Testing the number of components in finite mixture model with normal panel regression

IAAE Meeting 2019

Jasmine Hao and Hiro Kasahara[1]

Vancouver School of Economics

May 16, 2019

Table of contents

Background and motivation

Finite mixture model

- Finite mixture model: represent non-standard distribution density using the mixture of standard distributions ;
- Provide convenient way to model individual specific unobserved heterogeneity. Wide applications in economics area:
  - Labor economics: Keane and Wolpin [1997], Heckman and Singer [1984], Cameron and Heckman [1998];
  - Health economics: Deb and Trivedi [1997] ;
  - Industrial organization: Kamakura and Russell [1989], Andrews and Currim [2003];
  - Marketing: Dubé et al. [2010];
  - Dynamic discrete choice model: Arcidiacono and Miller [2011];

Motivation

- How to determine the number of components is not sufficiently discussed in past literature;
- Kasahara and Shimotsu [2012] propose a likelihood ratio test framework to test the number of components m under $H_0 : m = m_0$ against $H_1 : m = m_0 + 1$ under certain assumptions on mixture density functions;
- the normal panel regression model have the likelihood function that satisfies the assumptions.

Model: Finite mixture normal panel regression

Finite mixture normal panel regression model

Consider the panel regression:

- Data generating process: $y_{it} = \mu + x_{it}'\beta + z_{it}'\gamma + \epsilon_{it}\sigma$;
- observe $\{y_{it}, x_{it}, z_{it}\}_{t=1}^{T}$ for $i = 1, \ldots, N$, $t = 1, \ldots, T$;
- collect the observed data for firm i at time t to be $\omega_{it} = \{y_{it}, x_{it}, z_{it}\}$
- parameters: $\theta = (\mu, \beta', \sigma^2)'$, $\mu \in \mathbb{R}, \beta \in \mathbb{R}^q$, $\sigma \in \mathbb{R}_{++}$ and $\gamma \in\subset \mathbb{R}^q$;
- $\phi(t) = (2\pi)^{-1/2} \exp(-\frac{t^2}{2})$.

The joint density of panel data

$$f(\{\omega_{it}\}_{t=1}^{T}; \gamma, \theta) = \Pi_{t=1}^{T} \frac{1}{\sigma^j} \phi\left(\frac{y_{it} - \mu - x_{it}'\beta - z_{it}'\gamma}{\sigma}\right), \quad (1)$$

Mixture model

Finite mixture panel regression model:

- m unobserved class, each of which is characterized by θ^j;
- the mixing probability α^j for type $j = 1, \ldots, m$;
- $\sum_{j=1}^m \alpha^j = 1$ and $\alpha^j \in (0,1)$ for each type;
- collect the parameters in m-component mixture model and define $\vartheta_m = (\theta^1, \theta^2, \ldots, \theta^m, \alpha^1, \ldots, \alpha^{m-1}, \gamma)$.

Density function of finite mixture normal panel regression model:

$$f_m(\{\omega_{it}\}_{t=1}^T, \vartheta_m) = \sum_{j=1}^m \alpha^j f(\{\omega_{it}\}_{t=1}^T; \gamma, \theta^j). \tag{2}$$

Likelihood Ratio Test(LRT)

$H_0 : 1$ against $H_0 : 2$: Why traditional approach fails?

Usually, approximate likelihood ratio test statistics by quadratic expansion. The density function of two-component model is

$$f_2(\{\omega_{it}\}_{t=1}^{T}, \vartheta_2) = \alpha f(\{\omega_{it}\}_{t=1}^{T}; \gamma, \theta^1) + (1-\alpha) f(\{\omega_{it}\}_{t=1}^{T}; \gamma, \theta^2). \quad (3)$$

- Suppose null model true, then $\theta^1 = \theta^2$;
- α is not identified: $\nabla_\alpha \log f_2(\{\omega_{it}\}_{t=1}^{T};; \vartheta_2) = 0$;
- first order conditions w.r.t. θ^1 and θ^2 are collinear;
- therefore the Fisher Information matrix is deficient in $1 + \dim(\theta)$ degrees.

