University of Strathclyde, Glasgow

Department of Economics, Strathclyde Business School

Three applications using Panel Data

John Alexander Houston

Presented in fulfilment of the requirements for the degree of Doctor of Philosophy

Declaration of Authenticity

This Thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date: 27th April, 2018

Abstract

This Thesis comprises three applied studies related to labour economics that utilise different Panel Datasets, assembled by the Author from public sources. The common Aim in all three studies is to determine whether the conclusions drawn from regression-based modelling would have been significantly different had standard estimators, as opposed to Panel-based ones, been applied. The stance taken, is that of the policymaker wishing to either evaluate the effectiveness of an established policy, or to quantify an established trend in Society, with a view to designing and implementing Policy in the future. In so doing, they wish to adopt an evidence-based approach to this exercise, using historic data.

While the three studies are related to labour economics, the have been deliberately chosen to differ from each other, both in the scenario being studied, the variables involved, and the estimators applied. The first study is related to gendered policing and arrests made for gender-based violence in England and Wales; the second, estimates the returns to qualifications in the UK Labour market, evaluated at the quantiles of males' and females' wage-rate distributions and the third, the technical efficiency of a panel of US Airlines.

The overall conclusions from the studies are that Panel-based estimators *can* lead to significantly different conclusions being drawn from regression-based evaluations but this need not always be the case. Nevertheless, where data are available as a Panel, or can be rendered as such, the Policy-maker should apply both Pooled- and Panel based estimators before drawing their conclusions from the exercise.

Acknowledgments

I would like to thank Anne, my Partner and fellow Economist and Academic for her unwavering support, patience and perseverance throughout the production of this Thesis. It simply would not have been possible to complete it without her.

I would also like to express my thanks to my supervisor, Professor Robert Wright, who suggested the topic to me and read my early drafts of the three studies and gave useful advice and encouragement throughout.

I would also like to express my thanks to Professor Alex DeRuyter and Dr Stuart McChlery, who were my Heads of Department while I was employed at Glasgow Caledonian University and undertaking the Research for the Thesis on a part-time basis. Professor DeRuyter arranged for a one-Semester Sabbatical for me, and both he and Dr McChlery ensured that I had the necessary hardware and software necessary to complete the Research.

I would also like to thank my children, Saoirse, Grace and Jack for being 'them'.

'Roma non uno die aedificata est'

John Houston, 27th April, 2018

Table of Contents

1. THREE APPLICATIONS OF PANELS 1

1.1. Introduction and Motivation 1

1.2. Regression Modelling: general approach and problems 5

1.3. Panels & Econometric Issues 8

1.4. Software employed 20

References used in Chapter One 20

2. REPRESENTATIVE BUREAUCRACY: GENDERED POLICING AND GENDER-BASED VIOLENCE IN ENGLAND AND WALES 23

2.1. Introduction & Motivation 23

2.2. Gender-based Violence (GBV): 25

2.2.1 International Context and Definitions 25

2.2.2 The English/Welsh perspective 28

2.3. Representative Bureaucracy 35

2.3.1 The Concept of a Representative Bureaucracy 35

2.3.2 Critical Mass in Society and Bureaucracies 39

2.3.3 Gender in the workplace and the particular effect of Women in the Police 43

2.3.4 Panel-based studies of Crime: a brief review 47

2.4. The Police Service in England and Wales 50

2.4.1. Introduction 50

2.4.2 Women in the English and Welsh Police Forces 53

2.5. Scope and Aims of this Study and the Data 56

2.5.1. Scope and Aims of this Study 56

2.5.2 The Dependent Variable: number of GBV Arrests 57

2.5.3 The Independent Variables: an overview 61

2.6. Testing the Hypothesis: models' estimation 69

Table of Contents (continued)

2.6.1. The Hypothesis **69**

2.6.2. The question of Truncation **72**

2.6.3. The Econometric Models **74**

2.6.4. Aggregates' Models **75**

2.6.5. Rates Models **84**

2.7. Discussion and Future Work **94**

2.7.1. Discussion **94**

2.7.2. Future Work **97**

References used in Chapter Two **99**

3. ESTIMATING RETURNS TO EDUCATION: WAGE RATE PREMIA IN THE UK ECONOMY: A PANEL-BASED QUANTILES' REGRESSION APPROACH 104

3.1. Introduction **104**

3.2. Outline Empirical Literature **112**

3.3. Estimating the returns to qualifications: the case for Quantile Regression **116**

3.3.1 Quantile Regression: a brief overview **116**

3.3.2 Quantile Regression: Panel Data **120**

3.3.3. Applications of Quantile Regressions to Earnings. **121**

3.4. Aims / Modelling Strategy / Included Variables **122**

3.4.1. Aims of the Study **122**

3.4.2 Income Inequality in the UK **122**

3.5. The Data and variables considered for the Study **128**

3.5.1. The Data **128**

3.5.2. The Dependent variable: Hourly Wage Rate **130**

3.5.3. The candidate Independent Variables **132**

3.6. Econometric Issues and Models **146**

3.6.1. Introduction **146**

Table of Contents (continued)

3.6.2. Attrition Bias **147**

3.6.3 Sample Selection Bias **149**

3.6.4. The Econometric Models **150**

3.7. Results **153**

3.7.1 Introduction **153**

3.7.2. Assessing the significance of the Attrition and Sample Selection Corrections **156**

3.7.3. Pooled Regressions: Quantiles v OLS **158**

3.7.4 Panel-based Results: Powell's FE Quantiles versus Means-based (Prais-Winsten) **163**

3.7.5 Quantiles' estimators: Powell's Fixed Effects Panel (qregpd) versus Cross-sectional returns (qreg) **168**

3.7.6. Non-additive v Additive Fixed Effects Panel Models (Powell v Canay) **171**

3.7.7. Relative Premia **176**

3.8. Final Discussion and Further Work **179**

3.8.1. Discussion **179**

3.8.2. Future Work **183**

References used in Chapter Three **184**

Appendix to Chapter Three: detailed Regression Tables **189**

4. ASSESSING AIRLINE EFFICIENCY: A STOCHASTIC FRONTIER APPROACH USING PANEL DATA 223

4.1. Introduction & Aims of the Study **223**

4.2 The competitive and technical conditions in the airline industry **224**

4.2.1. An outline history of competition and regulation in the International Airline Industry **224**

4.2.2. Measuring Industry Competitiveness **234**

4.3. Quantifying airline efficiency: Methodology & Literature **241**

4.3.1. Introduction **241**

4.3.2. Previous Studies on Airline Efficiency **241**

Table of Contents (continued)

4.3.3 Production Functions as a measure of producer efficiency 244

4.4. The Data & Econometric strategy 254

4.4.1 The Data 254

4.4.2 The Econometric Strategy 259

4.5. Results 265

4.5.1. Introduction 265

4.5.2. Cobb-Douglas Models: Panel v Cross sections 266

4.5.3. Translog Models: Panel versus Cross sections 271

4.5.4. Panel models, with and without Attrition Correction 277

4.5.5 Are there any patterns in the technical efficiencies? 283

4.6. Discussion & Further work 285

References Used in Chapter Four 288

5. FINAL COMMENTS, CONCLUSIONS AND OBSERVATIONS 292

List of Tables

Table 1.1 Wage Rate premia associated with Highest Qualification: Scotland 2

Table 2.1 Types of GBV and English/Welsh Legal Definitions 29

Table 2.2 Summary Geography and Demography of the Police Regions of England & Wales, 2002 and 2013 51

Table 2.3 Home Office Offences assumed to be GBV related 58

Table 2.4 Independent Variables: partial correlations' distributions (Aggregates) 78

Table 2.5 Dependent Variable: GBV Arrests 81

Table 2.6 Independent Variables Partial correlations' distributions (Rates) 89

Table 2.7 Dependent Variable: GBV Arrest Rates v Rates' Variables 91

Table 3.1 Hourly pay - Gross (£) - For all employee jobs: United Kingdom, 2015 126

Table 3.2 Understanding Society Gender Counts by Wave & Age-group percentages 128

Table 3.3 Attrition from the US Panel (All Respondents between 18 and 60: Waves 1 - 7) 129

Table 3.4 Centile & Average Wage Rates: UK 2009 and 2015 131

Table 3.5 Highest Qualification by Gender (2009 and 2015) 134

Table 3.6 Mean Rates by Gender 135

Table 3.7 Public v Private Sector & Trade Union v Non-Trade Union members: Mean Wage Rates by Gender Understanding Society 2009 and 2015 137

List of Tables (continued)

Table 3.8 Average Wage Rates for Males and Females with and without a second job 138

Table 3.9 Median full-time gross weekly earnings by UK region, April 2015 141

Table 3.10 Average Regional Wage Rates by Gender: 2009 and 2015 142

Table 3.11 Occupational Wage Rates: US Waves 1 and 7 143

Table 3.12 Sectoral Wage Rates: US Waves 1 and 7 144

Table 3.13 Effects of Attrition and Sample Selection Corrections on Degree Premia 157

Table 3.14 Percentage boost to Wage Rates attributable to qualifications 177

Table A.3.1 Heckman OLS Regression: Sample Selection corrected: Males (Figure 3.20) 189

Table A3.2 Male Quantiles with sample selection correction (Figures 3.20 and 3.24) 192

Table A3.3 Heckman OLS Regressio : Sample Selection corrected : Females (Figure 3.21) 196

Table A3.4 Female Quantiles with sample selection correction (used in Figures 3.21 and 3.25) 199

Table A.3.5 Linear Prais-Winsten Estimator With heteroscedasticity and serial correlation corrections (Figure 3.22) 203

Table A.3.6 Male Quantiles Panel with attrition correction and sample selection corrections (Figures 3.22 and 3.24) 205

Table A3.7 Linear Prais-Winsten Estimator With heteroscedasticity and serial correlation corrections Females Figure 3.23 209

Table A3.8 Female Quantiles Panel with attrition and sample selection corrections Figure 3.23 and 3.25 211

List of Tables (continued)

Table A.3.9 Male Canay with Attrition and Sample Selection Corrections (Figure 3.27) 215

Table A.3.10 Female Canay with Attrition and Sample Selection Corrections (Figure 3.28) 219

Table 4.1 Development of the Boeing 747 (1970 to the present) 231

Table 4.2 Concentration measures for the World Airline Industry: Passengers carried 2011-2016 237

Table 4.3 Concentration measures for the World Airline Industry: Cargo carried 2011-2016 238

Table 4.4 US Airlines: Cobb Douglas Stochastic Frontier Production Function (no attrition correction) 267

Table 4.5 US Airlines: Translog Stochastic Frontier Production Function (no attrition correction) 272

Table 4.6 US Airlines: Stochastic Frontier Production Function (Panel specifications) 279

List of Figures

Figure 1.1. Regression Line fitted through observed x and y values 6

Figure 1.2 Schematic of a Panel Dataset 9

Figure 1.3 Fixed Effects Regressions: Unobserved Heterogeneity 10

Figure 2.1 Number of recorded arrests for GBV offences in England and Wales (2002 to 2013) 28

Figure 2.2 Proportions of GBV Arrests England & Wales 2002 – 2013 31

Figure 2.3 The 43 Police Force Areas in England and Wales 52

Figure 2.4 Total Officers and Female Officers: England and Wales 2002 – 2013 56

Figure 2.5. Indexes of Total Arrests for the four largest Police Forces 2002 – 2013 (2002 = 100) 59

Figure 2.6 Indexes of GBV Arrests for the four largest Police Forces 2002 – 2013 (2002 = 100) 60

Figure 2.7 Percentages of Arrests made for GBV: 2002:2013 61

Figure 2.8 Index of Forces' Strength: 2002 – 2013 64

Figure 2.9 Percentage of Female Officers in the two best- and two-worst gender-balanced Forces 65

Figure 2.10 Total Population and Households and Household Densities: England and Wales 2002 – 2013 66

List of Figures (continued)

Figure 2.11 Population Indexes for the two fastest and two slowest-growing Forces in England and Wales 67

Figure 2.12 Force Area Gender compositions (2002 – 2013). Highest and lowest % female 67

Figure 2.13 Greatest- and smallest Household Densities in England and Wales 2002 – 2013 69

Figure 2.14 Female-to-Male Officers Ratios: Forces with the highest, median and lowest average Ratios (2002 – 2013) 87

Figure 2.15 Female-to-Male Population Ratios: Forces with the highest, median and lowest average Ratios (2002 – 2013) 87

Figure 2.16 Per-Capita Arrest Rates: highest, median and lowest average rates 2002 – 2013 88

Figure 3.1 UK Unemployment - Count & Rate: 1922 – 1994 105

Figure 3.2. Economic Activity Rates (UK) 1992 – 2017 106

Figure 3.3. Accepted Application to UK HE Institutions and proportions of 18-24 year olds enrolled in Degree Programmes in the UK 107

Figure 3.4. Percentage of GDP spent on Public Education: G7 Economies (1980 - 2012) 108

Figure 3.5. UK Real spending on Public Education 1953 – 2015 108

List of Figures (continued)

Figure 3.6. HEFCE & SFC Grant Support to Universities 2007 – 2016 109

Figure 3.7. Student Loans' Takeup 1990 – 2014 110

Figure 3.8 Value of loans taken out 1990-2014 110

Figure 3.9. Graduate and Non-graduate Inactivity Rates UK 1993 – 2017 111

Fig 3.10 Conditional Quantiles 119

Figure 3.11. GINI Coefficients: G7 Economies (2002 - 2015) 123

Figure 3.12 UK Income Distribution (1999 - 2014) 124

Figure 3.13. Distribution of full-time gross weekly earnings, UK, (1997-2015) 124

Figure 3.14 UK Average Gross Weekly Earnings by Industrial Sector (2000 - 2016) 127

Figure 3.15 10th & 90th Centile Wage Rates as % of Median: UK 131

Figure 3.16 Wage-rate centile composition of each qualification level: Understanding Society Wave 4 (2012) 136

Figure 3.17 Wage Rates by Age by Gender: Understanding Society Wave 1 (2009) and Wave 7 (2015) 139

Figure 3.18 Gender Pay Gap and Age: Understanding Society Wave 1 (2009) and Wave 7 (2015) 140

Figure 3.19 Means-based and Quantile Returns 155

List of Figures (continued)

Figure 3.20 Pooled Regression Results: Males 159

Figure 3.21 Pooled Regression Results: Females 161

Figure 3.22 Panel Regression Results: Males 164

Figure 3.23 Panel Regression Results: Females 166

Figure 3.24 Panel and Pooled Quantiles' Results: Males 169

Figure 3.25 Panel and Pooled Quantiles' Results: Females 170

Figure 3.26 Estimated Fixed Effects, ln(wage rate) and [ln(wage rate)-α_i] distributions 172

Figure 3.27 Canay v Powell Panel Quantile Estimators (Males) 174

Figure 3.28 Canay v Powell Panel Quantile Estimators (Females) 175

Figure 4.1. Real Spot-price of aviation fuel (US Gulf Coast Kerosene) 234

Figure 4.2. Real and Nominal Average Airfares in the US: 1995-2016 239

Figure 4.3. US Consumer Prices Index for Airfares and all items: 1989-2018 240

Figure 4.4. The Stochastic Production Frontier and Inefficiency and Idiosyncratic elements 252

Figure 4.5 Central and extreme values of the dependent and independent variables 1990 – 2014 256

Figure 4.6. Output per aircraft and per employee: US Airlines 1990 – 2014 258

List of Figures (continued)

Figure 4.7. Highest and Lowest Average Productivities: Aircraft and Employees 259

Figure 4.8. Technical Efficiencies: Cobb-Douglas models (No attrition correction) 270

Figure 4.9. Technical Efficiencies: Translog models (No attrition correction) 276

Figure 4.10. Technical Efficiencies: Cobb-Douglas Panel models (with and without attrition correction 281

Figure 4.11 Technical Efficiencies: Translog Panel models (with and without attrition correction 282

Figure 4.12. All-time ranks for the top- and bottom-ranked airlines 285

1. Three Applications of Panels

1.1. Introduction and Motivation

This Thesis makes a contribution to knowledge, by assessing the role that empirical data and, in particular, regression-based models play in the evaluation and formulation of policy making. The Author has made a modest contribution in the past both within Scottish- and UK settings (Houston, et al (1999); Gasteen, et al (2000); Houston, et al (2001); Houston and Gasteen (2007), Houston (2016) and Johnson and Houston (2018)), using a variety of primary- and secondary data sources. The common theme of these publications was to assess and critique the effectiveness, or otherwise of policy, reflecting this Author's interest in applying standard quantitative to practical issues in Economics. It was in 2003, when working on estimating the returns to Further Education Qualifications in the Scottish Labour market (Houston, et al, 2005), that the Author experienced at first hand and for the first time, the application of *Panel Data* to a policy-related research question. It was striking that the estimates of the premia associated with particular levels of qualification differed when the *same data* (the now discontinued British Household Panel Survey (BHPS)) was used in conjunction with standard 'pooled' regression estimators, and also a variety of panel-based estimators. **Table 1.1** is extracted from that Study and indicates the differences in wage-rate premia accruing to different levels of qualifications, according to the estimator employed.

Although not explicitly addressed in that Study, the differences in the estimates obtained could have led to entirely opposite conclusions as to the 'value' of any of the qualifications to their holders, at least in the Scottish Labour market, at that time. Other studies, some of which were used in instructional texts, such as Greene (2008) illustrate significant differences between Pooled- and Panel-derived estimates. [1] This Author easily came to the conclusion that regression-based studies of this type should be, by default, based on Panel Data.

[1] Table 9.4, page 199.

Table 1.1 Wage Rate premia associated with Highest Qualification: Scotland Dependent variable: log of wage rate			
	Panel Models		Pooled Model
	Fixed Effects	Random Effects	
University Degree	0.3471 (0.5036)	0.6515 (0.0815)	0.7939 (0.0344)
College HNC/D	0.4558 (0.4652)	0.3888 (0.0761)	0.4540 (0.0315)
Other College	-	0.0565 (0.1533)	0.2349 (0.0444)
School leaving	0.4059 (0.4639)	0.2145 (0.0770)	0.3423 (0.0339)
Lower School	0.3565 (0.4254)	0.0771 (0.0766)	0.1793 (0.0320)
Source: Houston, et.al (2005), extracted from Table 12.1, pp 193-4			

This Thesis is not about developing either Economic, or Econometric Theory. Rather, it is a collection of three, otherwise unrelated investigations into very different scenarios. The interest is in whether or not the magnitude and statistical significance of panel-derived estimates of underlying relationships between a dependent variable and the included independent variables are, perforce, so different from the Pooled estimates, that the conclusions drawn would change as a result. The stance adopted in each Study is to imagine a Policymaker wishing to either evaluate the effectiveness of a Policy they set in train in the Past, or to quantify the effects of an established trend in Society, with a view to developing Policy designed to either nurturing it, or to battle against it, for that Society's benefit.

Each Study applies an orthodox economic model (where one exists) to each issue, assembles bespoke Panels using published Primary Data, then applies appropriate, existing Econometric estimators and statistical tests to the data. The contributions to knowledge stem from the application of the chosen techniques to areas not currently in the Literature, and in the case of one of the Studies, apply a little-used Panel estimator to *quantile*, as opposed to *means*-based estimates of the relationships between the selected variables.

The three Studies in this Thesis were selected on the basis that they involve real-life issues that affect particular sections of Society of interest to the Author, and also because the scenarios differ greatly from each other with respect to the outcome variables. If they do share anything in common, they are all concerned, directly or indirectly, with *labour markets.* Directly, in the sense that the effect of educational policy on a specific labour market outcome, wage rates, is examined. Indirectly, in the senses that (i) increasing the participation of females in a particular Public Service has implications for both female- and male participation in the future, (ii) altering the input-mix into a particular production process has implications for future recruitment into that industry. The three Studies, in outline are:

1. The effect of increasing the representation of female police has on the arrest rate for Gender-based Violence (GBV) in England and Wales (Chapter Two)
2. The economic returns to qualifications in the UK, evaluated at the quantiles of the male and female wage-rate distributions (Chapter Three)
3. The extent to which there are differences in commercial airlines' technical efficiency in the United States of America (Chapter Four).

The second- and third Studies have, at their core, established *economic theory* and associated quantitative models: a 'Mincer' model of the wage-rate and a Stochastic Frontier Production Function, respectively. The first Study has at its core, theory emanating from *Administrative Science* that relates the composition of an organisation's workforce to the effectiveness of the service it provides to Society. It does not use, nor does it develop, a criminologically-based theory as to the causes of GBV, or an economic one as to why females join the Police or, having done that, *why* they arrest or don't arrest someone on suspicion of having committed GBV. Instead, the work attempts to correlate the increased participation of females in the Service, against the number of GBV arrests made.

All three studies adopt a regression-based approach where a dependent variable is regressed against a number of variables deemed to be relevant, some of which represent one or more *policy instruments*. Each Study has its own, bespoke Panel, assembled by the Author from published sources. The first Panel (GBV) uses published Home Office data on crime and policing levels, combined with pertinent demographic data on the regions of England and Wales. The second (Wage-rate returns) uses the *Understanding Society* Panel, and the third (airlines' efficiency), two separate, public sources of inputs to, and outputs from a sample of US airlines. One of the frustrations of reporting the work done for applied research like this, is the difficulty in conveying to the reader the effort expended to assemble and 'clean' the data. Of the three Panels, the GBV one was probably the most straightforward to assemble, as the Home Office makes available large amounts of data, aggregated annually at the level of the Police Force 'Area' on the numbers of each category of crime, and detailed analysis of the composition by rank and gender of each Area's Police Force. [2] The Office for National Statistics does likewise for population statistics across the UK. Assumptions had to be made as to what constituted GBV, and there had to be some interpolation of the population data, given the Study Period straddled two decennial Censuses, and some realignment between the Police 'Areas' and ONS 'Regions'.

The wage-rate Panel was only more arduous due to its sheer size of the Understanding Society ('US') survey and care was required to link the different years' surveys together. The Author had the experience gained from working with its predecessor, the BHPS to fall back on, which assisted greatly in the process. The most difficult Panel to assemble was the Airlines' one, and is the one that, as discussed in Chapter Four, would benefit the most from either clerical resource to aid assembly or 'single-click' access to data that only requires some modest rendering to make it amenable to estimation.

[2] Unfortunately, the Scottish Government chooses not to provide the same aggregated data on crime and policing, either publicly or in response to requests by the Author, preferring to seek shelter behind Data Protection legislation to deny access to it.

The remainder of this chapter will consider the nature of panel data and associated estimators, the advantages Panel Estimators may have compared with the standard cross-sectional ones, as well as the problems associated with these data and estimators. Each subsequent chapter will describe and discuss each Study in detail, and a short concluding chapter will draw together what has been learned from the research.

1.2. Regression Modelling: general approach and problems

The origins of regression can be traced back to the early 19th Century, to mathematicians such as Johan Gauβ, who applied calculus and algebra to fit 'least-squares' lines to data in an effort to explain, *inter alia*, planetary motion. The term 'regression' refers to the tendency of sample data to reveal the underlying, 'true' population mean as the sample size is increased, in other words, to *regress* towards the mean. The tenet of regression is that a variable *y* derives its value entirely from one, or more other variables, labelled *x* ('co-variates'). As such, *y* is functionally *dependent* upon these x-variables:

$$\boldsymbol{y = f(x_1, x_2, x_3,, \dots x_k)}$$

This functional model can be converted into a statistical, or *econometric*, model as follows:

$$\boldsymbol{y = \beta_1 x_1 + \beta_2 x_2 + \beta_2 x_3 + \cdots + \beta_k x_k + \varepsilon} \qquad [1.1]$$

That is to say, there is a vector of parameters (***β***) that quantifies the per-unit relationship between each *x* and *y*. Summed together, they would generate an estimated ('fitted') value of y $(\hat{y})$, and, in this form, is deterministic in nature. There is, however, a random element (ε) that recognises that some part of *y's* value will come from apparently random shocks, and that it will have its own governing statistical distribution. A member of the population, *i* will be observed and measured with respect to all the variables and their observation would be modelled as:

$$\boldsymbol{y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_2 x_{3i} + \cdots + \beta_k x_{ki} + \varepsilon_i}$$

Other members of the population (*j, k, l…*) can be added and methods, such as Gauβ's Least-squares used to estimate the 'best-fitting' vector, ***β***. To this, may be added a constant term (β_0) that monotonically moves the regression function vertically and help it to better track its way through the observed points, as **Figure 1.1** illustrates in the single-x case.

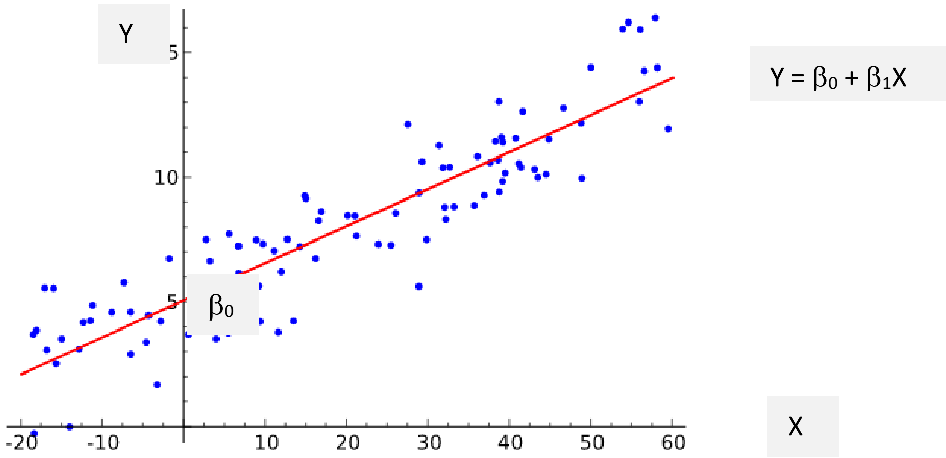


Figure 1.1. Regression Line fitted through observed x and y values [3]

β pertains to all those in the sample that derived it, and by implication all others in the population that were not included in it: an important issue when it comes to considering the advantages of Panels.

Economists adopted Regression Analysis as a key method in trying to understand and explain behaviour in economic systems. In 1930, the Econometric Society was founded to:

> '…promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative and the (latter's) approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate the natural sciences.'

(Frisch, 1933)

[3] By Sewaqu - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=11967659

In the absence of computers and associated software, researchers had to derive $\boldsymbol{\beta}$ manually. This, perforce, limited the size and scope of empirical regression-based research to models with a handful of variables and observations, until the 1980s when rapid progress on both hardware- and software fronts allowed much larger models to be tackled. That said, statisticians and econometricians did not wait until then to identify and attempt to tackle the standard 'problems' of econometrics. The issue of *collinearity* had been understood since the time of Euclid, but its potential effect on $\boldsymbol{\beta}$ not fully appreciated until the work of people like Longley (1967). The more severe problem of *heteroskedasticity* led to the development of several, competing detection tests (e.g. Breusch and Pagan, 1979) and corrective treatments suggested and encoded into emerging Econometric Software. The issue of *serial correlation* in time series residuals was first tackled in the 1950s (Durbin and Watson, 1950), leading to an extensive array of diagnostics tests and remedies, latterly extending into the issues of Unit Roots and spurious regressions.

All these issues, with the exception of multicollinearity, have the potential to either significantly bias the sample-derived estimates of $\boldsymbol{\beta}$. Or the associated standard errors, with potential implications for conclusions regarding the significance of key variables. To this, can be added the issue of *omitted variables' bias* ('OVB'), that is the omission of one or more significant *x* variables from **[1.1]**, either as a result of non-availability of data, lack of awareness that the variable(s) omitted should have been included, or modeller bias, emanating from a 'distrust' of the variable(s). The effect would be to produce estimates of $\boldsymbol{\beta}$ that are both based and inconsistent. Suppose the true relationship y has is embodied in this cut-down version of **[1.1]**:

$$\boldsymbol{y} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_1} + \boldsymbol{\beta_2 x_2} + \boldsymbol{\varepsilon} \qquad [1.2]$$

but that x_2 is omitted from the model, and **[1.3]**, by implication, is estimated instead:

$$\boldsymbol{y} = \boldsymbol{\gamma_0} + \boldsymbol{\gamma_1 x_1} + \boldsymbol{\upsilon} \qquad [1.3]$$

Substituting **[1.3]** into **[1.2]** will give:

$$y = (\beta_0 + \beta_2\gamma_0) + (\beta_1 + \beta_2\gamma_1)x_1 \quad + (\varepsilon + \beta_2\upsilon)$$

Accordingly, the marginal effect of x_1 on *y*, viz $\left(\frac{\partial y}{\partial x_1}\right)$ which should be β_1, would in fact be estimated as, $(\beta_1 + \beta_2\ \gamma_1)$. The residual is now also a compound of what it should be, plus what it would have been, factored by the missing parameter, β_2. The obvious remedy is to include the missing variable(s): easier said than done where there is no established, accepted theory, or the data are simply not available or unobservable (such as *ability* in a wage equation, for example). The fashion has been to use *instrumental variables* instead and hopefully allow the true marginal effects of the included x variables to be revealed. Allied to this is the hazard of assuming that relationship between y and a particular *x* variable is perforce, *linear*. Failure to include, for example, higher- or lower powers of *x* is a form of OVB, and a failure to replace an inappropriate, linear-x with a transformed x (say $\log_e(x)$), would constitute a mis-specification of the relationship of *y* with *x*.

Advocates of Panel-data methods claim that some of these issues, in particular multicollinearity and OVB can be avoided, or ameliorated with Panel Data. In the next section, what Panel Data are, and what potential advantages they offer the empirical researcher, such as the Author, will be considered.

1.3. Panels & Econometric Issues

Baltagi (2013) defines a Panel as:

> '… the pooling of observations on a cross-section of households, countries, firms, etc over several time periods. This can be achieved by surveying a number of households or individuals and following them over time.'

Figure 1.2 depicts what a Panel dataset would look like in practise. The dataset consists of *N* 'individuals', (people, firms, countries, often referred to as 'units' or 'groups') each observed on *T* occasions ('waves'), on *K* independent variables and one dependent variable. It is necessary, therefore, to identify each data element by the individual that generated it (i), what variable is being referred to (*k*) and in what time period it was observed (*t*).

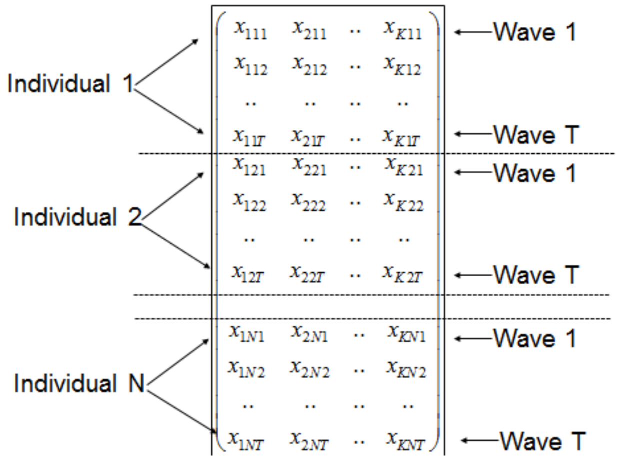


Figure 1.2 Schematic of a Panel Dataset

Contrast this with a dataset that can only distinguish when the data were collected (if it were a time series) and what variable is being referred to, but not what unit generated the data. This would be termed a *cross-section* and from it would be estimated models in the style of **[1.2]**. While the *N* dimension in the Panel could be ignored (as is done in the three Studies to allow a comparison with the Panel Estimators), it would now be possible to estimate regression models in the style of **[1.4]**.

$$y_{it} = \beta_{0i} + \sum_{j=1}^{k} \beta_j x_{jit} + \varepsilon_{it} \qquad [\mathbf{1.4}]$$

The main, ostensible difference between **[1.2]** and **[1.4]** is the ability to estimate a personal constant, or 'fixed effect' (β_{0i}) for each unit in the Panel, as illustrated in **Figure 1.3**. The general effect is to allow the regression to find an intercept term for each unit, rather than all having to share one. This also permits the (shared) slope of the regression equation to be β'_1, rather than β_1 (recall the differences in wage-rate returns in **Table 1.1**).

[1.4] is termed the Fixed-Effects Panel model, and the individual constants are understood to be the *omitted*, or *unobserved* effects, i.e. those characteristics possessed by the unit that were, for whatever reason, not included in the dataset and may exert a significant influence on the dependent variable. Where those omitted effects are correlated with the observed variables, then **[1.4]** is appropriate. It is easy to imagine situations where the fixed effect is correlated with the included variables, for example, *ability* and *highest level of qualification* and also that fixed effects may be so significant as to materially alter the marginal effect of any of the x-variables on the dependent variables (as **Table 1.1** and **Figure 1.3** appear to indicate).

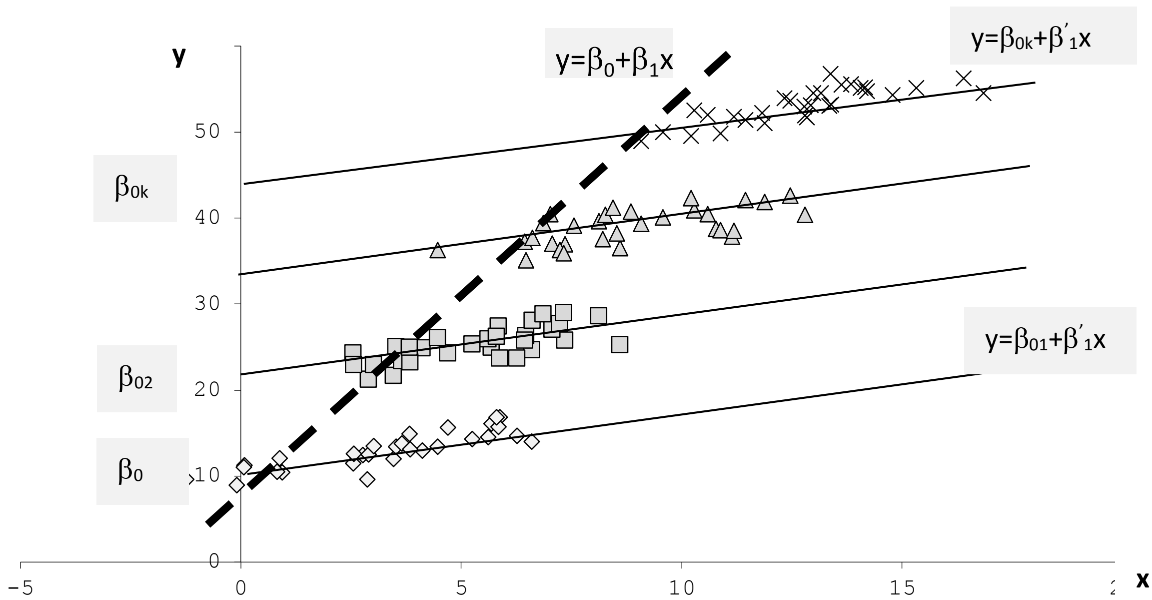


Figure 1.3 Fixed Effects Regressions: Unobserved Heterogeneity

Whether or not effect of estimating **[1.4]** instead of something like **[1.2]** is to significantly change the estimated parameters is conventionally tested via an F-test of the Null Hypothesis that the individual constants are all equal to the common constant (and therefore each other), where:

$$F = \frac{(R^2_{FE} - R^2_{OLS}) \big/ (n-1)}{(1 - R^2_{FE}) \big/ (nT - n - K)}$$

N is the number of units in the sample, *T* is the number of waves they have been observed and *K* the number of parameters being estimated.

