



# *University of* **HUDDERSFIELD**

## **University of Huddersfield Repository**

Johnes, Jill and Johnes, Geraint

Costs and efficiency in English higher education: An analysis using latent class stochastic frontier models

### **Original Citation**

Johnes, Jill and Johnes, Geraint (2015) Costs and efficiency in English higher education: An analysis using latent class stochastic frontier models. In: 5th Workshop on Efficiency and Productivity Analysis, 5 October 2015, Catholic University Porto, Portugal. (Unpublished)

This version is available at <https://eprints.hud.ac.uk/id/eprint/27830/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>

# COSTS AND EFFICIENCY IN ENGLISH HIGHER EDUCATION

University of  
**HUDDERSFIELD**  
Inspiring tomorrow's professionals

## AN ANALYSIS USING LATENT CLASS STOCHASTIC FRONTIER MODELS

*5th Workshop on Efficiency and Productivity Analysis*

Porto 5<sup>th</sup> October 2015



**Jill Johnes**, University of Huddersfield UK

**Geraint Johnes**, Lancaster University UK



Inspiring tomorrow's professionals



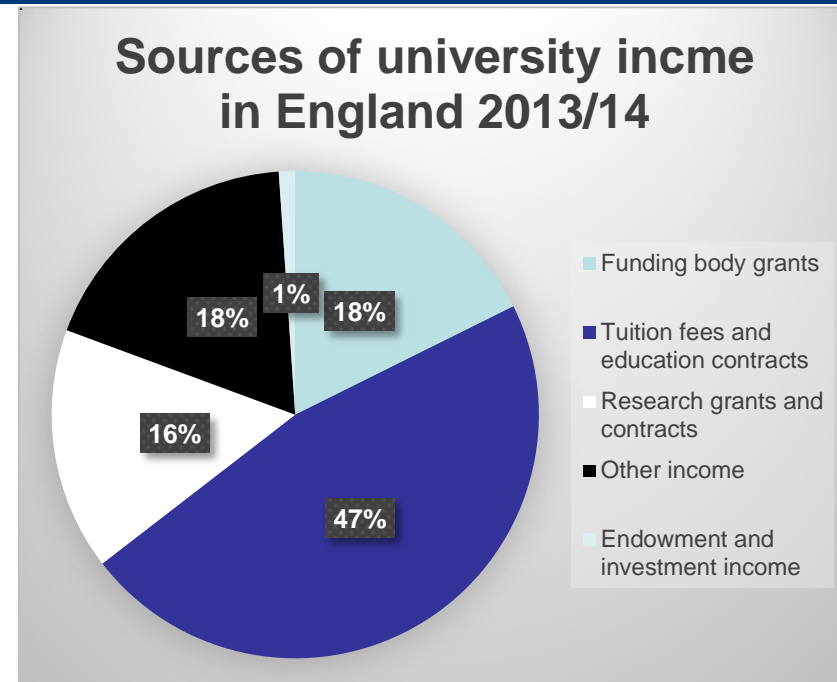
# Outline of talk

- 1. Introduction**
- 2. Literature review**
- 3. Conceptual issues**
- 4. Model specification**
  - Defining the variables
  - Estimation method
- 5. Results**
  - Estimated average costs
  - Economies of scale
  - Economies of scope
  - Efficiencies
- 6. Conclusions**

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/237411/bis-13-918-efficiency-in-higher-education-sector.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/237411/bis-13-918-efficiency-in-higher-education-sector.pdf)

# 1. Introduction

- HEIs receive public money
  - funding body grants
  - non-payment of tuition fees
- Reduced incentive to be efficient
- Need to assess efficiency of higher education institutions (HEIs)
- Cost functions provide information on efficiency, economies of scale and economies of scope



# 1. Introduction

The English higher education sector comprises very diverse groups of HEIs:

- ✓ Pre-1992 universities: degree programmes in all academic subjects; research mission
- ✓ Post-1992 universities: degree programmes in academic and vocational subjects; some have a research mission
- ✓ Former colleges of HE: small, specialist HEIs; most do not have a research mission

# 1. Introduction

## Questions

- What are average and marginal costs of outputs of English HEIs?
- Are there economies of scale and scope in English HE?
- How efficient are English HEIs?
- How does 'mission group' affect costs?
- Are there other factors which might affect HEIs' costs?

