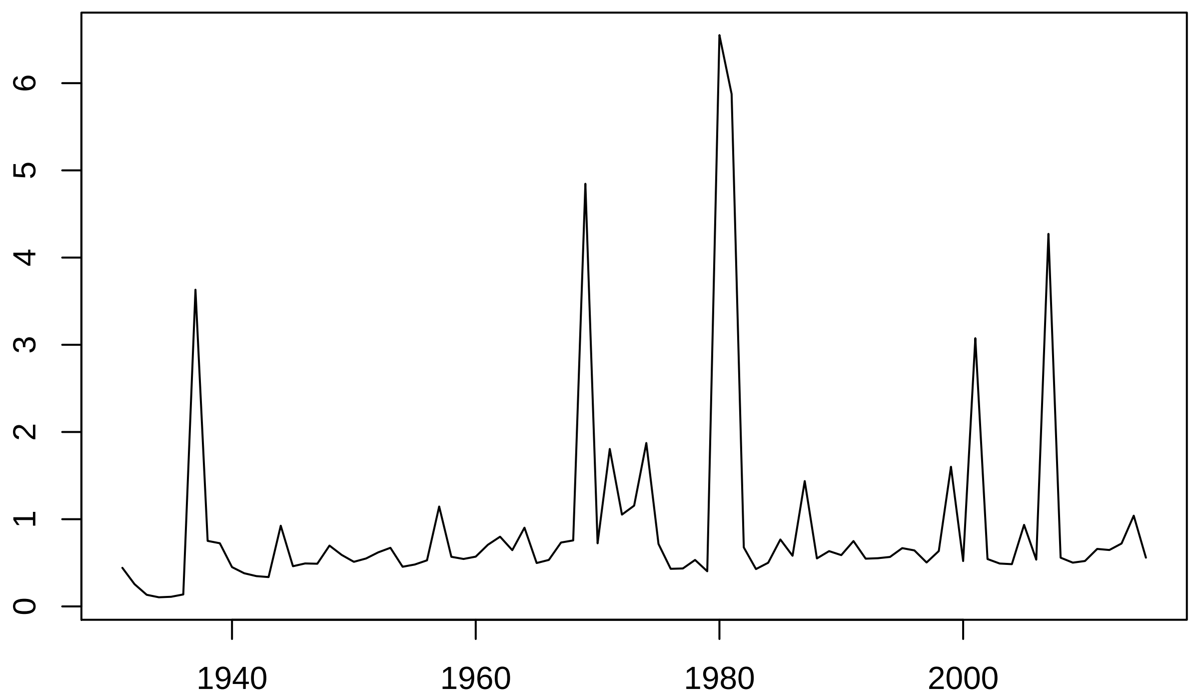
- View MRS $= (\text{MRS}_{0,1}, \ldots, \text{MRS}_{n-1,n})$ as a parameter to be estimated from data:
  - real returns including dividends on 25 Fama-French (1993) portfolios.
  - real returns on 30 day T-bills
  - real returns including dividends on the aggregate stock market
  - real, per-capita, consumption expenditure and labor income growth

- Form moment equations of Euler equation pricing errors of all 26 securities and interact with a constant and all lags – 754 moments.

- The prior for MRS has the form $f_{SNP}(\text{MRS} \,|\, \eta)\, \phi(\eta \,|\, \mu, \sigma)$, where μ is the is the location of the SNP parameter η and σ is its scale. The location parameter μ is determined by fitting $f_{SNP}(\theta \,|\, \mu)$ to a long simulation of $\{\text{MRS}_{t-1,t}\}$ in a Bansal and Yaron (2004) economy; σ chosen so prior is loose.

- Use moment induced Bayes to estimate MRS with $\psi = \phi$. Follows Gallant and Hong (2007).

Figure 1. The Posterior Mean of the MRS



Return to Wealth

- Multiply ex post $\text{MRS}_{t-1,t}$ by observed $\frac{C_t}{C_{t-1}}$

- Fit an SNP conditional density function to $\frac{C_t}{C_{t-1}}\text{MRS}_{t-1,t}$.

- Use simulation from the SNP conditional density function to compute the expectations in

$$R_{ct} = \frac{\frac{C_{t-1}}{C_{t-2}} \sum_{j=1}^{\infty} \mathcal{E}_t \prod_{k=1}^{j} \left(\frac{C_{t+k}}{C_{t+k-1}} \text{MRS}_{t+k-1,t+k} \right) + \frac{C_t}{C_{t-1}} \frac{C_{t-1}}{C_{t-2}}}{\sum_{j=1}^{\infty} \mathcal{E}_{t-1} \prod_{k=1}^{j} \left(\frac{C_{t+k-1}}{C_{t+k-2}} \text{MRS}_{t+k-2,t+k-1} \right)}$$

Table 1. Simple Statistics for the Data

Series	Mean	Standard Deviation	Skewness	Excess Kurtosis
s_t	0.08609	0.19501	-1.16717	2.14135
b_t	0.03211	0.02880	0.85028	0.25196
c_t	0.01996	0.02198	-1.40754	5.06603
w_t	0.01617	0.29628	0.40303	1.17462

The data are annual for the years 1930 through 2015; the sample size is $n = 86$. s_t is log real gross stock return. b_t is log real gross bond return. c_t is log real per capita consumption growth. w_t is log real gross wealth return.