[1.4] is known as a *One-way error component* model, where there is one unobserved (fixed) effect, pertaining to each unit included in the Panel that does not vary over time. This can be extended to include the unobserved effect of *time* on all the units, by including of time dummies (λ_t) in the specification, viz.

$$y_{it} = \beta_{0i} + \sum_{j=1}^{k} \beta_j x_{jit} + \sum_{m=1}^{T} \lambda_t + \varepsilon_{it}$$

which is known as the *Two-way error component model*. This specification is also used in the GBV Study, given the potential for unobserved time-based effects on societal attitudes to that particular type of crime.

One criticism of the Fixed Effects model is that it estimates lots of parameters (K regressors plus *N* Fixed effects) and where *K* is large and *T* is small, the degrees of freedom available may be uncomfortably tight. Also, the estimated fixed effects can pertain only to the units included in the Panel, and not to those not included in the sample. In the GBV Study, this is not an issue, *per se*, as all 43 Police Forces are included, and nothing is to be concluded about the excluded Scottish or Northern Irish Forces. However, it might be in the returns to qualifications' Study, as the assembled Panel of ~215,000 observation on ~42,000 individuals represent only about 0.14% of the UK's working population (**Table 3.2**). If, however, it proves to be the case that the unobserved effects are *uncorrelated* with the included variables, then it can be assumed that they are drawn randomly from an unknown population distribution, with a population mean α, and a unit-specific amount, u_i, where $E[u_i|X] = 0$ and $E[u_i^2 = \sigma_i^2]$. **[1.4]** can be re-cast as:

$$\boldsymbol{y_{it} = (\alpha + u_i) + \sum_{j=1}^{k} \beta_j x_{jit} + \varepsilon_{it}}$$

This is known as the *Random Effects* (RE) model and can be estimated by Generalised Least Squares. It has the advantage of having fewer parameters to be estimated than the Fixed Effects (FE) model. The RE model is assumed, by default to be the efficient estimator and is tested, as standard test by the *Hausman Test* computes a test statistic which follows a Chi^2 distribution, where:

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})'[V_{FE} - V_{RE}]^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{RE})$$

where $\hat{\beta}_{FE}\ and\ \hat{\beta}_{RE}$ are the vectors of parameters estimated by the FE- and RE estimators respectively, and V_{FE} and V_{RE} are the variance-covariance matrices of those estimators. Rejection of the Null is assumed to imply acceptance of the FE models. A failure to reject the Null also requires there to be no correlation between the unobserved effect and the included variables: a strong assumption.

Panels are not immune to either *heteroskedasticity* of the error terms, or *serial correlation*. The three studies in this Thesis bring together units of very different size. In the GBV study, the huge London Metropolitan Force appears alongside its tiny neighbour, the City of London Police (see **Table 2.2**). In the qualifications' returns Study, individuals on £100 an hour are combined with those on the National Minimum Wage. The airline study contains airlines that output billions of tonne-kilometres, alongside those outputting less than two million (see **Figure 4.2**). The detection of heteroskedasticity in the GBV model was achieved variously via (i) examining plots of the residuals, backed up with (ii) regressions of the squared residuals versus the x-variables and (iii) conducting Lagrange Multiplier tests of difference between models with and without heteroskedasticity. Models with robust standard errors were also estimated by default. Similarly, in the wages' study, robust standard errors and plots of the residuals were examined and, where possible, Breusch-Pagan tests performed. In the case of the airline Study, the strategy is to estimate models on the assumptions of homo- and heteroskedasticity. Stata allows the variance of both error components to vary with the 'size' of the unit and report the significance of the 'size' parameter on the variance of either/both errors. A significant 'size' parameter is understood to signal the presence of heteroskedasticity, and that the correction to the estimates, probably justified.

Panels are, perforce Time Series, and susceptible to serially-correlated errors. With the GBV study, serial correlation was tested for using the Wooldridge Test (Wooldridge, 2002) as encoded via the *xtserial* routine by Drukker (2003). Regression assuming Panel-specific and Common autocorrelation (AR1) were included and compared. In the case of the GBV study, the Panel was tested for Unit Roots, using standard Durbin-Watson tests and whether or not they are cointegratable via Kao-Pedroni tests.

Thus, Panels offer the opportunity to allow estimates of the unobserved component to be made, and perhaps to allow the true relationships between y and x to be revealed. This may overcome the problem of OVB discussed earlier, though efforts to explicitly include relevant variables should always be made. Other claimed virtues of Panels include:

- Lowering the incidence and magnitude of multicollinearity amongst the x-variables by the inclusion of the extra dimension of the individual, counterfactually loosening the links between the variables.
- The ability to separate out the between-units and within-units' variation.
- The opportunity the model *dynamics*, where units change their status from one period to the next, perhaps in response to changes in one or more of the x-variables. The key point is that it is the same units who are changing their status (or otherwise), ergo their unobserved heterogeneity is controlled for, and the 'pure' effect of the x-variable change identified (this, of course assumes that each unit's unobserved effect does not, itself, change through time). It is also possible to model the duration of these status changes, useful in, for example, studies of un/underemployment.
- Assumed by some to be a Panel-specific shortcoming, Panel-data are better able to reveal some types of *measurement error* than a single cross-section could. With a cross section, only patently absurd responses could be identified – such as negative hours worked, or the

> legendary four-year old High Court Judge that lurked in a large UK survey on earnings - otherwise, reporting bias and clerical errors may still yield feasible, though inaccurate data. With a Panel, there is the ability to first-difference consecutive wave's data, to reveal any massive changes from one period to the next. Depending upon the size of the Panel, and the resource available to clean it, these could then be investigated and where necessary, corrected or eliminated from the Panel.

This exercise was performed on all three Panels, though with some assurance in the case of the GBV and Wage Rate data, that the collecting agencies would have checked the validity of the data before making it public, or would have been alerted by other users, post-release and the data withdrawn, edited and re-released. With the airlines data, the differencing revealed some relatively large changes year-on-year in the inputs and outputs. A closer inspection revealed that most of these were genuine, particularly with the smaller airlines where the addition, or disposal of a single aircraft into/from a small fleet of two or three, would result in large relative changes in output.

There are, however, Panel-specific problems that emanate precisely from the practise of observing the Unit more than once. One, of course, is the *additional cost* that committing to repeating the survey more than once, with the consequent administration involved in keeping track of its participants. This can place a question marks over the feasibility of the Survey itself, particularly in times when public resources are stretched (such as I the last decade in the UK, for example). Baltagi (2013, op.cit) quotes studies that identify *responder fatigue* from being asked the same questions at each wave, particularly in high-frequency studies. There is also the issue of *responder amnesia* regarding things that may have happened to them months, even years ago, as well as a *duration bias*, where individuals that have been in the Panel for several waves may become blasé about the responses they give, defaulting to parroting previous responses to a given question.

Of the three Panels used in this Thesis, only the Earnings' Panel may suffer from any of these problems. The GBV Panel includes all 43 Police Forces in England and Wales and the 'responses' were not to individual questions but had been aggregated by Home Office statisticians from the administrative data originating at 'street-level'. Similarly, the population data were collated by ONS statisticians, albeit using from the decennial Census, administered on all individual households in the UK. The airlines' panel is not a census, *per se*, but collated from administrative data. Responder bias in the Understanding Society survey is inevitable however, and an accepted hazard facing researchers using it. Again, reasonableness checks can be carried out on contiguous responses, and responder fatigue mitigated by not asking *all* the questions in each wave (cycling whole sections a biennial basis, for example).

However, the major issue affecting Panels like Understanding Society is the potential of non-random *attrition*. **Figure 1.2** depicts a situation of perfect completion of the panel: all units are ever-present throughout its T periods' duration, responding in such a way that the K variables could be filled with data. However, given the voluntary nature of participation, respondents may simply refuse to participate, perhaps fearing that the actual aim of the Survey is not was indicated to them. People move address, failing to leave new contact details, or die or become otherwise, incapacitated. Individuals benefitting from a 'policy treatment' are more likely to remain attached to the Survey, than those whose welfare has been reduced by the same Policy. As discussed in section **3.5.1** there is a significant issue with non-random attrition from the Understanding Society Panel, particularly among ethnic minorities and certain age groups in the population. The Survey's statisticians attempt to counteract the problem by generating weights to boost the effect of 'under-represented' individuals in the sample and reduce the effect of 'over-represented' groups. The Author, however, preferred to follow the two-step procedure proposed by Becketti, et al (1988). The first step is to assess the significance of attrition from the Panel via a regression of the dependent variable against the independent variables, an attrition dummy and interactions with it and the independent variables. The data are restricted to only the first observation for each individual, whenever it was observed in the seven waves. An F-test of the joint significance of parameters involving the attrition dummy is performed and if that proves to be significant, the second stage computes an attrition weight.

This involved the computation of a subjective 'attrition risk' score, where observed attributes regarded as likely to lead to attrition from the Panel were combined to yield the score. Two probit regression of the attrition dummy against (i) the independent variables used in the main regression plus the attrition score and, (ii) the independent variables only are carried out. The fitted value from (ii) is divided by the one from (i) to generate the attrition weight. More detail on this is given in section **3.5.2** in the GBV study, and results with, and without the attrition weights reported from some of the Panel regressions. The same approach was used in the Airlines' Study and models with- and without the attrition correction estimated and compared (see sections **4.4.2** and **4.5.4**).

Panels are equally susceptible to significant non-random *sample selection bias*, where the effect is to render the sample unrepresentative of the population, therefore producing biased results. This is obviously not an issue with the GBV Study but will likely be a major issue that will confront the on wage returns study. In this situation, the dependent variable will be the (log of the) hourly wage rate at time *t*, w_{it}, to be regressed against a set of observables deemed to determine its level, i.e.

$$w_{it} = \beta \mathbf{x}_{it} + u_{it} \qquad \textbf{[1.6]}$$

where $\mathbf{x}_{it}$ is a vector of observations on the selected 'independent' variables, β are the parameters to be estimated and u_{it} is the disturbance term [u_{it} ~N(0,σ)].

The problem, however, is that a wage rate is only observed iff the individual chooses to work (for money), and that they will only do so if their minimum, or *reservation wage* (rate) has been at least met by an employer. Where it has not been met, they will not work, and they will be unobserved, *de facto*. In effect, there is a prior stage, called the *selection stage* in which the individual is considering the opportunity cost of any work, and the associated wage rate(s) currently available to them. The greater the opportunity cost associated with the non-work alternatives, the higher will be the *reservation wage*, the more likely that they will be unobserved in the wage study as they engage, instead, in 'leisure'. [4] One way to model this 'Work / Not work' binary decision is via a Probit regression of the discrete choice variable, y_{it} on a vector of variables, $\boldsymbol{z}_{it}$ i.e.

$$y_{it} = \psi \boldsymbol{z}_{it} + v_{it} \quad \textbf{[1.7]}$$

ψ are the parameters to be estimated and v_{it} is the disturbance term [v_{it} ~N(0 1)].

The variables included in $\boldsymbol{z}_{it}$ should be those that significantly affect the work/no work decision, in other words significantly determine the probability of individual *i* being observed work (and receiving a wage) in time *t.* The wage rate w_{it} in the 'main equation' is observed iff:

$$\psi \boldsymbol{z}_{it} + v_{it} > 0$$

It is possible to estimate the correlation coefficient, ρ between u_{it} and v_{it} and where it ≠ zero, any standard regression on **[1.6]** will produced a biased estimate of β. Heckman (1979) derived a correction method that can eliminate the bias, via a two-step process, where **[1.7]** is estimated and a sample-bias variable, λ constructed, where:

$$\lambda = \frac{\phi(\psi \boldsymbol{Z})}{\Phi(\psi \boldsymbol{Z})}$$

λ is also termed the *inverse of the Mills Ratio*. A revised version of **[1.6]** is estimated incorporating λ, as follows:

$$[w_{it} \,|(\psi \boldsymbol{z}_{it} + v_{it} > 0)] = \beta \boldsymbol{x}_{it} + \rho \sigma_u \lambda(\psi \boldsymbol{z}_{it}) \qquad \textbf{[1.8]}$$

[4] This is, perhaps, an unfortunate, even outmoded term, as it implies that the individual must therefore be 'at ease', when in fact they may be active in household production (child rearing, looking after their spouse or partner and/or elderly dependents, for example).

Of course, the efficacy of the correction (i.e. its ability to 'unbias' β) is dependent upon the variables included in z, but a large body of literature has developed over the last four decades that has used the same choice variables. The process used in the returns Study will be discussed in more detail at the appropriate point.

The issue is discussed, and a correction attempted in the Returns to qualifications' Study (see sections **3.5.2** and **3.6.2**) and discussed, but is, in effect, assumed away in the Airlines' Study. The sample selection correction on the means-related wage rate returns had the effect of greatly increasing the premia, well in excess of the premia estimated without the correction and the quantile returns (see sections **3.5.2** and **3.6.3**). It is clear that any follow-up Study would have to consider the variables included in the correction, trialling different (combinations of) variables explain the reasons for non-participation in paid work.

A final issue to consider is that of *structural instability,* or rather, the absence of structural stability in time-series data, such as the Gender Based Violence (GBV) data that are used in Chapter Two. It is assumed, implicitly, that the parameters estimated from a regression using time series data apply to the whole period spanned by it. However, where one or more of the time series-related variables exhibit a sudden change in trajectory during the period, it may no longer be safe to assume this. The change in trajectory may be taken as evidence of a 'regime change', or policy change and that the underlying relationship between the dependent and included independent variables may have changed, significantly. In effect, there may be a set of parameters that applies before the change, and another that apply after (and, it is assumed. as a result of) the change. As will be described in Chapter Two, there is evidence of a change in policy relating to the judicial system's response to GBV in England and Wales between 2002 and 2013 that led to a significant increase in the number of arrests made on suspicion of GBV in 2013 (**Figure 2.1**). Thus, it is imperative that appropriate testing of the stability of the parameters estimated from the data (that purport to explain the number of GBV arrests) is performed.

The classic test for structural stability is the *Chow Test,* after Chow (1960) where the candidate model is regressed against the full dataset, the data before the regime change and, the data after the regime change. An F-test combining the SSRs from all three regressions is used, with the Null of stability assumed. A significant F-statistic allows the Null to be rejected and the conclusion that there was a significant regime/policy change at the selected point in time. While this test has considerable currency, it has several shortcomings, pertaining to the assumptions made about the model's error structure, its applicability to Panel Models, its ability to detect gradual but persistent changes in the data (as opposed to sudden 'spikes'). It also cannot be used to test early, or late-occurring changes in trend, as there is simply not enough data to estimate the model (i.e. $n_{after} < k$). As it happens, the spike in the aggregate GBV data takes place in the final year of the Study Period and implies that the Chow Test cannot be applied to even the aggregate cross-sectional data. The solution to this, as outlined in Greene, (2008, op.cit.) is to regress on the full data set and the data before the regime change and to compute an F-statistic based on the RSS of both, as follows:

$$F = [n_{after}, n_{before} - k] = \frac{(RSS_{full} - RSS_{before}) / n_{after}}{RSS_{before} / (n_{before} - k)} \qquad \textbf{[1.5]}$$

Although the GBV study is Panel based, this test will be carried out on the all-England/Wales aggregate data to give an indication of the stability of the parameters at that level. The method for testing the stability of parameters from a Panel starts with creating an 'indicator' dummy variable that takes the value '0' before the assumed regime change, and '1' thereafter. A regression that includes the indicator dummy and it interacted with the independent variables is carried out and a Wald Test performed to assess whether the pre- and post-change parameters are significantly different from each other. This might normally raise the objection that the assumption of when the change occurred introduces subjectivity into the test, but in this case, it is fairly obvious that a change took place in 2013, broadly mirrored throughout all members of the Panel.

1.4. Software employed

All the data were collated and cleaned using MicroSoft Excel. The Author was more confident in using it to combine survey waves, crimes' and airlines' data, etc and to run checks for possibly mis-reporting. Most of the graphs in the Thesis were produced via Excel as well, with the quantiles' results (Chapter Four) linking directly to the detailed regression tables produced by the other software. All the econometric modelling was done using Stata, variously versions 11, 13 and 15. Where necessary, user-written routines, such as *xtserial* and *qregpd* were downloaded and installed and some estimators (such as Canay's Quantile Fixed Effects Panel) written by the Author from first principles. Results' tables were written out from Stata into Rich-Text Format files and incorporated and edited in Word. All output, including graphs of residuals, etc, are contained in an Appendix to the Thesis.

References used in Chapter One

Baltagi, B.H., (2013), 'Econometric Analysis of Panel Data', 5th ed., Wiley, ISBN 978-1-118-67232-7

Becketti, S., Gould, W., Lillard, L. and Welch, F. (1988), 'The Panel Study of Income Dynamics after fourteen years: an Evaluation', J. Labor Economics, 6, 472-92

Breusch, T. S. and Pagan, A. R. (1979). 'A Simple Test for Heteroskedasticity and Random Coefficient Variation', Econometrica. 47, 5, pp 1287–1294

Chow, (1960), 'Tests of Equality Between Sets of Coefficients in Two Linear Regressions', Econometrica, 28, 3, pp 591-605

Drukker, D. M, (2003), 'Testing for serial correlation in linear panel-data models', Stata Journal, 3, pp 168–177

Durbin, J. and Watson, G. S. (1950). 'Testing for Serial Correlation in Least Squares Regression, I', Biometrika, 37 (3–4): 409–428.

Frisch, R., (1933), 'Editorial', Econometrica, 1, pp1-4

References used in Chapter One

Gasteen, A., Houston, J. and Asenova, D., (2000), 'UK Monetary Policy, Earnings Growth and Labour Market Structure 1989-1998', Economic Issues, 5, 2

Greene, W.H, 'Econometric Analysis', 9th Edition, Pearson-Prentice Hall, 2008, ISBN 978-0-13-513245-6

Heckman, J.J., (1979), 'Sample selection bias as a specification error', Econometrica, 47(1): 153–61

Houston, J. and Gasteen, A., (2007) 'Employability and Earnings Returns to Qualifications in Scotland, 1999-2001', Regional Studies, 44, 4, pp1-9

Houston, J., (2016), 'Estimating the Employment and Earnings' value of Qualifications in the Scottish Labour Market 2006 – 2015', report to the Scottish Qualifications Authority

Houston, J., Gasteen, A. and Asenova, D., (1999), 'Structural change in the Scottish Labour Market, the tax take and the new Parliament's Income Tax-varying powers', J.Public Money and Management, 19, 4

Houston, J., Gasteen, A. and Asenova, D., (2001), Labour Market Flexibility in Scotland and The New Parliament's Income Tax Varying Powers, Regional Studies, 35, 4, June 2001

Houston, J, Gasteen, A. and Davidson, C., (2005), 'Earnings Returns to Further Education', in 'Living in Scotland: Findings from the British Household Panel Survey', Ermisch, J.F. and Wright, R.E. (eds.), Bristol, Policy Press, 2005, (ISBN 1 v86134 593 3)

Johnson, K. and Houston, J. (2018), 'Does Female Police Leadership Affect Gender Based Violence Arrests?', Int. Review of Administrative Sciences 84, 1, 3 – 20

Longley, J., (1967), 'An appraisal of Least Squares programs from the point of the user', J.of the American Statistical Association, 62, pp 819-841

References used in Chapter One (continued)

Ramsey, J. B. (1969), 'Tests for Specification Errors in Classical Linear Least Squares Regression Analysis', J. Royal Statistical Society, Series B, 31, 2, pp350–371

Wooldridge, J. M. (2002), Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press, ISBN 978-0-262-23219-7

2. Representative Bureaucracy: Gendered Policing and Gender-based Violence in England and Wales

2.1. Introduction & Motivation

This Study tests the hypothesis that private- and public-sector organisations ('bureaucracies') better serve their clients only when they make it their primary aim to provide a service that best serves those needs. It is, perhaps axiomatic, that it will be easier for a bureaucracy whose staff composition resembles the client population, in relation to some key characteristic (age, gender, race, for example) to provide this improved service. However, this may be to deny that employees who do not explicitly have the same relevant characteristics as the client, but are at least capable of *sympathy*, or even *empathy* for them, would not strive to meet the clients' needs as well as those that do (male feminists, for example). It is perhaps even more axiomatic that those services that require a high degree of education and training can only be delivered by people who are very different to their clients (for example, medical care to mentally-retarded children or adults). They may be motivated by more than the financial and status rewards of their efforts, deriving utility from improving their clients' lives.

Sociologists and Administrative Scientists have tested the hypothesis in a wide variety of settings: the focus in their literature has been mostly on public bureaucracies and national political systems and the criteria used to measure their quality focussed on *social* outcomes (inclusivity and social justice), rather than explicitly *economic* ones. Economics can, of course, include such factors in considering the *welfare* implications of bureaucracies making efforts to identify and rectify its shortcoming in this respect: relating to both the direct clients, its employees and Society in general.

This Study tests the hypothesis that making the 43 separate English and Welsh Police Forces more represented in terms of the gender balance of its workforce, resulted in them providing a better service to a section of the general public, namely, women. Service, in this context, is the number (or rate) of arrests on suspicion of having committed criminal acts of *gender-based violence* (GBV) upon mostly, though not exclusively, women. This Study takes the stance that a *higher number*, or a *greater rate*, of arrests for GBV implies a better quality of service, insofar as the Service is more focussed than before on tackling that type of crime. Any demonstration of a positive association between the female composition of the Forces and the number or the rates of GBV arrests would indicate the existence of a positive policy instrument (the recruitment and promotion of women into the Police), leading to a positive welfare outcome (reduced propensity for GBV) and labour market outcomes for women (increased demand for female workers). We should also note however, the potentially adverse labour-market outcomes for both the males who might otherwise have been recruited by the Police and the males who acquire criminal records as a result of the focus on their crimes.

There have been several, mainly qualitative studies, many using 'vignettes', on the differences between female- and male officers and aspects of police service provision. Some of the more relevant literature is summarised in **2.3.3**. One thing that these studies share in common, is that they purport to test a particular sociological or organisational axiom, involving a narrowly-focussed, highly-specific arena, using qualitative data gathered over a short timeframe, analysed with largely descriptive statistical methods. By their very nature, they act as an encouragement to other researchers to re-test the axiom using another arena of specific interest to them. Previous literature is used to establish the hypothesis being tested, and results and conclusions contrasted with this earlier literature. The author has made a modest contribution to this literature (Miller and Houston, 2018), adopting a quantitative, multivariate approach to assess the effect of gendered policing in the junior and senior Ranks, on GBV arrests in England and Wales. That Study assembled data to test the hypothesis that increasing the numbers of female officers at either or both levels increased the number of arrests made for gender-related crimes. It augmented the then-existing literature by placing more emphasis on linking and analysing published data and constructing a simple multivariate model to assess the impact of change in workforce composition on arrest made for these crimes. This Study exploits the same data but develops and tests far more rigorous econometric models. A major element of this increased sophistication is the contrasting of Pooled- versus Panel-based estimators in the identification of the specific effect of female police numbers on GBV arrests. To date, there has been no other published study that adopts this approach on the English and Welsh police, nor anywhere else in the World. The chapter is structured as follows:

- An outline definition of what constitutes acts of Gender-based Violence (GBV);
- A review of the literature on Representative Bureaucracy and women in the Police and quantitative models of crime;
- The scope of the Study;
- An outline of the English and Welsh Police Forces and the overarching justice process;
- The data used and the econometric strategy: variable selection; models' specification, pre- and post-estimation testing;
- The results and future work.

2.2. Gender-based Violence (GBV):

2.2.1 International Context and Definitions

Russo and Pirlott (2006) assert that Society consciously organises itself in ways that condones, even encourages violence by men against women, through combinations of religious dogma, associated marriage and property rights and social norms regarding the allocation of work in the household. Physical strength and social position are clearly factors in this process, whereby males have more scope to use their physical power and pre-eminent positions in Society to dominate females. A multitude of studies have demonstrated that men have significantly more aerobic power in their upper- and lower bodies than do 'equally-conditioned' women (e.g. Bishop, et.al, 1987). This, by itself, gives males an inherent advantage over females, either in the context of face-to-face (domestic) conflict situations or more broadly, in Society where production processes require men's superior physical strength.

Despite, or perhaps because of, technological development reducing the need for this workplace strength and the ever-reducing tolerance most modern societies have for all forms of discrimination, GBV is now regarded as a major, but correctable fault-line running through societies. The European Institute for Gender-based Violence (EIGE, 2014) declares GBV to be:

'.. one of the most persistent human rights violations of modern times…. connected with the unequal distribution of power between women and men.'

It recognises the conflation of GBV and violence against women in general, by stating that:

'The terms gender-based violence and violence against women are often used interchangeably as most violence inflicted against women and girls is based on their gender.'

The Council of Europe (2011) also equates GBV with violence against women, as:

'… all acts of gender-based violence that result in, or are likely to result in, physical, sexual, psychological or economic harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life.'

In a summary of the various publicly-stated definitions of GBV, the Commission of the European Union (2012) defined GBV as:

'… violence that is directed against a <u>person</u> because of that person's gender … that affects persons of a particular gender disproportionately.'

The United Nations in its 1993 'Declaration on the Elimination of Violence against Women', stated clearly that it viewed GBV as being almost exclusively a problem of violence perpetrated by males against females, i.e. women and girls (UN, 1993). Its Population Fund Organisation (UNPF) devotes almost its entire website on Gender-based Violence to discussions on the various forms of violence perpetrated on females by males. [5] Other organisations, such as the Overseas Development Organisation, similarly cite this as a problem affecting mostly women, but point out that males can also be affected, though generally at the hands of other males, particularly in times of conflict. [6] These crimes are also gender-related, in the sense that the males are perceived as a military threat and are singled out from the females for (illegal) imprisonment, or execution (such as the conflict in Kosovo in the 1990s and subsequent and ongoing conflicts in parts of Africa and the Middle East).

The common theme in all these definitions is that the *victim's gender* is the key, perhaps sole, factor in their treatment at the hands of their perpetrators: it is not necessary, therefore, that the perpetrators are of *different genders*. It is also possible that acts of GBV could be carried out by women on men, possibly in public conflict scenarios, but also in intimate, domestic situations. Some studies report instances of domestic violence perpetrated by women on men (e.g. Frieze and McHugh, 2005). Anecdotally, these are assumed to be especially under-reported, due to an exaggerated social stigma stemming from men's shame in being physically assaulted by a woman.

It is clear that females are much more likely than males to be subjected to GBV perpetrated by males. For example, the United Nations quotes Vanderschueren (2000) who estimates that women are, on average, twice as likely as men to be victims of violence, including GBV (United Nations-Habitat, 2006). Surveys, such as those undertaken by the World Health Organisation (WHO, 2013) claim that 36% of women alive in 2005 had been subjected to some form of intimate/sexual violence, with Europe exhibiting the lowest rate (27.2%) and Africa, the highest (45.6%). What is less clear, are the comparators for men, but the sheer weight of such surveys, and their predisposition to regard this, from the outset, as a 'female' issue, suggests that this is predominantly an issue relating to way men mistreat women

[5] http://www.unfpa.org.gender-based-violence: accessed 2nd June, 2015
[6] http://www.odi.org/comment/8502-male-gender-based-violence-conflict-humanitarian-response: accessed 2nd June 2015.

throughout the World. A standard world-wide classification of different categories of GBV has emerged. *Inter alia*, the significance of this classification is to align each one with specific national- and international legislation outlawing specific activities.

2.2.2 The English/Welsh perspective

Table 2.1 summarises the English/Welsh legal definitions, categorised in alphabetical order by type, and the current enabling legislation in England and Wales. [7] While this table focusses on English/Welsh definitions and legislation, most articles and reports examining aspects of GBV refer to this classification (including Russo and Pirlott, 2006, op.cit). **Figure 2.1** plots the total number of recorded arrests for GBV offences in England and Wales between 2002 and 2013, using published Home Office data.

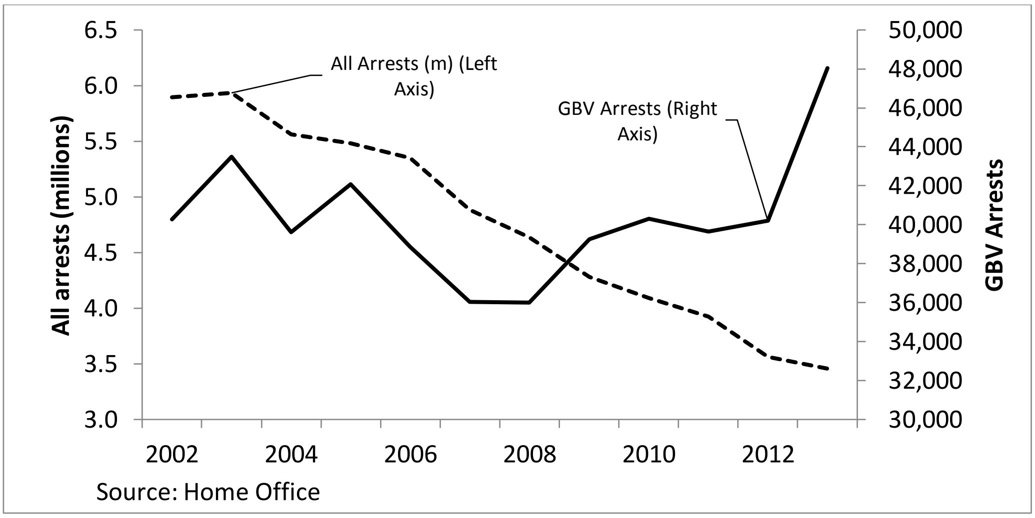


Figure 2.1 Number of recorded arrests for GBV offences in England and Wales (2002 to 2013)

[7] Source: Extracted by the Author from http://eige.europa.eu/gender-based-violence/legal-definitions?c%5B%5D=345 (accessed 16th February 2015).

Table 2.1
Types of GBV and English/Welsh Legal Definitions

Type of GBV	Legal Definition	England/Wales Legal Sources
Intimate Partner Violence	Any incident of threatening behaviour, violence or abuse (psychological, physical, sexual, financial or emotional) between adults who are or have been intimate partners or family members, regardless of gender or sexuality.	The Domestic Violence Act 2004/Protection from Harassment Act 1997
Rape	A person commits an offence if (a) he intentionally penetrates the vagina, anus or mouth of another person (B) with his penis, (b) B does not consent to the penetration, and (c) A does not reasonably believe that B consents.	Sexual Offences Act 2003, Section 1.
Sexual Assault (excluding rape)	"Sexual assault" is the sexual touching of a person without their consent: A commits an offence if (a) he intentionally penetrates the vagina or anus of another person 'B' with a part of his body or anything else, (b) the penetration is sexual, (c) B does not consent to the penetration and (d) A does not reasonably believe that B consents. Person A commits an offence if (a) he intentionally touches another person B, (b) the touching is sexual, (c) the person does not consent to the touching and (d) A does not reasonably believe that B consents.	Sexual Offences Act 2003 UK, Section 2 and 3
Sexual Harassment	Person (A) harasses another (B) if (a) A engages in unwanted conduct related to a relevant protected characteristic, and (b) the conduct has the purpose or effect of (i) violating B's dignity, or (ii) creating an intimidating, hostile, degrading, humiliating or offensive environment for B. A also harasses B if (a) A engages in unwanted conduct of a sexual nature, and (b) the conduct has the purpose or effect referred to in subsection (1)(b) of the Act.	Equality Act 2010, Section 26A
Stalking	Stalking is a term used to describe a particular kind of harassment. Generally, it is used to describe a long-term pattern of persistent and repeated contact with, or attempts to contact, a particular victim.	Protection from Harassment Act 1997 (PHA), Protection of Freedoms Act 2012.S125(2) of the SOCP Act 2005

While the number of women in England and Wales rose from 26.3 million in 2002 to 28.9 million in 2013, there was an increased propensity to makes arrests for GBV. In 2002, it took an average of 653 women to generate a single arrest for GBV: by 2013, it required only 593 women. The percentage of arrests made that were for GBV doubled from 0.7% to 1.4% in the same period. [8]

In the context of a population of more than 26 million females, and total arrests of at least 3.5 million per annum, the number of GBV arrests is relatively small. However, even if we regard one arrest for GBV as unacceptable, the concern has to be that these observed levels may mask a much larger number of unreported GBV incidents or reported incidents that do not lead to an arrest. A combination of victim-reluctance and an unwillingness of a historically male-dominated police force to arrest suspected perpetrators may have acted to significantly understate the extent of the problem.

An examination of the components of GBV arrests (**Figure 2.2**) reveals that there was a steady rise in arrests for various offences against (mostly) female minors. This might be due to any combination of changes to 'offender preferences', a rapidly heightened societal intolerance for such crimes, or national initiatives designed to bear down on such offences, and/or an increased propensity for victims to come forward and report the crime perpetrated against them. It is evident that although the reported numbers of arrests for GBV in England and Wales during the period was relatively small, it was growing towards the end of the Study period. Note again, that these figures probably mask significant and persistent under-reporting and/or a reluctance by Police Officers to arrest a suspected GBV perpetrator (this will be considered later).

[8] Note that the number of arrests for GBV continued to rise after the Study Period to reach 89,000 by 2016, 2.1% of all arrests made. Clearly, this 2013 'spike' heralded a significant change in the Judicial System's attitude towards GBV.

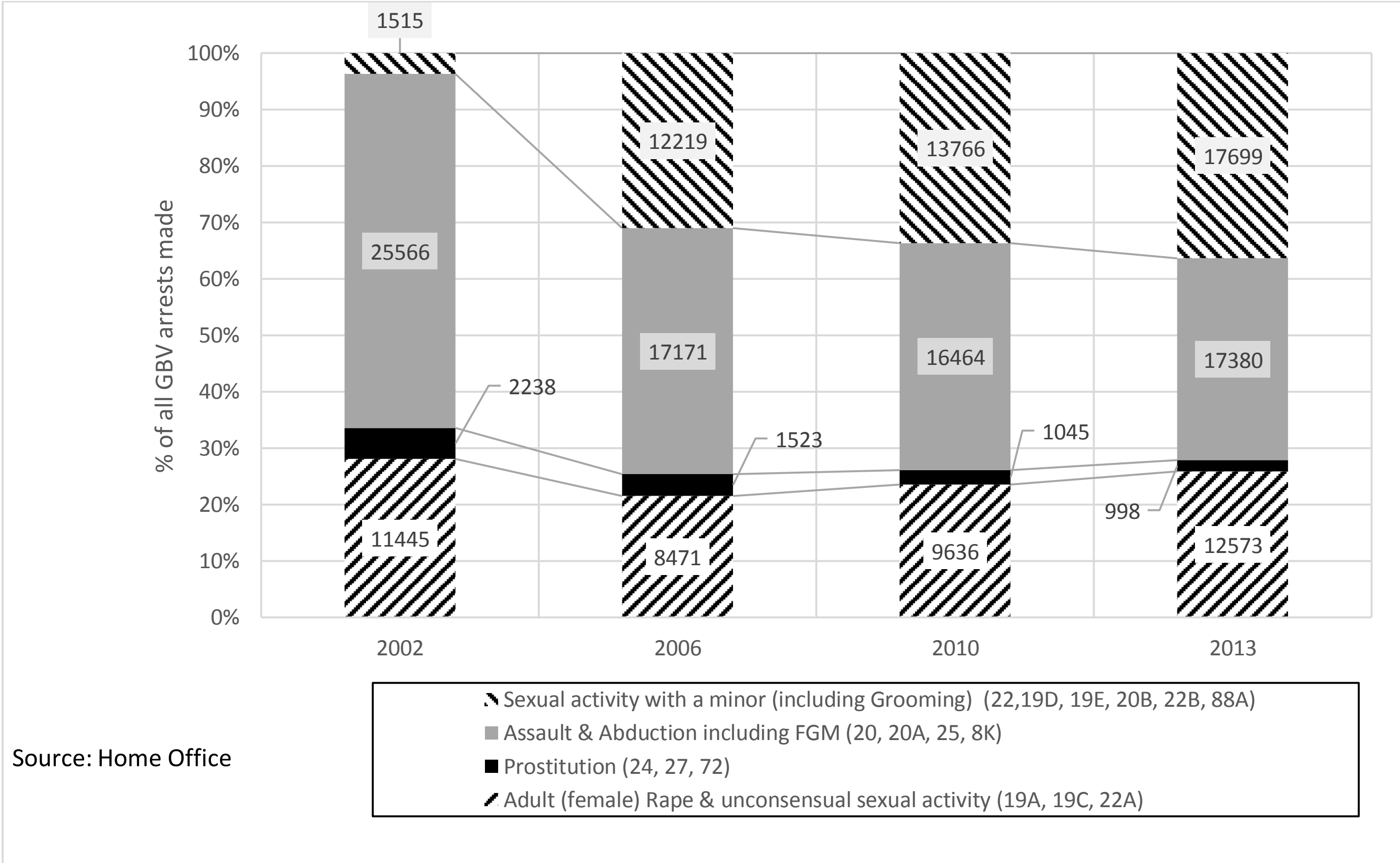


Figure 2.2 Proportions of GBV Arrests England & Wales 2002 – 2013

It can also be seen that the underlying composition of the annual totals changed between 2002 and 2013, with a much-increased focus on sexual offences against minors, and less on Prostitution and Sexual Assault. As this Study looks only at total GBV Arrests, no detailed causal explanations are needed here for these changing propensities, except to note that these changes may be as a result of changes in 'taste' for specific forms of GBV (for example, the Internet makes on-line 'grooming' of minors much easier now than was previously the case). In addition, specific National initiatives may have resulted in crackdowns on specific sub-categories (like Grooming) emanating from newly-acquired Societal disgust, leading to political pressure and, from there, to changes in policing emphasis.

