Estimates of the long-run growth rate of Singapore with a CES production function

Rao, B. Bhaskara and Shankar, Sriram

University of Western Sydney

13 June 2011

Words:1,995

13/6/11

Estimates of the Long-run Growth Rate of Singapore with a CES Production Function

Sriram Shankar

S.Shankar@uws.edu.au

School of Economics and Finance

University of Western Sydney, Sydney (Australia)

B. Bhaskara Rao

raob123@bigpond.com

School of Economics and Finance

University of Western Sydney, Sydney (Australia)

Abstract

This paper estimates with the Bayesian methods a CES production function for Singapore for 1960-2009. It is found that the elasticity of substitution is 0.6, technical progress is labour augmenting and the steady state growth rate of Singapore is about 1.8%.

Keywords: Bayesian methods, CES production function and Technical progress.

JEL: O40 O47

1. Introduction

Singapore, China and India have been growing rapidly at 5%, 7% and 10% per annum respectively during 2000-2009. Are these growth rates permanent? If they are transitory, what are their long run growth rates? This paper tries to answer these questions with data from Singapore. Singapore is an interesting example. Its average rate of 8% during 1970-2000 has slowed down to 5% since 2000. India and China are also likely to grow eventually at a much slower rate like Singapore.

Long run growth rate means the steady state growth rate (SSGR). The existence of a SSGR depends on the nature of the production function. We assume constant returns to avoid the adding up problem. If the production function is Cobb-Douglas (CD) SSGR always exists. However, if it is the constant elasticity of substitution (CES) type, SSGR exists only if the elasticity of substitution is less than unity and technical progress is labour augmenting. Therefore, answers to our questions lie in the parameters of a CES function.

There are alternatives to estimate a CES, but we shall use the Bayesian approach from Luoma and Luoto (2010).The advantage is that it can be estimated with data on factor shares instead of the more difficult to obtain data on factor returns. Furthermore, the finite sample properties of the Bayesian estimators are well-articulated by Poirier (1995) and the exact finite sample properties for nonlinear functions like CES can be computed. Section 2 briefly discusses specification and estimation and empirical results are in Section 3. Section 4 concludes.

2. Specification and Estimation

Our specification and the underlying technology is similar to Luoma and Luoto (2010):

$$\text{(1)} \qquad \frac{Y_t}{\overline{Y}} = \zeta \left[\delta \left(\frac{K_t}{\overline{K}} e^{g_K(t,\bar{t})} \right)^{\frac{\sigma-1}{\sigma}} + (1-\delta) \left(\frac{H_t \times L_t}{\overline{H \times L}} e^{g_L(t,\bar{t})} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

where $\overline{Y}$, $\overline{K}$, $\overline{H \times L}$ and $\bar{t}$ are sample geometric averages of output, capital, labour embodying human capital and time respectively. These baseline averages are used to normalise the CES.

The parameter δ is the average capital income share, σ is the elasticity of substitution between capital and labour and ζ accounts for deviation of CES production function from a log-linear function; see Klump, McAdam and Willman 2007. This deviation primarily arises because for a non-linear function the sample geometric averages of the variables defined above may not coincide with the true fixed point. $g_K(t,\bar{t})$ is the rate of capital augmenting technical progress and similarly $g_L(t,\bar{t})$ is the rate of labour augmenting technical progress. They are defined as:

$$(2) \qquad g_j(t,\bar{t}) = \frac{\gamma_j \bar{t}}{\lambda_j}\left[\left(\frac{t}{\bar{t}}\right)^{\lambda_j} - 1\right] \quad t > 0, \quad j \in \{K, H \times L\}$$

where λ_j and γ_j measures the shape and the speed of adjustment of factor augmenting technological progress. When the shape parameter takes the values 1, 0, or less than 0, the technological progress function $g_j(t,\bar{t})$ is linear, log-linear and hyperbolic respectively. Thus estimates of these parameters can be used to understand the nature of the production technology.