Distribution of Z – Likelihood and Technical Prior

- BIC selected

$$f_{SNP}\big(x_t \,|\, x_{t-1}, \theta_{(1)}\big) = n(x_t \,|\mu_{t-1}, \Sigma_{t-1}) \tag{10}$$

where

$$\begin{aligned} \mu_{t-1} &= b_0 + Bx_{t-1} \\ \Sigma_{t-1} &= \Sigma_0 + q^2\Sigma_{t-2} \\ &\quad + [\mathrm{diag}(p_1, p_2, p_3, p_4)](x_{t-2} - \mu_{t-2})(x_{t-2} - \mu_{t-2})'[\mathrm{diag}(p_1, p_2, p_3, p_4)] \end{aligned}$$

- Prior

$$p(\theta) = p\big(\theta_{(1)}, \theta_{(2)}\big) = p\big(\theta_{(1)}\big)\, p\big(\theta_{(2)}\big)$$

- Technical prior $p\big(\theta_{(1)}\big)$

An indicator that the largest eigen values of the companion matrices of the mean μ_{t-1} and variance function Σ_{t-1} are less than one times $p(q) = n\left[\,q \,|\, 0.88, (0.01/1.96)^2\,\right]$.

Distribution of Z – First Substantive Prior

- Prior
$$p(\theta) = p\big(\theta_{(1)}, \theta_{(2)}\big) = p\big(\theta_{(1)}\big)\, p\big(\theta_{(2)}\big)$$

- First substantive prior $p\big(\theta_{(2)}\big) = p(\gamma)p(\psi)p(\delta)$

$$\begin{aligned}
p(\gamma) &= I_{\{\gamma>0|\}}(\gamma)\, n[\gamma\,|\,10.0, (5.0/1.96)^2] \\
p(\psi) &= I_{\{1.1<\psi<2.0\}}(\psi)\left[1+\cos\left(\pi + 2\pi\frac{\psi-1.1}{2.0-1.1}\right)\right] \\
p(\delta) &= I_{\{0.9615385<\delta<0.9900990\}}(\delta)\left[1+\cos\left(\pi + 2\pi\frac{\delta-0.9615385}{0.9900990-0.9615385}\right)\right]
\end{aligned}$$

0.9615385 corresponds to 4 per cent per annum and 0.9900990 to 1 per cent.

Distribution of Z – Likelihood and λ-Prior

- Likelihood $f\big(x \,|\, \theta_{(1)}\big)$ has transition density $f_{SNP}\big(x_t \,|\, x_{t-1}, \theta_{(1)}\big)$

- λ-prior

$$p_\lambda(\theta) = p(\theta) \times \exp\left[-\lambda\, n\, \rho'\big(\theta_{(1)}, \theta_{(2)}\big)\, W^{-1}\big(\theta_{(1)}, \theta_{(2)}\big)\, \rho\big(\theta_{(1)}, \theta_{(2)}\big)\right].$$

- Moment constraint

$$\rho\big(\theta_{(1)}, \theta_{(2)}\big) = \int m(x_t, x_{t-1}, \theta_{(2)})\, f_{SNP}(x|\theta_{(1)})\, dx$$

- Largest λ for which θ-chain will mix is $\lambda = \exp(3)$, length $R = 200000$

- For every tenth θ in the θ-chain a sample x from $f\big(x \,|\, \theta_{(1)}\big)$ of size $n = 86$ was generated and $z = Z\big(x, \theta_{(2)}\big)$ evaluated to form a z-chain of length 20000.

Table 2. Moment Constrained Estimates, First Prior

	Estimates			Z Prior		
Parameter	Mean	Mode	Std. Dev.	Mean	Mode	Std. Dev.
γ	12.002	9.8193	2.2623	10.018	9.9878	2.5513
ψ	0.7148	0.6893	0.0757	1.5487	1.5541	0.1632
δ	0.9760	0.9767	0.0051	0.9759	0.9759	0.0051

The estimation method is moment constrained Bayes using the lambda prior computational method with $\lambda = \exp(3)$ as described in Section 1. Data are real, annual, per capital consumption for the years 1930–2015 and real, annual stock, bond, and wealth returns for the same years from BEA (2016) and CRSP (2016) that are used to form the moment functions (6) through (6). In the columns labeled Z Normal the distribution of Z given by (10) is presumed to be Normal. In the columns labeled Z Empirical the distribution of Z given by (10) is presumed to be the distribution derived in Section 3; see also Figure 6. The prior is an independence prior with γ normal with mean 10 and standard deviation 2.55 constrained to have positive support, ψ with a cosine density with support (1.1,2.0); and δ a cosine density with support 1% to 4% per annum; see (23). The columns labeled mean, mode, and standard deviation are the mean and standard deviations of an MCMC chain (Gamerman and Lopes (2006) of length 200,000 collected past the point where transients have dissipated. The proposal is move-one-at-a-time random walk.

