KERP

2002/07

On the Variance Covariance Matrix of the Maximum Likelihood Estimator of a Discrete Mixture

Gauthier Lanot

Keele Economics Research Papers

KEELE UNIVERSITY

Keele ▪ July 2002

KERP **Keele Economics Research Papers**

The Keele Economics Department produces this series of research papers in order to stimulate discussion and invite feedback. Copyright remains with the authors.

All papers in the KERP series are available for downloading from the Keele Economics website, via *www.keele.ac.uk/depts/ec/kerp*.

ISSN 1352-8955

Department of Economics ▪ Keele University, Keele, Staffordshire, ST5 5BG, UK
▪ tel: (44) 1782 583091, fax: (44) 1782 717577 ▪ email: economics@keele.ac.uk

On the Variance Covariance Matrix of the Maximum Likelihood Estimator of a Discrete Mixture

by

Gauthier Lanot (Keele University)

Date July 2002

Abstract The estimation of models involving discrete mixtures is a common practice in econometrics, for example to account for unobserved heterogeneity. However, the literature is relatively uninformative about the measurement of the precision of the parameters. This note provides an analytical expression for the observed information matrix in terms of the gradient and hessian of the latent model when the number of components of the discrete mixture is known. This in turn allows for the estimation of the variance covariance matrix of the ML estimator of the parameters. I discuss further two possible applications of the result: the acceleration of the EM algorithm and the specification testing with the information matrix test.

Keywords Discrete Mixtures; EM Algorithm, Variance Covariance Matrix; Observed Information

J.E.L. Class C1-C4

Address Department of Economics, Keele, Staffordshire ST5 5BG, United Kingdom, E-mail: g.lanot@keele.ac.uk; Fax: +44 -1782 717577; Tel: +44 1782 583102.

Download www.keele.ac.uk/depts/ec/web/wpapers/kerp0207.pdf

1 Introduction

In practice the specification of many microeconometric models requires the introduction of components capturing unobserved heterogeneity to account, inter alia, for overdispertion (in Poisson models, in duration models), for specific effects (when dealing with repeated measurements, or in the case of cluster specific effects) , for uncertainty about the model that should be considered (in the case of switching regression with unobserbed regime) etc…

The specification of the model of the observed quantities of interest can often be decomposed into two parts. The first part describes the distribution of the unobserved component, and given the unobserved element and in general some covariates, the second part describes the conditional distribution (density) of the quantity of interest. Formally the distribution of the observed variable of interest, say Y , given a vector of covariates, say X , takes the general form of a mixture :

$$\Pr\left[Y = y \mid X = x\right] = \int_{D_\varepsilon} \Pr\left[Y = y \mid X = x, \varepsilon = e\right] dF\left(e\right), \tag{1}$$

where ε is the unobserved random variable which describes unobserved heterogeneity over some domain of definition D_ε, and where $\Pr\left[Y = y \mid X = x, \varepsilon = e\right]$ is usually fully specified. $F\left(e\right)$ is the cumulative distribution function of ε. At this stage $F\left(e\right)$ can be assumed to belong to a well specified parametric family in which case the calculation of the observed likelihood is straightforward and the estimation proceeds, more or less directly, from there. Alternatively it can be left unspecified and the estimation problem can be thought essentially as a non-parametric problem. However, in this latter case it can be shown that the maximum likelihood estimator of the mixture takes the form of a finite discrete mixture (i.e. a list of discrete locations, i.e. values of ε in D_ε and probability weigths, see Lindsay (1983)), which in practice allows/demands the use of conventional maximum likelihood arguments.

In this context the estimation problem is often solved using the EM algorithm (see for example Gouriéroux & Monfort, 1995). In this note, following Oakes (1999), I show that the EM algorithm in the discrete mixture case allow for a relatively simple evaluation of the matrix variance covariance of all the parameters of interest. I give a general analytic expression for the hessian of the observed likelihood of a model with finite discrete mixture with known number of types in terms of the gradient and the hessian of the latent likelihood.

Finally I discuss how this analytical expression allows for quasi-Newton acceleration of the EM algorithm, and for an Information Matrix specification test.