Somewhere, amongst all this, may be a change in Officers' tolerance for such offences, emanating from the changing gender composition of the Forces: identifying the existence and significance of this is the overarching aim of this Study, in particular whether or not the conclusions are dependent upon either or both the model specification and estimator employed.

However, it is the case that there was a change in Government Policy relating to GBV in general in England and Wales and not just a (possible) change in Societal attitudes to it. In 2010, the Home Office published a document laying out the (then new) UK Government's strategy to pertaining to Violence against Women and Girls (VaWG) (Cabinet Office, 2010). This, and the associated Action Plan document acknowledged the change in tone and aspiration emanating from the 1993 UN Declaration (UN, 1993, op.cit), and laid out a framework for action against VaWG. The Action Plan consisted of three main strands:

1. <u>Preventing GBV</u>, by challenging and changing the attitudes, behaviours and practices in UK Society, as well as raising awareness of the services available to help those vulnerable to GBV to access interventions designed to prevent GBV from taking place.
2. <u>Providing adequate support</u> where and when an act of GBV has actually occurred (or alleged to have occurred). *Inter alia*, Local Authorities were

required to provide adequate refuge facilities to which (alleged) victims could seek safety. Interestingly, priority was also to be given to training 'frontline partners' to recognize VaWG in all its manifestations. This, perforce, includes attending Police Officers, insofar as they are often the first point of contact with Domestic GBV, in particular.

3. Altering Justice Outcomes and Risk Reduction. The aim here is to create an environment where the (alleged) victims feel more confident that the criminal justice system will respond appropriately where they have come forward to report acts of GBV perpetrated against them. Also, priority was to be given to assist previously-convicted GBV perpetrators from *reoffending*.

With reference to the third strand, the Action Plan spelt out 11 separate changes to legislation and justice outcomes, including the introduction of Domestic Violence Protection Notices and Protection Orders, to cover the period between the (alleged) GBV and the stage at which a prosecution may actually be made. Alleged victims were also to be supported through what would otherwise be a harrowing and intimidating Court Process where a Trial is actually initiated. In doing this, the Government's Home Office was implementing recommendations made in the 2010 Stern Review (Home Office, 2010) that was charged with looking at the way that (alleged) Rape Incidents were handled by the Judicial System in England and Wales. The Review found that, historically, Police Forces had demonstrated '… disbelief, disrespect, blaming the victim, not seeing rape as a serious violation, and therefore deciding not to record it as a crime.' It did acknowledge that *some* Forces had, since 2002, trained and deployed 'rape specialists' to handle such cases, and to otherwise alter the specific steps taken immediately following the (alleged) incident, and allowing the victims to give evidence video-recorded, *in camera*. The Review specifically recommended that every Force establish one or more Sexual Assault Referral Centres by 2011 and to work in partnership with the local NHS Trust(s), and that each Force's Senior Officers strive to further integrate and otherwise improve its service with respect to such (alleged) crimes.

In 2012, the UK Government signed the European Council Convention on combatting violence against Women and Domestic Violence (‘the Istanbul Convention’) that espoused the same aims as the emerging domestic policy (Integration – Prevention – Protection – Prosecution), that led to the Preventing and Combating (sic) Violence against Women and Domestic Violence (Ratification of Convention) Act, 2017 being passed as an Act of Parliament in England and Wales. The UK Government will also bring forward a Domestic Abuse Bill that will define Domestic Violence as a specific crime (as opposed to ‘ordinary’ Breaches of the Peace).

It is noticeable in **Figures 2.1** and **2.2** that there appears to have been a ‘Policy Lag’ where the actions taken, unilaterally by some Forces in 2002 had some immediate impact, while the more concerted and co-ordinated, National Action against VaWG started to impact in 2013. This is to be expected, as Police Forces and the other agencies had to train their staff, and new procedures learned and adopted. The (alleged) victims, for their part, needed time to realise that the improved support is available to them, and that they should be more willing and able to come forward and report the crime perpetrated against them. Clearly, the co-ordinated Policy is making a real impact, given the increased number of incidents reported.

This, however, raises the possibility of there being an Econometric Issue that may affect the Study, namely *Structural Instability* brought about by a significant ‘regime’ (policy) change. The effect may be to, in effect, split the Study Period into two (or more) ‘eras’: one before the policy change, the other after. Accordingly, there may actually be two, distinct, sets of parameters, each one applying to its ‘era’, and not just one applicable to the whole period. This issue was discussed, in outline in 1.3 and will be tested for in the model estimation sections of this chapter.

2.3. Representative Bureaucracy

2.3.1 The Concept of a Representative Bureaucracy

Through History, societies developed systems of organisations through which the production, distribution and consumption of goods and services were enabled. As the systems became larger and more varied, covering larger populations and territories, the need to administer their functioning became ever more important. Organisations developed *bureaucracies* to organise production and delivery, and to ensure that sums due and sums owed were recorded and paid or collected. The word '*bureaucracy'* is generally regarded as having come into everyday use during the 18th Century and is a combination of the French word *bureau*: a desk, or a physical or conceptual 'office', and *kratos*: Ancient Greek for political rule or *power*. The idea, however, of a conceptual 'office' exercising power and influence over people is much older and can be traced back to the advent of writing and the evolution of tribal societies into city states. The desire to build empires, and the consequent wars that resulted, required an effective bureaucracy to conquer people, maintain rule over them and, of course, to levy taxes. From the British perspective, the Romans were probably the first to impose on them, a powerful, hierarchical and centralised bureaucracy. The British perfected the Art, with the development of their 'Home' and 'Foreign' Civil Services, of a type now referred to as a *Weberian Civil Service* after the Sociologist Max Weber, who advocated the use of intensely hierarchical bureaucracies as the 'best way' to organise human activity (Weber, 1922). He is still credited, today, as being the first to study organisations from an academic perspective, giving rise to what has become known as, *Organisational Theory*, now referred to as *Administrative Science* (AS).

Both public- and private organisations require a bureaucracy to allow them to function at their strategic and operational levels. The way in which bureaucracies do this is at the core of AS, and a detailed consideration is outwith the scope of this Study. Nevertheless, it can be said that organisations differ in the way they structure themselves, as well as the extent to which they centralise, or devolve power through their levels. The structure may be *voluntary*, in the sense that it embodies the instinct and personality of the person that created it. It may instead be objectively *necessary* (e.g. 'command-and-control', as in the Armed Forces), or *maximally-devolved* (such as in the Creative Arts). As might be expected, the general structure of the Police Forces is very much of the rigidly hierarchical, 'command and control' type.

Weberian bureaucracies were criticized by some of those who came after, such as Merton (1957), who bemoaned their tendency to become detached from the client's objectives and needs, and to seek to protect themselves. In effect, they function largely to allow themselves to continue to function, exploiting their natural advantage of having control of resource allocation. [9] However, it might be argued that there will always be natural checks-and-balances that should bring the bureaucracy back to its primary purpose. Private-sector organisations exist to sell products and services at a profit, and deliver to their owners, these profits. The cost of the bureaucracy is, perforce, a component of total cost, and reduces profits. The owners have an incentive, therefore, to optimise the cost-effectiveness of bureaucracy, not allowing it to become overly bloated. Public-sector bureaucracies, on the other hand, exist to serve particular sections of Society and are not otherwise subject to the same competitive pressures, having no profit motive. Therefore, it is argued, these bureaucracies are more able to burgeon and then maintain a bloated state. Only in times of austerity and/or the election of governments naturally hostile to Public Sector provision, may they be pruned back and re-focussed onto their primary purpose (though not necessarily congruent with its Client's needs).

Since the time of Weber and Merton, AS has progressed to consider the concept of *Representative Bureaucracy*, particularly in the Public Sector. This refers to the extent to which a particular bureaucracy succeeds in serving Society, in particular its willingness to meet Society's various needs and objectives. Even with ever-increasing use of Information Technology, bureaucracies function through their

[9] This involves to the so-called 'Parkinson's Law' which asserts that ' ...*work expands so as to fill the time available for its completion*.' (Parkinson, C.N. 1958).

employees, who bring to it, their own talents, ambitions, attitudes and prejudices. Where the bureaucracy is long-established, it will have already developed its own 'culture'. As a result, workers with particular characteristics and associated values will be attracted to seek employment with it. This helps to perpetuate, even deepen, the prevailing culture and make it more difficult for others not sharing its values, from gaining employment with it. All the while, the bureaucracy develops an 'attitude' to the Client and offers it the service it believes to be appropriate. The bureaucracy may become closed to this Society, in the sense that it is no longer interested in learning what its current needs are, how well it is serving even the perceived needs, or even whether it needs to better align its objectives, when it is aware of a mismatch. The bureaucracy may even develop an active 'dislike' of this Society, and *vice versa*. Society, in these circumstances, is said to be *under-served* by the bureaucracy. In the presence of a significant democratic deficit, whereby Society cannot force the bureaucracy to serve it better, this perceptual gap will persist and even widen through time. Where democracy is strong, Society, as the Electorate, can remove the bureaucrats directly, or at least, the politicians that appointed them.

It is argued that one way in which a bureaucracy may remain open to Society, is to ensure that the composition of its workforce reflects that of the Society it purports to serve. In the absence of practical reasons for not doing so precisely (such as in the Armed Services), the composition should take into account the gender, race, religion and age structure of the Client. Naively, perhaps, it might be assumed that this, alone, will be sufficient to alter the bureaucracy's behaviour and to bring it (back) into step with the Client (Meier and Nigro, 1976). Thus, a bureaucracy that wishes, or is forced to serve Society better, needs to consider the composition of its workforce, and for them to (re)align the bureaucracy's priorities.

Johnson and Houston (2018, op.cit.) summarised the concept of representative bureaucracy, citing the work of Mosher (1982), Riccucci and Saidel (1997) and Bradbury and Kellough (2011), all of whom distinguish between *passive* and *active* representation of societal groups within it. *Passive representation* refers to the efforts that bureaucracies make to align the makeup of its workforce and the Client, with reference to their corresponding age, ethnicity and gender structures. The mere presence of employees in the correct proportions is assumed to be sufficient for the bureaucracy to be representative. Where a particular group is numerically under-represented, there needs to be a period of skewed recruitment in an effort to realise the correct composition. However, these employees may not be in any position to alter the bureaucracy's culture sufficiently quickly, given that most of them are likely to have been recruited into its lower levels. For the time being at least, the representativeness is *token* and not certain, on its own, to better align the bureaucracy with the Client.

Active representation, on the other hand, describes the efforts made to ensure that the Client's specific needs are constantly at the heart of the bureaucracy's mission and aims. It is neither necessary nor sufficient to assume that the presence of employees in correct proportions will guarantee a proper service, only that its employees must be focussed correctly on satisfying the Client's objectives. Internal democracy is sufficiently established and strong to ensure that the employees are constantly and appropriately focussed and that a minority clique is unable to enforce its incompatible will on the bureaucracy and therefore, the Client (Thielemann & Stewart (1996) cited in Johnson and Houston, 2018, *op.cit*). This may be achieved, via programmes of compulsory race- and gender-awareness training, for example.

Other research in this area has considered whether or not *passive* representation results, inevitably, in *active* representation, in particular as a critical mass of a particular minority group is attained, and representatives from among them are promoted to resource-allocation and policy-making levels. In addition to active representation at the upper, policy-making levels, bureaucracies that devolve more power down the hierarchy, to those who directly interface with their clients, are more likely to act in harmony with them (provided that those interfacing are not significantly prejudiced against the clients). The same applies to bureaucracies that adopt a flatter, wider structure. Lipsky (1980) discussed the notion of the empowered *street-level bureaucrat*, playing a significant role in making policy effective, by actively implementing it. For example, when considering women as a minority, Keiser, et al (2002) concluded that US Public schools with a greater percentage of female teachers, resulted in better test scores for girls. They stressed the importance of the teacher, as a client-facing bureaucrat, identifying more with the female pupils (and *vice versa*), altering the relationship between them and resulting in a better learning environment and performance. The corollary is that *unempowered* street-level bureaucrats will be less able to facilitate active alignment, even where they either share or are sympathetic to the Clients' needs: they are merely 'obeying orders'.

2.3.2 Critical Mass in Society and Bureaucracies

An important issue to be considered is the ability of an under-represented group to influence, in any significant way, the bureaucracy's objectives. Adherents of the 'passive school' assume that the mere infiltration of a particular 'type' will guarantee the necessary changes, albeit at a slower pace. Those who do not believe that this is inevitable, consider the concept of a *critical mass* of employees as being key. In doing so, they are borrowing from nuclear physics, where a critical, or minimum amount of fissile material is necessary to trigger self-sustaining nuclear fission. Once that critical minimum has been achieved, then a 'big bang' is triggered and everything changes irrevocably. In the organisational world, the equivalent 'Big Bang' ushers in a period of rapid, permanent change in the bureaucracy's culture. Prior to that, however, changes in the demographic of the workforce towards representativeness yield little or no, even regressive, changes: the argument being that the minority is simply too small to make any palpable difference. Taken to its extreme, some of the minority may even adopt the characteristics of the majority ('go native') in an effort to gain acceptance and personal advancement within the bureaucracy.

Much of the literature in this area is in the political science arena, where the focus has been on the effectiveness, or otherwise, of females in particular legislative (parliamentary) arenas. In a study of women in Science, Etzkowitz, et al, (1994) quoted a 'rule-of-thumb' of at least 15% workplace representation as a critical mass, though warning that any tendency for the growing minority to simply form its own shadow groups could thwart the idea of a single, unified bureaucracy ever emerging. In a study of progress into Senior Management at a number of Australian Universities, Chesterman and Ross-Smith (2006) quoted a figure of 35%, at which point the management culture began to change. Women in the lower echelons felt they had a clear career path if they wished to pursue it. Kanter (1977) argued that 25% (female) representation is 'sufficient for them to have enough mutual support in the organisation'. The fact that these quoted percentages are less than 50%, indicates that female minorities do not feel the need to wait until they have achieved numerical equality, but can already feel sufficiently emboldened to reveal their true selves, in the belief that they will always have a sufficient constituency with which they can at least confide.

Dahlerup (1988) examined the tacit acceptance and promulgation of the notion that female elected parliamentarians required to be at least *30%*, before a significant change in the legislative priorities and *modus operandi* (of the New Zealand Parliament) would be realised. She concluded that the empirical evidence for this asserted level is tenuous, and possibly little more than folklore. Even assuming that there was such a thing as a 'lift-off' percentage, it could not even be assumed that 50% would be the point at which attitudinal change would be affected. She linked the general position of women in wider Society (~50% of the population, but in an otherwise subservient position to men) to the subservient position that even a numerical majority of women in an organisation would remain in. [10] In tracing the emergence of this 30:70 lift off, she cited Kanter (1977, *op.cit*.), who identified four stages through which an organisation may transit as the minority-majority ratio changes (the indicated ranges assume that 50:50 is the appropriate balance, insofar as this reflects the actual structure of the underlying population that the bureaucracy is purported to serve):

[10] Similar issues are noted for *racial groups*, albeit with different percentages applying.

- Uniform Group (0:100 range) i.e. only one 'type' in the group, with a single, all-dominating culture. There is no inter-group friction, by definition, but such a group is most likely to have developed antipathy for its Client, where they do not share any of its characteristics, or are otherwise unable to demonstrate any sympathy for its members.
- Skewed Group (15:85 range) The minority are dominated by the majority's culture and are regarded as being *tokens*, tolerated by the majority and perhaps useful as a propaganda tool. It is at this range that the minority are at most risk of being discriminated against and also likely to 'go native'. There is also the risk of 'kickback' as the majority perceive a direction of travel towards balance and aim to stem it, while they are still in the ascendant.
- Tilted Group (35:65 range) The minority is now large enough to form 'protective' alliances and to otherwise effect change, to the *modus operandi* and priorities within the organisation. The concept of tokenism is replaced by one of (reluctant) acceptance by the majority and acknowledgement of its members as important members of the bureaucracy. There is the danger that the minority may form a shadow group and frustrate progress towards a genuinely unified bureaucracy.
- Balanced Group (45:55 - 50:50 range) Taking into account the propensity for Societal discrimination against the former minority to influence the power balance within the organisation, it may now be in a position where an employee's gender will no longer be a factor, though other factors (such as age, race, social class) may come into play.

Subsequent research has gone on to document the disappointment, even puzzlement that no significant change, or 'lift-off' has been observed as the bureaucracies studied moved towards 'balance'. Grey (2006) summarised several studies on women politicians and concluded that the broad ranges Kanter advocated were very specific, and that, in fact, different critical masses may apply at different breakpoints. For instance, it may be that at something like 30:70, women politicians may be able to 'feminize' the legislative programme to some extent (to the benefit of women in Society), but the way in which the Assembly debates and otherwise operates may not change (for instance, through the adoption of 'family friendly' hours). For that to change, a ratio higher than 30:70 would be required. The result is that change is less than expected, with no real idea of what point 'real' change will be experienced, if ever. The danger is that this disappointment translates into disillusion and a withdrawal by the minority from the organisation. She examined the literature that assumes that minority agents will actually actively represent their Societal Counterparts, once elected. She also noted that newcomers to the Uniform- or Skewed bureaucracies may crave acceptance by the majority, eschewing the formation of alliances with other tokens, even exaggerating the characteristics of the majority, acting against the incumbent tokens. This helps to reinforce the majority's cultural grip on the bureaucracy and to inhibit its progress toward the Tilted- and Balanced stages. In politics, the practice of 'whipping', whereby elected members are often forced to take the Party Line against their, and their electorate's wishes, exacerbates this issue. While this process may not formally operate in other bureaucracies, minorities may feel pressured into 'towing the line' and supporting the will of the majority. In the context of this Study, there is therefore the possibility that the increasing percentage of women in the Police Force, will not result in an improved outcome for female members of the Public, and may actually make it worse for them: dependent upon what 'Kanter-stage' each Force has reached: i.e. it cannot be assumed that increasing the number and proportion of female police officers in the Force will guarantee an increase in the number of arrests made for GBV.

2.3.3 Gender in the workplace and the particular effect of Women in the Police

There is an extensive literature on gender differences in the workplace. Some of it has focussed on differences in pay and working hours and mode between women and men. Other literature has focussed on the sociology of gender in the workplace, looking at gendered sectoral- and occupational herding, work mode selection, as well as collegial work-based relationships. A consensus has emerged that men prefer competitive environments and linear, 'systemized' thinking, whereas women exhibit preferences for more socialised, empathetic thinking (Kanter, 1977, *op. cit.*). As a consequence, males are attracted to bureaucracies where competitiveness and rules-based procedures are emphasized and valued. Status within these organisations is determined by the levels of aggression and competitiveness displayed at key moments in their careers. As a result, males tend to progress more quickly to the higher levels of the bureaucracy than females (and other 'feminized' males): typical examples would be the Military and Foreign Exchange trading rooms. Females, on the other hand, select into roles and organisations where 'caring', consensus, inclusion and consultation are more important, such as nursing, counselling and early-years teaching. Debate surrounds the reasons for this bifurcation, housed within the overarching Nature-Nurture debate.

There has emerged, however, a counter-argument that purports that different roles demand different *blends* of systematic- and empathetic thinking, but that men and women are first attracted to the *job*, then the bureaucracy, second. Thus, the men and women that work for it are actually quite alike, having been screened by the employer prior to entry and selected on the basis that they are the 'right kind of person'. This *Person-Fit-Environment Theory* has been used to support the idea that 'feminised' men will work happily in organisations alongside women, where empathy is valued, while 'masculine' women will locate elsewhere, alongside men in these systematized roles. The consequence of this argument is that gender composition is therefore irrelevant to the way in which a 'service' is delivered, given the personal self-selection and pre-entry filtering processes that have taken place. Several studies have tested this hypothesis, for example, Nielsen (2014) who examined the extent of inter-gender differences amongst 1,320 public sector employees in Denmark, employed in five different (professional) occupations. Strong, though not universal, support was found for the self-selection and employer-screening process reducing and, in some cases, eliminating perceptual and work-practice differences between men and women in the same occupations.

Turning to policing, an easy assumption to make would be that street-level, male officers working in an essentially ‘masculine’ bureaucracy may be relatively more tolerant of suspected (male) GBV offenders, in what is still regarded as a patriarchal Society. This is not to say that the suspect would not escape censure, but that the probability of them being arrested on suspicion of the offence would be lower, counterfactually, than would be the case in a more representative Force. Add to this, a similar *laissez-faire* attitude at the senior levels, and it would be tempting to conclude that, as women feature more and more in policing, that the arrest rate for GBV should increase, and thus increase the cost to the perpetrators of committing these crimes.

Counter to this, runs the argument that the presence of women of in the Force might actually *reduce* the rate. Holdaway and Parker (1998) reported that female police officers believed that they had to alter their behaviours in unwanted ways, in order to gain and maintain acceptance amongst their male colleagues, and in order to not jeopardise their promotion prospects. The female officers also perceived that they were not being encouraged to apply for promotions to the same extent as the men, particularly in Kanter’s Skewed Group phase, where the minority is still small (5–15%, perhaps). As a result, any inherent prejudices that the male officers may have about the female victims, are adopted by the female officers. The fear of appearing to run against the grain by, for example, zealously pursuing male suspects for GBV, may force female officers to be even more tolerant of GBV than their male colleagues. However, the argument continues, as the bureaucracy enters the Tilted

Stage, female officers will feel freer to act as they wish to, and even persuade male colleagues to reconsider their own attitudes. This process would also be expected to apply at the senior levels, the difference being that there is more power to achieve change by ‘command’, rather than persuasion and demonstration.

Turning back again to Police Officers’ perception of GBV-related incidents. The easy assumption to make is that members of a group will tend to identify more with other members of that group, whether as a majority or minority. Arguably, members of a minority may be more protective of their own type, unless they are actively trying to distance themselves from them in order to gain ‘admission’ to the majority (most likely at the *Tilted* stage). When it comes to direct contact between Police officers and Society, account has to be taken of the pre-determined attitudes to both the alleged victim, the alleged perpetrator, and the nature of what is supposed to have happened between them. Thus, incidents to which officers have been called, will involve victims and perpetrators with whom the officers will negatively, or positively identify. In extreme cases, a ‘polarised’ officer may be confronted with a victim with whom they identify very positively, but extremely negatively with the perpetrator, for example. A previously-established set of rules governing this process may suddenly change as the bureaucracy transits through Kanter’s stages, towards Balance. Incidents that were once dismissed as trivial, and not worth anything beyond a verbal warning, may now be regarded as worthy of arrest and detention. While there are explicit and detailed Laws, with which all officers are conversant, they, as the first point of contact between the Public and the Judicial System effectively decide upon the course of action the System is to take. Thus, their attitude to the Client will be highly influential in determining whether or not, an arrest is made. As it is straightforward to come up with ‘common sense’ expectations of the outcome of these various encounters, researchers working in the area of policing and crime have had to engage in highly specific studies, drawing conclusions that can pertain only to similar scenarios, and cannot be generalised to specific theories (and solutions). The sheer combination of victim/perpetrator/officer types, the societal pre-conditions that determine their attitudes to each other make it very difficult to generalise across territories, incident type, or Time. It is as if the licence that the research has to draw a conclusion and policy recommendation, is both short-lived with a narrow applicability. It main value is, therefore, to inform future research, particular relating to data and methodological issues as well as keeping to the fore, that different Societies suffer from different imperfections in their judicial processes. Several studies have considered the extent of the engenderdness of police officers' responses to many different crime scenarios. Visher (1983) summarised research dating back to the 1950s. Her own US-based Study confirmed the continued existence of male Officer 'chivalry' when confronted with incidents involving *female* perpetrators, whereby the Officer was less likely to make an arrest than where the perpetrator was male, provided that certain expected 'female behaviours' were demonstrated. That said, where the woman was suspected of committing an act of violence, she was more likely to be arrested, presumably as that behaviour could not possibly fit the established stereotypes regarding women's acceptable behaviour (the classic *femme fatale* phenomenon).

Other studies have focussed more on specific types of crime. Stewart and Maddren (1997) looked at Australian Police Officers' reactions to domestic violence, where alcohol was a factor. They exposed a sample of male and female Officers to a series of vignettes, based on real incidents. All possible combinations of victim-perpetrator-inebriation were included, and the Officers asked to make and justify their (simulated) decision. While Officers (as people) were influenced by the gender combination and who was/wasn't inebriated, the Study failed to find any significant differences between the male and female Officers, though the small sample sizes used made it difficult to allow these findings to be generalised across space and time. Kite and Tyson (2004) adopted a similar approach to, exposing a larger sample of Australian Police Officers to a single vignette of child-related sex abuse, where the gender of the perpetrator was varied. Officers were asked to rate the seriousness of the incident, the impact on the child and what, if any, action they would have taken. Their findings contradicted other earlier studies conducted outside Australia, insofar as it failed to find any significant difference between male- and female officers, but that officers of both genders were more likely to take action against a male perpetrator.

These findings illustrate a trend in the literature, where older studies tend to confirm gender differences between police officers, but more recent ones find it difficult to do so. Given that the methods employed by the various studies appear to be largely unchanged, this is perhaps a testament to the efforts made by these societies and police forces to eliminate gender as a factor in crime detection. Credit may also be due to the (younger) male officers, whose attitudes to women, and the crimes that affect them, may have become progressively more enlightened with the passing of Time. It may also be credited to the increased presence and normalised roles of female officers at street level, and into the higher, policy-making echelons. To this, might be added the contribution of older, senior female Officers who, having survived and prospered in a Service where they were 'tokens', have not turned their backs on the issues that affect women. Society will also have been partly reformed, to become more progressive and inclusive of previously under-represented groups, such as women and racial / sexual minorities.

2.3.4 Panel-based studies of Crime: a brief review

The above studies variously examined very specific scenarios, using a variety of elementary descriptive- and inferential methods testing various hypotheses concerning gender of victims, perpetrators and Police Officers. Given their narrow focus, combined with small sample sizes and, in some cases, questionable statistical methods, it is difficult for them to conclude beyond the particular to the general, as many authors freely acknowledge. That said, each Study is valuable insofar as it highlights methodological issues that allow subsequent Studies to learn from mistakes and deficiencies. Following Becker (1968), there flowed a steady stream of papers focussing on different crime deterrent scenarios, in a variety of jurisdictions, for example, Carr-Hill and Stern (1973) and Erlich (1975). Others looked at aspects of Police performance, for example in converting arrests made into successful prosecutions (Cameron, 1987).

While the quality of the Econometric techniques improved through time (along with the size of the sample datasets), it was not until 1994 that the first Crime Study based on a Panel Dataset was published by Cornwell and Trumbull (1994). They examined the effect of deterrents on the probabilities of arrest, conviction and imprisonment based on a Panel of 90 counties throughout the United States, spanning seven years. They derived parameter estimates from Pooled Cross-sectional models and between- and within- Panel models, demonstrating significantly weaker deterrent effects from the Panel models. The explanation they offered was that there were unobserved county characteristics significantly correlated with the included (crime-related) variables, resulting in inconsistent parameter estimates. By way of an example, they cited differences in crime-recording practices in the jurisdictions, and Officers' zeal in arresting and prosecuting suspected felons, affecting both the numerators and denominator of the composite dependent variable.

Other, more recent Panel-based crime Studies include Andresen (2012) who looked for evidence of a link between unemployment and different crime rates in Vancouver, Canada. He used a small three-wave Panel to model burglary, car theft and violent assault rates against a number of plausible independent variables, including the unemployment rate, percentage of young males in the population, percentage of single families and population density. He ascertained that the significance of the independent variables varied according to the dependent variable modelled. For example, the rate of violent crime was positively related to the unemployment rate, population growth rate and population density, but only the unemployment rate was significant in explaining the robbery rate.

Falls and Thompson (2014) examined the presence of newly-established Casinos in Michigan in the USA and their effects on different crime rates in the State (as per Andresen). Using a Random Effects specification, they derived results from a Panel, comprising the 83 counties in the State over 21 years. Independent variables included average income per capita, population density, the number of police officers per capita, ethnic, gender and age population compositions and measures of general and casino-based gambling. Similar to the Andresen Study, they reported different combinations of significant variables according to the crime rate being modelled. Interestingly, only vehicle theft was affected by the police variable (negatively), though population *density* was a positively associated with most types of crime.

Iparraguirre (2014) used a single cross-section of Crown Prosecution (CPS) Data from 2011 pertaining to prosecutions for 'Hate Crimes' against older (>65) people in England and Wales, to estimate a structural model. The data were measured at the Local Authority Level (which has some broad correspondence with the Police Force Areas used in this Study). The dependent variable in the main equation was the number of Hate Crimes each year in each Area, with the independent variables including per-capita income and employment measures, as well as education and local conviction rates and the area's general level of all Hate Crime. This general Hate Crime level was modelled, separately, by an equation that was regressed against an income variable and the number of people between 15 and 24 in the area (the modal 'attack' group). However, no explicit policing variables were included. Virtually all of the included variables indicated a significant, positive effect on the age-related Hate Crime count.

These, and other multivariate Studies are, perforce, superior to those summarised in **2.3.3**, insofar as they attempted to model the interaction of several regressors, some based on Panel-based specifications and some using larger sample sizes. That said, their licence to generalise is similarly restricted to specific territories and time-periods. Their value is to contribute to a body of evidence which may, or may not indicate particular trends through time, or indicate broader territorial (e.g. continental) congruence with other like Studies.

2.4. The Police Service in England and Wales

2.4.1. Introduction

The modern English and Welsh Police Service started with the establishment of the London Metropolitan ('Met') Force in 1829. Given its success in curbing crime, other cities and counties quickly established their own Forces, with it becoming a legal requirement to do so throughout England and Wales by 1857. After many structural changes, England and Wales currently operates with 43 Territorial Forces, with the last major reorganisation taking place in 1974. **Figure 2.3** maps the physical boundaries of the Forces. [11] While all Forces operate under the same legal jurisdiction, they are independent organisations, each operating with broadly similar command-and-control structures, though with differentiated allocations of responsibilities at each level (particularly the two London Forces). Each Force has the discretion to organise its own workforce within its structure and how to detect and prosecute crimes in its Territory. The Forces are also grouped into Regions in an effort to resource- and share information on a basis of proximity to each other. As **Table 2.2** indicates, they are clearly not standardised to be the same physical size, or to have the same population residing within them.

For instance, notice the large differences in populations in 2013, ranging from the Met Force with almost 8.5millon people to the minute City of London Force, an area of 3km^2 with only 7,600 residents. The population densities ranged from 4,670 people-per-hectare (pph) in the Met in 2013 to 45pph in the physically larger, Dyfed-Powys. These imbalances can be explained, in part, by nostalgic ties to old County Boundaries as well as to an unwillingness to incur the cost of a radical redrawing of Force boundaries, beyond the very small adjustments that have been made, from time to time.

[11] The Welsh Forces have some emerging differences in Law (for example with Drink-Driving legislation), with the advent of the Welsh Assembly in 1998, but are otherwise apply the same Laws as England. In addition to the Territorial Forces, there are a number of all-UK Forces, such as the Border Force, Transport Police, and the Ministry of Defence Police, that are not included in this Study: neither is the small City of London Force, given the highly unusual territory and role it is involved with.

Table 2.2
Summary Geography and Demography of the Police Regions of England & Wales, 2002 and 2013

	Area size (km^2)	Population		People per hectare (pph)		% population change 2002 - 2013
		2002	2013	2002	2013	
Maximum	10976 km^2 (Dyfed-Powys Police)	7.369m (Metropolitan Police Service)	8.409m (Metropolitan Police Service)	4670 (Metropolitan Police Service)	5329 (Metropolitan Police Service)	14% (Cambridgeshire Constabulary)
Minimum	2.6 km^2 (City of London Police) 5.97 km^2 (Cleveland Police)	7,100 (City of London Police) 488,000 (Cumbria Police)	8,000 (City of London Police) 498,000 (Cumbria Police)	44.7 (Dyfed-Powys Police)	47.1 (Dyfed-Powys Police)	1% (Cleveland Police)
Median	2713 km^2 (Staffordshire Police)	0.986 m (Cheshire Constabulary)	1.035 m (Cheshire Constabulary)	368.3 (Leicestershire Police)	403.4 (Thames Valley Police)	8% (West Yorkshire Police)
England & Wales Average	3527 km^2	1.223 m	1.324 m	347	376	8%

Sources: Home Office; ONS

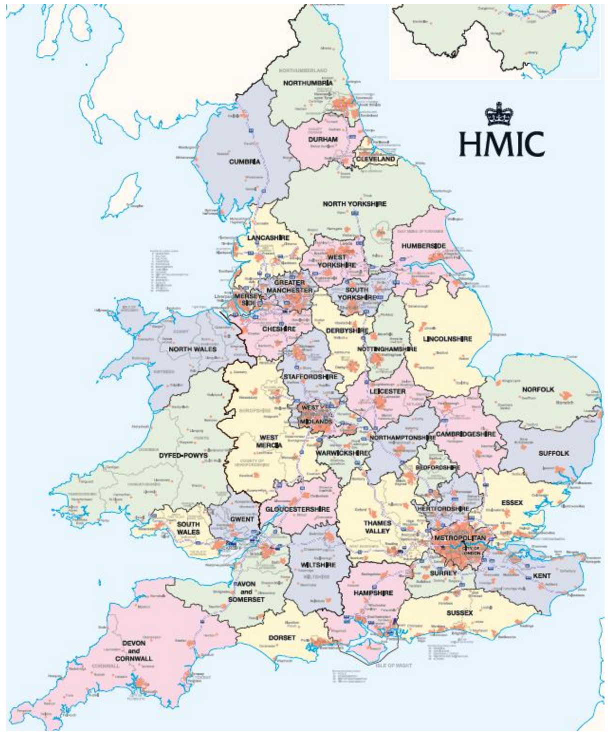


Figure 2.3 The 43 Police Force Areas in England and Wales

(www.latitudemaps.co.uk Product includes mapping data licensed from Ordnance Survey with the permission of the Controller of Her Majesty's Stationary Office ©2013 All rights reserved. Map image is incorporated in this thesis under an Open Government Licence, 2019. Available https://www.justiceinspectorates.gov.uk/hmicfrs/media/police-force-map.pdf Date last accessed: 11th March, 2019)

From their inception, all Forces adopted and have since maintained, a quasi-militaristic structure and ethos, with a strict pyramidal command-and-control structure: all officers pledge loyalty to the Head of State ('the Crown') and must obey orders given to them by a superior. Although they all have the power to investigate alleged crimes and, if necessary, arrest suspects, the higher ranks have additional powers and responsibilities. As with other bureaucracies, the higher up the hierarchy the officer is located, the further away they are from the interface with the client. Street patrols tend to be staffed with Constables and Sergeants and occasionally, Inspectors. Set piece 'operations', like drugs' raids and football crowd-control will have on-site commanders up to Superintendent level present, to monitor proceedings and, where necessary, adjust tactics. The upper-echelons are primarily concerned with supervising the lower ranks, policy setting and resource allocation. All Forces operate through a system of stations and a headquarters, staffed with a mixture of uniformed and detective Officers working with civilian staff who provide clerical and specialist administrative support, such as HR and data entry. Through a long process of evolution and reorganization, each Force has created various Divisions and sub-commands, each tasked with detailed oversight of either a geographic- or functional area.