Taking logarithm on both sides of (1), it can be written as

$$(1a) \qquad \ln\frac{Y_t}{\bar{Y}} = \ln\zeta + \frac{\sigma}{\sigma-1}G\left(Y_t, H_t \times L_t, \delta, \sigma, \gamma_K, \lambda_k, \gamma_{H\times L}, \lambda_{H\times L}\right) + \varepsilon$$

where ε is a *iid* random error variable and

$$G\left(Y_t, H_t \times L_t, \delta, \sigma, \gamma_K, \lambda_k, \gamma_{H\times L}, \lambda_{H\times L}\right) = \ln\left[\delta\left(\frac{K_t}{\bar{K}}e^{g_K(t,\bar{t})}\right)^{\frac{\sigma-1}{\sigma}} + (1-\delta)\left(\frac{H_t \times L_t}{\overline{H \times L}}e^{g_L(t,\bar{t})}\right)^{\frac{\sigma-1}{\sigma}}\right].$$

Equation (1a) can be estimated using maximum likelihood, but as the likelihood function for CES technology is multi-modal, the parameter may not converge to the global optimum. Therefore, we estimate (1a) with the Bayesian methodology and priors guided by economic theory. The likelihood function for a sample of T output observations $\mathbf{Y} = \left(Y_1, ..., Y_T\right)$ can be written in terms of inputs $\mathbf{K} = \left(K_1, ..., K_T\right)$, $\mathbf{H \times L} = \left(H_1 \times L_1, ..., H_T \times L_T\right)$ and the parameters of technology as follows:

(3) $$L(\boldsymbol{\theta}|\mathbf{Y},\mathbf{K},\mathbf{H}\times\mathbf{L}) \propto h^{\frac{T}{2}} \exp\left\{-\frac{h}{2}\sum_{t=1}^{T}\left(\ln\frac{Y_t}{\bar{Y}} - \ln\zeta + \tfrac{1-\sigma}{\sigma}G(Y_t, H_t \times L_t, \delta, \sigma, \gamma_K, \lambda_k, \gamma_{H\times L}, \lambda_{H\times L})\right)^2\right\}$$

where $\boldsymbol{\theta} = (\zeta, \delta, \sigma, \gamma_K, \lambda_k, \gamma_{H\times L}, \lambda_{H\times L})$, and h (the inverse of variance) is commonly referred to as the precision parameter in the Bayesian literature. We use the following priors for $\boldsymbol{\theta}$ and h:

(4) $$p(\boldsymbol{\theta}, h) \propto p(\boldsymbol{\theta}|h)p(h)$$

We elicit a prior for $\boldsymbol{\theta}$ conditional on h of the form[1]

(5) $$p(\boldsymbol{\theta}|h) \quad = \quad f_N(\boldsymbol{\theta}|\underline{\boldsymbol{\theta}}, h^{-1}\underline{\mathbf{V}})\, I(\boldsymbol{\theta} \in R)$$

(6) $$p(h) \quad = \quad f_G(h|\underline{s}^{-2}, \underline{\nu})$$

where $I(.)$ is an indicator function taking the value 1 if the argument is true and 0 otherwise, and R is the region in the parameter space where $0 < \zeta < \infty, 0 < \delta < 1$, $0 < \sigma < \infty$ and the remaining parameters may take any value on the real line. In (5) $\underline{\boldsymbol{\theta}}$ and $\underline{\mathbf{V}}$ are prior mean and covariance matrix respectively for the parameter $\boldsymbol{\theta}$ and in (6) $\underline{s}^{-2}$ and $\underline{\nu}$ are the prior shape and scale parameter respectively for h.