## 2. Literature Review

- USA: Cohn *et al* (1989)
- UK: Glass *et al* (1995a; 1995b); Johnes (1996; 1997; 1998); Izadi *et al* (2002); Stevens (2005); Johnes *et al* (2005; 2008); Thanassoulis *et al* (2011)
  - ✓ Allow for economies of scale and scope
  - ✓ Disaggregate output by subject and by type of HEI
  - ✓ Limited analysis of additional variables
  - ✓ Increasing use of panel data
  - ✓ Generally cover a subset of the English HE sector

## 2. Literature Review

- USA: Cohn *et al* (1989)
- UK: Glass *et al* (1995a; 1995b); Johnes (1996; 1997; 1998); Izadi *et al* (2002); Stevens (2005); Johnes *et al* (2005; 2008); Thanassoulis *et al* (2011)
  - ✓ Relatively low efficiency in panel data studies
  - ✓ Efficiency varies by type of university
  - ✓ Ray economies of scale; diseconomies of scope
  - ✓ Student quality, location of HEI are not important determinants of costs



## 2. Literature Review

### Most recent developments

- UK: Johnes & Johnes (2009) use a random parameter model (RPM) with stochastic frontier analysis (SFA)
  - ✓ Allows HEIs to have different objectives; the model allows the coefficient on one output *to vary by HEI*
  - ✓ Can be difficult to estimate the parameters of a RPM SFA
- Findings:
  - ✓ HEIs are heterogeneous in terms of both cost structure and efficiency


## 2. Literature Review

### Most recent developments


- USA: Agasisti & Johnes (2009) use latent class model (LCM) with SFA
  - ✓ Rather than calculating cost functions by pre-defined groups, they use the LCM method *to let the data* suggest distinct groups
- Findings:
  - ✓ HEIs are heterogeneous in terms of both cost structure and efficiency

# 3. Conceptual Issues

## Functional form of cost function

a) Linear:  $C = \alpha_0 + \sum_i \beta_i y_i$  

b) CES:  $C = \alpha_0 + \left[ \sum_i \beta_i y_i^{\delta_i} \right]^{\rho} + v$

c) Quadratic:  $C = \alpha_0 + \sum_i \alpha_i F_i + \sum_i \beta_i y_i + \left( \frac{1}{2} \right) \sum_i \sum_j \gamma_{ij} y_i y_j + v$  

d) Hybrid translog:

$$\ln C = \alpha_0 + \sum_i \alpha_i \ln(w_i) + \sum_k \beta_k [(y_k^{\theta} - 1) / \theta] + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \ln w_i \ln w_j +$$
$$\frac{1}{2} \sum_k \sum_l \delta_{kl} [(y_k^{\theta} - 1) / \theta] [(y_l^{\theta} - 1) / \theta] + \sum_i \sum_k \rho_{ik} \ln w_i [(y_k^{\theta} - 1) / \theta] + v$$

### 3. Conceptual Issues

- **Average incremental cost (AIC)**

$$AIC(y_i) = [C(y) - C(y_{N-i})]/y_i$$

- **Ray economies of scale**

$$S_R = \frac{C(y)}{\sum_i y_i C_i(y)}$$

where  $C_i(y) = \frac{\partial C(y)}{\partial y_i} = MC_i$

- ✓ If  $S_R > 1$  ( $< 1$ ) then there are economies (diseconomies) of scale

Note that  $C(y)$  is the total cost of producing all  $N$  outputs.