Each Force can conditionally determine its priorities for policing, subject to national initiatives emanating from the Government. Senior Officers are required to constantly review the Crime Statistics for their Force, and to resource initiatives designed to curb specific types of crime that are currently at unacceptably high levels. While all Forces operate the same basic structures at their lower levels, it is at their senior levels that they choose to vary their structure and relative preponderance.

2.4.2 Women in the English and Welsh Police Forces

In 1914, the London Met was the first Force to allow women to become police officers and undertake limited duties, including some patrolling. This was at a time when large numbers of suitable men were conscripted into the Armed Services during the First World War, though it was not until 1923 that women were given the status and powers of Constables. Until 1946, women were required to resign from the Service when they married. Their uniforms were very different from the men's, being required to wear skirts and different hats, and clearly identified as a 'WPC', as opposed to a 'PC'. Women were not allowed to join the Police Federation (the *de facto* Trades Union) until 1948. Each 'Women's Force' had to be headed by a female Officer, with the middle-management rank of Chief Superintendent. This was, *de facto*, the highest rank a woman could attain, until 1969, when the Met promoted a female officer to the rank of Commander.

The Right to equal pay was not granted until 1974 following the Equal Pay Act of 1970, and as Brown, et.al (1993) report, the 1975 Sex Discrimination Act was a 'watershed in the evolution of Policewomen's role within the British Police Force', that heralded the end of the separate Women's Forces (Martin, 1996). However, women still formed only 6% of Forces' strength in 1980 [12], in contrast to 46% of the workforce in general. [13] They were kept well away from 'unpleasant' situations and/or those requiring brute force. A Study based on a small sample of English Police Officers (Brown, et.al, 1993, *op.cit*) reported bias in the deployment of female Officers even where 'gender-specific' skills were not required. In a follow-up Study, Brown (1998) reported that female Officers were more likely to be confined to 'in-station' work, dealing with sex offence- and juvenile-related incidents, while male Officers were deployed to public (dis)order incidents and to deal with violent offenders. That said, there was no evidence of differential deployment to gender-paired patrols, or that women were more likely to be assigned to sole patrols in relatively 'safe' areas. Holdaway and Parker (1998, op cit) investigated the deployment patterns of female- and male Officers and confirmed that there were significant differences. Dispatching Officers assumed that women Officers were better deployed in incidents involving young people and domestic disputes. As a consequence, female Officers were inevitably more involved in domestically-related sexual crime. However, they also found evidence of *gendered preferences* for particular activities, with females being more interested in arresting and interviewing suspects, though they also felt they were being too narrowly focussed by their superiors on this 'kind of work'. Similarly-gendered differences were reported by Rabe-Hemp (2008), based on a US Study, particularly that male Officers have a greater preference for 'controlling' behaviour, i.e. threats, physical restraint and arrest. Her study found that female Officers were 27% less likely to use controlling behaviour but

---

[12] Source: History of Met. Women Police Officers, Metropolitan Women Police Association (http://www.metwpa.org.uk/history.html, accessed 12th April, 2015)

[13] http://data.gov.uk/dataset/labour_force_survey_employment_status_by_occupation, accessed 15th May, 2015

were equally comfortable in issuing direct commands and advice: evidence perhaps of them 'aping' male behaviour in an 'unwanted way'. In another US-based Study, Garcia (2003) stated that female Officers were 'repeatedly channelled into the least desirable jobs', also quoting other Studies that made the same claim. In an analysis of what officers actually do, she asserted that it is 'common knowledge' that policing mostly involves 'social work interventions' with the Public, and that only approximately 20% of the time is actually spent enforcing suspected breaches of the Law. However, she suggested that people are attracted into the Service in the belief that they will actually spend the vast majority of the time literally pursuing and arresting suspects and collecting evidence for subsequent prosecution. This emphasis on 'male skills' (aggression and violence) is at the expense of the 'female skills' of empathy and negotiation. Given the minority presence of women in the Force, there is therefore a skills' mismatch. In analysing the apparent lack of progress female officers were making up it to the Forces' higher echelons, Dick and Metcalfe (2007), found no evidence in a study of two (large) UK Forces that women displayed any less commitment to the Service than men, nor were they inclined to leave it any sooner. That said, they concluded that their career progression was slower and terminated at lower levels. The implicit conclusion was, therefore, that discrimination against women in the Force was a significant factor.

Figure 2.4 charts the progress made by women across the Service during the study period (2002 – 13). Total numbers rose by about 13,000, matched almost exactly by a reduction in the male headcount. The percentage rose from 18% to 28%, still well short of 50% 'balance' but, if Kanter is to be believed, possibly moving the Service out of the *Skewed* into the *Tilted* phases, whereby the female officers may have become able to alter the Service's priorities towards, *inter alia*, crimes that affect mainly, women.[14]

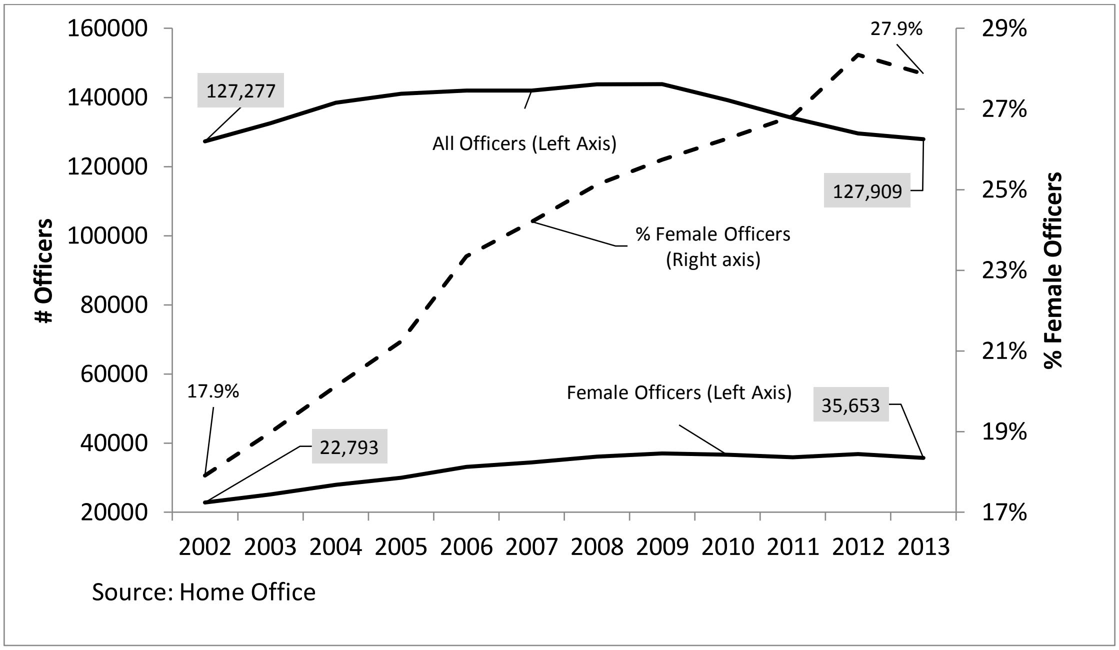


Figure 2.4 Total Officers and Female Officers: England and Wales 2002 - 2013

2.5. Scope and Aims of this Study and the Data

2.5.1. Scope and Aims of this Study

As stated in the introduction to this chapter, this Study adopts a distinctive quantitative stance in identifying the existence, or otherwise, of a specific gendered effect on both the number and rates of GBV arrests in England and Wales between 2002 and 2013. This, perforce, makes its scope, both spatially and temporally, limited as with the other studies. It is to be hoped, however, that the estimation of well-specified econometric models will identify the impact of more-gender balanced policing in England and Wales on a crime type that predominantly affects females. From this, it may be possible to further critique Kanter's organisational stages and at what stage(s) the English and Welsh police forces were at in that time period.

[14] Note that the slight downturn in the percentage of female officers in 2013 was revered thereafter and reached 29.1% across England and Wales by 2016.

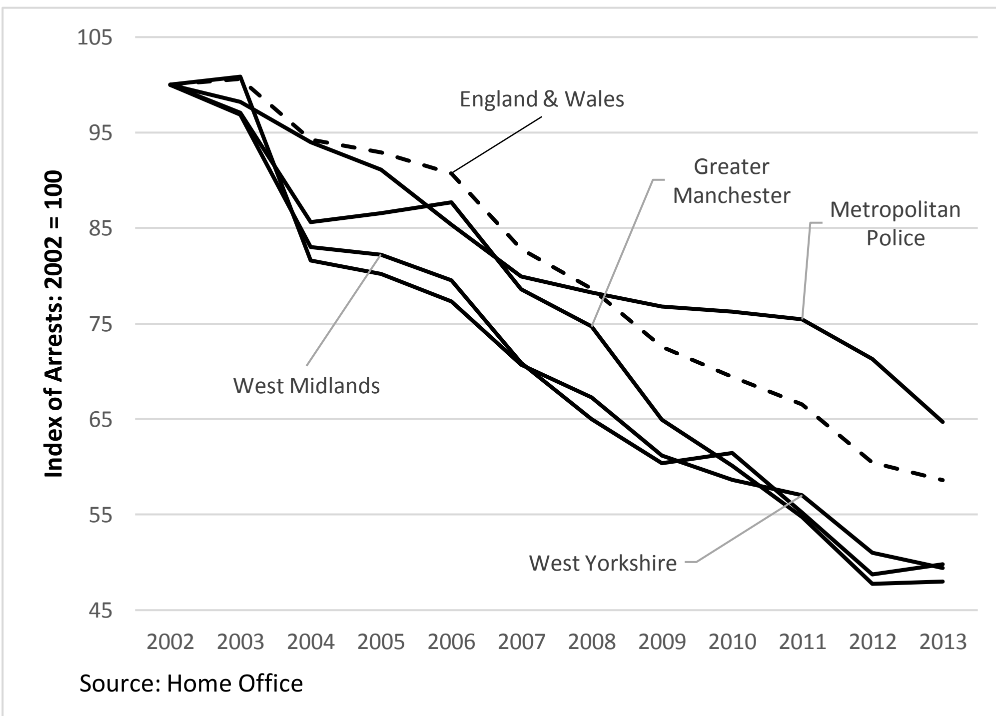
The data used will be outlined in more detail shortly**,** but it can be said here that they are semi-aggregated, in the sense that they are published Home Office data on the numbers of arrests for specific categories of crime, and sub-totals of office numbers, by rank and gender. The data are enumerated at *Force* level. In an ideal world, the officers' notebooks and station records would be used as the primary data source, to include the witnessing officer's rank, age and, of primary relevance, gender. The reasons for arresting (or otherwise), would also be recorded, as well as information on both the (suspected) perpetrator and (suspected) victim. These data are not available to the Author, but authorised access to it would form part of any subsequent work in this area. The unit of analysis is therefore not the *Officer*, but the *Force* in which the arrests were made.

2.5.2 The Dependent Variable: number of GBV Arrests

The Home Office currently identifies 1,498 separate 'notifiable offences', sub-categorised into 128 high-level offence codes and nine 'super-classes': Burglary; Criminal Damage; Drug Offences; Fraud and Forgery; Robbery; Sexual Offences; Theft and Handling of stolen property; Violence against the Person, and; Miscellaneous. Over 200 Acts of Parliament (some dating back to the 14^{th} Century), Local 'bye-laws' and the Common Law define what constitutes a particular offence and specify the range of punishments (including custodial sentences) available to the Courts. For the purposes of this Study, the offences listed in **Table 2.3** are regarded here as being GBV-related. As might be expected, these belong to the Sexual Offences and Violence against the Person, super-categories. The table indicates the wide range of offences that constitute acts of GBV, some affecting minors, other adults, and also that they are not necessarily related to *unconsensual* sex acts. In addition, there are several identical offences pertaining to *male* victims, for example, 19F: Rape of a male aged 16 and over, correspondent with 19C. However, given the earlier discussion that GBV has to be seen as predominantly perpetrated by males on females where gender is *the* key factor, offences against males of any age are excluded from this Study even though a small number of these will have been perpetrated by a female. There is some correspondence between these categories and the Home Office categories of violence against the person with injury and sexual offences.

Table 2.3 Home Office Offences assumed to be GBV related			
Home Office Offence Code	**Definition**	**Home Office Offence Code**	**Definition**
8K	Poisoning or female genital mutilation (FGM)	19E	Rape of a female child under 13
20	Indecent assault on a female	20B	Sexual assault on a female child under 13
20A	Sexual assault on a female aged 13 and over	22	Unlawful sexual intercourse with a girl under 16
25	Abduction of female	22B	Sexual activity involving child under 16
19A	Rape of a female	88A	Sexual grooming
19C	Rape of a female aged 16 and over	24	Exploitation of prostitution
22A	Causing sexual activity without consent	27	Soliciting for the purposes of prostitution
19D	Rape of a female child under 16	72	Trafficking for sexual exploitation

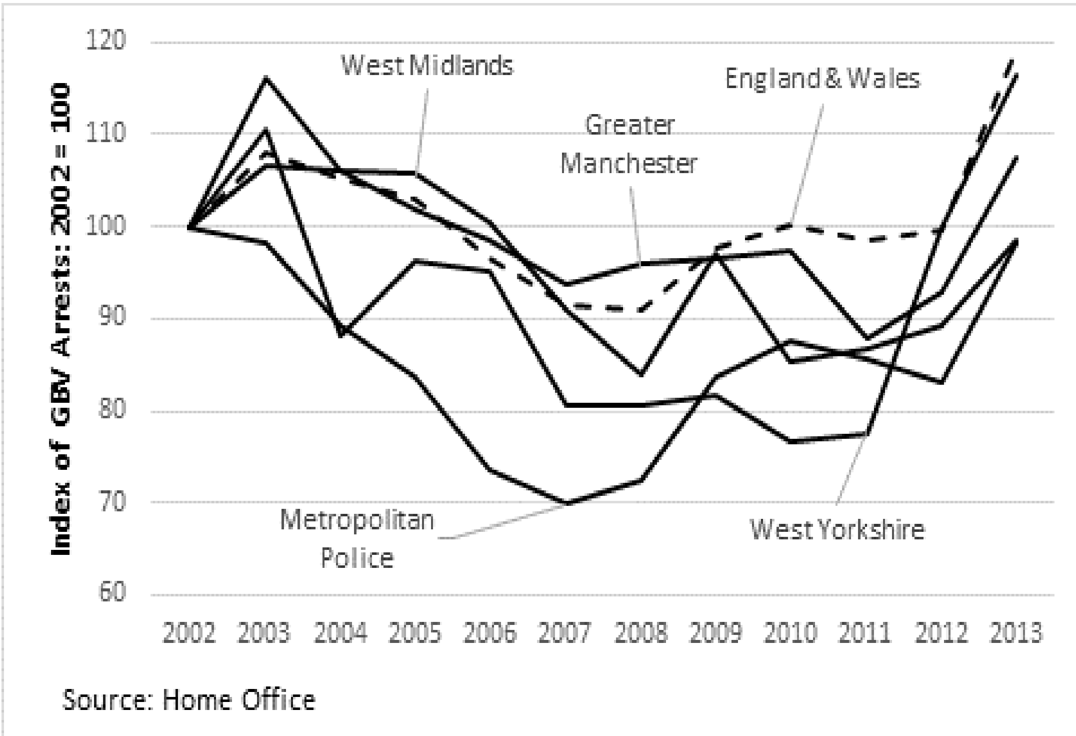
While the table includes ‘Stalking’ and ‘Sexual Harassment’ as categories of GBV, and the Home Office schema includes ‘Harassment’ as a crime type, it does not clearly specify whether the harassment is, perforce ‘sexual’, or otherwise perpetrated because of the (alleged) victim’s gender. Consequently, that crime category is not included in the calculation of the total numbers of GBV crimes during the period of this Study. **Figure 2.1** shows that despite a net increase in the population of England and Wales of 4.4 million people during the Study Period, the total number of all arrests *per annum* fell. Considering the detailed reasons for that are outwith the scope of this Study, except to note that changes in per-capita resourcing of the Service, as well as in the Law and procedures relating to arrest, lie at the heart of this phenomenon. Conversely, the number of *GBV* arrests, while falling initially, increased steadily from 2008 to 2012 and then rapidly during 2013. Consequently, the proportion of arrests that were for GBV doubled from 0.68% to 1.39%, corresponding approximately to the increase in arrests on suspicion of *sexual* offences. We should recall, however, that the data are a Panel of 43 separate Forces, each possessing a high degree of autonomy in the setting and pursuance of its priorities: it is not necessarily the case that this aggregate downward trend is echoed across the entire Service. **Figure 2.5** plots the indexed series of Total Arrests for the four largest Forces (by population), along with the all-England and Wales series, using 2002 as the Base Year. These series do indicate some differences in a shared downward trajectory, but that all ended the period with lower numbers of arrests than at the start, with Forces like the West Midlands and West



Yorkshire making less than half the number made in 2002 (NB: all the other 39 Forces behaved, broadly, likewise).

Figure 2.5. Indexes of Total Arrests for the four largest Police Forces 2002 – 2013 (2002 = 100)

Figure 2.6 plots the indexes of *GBV* Arrests for the same four Forces. As with Total Arrests, there are observable differences between them. The larger Forces generally followed each other on a downward trend from 2002 to around 2007 then made a sharp turn upwards in 2012 to at least get back to their 2002 levels (in Greater Manchester and the West Midlands cases, well above that). There was no fall below the 2002 levels for the smaller Forces and, in fact they ended the period with more



than 70% more arrests per annum.

Figure 2.6 Indexes of GBV Arrests for the four largest Police Forces 2002 – 2013 (2002 = 100)

With GBV arrests rising through a period in which total arrests fell, it is the case that the percentage of arrests made for GBV increased. **Figure 2.7** plots the series for the two regions with the highest average percentage, along with the two lowest and the average for England and Wales. Again, differences between the Forces can be seen, as well as one or two short-term downturns in the rates, but all Forces (including those not shown here) finished the period with significantly higher GBV arrest rates than at the start.

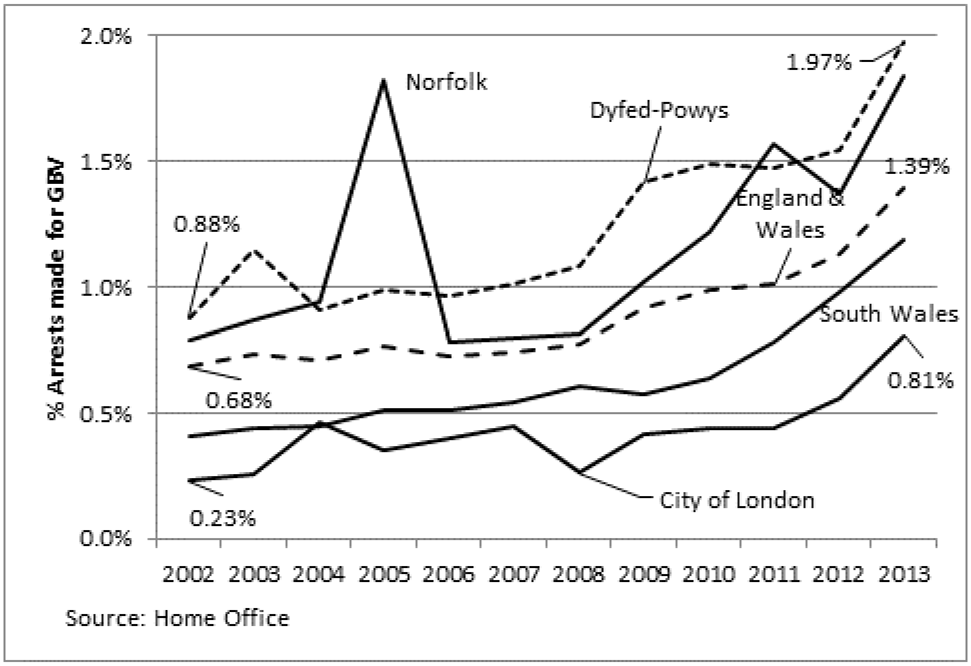


Figure 2.7 Percentages of Arrests made for GBV: 2002:2013

2.5.3 The Independent Variables: an overview

In the absence of an established quantitative theory as to the specific causes of arrests for GBV, we are left with the challenge of martialling a set of plausible, observed co-factors that help explain either/both the level and rate. Clearly, the number of arrests will be a function of:

(i) the number of GBV incidents,

(ii) the alleged victims' willingness to make the Police aware that these incidents have taken place and/or,

(iii) the Forces' abilities to actually witness the incidents taking place, and,

(iv) the Officers' appetites for making arrests.

The number of incidents will be a function of many factors, such as the overall size of the population and its age-gender composition. On the age front, it is simple to imagine that an older population will be less inclined to commit any acts of violence upon itself. On the gender dimension, a Society that was comprised entirely of only one gender would, by definition, almost be incapable of committing acts of GBV upon itself. Any Society that is dominated by females would be less likely to commit acts of violence, compared to one that was male-dominated. The issue of whether acts of *GBV* would be less likely in a female-dominated Society is moot, given the physiological and psychological differences between the sexes. One could add to this, objective socio-economic factors, such as income (distribution), employment rates, housing density, as well as subjective factors, like local cultural norms and educational initiatives, any of which could be significant co-factors.

Of course, this Study is concerned with modelling the arrests and arrests'-rate for GBV. It might be assumed that there is a direct relationship between the number of GBV incidents and the arrests carried out. This may be true, but we must take into account the Service's ability to observe and, if deemed appropriate, detain and arrest the suspected perpetrators. An increased number of street-level officers should have the immediate effect of allowing a greater percentage of incidents to be observed, with an increased willingness to make arrests and to cope with the consequent administration. This should, in theory, have the longer-term effect of reducing the incidents of all crime, including GBV. Reducing officers' tolerance for GBV, *ceteris paribus*, should increase the arrest rate for it. This tolerance may be reduced by introducing more women officers into the policing effort and also by improving cultural awareness, via training of all officers. Encouraging a culture where the alleged victims of GBV feel they can make formal complaints to the Police and be taken more seriously would improve the under-reporting situation, but that is a macro-level Societal issue, only partly within the control of the Service.

Consideration of the gender-composition of each Force through Time may tell us something about the extent to which at least, the *passive* representation of females in the Service is facilitating the arrests for GBV. To this end, what effect, if any, female officers have had at the higher policy-setting and resource allocation echelons of the Service should be considered. Referring again to Kanter's classification of the stages that a bureaucracy can go through, these time-served Senior Officers would have entered the Service as Constables in the 1980s and 1990s, as 'tokens' in a *Skewed* bureaucracy. By the time of the start of the Study period, they would have become 'tokens' in the upper echelons but presiding over a bureaucracy that has moved towards better balance, where their junior female officers are less 'token', perhaps part of a *Tilted* group at the lower levels.

As the author has no special access to the highly-confidential officer and station-level data that record the genders of the victims and reasons for the arrest and (indirectly) the gender(s) of the arresting officer, publicly-available data, aggregated annually at Force level were used instead to test the representative bureaucracy hypothesis. Accordingly, the independent variables included in this Study are:

- The numbers of Police Officers (Total & Female) deployed by each Force each year (Home Office Data);
- The numbers of arrests made by each Force each year (Home Office Data);
- The population (by Gender) and number of households in each Force, each year (Central Statistical Office Census Data);
- Dummy Variables for Force and Year.

<u>Police Officers deployed and Gender composition in the English and Welsh Forces</u>. **Figure 2.4** records that the total Force strength rose from, then fell back to, ~127,000 by the end of the Study period. **Figure 2.8** shows that some Forces increased their strength significantly, while others saw theirs fall, equally significantly. While the

numbers of officers are suggestive of a decline in ability to observe all crime (perhaps helping to explain the reduction in the number of arrests?), Forces have deployed more technology and been aided by the recruitment of low-cost Community Support Officers in an effort to maintain effectiveness. Nevertheless, it is the officers 'on the ground' that actually arrest people, so these numbers should be significant.

A precis of the history of women in the English and Welsh police forces was given in **2.4.2**. By the start of the Study Period, women had made significant progress in all the Forces, including an increased presence in their upper-echelons. In 2012, there was, for example, a female Assistant Commissioner in the London Metropolitan Force. [15] **Figure 2.4** illustrated how the numbers and proportions of females across the whole Service rose through the Study Period.

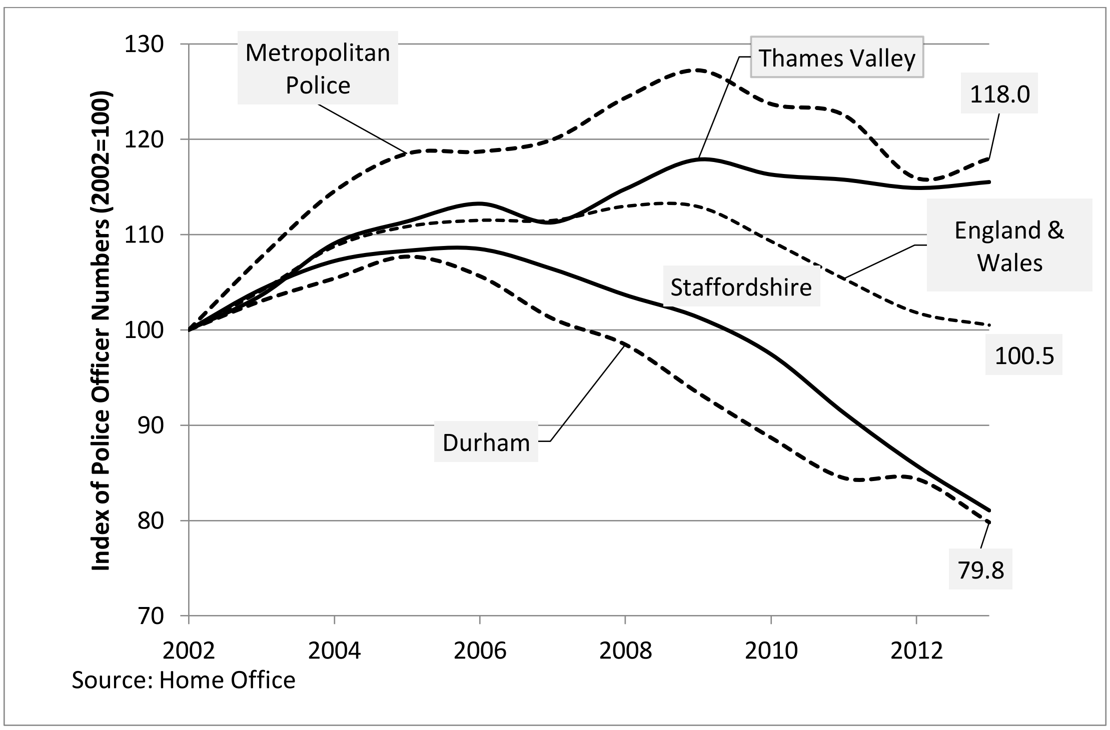


Figure 2.8 Index of Forces' Strength: 2002 – 2013

[15] Subsequently promoted in February 2017, to become the first-ever female Commissioner of the London Met, the second-ever female at this level in England and Wales.

That said, while all Forces improved their gender-balance, they did not do so to the same degree (**Figure 2.9)**. This confirms that all Forces (including those not shown here) increased the relative numbers of female officers in their ranks, though there remained considerable disparities between them. This raises the possibility of Forces being at different 'Kanter' phases, with consequences for the emphasis placed on GBV, for example.

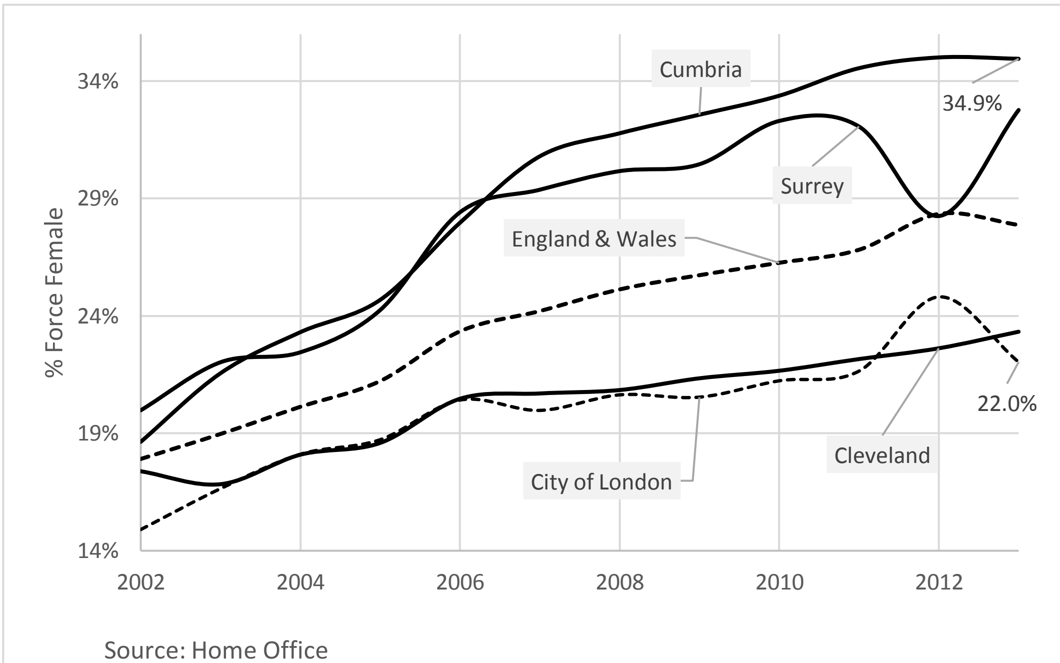


Figure 2.9 Percentage of Female Officers in the two best- and two-worst gender-balanced Forces

<u>Population and Household numbers in England and Wales</u> The population of England and Wales rose by 4.4million people (8.3%) through the Study period. At the same time, the number of households formed rose by just over two million (9.4%), implying a modest fall in the average density from 2.42 people per household (pph) to 2.40pph (**Figure 2.10**).

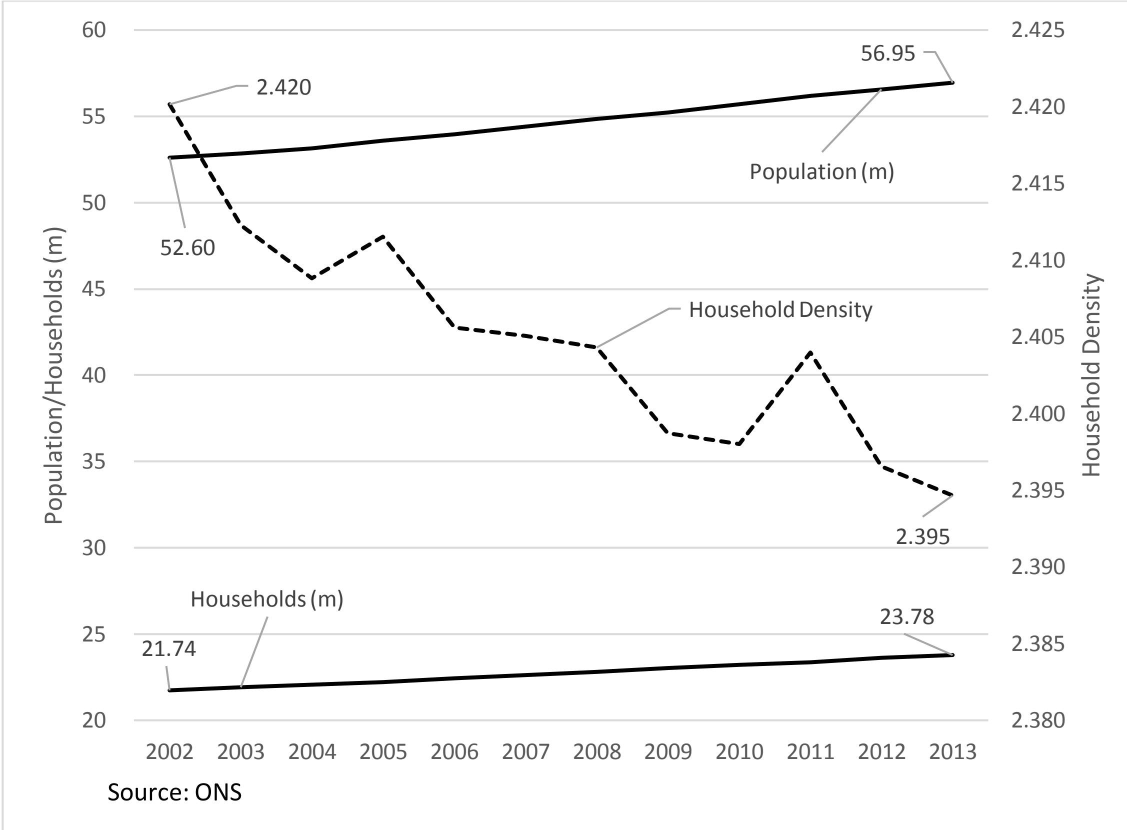


Figure 2.10 Total Population and Households and Household Densities: England and Wales 2002 - 2013

However, as **Figures 2.11** illustrates, the rates of population increase were not uniform across all Areas, rising more rapidly in the South and West of England) than in the North and in Wales. Note that the London Met area's growth rate was third, though its absolute growth was the greatest (~ one million extra people). **Figure 2.12** plots the percentages of the Area populations that were female, for Merseyside (the highest on average) and Bedfordshire (the second lowest). As might be expected, there was both relatively little difference between these areas, and little over the Study period. Thus, there is no ostensible reason why all Forces should not be aiming for the same gender profile in their staffing, though that this may not necessarily have to reflect, exactly, the population balance.