The posterior is proportional to the product of likelihood function given by (3) and the joint prior given by (4). In order to sample from posterior density, it is helpful to use a Gibbs sampler. In Gibbs sampling algorithm, (see Gelfand and Smith, 1990) draws from joint posterior density are generated by sampling from a series of conditional posteriors. However, to impose constraints on the parameters requires a method for drawing observations from a truncated multivariate distribution. In Bayesian literature accept-reject (A-R) and Metropolis-Hastings (M-H) algorithms are the most commonly used to draw observations from a target distribution. A disadvantage of A-R algorithm (see Gelfand and Lee, 1993) is that for some complex conditional posterior densities one would need to generate a very large number of candidate draws before finding one that is acceptable. Therefore in our simulation we prefer

---

[1] $f_N(\boldsymbol{\alpha}|\boldsymbol{\beta}, \mathbf{C})$ indicates that $\boldsymbol{\alpha}$ is a multivariate normal vector with mean $\boldsymbol{\beta}$ and covariance matrix $\mathbf{C}$ and $f_G(a|b, c)$ indicates that a has a gamma distribution with shape parameter b and scale parameter c.

to use a more efficient random-walk Metropolis-Hastings (M-H) algorithm (see Chib and Greenberg, 1995).

The Gibbs sampler necessitates drawing sequentially from the following conditional posteriors:

$$(7)\quad p(\boldsymbol{\theta}|\mathbf{Y},\mathbf{K},\mathbf{H}\times\mathbf{L},h)\propto \exp\left\{-\frac{h}{2}\sum_{t=1}^{T}\left(\ln Y_t-\ln\zeta+\tfrac{1-\sigma}{\sigma}G\left(Y_t,H_t\times L_t,\boldsymbol{\theta}\right)\right)^2\right\}\times \exp\left\{-\frac{1}{2}(\boldsymbol{\theta}-\underline{\boldsymbol{\theta}})'\underline{\mathbf{V}}^{-1}(\boldsymbol{\theta}-\underline{\boldsymbol{\theta}})\right\}I(\boldsymbol{\theta}\in R)$$

and

$$(8)\quad p(h|\mathbf{Y},\mathbf{K},\mathbf{H}\times\mathbf{L},\boldsymbol{\theta})\propto f_G(h|\bar{s}^{-2},\bar{\nu})$$

where

$$(9)\quad \bar{\nu} = T+\underline{\nu}$$

and

$$(10)\quad \bar{s}^2 = \frac{\left[\sum_{t=1}^{T}\left(\ln Y_t-\ln\zeta+\tfrac{1-\sigma}{\sigma}G\left(Y_t,H_t\times L_t,\delta,\sigma,\gamma_K,\lambda_k,\gamma_{H\times L},\lambda_{H\times L}\right)\right)^2\right]+\underline{\nu s^2}}{\bar{\nu}}$$

Draws from these conditional posteriors will converge to draws from the posterior $p(\boldsymbol{\theta},h|\mathbf{Y},\mathbf{K},\mathbf{H}\times\mathbf{L})$. Simulating from the gamma density (8) is straightforward using random number generators available in most softwares. However, simulating from (7) is slightly complicated because it is a truncated probability density function (pdf). To simulate from (7) we used a random walk Metropolis-Hastings algorithm with a multivariate normal proposal density; see Koop (2003, p.92). During the transition, or burn-in phase of the algorithm, the covariance matrix of the proposal density was set to a scalar multiplied by an identity matrix. The scalar was set by trial and error to yield an acceptance rate in the range 0.3-0.5.[2] After

[2] There is no rule for the best acceptance rate. Roberts et al. (1997) show that if the target and proposal densities are normal pdfs, the optimal acceptance rate is between 0.45 in one-dimensional problems and approximately 0.23 as the number of dimensions becomes infinitely large.

the burn-in, to improve the efficiency of the algorithm, we used the covariance matrix of the burn-in observations as the covariance matrix in the proposal density.