2 The Model

In what follows I assume that there are N (independent and identically distributed) observations/clusters, indexed $i, i = 1...N$, and F (given) number of components to the mixture. However the allocation of observations to types is unobserved. For a given observation i and type f the contribution to the likelihood is $L_{if}(\theta)$ (I omit to indicate the values Y and X take…) where θ is a vector of k parameters. All the unknown locations γ_f, $f = 1...F$, are elements of θ, hence $k \geq F$. Indeed, in some case the model of interest will include some covariates and the parameters (type invariant or not) associated with the covariates are included in θ. On the other hand the probabilities of each type p_f, $f = 1...F$, $\sum_{f=1}^{F} p_f = 1$, are not collected in θ. Assuming that we observe the type of each observation i the (latent) likelihood can be written as

$$\mathfrak{L}(\phi) = \prod_{i=1}^{N}\prod_{f=1}^{F}\left(p_f L_{if}(\theta)\right)^{\delta_{if}}, \qquad (2)$$

where $\phi' = (\theta', p_1, ..., p_{F-1})$, and $\delta_{if} = 1$ if observation i is of type f, and 0 otherwise. Note that the latent likelihood given complete observation $\mathfrak{L}(\phi)$ is to be distinguished from the observed likelihood given partial observation, $L(\phi)$. In principle $L(\phi)$ is easily defined:

$$L(\phi) = \prod_{i=1}^{N}\sum_{f=1}^{F} p_f L_{if}(\theta),$$

however in practice it may be difficult to evaluate (see Lee, 2000) and/or difficult to maximise. In what follows I'll assume that the evaluation of $L_{if}(\theta)$ and its derivative is "straightforward" (or known, or at least easier to obtain than the equivalent quantities from the observed likelihood).

The latent log-likelihood can therefore be written as

$$\ln \mathfrak{L}(\phi) = \sum_{i=1}^{N}\sum_{f=1}^{F}\delta_{if}\left\{\ln p_f + \ln L_{if}(\theta)\right\}. \qquad (3)$$

For a given value of the parameters, collected in $\psi' = \left(\theta_0{}', p_{01}, \cdots, p_{oF-1}\right)$, the EM algorithm proceeds first (Expectation step) by calculating the Expected latent log-likelihood given what is observed (which we represent by $\mathcal{O}_i$), we have

$$\begin{aligned} \ln L(\phi;\psi) &= \mathrm{E}_\psi\left[\ln \mathfrak{L}(\phi) \mid \{\mathcal{O}_i\}_{i=1}^N\right] \\ &= \sum_{i=1}^N \sum_{f=1}^F \left\{\mathrm{E}_\psi\left[\delta_{if} \mid \mathcal{O}_i\right] \ln p_f + \mathrm{E}_\psi\left[\delta_{if} \mid \mathcal{O}_i\right] \ln L_{if}(\theta)\right\}, \end{aligned} \tag{4}$$

where $\mathrm{E}_\psi[.|.]$ stand for the conditional expectation calculated with the joint distribution indexed by the vector of parameters ψ. In particular it can be shown that

$$\mathrm{E}_\psi\left[\delta_{if} \mid \mathcal{O}_i\right] = \frac{p_{0f} L_{if}(\theta_0)}{\sum_{g=1}^F p_{0g} L_{ig}(\theta_0)} = \pi_{if}(\psi). \tag{5}$$

In the second stage (Maximisation step) $\ln L(\phi;\psi)$ is maximised with respect to ϕ. This procedure is repeated until convergence (i.e. until $\psi = \arg\max_\phi \ln L(\phi;\psi)$). The algorithm is known to be monotonic (i.e. the observed likelihood increases), and if convergent it solves the likelihood equations. Furthermore in some cases it can be shown to yield the maximum likelihood estimator, i.e. $\psi = \arg\max_\phi \ln L(\phi)$.