Reparameterization

Reparameterize θ^1 and θ^2 as $\psi_\alpha := (\gamma', \nu', \lambda')$:

$$\begin{pmatrix} \lambda \\ \nu \end{pmatrix} := \begin{pmatrix} \theta^1 - \theta^2 \\ \alpha\theta^1 + (1-\alpha)\theta^2 \end{pmatrix} \text{ so that } \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \nu + (1-\alpha)\lambda \\ \nu - \alpha\lambda \end{pmatrix}. \quad (4)$$

Redefine the density:

$$\begin{aligned} g(\{\omega_{it}\}_{t=1}^T; \psi_\alpha, \alpha) &= \alpha f(\{\omega_{it}\}_{t=1}^T; \gamma, \nu + (1-\alpha)\lambda) \\ &\quad + (1-\alpha) f(\{\omega_{it}\}_{t=1}^T; \gamma, \nu - \alpha\lambda); \\ l(\{\omega_{it}\}_{t=1}^T; \psi_\alpha, \alpha) &= \log g(\{\omega_{it}\}_{t=1}^T; \psi_\alpha, \alpha), \end{aligned}$$

Score function

$\nabla_\lambda l(\{\omega_{it}\}_{t=1}^{T};\psi_\alpha^*,\alpha)=0$, use second order derivative instead. Score function defined as:

$$
\begin{aligned}
s_i &= \begin{pmatrix} \nabla_\gamma l(\{\omega_{it}\}_{t=1}^{T};\psi_\alpha^*,\alpha) \\ \nabla_\nu l(\{\omega_{it}\}_{t=1}^{T};\psi_\alpha^*,\alpha) \\ \frac{1}{(1-\alpha)\alpha}\nabla_{\lambda\lambda'} l(\{\omega_{it}\}_{t=1}^{T};\psi_\alpha^*,\alpha) \end{pmatrix} \\
&= \begin{pmatrix} \nabla_\gamma f(\{\omega_{it}\}_{t=1}^{T};\gamma^*,\theta^*) \\ \nabla_\theta f(\{\omega_{it}\}_{t=1}^{T};\gamma^*,\theta^*), \\ \nabla_{\theta\theta'} f(\{\omega_{it}\}_{t=1}^{T};\gamma^*,\theta^*) \end{pmatrix} / f(\{\omega_{it}\}_{t=1}^{T};\gamma^*,\theta^*).
\end{aligned} \tag{5}
$$

Fourth order expansion of LRT statistics

- $L_n(\psi_\alpha, \alpha) := \sum_{i=1}^{n} l(\{\omega_{it}\}_{t=1}^{T}; \psi_\alpha, \alpha)$;
- $S_n := n^{-1/2} \sum_{i=1}^{N} s_i$, $S_n \to_d S \sim N(0, \mathcal{I})$;
- $\mathcal{I}_n := \frac{1}{n} \sum_{i=1}^{N} s_i s_i'$,, $\mathcal{I}_n \to_p \mathcal{I}$;
- $t_n(\psi_\alpha, \alpha) := \begin{pmatrix} n^{1/2}(\gamma - \gamma^*) \\ n^{1/2}(\nu - \theta^*) \\ n^{1/2}\alpha(1-\alpha)v(\lambda) \end{pmatrix}$.

$$\begin{aligned} & L_n(\psi_\alpha, \alpha) - L_n(\psi_\alpha^*, \alpha) \\ & = t_n(\psi_\alpha, \alpha) S_n - \frac{1}{2} t_n(\psi_\alpha, \alpha)' \mathcal{I}_n t_n(\psi_\alpha, \alpha) + R_n(\psi_\alpha, \alpha) \end{aligned} \tag{6}$$

where $R_n(\psi_\alpha, \alpha) \to_p 0$.

Asymptotic distribution of LRT

- $\hat{\psi}_\alpha = \arg\max_{\psi_\alpha} L_n(\psi_\alpha, \alpha)$;
- $\hat{\phi}_0 = (\hat{\gamma}_0, \hat{\theta}_0) = \arg\max_\phi L_{0,n}(\phi)$;
- $L_{0,n}(\phi) = \sum_{i=1}^N \log f(\{\omega_{it}\}_{t=1}^T; \gamma, \theta)$;
- $LR_{n,1}(\epsilon) = \max_{\alpha \in [\epsilon, 1-\epsilon]} 2\{L_n(\hat{\psi}_\alpha, \alpha) - L_{n,0}(\hat{\phi}_0)\}$, where ϵ arbitrary;
- $\eta = (\gamma, \theta)$;
- partition $S = \begin{bmatrix} S_\eta \\ S_\lambda \end{bmatrix}$, then $\mathcal{I} = SS' = \begin{bmatrix} I_\eta & I_{\eta\lambda} \\ I_{\lambda\eta} & I_\lambda \end{bmatrix}$;
- $\mathcal{I}_{\lambda,\eta} = \mathcal{I}_{\lambda\lambda} - \mathcal{I}_{\lambda\eta}\mathcal{I}_\eta^{-1}\mathcal{I}_{\eta\lambda}$.