3. Empirical Results:

We choose $\underline{\boldsymbol{\theta}} = (1, 0.33, 1, -1, 0.02, 1, 0.01)$, $\underline{\mathbf{V}} = 0.25 I_7$, $\underline{\nu} = 12$, and $\underline{s}^{-2} = 10$.[3] Given the likely magnitudes of the marginal products of effective labour and capital and the way we have normalized the data, these choices are sensible, but relatively non-informative. We simulated 400,000 observations from the conditional posteriors (7) and (8), and discarded the first 200,000 draws as burn-in.[4] Figures A1 and A2 in the appendix present convergence plots for each of the elements of $\boldsymbol{\theta}$. They clearly indicate that the Markov chain Monte Carlo (MCMC henceforth) sequence for all the parameters is stationary. We formally checked convergence for each of the parameters using Gelman and Rubin's (1990) diagnostic R.

Estimates of the unknown parameters for the periods 1960-2009 and 1960-2004 are presented in the first and second columns of Table 1. We added data for five years to our original sample of 1960-2004 to see how sensitive our estimates to the sample size are. It can be seen that parameter estimates in both samples are close. The point estimates are the means of the MCMC samples and are optimal Bayesian point estimates under quadratic loss. The inequality restrictions in the prior (5) ensure that all the estimates in Table 1 are theoretically plausible.

[3] We used MATLAB for simulating the posterior conditional densities.

[4] The prior variance-covariance matrix is a diagonal matrix since it is not easy to guess what they might be.

Table 1 Estimated results for Singapore

$$\frac{Y_t}{\overline{Y}} = \zeta\left[\delta\left(\frac{K_t}{\overline{K}}e^{g_K(t,\overline{t})}\right)^{\frac{\sigma-1}{\sigma}} + (1-\delta)\left(\frac{H_t \times L_t}{\overline{H \times L}}e^{g_L(t,\overline{t})}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

$$g_j(t,\overline{t}) = \frac{\gamma_j \overline{t}}{\lambda_j}\left[\left(\frac{t}{\overline{t}}\right)^{\lambda_j} - 1\right] \quad t > 0, \quad j \in \{K, H \times L\}$$

Parameter	1960-2009	1960-2004
ζ	0.8976 [0.8266, 0.9801]	0.9124 [0.8364,1.0087]
δ	0.2739 [0.0377, 0.6042]	0.2663 [0.0124,0.6445]
σ	0.6318 [0.3646, 1.1913]	0.6155 [0.3631,1.1427]
λ_k	0.8584 [-0.1580,1.8434]	0.8735 [-0.0866,1.8845]
γ_K	-0.0121 [-0.0899, 0.1049]	-0.0024 [-0.1266,0.2152]
$\lambda_{H\times L}$	1.0668 [0.4777, 1.8450]	1.1528 [0.4560,1.9579]
$\gamma_{H\times L}$	0.0345 [-0.0222, 0.0696]	0.0392 [-0.0077, 0.0918]
h	42.5588 [28.5151, 59.2713]	34.7397 [21.6795, 50.8327]
TFP Growth	0.0182 [-0.0085, 0.0448]	0.0199 [-0.0083, 0.0474]