3 The Hessian of the Observed Likelihood

In a recent paper Oakes (1999) shows that the Hessian of the observed log-likelihood, $\frac{\partial^2 \ln L(\phi)}{\partial\phi\partial\phi'}$, can be obtained from derivatives of $\ln L(\phi;\psi)$. For all ϕ, we have

$$\frac{\partial^2 \ln L(\phi)}{\partial\phi\partial\phi'} = \left\{\frac{\partial^2 \ln L(\phi;\psi)}{\partial\phi\partial\phi'} + \frac{\partial^2 \ln L(\phi;\psi)}{\partial\phi\partial\psi'}\right\}\Bigg|_{\phi=\psi}. \tag{6}$$

While we would expect the first term (the information matrix for the parameters of the latent model) to be definite negative, the second term which represents the missing information is likely to be definite positive. In the multivariate context we may expect that the first term dominate the second term (in the sense that the difference between the two terms is a negative definite matrix).

In the context of discrete mixtures the evaluation of the quantities of interest is straightforward. Direct calculations of the gradients with respect to the components of ϕ give

$$\frac{\partial \ln L(\phi,\psi)}{\partial \theta_1} = \sum_{i=1}^{N}\sum_{f=1}^{F} \pi_{if}(\psi) G_{if}(\theta_1) = \sum_{i=1}^{N} \bar{G}_i(\theta_1;\psi), \tag{7}$$

with $G_{if}(\theta_1) = \dfrac{\partial \ln L_{if}(\theta_1)}{\partial \theta_1}$, and $\bar{G}_i(\theta_1;\psi) = \sum_{f=1}^{F} \pi_{if}(\psi) G_{if}(\theta_1)$.

$$\frac{\partial \ln L(\phi;\psi)}{\partial p_{1f}} = \sum_{i=1}^{N}\left(\frac{\pi_{if}(\psi)}{p_{1f}} - \frac{\pi_{iF}(\psi)}{p_{1F}}\right), \text{ for all } f = 1 \ldots F-1. \tag{8}$$

The required second derivatives are then easy to obtain, we have

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial \theta_1 \partial \theta_1{}'} = \sum_{i=1}^{N}\sum_{f=1}^{F} \pi_{if}(\psi) H_{if}(\theta_1) = \sum_{i=1}^{N} \bar{H}_{if}(\theta_1), \tag{9}$$

with $H_{if}(\theta_1) = \dfrac{\partial^2 \ln L_{if}(\theta_1)}{\partial \theta_1 \partial \theta_1'}$ and $\bar{H}_{jf}(\theta_1;\psi) = \sum_{f=1}^{F} \pi_{if}(\psi) H_{if}(\theta_1)$.

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial p_{1f} \partial p_{1g}} = \sum_{i=1}^{N}\left(-\frac{\pi_{if}(\psi)}{p_{1f}{}^2}\mathbf{1}_{[g=f]} - \frac{\pi_{iF}(\psi)}{p_{1F}{}^2}\right), \tag{10}$$

for all $f, g = 1 \ldots F-1$.

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial \theta_1 \partial p_{1f}} = 0, \text{ for all } f = 1 \ldots F-1. \tag{11}$$

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial \theta_1 \partial \theta_0{}'} = \sum_{i=1}^{N} Cov_{\pi_i(\psi)}\left[G_{if}(\theta_1), G_{if}(\theta_0)\right], \tag{12}$$

where

$$Cov_{\pi_i(\psi)}\left[G_{if}(\theta_1), G_{if}(\theta_0)\right] =$$

$$\sum_{f=1}^{F}\left\{\pi_{if}(\psi) G_{if}(\theta_1) G_{if}(\theta_0)'\right\} - \sum_{f=1}^{F}\left\{\pi_{if}(\psi) G_{if}(\theta_1)\right\} \sum_{f=1}^{F}\left\{\pi_{if}(\psi) G_{if}(\theta_0)'\right\},$$

hence if $\theta_0 = \theta_1$,

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial \theta_1 \partial \theta_0{}'} = \sum_{j=1}^{J} Var_{\pi_j(\psi)}\left[G_{jf}(\theta_0)\right], \text{symmetric and p.s.d.} . \tag{12$'$}$$