Assumptions on density

Assumption 1

(a) If $(\gamma, \theta) \neq (\gamma^*, \theta^*)$, then $f(\{\omega_{it}\}_{t=1}^{T}; \gamma, \theta) \neq f(\{\omega_{it}\}_{t=1}^{T}; \gamma^*, \theta^*)$; (b) Θ_θ and Θ_γ are compact; (c) $\log f(\{\omega_{it}\}_{t=1}^{T}; \gamma, \theta)$ is continuous at each $(\gamma, \theta) \in \Theta_\gamma \times \Theta_\theta$ with probability one; (d) $E[\sup_{(\gamma,\theta)\in\Theta_\gamma\times\Theta_\theta} |\log f(\{\omega_{it}\}_{t=1}^{T}; \gamma, \theta)|] < \infty$.

Assumptions on density

Assumption 2

(a) γ^* and θ^* are in the interior of Θ_γ and Θ_θ. (b) For every x, $f(\{\omega_{it}\}_{t=1}^T; \gamma, \theta)$ is four times continuous differentiable in a neighborhood of (γ^*, θ^*). (c) For $\alpha \in (0,1)$, $E \sup_{\psi_\alpha \in \mathcal{N}} \|\nabla^{(k)} l(\{\omega_{it}\}_{t=1}^T; \psi_\alpha, \alpha)\| < \infty$ for a neighborhood $\mathcal{N}$ of ψ^* and for $k = 1, \ldots, 4$, where $\nabla^{(k)}$ denotes the k-th derivative w.r.t. ψ. (d) For $\alpha \in (0,1)$, $E\|\frac{\nabla^{(k)} g(\{\omega_{it}\}_{t=1}^T; \psi_\alpha^*, \alpha)}{g(\{\omega_{it}\}_{t=1}^T; \psi_\alpha^*, \alpha)}\|^2 < \infty$ for $k = 1, 2, 3$.

Assumption 3

$\mathcal{I}$ is finite and positive definite.

Asymptotic distribution of LRT

Kasahara and Shimotsu [2012] shows under A1-A3,

- $2\{L_n(\hat{\psi}_\alpha, \alpha) - L_n(\psi^*_\alpha, \alpha)\} \rightarrow_d \hat{t}'_\lambda \mathcal{I}_{\lambda,\eta} \hat{t}_\lambda + S'_\eta \mathcal{I}^{-1} S_\eta$;
- $2\{L_n(\hat{\psi}_\alpha, \alpha) - L_{0,n}(\hat{\gamma}_0, \hat{\theta}_0)\} \rightarrow_d \hat{t}'_\lambda \mathcal{I}_{\lambda,\eta} \hat{t}_\lambda$;
- $LR_{n,1}(\epsilon_1) \rightarrow_d \hat{t}'_\lambda \mathcal{I}_{\lambda,\eta} \hat{t}_\lambda$;
- $\hat{t}'_\lambda$ is not standard distribution, is a projection of $W_{\lambda,\eta} \sim N(0, \mathcal{I}_{\lambda,\eta})$ on to a cone;
- simulate and get $\hat{t}'_\lambda$ and $\hat{t}'_\lambda \mathcal{I}_{\lambda,\eta} \hat{t}_\lambda$.

Modified EM

Unbounded likelihood:

- H_0 true, when estimate H_0;
- $\alpha \to 0$, $\hat{\sigma^1} \to 0$, $L_n(\psi_\alpha, \alpha) \to \infty$.

Need to bound $\hat{\sigma}^1$ or $\hat{\sigma}^2$ from 0.

E-step, calculate the weight matrix for each firm:

- $w_{i1}^k = \frac{\alpha^{(k)} f(\{\omega_{it}\}_{t=1}^T; \gamma^{(k)}, \theta^{1(k)})}{f_2(\{\omega_{it}\}_{t=1}^T; \vartheta_2^{(k)})}$;
- $w_{i2}^k = \frac{(1-\alpha^{(k)}) f(\{\omega_{it}\}_{t=1}^T; \gamma^{(k)}, \theta^{1(k)})}{f_2(\{\omega_{it}\}_{t=1}^T; \vartheta_2^{(k)})}$;