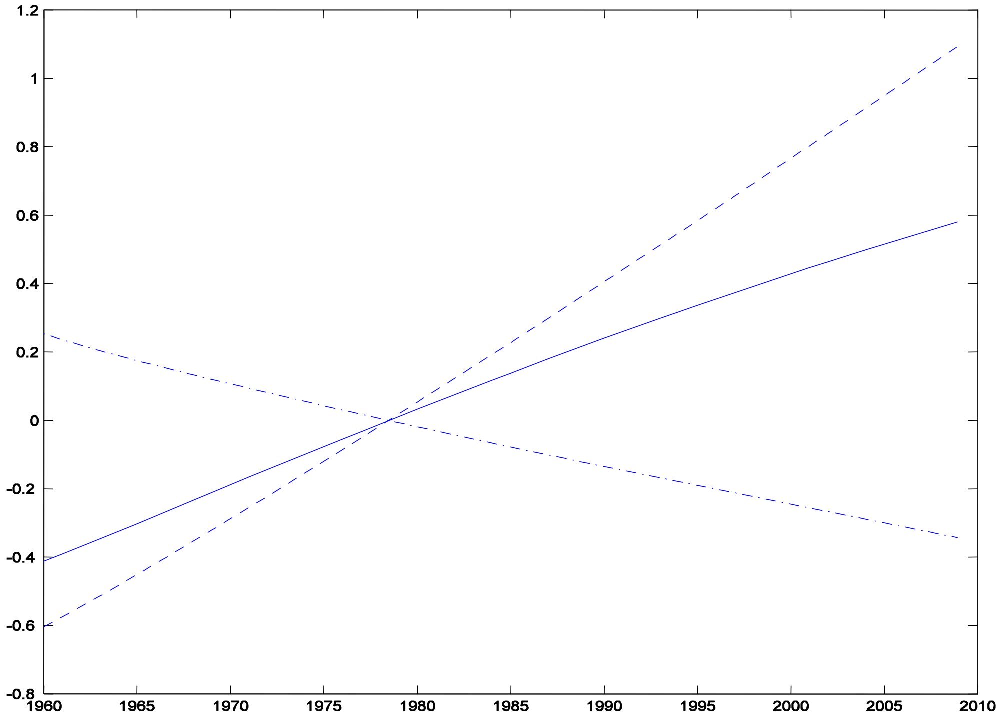
Notes: Labour-augmenting $g_{H\times L}(---)$, capital-augmenting $g_K(-.-.-)$, and $\ln TFP(-)$. Total factor productivity is calculated by applying Kmenta (1967) approximation around the fixed points where:

$$\ln TFP_t = \delta g_K(t,\overline{t}) + (1-\delta)g_{H\times L}(t,\overline{t}) - \tfrac{(1-\sigma)\delta(1-\delta)}{2\sigma}\left[g_K(t,\overline{t}) - g_{H\times L}(t,\overline{t})\right]^2.$$

The confidence intervals in the square brackets are for the MCMC samples and suggest that all the parameter estimates are reliable. A more complete picture of the level of uncertainty surrounding the unknown parameters is presented in Figures A2. These figures present estimated marginal posterior pdfs for each of the parameters for Singapore. A feature of these pdfs is that the estimated pdfs for δ and σ are asymmetric. This is a due to the inequality restrictions incorporated in the prior.

Figure 1

Technological progress for Singapore in log levels



Notes: Labour-augmenting $g_{H\times L}(---)$, capital-augmenting $g_K(-.-.-)$, and $\ln TFP(-)$. Total factor productivity is calculated by applying Kmenta (1967) approximation around the fixed points where:

$$\ln TFP_t = \delta g_K(t,\bar{t}) + (1-\delta) g_{H\times L}(t,\bar{t}) - \tfrac{(1-\sigma)\delta(1-\delta)}{2\sigma}\left[g_K(t,\bar{t}) - g_{H\times L}(t,\bar{t})\right]^2.$$

Estimates of the two crucial parameters viz., share of capital (δ) and elasticity of substitution (σ) are close in both samples. The share of capital at 0.27 is virtually the same in both samples and not significantly different from its stylised value of one third. On the other hand the elasticity parameter at about 0.63 in the larger sample is significantly less than unity. This implies that estimates of TFP with the CD production function are somewhat overestimated as can be seen from the equation for $\ln TFP$ in the notes for Table 1. The nature of TFP i.e., whether it is labour or capital augmenting is plotted in Figure 1. It can be seen that up to the mid 1970s, capital augmenting TFP has dominated and since then labour augmented TFP has dominated. This is to be expected because it is relatively quicker to invest in technically superior capital equipment in the initial stages of a country's development. However, considerable time is needed to train and educate the labour force. Towards the end of our sample period in 2009, labour augmented TFP is about 4.5 times higher than capital augmented TFP. Therefore, it is reasonable to conclude that TFP in Singapore since the mid 1970s is labour augmenting and our estimate for the TFP growth of 1.8% is a reasonable estimate of Singapore's SSGR.