Figure 2. The Posterior Distribution of Z

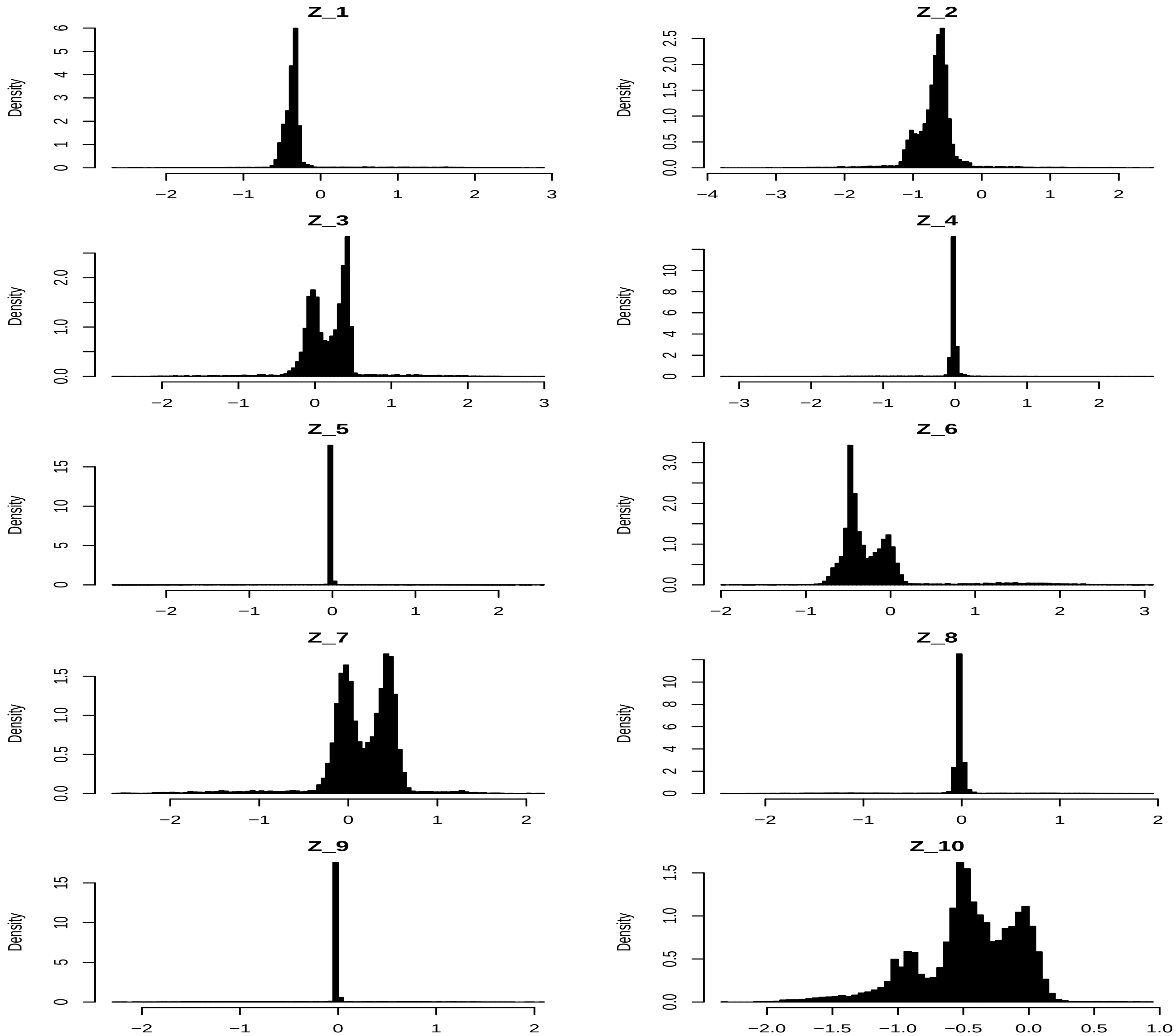
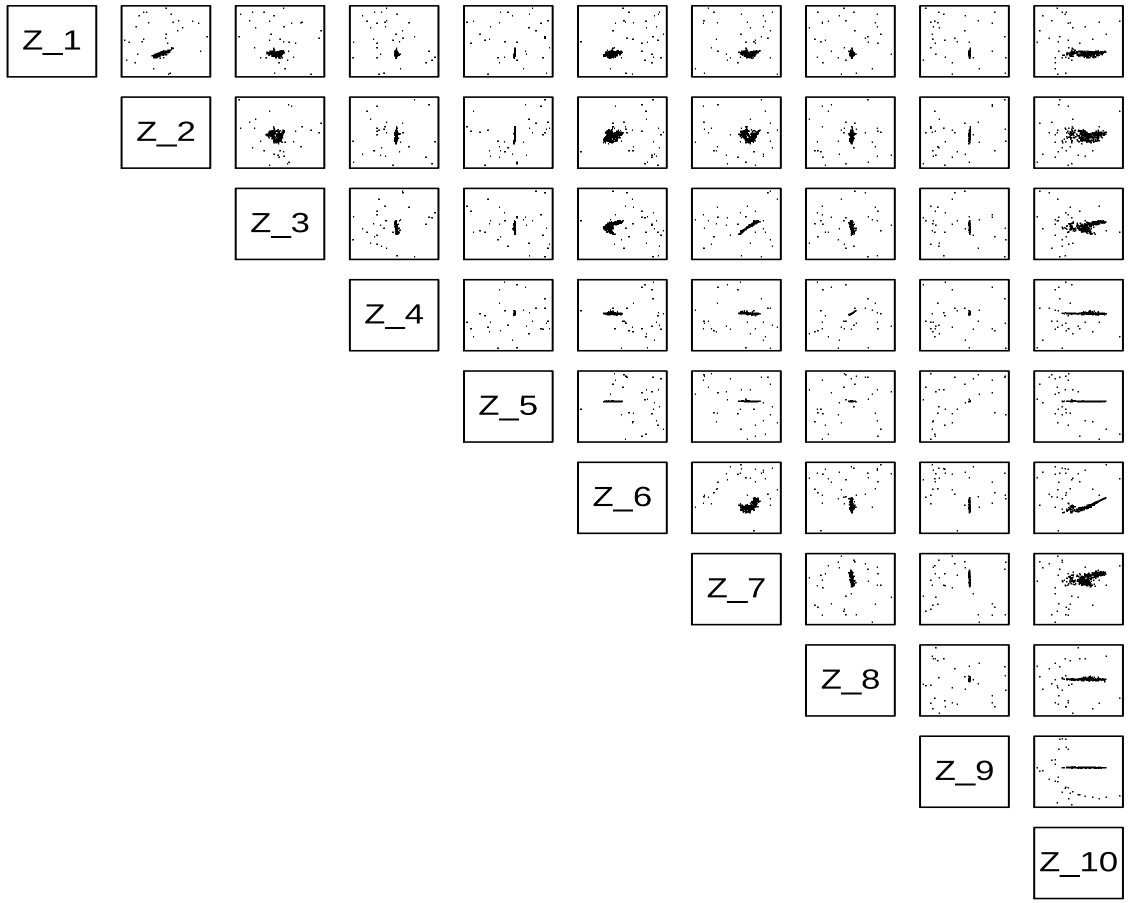