Modified EM

M-step, update parameters:

$$\gamma^{(k+1)} = \left(\sum_{i=1}^{N}\sum_{t=1}^{N} z_{it} z_{it}'\right)^{-1} \left(\sum_{i=1}^{N}\sum_{t=1}^{N} z_{it}\left(y_{it} - \sum_{j=1}^{2} w_{ij}^{(k)} \tilde{x}_{it}' \beta^{j(k)}\right)\right);$$

$$\begin{pmatrix}\mu^{j(k+1)} \\ \beta^{j(k+1)}\end{pmatrix} = \left(\sum_{i=1}^{N} w_{ij}^{(k)} \sum_{t=1}^{N} \tilde{x}_{it}\tilde{x}_{it}'\right)^{-1} \left(\sum_{i=1}^{N} w_{ij}^{(k)} \sum_{t=1}^{N} \tilde{x}_{it}\left(y_{it} - z_{it}'\gamma^{(k+1)}\right)\right), \; j = 1, 2; \tilde{x}_{it} = (1, x_{it}')';$$

$$\sigma^{(k+1)} = \left(\frac{\sum_{i=\mathbf{1}}^{N} w_{ij}^{(k)} \sum_{t=\mathbf{1}}^{N} (y_{it} - \mu^{j(k+\mathbf{1})} - z_{it}'\gamma^{(k+\mathbf{1})} - x_{it}'\beta^{j(k+\mathbf{1})})^{\mathbf{2}} + 2a_n\hat{\sigma}_{\mathbf{0}}^{\mathbf{2}}}{T\sum_{i=\mathbf{1}}^{N} w_{ij}^{(k)} + 2a_n}\right)^{1/2};$$

$$\alpha^{(k+1)} = \frac{\sum_{i=\mathbf{1}}^{N} w_{ij}^{(k)}}{N}. \; a_n = a_n(N, T, misclassification; m);$$

Generalizing the test

- Testing the general m_0-component model against a (m_0+1)-component model is an easy extension of 1-component model against a 2-component model.
- can express m_0-component model as a (m_0+1)-component model;
- estimate ϑ_{m_0+1} by using $\tau \in (\epsilon, 1-\epsilon)$ to split one of the mixing probability in ϑ_{m_0};
- do the modified EM for each of the m_0 components;
- $LR_{n,1}^{m_0}(\epsilon) \rightarrow_d \max\{(\hat{t}_\lambda^1)' \mathcal{I}_{\eta,\lambda}^1 (\hat{t}_\lambda^1), \ldots, (\hat{t}_\lambda^{m_0})' \mathcal{I}_{\eta,\lambda}^{m_0} (\hat{t}_\lambda^{m_0})\}$.

Simulation result

Sizes

Table: Sizes(in %) in modified EM size test of $H_0 : m_0 = 2$ against $H_A : m_0 = 3$ at 5% level

	N = 100			N = 500		
Model	T = 2	T = 5	T = 10	T = 2	T = 5	T = 10
(A, C)	5.25	5.4	4.2	4.55	4.5	4.0
(A, D)	2.6	3.55	3.4	3.4	3.15	4.3
(B, C)	4.05	5.75	3.5	4.35	4.85	4.1
(B, D)	2.45	3.25	3.4	4.5	3.95	5.25

Use A, B to denote $(\alpha_1, \alpha_2) = (0.5, 0.5)$ and $(0.2, 0.8)$; use C, D to denote $(\mu_1, \mu_2) = (-1, 1)$ and $(-0.5, 0.5)$; and set the variance $(\sigma_1, \sigma_2) = (0.8, 1.2)$.

Powers

Table: Powers (in %) of modified EM test of $H_0 : m_0 = 2$ against $H_A : m_0 = 3$ at 5% level

α	A				B			
	N = 100		N = 500		N = 100		N = 500	
(μ, σ)	T=2	T=5	T=2	T=5	T=2	T=5	T=2	T=5
(C, G)	8.8	64.8	18.2	100	8.2	74.0	26.6	100
(C, H)	56.4	100	100	100	40.6	99.8	99.8	100
(C, I)	73.2	100	100	100	84.6	100	100	100
(D, G)	34.8	100	98.0	100	40.0	100	99.8	100
(D, H)	93.4	100	100	100	84.0	100	100	100
(D, I)	99.8	100	100	100	100	100	1.0	1.0

A and B refers to $(\alpha_1, \alpha_2, \alpha_3) = (1/3, 1/3, 1/3)$ and $(1/4, 1/2, 1/4)$, respectively; C, D refers to $(\mu_1, \mu_2, \mu_3) = (-1, 0, 1), (-1.5, 0, 1.5)$; G, H, I refers to $(\sigma_1, \sigma_2, \sigma_3) = (1, 1, 1), (0.6, 1.2, 0.6), (0.6, 0.6, 1.2)$.