### 3. Conceptual Issues

- **Product-specific economies of scale**

$$S_i(y) = AIC(y_i)/C_i(y)$$

- ✓ If  $S_i > 1$  ( $< 1$ ) then there are economies (diseconomies) of scale for product  $i$

- **Economies of scope**

$$S_G = [\sum_i C(y_i) - C(y)]/C(y)$$

- ✓ If  $S_G > 0$  ( $< 0$ ) then global economies (diseconomies) of scope exist for producing the outputs jointly rather than in separate institutions

# 4. Model Specification

## a) Outputs

### TEACHING

- **UGMED** FTE undergraduates in medicine and dentistry (000s)
- **UGSCI** FTE undergraduates in sciences other than medicine and dentistry (000s)
- **UGARTS** FTE undergraduates in non-science subjects (000s)
- **PG** FTE postgraduates in all subjects (000s)

## 4. Model Specification

### a) Outputs

#### RESEARCH

- **RESEARCH** Quality related funding and research grants

#### THIRD MISSION

- **IPINCOME** Income from third mission activity

**Note that** all squares and interactions of UGMED, UGSCI, UGARTS, PG and RESEARCH are included; the square of IPINCOME and interaction of IPINCOME only with RESEARCH are included.

## 4. Model Specification

### b) Additional factors

#### QUALITY OF STUDENTS

- **MEANSAL** Mean salary of graduates 6 months after graduation

#### QUALITY OF TEACHING

- **NSS** Percentage saying yes to the question: 'Overall, I am satisfied with the quality of the course' from the National Student Survey



## 4. Model Specification

### b) Additional factors

#### **WIDENING PARTICIPATION**

- **LOWPNO** Number of FT UG entrants from 'low participation' neighbourhoods

#### **ESTATES COSTS**

- **LISTED** The total area of the HEI identified as a listed building

## 4. Model Specification

### b) Additional factors

#### DUMMY VARIABLES

- **OXBRIDGE** Dummy variable: 1 if HEI is Oxford or Cambridge
- **YEAR** Dummy for each year in the study (apart from the last)

# 4. Model Specification

- SFA

For HEI  $i$  at time  $t$ :

$$C_{it} = f(y_{1it}, \dots, y_{kit}) + v_{it} + u_{it}$$

- SFA with latent class model (LCM)

For HEI  $i$  at time  $t$ ,  $m$  classes:

$$C_{it} = f_m(y_{1it}, \dots, y_{kit}) + v_{it,m} + u_{it,m}$$

# 5. Results

- Panel data from 2003/04 to 2010/11 covering around 120 HEIs
- Model estimates for 3 time periods: 2003/04 to 2004/05, 2005/06 to 2007/08 and 2008/09 to 2010/11
- Comparison of results from applying SFA and SFA LCM
- Efficiency is allowed to vary over time within any given model
- Data are largely from the Higher Education Statistics Agency
- All money units are in 2011 values

## 5. Results

### AIC from SFA linear model (2011 £)

AICs	2008/09 to 2010/11	2005/06 to 2007/08	2003/04 to 2004/05
UGMED	13484	13866	9748
UGSCI	7775	7040	5609
UGARTS	4574	6657	3951
PG	13953	9409	9818

Other outputs included: RESEARCH, IPINCOME  
Controls for: LISTED, LOWPNO, YEAR dummies,  
OXBRIDGE