Moreover it is straightforward to show that

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial\theta_1\partial\theta_1{}'} + \frac{\partial^2 \ln L(\phi;\psi)}{\partial\theta_1\partial\theta_0{}'} =$$
$$\sum_{i=1}^{N}\sum_{f=1}^{F}\pi_{if}(\psi)\left\{\frac{\mathcal{H}_{if}(\theta_0)}{L_{if}(\theta_0)} - G_{if}(\theta_0)G_{if}(\theta_0)'\right\} +$$
$$\sum_{i=1}^{N}\sum_{f=1}^{F}\pi_{if}(\psi)G_{if}(\theta_0)G_{if}(\theta_0)' - \sum_{i=1}^{N}\bar{G}_i(\theta_0)\bar{G}_i(\theta_0)' = \tag{13}$$
$$\sum_{i=1}^{N}\sum_{f=1}^{F}\frac{p_{0f}\mathcal{H}_{if}(\theta_0)}{\sum_{g=1}^{F}p_{0g}L_{ig}(\theta_0)} - \sum_{i=1}^{N}\bar{G}_i(\theta_0)\bar{G}_i(\theta_0)'$$

where $\mathcal{H}_{if}(\theta_0) = \dfrac{\partial^2 L_{if}(\theta_0)}{\partial\theta\partial\theta'}$.

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial p_{1f}\partial\theta_0{}'} =$$
$$\sum_{i=1}^{N}\left(\frac{\pi_{if}(\psi)}{p_{1f}}\left(G_{if}(\theta_0) - \bar{G}_i(\theta_0)\right)' - \frac{\pi_{iF}(\psi)}{p_{1F}}\left(G_{iF}(\theta_0) - \bar{G}_i(\theta_0)\right)'\right) \tag{14}$$
$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial\theta_1\partial p_{0f}} =$$
$$\sum_{i=1}^{N}\left(\frac{\pi_{if}(\psi)}{p_{0f}}\left(G_{if}(\theta_1) - \bar{G}_i(\theta_1)\right) - \frac{\pi_{iF}(\psi)}{p_{0F}}\left(G_{iF}(\theta_1) - \bar{G}_i(\theta_1)\right)\right) \tag{15}$$

such that whenever $\phi = \psi$, $\dfrac{\partial^2 \ln L(\phi;\phi)}{\partial p_{1f}\partial\theta_0{}'} = \dfrac{\partial^2 \ln L(\phi;\phi)}{\partial\theta_1\partial p_{0f}}'$.

Finally for all $f, g = 1 \ldots F-1$, we have :

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial p_{1f}\partial p_{0g}} =$$
$$\sum_{i=1}^{N}\left(\frac{\pi_{if}(\psi)}{p_{1f}p_{0f}}\mathbf{1}_{[g=f]} + \frac{\pi_{iF}(\psi)}{p_{1F}p_{0F}}\right) - \tag{16}$$
$$\left(\frac{\pi_{if}(\psi)}{p_{1f}} - \frac{\pi_{iF}(\psi)}{p_{1F}}\right)\left(\frac{\pi_{ig}(\psi)}{p_{0g}} - \frac{\pi_{iF}(\psi)}{p_{0F}}\right),$$

such that whenever $p_{1f} = p_{0f}$ for all $f = 1 \ldots F-1$,

$$\frac{\partial^2 \ln L(\phi;\psi)}{\partial p_{1f}\partial p_{1g}} + \frac{\partial^2 \ln L(\phi;\psi)}{\partial p_{1f}\partial p_{0g}} = -\sum_{i=1}^{N}\left(\frac{\pi_{if}(\psi)}{p_{0f}} - \frac{\pi_{iF}(\psi)}{p_{0F}}\right)\left(\frac{\pi_{ig}(\psi)}{p_{0g}} - \frac{\pi_{iF}(\psi)}{p_{0F}}\right) \tag{17}$$

which is symmetric in the indices f and g.
These expressions define the observed information, and therefore can be used together to obtain an estimate of the variance covariance of the parameters (i.e. by taking the inverse of the observed information matrix).

4 Implications

A first direct consequence of the expression above is that it is in principle possible to accelerate the convergence of the EM algorithm. At least when the likelihood gradient is small enough (this would have to be determined in practice), the calculation of the hessian of the observed likelihood allows one or more "safe" Newton Raphson step (safe in the sense that it does not lead to a reduction in the likelihood) with the advantage of a quicker convergence. Moreover, the expressions above can be used in conjunction with the acceleration methods proposed elsewhere (on this topic see for example Louis, 1982 and Jamshidian & Jennrich, 1997) to lead to faster convergence of the EM algorithm.