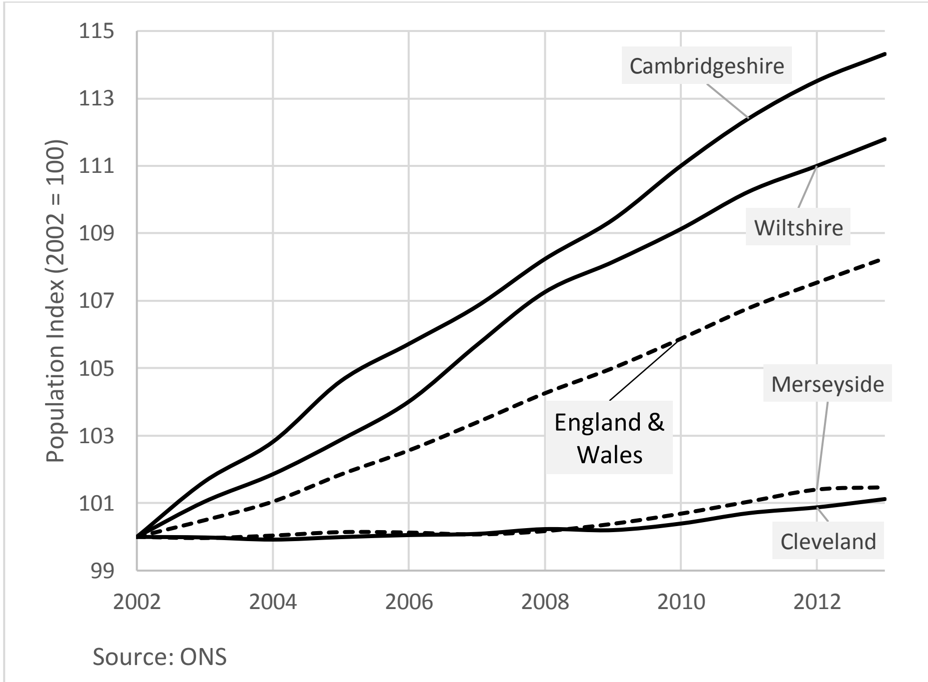


Figure 2.11 Population Indexes for the two fastest and two slowest-growing Forces in England and Wales

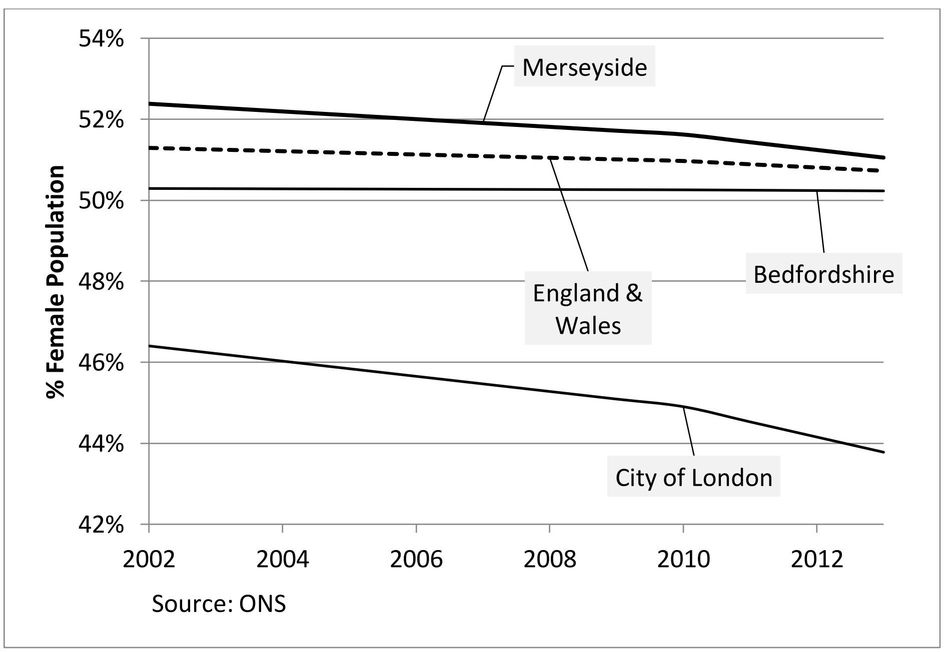


Figure 2.12 Force Area Gender compositions (2002 – 2013). Highest and lowest % female

Note the outlying City of London Force Area, where females were always a declining minority in the population. This Area lies within the boundary of the Met Area, where the UK's Financial Sector is located. While it has only an official population of ~7,000, many more workers commute in-and-out daily from the surrounding area, as they simply cannot afford to live there. It appears that those who could were predominantly male, über-wealthy Financiers living 'off-street' in high-rise blocks. Clearly, its small size, distinctive gender composition and socio-economic profile make this a rather unusual area. **Figure 2.13** plots the household densities for the two most- and two least dense Areas. This illustrates some differences between the Force Areas, but that all shared a small, but steady trend towards *lower* densities. The ONS notes upward trends in single-occupancy households (implying a *lower* density) but also in 20 to 34-year olds continuing to live in the parental home (implying a *higher* density).[16] This may signal the formation of smaller family units and, assuming a constant dwelling size, more physical space per person.

The effect, *ex-ante* of density on GBV is unclear. On the one hand, a reduced density may be concomitant with reduced social pressures, reducing the motivation for intimate-partner GBV, for example. On the other hand, the same reduced density would allow for more behind-doors privacy and may encourage perpetrators in the absence of third-party intervention. Only as the arrest-rate models are estimated in section **2.6.4**, will the net-effect become apparent.

[16] Families and households in the UK: 2016, Statistical Bulletin, ONS, https://www.ons.gov.uk/ people populationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2016: accessed 14th April 2018.

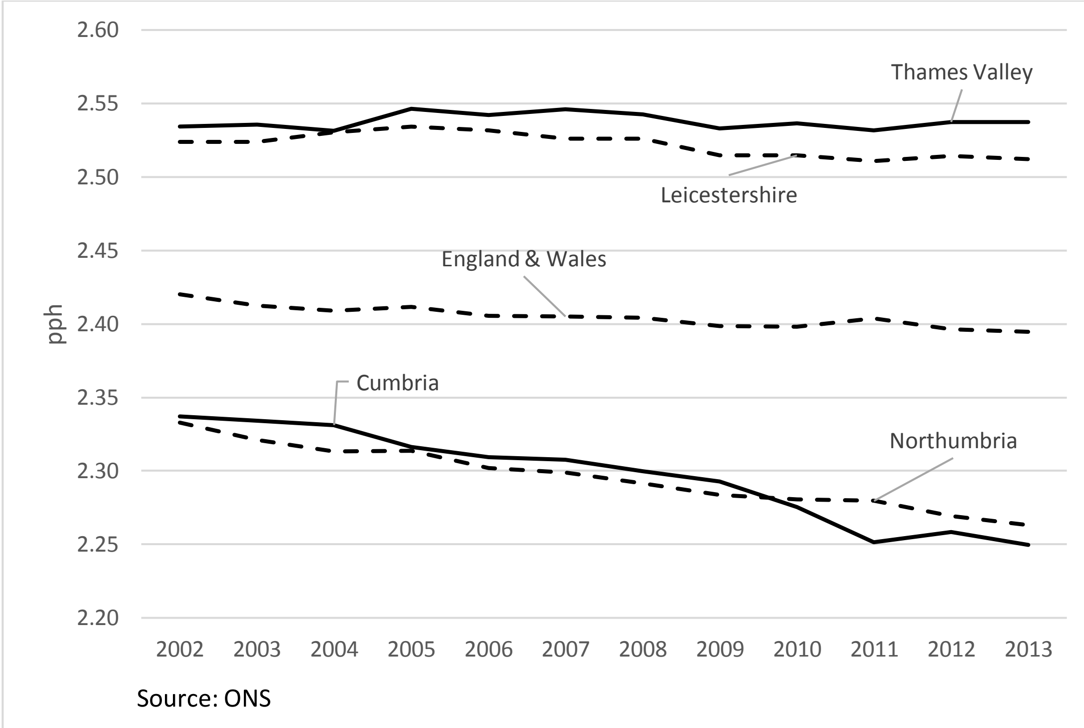


Figure 2.13 Greatest- and smallest Household Densities in England and Wales 2002 – 2013

2.6. Testing the Hypothesis: models' estimation

2.6.1. The Hypothesis

A fundamental issue with the Study is not knowing if any crime type including GBV, is being over- or under-reported and, as a second filter, resulting in too many, or too few arrests. The suspicion has to be that GBV, historically, has been under-reported given its behind-doors nature, the ability for the perpetrator to further intimidate the victim to remain silent and the engendered reluctance of the police to become involved in what were regarded as low-level domestic problems. This then fuelled and sustained reluctance on the part of the victims to report the crimes against them in the expectation that they would not be believed and possibly mistreated more by the perpetrator. Otherwise sympathetic officers would be reluctant to intervene to the point of arrest, anticipating a low probability of their superiors allowing the case to proceed towards prosecution, given the poor prospects of obtaining sufficient evidence to allow the case to come to Court. Given the budget constraints that the higher ranks need to manage, they simply do not want to incur the fixed costs of making an arrest that will, in all probability, not result in a conviction. Accordingly, this Study cannot conclude on the *sufficiency* of the level/rate of GBV arrests, only that it may have been influenced, or otherwise, by the changed gender composition of the Forces.

The fundamental premise of this Study is that GBV in England and Wales, a crime mostly perpetrated by males on females has been, counterfactually, under-investigated by a historically male-dominated Police Service. Accordingly, English and Welsh females have been under-served by the 43 Forces, as they have not properly aligned their staffing profile with that Society (in the gender dimension, at least). This implies, therefore that those Forces, as bureaucracies, were *unrepresentative*. That said, all the Forces, to differing degrees, steadily increased the participation of females in all their ranks (**Figures 2.4** and **2.9**). Contemporaneous with this, was the increase in the number of arrests on suspicion of GBV offences (**Figures 2.1** and **2.7**), against a backdrop of a decline in arrests in general (**Figures 2.1** and **2.5**), in the face of fiscal pressures affecting policing intensity and, perhaps other unobserved/inscribable reasons. The increase in GBV arrests may be a result of:

- an increase in males' violence against females
- a reduced Societal toleration of such acts, leading to a greater willingness and ability of the victims and others to report them, or
- a reduced tolerance by the police officers making it more likely than before that such incidents to which they have been called, will lead to at least an arrest being made.

Even if it can be demonstrated that the increased participation of female officers in the Service has been a significant factor in increasing the numbers and rates of GBV arrests, it will not be possible, given the data, to determine whether this is a result, wholly of their *passive* representation (the presence of the female officers altering the behaviour of their male colleagues) or *active* representation (senior and junior female officers actively targetting GBV with their (assumed) naturally-lower tolerance). Neither would it be possible to determine, precisely, where each Force or the entire Service is 'at' with reference to Kanter's four stages, except, perhaps to rule out certain stages on both numeric- and behavioural bases. For instance, evidence that the presence of more female officers is a significant factor would indicate that the Service has progressed through the *Skewed* Stage, where tokenism, 'male kickback' and female officers 'going native' would combine to even *reduce* GBV arrests. This would point the finger at either the *Tilted* or *Balanced* stages. Arguments against Balance would be that (i) none of the Forces had attained numerical equality in its ranks by 2013 (**Figure 2.9**), and (ii) the British Association for Women in Policing was formed in 1987, that still claims to be:

'…the only national organisation to embrace women of all ranks and grades within the Police Service. Our mission is to ensure that those women are heard and work toward gender equity in policing' [17]

As noted in **2.3.2**, the formation of such self-interest groups may seek to hinder progress towards a truly unified bureaucracy. Should it <u>not</u> be possible to demonstrate any positive association between female officer participation and GBV arrests, then this would point the finger at the Service being somewhere in the Skewed-Tilted border. Should there be a *negative* association, then this would be strongly suggestive that the Service overall was located in the lower reaches of the Skewed Phase, where the additional females still feel they have to behave like prejudiced male colleagues, and/or these males are actually expressing 'kickback' against the growing minority becoming more tolerant of GBV.

It should also be borne in mind that any valid regression results must pertain to the whole of England and Wales, and cannot be generalised to any Force, in particular.

[17] https://www.bawp.org, last accessed, 2nd March 2018.

Bearing in mind that the Forces did not all have the same percentages of females on their respective Strengths in 2013 (**Figure 2.9**), it is possible that they may not all have evolved to the same 'Kanter' stage by that time. It would be only once all Forces had achieved the unknown minimum critical mass (not necessarily 50%) that the general result could be inferred to all of them.

2.6.2. The question of Truncation

It is clear that *arrests* made on suspicion of any crime are, by no means, a record of all incidents that occurred of that particular crime in an area over a particular period of time. At the fundamental level, many crimes will go unobserved completely, where even the victim does not realise that they have been predated upon (for example, some instances of data cyber-theft). They may, however, come to realise that they have been predated upon, but only after a significant period of time, by when reporting the crime is seen as being a waste of time (for example, being pickpocketed while abroad on holiday in a busy City). Other victims may simply be too embarrassed to report the crime (for 'honour' reasons, for example), or worried that actually reporting the crime may only lead to even greater utility losses (for example, being dismissed from employment). [18] GBV is not particularly different to other types of crime as the perpetrator believes they are gaining something out of it, they are unconcerned about the utility losses experienced by the victim and that they would prefer not to be punished for having committed it. To that end, they would prefer that the crime remained unreported, which would be aided by it being unobserved by any third parties.

Of course, this Study is not concerned, directly with the number of GBV incidents, but the number of arrests made on suspicion of it having been committed. It is reasonable

[18] The introduction of the General Data Protection Regulation Act in the UK in May 2018 allows the Regulator to judicially punish not only the cyberhackers illegally acquiring data from organisations, but also the Organisations themselves: the aim being to incentivise the latter to better defend its data. The unintended consequence, however, may be to discourage many of these organisations from reporting the theft, leaving it to the Judicial System to take on a greater burden of the policing.

to expect that this variable should be positively correlated with the number of incidents (zero incidents should mean zero arrests, presumably), but that there are several impediments to arrests being made:

- The incident has to be either observed by a Police Officer or another third party, who is willing to report the incident to the Police.
- This Third Party, or the (alleged) victim has to be sufficiently plausible to result in the Police investigating the incident further.
- There is sufficient evidence of the incident having taken place to allow the Police to consider making an arrest
- The officer(s) in question believe that the incident is sufficiently serious to warrant an arrest being made (as opposed to no action, or a verbal warning being given to the perpetrator regarding their future conduct)
- The Custody Officer is similarly convinced that the Perpetrator should be arrested and, if necessary, confined pending the next stage in the Judicial Process. [19]
- The Police Service is resourced sufficiently to fulfil its role properly, namely observation, investigation and (where appropriate) arrest.

The fact that even observed/reported incidents do not result in an arrest being made, implies a degree of *truncation* of the dependent variable (GBV Arrests) where an unknown proportion of GBV incidents did not result in an arrest, and that a Tobit-type ('Heckit') specification may be appropriate. Here, the arrest/no arrest decision would be modelled via a Probit equation and used to inform the regression of GBV arrests on the selected independent variables. The judgement here, however, is not to include this first arrest-selection phase, for the following reasons:

[19] There are, of course, stages after that, the first of which is a decision to prosecute or not: these, of course follow the arrest stage, which is the variable o interest I this Study, *not prosecutions or convictions.*

- The arrest/no arrest decision is a complex and, it would appear, a largely un-researched process. There are 'stories' that female officers, for example, are less inclined to make arrests than male counterparts, preferring to 'talk down' conflict situations rather than provoking arrests. This may apply to situations involving alcohol-fuelled breaches of the peace, where the police are effectively asking for 'good behaviour', rather than GBV-type situations where the crime may have occurred already, prior to police attendance. The understanding of the arrest decision is a (large) Study in itself and well beyond the scope of this one. There would be little to be gained from second-guessing what observable factors underlie the decision, and that the resulting parameter estimates may be no better than standard regression ones.
- In any case, even if the process were understood in the literature, and the key variables observable, the data would not be available to this researcher, in the absence of an agreement to access it. This would also imply a move away from the Area-aggregated data used in this Study to Incident-level data. As outlined in **2.7.2**, this would require a trust relationship to be developed with the Home Office and/or all 43 Forces, and the necessary computing and clerical resources available to undertake such a large Study.

2.6.3. The Econometric Models

The significance and, if appropriate, direction of the association between GBV arrests and female participation in the Service will be estimated via a number of regression-based Panel models using the assembled data. Three main types of models will be estimated:

1. OLS Models, with and without year dummies: to provide benchmarks against which the Panel models can be compared and contrasted.
2. One-way Panel Models: Fixed Effects, Random Effects, GLS (including Prais-Winsten estimators).

3. Two-way Panel Models: ditto, with Year Dummies added.

For each model category, two versions of the dependent variable and associated independent variables are modelled:

i. *Aggregate GBV arrests*: regressed against aggregates of female officer numbers, female population and total arrests.
ii. *GBV Arrest rates:* i.e. the *proportion* of arrests made that were for GBV regressed against the *proportions* of females in the police forces and the general population, arrests per capita and household densities.

The data and models are subjected to pre- and post-estimation testing as outlined in Chapter One, and the headline results are reported in each section and in the tables. Note that as this Panel is perfectly balanced, there are no issues with *attrition*.

2.6.4. Aggregates' Models

The Aggregates' Models aim to explain the total number of arrests for GBV in each area using aggregates of the independent variables. [20] Three main models are estimated:

1. OLS models (with heteroskedasticity-corrected errors)

$$GBV\_Arrests_i = \beta_0 + \beta_1 Female\_Officers_i + \beta_2 All\_Arrests_i + \beta_3 Female\_Population_i + u_i$$

2. One-way Panel model (no time effect)

$$GBV\_Arrests_{it} = (\alpha + \beta_{0i}) + \beta_1 Female\_Officers_{it} + \beta_2 All\_Arrests_{it} + \beta_3 Female\_Population_{it} + u_{it}$$

3. Two-way Panel model (time effect included in the error term)

[20] Models with ln(GBV Arrests) and higher powers (squares and cubes) of the dependent variables were also estimated, but not reported here (output and test statistics are available in a separate document).

$$GBV\_Arrests_{it} = (\alpha + \beta_{0i}) + \beta_1 Female\_Officers_{it} + \beta_2 All\_Arrests_{it} + \beta_3 Female\_Population_{it} + (\lambda_t + \varepsilon_{it})$$

Of course, GBV Arrests appear on both sides of the equation: it is the dependent variable itself and is also embedded within the *All Arrests* variable on the right-hand side. However, as **Figure 2.1** reveals, the proportion of arrests made for GBV was small (peaking at 1.4% in 2013). Subtracting the GBV component from the total would have no appreciable effect on the results and otherwise infer that the models were seeking the relationship between GBV arrests and all others, as if the Forces were making some sort of (constrained) choice between them: in fact, what is sought is the effect of all arrests on the numbers made for GBV.

Absent from these specifications are any explicit socio-economic variables, such as unemployment rates, income distribution, Indices of Multiple Deprivation[21], etc. This is, in part, due to the lack of an accepted theory as to the causes of GBV itself, comprised of observable variables. Apart from anything else, the huge range of GBV 'types' (see **Table 2.1**) introduces a huge range of complex, inter-related factors not confined to observable socio-economic ones, such as 'culture' and religion, even the weather and football scores! (Card and Dahl, 2011). One of the acclaimed benefits from estimating models based on Panel Data is, of course, that the unobserved attributes pertaining to the 'unit' are contained within the Fixed Effects. The units, in this case, are the Police Force Areas (**Figure 2.3**) that map almost exactly to unitary Local Authorities in England and Wales. While there will be huge degrees of inter-Area heterogeneity with respect to socio-economic factors, the judgement is that having these multitude of conflicting factors wrapped up into the Fixed Effects is better than trying to disentangle these and expand the number of independent variables, with the

[21] See, for example, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464430/English_Index_of_Multiple_Deprivation_2015_-_Guidance.pdf (accessed 19 July, 2018).

consequent loss of degrees of freedom in the models. Of course, were the data to be at the level of Officers' notebooks (where every incident is recorded), then more variables could be included, while still retaining the benefits of Panel data.

Econometric Issues: It is evident from the descriptives that there is considerable scope for the standard econometric issues. The underlying population grew overall, and in most of the Force Areas (**Figures 2.10** and **2.11**). The number of arrests carried out fell steadily throughout England and Wales but those made for GBV (**Figures 2.1** and **2.5**), rose (**Figures 2.1** and **2.6**). The numbers of police officers deployed remained broadly static overall (**Figure 2.4**), but fell in some Force Areas, while rising in others, in some cases by ±20% (**Figure 2.8**). The numbers of female officers in the Service increased steadily throughout the Study period, as did their percentage of all officers (**Figure 2.4**) but there were significant differences in representation between the Forces (**Figure 2.9**). Thus, we have to be alert to serially correlated errors, even the possibility of a Unit Root and any of the regressions being spurious. **Table 2.2** gives an indication of the huge range in the sizes of the Forces which, with an aggregates-based model, makes it almost certain that the errors from the models will be heteroskedastic. We might also expect there to be a degree of correlation between the independent variables, on the basis that a given Force's strength will be determined, to some extent, by the size of the population it is supposed to police (though this may be a lagged response). The number of arrests carried out should also be correlated with the population and the number of officers deployed.

Given the parsimony of the models, the lack of an accepted universal theory as to the causes of GBV, and arrests for it and the aggregated nature of the data, there is also scope for significant omitted variables' bias in the models. Finally, although the 43 Forces operate autonomously from each other and exhibit different demographics (and changes), they are operating within a common Legal Framework and National Political system and Government. As part of a single Society, they have much in common regarding, for example, attitudes to GBV and the inclusion of females in all levels of their respective bureaucracies. Thus, there is scope for significant inter-Panel correlation, evident in **Figure 2.5**, for example. Prior to estimating any of the models, the independent variables were assessed for multicollinearity (**Table 2.4**). For Total Arrests v Female Population, there is a spread across the full spectrum, with some bunching towards the negative end, a reasonably even distribution across the spectrum for Total Arrests v Female Officers, with some bunching towards the positive end of the scale for Females in the Population v Female Officers. The variance inflation factors for the models with these variables was 20.75 and only 6.77 if year dummies are included. The potential for inefficiency of the standard errors is noted, though no other corrections will be made to counteract this issue.

In order to assess the stationarity of the 258 constituent time series, Dickey-Fuller tests for random walks with no drift, with drift and with drift and a time trend were performed: the Null Hypothesis in all cases is that there is a Unit Root in the Data. Of the 774 computed test statistics, 92% failed to reject Null ($p < 0.01$).

Table 2.4
Independent Variables: partial correlations' distributions (Aggregates)

Offences v Female Population	f	Offences v Female Officers	f	Fem Pop v Fem Officers	f
-0.991	1	-0.931	1	-0.919	1
-0.806	14	-0.744	3	-0.731	1
-0.622	4	-0.557	5	-0.542	1
-0.438	2	-0.370	2	-0.354	2
-0.254	2	-0.183	2	-0.165	1
-0.070	2	0.004	5	0.023	3
0.114	2	0.190	5	0.212	5
0.298	1	0.377	3	0.400	2
0.482	3	0.564	4	0.588	4
0.666	4	0.751	7	0.777	8
0.851	5	0.938	6	0.965	15

Pooled Regressions of GBV arrests against various combinations of female and total officers, population and households along with dummies for Force and Year all displayed the tell-tale signs of spuriousness: R^2s in excess of 98%; highly significant F-statistics and many variables with highly significant t-statistics. The Levin-Lin-Chu test for a Unit Root on data as a Panel, was performed and revealed a mixed set of results, whereby the Null Hypothesis of a Unit Root cannot be rejected for some of the series but can for others (depending upon the type of series assumed). Kao- and Pedroni tests for the existence of long-run cointegrating relationships between the Panel series were carried out and unanimously rejected the Null Hypotheses of no Cointegrating relationships.

The models were estimated using STATA 15, with Robust Standard errors specified. Both Random Effects models were rejected by the adapted Hausman test (160.64, $p<0.001$; 147.35, $p < 0.001$) and are not reported here. The residuals for all the Panel models are mostly approximately normally distributed around a mean of zero, though there is some positive skew in the OLS models. The Ramsey Reset test on the OLS model indicates that there are missing variables ($F = 47.61$, $p<0.001$), that even the inclusion of the higher powers of the independent variables could not overcome ($F = 28.57$, $p<0.001$). There is serial correlation present in the residuals of the Fixed Effects model (possibly to lag 2; $F = 53.447$, $p<0.001$). Both the Breusch-Pagan (χ^2= 1180.67, $p<0.001$), White's ($\chi^2$398.25, $p<0.001$) and LR tests (χ^2=604.93, $p<0.001$) indicate the presence of significant heteroskedasticity in the residuals, as do scatter plots of the residuals against the fitted values. The Fixed Effects models also report significant correlation between the residuals and the fitted values (>0.99).

It was at this stage, that a closer examination of the residuals' plots revealed that both the City of London and the surrounding London Met area were contributing significantly to the non-normality and heteroskedasticity issues. The City of London, as previously discussed, is a highly unusual place, with a very small, male-dominated resident population, and with a very large transient population of workers and tourists. This appears to have had the effect of greatly inflating its arrest rates (and therefore the number of arrests) in comparison to the rest of England and Wales – only its small size did not make this obvious from a review of the data. It also had the poorest gender balance of all 43 Forces (**Figure 2.13**) and a relatively low household density. The Met area is a very different place to the rest of the country, in the way that Capital Cities are, with very high population densities (**Table 2.2**) and otherwise 'pressurised' in the way that other areas are not. Accordingly, these were excluded from the models' estimation, and it is the case that while there are still some issues remaining regarding heteroskedasticity and serial correlation, these were less pronounced than in the models with all 43 areas included. The results are presented in **Table 2.5.** [22]

Structural stability tests, comparing the 2002 – 2013 against the 2002 – 2013 sub-period were unable to reject the null hypothesis of stability for all specifications, indicating that, in spite of the policy changes outlined in **2.2.2** appearing to translate into an increase in the number of GBV arrests throughout England and Wales in 2013 (**Figure 2.1**), the increases at Area level (**Figure 2.7**) were not sufficiently significant to suggest that there was (yet) a 'before' and 'after' to consider. [23]

The Fixed-effects models suffer from non-normal and heteroskedastic errors and are not considered further. Attention therefore focusses on the GLS and Prais-Winsten (PW) results, with the OLS as a comparator. The models agree that the more arrests a Force makes, the more it will make for suspected GBV offences. As the numbers of arrests made each year across England and Wales fell by an average of 203,000, this would translate to a fall in GBV arrests of about 500 per annum.

---

[22] The numbers associated with each model are for ease of reference and used in the STATA 'Do' file that generated the results. Regional Dummies were not specified for these models as the Fixed Effect (the Police Force) map almost perfectly to these.

[23] That said, it was possible to reject the Null of stability at the *aggregate* (i.e. all England & Wales) level. Using the F-test **[1.5]** specified in **1.3**, to test the Null of a Structural Break after 2012 resulted in $F(1,9) = 9.62$, $p < 0.05$.

Table 2.5

Dependent Variable: GBV Arrests

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (N=516) (London Met & City of London excluded 2013 is the excluded category)

Model	**OLS (Het-corrected)**		**Fixed Effects**		**GLS : Panel-specific AR(1)**		**GLS : Common AR(1)**	
	No Time Dummies (1.1.1)	Time Dummies (1.2.1)	One-way (2.1.1)	Two-way (3.1.1)	One-way (2.1.5)	Two-way (3.1.5)	One-way (2.1.6)	Two-way (3.1.6)
Female Officers [24]	0.195***	0.184**	-0.332**	-0.036	0.147*	0.084	0.097	-0.106
All Arrests	0.002***	0.003***	0.002***	0.003***	0.002***	0.003***	0.002***	0.003***
Females Population	0.001***	0.001***	0.005***	0.003**	0.001***	0.001***	0.001***	0.001***
Y 2002		-291.988***		-237.975***		-222.550***		-224.233***
Y 2003		-223.804***		-167.274***		-163.582***		-165.425***
Y 2004		-216.099***		-157.686***		-148.230***		-146.756***
Y 2005		-230.894***		-174.438***		-163.783***		-154.304***
Y 2006		-281.709***		-219.787***		-191.194***		-171.904***
Y 2007	-	-298.080***	-	-243.171***	-	-192.181***	-	-161.389***
Y 2008		-297.416***		-247.161***		-191.968***		-157.356***
Y 2009		-233.597***		-189.958***		-144.478***		-111.858***
Y 2010		-202.377***		-171.636***		-112.818***		-84.879***
Y 2011		-202.301***		-184.979***		-110.381***		-92.083***
Y 2012		-159.313***		-152.862***		-101.646***		-89.342***
Constant	-59.742***	179.236***	-2242.390***	-1283.498	-71.369***	102.421***	-59.815*	69.009**
R^2	0.8843	0.9049	0.1498	0.3679	-			
Wald X^2	-				1252.62***	2839.96***	736.32***	1435.18***
F	929.057***	271.486***	18.406***	13.406***	-			
r(u_i, Xβ)	-		-0.9902	-0.9796				
F (Structural Stability)			0.19	0.75	0.14	1.24	0.41	0.69

[24] Of course, adding or removing a *male* officer may also have an impact that could, for instance, be even greater than the female officer. Regressions using male and all officers were estimated to reveal that in almost all instances, female officers had a significantly different impact (either in magnitude or both magnitude and sign to their male counterparts).

Table 2.5 (continued) Dependent Variable: GBV Arrests * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (N = 516) (London Met & City of London excluded 2013 is the excluded category)				
Model	**Prais-Winsten Panel-specific AR(1)**		**Prais-Winsten Common AR(1)**	
	One-way (2.1.8)	Two-way) (3.1.8)	One-way (2.1.7)	Two-way (3.1.7)
Female Officers	0.163	0.175*	0.188*	0.138
All Arrests	0.002***	0.003***	0.002***	0.003***
Females Population	0.001***	0.001***	0.001***	0.001***
Y 2002	-	-283.900***	-	-291.360***
Y 2003		-215.382***		-221.341***
Y 2004		-206.723***		-212.801***
Y 2005		-220.882***		-226.440***
Y 2006		-271.371***		-274.798***
Y 2007		-289.396***		-291.900***
Y 2008		-292.189***		-291.090***
Y 2009		-229.864***		-228.212***
Y 2010		-198.241***		-198.459***
Y 2011		-197.135***		-200.132***
Y 2012		-160.029***		-158.969***
Constant	-81.069**	161.552***	-66.482**	144.370***
R^2	0.8063	0.9448	0.9079	0.8540
Wald X^2	1092.31***	3423.95***	1694.69***	1625.47***
F	-			
$r(u_i, X\beta)$				
F (Structural Stability)	0.35	1.99	0.31	3.82

The percentage of arrests made for GBV was 1.39% in 2013, which would rise to 1.46% in 2014 *if* the averaging behaviour were projected forwards. In other words, the propensity to arrest for GBV would continue to increase, despite a continued fall in the numbers of arrests made.

The models also agree that the more females there are in the population, the more arrests there will be for GBV, with one more arrest for every additional 1,000 females in the population. Looking again at the historic data, the population of England and Wales grew at an average of ~360,000 people per annum, while the percentage of the population that were female fell by an average of 0.5% per annum. This would translate into an additional ~157,000 females in 2014 implying an additional 157 arrests for GBV. In 2013 it required 603 females to generate one arrest for GBV, but in 2014 this would have fallen to 601 females per arrest. This might be seen as either evidence that females are being predated on more, and/or that they are more willing than before to report such offences

The year dummies in the two-way models indicate that the number of GBV arrests increased as time passed, though there was some reversal in the mid-2000s: evidence, perhaps, that Society and the Judicial System are altering their attitude to this crime type, and/or that Society is developing more of a taste for committing GBV against itself.

Of course, the primary variable of interest is the numbers of Female Officers deployed by each Force. Here, the consensus breaks down somewhat. The OLS models are persuaded that ~ five additional female officers deployed would result in an extra GBV arrest being made, as are the one-way GLS-PSAR1 and PW-AR1 models (though the GLS-PSAR1 model would be looking for ~seven additional female officers). The other GLS models do not, however, discern any significant effect. Looking at the historic data, the Forces combined added an average net 1,070 female officers per annum to the Service's Strength. Assuming any of the models that found a significant relationship applied, then this would have resulted in an additional 150 – 200 arrests for GBV per annum. This would imply a *fall* in the average number of GBV arrest per female officer from 1.35 to 1.31 were the Service to continue building up its female complement at this rate. The models that do find a significant, and positive link also imply that the presence of an extra female officer had between 150-200 times the impact that the presence of the additional female citizen did.

The evidence supporting the hypothesis is therefore mixed, with statistical significance depending upon the variables included and the assumptions made about the error structure. However, it seems that the existence of any *negative* relationship overall, between the numbers of female officers and GBV arrests, an indication that the Service may have progressed beyond Kanter's *Skewed* stage: this will be considered again in **2.7.1**.

2.6.5. Rates Models

The aggregates models indicate some evidence for improved representativeness in the Police Bureaucracy over the Study Period. However, the use of *absolute*, rather than *relative* measures may be criticized for not capturing the effect of the balance of forces that will have come together in both the general populations and workplaces of the different Forces. For instance, the number of female officers present in a Force does not necessarily imply that they are stronger as a sub-group if the numbers of male officers, for example, kept pace. Similarly, in the population, the number of females in the population does indicate something about the number of potential targets there are for GBV, but not how many of them there were in relation to the (potential) perpetrators. The total number of arrests carried out in an Area does not necessarily imply anything about the general levels of lawlessness there: the population may have grown or shrunk, or the police more or less zealous in making arrests. Intuitively, expressing the dependent variable, GBV Arrests, as a percentage of all arrests made may give a better insight to the *relative* priority the Service in England & Wales attached to this crime type than the sheer number of such arrests would. Accordingly, the aggregates variables were combined together, as follows:

1. OLS models (with heteroskedasticity-corrected errors)

$$\frac{GBV\ Arrests_i}{Total\ Arrests_i} = \beta_0 + \beta_1 \frac{Female\ Officers_i}{All\ Officers\ _i} + \beta_2 \frac{Total\ Arrests_i}{Population_i} + \beta_3 \frac{Population_i}{Households_i} + \beta_4 \frac{Female\ Population_i}{Total\ Population_i} + \upsilon_i$$

2. One-way Panel Model (no time dummies)

$$\frac{GBV\ Arrests_{it}}{Total\ Arrests_{it}} = (\alpha + \beta_{0i}) + \beta_1 \frac{Female\ Officers_{it}}{All\ Officers\ _{it}} + \beta_2 \frac{Total\ Arrests_{it}}{Population_{it}} + \beta_3 \frac{Population_{it}}{Households_{it}} + \beta_4 \frac{Female\ Population_{it}}{Total\ Population_{it}} + \upsilon_{it}$$

3. Two-way Panel model (time effect included in the error term)

$$\frac{GBV\ Arrests_{it}}{Total\ Arrests_{it}} = (\alpha + \beta_{0i}) + \beta_1 \frac{Female\ Officers_{it}}{All\ Officers\ _{it}} + \beta_2 \frac{Total\ Arrests_{it}}{Population_{it}} + \beta_3 \frac{Population_{it}}{Households_{it}} + \beta_4 \frac{Female\ Population_{it}}{Total\ Population_{it}} + (\lambda_t + \upsilon_{it})$$

The GBV Arrest rate is expressed as a *percentage* of the total arrests, as a proxy measure of the *relative* priority given to that crime type. Percentages of females in the

Forces and in the general population are included as independent variables to capture (i) any effect the female officers may be having on the propensity to make GBV arrests, and (ii) the effect of the gender composition of the Area. The arrests per capita are included to measure the apparent 'lawlessness' of each Area (or the Force's willingness to make arrests. The household density, as measured by the mean number of people in a household was included in order identify what effect, if any, there is from the extent to which people in a given area are corralled together in domestic dwellings. This is a nod towards the socio-economic conditions that may have prevailed, though it should be stressed that this is likely to influence different sub-categories of GBV in different ways.