## 5. Results

### AIC from the linear SFA LCM (2011 £)

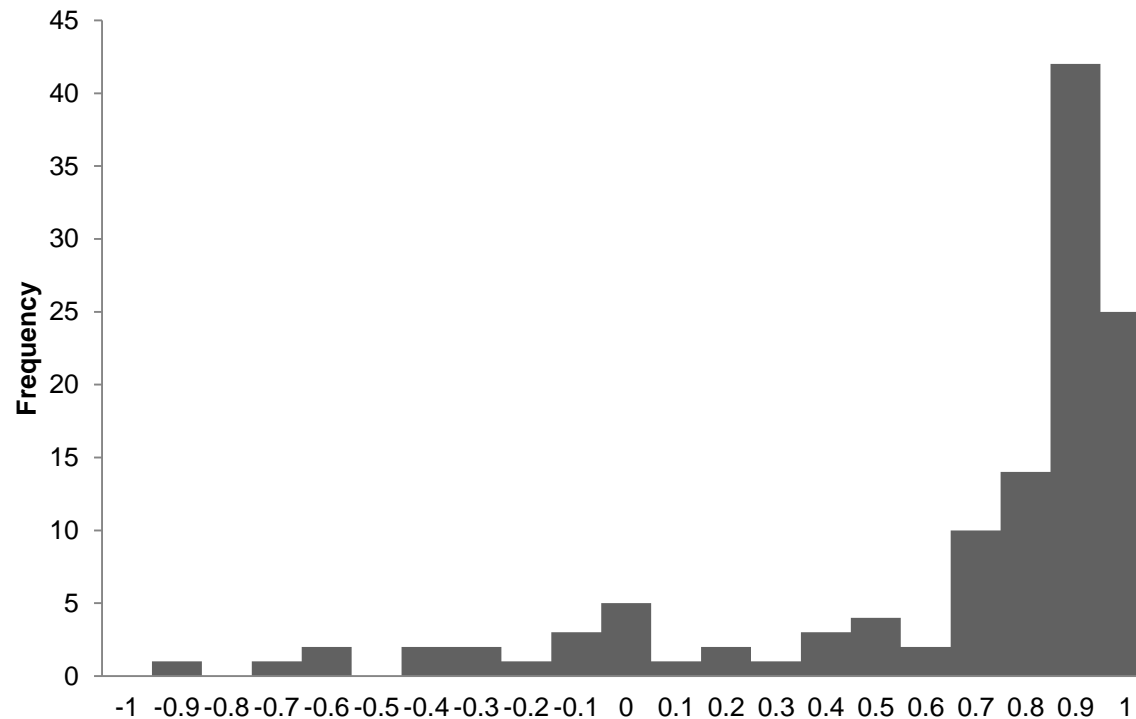
	2008/09 to 2010/11		2005/06 to 2007/08		2003/04 to 2004/05	
AICs	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2
UGMED	10865	7774	9732	6623	2406	9446
UGSCI	1931	8472	1748	8641	2538	7055
UGARTS	9353	2757	8166	4659	6502	4427
PG	246	18694	10459	5754	13432	8614
No. in each class	121	234	111	216	60	136

Controls for: LISTED, LOWPNO, YEAR dummies,  
OXBRIDGE

# 5. Results

## Histogram of efficiency scores

**Final year of linear 2008/09 to 2010/11 model**

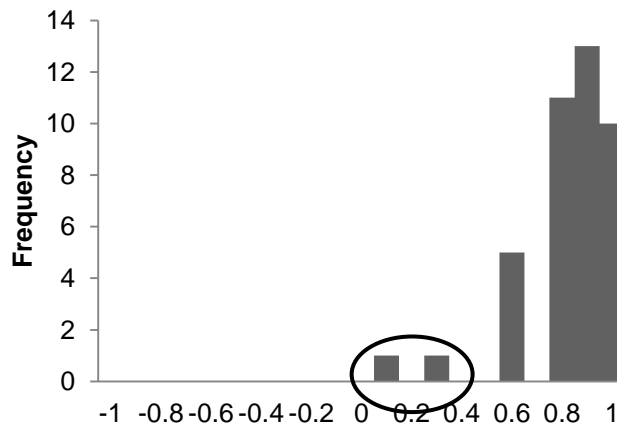


# 5. Results

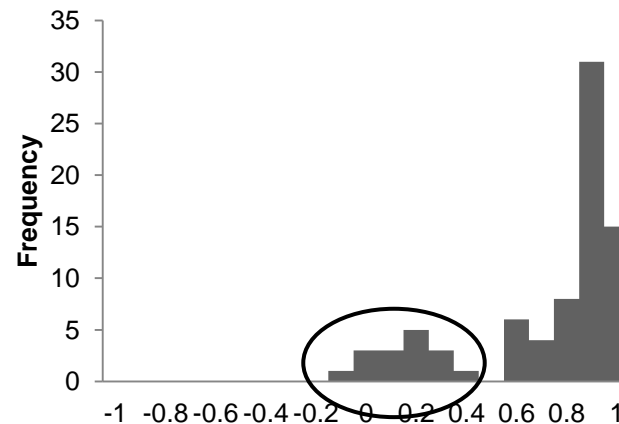
## Histogram of efficiency scores

**Final year of 2008/09 to 2010/11 linear latent class model**

**Latent class 1**



**Latent class 2**





## 5. Results

# Akaike Information Criterion (AkIC)

$$AkIC = -2.\log LF(m) + 2.k$$

where k is the number of estimated parameters

No. of classes	2008/09 to 2010/11	2005/06 to 2007/08	2003/04 to 2004/05
1	8393.3	7574.0	4356.5
2	7711.9	7019.1	4119.4
3	7637.9	6989.6	4081.7
4	7561.9	6921.7	4037.7