Furthermore, the formulae above provide expressions, in terms of the latent likelihood, for the restrictions that have to hold under correct specification, i.e. when the information matrix equality holds. In particular, given correct specification and for any value of the parameters we have (allowing for the covariates)

$$\mathrm{E}_{\mathrm{X}}\left[\mathrm{E}_{\theta,p}\left[\sum_{f=1}^{F}\frac{p_f \mathcal{K}_f(\theta)}{\sum_{g=1}^{F} p_g L_g(\theta)}\right]\right] = \mathrm{E}_{\mathrm{X}}\left[\mathrm{E}_{\theta,p}\left[\sum_{f=1}^{F}\pi_f(\psi)\left\{H_f(\theta) + G_f(\theta)G_f(\theta)'\right\}\right]\right] = 0, \tag{18}$$

as a consequence of $\int_{D_Y}\sum_{f=1}^{F} p_f L_f(y;\theta)\,dy = 1$ and differentiating twice with respect to θ. Furthermore, we can easily see that (again as a consequence of $\int_{D_Y} \sum_{f=1}^{F} p_f L_f(y;\theta)dy = 1$ this time differentiating with respect to θ first and then with respect to p_f), for all $f = 1 \ldots F-1$, and for $\phi = \psi$:

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[\frac{\pi_f(\psi)}{p_f}G_f(\theta) - \frac{\pi_F(\psi)}{p_F}G_F(\theta)\right]\right] =$$

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[\frac{\frac{\partial L_f(\theta)}{\partial\theta} - \frac{\partial L_F(\theta)}{\partial\theta}}{\sum_{g=1}^{F} p_g L_g(\theta)}\right]\right] = 0. \tag{19}$$

These restrictions are the basis of the Information Matrix test (for an introduction see Gouriéroux & Monfort, 1995, for more details see White, 1994). As an illustration of the type of restrictions obtained, consider the example of a mixture involving two types, and such that, given the type, the latent distribution of Y given a vector of covariates X, is Poisson with parameter $\lambda_f(x) = \exp(x\beta + \gamma_f)$. In this context $\theta' = (\beta', \gamma_1, \gamma_2)$.

We have the following expressions for $G_f(\theta)$ and $H_f(\theta)$

$$G_f(\theta) = \begin{pmatrix} x \\ e_f \end{pmatrix}\{y - \lambda_f(x)\}, \tag{20}$$

$$H_f(\theta) = -\begin{pmatrix} x \\ e_f \end{pmatrix}\begin{pmatrix} x' & e_f' \end{pmatrix}\lambda_f(x), \tag{21}$$

and therefore

$$H_f(\theta) + G_f(\theta)G_f(\theta)' = \begin{pmatrix} x \\ e_f \end{pmatrix}\begin{pmatrix} x' & e_f' \end{pmatrix}\left\{(y - \lambda_f(x))^2 - \lambda_f(x)\right\}, \tag{22}$$

where e_f is a vector of zeros with a 1 in position f.

Substituting (22) in (18) and making the required simplifications I obtain the restrictions:

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[\mathrm{vech}(x\,x')\{\mathrm{E}_{\pi(\psi)}[Var_f[Y]] - \mathrm{E}_{\pi(\psi)}[\mathrm{E}_f[Y]]\}\right]\right] = 0, \tag{23}$$

where

$$\mathrm{E}_{\pi(\psi)}[\mathrm{E}_f[Y]] = \pi_1(\psi)\lambda_1(x) + \pi_2(\psi)\lambda_2(x), \tag{24}$$

and

$$\mathrm{E}_{\pi(\psi)}[Var_f[Y]] = \pi_1(\psi)(y - \lambda_1(x))^2 + \pi_2(\psi)(y - \lambda_2(x))^2 \tag{25}$$

and where $\mathrm{vech}(.)$ is the operator which stacks the functionally independent elements of a symmetrix matrix.