4. Conclusions

We estimated a CES production function for Singapore with the Bayesian methods and found that the elasticity of factor substitution is well below unity. Since Singapore's TFP is mainly labour augmenting, especially since the mid 1970s, it is reasonable to conclude that the long run growth rate for Singapore, as we found, is 1.8%.

Appendix

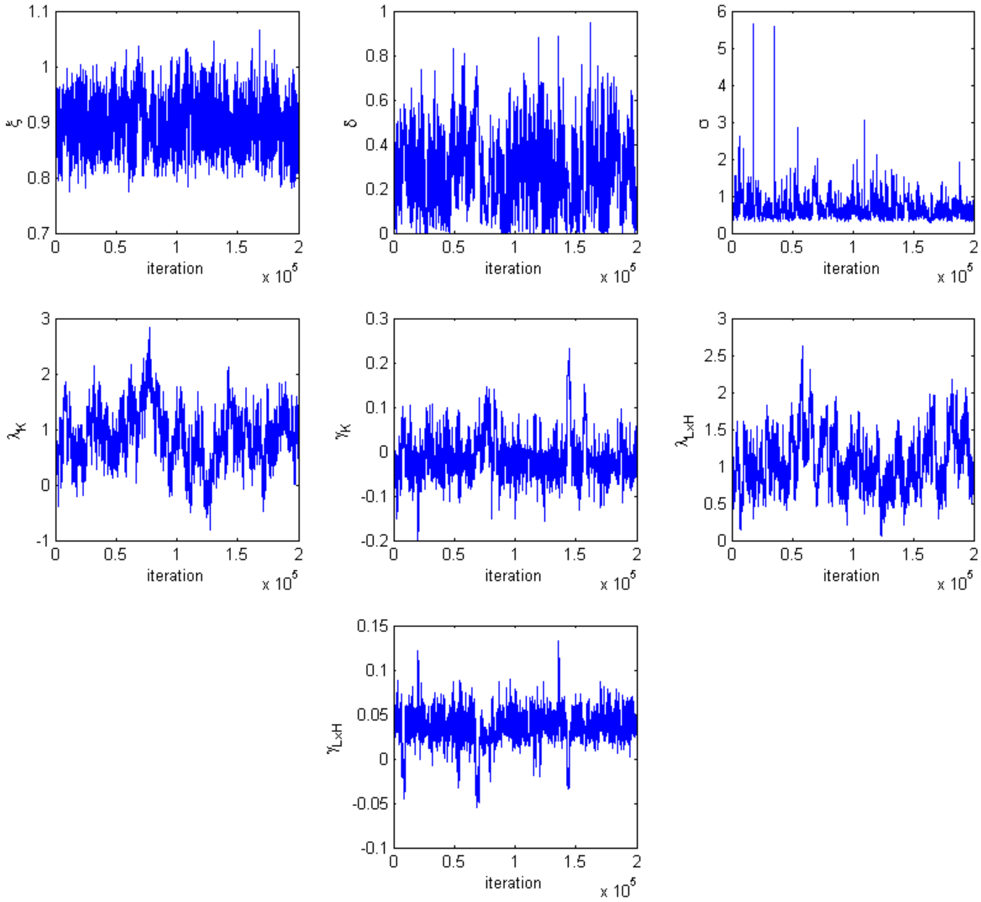


Figure A1: Convergence plots for Singapore

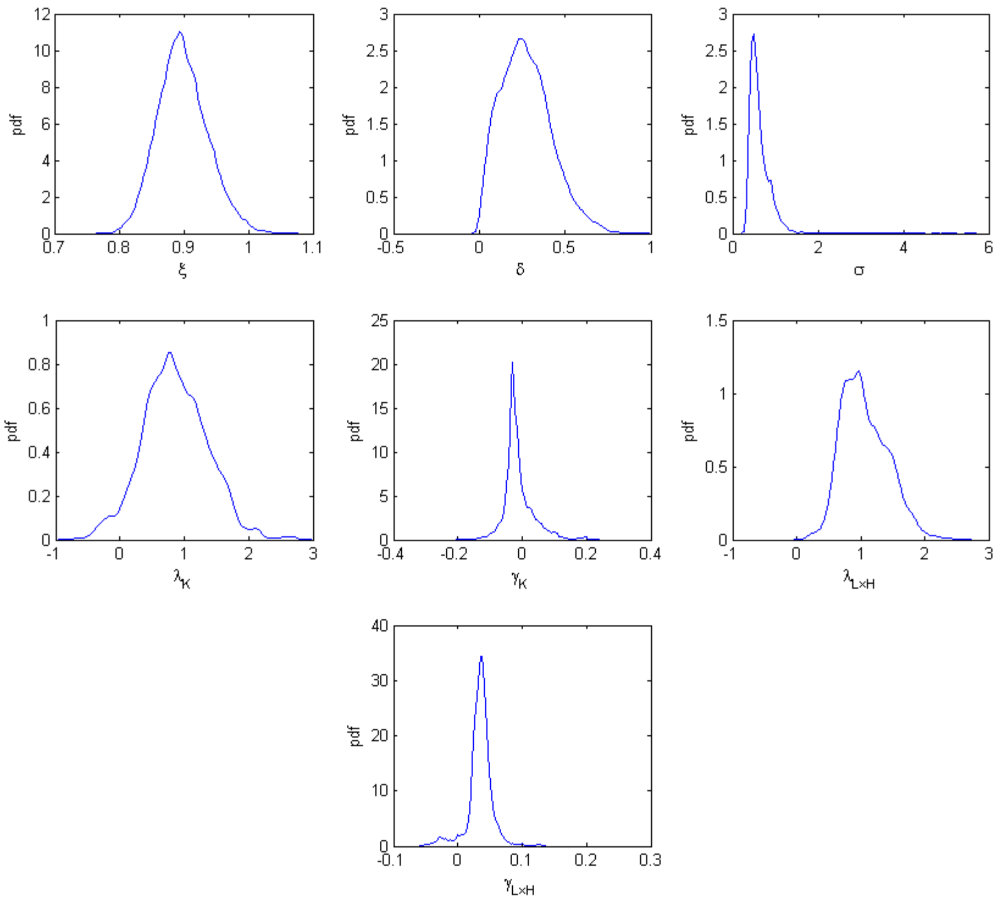


Figure A2: Estimated Posterior Pdfs for Singapore

References

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-H
astings algorithm, The American Statistician, 49: 327-335.

Gelfand, A. E. and Lee, T.-M. (1993). Discussion of the meeting on the Gibbs sampler and other Markov chain Monte Carlo methods, Journal of the Royal Statistical Society, Series B 55: 72-73.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, Journal of the American Statistical Association 85: 398-409.

Gelfand, A. E., Smith, A. F. M. and Lee, T.-M. (1992). Bayesian analysis of constrained parameter and trucated data problems using Gibbs sampling, Journal of American Statistical Association 87: 523-532.

Gelman, A. and Rubin, D. B. (1992) Inference from Iterative Simulation Using Multiple Sequences, Statist. Sci. 7: 457-472.

Koop, G. (2003). *Bayesian Econometrics*, John Wiley and Sons, Hoboken NJ.

Poirier, D. (1995) *Intermediate Statistics and Econometrics: A Comparative Approach* The MIT Press, Cambridge MA.

Roberts, G.O., Gelman A. and Gilks W.R. (1997) "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms" Annals of Applied Probability 7:110-120.