## 5. Results

### AIC from SFA quadratic model (2011 £)

AICs	2008/09 to 2010/11	2005/06 to 2007/08	2003/04 to 2004/05
UGMED	16034	15000	9195
UGSCI	7858	9444	4591
UGARTS	5459	4587	329
PG	5275	2601	7073

Other outputs included: RESEARCH, IPINCOME  
Controls for: LISTED, LOWPNO, YEAR dummies,  
OXBRIDGE

## 5. Results

### AIC from quadratic SFA LCM (2011 £)

	2008/09 to 2010/11		2005/06 to 2007/08		2003/04 to 2004/05	
AICs	Class 1	Class 2	Class 1	Class 2	Class 1	Class 2
UGMED	8720	19595	8351	8933	3958	4962
UGSCI	5260	7185	7708	11109	860	8753
UGARTS	5883	2176	-2354	6146	764	6576
PG	7839	1242	-10071	306	-4895	376
No in each class	236	119	132	195	100	96

Other outputs included: RESEARCH, IPINCOME

Controls for: LISTED, LOWPNO, YEAR dummies,  
OXBRIDGE

# 5. Results

## Economies of scale

**Quadratic model (for a HEI with mean levels of output) 2008/09 to 2010/11**

	SFA	SFA class 1	SFA class2
Ray economies	1.01	0.95	0.97
UGMED	1.25	1.11	1.23
UGSCI	1.00	1.26	0.75
UGARTS	1.23	0.84	0.46
PG	0.78	0.60	0.25
RESEARCH	1.13	0.97	1.00
IPINCOME	1.09	1.12	1.00