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[x\pi_f(\psi)\left\{(y - \lambda_f(x))^2 - \lambda_f(x)\right\}\right]\right] = 0, \text{ for } f = 1,2. \tag{26}$$

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[\pi_f(\psi)\left\{\left(y-\lambda_f(x)\right)^2-\lambda_f(x)\right\}\right]\right]=0, \quad \text{for } f=1,2. \quad (27)$$

The first set of restrictions (23) assesses the heteroscedasticity of an average second order residual, $\mathrm{E}_{\pi(\psi)}[Var_f[Y]]-\mathrm{E}_{\pi(\psi)}[\mathrm{E}_f[Y]]$, while the second set of restrictions in (26) and (27) assesses the heteroscedasticity of each type specific second order residual. In particular, the last set of restrictions demands that the posterior allocation of an observation to a type, $\pi_f(\psi)$, and the type specific second order residual, $\left(y-\lambda_f(x)\right)^2-\lambda_f(x)$, are uncorrelated.

Following an identical process of substituting (22) in (19) and simplifying leads to the following restrictions

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[x(\alpha_1-\alpha_2)\left(y-\frac{\lambda_1(x)\alpha_1-\lambda_2(x)\alpha_2}{\alpha_1-\alpha_2}\right)\right]\right]=0, \qquad (28)$$

$$\mathrm{E}_X\left[\mathrm{E}_{\theta,p}\left[\pi_f(\psi)\left(y-\lambda_f(x)\right)\right]\right]=0, \text{ for } f=1,2. \qquad (29)$$

where $\alpha_f=\dfrac{\pi_f(\psi)}{p_f}$ for $f=1,2.$

The first of this latter set of restrictions (28) requires that $y-\dfrac{\lambda_1(x)\alpha_1-\lambda_2(x)\alpha_2}{\alpha_1-\alpha_2}$ and $x(\alpha_1-\alpha_2)$ are uncorrelated. For this restriction to have any power, $\alpha_1-\alpha_2$ must be different from zero, which is equivalent to require that $\lambda_1(x)\neq\lambda_2(x)$ for all values of x, i.e. the two types must be different. The second set (29) requires that the posterior allocation to type, $\pi_f(\psi)$ is uncorrelated with the type-specific residual $\left(y-\lambda_f(x)\right)$.

The generalisation to a larger number of types seems straightforward, however how well would such a test perform in practice remains to be studied. In particular, it would be of interest to understand how such a test performs when the number of type is too small or too large.

5 References

Gouriéroux, C. , and Monfort, A. 1995. *Statistics and Econometric Models*. Cambridge: Cambridge University Press.

Jamshidian, M. and R.I. Jennrich. 1997. Accelaration of the EM algorithm by using quasi-Newton methods. *Journal of the Royal Statistical Society* B 59, no. 3: 569-587.

Lee, L.F. 2000. A numerically stable quadrature procedure. *Journal of Econometrics* 95, no. 1: 117-129.

Lindsay, B. 1983. The Geometry of Mixture Likelihood : A General Theory. *Annals of Statistics* 11, no. 1: 86-94.

Louis, T.A. 1982. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society* 44 : 226-233.

Oakes, D. 1999. Direct calculation of the information matrix via the EM algorithm. *Journal of the Royal Statistical Society* 61, no. 2: 479-482.

White, H. 1994. *Estimation, Inference and Specification Analysis*. Cambridge: Cambridge University Press.

KERP Keele Economics Research Papers — Recent Contributions

All papers in the KERP series are available for downloading from the Keele Economics website, via *www.keele.ac.uk/depts/ec/kerp*.

2002/11 *Dissipation In Rent-Seeking Contests With Entry Costs*
Richard Cornes and Roger Hartley

2002/10 *Advantageous or Disadv. Semi-collusion Licensing in a Vert. Separated Industry*
Arijit Mukherjee

2002/09 *Licensing in a Vertically Separated Industry*
Arijit Mukherjee

2002/08 *U-shaped Paths of Consumption and Phys. Capital in Lucas-type Growth Models*
Farhad Nili

2002/07 *On the Variance Covariance Matrix of the M.L. Estimator of a Discrete Mixture*
Gauthier Lanot