Econometric Issues: A number of econometric issues were identified with the aggregate-level data. Whether or not any of these will be ameliorated with the rates' variables remains to be seen, but it is to be hoped that the issue of *heteroskedasticity* in the error terms may be avoided by the effective de-scaling of the data. An inspection of the dependent variable, GBV Arrest% (**Figure 2.7**) confirms a steady increase in the percentage of arrests made for GBV over the period (all other Forces exhibited similar trajectories). This, of course, is not surprising, given that we already know that the absolute number of GBV arrests (**Figure 2.6**) rose against a backdrop of falling numbers of arrests in general (**Figure 2.5**), and indicates that, as before, we may anticipate an issue with serially-correlated error terms in the regressions.

Neither is it any surprise that the gender ratio in the 43 forces tilted, generally towards balance (**Figure 2.9** and **Figure 2.14**, below). Again, the upward trend and, possibly, the differences in the magnitude in the ratios between the Forces may result in serial correlation and/or heteroskedasticity. **Figure 2.15** plots the gender ratios in the general population of three of the 43 Areas. Given that most populations throughout the World comprise more women than men, in part due to naturally-occurring longer life expectancies, we would not expect the regional ratios to be exactly 1:1. As noted in

2.5.3 and in **Figure 2.12**, there was gentle movement towards balance due, perhaps to relatively improved male life expectancies and the tendency for young males to immigrate into the UK. **Figure 2.13** noted the moderate, downward movements toward lower household densities in all Areas, and their relative proximity.

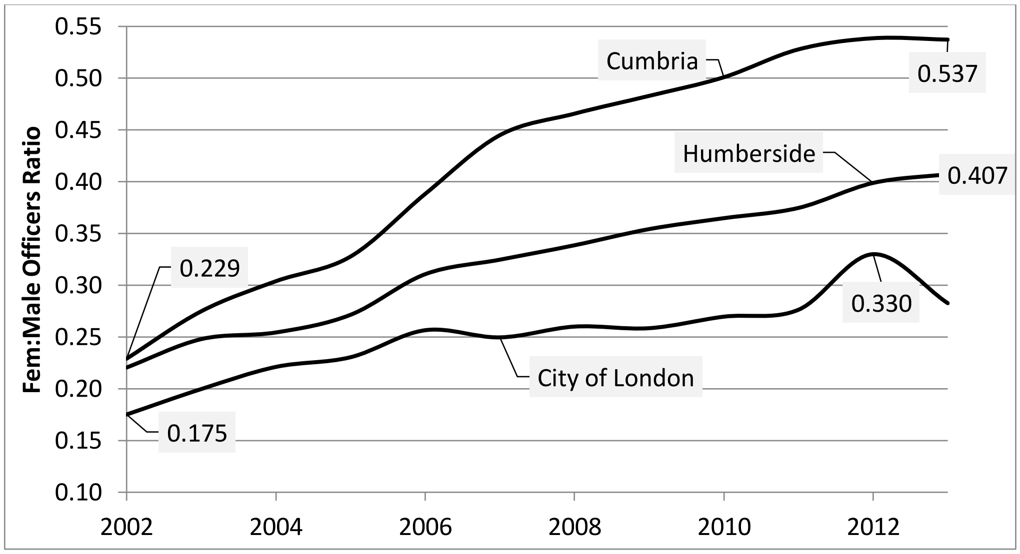


Figure 2.14 Female-to-Male Officers Ratios: Forces with the highest, median and lowest average Ratios (2002 – 2013)

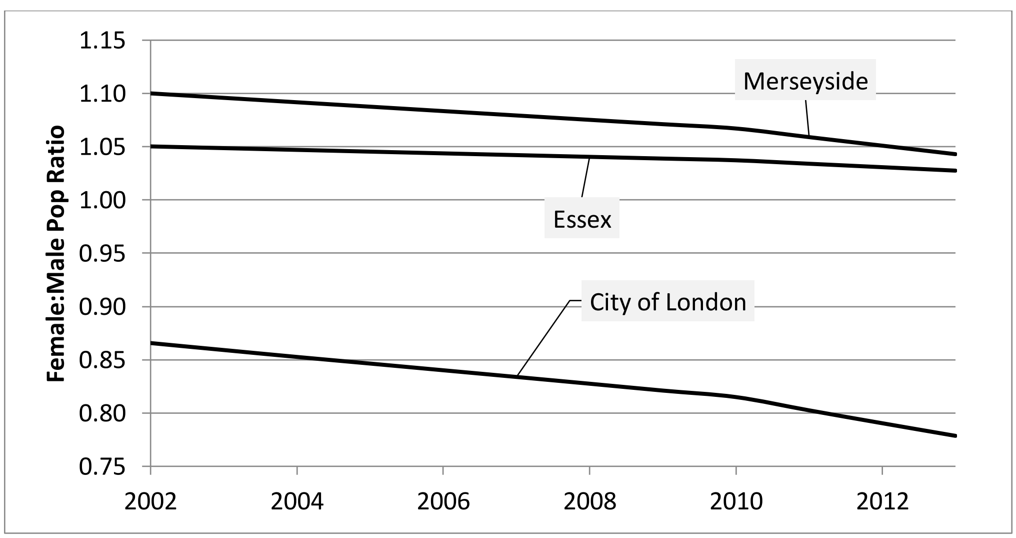


Figure 2.15 Female-to-Male Population Ratios: Forces with the highest, median and lowest average Ratios (2002 – 2013)

Figure 2.16 plots the arrest rates for the highest, median and lowest average performing Forces. The City of London rate is so much higher (11-12 times the average in rest of England and Wales), that its series is plotted against the right-hand scale. The graph shows all Areas reducing their arrest rates, albeit at different speeds: the London Met effectively halving its, while Dyfed-Powys managed a more modest 29%. Again, a progressive, downward-pointing series that might be expected to contribute to a serial correlation issue. **Table 2.6** contains the frequency distributions for the partial correlations between the dependent variables. The general impression is that multicollinearity is less of an issue with the rates' models than with the aggregates (**Table 2.4**), though some positive correlations are noticeable between Arrests per capita and the population gender ratio; household density and arrests per head of population (though all less than +0.7), and; household density and the population gender ratio. These looser correlations are supported by the vifs of 3.66 and 3.18, if year dummies are included. Of the 645 Dickey-Fuller tests performed, 85% failed to reject the existence of Unit Roots in the data, and a similarly mixed picture emerged from the Lenin-Liu-Chu, Kao and Pedroni tests.

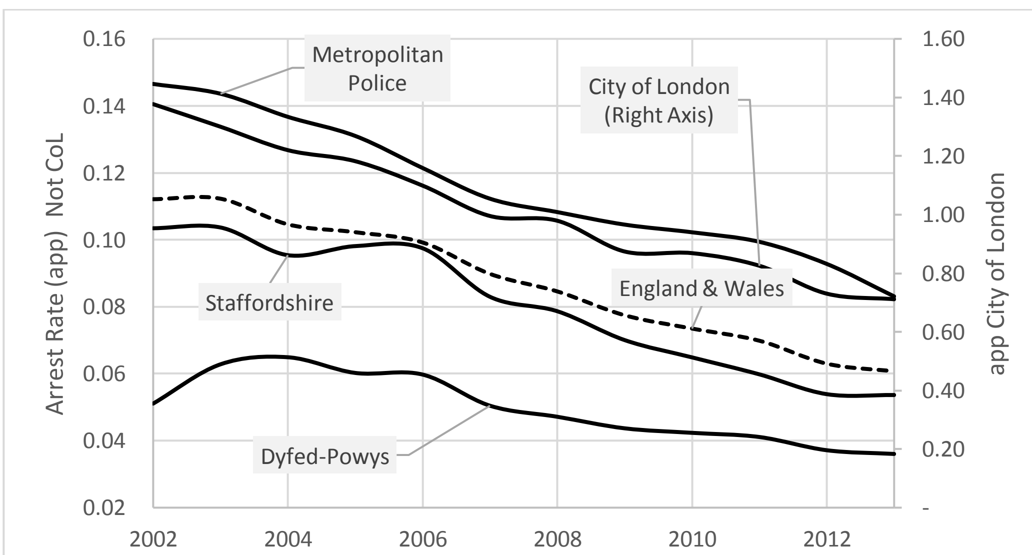


Figure 2.16 Per-Capita Arrest Rates: highest, median and lowest average rates 2002 - 2013

Table 2.6 Independent Variables Partial correlations' distributions (Rates)											
Arr per cap-Fem:Male Officers		Fem:Male Pop-Fem:Male Officers		Arr per cap-Fem:Male Pop		House Dens-Arr per cap-		House Dens-Fem:Male Officers		House Dens-Fem:Male Pop	
r	f	r	f	r	f	r	f	r	f	r	f
-0.863	1	-0.873	1	-0.057	1	-0.539	1	-0.795	1	-0.848	1
-0.710	5	-0.724	4	0.043	2	-0.420	2	-0.662	2	-0.674	0
-0.558	7	-0.575	6	0.143	2	-0.301	2	-0.530	5	-0.501	2
-0.405	10	-0.426	6	0.244	5	-0.181	0	-0.397	4	-0.328	3
-0.253	2	-0.277	8	0.344	4	-0.062	6	-0.265	6	-0.155	5
-0.100	5	-0.128	8	0.444	2	0.057	4	-0.133	3	0.019	6
0.053	3	0.022	1	0.544	5	0.176	2	0.000	8	0.192	3
0.205	5	0.171	5	0.644	6	0.296	6	0.132	2	0.365	9
0.358	2	0.320	1	0.745	4	0.415	7	0.265	7	0.538	5
0.510	0	0.469	1	0.845	4	0.534	6	0.397	3	0.712	6
0.663	3	0.618	2	0.945	7	0.653	6	0.530	2	0.885	3

Structural Stability The aggregates' models of the previous section were unable to reject the null hypotheses of structural stability at the Area level, though not at the all-England/Wales level. The same tests were performed on the rates' variables assuming a structural break at 2013 at the Area Level, and all bar model 9.1.7 (Prais-Winsten AR)1) were able to reject the null of stability to at least p = 0.01. [25] It appears that the upturn in *rates* as opposed to *aggregates* in 2013 was sufficiently marked to signal a regime change in the policing of GBV. Accordingly, for the purposes of this Study, the data for 2002 to 2012 were used to estimate the rates' models summarised above. The regressions results are presented in **Table 2.7**. Both Random Effects models are, once more, rejected by the adapted Hausman test (96.51 (p<0.001; 26.97 (p < 0.001)) and are not reported here. The residuals for all the Panel models are mostly positively skewed. The Ramsey Reset test on the OLS model indicates that there are missing

[25] The all-England/Wales test indicated F(1,8) = 1.433, implying a failure to reject the null of Structural Stability. Interestingly, the same test performed on the data to 2016 (still assuming a break in 2013) results in F(1,11) = 1088.41 (p<0.000) implying that a significant break was coming at the time of this Study's endpoint. The F-statistics shown in the last lines of both parts of **Table 2.7** are based on the full period, 2002-2013.

variables (F = 48.26, p<0.001) that the inclusion of the higher powers could not overcome (F = 41.93, p<0.001). There is serial correlation present in the residuals of the Fixed Effects model (possibly to lag 2; F = 40.434, p<0.001). Both the Breusch-Pagan (χ^2= 12.02, p<0.001), White's (χ^2=112.67, p<0.001) and LR tests (χ^2=173.32, p<0.001) indicate the presence of significant heteroskedasticity in the residuals, as do scatter plots of the residuals against the fitted values. The One-way Fixed Effects models also report significant correlation between the residuals and the fitted values (>0.99), but the Two-way models fare better (minus-0.253 and minus-0.362).

As with the Aggregates' Models, a closer examination of the residuals' plots revealed that both the City of London and the surrounding London Met areas were again, outliers: the City of London appears to have had its arrest rates inflated in comparison to the rest of England and Wales (as noted in **Figure 2.16**). It also had the poorest gender balance of all 43 Forces (**Figure 2.13**) and a relatively low household density.

Also, as discussed, the Met, although de-scaled in these rates' models, is itself, a very different place to the rest of the country. Accordingly, both were, again, excluded from the estimation of the Rates' models. Doing so, improved the diagnostics considerably compared to those reported above, although heteroskedasticity and serial correlation are still issues.

As with the aggregates' models, the direction and significance of some of the key parameters are dependent upon the assumptions made about the structure of the error terms. Given the persistent presence of some of the econometric issues, the focus is on the GLS/PW models. Looking first at the Arrests-per-capita (app) variable, there is consensus across virtually all the models that the relationship between it and the GBV arrest rate is *negative*. Given that the Arrests-per-capita fell consistently in all areas over the Study Period, it appears that one consequence of a lowered general arrest rate, is an increased focus on GBV.

Table 2.7 **Dependent Variable: GBV Arrest Rates v Rates' Variables** * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (N = 451) (London Met & City of London Excluded 2012 is the excluded category)								
Model	**OLS (Het-corrected)**		**Fixed Effects**		**GLS: Panel-specific AR(1)**		**GLS: Common AR(1)**	
	No time Dummies (7.1.1)	With time Dummies (7.2.1)	One-way (8.1.1)	Two-way (9.1.1)	One-way (8.1.5)	Two-way (9.1.5)	One-way (8.1.6)	Two-way (9.1.6)
% Female Officers	0.003	-0.014***	0.002	0.005	0.003	-0.004	0.004	-0.002
Arrests per capita	-0.046***	-0.033***	-0.035***	-0.004	-0.052***	-0.019***	-0.049***	-0.018***
Household Density	-0.009***	-0.005***	-0.011*	-0.008	-0.008***	-0.006***	-0.008***	-0.004**
% Females Pop	-0.171***	-0.127***	-0.460*	-0.036	-0.198***	-0.088**	-0.239***	-0.059
Y 2002	-	-0.004***	-	-0.004***	-	-0.004***	-	-0.004***
Y 2003		-0.003***		-0.003***		-0.003***		-0.003***
Y 2004		-0.003***		-0.003**		-0.003***		-0.003***
Y 2005		-0.003***		-0.003***		-0.003***		-0.003***
Y 2006		-0.003***		-0.004***		-0.003***		-0.003***
Y 2007		-0.003***		-0.003***		-0.003***		-0.003***
Y 2008		-0.003***		-0.003***		-0.003***		-0.003***
Y 2009		-0.002***		-0.002***		-0.002***		-0.002***
Y 2010		-0.001**		-0.001***		-0.001***		-0.001***
Y 2011		-0.001*		-0.001***		-0.001***		-0.001***
Constant	0.120***	0.095***	0.273*	0.049	0.132***	0.073***	0.153***	0.053**
R^2	0.4627	0.5699	0.6384	0.7483				
Wald X^2								
F	76.792	32.727	53.331	46.185				
F (Stability)	10.62**	-	22.21***	-	34.05***	-	23.74***	-

Table 2.7 (continued) **Dependent Variable: GBV Arrest Rates v Rates' Variables** * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (N = 451) (London Met & City of London Excluded 2012 is the excluded category)				
Model	**Prais-Winsten Panel-specific AR(1)**		**Prais-Winsten Common AR(1)**	
	One-way (8.1.8)	Two-way (9.1.8)	One-way (8.1.7)	Two-way (9.1.7)
% Female Officers in Force	0.003	-0.013^{**}	0.003	-0.010^{*}
Arrests per capita	-0.053^{***}	-0.026^{***}	-0.051^{***}	-0.028^{***}
Household Density	-0.010^{***}	-0.005^{**}	-0.008^{***}	-0.004^{*}
% Females in Population	-0.171^{***}	-0.109^{**}	-0.192^{***}	-0.107^{**}
Y 2002	-	-0.005^{***}	-	-0.004^{***}
Y 2003		-0.004^{***}		-0.003^{***}
Y 2004		-0.003^{***}		-0.003^{***}
Y 2005		-0.003^{***}		-0.003^{***}
Y 2006		-0.004^{***}		-0.003^{***}
Y 2007		-0.003^{***}		-0.003^{***}
Y 2008		-0.003^{***}		-0.003^{***}
Y 2009		-0.002^{***}		-0.002^{***}
Y 2010		-0.001^{***}		-0.001^{***}
Y 2011		-0.001^{***}		-0.001^{***}
Constant	0.124^{***}	0.084^{***}	0.130^{***}	0.081^{***}
R^2	0.8082	0.8813	0.3311	0.4811
Wald X^2				
F				
$r(u_i, X\beta)$				
F (Struc.Stability)	41.10^{***}	176.67^{***}	28.69^{***}	7.64

There is a similar degree of consensus between the models that a lower household density *increases* the rate of GBV arrests, though the amount varies according to the specification used. This seems to support the idea that perpetrators are taking advantage of the reduced density to isolate their victims and to commit GBV with less fear of third parties witnessing it, to the extent that any benefits implied from having more space to live in are unable to counteract this tendency.

There is also strong consensus that there is a *negative* association between the GBV arrest rate and the female-to-male population ratio. This translates to a tilt in the numerical domination away from females towards parity, resulting in an increase in the percentage of arrests made for GBV. For example, if model 8.1.5 were selected, an area with 51% females, falling to 50%, would see a rise of approximately 0.34% in the percentage of arrests that were made for GBV. Assuming the arrest rate is correlated with the incidence rate, then this might, for example, be explained by a fall in the number of un-partnered females being subjected to intimate-partner GBV.

The models with year dummies included reveal that the passing of the years since 2002 *increased* the arrest rate for GBV, reflecting, perhaps, that Society and the Forces are adopting an evolving intolerance for this type of crime in line with the concerted International crackdown on VaWG and the UK Government's efforts to implement the provisions of the Istanbul convention outlined in **2.2.2**.

Of course, the variable of most interest is the Female-to-Male Officers ratio. With the aggregates' models (**Table 2.5**, earlier), the evidence of the effect was a little mixed, and depended upon the specification chosen. Here, things are a little more clear-cut, in the sense that it does not appear to be the case that increasing the ratio of female to male officers in a Force *increased* the proportion of arrests made for GBV, either leaving it unchanged, or actually lowering it. The numbers of officers in service was broadly the same at either end of the Study Period (**Figure 2.8** with a fall of 4,100 officers in total) though the number of female officers increased by around 13,000 while the number of male officers fell by around

17,000. Accordingly, the Forces can be said to be better gender-balanced, but it is not (yet) the case that this is increasing the GBV arrest rate.

2.7. Discussion and Future Work

2.7.1. Discussion

The Study has used a Panel of 12 years' aggregate-level data of arrests, police numbers and demographics for England and Wales to test the hypothesis that increasing the number and the proportion of female officers, has resulted in a counterfactual increase in the numbers and rate of arrests for GBV. The results from the analysis are far from conclusive: the significance and direction of the effect of the female officers is dependent on the specifications of the models, the variables included, and assumptions made about the error structure.

In the case of the Aggregates' models, the presence of more females in the general population increases the numbers of GBV arrests made, as do the numbers of arrests made in general. The effect of the number of females is not surprising, as there are more 'targets' on which males can predate. The relationship with the number of arrests made in general indicates that Forces that making more arrests implies that more will also be made for GBV: of course, as **Figure 2.1** shows, the number of arrests made across England and Wales actually fell from ~6million to 3.5million between 2002 and 2013, implying that the number of GBV arrests would have been reduced, as a result: the fact that they rose by about 10,000 per annum over the period (also **Figure 2.1**) indicates that other factors, including the extra number of females in the general population more than counteracted this downwards effect. The problem, however, is not being able to discern whether that is because part of Society is increasing its appetite for GBV, while another part is increasingly reviled by it, or that the alleged victims are more willing to come forward to report it, or that the Service is better able to detect it. None of these reasons are mutually exclusive, but the relative impact of each cannot be distinguished.

Turning to the key variable of interest: the number of female officers serving in each Force, only the OLS models and the one-way GLS model indicate a significant, positive relationship with GBV, while the one-way Fixed Effects actually establishes that more female officers actually lowers the number. All two-way Panel specifications fail to establish any significant link, with the population and arrest counts explaining the variation, along with the year effects. As noted at the end of **2.6.3**, the location of the Service, with respect to the gender balance of the Forces in Kanter's framework was at the *Skewed* stage, where female officers, as a minority may have actively sought to lower the GBV arrest rate in an effort to blend in with the male majority, can probably be ruled out. Kanter asserted that a bureaucracy would transit from this stage to the *Tilted* stage when the minority group had attained around 35% of total workforce count. Looking at the Gender composition of the Forces, only one of them (Cumbria) had achieved that level by 2013, with the England/Wales average at 28% (**Figure 2.4**). Thus, the transition to a *Tilted* bureaucracy, if it ever took place, did so below 35% representation.

Looking at the Senior ('ACPO') levels of the Forces, they lag behind in terms of their gender balance, though there has also been progress. Between 2002 and 2013, the average percentage of female ACPOs rose from 7% to 16%: the highest average representation in 2013 was 33% (Cheshire). This lag is to be expected, of course, given the time it takes for officers to reach that level. In any case, it might be argued that given the number of people involved at the higher levels is small compared with those on the street, it would be easier for one or two senior female voices to be heard at the top table, implying that the transition there might actually take place even sooner.

Turning to the Rates' models (**Table 2.7**), the proportion of GBV arrests falls as the percentage of females in a population increases. Of course, as **Figure 2.12** shows, the percentages actually fell through the Study period, implying a counterfactual rise in the GBV arrest rate. Similarly, the GBV Arrest rate moves in opposition to the general arrest rate, which as figures like **2.1** and **2.5** imply, also fell throughout the period, therefore implying a counterfactual increase in the GBV rate. It is tempting to conclude that the Forces are, perhaps, focussing more on GBV arrests, at the expense of other crime-types, having to do so given that their resources have not expanded to at least maintain their efforts on all fronts.

All specifications, except the one-way Fixed Effects model find a negative relationship between household density and the GBV arrest rate (it finds no significant effect at all). It was left moot what the net effect might be, with it being possible to argue, intuitively, that it could go either way: an increased density could increase domestic 'pressures' leading to more GBV, or it could lower it, due to fewer opportunities to predate, unwitnessed. The evidence appears to favour the latter causation (though this could be the result of a netting-off effect). Noting that **Figure 2.10** indicates that household density fell on average from 2.42 people to 2.39 between 2002 and 2013, this implies that there were more chances to be 'alone' and that the GBV rate would have increased, counterfactually.

Apart from the OLS model with time dummies, all models failed to find a significant relationship between the GBV rate and the percentage of females in the Forces. Given that the Forces were at the *Skewed* Stage, it might be concluded that the females officers were continuing to ape male officers' behaviour (assuming that, as women, they do have a different attitude from them to GBV), or that, in fact, they don't have a different attitude and that gender is irrelevant, in this respect. Final confirmation of this might only come when all Forces reach 'balance', which at the rates witnessed by 2013 will not come until 2040.

While the Study has focussed on the effect on GBV arrests, there are also several benefits to the Police Service from working towards gender parity in its staffing. As well as potentially offering an improved and otherwise more relevant service to the Public it serves, it is clearly better for young females interested in a career in the Service to know that their gender is no longer an impediment to them. From the Service's (and Public's) perspective, it must be better that it can choose from a wider pool of suitable talent, no longer limiting itself to the smaller pool of predominantly young males. As a consequence, the average quality of the officers should rise, assuming that the females admitted at the margin are more able than the males excluded at the margin.

Finally, Societal Welfare must be being increased where the safety and security of the majority (women and children) is being given a greater priority than before. The longer-term challenge for Society is to eliminate the incentives for males (in particular) to commit GBV, and to allow them to avoid the negative personal consequences from doing so.

2.7.2. Future Work

<u>Data</u>: this Study used publicly-available, free data covering the period 2002 – 2013. Four years have passed since the initial collection of the data, and it would be relatively straightforward to add more data to the Panel, with the option to roll off the older years. Given the extent of omitted variables' bias in the models, it seems necessary to look for other variables that appear to be influencing the numbers and rates of arrests for GBV. These may pertain to socio-economic data, such as the levels and distribution of wealth and income, more refined data on the composition of the population beyond mere aggregates. Including variables on the levels of education may be key here. However, un-aggregated incident-level data is always going to be preferable, allowing the genders, ages and socio-economic profiles of the victim, the perpetrator and the officers to be controlled for (it may also allow the London Forces to be reattached to the Study). Data like this, are highly confidential, and would require a trust relationship to be initiated and developed with the Home Office and devolved Administration in Wales. As discussed in **2.6.2**, this would facilitate a two-stage Heckman-style regression approach, with the first stage being the arrest-decision Probit, then the main regression along the lines of those estimated in this Study. This, of course, assumes that there is some understanding about the Arrest Decision which may, in turn require to be researched.

As well as this overtly quantitative data, a *qualitative* study could be designed to elicit, in particular, the arrest decision-process to be better understood and to ascertain the extent to which, if any, that the officers' gender played a role. Also, the developing relationship between officers as the female contingent has grown at street- and senior levels could be explored. This would elicit a better idea of to what extent female representation is truly *active*, in the sense that the female officers are explicitly agitating for a change in attitude, or it is *passive* i.e. it is just the increased numbers that is slowly tilting the culture of the Service. Within that, a study of female officers at the senior, decision-making levels would be key. Again, relationships and mutual trust would have to be developed, and funds acquired, to undertake the research.

<u>Model Specification</u>: The models estimated here have all been single-variable models, with no instrumentation. The specification could be expanded to a simultaneous equations' structure, with separate equations for the generation of GBV incidents, the decision to staff each Force with a given number of officers, the arrest decision, etc. There is also the issue of modelling the female-male officer relationship to reflect, better, the non-linear relationship of the changing minority-majority balance. As it stands, the straight ratio of female to male numbers apes a gradually acceleration in the balance but does not capture the logistic-type trajectory where change occurs in mid-range, tapering off as parity is reached.

<u>Scope</u>: The Study could be easily extended to any other territories that collect and make available similar statistics. The author wanted to include Scotland in the Study, where the arrest laws are different to those in England and Wales. In particular, two officers must be present to make an arrest, reflecting the higher levels of corroboration required in its Legal System. Unfortunately, the data are not readily available on-line in the way that the Home Office data are. When the Scottish Government's Justice Department were contacted, it placed several Freedom of Information hurdles in the way of the author, preventing acquisition of even any aggregated data. Again, a trust-and-share relationship would have to be established to allow that region to be involved in any future Study.

References used in Chapter Two

Andresen, M.A. (2012) 'Unemployment and Crime: a neighborhood level panel data approach', Social Science Research 41, 1615 - 1628

Becker, G. (1968) 'Crime and Punishment', Journal of Political Economy, University of Chicago Press 76, 2, 169 - 217

Bishop, P., Cureton, K. and Collins, M. (1987) 'Sex difference in muscular strength in equally-trained men and women', Ergonomics 30, 675-687

Bradbury, M. and Kellough, J.E. (2011) 'Representative Bureaucracy: Assessing Current and Additional Evidence of Active Representation', American Review of Public Administration 41, 2

Brown, J., Maidment, A. and Bull, R. (1993) 'Appropriate Skill-Task Matching of gender bias in deployment of male and female police officers?', Policing and Society, Harwood 3, 121-136

Brown, J. (1998) 'Aspects of discriminatory treatment of women police officers serving in Forces in England and Wales', British Journal of Criminology, Oxford Publishers 38, 2, 265-282

Cabinet Office (2010) 'Call to end violence against women and girls: taking action – the next chapter,' HMSO, 978-0-10-851143-1

Cameron, S. (1987) 'A disaggregated study of Police Clear-up rates for England and Wales', Journal of Behavioral Economics 16, 4, 1 – 18

Card, D and Dahl, G.B (2011), 'Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior', *Quarterly Journal of Economics*, 126, 1, 103–143

References used in Chapter Two (continued)

Carr-Hill, R.A and Stern, N.H. (1973) 'An econometric model of the supply and control of recorded offences in England and Wales', Journal of Public Economics 2, 289 - 318

Chesterman, C. and Ross-Smith, A. (2006) 'Not tokens: reaching a 'critical mass' of senior women managers', Employee Relations 28 6, 540-2

Commission of the European Union (2012) 'Establishing minimum standards on the rights, support and protection of victims of crime', Official Journal of the European Union, L315/57

Cornwell, C. and Trumbull, W.N. (1994) 'Estimating the Economic Model of Crime with Panel Data', Review of Economics and Statistics 76, 2, 360 - 366

Council of Europe (2011), 'Convention on preventing and combating violence against women and domestic violence', Report on the Istanbul Convention, Council of Europe Treaty Series - No. 210

Dahlerup, D. (1988) 'From a small to a large minority: Women in Scandinavian Politics', Scandinavian Political Studies 11, 4, 275 – 298

Dick, G. and Metcalfe, B. (2007) 'The progress of female police officers? An empirical analysis of organisational commitment and tenure explanations in two UK Police Forces', International Journal of Public Sector Management, Emerald 20, 2, 81-100

Ehrlich, I. (1975) 'The deterrent effect of Capital Punishment: a question of Life or Death', American Economic Review 65, 397 - 417

Etzkowitz, H., Kemelgor, C., Neuschatz, Mi., Uzzi, B. and Alonzo, J. (1994) 'The paradox of critical mass for women in Science', Science 266, 5182, 51-54

References used in Chapter Two (continued)

European Institute for Gender-based violence (EIGE) (2014) ‘Administrative data sources on gender-based violence against women in the EU: current status and potential for the collection of comparable data’.

Falls, G.A., Thompson and Philip B. (2014) ‘Casinos, casino size and crime: a Panel data analysis’, Quarterly Review of Economics and Finance 54, 123 - 132

Frieze, I.H. and McHugh, M. (2005) ‘Female violence against intimate partners: an introduction’, Psychology of Women Quarterly, Blackwell Publishing 29, 3,

Garcia, V. (2003) ‘"Difference" in the Police Department’, Journal of Contemporary Criminal Justice, Sage 19, 3, 330-344

Grey, S. (2006) ‘Numbers and Beyond: the relevance of Critical Mass in Gender Research’, Politics and Gender, Cambridge 2, 491 - 530

Holdaway, S. and Parker, S.K. (1998) ‘Policing Women Police’, British Journal of Criminology, Oxford Publishers 38, 1, 40-60

Home Office (2010) ‘The Stern Review: A Report by Baroness Vivien Stern, CBE, of an independent review into how rape complaints are handled by Public Authorities in England and Wales’, HMSO

Iparraguirre, J. (2014) ‘Hate crime against older people in England and Wales – an econometric enquiry’, Journal of Adult Protection 16, 3, 152-165

Johnson, K and Houston, J. (2018) ‘Does Female Police Leadership Affect Gender Based Violence Arrests?’, Int. Review of Administrative Sciences 84, 1, 3 - 20

References used in Chapter Two (continued)

Kanter, R.M. (1977) 'Men and women of the Corporation', Basic, NY

Keiser, L., Wilkins, V., Meier, K. and Holland, C. (2002) 'Lipstick and Logarithms: Gender, Institutional Context, and Representative Bureaucracy', American Political Science Review 96, 3, 553 - 564

Kite, D. and Tyson, G.A. (2004) 'The impact of perpetrator gender on Male and Female Police Officers' perceptions of Child Sexual Abuse', Psychiatry, Psychology and Law 11, 2, 308 - 318

Lipsky, M. (1980) 'Street-Level Bureaucracy', New York: Russell, Sage

Martin, C. (1996) 'The impact of equal opportunities policies on the day-to-day experiences of women police constables', British Journal of Criminology, Oxford Publishers 36, 4, 510-528

Meier, K.J. and Nigro, L G (1976) 'Representative bureaucracy and policy preferences', Public Administration Review 36, 458-470

Merton, R.C. (1957) 'Social Theory and Social Structure', New York: Free Press.

Mosher, F. (1982) 'Democracy and the public service', New York: Oxford University Press. 2nd edition.

Nielsen, V.L. (2014) 'Differences in Male and Female Employees' Personal Attributes? Myth or a reasonable assumption: even within professions?', Gender Issues, Springer 31, 163-184

Parkinson, C.N. (1958) 'Parkinson's Law: The Pursuit of Progress', John Murray, London

References used in Chapter Two (continued)

Rabe-Hemp, C. E. (2008) 'Female Officers and the ethic of care: does Officer gender impact Police behaviors?', Journal of Criminal Justice, Elsevier 36, 426-434

Riccucci, N.M. and Saidel, J.R. (1999) 'A response to James F. Guyots' "Representative Bureaucracy Recast"', Public Administration Review 59, 4, 369

Russo, N.F. and Pirlott, A. (2006) 'Gender-based Violence: Concepts, Methods and Findings', Annals of the New York Academy of Sciences 1087, 178-205

Stewart, A. and Maddren, K. (1997) 'Police Officers' judgements of blame in family violence: the impact of gender and alcohol, Sex Roles', Plenum Publishing 37, 11/12, 921- 933

Thielemann, G. and Stewart, J. (Jr) (1996) 'A demand-side perspective on the importance of representative bureaucracy: AIDS, Ethnicity, Gender and Sexual Orientation', Public Administration Review 56, 168-173

UN, 48/104. (1993) 'Declaration on the Elimination of Violence against Women', General Assembly, 48/104, December (http://www.un.org/documents/ga/res/48/a48r104.htm).

UN-HABITAT (2006) 'State of the World's Cities'

Visher, C.A. (1983) 'Gender, Police Arrest decisions and notions of Chivalry, Criminology', American Society of Criminology 21, 1, 5 - 28

Weber, M. (1922) 'Economy & Society', University of California Press

3. Estimating returns to education: wage rate premia in the UK Economy: A Panel-based Quantiles' Regression Approach

3.1. Introduction

Jobs were once plentiful throughout the Industrialised World, the majority of which required nothing more than a rudimentary grasp of basic reading, writing and arithmetic: only the Professions required anything substantially beyond this. There was no obvious need to estimate the economic value individuals and economies obtained from education, particularly, post-compulsory education. Until 1918, children in the UK were only required to complete some formal education to the age of 12 when it rose to 14, then again to 15 in 1947, and finally to 16 from 1972. School Leavers mostly went into a life of stable work or homemaking, largely according to their gender. In 1953, only 30% of 16-17 year olds in the UK were still at School, rising slowly to 47% by the late 1960s. In 1953 only 11% of School Leavers in England and Wales left with five or more 'good' Ordinary-level awards (21% by 1969) and only 6% with at least one Advanced-level award (18% by 1969). There was a rapid rise in the provision of places in Further Education (FE) Colleges after 1945, but these were intended, initially, for de-mobilised ex-Service personnel who required re-skilling. In 1920, only 5,060 people obtained a University Degree, rising to 19,747 in 1950, then 25,699 in 1960 (all statistics quoted by Bolton, (2012)).

Using data collected by the UK's Office of National Statistics (ONS), Denman and McDonald (1996) noted that the increase in unemployment in the Great Recession of the early 1930s did not result in any significant increase in the uptake of post-compulsory education, but corrected itself via market forces, with the considerable assistance of World War 2 (**Figure 3.1**). The mid-to-late 1970s can now be seen as the time during which the UK Labour Market transited to a higher level of unemployment as a process of rapid de-industrialisation began. This led, *inter alia*, to the breakdown of the Apprenticeship system and a consequent rise in (male) youth unemployment. In a 1976 study, it was noted that:

'longer term unemployment among school leavers [had become] a serious problem and … to have become a more permanent feature of the economy'

- highlighting a near quadrupling in School Leavers' unemployment in July 1976 compared with July 1975 (Dean, 1976).