Figure 3. Scatter Plots of Z



Distribution of Z – Conclusions

- The first substantive prior, although reasonable, is too influential. We shall consider a looser prior a few slides hence.

- Scatter plots suggest that there is no need for a copula, can model marginals of PZ where $\text{Var}^{-1}(z) = P'P$

- Using $\psi(z)$ instead of $\phi(z)$ will have the effect of trimming the tails of Z thereby robustifying moment induced Bayes.

- The long, thin tails are due to the instability of the Epstein-Zin-Weil MRS rather than instability of the weighting matrix.

- The tails are exponential, as seen next two slides.

- Q-Q plots of degree ten SNP determination of $\psi(z)$ third slide hence.

Figure 4. Left Q-Q Plots, Z vs. 2 d.f. t

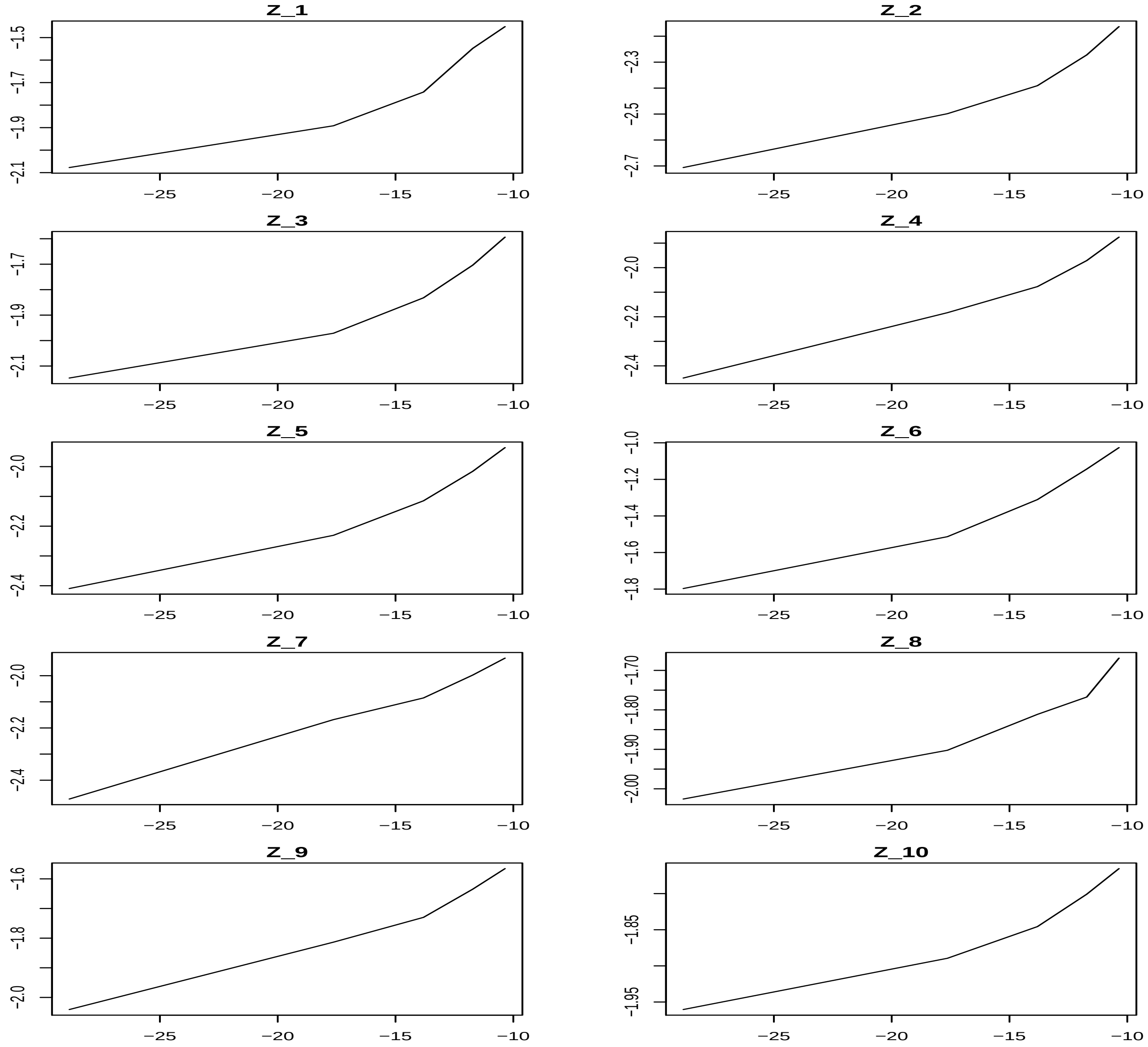


Figure 5. Left Q-Q Plots, Z vs. 30 d.f. t

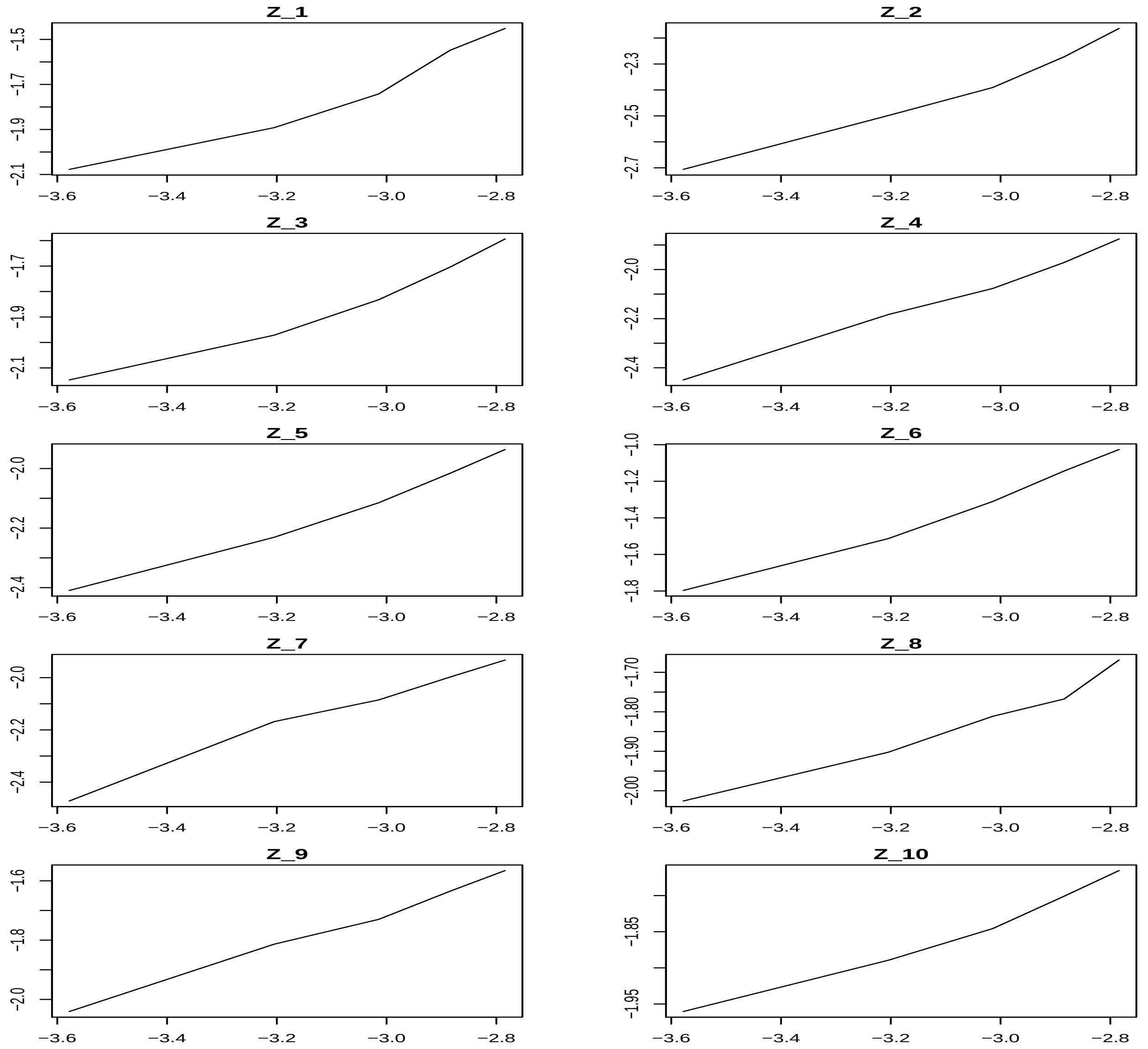
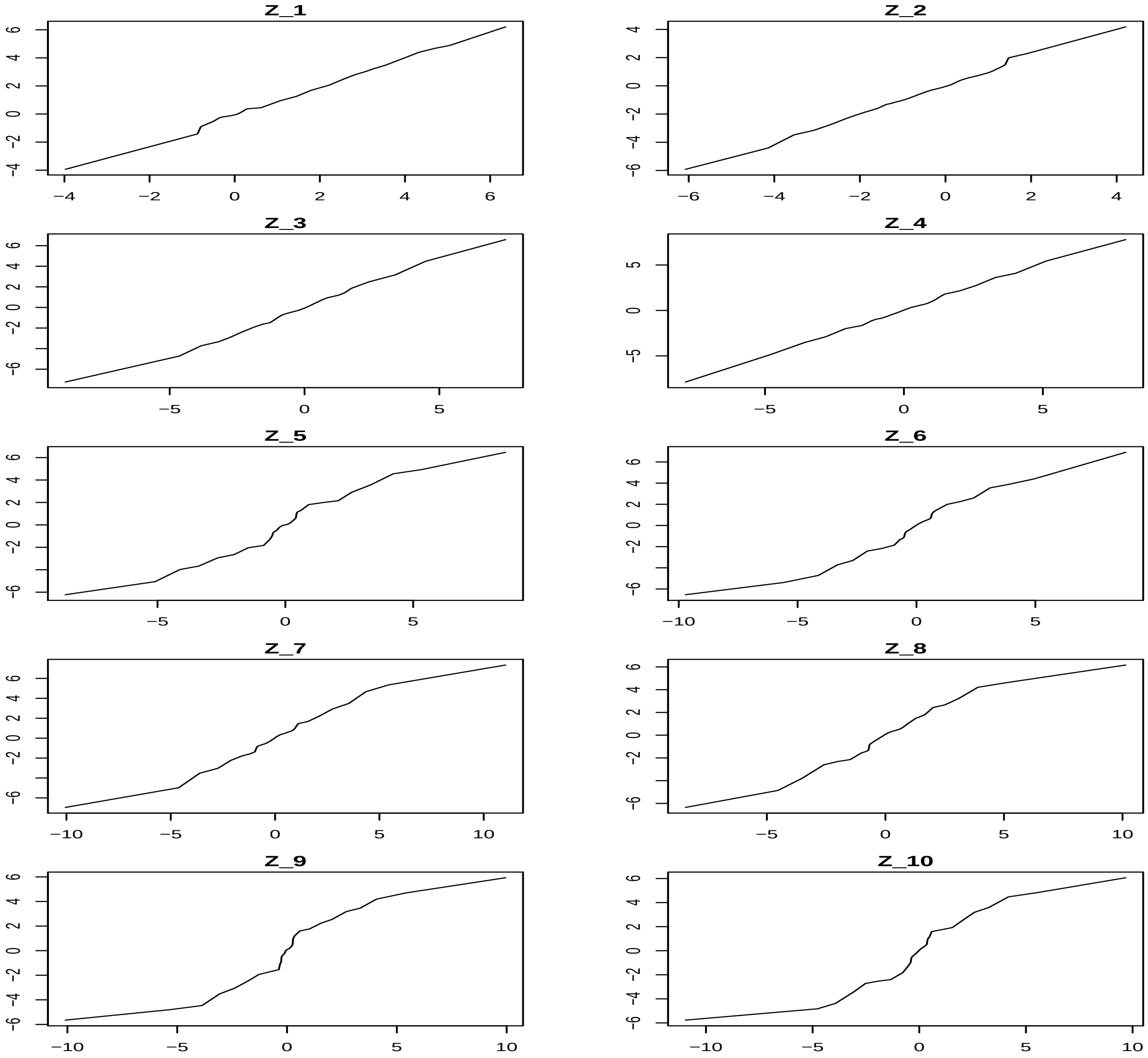