Application

Application: Identification of the flexible input share

Cobb-Douglas production function implied:

$$y_{it} = \beta_0 + \beta_m m_{it} + \beta_k k_{it} + \beta_l l_{it} + \epsilon_{it}, \quad (7)$$

Firms observe input and output prices and choose material input flexibly each time period:

$$M_{it} = \arg\max_{M} P_{Y,t} Y_{it} - P_{M,t} M. \quad (8)$$

The model implies

$$s_{it} = \log \frac{P_{M,t} M_{it}}{P_{Y,t} Y_{it}} = \log \beta_m + \epsilon_{it}. \quad (9)$$

$(\beta_m, \beta_k, \beta_l)$ are assumed to be same across firms within a narrowly defined industry.

Application

- Gandhi et al. [2013] found s_{it} is heterogeneous across firms and persistent through time for one firm.
- Kasahara et al. [2015] non-parametrically identify production function with heterogeneity using finite mixture model.
- The data of Japanese producer is compiled by the Development Bank of Japan (DBJ). This dataset contains detailed corporate balance sheet/income statement data from 1980 to 2008 for the firms listed on the Tokyo Stock Exchange.

Empirical results 1

Table: Estimated Likelihood Ratio for Japanese producer in Machine industry

Time	$m_0 = 1$	$m_0 = 2$	$m_0 = 3$	$m_0 = 4$	$m_0 = 5$
1	93.309***	19.851***	6.849*	0.163	0.612
2	297.487***	111.438***	90.417***	37.875***	32.410***
3	524.024***	205.256***	121.135***	78.498***	59.610***
4	732.667***	320.623***	159.470***	124.478***	89.004***
5	934.599***	419.310***	214.137***	156.153***	126.073***

The estimation is based on Machinery industry, with null model of $m_0 = 1, 2, 3, 4, 5$, respectively.
For $T = 1$, I use data from the latest year 2008. For $T = 2$, I use data from 2007-2008. For $T = 3$, I use data from 2006-2008. For $T = 4$, I use data from 2005-2008. For $T = 5$, I use data from 2004-2008.
* indicate the result is significant at 10% level.
** indicate the result is significant at 5% level
*** indicate the result is significant at 1% level

Empirical results 2

Table: Estimated LR for Japanese producer(T = 1)

Industry	$H_0 : m_0 = 1$	$H_0 : m_0 = 2$	$H_0 : m_0 = 3$	$H_0 : m_0 = 4$	$H_0 : m_0 = 5$
Chemical	39.119***	4.971	0.071	1.572	1.539
Ceramics	1.898	0.010	0.142	0.203	0.223
Other	29.802***	2.267	5.488	0.905	0.542
Food	40.396***	11.700	0.387	3.306	3.531
Othermetal	31.324***	0.411	0.079	22.547***	22.547***
Textile	15.857***	1.185	0.010	1.455	0.639
Paper	5.485	0.051	1.913	0.088	0.163
Steel	15.221***	0.007	0.252	0.100	0.278
Electronics	122.642***	24.106	3.579	0.055	0.162
Transportation equipment	53.170***	19.220	0.481	0.068	0.110
Precision instrument	1.314	0.571	0.074	0.134	42.210
Metal product	9.373	2.756	0.064	1.125	0.214
Plastic	0.036	0.019	0.040	0.396	5.454

The estimation is based on revenue share of intermediate material.
* indicate the result is significant at 10% level.
** indicate the result is significant at 5% level
*** indicate the result is significant at 1% level

Empirical results 2

Table: Estimated LR for Japanese producer(T = 2)

Industry	$H_0 : m_0 = 1$	$H_0 : m_0 = 2$	$H_0 : m_0 = 3$	$H_0 : m_0 = 4$	$H_0 : m_0 = 5$
Chemical	226.387***	95.997***	95.445***	61.846***	33.470***
Ceramics	47.393***	19.939***	19.938***	9.257**	7.714
Other	118.481***	39.467***	29.780***	21.311***	6.454
Food	138.022***	96.038***	60.252***	42.777***	45.893***
Othermetal	21.011***	16.365***	7.532	5.454	1.505
Textile	65.113***	38.819***	22.184***	28.098***	14.090
Paper	40.466***	25.543	15.532	17.877***	17.901**
Steel	56.517***	12.050***	7.961*	2.776	4.742
Electronics	380.111***	146.134***	60.844***	48.359***	48.386***
Transportation equipment	171.980***	72.937***	51.851***	43.700***	41.408***
Precision instrument	27.780***	17.896***	13.987	10.530	7.699
Metal product	60.267***	32.172***	25.641***	19.194***	14.914***
Plastic	8.578**	19.782	8.568	9.179	2.617