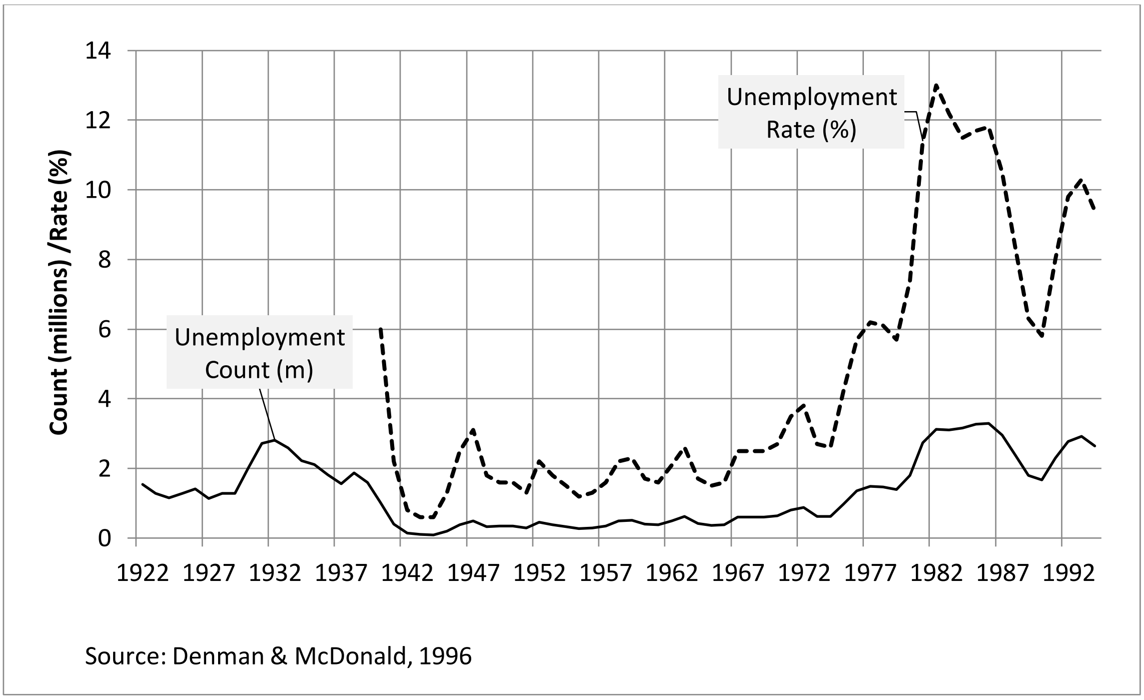
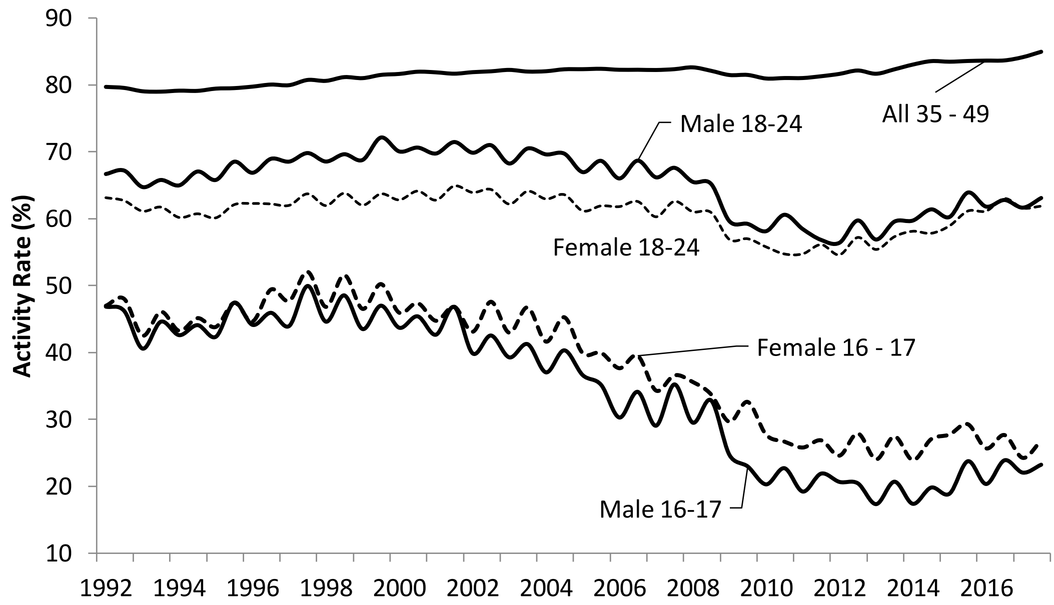


Figure 3.1 UK Unemployment - Count & Rate: 1922 - 1994

Focussing on the more recent past, the rates of young persons' economic *activity* (i.e. either employed or seeking work) stayed well below that of middle-aged workers (**Figure 3.2**). While the rate for 18-24 year olds improved slightly from its post-2008 recession low, they still represented over 30% of that age group. Unsurprisingly, the younger 16-17 group had a lower activity rate compared to all others, but the fall in their activity rate was rather more marked and persistent. Also noticeable was the closing of the gap between males and females in the 18-24 group, and the more marked decline in activity rates for 16-17 aged males. The pre-Thatcher (1979) policy response to unemployment was to stimulate the macroeconomy with a package of fiscal measures and to create jobs for 'young' people, either directly in the Public Sector, or via training and employment subsidies to private-sector employers. From the 1980s onwards however, the emphasis switched to using post-compulsory, formal Education as the means of both increasing young persons' and redundant older workers' employability, against a background increasingly insecure employment and rising underemployment. Bolton (2012, op.cit.) estimated that by 2009, 86% of 16-year-olds and 76% of 17-year-olds were still at School (or Sixth Form College) in England and Wales, while more than 80% left with five or more 'good' Ordinary-level awards and 36% with at least one Advanced-level qualification.



Source: (ONS, using the LFS)

Figure 3.2. Economic Activity Rates (UK) 1992 - 2017

The English and Welsh Department for Education and Skills (DfES), charted the rise in the numbers of students enrolled in FE Institutions, from 3.3 million in 1980 to 4.3 million in 2010 (cited in Bolton (2012, op.cit.). Participation in University-level Higher-level (HE) education had increased slowly from 3.4% of School Leavers in 1950, to 8.4% in 1970, thanks to the establishment of a number of new Institutions and the conversion of existing Technical Colleges into Universities.

This was followed by a further process of College-to-University conversion and expansion of existing Universities in the early 1990s. By 2014, the post-School participation rate for School Leavers was estimated to be 48% in England and Wales (DfES, 2016) and 55% in Scotland (Scottish Funding Council, 2015). **Figure 3.3** charts the increase in the number of accepted applications between 1994 and 2017, with a near doubling in the proportion of 18-24 year olds enrolled in *HE* programmes.

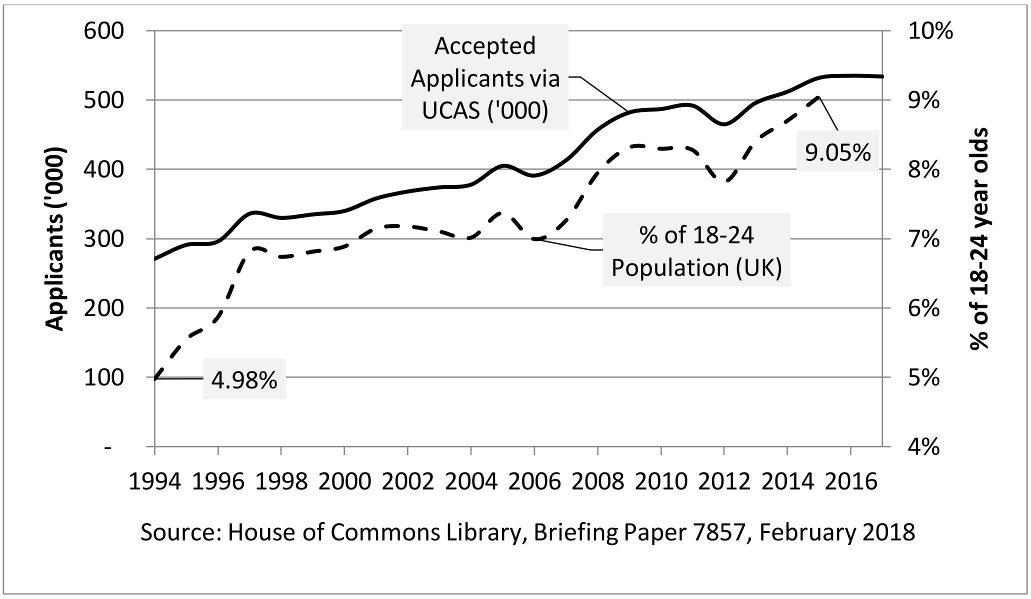


Figure 3.3. Accepted Application to UK HE Institutions and proportions of 18-24 year olds enrolled in Degree Programmes in the UK

The question arises: who paid for this ever-increasing time spent in formal post-compulsory education? While there was a significant private provision of School Education in the UK, the bulk of the spending came from the Public Purse. **Figure 3.4** plots World Bank Data for the G7 Economies from 1980 to 2012 (World Bank, 2013). The UK devoted between 4.5% to 5.6% of its GDP on Education in the period, placing it towards the upper end of the range of the G7 in 2012. The Institute for Fiscal Studies calculated the real-terms spending on Public Education from 1953 to 2015 (**Figure 3.5**), noting both a rapid relative and real-terms rise through the 1950s and 1960s that peaked in the mid-1970s. There then followed a period of decline in the GDP share through the 1980s and 1990s, although real spending actually increased (IFS, 2015). Spending on education as a percentage of GDP rose from the late 1990s, with relative spending peaking around the mid-1970s level in 2008. The Global Financial Crisis of 2008 and the consequent austerity measures taken by the UK Government has resulted in a fall in both real and relative expenditure on Education that is being experienced still, at the time of writing.

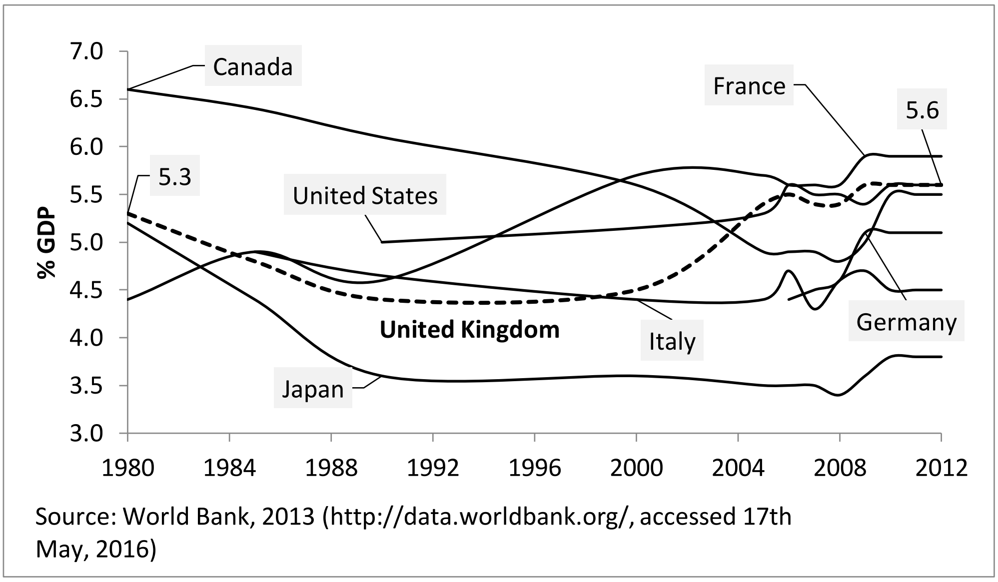


Figure 3.4. Percentage of GDP spent on Public Education: G7 Economies (1980 - 2012)

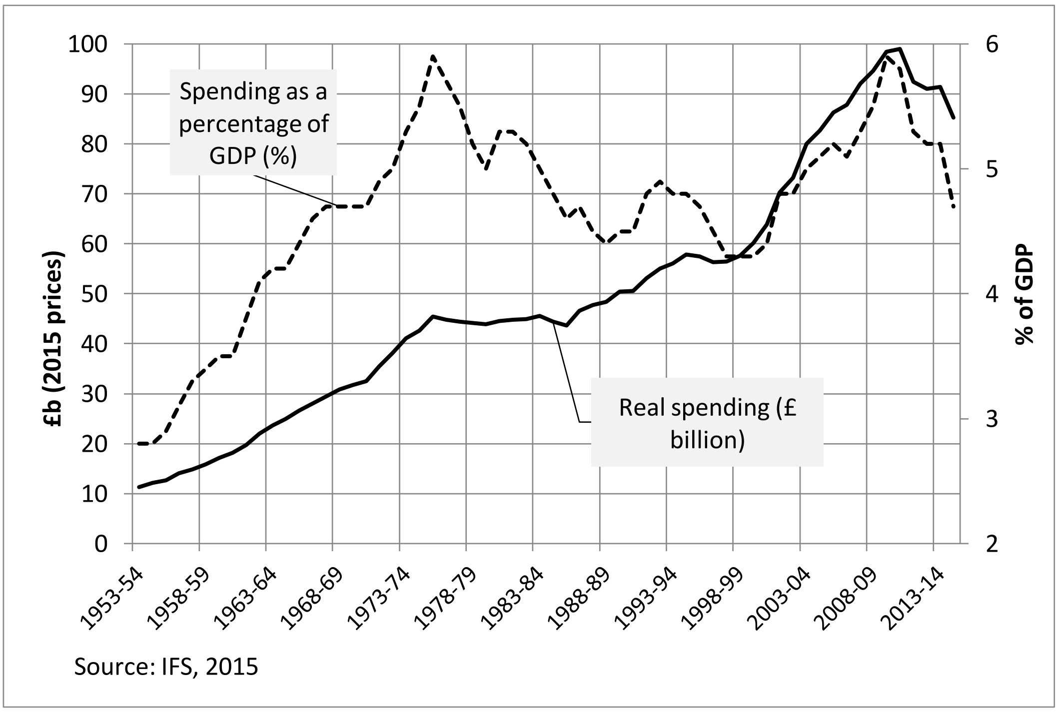


Figure 3.5. UK Real spending on Public Education 1953 - 2015

The IFS noted however, that the sectoral composition of the spending changed, as the UK's Central and Devolved Governments sought to protect politically-sensitive spending on State Schools, at the expense of Further- and Higher Education. As **Figure 3.6** shows, the nominal funding provided by the English/Welsh Funding Council (HEFCE) to Universities declined rapidly from just under £8billion in 2009/10 to less than £4billion in 2016/17.

Scotland's *pro rata* support via its Funding Council (SFC) was higher, historically, than England and Wales', reflecting both its greater number of HE Institutions per capita and a greater uptake of post-compulsory education. While the SFC did not reduce its support to the same extent as HEFCE, spending was reduced in 2015 and again in 2016 as, the Scottish Government sought to prioritise early-years' Schooling.

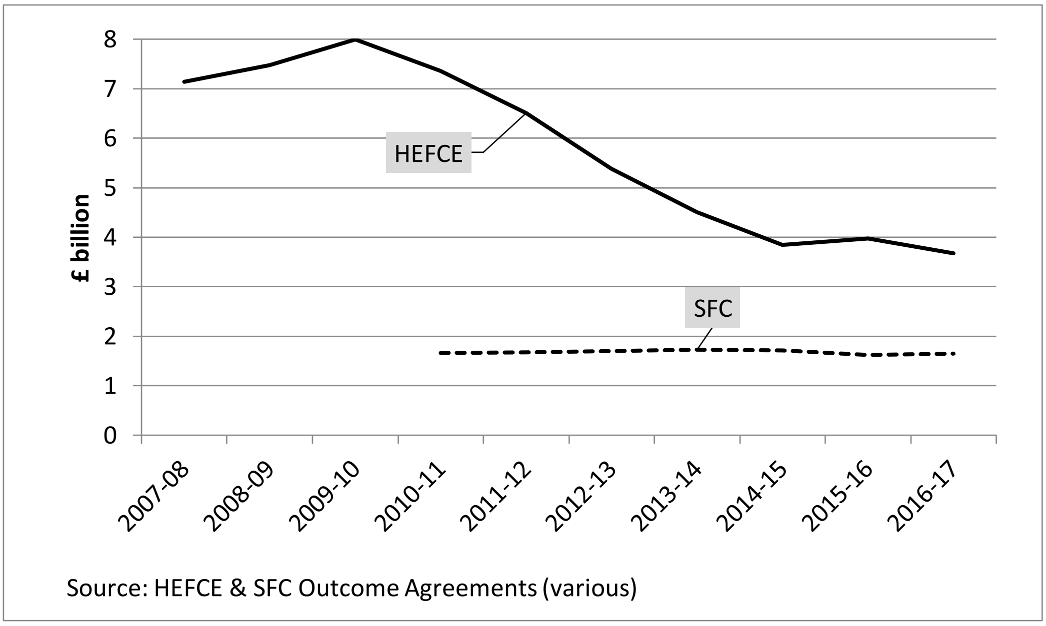
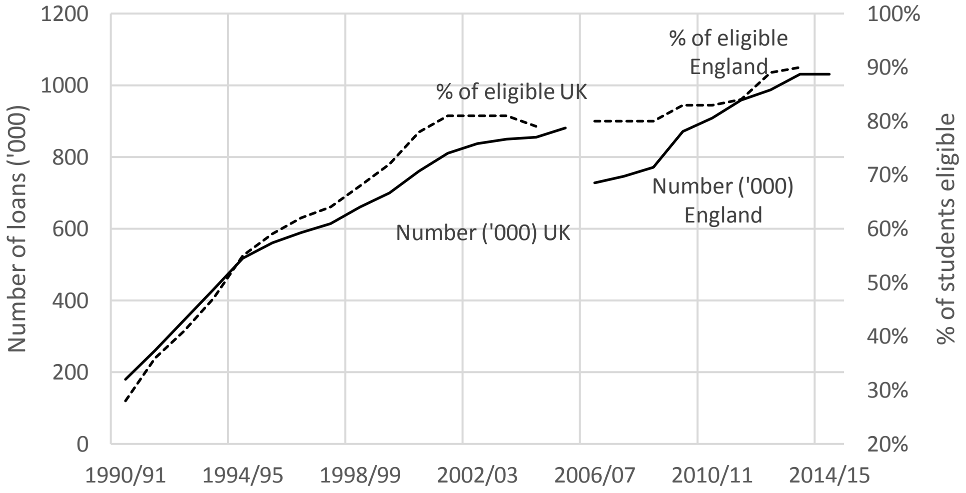


Figure 3.6. HEFCE & SFC Grant Support to Universities 2007 – 2016 [26]

As **Figures 3.7** and **3.8** show, the uptake of student loans increased, as grants were replaced with maintenance loans (1999) and tuition fee caps were raised in England (£3000 = 2006; £3225 = 2009; £9,000 = 2010, to rise with inflation from 2017). The devolved Governments in Scotland, Wales and Northern Ireland made efforts to

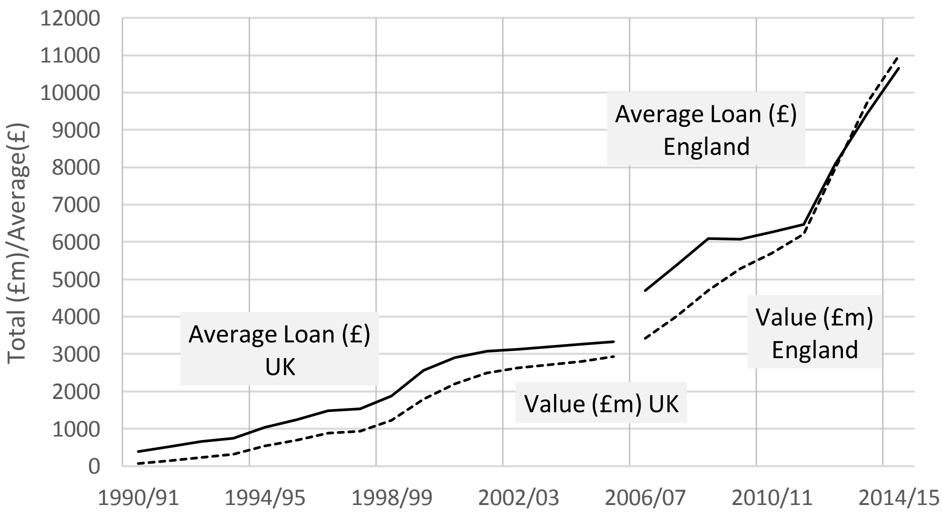
[26] Comparable Scottish data not available prior to 2010/11.

ameliorate the impact of the transfer of costs onto students. Scottish students do not (yet) pay tuition fees if they study in Scotland, while fees in Northern Ireland are currently capped at £4,160. Welsh Universities can charge the same as their English counterparts but grants to support living costs are available to Welsh students studying anywhere in the UK.



Source: Bolton, 2016

Figure 3.7. Student Loans' Takeup 1990 - 2014



Source: Bolton, 2016

Figure 3.8 Value of loans taken out 1990-2014

Proponents of the value of post-School education point to data such as those depicted in **Figure 3.9** that show much lower inactivity rates for graduates compared to non-graduates. However, it is possible to discern a rise in the rate between 1993 and 2017, particularly for 'non-recent' graduates [27], at a time when non-graduate inactivity had edged down, slightly. This says nothing about the type of work that the active graduates do, and there a consensus has emerged that there is significant and rising, graduate *underemployment*, i.e. graduates in occupations that do not require a degree. Green and Henseke (2017) estimate that the rate of graduate under-employment in England and Northern Ireland in 2011 was around 30% (though this was not only an issue in those labour markets, noting a rate of 50% in Japan at the same time).

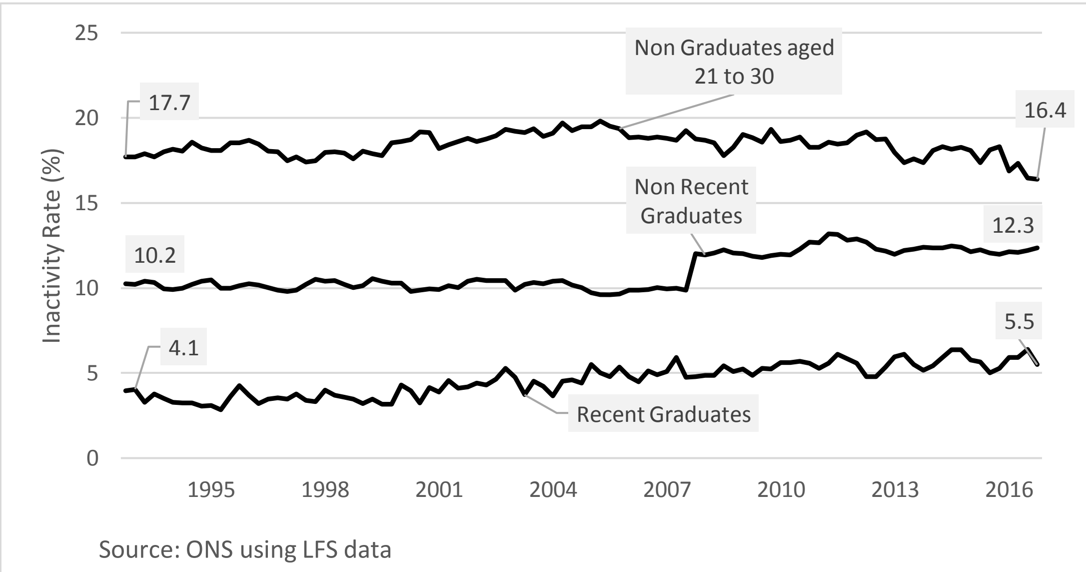


Figure 3.9. Graduate and Non-graduate Inactivity Rates UK 1993 – 2017

The increased personal cost of post-compulsory education and the weakening of the guarantee of true graduate level work soon after graduation, has focussed minds on the economic returns to degrees, and other post-compulsory qualifications. This, in turn, has continued to fuel the academic interest in estimating what these returns are, how they compare with those in the past and,

[27] ONS define 'recent' as someone who had graduated within the last five years at the time of observation.

crucially, what the trends appear to be into the future. This is considered in the next section.

3.2. Outline Empirical Literature

While the academic interest in estimating the economic value of formal qualifications pre-dates the 1980s, the switching of the onus from the Public Purse to the recipient (and assumed benefactor) of the qualification, combined with a much more competitive labour market, has focussed the interest on the *private*, personal returns to an individual's education. This, in turn, has both sustained and sharpened the research intensity in this area, as interested parties seek assurance that the investment is still worthwhile, both from the recipient's and Public Purse's different perspectives. The economic 'return' can be viewed as some combination of:

(i) The increased probability of earning an income from (any) employment, and

(ii) The extra (lifetime) earnings the individual gets from acquiring the education, given that they are employed.

- assumed to have been as a result of acquiring the extra education (certificated, or otherwise). The post-compulsory education decision can be viewed in a similar fashion to investing in a piece of capital equipment, where the individual decides to take the education if, and only if, the lifelong net present value of the marginal benefits (ΔB) accruing to the certification exceed the discounted present value of the costs of acquiring it (C), i.e.

$$\sum_{t=x}^{n} \frac{\Delta B_t}{(1+r)^t} > \sum_{t=x-k}^{x} \frac{C_t}{(1+r)^t} \qquad \textbf{[3.1]}$$

where:

- ΔB_t is the (private) marginal benefit accruing to the individual at time t.

- C_t is the cost of the education required to obtain the certification, at time t.
- r is the rate of discount used to equate future costs and benefits to a 'present-day' value.

This assumes that the certification is obtained in time x following a course of education, duration k; that no re-education is required thereafter to maintain the certificate's currency and; that the discount rate (r) applies throughout the individual's working life (finishing at n). Of course, this can be altered to account for relaxations in these assumptions. The cost of the certification is relatively easy to estimate, given its immediacy and the predictability of at least some of the opportunity costs of acquisition. Much harder to estimate are the marginal benefits of the certification, so much so that individuals are, even to this day, happy to make the decision based on partial and otherwise hazy information.

While the benefits will include intangibles, such as 'power' in the workplace, social contact and status and the inherent quality of the work, the focus is usually on the tangible, private monetary benefits accruing to the individual, insofar as these can be estimated objectively. To this might be added some notion of the value of any retirement pension, though given that most people still in education are 'young', this may not be expected to feature large in their calculations. While many jobs are not overtly remunerated at an hourly rate, but instead on the basis of an annual wage with a contracted, or notional number of associated hours, it is assumed that individuals 'guessstimate' an implied hourly wage rate. Given that, they allocate their time to labour, leisure and homemaking offering a certain number of hours to the employer (Becker, 1965). From this standpoint, there developed a literature based on stylized Econometric models of various Labour Markets, that sought to explain the (hourly rate of) pay as a function of observed characteristics, viz.

$$w_i = \alpha + \beta X_i + \gamma Q_i + \varepsilon_i \qquad \textbf{(3.2)}$$

where:

- w_i is individual i's wage (rate)

- X_i is a vector of observations of individual *i*, on a number of 'non-educational' variables assumed to influence their wage rate.
- Q_i is a vector of observations of individual *i*, on a number of 'educational' variables assumed to influence their wage rate.
- α is a constant term. In this context, that part of the wage rate that cannot be assigned to the observed characteristics, applicable to all.
- β and γ are vectors of parameters to be estimated.
- ε_i *is the residual / error term*

This equation is usually termed a *Mincer Equation* as a result of Jacob Mincer's application of this approach to wage-rate estimation and the effect upon it from an extra's year post-compulsory schooling (Mincer, 1958). Of particular interest is the vector γ as its elements are assumed to indicate how the hourly rate of pay is affected by a particular amount or level of certificated education.

The relevant literature falls under two categories:

(i) general theory of human capital and signalling, and

(ii) specific literature on the Returns to Education.

Much of the literature focusses on the wage gains associated with education. [28] High levels of education and the positive attributes that potential employers are inclined to infer from them, act as a signal to the labour market about the holder's productivity (i.e. their skill or level of ability) relative to others without that qualification. This signal, based on the association between productivity, intelligence, skills and other characteristics associated with a degree, may indicate the extent to which a perspective employee is more or less qualified when compared with other potential candidates (Spence, 1973).

More specifically, the value associated with qualifications, expressed in terms of their ability to increase graduates' earning potential, represents a significant area of interest in which a considerable amount of empirical research continues to be conducted. In the context of this research area, the 'value' associated with a

[28]For a thorough review of returns to education, see Blundell et al (1999).

qualification is expressed in terms of the extent to which the qualification marginally increases the graduate's employability and/or their wage, taking into account demographic variables such as age, gender and geographical location, etc. Davies, et al. (2013) highlighted the emphasis prospective students attach to the potential wage premium associated with their degree, thereby amplifying the importance of wage returns in their course selection process. Earlier research conducted on UK graduates focussed primarily only on the returns associated with degree qualifications, i.e. those obtained at the University level. A more diverse approach was applied as researchers acknowledged the inherent heterogeneity of qualifications. As a result, other research distinguished between different types and combinations of qualifications that can be attained from a variety of educational routes: Blundell et al, (1997); Robinson, (1997), Dearden et al, (2002) and, for Scotland, Gasteen and Houston, (2007). These studies confirmed that degree qualifications are the most effective at enhancing earnings.

Additional research employed a regional perspective in assessing the returns to qualifications within specific countries: O'Leary and Sloane (2005); Strauss and De La Maisonneuve (2009); Kelly et al (2010); Chevalier (2011); Carnoy et al (2012); and Walker and Zhu (2011). Lindley and McIntosh (2015) also found that there is a growing wage inequality between graduates in different disciplines. They all found, to varying degrees, positive returns associated with higher education within the specific labour markets studied. The approaches of these studies differed insofar as their analysis quantified the returns associated with specific degrees or fields of study, rather than a broad analysis of qualifications, irrespective of the field to which they apply. Such research is a first step towards quantifying the value of specific degrees. A caveat regarding these studies is that wage returns are calculated on the basis of *general* employment. In applying this approach, they fail to take into account any wage differentials caused by different employment outcomes. Some recent studies, such as Freier et al. (2015) who examined returns to Law degrees, failed to account for the potential impact of occupation on these returns, although Lalley, et.al (2018) considered the effect of employment 'setting' on returns to specific degree-disciplines, and found that these can vary greatly. In effect, the traditional approach of the returns to education literature implies that the returns associated with a given degree are equal across all possible employment outcomes.

3.3. Estimating the returns to qualifications: the case for Quantile Regression

3.3.1 Quantile Regression: a brief overview

Least-squares or Maximum-likelihood methods are usually used to estimate the parameters in equations like **[3.2]**. They focus on the mean value of w conditional on the expected values of the independent variables (in *X* and *Q*), i.e. *E[w|X, Q].* As Angrist and Pischeke observed:

> *'…95% of Econometrics is concerned with averages… the focus on averages is partly because it's hard enough to produce good estimates of average causal effects….. But many variables, such as earnings…have continuous distributions. These distributions can change in ways not revealed by an examination of averages; for example, they can spread out, or become more compressed [around their 'local' means]. Applied economists increasingly want to know what is happening to an entire distribution, to the relative winners and losers, as well as to the average.'*

(Angrist and Pischke, 2009)

Quantile Regression (QR) aims to estimate the parameters at key quantiles of the dependent variable's distribution. These are often at the median, and the first and third quartiles, but many studies, like this one, also include the key quantiles such as the 10th and the 90th centiles. These, and the quartiles, are often points of interest in income/wealth studies, as the dependent variable's distribution is often significantly skewed towards the lower end of the relevant distribution. As a result, the *mean* may not truly represent the 'middle' ground of the distribution in terms of the individuals comprising the sample (or population) but may be better indicated by the *median*. Some welfare benefits may be available to individuals whose earnings are below a certain decile (i.e. relatively, and maybe absolutely low-paid). Workers in receipt of such benefits might behave differently to those who do not qualify for such the loss of entitlement, as it may lead to a higher marginal rate of tax, reducing their total income, while the high-paid group may be more inclined to accept it. Low-paid workers may benefit much more from successfully completing a certificated course and will have a greater incentive to undertake the associated course of instruction. Finally, a Degree might be essential for someone in a higher-paid job (such as Medicine), allowing them to benefit from the higher associated earnings, while it would only increase the level of *over-education* for someone in a low-paid job. It can be anticipated, therefore, that any earnings-related study would better informed by a quantile, as opposed to a mean-based regression analysis.

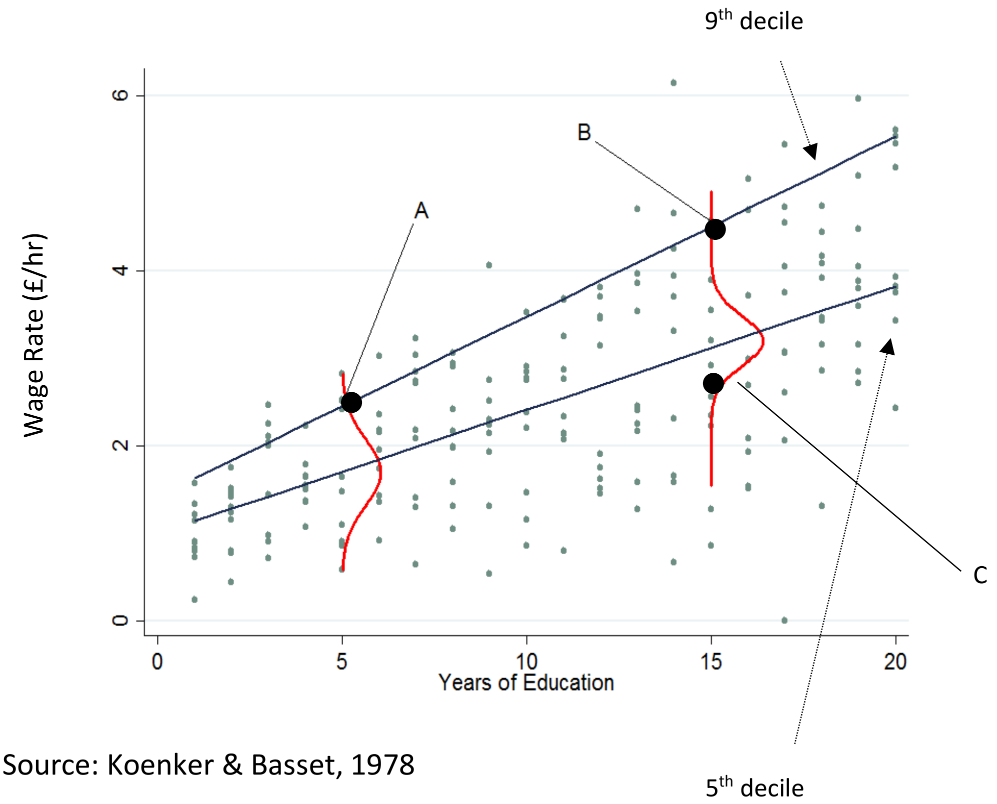
The Conditional Quantile Regression (CQR) model was first proposed by Koenker and Bassett (1978), building on work by Mosteller and Tukey (1977). Its underlying principle is basically that of the Least-squares (LS) model, namely that there is a dependent variable, *y*, that is assumed to be a function of one or more independent variables, *X*, plus a constant (α) and an error term, ε, Whereas LS methods look to minimise the sum of the squared errors ($\Sigma\varepsilon_i^2$), CQR aims to minimize the absolute sum of ε_i, i.e.

$$\sum|\varepsilon_i| = \sum\left|y_i - \beta_0 - \sum_{k=1}^{n}\beta_k X_i + \varepsilon_i\right|$$

by finding the optimal $\boldsymbol{\beta}$. This is known as the Least-absolute deviations estimator (LAD) and, in its simplest form, would be the median regression function, where 50% of the observations lie ‘below’ the function, and 50% above, each ε_i being equally weighted. The LAD estimator can be generalised to any quantile, *q*, by altering the weights given to the ε_i. Each quantile (τ), therefore, has a $\beta(\tau)$ that minimises the LAD. These are found by numerical search or Linear Programming Methods and are, in effect, a series of planes that are above an ever-increasing percentage of the data points as you move up to higher and higher quantiles. Note that *all* the data are used in estimating these, and not just the data at the quantile in question. The estimates are semi-parametric, insofar as there is no particular assumption made about the distribution of the error terms, often heralded as a considerable benefit in applied work, such as this Study.