Figure 6. Q-Q Plots, PZ vs. SNP Density



Distribution of Z – Second Substantive Prior

- Second substantive prior $p^*\left(\theta_{(2)}\right) = p^*(\gamma)p^*(\psi)p^*(\delta)$

$$
\begin{aligned}
p^*(\gamma) &= I_{\{\gamma>0|\}}(\gamma)\, n[\gamma \,|\, 10.0, (500.0/1.96)^2] \\
p^*(\psi) &= I_{\{1.1<\psi<1.5\}}(\psi) \left[1 + \cos\left(\pi + 2\pi \frac{\psi - 1.1}{1.5 - 1.1}\right)\right] \\
p^*(\delta) &= I_{\{0.9615385<\delta<0.9900990\}}(\delta) \left[1 + \cos\left(\pi + 2\pi \frac{\delta - 0.9615385}{0.9900990 - 0.9615385}\right)\right]
\end{aligned}
$$

- The resulting z-chain is not qualitatively different than the foregoing.

Table 3. Moment Constrained Estimates, Second Prior

	Estimates			Z Prior		
Parameter	Mean	Mode	Std. Dev.	Mean	Mode	Std. Dev.
γ	53.093	56.720	22.845	206.42	7.3546	156.14
ψ	1.3110	1.2155	0.0757	1.3001	1.2996	0.0721
δ	0.9758	0.9730	0.0053	0.9759	0.9759	0.0051

As in the legend for Table 2 except that the prior is an independence prior with γ normal with mean 10 and standard deviation 255 constrained to have positive support, ψ with a cosine density with support (1.1,1.5); and δ a cosine density with support 1% to 4% per annum; see (24).

Moment Induced Bayes – Underlying Probability Space

- A likelihood $f\left(x \,|\, \theta_{(1)}\right)$, a prior $p(\theta) = p\left(\theta_{(1)}\right) p\left(\theta_{(2)}\right)$, and a restriction $0 = \int m(x_t, x_{t-1}, \theta_{(2)})\, f(x|\theta_{(1)})\, dx$ determine a joint probability space $(\mathcal{X} \times \Theta, \mathcal{C}^o, P^o)$.
  - Θ is singular wrt Lebesgue
- Marginalizing gives $(\mathcal{X} \times \Theta_{(2)}, \mathcal{C}^o_{(2)}, P^o_{(2)})$ with prior $p\left(\theta_{(2)}\right)$
  - $\Theta_{(2)}$ is not singular
  - Note that $p\left(\theta_{(2)}\right)$ is the marginal distribution of $\theta_{(2)}$
- Both the joint and marginal probability spaces determine the same same density $\psi(z)$ for the random variable $Z(x, \theta_{(2)})$

Moment Induced Bayes – Induced Probability Space

- Conversely, an assumption that $Z(x, \theta_{(2)})$ has density $\psi(z)$ and $\theta_{(2)}$ has marginal density $p(\theta_{(2)})$ induces a probability space $(\mathcal{X} \times \Theta_{(2)}, \mathcal{C}^*_{(2)}, P^*_{(2)})$. (Gallant, 2016a).

- The two probability spaces are equivalent in the sense that $P^o_{(2)}(C) = P^*_{(2)}(C)$ for every $C \in \mathcal{C}^*_{(2)}$.