The estimation is based on revenue share of intermediate material.
* indicate the result is significant at 10% level.
** indicate the result is significant at 5% level
*** indicate the result is significant at 1% level


tenki
takk
спасибо
kam sah hamnida
dank je
hvala
дзякуй
gracias
dziękuję
bedankt
תודה
dhanyavadagalu
tack
blagodaram
mèsi
xièxie
tanemirt
rahmet
enkosi
mochchakkeram
trugarez
arigatô
danke
kop khun krap
faafetai lava
shukriya
ありがとう
kia ora
dankon
děkuji
manana
diolch
akun
grazzi
mamnun
gràçie
kiitos
spas
ačiū
thank you
barka
शुक्रिया
tau
dankie
tapadh leat
chnorakaloutioun
dhanyavad
gratias ago
gràcies
bayarlalaa
teşekkür ederim
obrigada
σας ευχαριστώ
sagolun
murakoze
taiku
mahalo
didi madloba
sukriya
obrigado
chokrane
rahmat
dakujem
misaotra
welalin
terima kasih
mercé
najis tuke
شكرا
köszönöm
asante
grazie
nandri
謝謝
mersi
sobodi
merci
nanni
vinaka
mauruuru
matondo
cảm ơn bạn
go raibh maith agat
paldies
ngiyabonga

Rick L Andrews and Imran S Currim. Retention of latent segments in regression-based marketing models. *International Journal of Research in Marketing*, 20(4):315–321, 2003. ISSN 01678116. doi: 10.1016/j.ijresmar.2003.04.001.

Peter Arcidiacono and Robert A. Miller. Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica*, 79(6):1823–1867, 11 2011. ISSN 0012-9682. doi: 10.3982/ECTA7743. URL http://doi.wiley.com/10.3982/ECTA7743.

Stephen V. Cameron and James J. Heckman. Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males. *Journal of Political Economy*, 106(2):262–333, 4 1998. ISSN 0022-3808. doi: 10.1086/250010. URL http://www.journals.uchicago.edu/doi/10.1086/250010.

Partha Deb and Pravin K Trivedi. Demand for medical care by the elderly: a finite mixture approach. *Journal of Applied Econometrics*, 12(3): 313–336, 1997. ISSN 08837252. doi: 10.1002/(SICI)1099-1255(199705)12:3<313:: AID-JAE440>3.0.CO;2-G. URL http://www.jstor.org/stable/2285252http://www.jstor.org/http://www.jstor.org/action/showPublisher?publisherCode=jwiley.http://www.jstor.orghttp://doi.wiley.com/10.1002/(SICI)1099-1255(199705)12:3%3C313::AID-JAE440%3E3.0.CO;2-G.

Jean-Pierre Dubé, Günter J. Hitsch, and Peter E. Rossi. State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, 41(3):417–445, 8 2010. ISSN 07416261. doi: 10.1111/j.1756-2171.2010.00106.x. URL http://doi.wiley.com/10.1111/j.1756-2171.2010.00106.x.

Amit Gandhi, Salvador Navarro, and David Rivers. On the identification of production functions: How heterogeneous is productivity? 2013. URL http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1071&context=economicscibc.

J. Heckman and B. Singer. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320, 3 1984. ISSN 0012-9682. doi: 10.2307/1911491.

Wagner Kamakura and Gary Russell. A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26(4):379–390, 1989. ISSN 00222437. doi: 10.2307/3172759. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2428442http://www.jstor.org/stable/3172759.

Hiroyuki Kasahara and Katsumi Shimotsu. Testing the number of components in finite mixture models. 2012.

Hiroyuki Kasahara, Paul Schrimpf, and Michio Suzuki. Identification and estimation of production function with unobserved heterogeneity. 2015. URL http://faculty.arts.ubc.ca/pschrimpf/research/IdentificationAndEstimationOfProduct_preview.pdf.

Michael P. Keane and Kenneth I. Wolpin. The career decisions of young men. *Journal of Political Economy*, 105(3):473–522, 1997. ISSN 00223808. doi: 10.1086/262080. URL http://www.journals.uchicago.edu/doi/pdfplus/10.1086/262080.