2002/06 *Monotonicity and the Roy Model*
Arnaud Chevalier and Gauthier Lanot

2002/05 *Capacity Commitment and Licensing*
Arijit Mukherjee

2002/04 *Household Credit and Saving: Does Policy Matter?*
Peter Lawrence

2002/03 *Innovation, Licensing and Welfare*
Arijit Mukherjee

2002/02 *Historical Schools of Economics: German and English*
Keith Tribe

2002/01 *R&D, Licensing and Patent Protection*
Arijit Mukherjee

2001/09 *Export and Direct Investment as a Signal in Global Markets*
Arijit Mukherjee and Udo Broll

2001/08 *The Welfare Effects of Quality Degradation with Network Externalities*
Jong-Hee Hahn

2001/07 *Cost Padding in Regulated Monopolies*
Spiros Bougheas and Tim Worrall

2001/06 *Is the Unskilled Worker Problem in Developing Countries Going Away?*
Ed Anderson

2001/05 *Does Society Prefer Small Innovation?*
Arijit Mukherjee

2001/04 *Bilateral Merger in a Leadership Structure*
Tarun Kabiraj and Arijit Mukherjee

2001/03 *Imitation, Patent Protection and Welfare*
Arijit Mukherjee and Enrico Pennings

2001/02 *R&D Organization and Technology Transfer*
Arijit Mukherjee and Sugata Marjit

2001/01 *International Joint Venture and the Technology of the Future*
Sugata Marjit, Arijit Mukherjee and Tarun Kabiraj

KERP Keele Economics Research Papers — Recent Contributions

All papers in the KERP series are available for downloading from the Keele Economics website, via *www.keele.ac.uk/depts/ec/kerp*.

2000/20 *Gift-Giving, Quasi-Credit and Reciprocity*
Tim Worrall and Jonathan P Thomas

2000/19 *Land Rents and Competitive Equilibrium*
Martin E. Diedrich

2000/18 *Monopoly Quality Differentiation with Top-quality Dependent Fixed Costs*
Jong-Hee Hahn

2000/17 *Time Consistency and Intergenerational Risk Sharing*
Tim Worrall

2000/16 *The Maximum Interest Rate on an Unbalanced Growth Path*
Martin E. Diedrich

2000/15 *Nonlinear Pricing of a Telecomm. Service with Call and Network Externalities*
Jong-Hee Hahn

2000/14 *Rent- seeking by Players with Constant Absolute Risk Aversion*
Richard Cornes and Roger Hartley

2000/13 *Differential Interest Rates on Unbalanced Growth Paths*
Martin E. Diedrich

2000/12 *Functional Quality Degradation of Software with Network Externalities*
Jong-Hee Hahn

2000/11 *Network Competition and Interconnection with Heterogeneous Subscribers*
Jong-Hee Hahn

2000/10 *On Monetary Policy Implications of Credit Rationing under Asymmetric Information*
Frédérique Bracoud

2000/09 *Cost of Regulation in Education: Do School Inspections Improve School Quality?*
Leslie Rosenthal

2000/08 *Intertemporal Substitution and Gambling for Long-Lived Agents*
Roger Hartley and Lisa Farrell

2000/07 *Joint Production Games and Share Functions*
Richard Cornes and Roger Hartley

2000/06 *The Value of Secondary School Quality in England*
Leslie Rosenthal

2000/05 *The Effects of Pollution and Energy Taxes across the European Income Distribution*
Elizabeth Symons, Stefan Speck and John Proops

2000/04 *Heterogeneous Demand Responses to Discrete Price Changes*
Roger Hartley and Gauthier Lanot

2000/03 *The Long-Run Labour Market Consequences of Teenage Motherhood*
Arnaud Chevalier and Tarja K. Viitanen

2000/02 *Economic Societies in Great Britain and Ireland before 1902*
Keith Tribe

2000/01 *Financial Transfers and Educational Achievement*
Arnaud Chevalier and Gauthier Lanot

ISSN 1352-8955

Department of Economics
Keele University
Keele, Staffordshire ST5 5BG
United Kingdom

tel: (44) 1782 583091
fax: (44) 1782 717577
email: economics@keele.ac.uk
web: www.keele.ac.uk/depts/ec/web/