- Therefore, $(\mathcal{X} \times \Theta_{(2)}, \mathcal{C}^*_{(2)}, P^*_{(2)})$ can be used as a substitute for $(\mathcal{X} \times \Theta, \mathcal{C}^o, P^o)$ for the purpose of Bayesian inference regarding $\theta_{(2)}$.

- The likelihood $f\left(x \,|\, \theta_{(2)}\right)$ on $(\mathcal{X} \times \Theta_{(2)}, \mathcal{C}^*, P^*)$ is $f\left(x \,|\, \theta_{(2)}\right) = \psi[Z(x, \theta_{(2)})]$.

- Semi-pivotal condition: $\left\{x : Z(x, \theta_{(2)}) = z\right\}$ not empty for any $\theta_{(2)}$ in the parameter space $\Theta_{(2)}$ and z in the range of Z

- There are footnotes, see Gallant (2016b, 2016c).

Moment Induced Bayes – Summary

- To summarize, moment induced Bayes assumes that

$$z = Z\big(x, \theta_{(2)}\big) = \sqrt{n}\left[W(x, \theta_{(2)})\right]^{-\frac{1}{2}}\left[\bar{m}(x, \theta_{(2)})\right].$$

  follows a distribution $\Psi(z)$ with density $\psi(z)$.

- One uses

$$p(x \,|\, \theta_{(2)}) = \psi(z)$$

  as the likelihood and proceeds directly to Bayesian inference using a prior $p(\theta_{(2)})$.

- Results for the example under the first and second priors follow.

Table 4. Moment Induced Estimates, First Prior

	Z Normal			Z Empirical			Z Prior		
Parameter	Mean	Mode	Std. Dev.	Mean	Mode	Std. Dev.	Mean	Mode	Std. Dev.
γ	12.460	11.669	2.2351	14.049	12.706	2.3716	10.018	9.9878	2.5513
ψ	1.3425	1.3497	0.1275	1.3494	1.3953	0.1197	1.5487	1.5541	0.1632
δ	0.9758	0.9758	0.0051	0.9759	0.9763	0.0051	0.9759	0.9759	0.0051

The estimation method is moment induced Bayes as described in Section 4. Data are real, annual, per capital consumption for the years 1930–2016 and real, annual stock, bond, and wealth returns for the same years from BEA (2016) and CRSP (2016) that are used to form the moment functions (6) through (6). In the columns labeled Z Normal the distribution of Z given by (10) is presumed to be Normal. In the columns labeled Z Empirical the distribution of Z given by (10) is presumed to be the distribution derived in Section 3; see also Figure 6. The prior is an independence prior with γ normal with mean 10 and standard deviation 2.55 constrained to have positive support, ψ with a cosine density with support (1.1,2.0); and δ a cosine density with support 1% to 4% per annum; see (23). The columns labeled mean, mode, and standard deviation are the mean and standard deviations of an MCMC chain (Gamerman and Lopes (2006) of length 200,000 collected past the point where transients have dissipated. The proposal is move-one-at-a-time random walk.

Table 5. Moment Induced Estimates, Second Prior

	Z Normal			Z Empirical			Z Prior		
Parameter	Mean	Mode	Std. Dev.	Mean	Mode	Std. Dev.	Mean	Mode	Std. Dev.
γ	87.490	138.15	40.917	14.889	13.357	3.1122	206.42	7.3546	156.14
ψ	1.3157	1.2964	0.0703	1.2605	1.2233	0.0707	1.3001	1.2996	0.0721
δ	0.9759	0.9759	0.0052	0.9759	0.9759	0.0053	0.9759	0.9759	0.0051

As in the legend for Table 4 except that the prior is an independence prior with γ normal with mean 10 and standard deviation 255 constrained to have positive support, ψ with a cosine density with support (1.1,1.5); and δ a cosine density with support 1% to 4% per annum; see (24).

Conclusions

- Specification of Ψ can make a substantial difference to moment induced Bayes in a representative application.
  - If Z has long, thin tails, the use of the correct Ψ robustifies estimates by making them less sensitive to Z outliers.
- Determination of $\Psi(z)$ from a z-chain is difficult and our determination was not as satisfactory as one might hope.
  - For this paper that does not matter because we did capture the qualitative feature of long thin tails.
  - Other approaches we tried such as a mixture of a large variance multivariate normal with a multivariate SNP density were less satisfactory
- Most likely an attempt to find a better way to determine $\Psi(z)$ is the wrong approach.
  - It would probably be more fruitful to investigate more easily implemented robustification strategies.
  - One that comes to mind is to fit a mixture such as $\psi(z) = p\, n(z \,|\, 0, \Sigma) + (1-p)\, n(z \,|\, 0, \alpha\Sigma)$ to the z-chain