# 5. Results

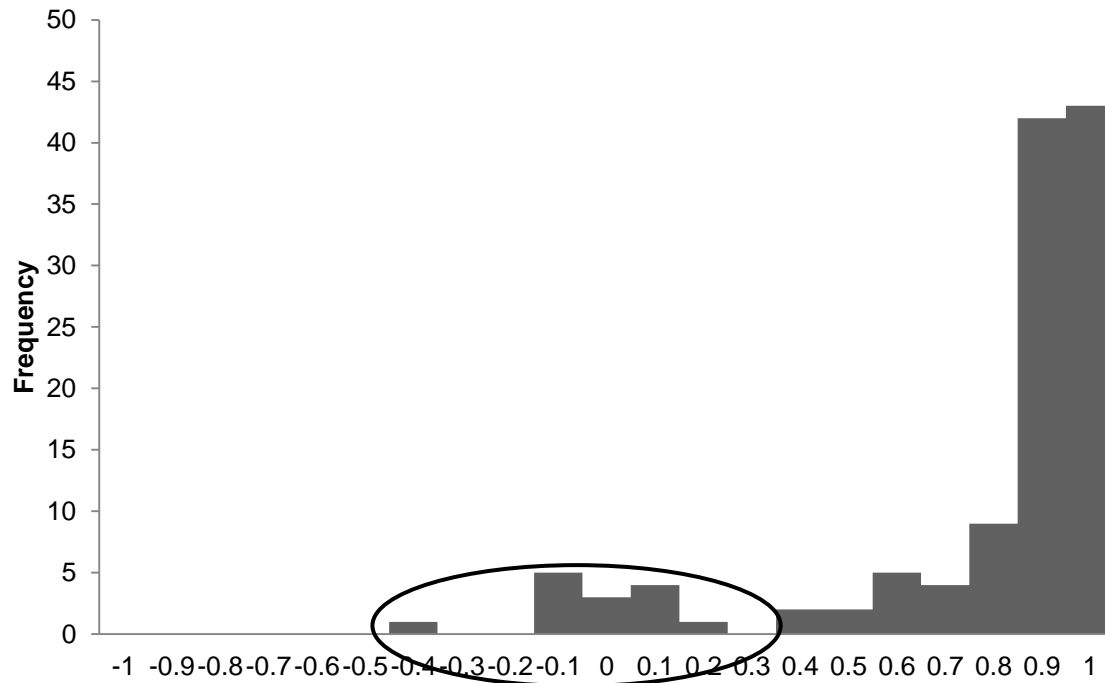
## Economies of scope

### Economies of scope (for a HEI with mean levels of output)

	SFA	SFA class 1	SFA class2
Global economies	-0.01	-0.13	-0.01

# 5. Results

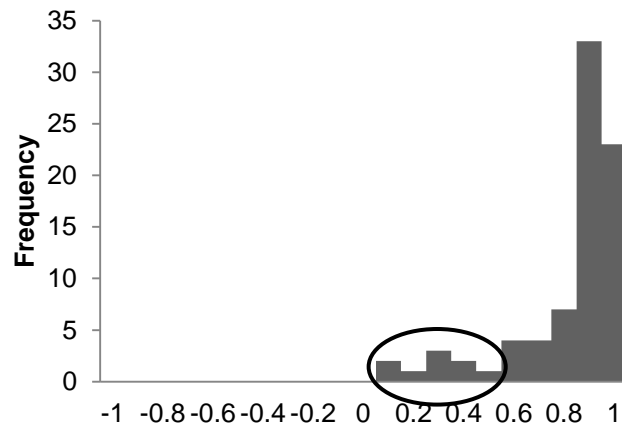
## Histogram of efficiency scores – final year of 2008/09 to 2010/11 quadratic model



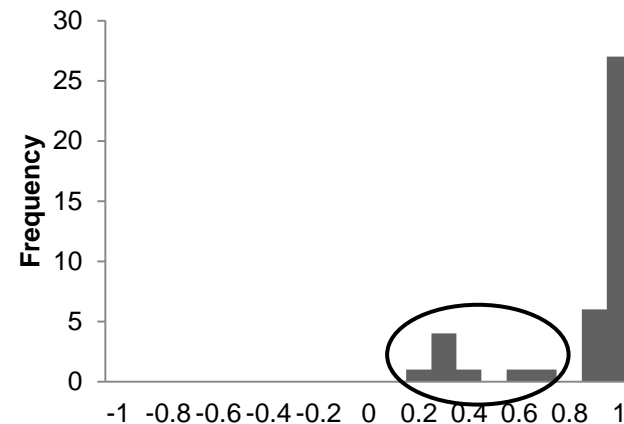
# 5. Results

## Histogram of efficiencies – final year of 2008/09 to 2010/11 quadratic latent class model

### Latent class 1



### Latent class 2



# 5. Results

## Comparison of Models with Akaike Information Criterion (AkIC)

$$\text{AkIC} = -2.\log\text{LF}(m) + 2.k$$

where k is the number of estimated parameters

No. of classes	2008/09 to 2010/11	2005/06 to 2007/08	2003/04 to 2004/05
1	-661.0	-367.7	-326.4
2	-848.9	-770.0	-567.8
3	-915.9	-922.8	-579.4



# 5. Results

	2008/09 to 2010/11			
AICs	Specialist	High tariff	Medium tariff	Low tariff
UGMED	12178	8265	8414	8839
UGSCI	2080	9827	8085	5024
UGARTS	12263	14850	3227	6925
PG	6411	11358	14609	11087
No. in each class	111	84	96	87
Is $\lambda$ significantly different from zero at the 5% significance level?	YES	NO	NO	YES

## 6. Conclusions

- Results for the earliest time period seem unreliable
- Estimates of AICs from SFA models (linear and quadratic) for the remaining periods seem plausible
- Estimates of AICs from SFA LCM seem less precise
- Ray economies of scale are exhausted; there are product specific economies in UG teaching and in research
- There are diseconomies of scope

## 6. Conclusions

- It is important to take into account other characteristics of universities (observable and unobservable) in estimating cost functions – efficiency differences are much lower once this is done
- Where a HEI has a low efficiency score this is usually explained by reference to special features observed in that HEI (eg. small, specialist)
- Can the LCM adequately deal with the heterogeneity observed in English higher education?