CQR estimates the *relative* position of an individual within a group of other individuals who are identical with reference to a particular variable. For example, suppose there are three workers, A, B and C. A and C both earn £2.50 per hour, while B earns £4.50. A has acquired five years of education; B and C, 15 years each. Clearly, B earns more than A and C and has more education than A, but the same as C. A and B are located at the 9th deciles of their respective distributions, while C is located at the 5th decile. A least-squares regression based on means, might allow us to conclude that more education increases the earnings' performance. If all the individuals with five years' education were gathered together into one group, and all those with 15 years into another, two virtual distributions could be formed. Membership of a particular group is *conditional upon* having the appropriate number of years of education. **Figure 3.10** indicates the virtual distribution for both groups. Although many of those in the 5-year group are relatively low-paid compared to the 15-year group, individual A is relatively highly paid amongst her group peers. Person B is also relatively highly paid compared with his group peers: C, however, is relatively poorly paid in their peer group, but would have appeared to be well paid if they had belonged, instead, to the 5-year group.

It may be that A differs in respect to some other attribute (for example, observed factors like location, gender, age, sector, occupation and unobserved ones like ability) that places her at the upper end of their conditional distribution (they are a 'big fish, in a small pond'). B may also possess that same 'other' attribute(s) that makes him different to most of the others in the 15-year group, while C does not.



Source: Koenker & Basset, 1978

Fig 3.10 Conditional Quantiles

It is possible to try and answer empirical questions of interest with a CQR model, such as 'what would be the effect on the wage rate of obtaining a degree?' In this example, person A would leave the 'five-year' group and join a group to the right, say the 'eight-year' one. By the same process, B and C would join the '18-year' group. The critical assumption with CQR, however, is that the individual maintains the *same relative decile* in their new group as they had in the old. Thus, A and B who were at the 9th decile before the extra education, would be assumed to locate at the 9th decile in the 18-year group (indicated in **Figure 3.10** with the dark circles). The effect on earnings would be measured by the movements up the y-axis resulting from their respective shifts along the x-axis (this is what would be reported in any results' table generated). This is a very strong assumption and there is probably no reason to expect why it must hold. For example, the individual acquiring the extra education may simultaneously be given a significant promotion and be relocated in a higher quantile in their new group. Alternatively, nothing may happen to them other than they move to the new group, and therefore down to a lower quantile within it.

Firpo, et al (2009) proposed the *Unconditional Quantile Regression* model (UQR), where individuals take their place in the *entire* distribution of the dependent variable. The focus, now, is on the effect on the division of the *new* distribution of the treatment being applied to an individual, holding all their other attributes constant. So, for example, A is given another year's education and they join the 'six-year' group. Assuming that this group has higher median earnings than the 'five-year' (as seems likely), the marginal effect will be to slightly increase the median *for the whole group*. In fact, as one individual moves up the distribution, all the quantiles change position, as they have to represent a given percentage of the changed group. As STATA uses CQR estimators, this will not be considered further in this Study.

3.3.2 Quantile Regression: Panel Data

The extension to Panel Data dates back to Koenker (2004), who established a Fixed-effects QR Model, the main aim being to identify the unobserved α_i assumed to be independent of the quantile they inhabit, while estimating the β for each quantile. The approach was developed by Canay (2011), who proposed a two-stage estimator. Stage One involves a means-based panel regression of the dependent variable upon the selected independent variables (Random Effects assumed), i.e.

$$y_{it} = \theta_{\mu} X_{it} + \alpha_i + u_{it}$$

The mean of the observed fixed component for each individual are retrieved from the regression ($\hat{\hat{\alpha}}$), and a transformed dependent variable, $\hat{y}_{it}$ is computed, where $\hat{y}_{it} = y_{it} - \hat{\hat{\alpha}}$. In Stage Two, a quantile regression is performed of $\hat{y}(\tau)_{it} = \beta_0(\tau) + \sum_n \beta_n(\tau) X_{it} + \varepsilon_{it}$

Powell (2015) criticized this approach, and other estimators with a similar approach, as they separate the estimated fixed effect ($\hat{\hat{\alpha}}$) from the disturbance term, u_{it}, when they are in fact, inseparable. The effect, he argues, is to provide estimates related to the variable $(y_{it} - \alpha_i)|X_{it}$ and not the actual variable of interest, y_{it}. There is no guarantee that an individual *i*'s relative position in the distribution $(y_{it} - \alpha_i)$ will be the same as it is in y_{it}. For example, an individual, *j* has a low wage rate (w_{jt}), and may also have a low (unobserved) α_j, while a highly-paid individual, *k*¸ earning w_{kt} *(>* w_{jt}*)* has a high fixed effect α_k *(>* α_j*).* Assuming this fixed effect actually measures self-esteem, motivation etc, the subtraction of the fixed effect will be to at least bring *j* and *k* closer together and may actually reverse their positions in the pecking order. As a consequence, the parameters estimated from the regression do not refer to the 'generic' dependent variable, but to some other version, shorn of fixed effects, pertaining to the included sample. He proposed, instead an estimator that does not require the identification of fixed effects, per se, but uses the observed within-group (individual) variation for identification. His model is encoded in Stata via the *qregpd* routine (described in Powell, 2014) and used in section **3.7.4**.

The Canay estimator is not (currently) supported in Stata, so will be derived from (semi-) first principles and the results compared with Powell's in section **3.7.6.** The significance, if any, of the distorting effects of the fixed effects' identification process should be observable from a direct comparison of the parameter estimates.

Arellano and Bonhomme (2013) proposed a Random Effects version:

$$Q\left(y_{ij} \middle| X_{ij}, \alpha_i, \tau\right) = \alpha_i \gamma(\tau) + X_{ij}^T \beta(\tau)$$

This is supported in Stata and will be estimated and then tested to assess whether or not it is the efficient quantile estimator for this data.

3.3.3. Applications of Quantile Regressions to Earnings.

As awareness of the QR method and advances in the underlying methodology and availability of software have progressed, so has the empirical literature, both within and outwith Economics, for example, Budria and Swedberg (2014). One paper of particular relevance to this study is Fournier and Koske (2012), who estimated conditional- and unconditional quantile regressions using data blended together from different countries' household surveys to assess the effect of qualifications on earnings' equality. They concluded that increasing the proportion of workers with School-Leaving and FE qualifications, allied with an increase in permanency of contracts in the workplace, reduced the earnings' gaps. They were unable to conclude that the possession of a Degree had the same effect, nor was sectoral shift, towards services and away from manufacturing, a significant factor.

3.4. Aims / Modelling Strategy / Included Variables

3.4.1. Aims of the Study

This Study aims to compare and contrast wage rate premia accruing to degrees, FE qualifications, School-Leaving and School-Intermediate level qualifications in the UK, derived from Pooled- and Panel-based estimators, and, in turn to compare the premia at the means, with those experienced at various points across the wage rate distribution. In addition, the *relative* importance of these premia at key quantiles will be estimated and commented upon, with a view to facilitating discussions as to the ongoing role qualifications may have in narrowing the inter-quantiles' and gender gaps in pay rates evident in the UK's labour markets.

3.4.2 Income Inequality in the UK

The extent to which income (and underlying wage rates) are unequal in the UK economy is considered here, and whether or not estimating quantile-based returns to qualifications may elicit additional information as to their relative value. The Gini Coefficient is a long-established metric, used to estimate the overall equality of income distribution (though not without its critics). The minimum possible value of zero would result if all income were shared absolutely equally among everyone, while the maximum value of one would be obtained when all the income was 'earned' by only one person. Values in between indicate the relative inequality in the distribution of income in the economy at a particular point in time. As **Figure 3.11** shows, none of the six 'G7' Economies featured came anywhere near either polar extremes, locating in the 0.35 – 0.45 range, indicating a significant degree of incomes' inequality. [29] The UK tended to have a less equal distribution than all G7 members, except the United States. There was perhaps a slight upward drift towards greater inequality, post the 2008 'Crash', followed by two years of downward shift, before resuming the upwards movement in 2010 – 11.

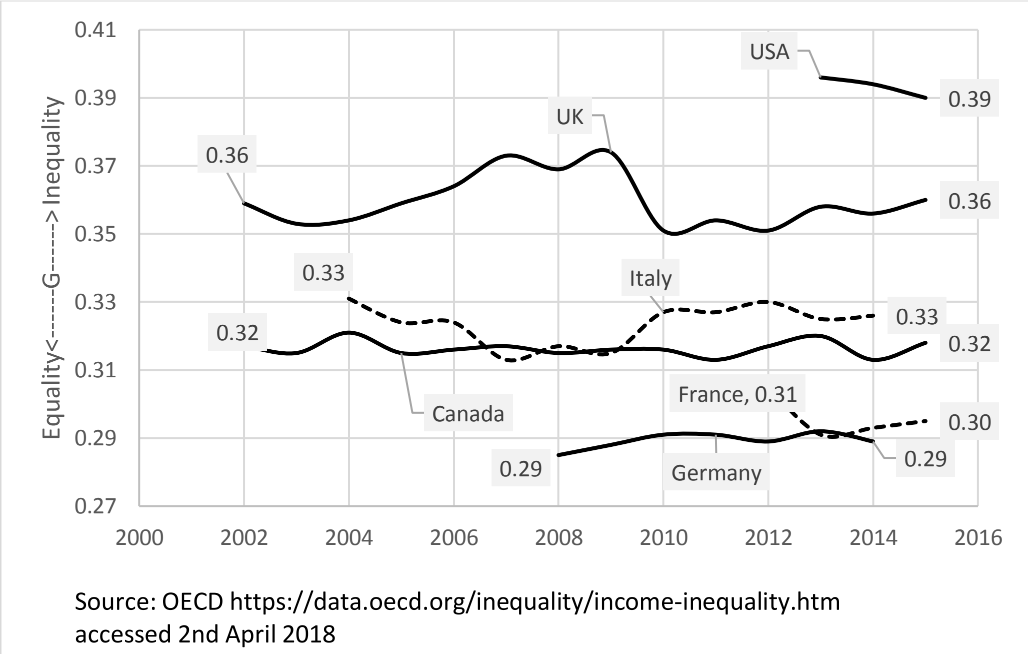


Figure 3.11. GINI Coefficients: G7 Economies (2002 - 2015)

Further evidence of incomes' inequality can be seen from an analysis of income tax data (HMRC, 2016). **Figure 3.12** charts the ratios of the 90th-to-10th and 50th-to-10th percentiles from 1999 to 2014. Those Income Tax payers with gross (nominal) income at the 90th percentile earned five times as much as those in the 10th, while those at the 50th earned twice as much as them. There is maybe a slight downward trend in both ratios over the period, particularly in the 90th-to-10th ratio, but the gap is wide, throughout. [30]

The UK's Office for National Statistics (ONS) uses the Annual Survey of Hours and Earnings (ASHE) [31] to, *inter alia*, estimate labour earnings and their distribution around the mean for taxpayers and non-taxpayers. **Figure 3.13.** plots gross *weekly*

[29] This is based on the OECD's estimate of gross incomes earned in each economy. A coefficient of 0.352 was estimated for Japan in 2012.

[30] These data include only those liable to pay income tax and do not, perforce, include those not earning enough to be taxed.

[31] Formerly called the *New Earnings' Survey*.

earnings at the 10th, 25th 50th, 75th and 90th percentiles from 1997 to 2015. This plot confirms the large and persistent gap between the percentiles and a general widening of the gap between the higher and lower deciles.

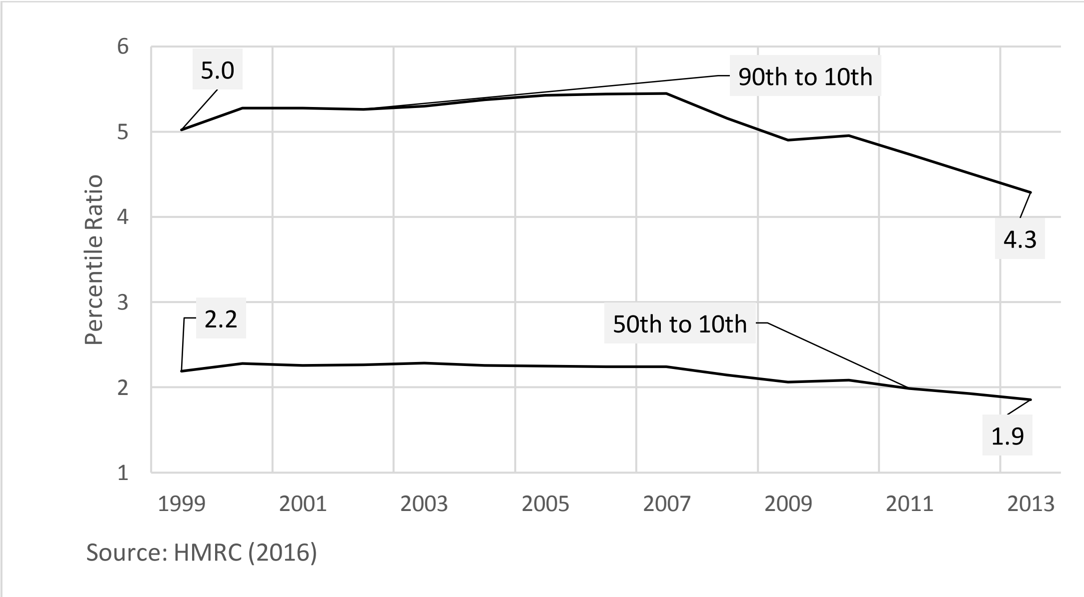
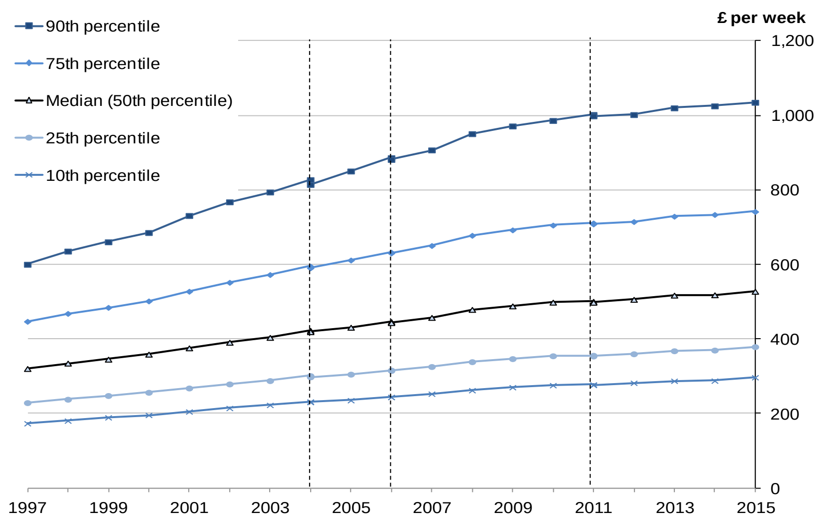


Figure 3.12 UK Income Distribution (1999 - 2014)

Of course, this inequality could be due, entirely, to differences in hours worked, which are reflected in the distribution of the wage rates used in this Study as will be shown later in **3.5.1**.



(Source: ONS, ASHE (2015), http://www., accessed 16th May, 2016)

Figure 3.13. Distribution of full-time gross weekly earnings, UK, (1997-2015)

It is also to be expected that the wage rate will vary according to *what* people do, insofar as occupations differ from each other in terms of their inherent complexity, responsibility and risk of harm to the individual. While there may be intangible benefits to undertaking such work, it is necessary to incentivise the taking up of the 'pressured' jobs via higher rates of remuneration. **Table 3.1** summarises the key quantile wage rates for 2015 by Standard Occupational Classification (SOC2000) 'supercategory'. [32] Unsurprisingly, Professional Occupations head the list at every quantile, earning more at the 10th percentile than those in the 90th percentile of the poorest paid Elementary category. They are followed by Managers and Associate Professionals and in turn, Skilled Trades' workers and Administrators. Those in Caring, Sales and, as already mentioned, Elementary occupations are relatively poorly paid. There is some variation in the location of the means of each distribution, though it is noticeable that they are all well above the median in, at least, the 60th percentile. Neither is the relative distance between the percentiles uniform across all the categories. Clearly, there are those in the second-best paying occupation, Managers and Directors, who are not particularly well paid, reflected in the 90:10 percentile ratio of 5.11. The most egalitarian are the low paid Caring, Sales and Elementary occupations.

Different industrial sectors pay their workers different rates, according to the labour demand and supply conditions that pertain. The ONS uses the Standard Industrial Classification of Economic Activities (SIC) system to encode the reported sectors both for the ASHE and the Understanding Society dataset used in this study. [33]

[32] The ONS uses the SOC2000 version with both the ASHE and Understanding Society data used in this Study.

[33] In particular, the data are encoded using the SIC2007 classification system. Detail on the codes used are in https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/455263/SIC_codes_V2.pdf (accessed 16th June, 2016).

Table 3.1 Hourly pay - Gross (£) - For all employee jobs: United Kingdom, 2015								
Percentiles						Ratio 90:10	Ratio 90:50	Mean is at percentile:
Occupation	**10**	**25**	**50**	**75**	**90**			
All employees	£6.90	£8.34	£11.80	£17.88	£25.64	3.72	2.17	66
Professional occupations	£12.21	£15.28	£19.50	£24.92	£32.20	2.64	1.65	60
Managers, directors & senior officials	£8.93	£12.64	£19.44	£30.66	£45.66	5.11	2.35	66
Associate professional and technical occupations	£9.48	£11.77	£15.04	£19.44	£25.94	2.74	1.72	63
Skilled trades occupations	£7.22	£8.79	£11.38	£14.47	£18.04	2.50	1.59	60
Administrative and secretarial occupations	£7.43	£8.63	£10.30	£13.07	£16.85	2.27	1.64	62
Process, plant and machine operatives	£6.88	£8.00	£9.94	£12.44	£15.81	2.30	1.59	61
Caring, leisure and other service occupations	£6.65	£7.29	£8.49	£10.30	£12.61	1.90	1.49	61
Sales and customer service occupations	£6.50	£6.78	£7.64	£9.55	£12.32	1.90	1.61	64
Elementary occupations	-	£6.59	£7.49	£9.09	£11.50		1.54	63
Source: ASHE, ONS 2015								

As **Figure 3.14** shows, the Financial and Business Services and Construction Sectors paid the most in the UK, though the latter suffered more from the 2008 Recession. The tendency for the UK's Public Sector to pay more than its Private Sector is well understood, but it can be seen here that it occupied a central position in terms of gross pay, with the gap between it and the top-paying sectors growing from 2014 onwards. Unsurprisingly, the Hotel/Restaurant sector is paid the least, characterised by low rates of pay, part-time and otherwise relatively precarious terms of employment. Accordingly, sectoral dummies are required to trap the specific sectoral effects on pay rates (note also, the precarious nature of employment in the Construction Sector).

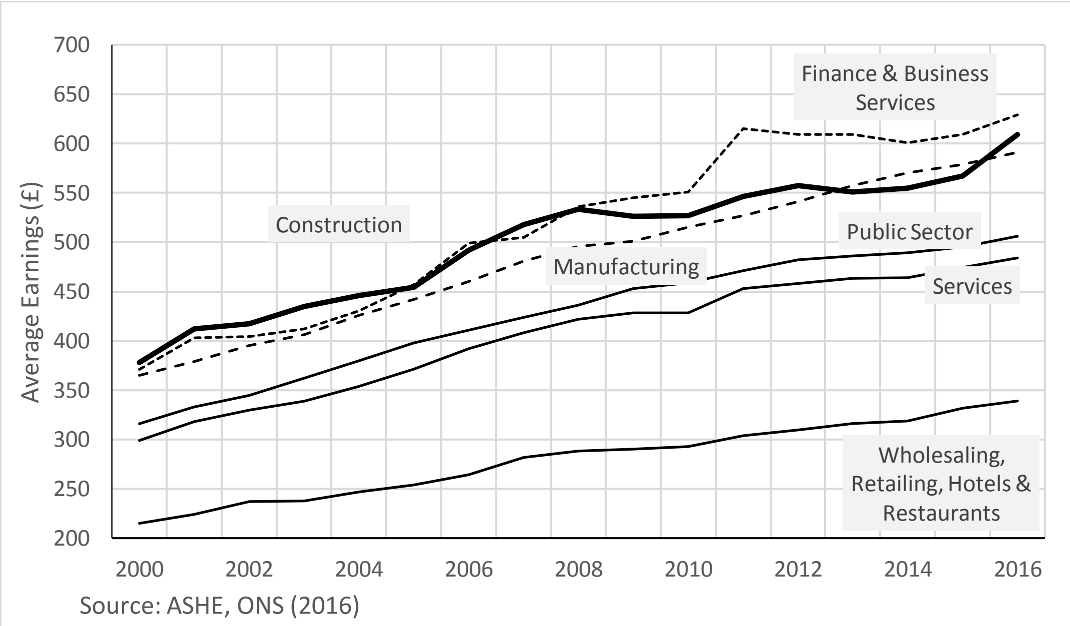


Figure 3.14 UK Average Gross Weekly Earnings by Industrial Sector (2000 - 2016)

Taken together, there appears to be a compelling case for the application of quantile- as opposed to means-based regression of the wage rate, in order to learn more about the relative effect of, amongst other things, qualifications on the wage rate that would simply not be revealed in a means-based regression.

3.5. The Data and variables considered for the Study

3.5.1. The Data

The data used for the Study are the seven available waves (at the time of writing), of the UK 'Understanding Society' (US) Panel Survey, successor to the British Household Panel Survey, covering the period 2009 to 2015. While other surveys, such as the Labour Force Survey and the ASHE have more respondents and variables on personal characteristics, earnings and occupation, US has the virtue of being a Panel. The dataset was restricted to those aged between 18 and 60 years of age. While people enter the workforce at 16, and remain beyond 60, the judgement is that their labour market behaviours are very different to the mainstream and that their inclusion would act to significantly alter the results, distorting the conclusions about the effect of qualifications on the normal-aged worker. **Table 3.2** summarises the number of respondents aged between 18 and 60 by gender in each Wave, and proportions surveyed in each of the three broad age groups, 18-24; 25-45 and 45-60.

Table 3.2
Understanding Society Gender Counts by Wave & Age-group percentages

Wave (Year)	Frequency (% of Wave Total)		Total Count (% of Grand Total)	Age Bands (% of Wave Total)		
	Females	Males		18-24	25-45	46-60
1 (2009)	18,670 (55.7%)	14,873 (44.3%)	33,543 (14.8%)	14.5%	53.0%	32.5%
2 (2010)	20,090 (54.8%)	16,557 (45.2%)	36,647 (16.1%)	14.3%	51.1%	34.6%
3 (2011)	18,664 (54.3%)	15,678 (45.7%)	34,342 (15.1%)	14.4%	49.7%	35.9%
4 (2012)	17,517 (53.9%)	14,963 (46.1%)	32,480 (14.3%)	14.8%	48.8%	36.4%
5 (2013)	16,397 (53.8%)	14,092 (46.2%)	30,489 (13.4%)	15.0%	47.6%	37.4%
6 (2014)	16,704 (53.9%)	14,293 (46.1%)	30,997 (13.6%)	15.3%	47.7%	37.0%
7 (2015)	15,552 (54.2%)	13,126 (45.8%)	28,678 (12.6%)	14.8%	47.1%	38.1%
Total	123,594 (54.4%)	103,582 (45.6%)	227,176 (100.0%)	14.7%	49.5%	35.8%

Despite there being approximately equal numbers of males and females in the UK throughout the Study Period, there were more women surveyed in the seven waves. There was also some variation in the numbers in each age, though the groups are approximately properly represented *pro rata*.

Some variation in the numbers observed at each wave is to be expected, given that participation in the surveys is voluntary. Considerable efforts are made by the survey's statisticians to both make the sample as representative of Society, and to keep as many participants included in successive waves, as possible. It is inevitable, however, that they will require to rebalance and replenish the Panel as the underlying structure of Society changes, and also as sight is lost of existing participants. As a result, the Panel assembled for this Study is *unbalanced.* As **Table 3.3** shows, less than one quarter of all participants were ever-present from the first wave in which they were included (not necessarily #1) and average possible participation was 55% for males and 60% for females.

Table 3.3
Attrition from the US Panel (All Respondents between 18 and 60: Waves 1 - 7)

	Females	Males	Total
Total	22,491	19,022	41,513
Ever-present	4,940	3,886	8,826
Only there in Wave 7	1,042	850	1,892
% of total	26.6%	24.9%	25.8%
Left but came back by end	2,683	2,163	4,846
% of total	11.9%	11.4%	11.7%
Left and did not come back	13,826	12,123	24,770
% of total	61.5%	63.7%	59.7%
Average possible participation / (SD)	60% (31%)	55% (34%)	-

That said, the statisticians were able to track down and re-include about 12% of participants who had previously left the survey. Significant, *non-random* attrition from a Panel can have significant implications for the validity of results and should be tested for, and where judged to be significant, corrected for at the estimation stage. Respondents are asked many questions pertaining to their background, attitudes, etc, but the focus in this Study is on those variables that tend to feature in Mincer-type studies of the wage rate.

3.5.2. The Dependent variable: Hourly Wage Rate

The Survey's respondents are asked to provide information about their current employment status, i.e. whether or not they are currently in paid employment (variable *jbhas*). Those replying 'No' to this question are perforce, excluded from the main wage rate regressions, but are key in the correction for sample selection bias. Those who are in employment are asked to report their hourly wage rate (*basrate*) if that is how they are paid, while some of them also reported their normal gross pay from their main job (*paygl*), the number of hours they normally work in each pay-period (*jbhrs*) and how many weeks this period pertains to (*paygwc*). Although there were 231,382 observations in total in the Panel, only 36,186 of the 161,237 that reported they were in some form of paid employment, were able to report an actual wage rate. In order to boost the number of respondents included in the regressions, an implied basic hourly wage rate ($\widehat{w}_{basic}$) was calculated as:

$$\widehat{w}_{basic} = \frac{paygl}{(jbhrs * paygwc)} \quad [\mathbf{3.3}]$$

Some respondents reported both a *basrate* and all the variables used in **[3.4]**. In general, these resulted in similar hourly rates, and the higher of the two rates was used.

In addition, working respondents were asked to report normal (weekly) hours of overtime, and what rate of pay these hours attracted, if any (w_o). These were combined with the basic pay calculated in **[3.3]** as follows: [34]

[34] As may people work unpaid overtime, the general effect of this, for them is to lower the implied hourly rate. For those fortunate to be paid at premium rates, the effect was to increase their implied rate.

$$w = \frac{(\hat{w}_{basic} * Normal\ Hours) + (w_o * Normal\ OT\ Hours)}{(Normal\ Hours + Normal\ OT\ Hours)} \quad [\mathbf{3.4}]$$

Following convention, any pay and hours involving a *second job* or income from *self-employment* was excluded. However, a dummy indicating employment in a second job was included to capture any effect it may have on the rate for the main job.

Those individuals reporting wage rates below the official National Minimum Wage rate in each year were excluded from the analysis. This yielded a potential 115,668 observations for the regressions, though only 107,408 of them appeared to have been paid at least the National Minimum Wage Rate, given the year and their age. As **Figure 3.15** indicates, there is a wide and skewed distribution of wage rates in the UK Labour Market, echoing those reported from the ASHE (**Figure 3.13**).

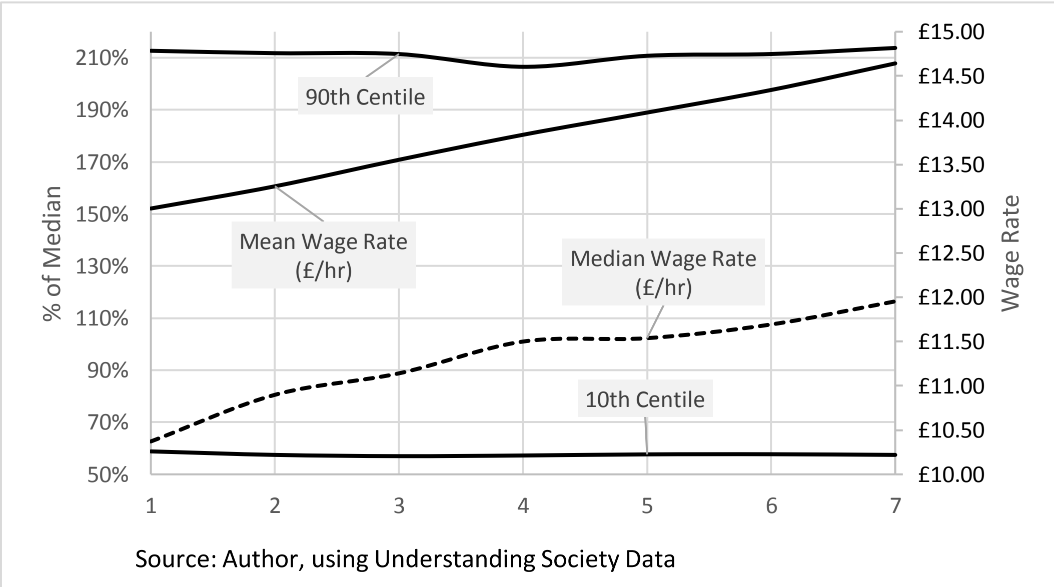


Figure 3.15 10th & 90th Centile Wage Rates as % of Median: UK

Table 3.4 summarises the centile and mean wage rates for the first and last years of the survey data. Amongst other things, it is clear that there is significant positive skew in both distributions, with the 10^{th}-40^{th} centiles located close to each other, while the median and the 60^{th} centiles are below the mean. There is also some change in the shape of the distributions, but both are clearly skewed to below the mean to a significant degree.

Table 3.4 Centile & Average Wage Rates: UK 2009 and 2015 (excluding overtime)				
Centile	**Wave 1 (2009)**		**Wave 7 (2015)**	
	Wage Rate	**% of Year's Median**	**Wage Rate**	**% of Year's Median**
10th	£6.06	64.0%	£7.20	65.5%
20th	£6.74	71.2%	£7.88	71.6%
30th	£7.50	79.2%	£8.67	78.8%
40th	£8.37	88.4%	£9.73	88.5%
Median	£9.47	100.0%	£11.00	100.0%
60th	£10.81	114.1%	£12.50	113.6%
Mean	£11.53	121.8%	£13.32	121.1%
70th	£12.52	132.2%	£14.38	130.7%
80th	£14.93	157.7%	£16.97	154.3%
90th	£18.88	199.4%	£21.91	199.2%
Source: Author, using Understanding Society Data				

3.5.3. The candidate Independent Variables

A restricted number of, mostly dummy *independent variables* pertaining to work and pay outcomes were extracted from the Survey. The Study does not introduce radically different variables or functional forms given that the focus is on *comparisons* of wage rate premia. That said, it is worthwhile to consider those that are included and to consider the appropriateness of their inclusion in the models. The independent variables considered in the first instance were:

- Highest Educational Qualification possessed
- Public Sector employment
- Whether the respondent had more than one job (i.e. a *main* job and one, or more *other jobs*)
- Age at time of observation
- Region of the UK in which employed
- Industrial Sector of employment (SIC2007) dummy variables
- Occupation engaged in (SOC2000)

The ones of most interest to this study, are those pertaining to the *respondents' highest qualification*. In this Study, four specific levels were included: [35]

- Degree-level qualifications, ranging from Bachelor's to Master's to Doctorates, regardless of the mode, duration on Institution awarding the Degree
- College-level: one- and two-year Certificates and Diplomas
- School Leaving: Advanced ('A')-levels, Scottish Highers and Advanced Highers, National-5
- School Intermediate: Ordinary ('O')-levels, Scottish Standard Grades, National-4.

Table 3.5 summarises the percentages of females and males possessing a particular highest level of qualification in Waves One and Seven of the Survey. The persistent effect of government policy whereby 50% of School Leavers are expected to engage in some form of post-School education can be seen, with the percentage reporting a degree qualification increasing from 26-28% in wave 1 to 28 – 30% in wave 7. Also noticeable is the reversal of the gender gap, formerly in favour of males, in favour of females by the end of the Study period, perhaps as a result of superior performance at School. This concords with findings from research into school- and post School performance in England (DfES (2007) and Hillman and Robinson (2016). The DfES work found that in the 2006-7 session, there was a large gap in school performance in favour of females in English, Arts & Humanities and a smaller gap (2%) in Mathematics. They noted that in 1991, the gap had generally been in favour of males, when they had a 4% gap in mathematics over female students.

[35] US contains another category, 'Other Qualifications' that gather together a miscellany of formal qualifications (e.g. City and Guilds, RSA, RYA, etc) that are designed to allow their holders to perform highly-specific work roles (such as speed typing, language interpreting, ship piloting). While these should, in theory, affect earnings, their wide-ranging nature imply that any conclusions to be drawn about their labour market performance would be difficult to generalise across their gamut.

Table 3.5 Highest Qualification by Gender (2009 and 2015)							
		Degree	**FE**	**School Higher**	**School Intermediate**	**Other**	**None**
Wave 1 (2009)	Female	25.5%	13.6%	19.9%	23.5%	7.6%	9.9%
	Male	28.2%	11.1%	22.9%	20.0%	9.3%	8.6%
Wave 7 (2015)	Female	30.5%	13.9%	22.5%	19.1%	8.6%	5.4%
	Male	28.6%	10.7%	25.6%	19.6%	10.2%	5.2%
Source: Author, using Understanding Society Data							

Hillman and Robinson noted from their analysis of UK University data, that there were more female entrants to University than males, and also that males had a significantly higher dropout rate. They opined that a major reason for this is that females perceived that the relative returns to them from a degree were higher than those perceived by males for themselves. This assertion will be examined in section **3.7.7** when the relative importance of the qualifications to the wage rate will be considered.

The numbers reporting FE qualifications as their highest level remained at around 11% for males, 14% for females. This is surprising, given the expansion of that sector alongside HE, but may be explained by the increased numbers going on from HNC/D programme, to degree programmes. In Scotland, for example, the numbers known to have done so there rose from 5,928 in 2008/9 to 8,251 in 2013/14.[36] The percentages reporting School leaving qualifications as their highest level, edged up from 20-23% to 23-26%, with a gap in favour of males, maintained. Given that the percentages reporting the lower, Intermediate School level as their highest attained fell by 1% for males and 4% for females, we may have witnessed upwards pressure in the School Sector to take more students to their higher level in an effort to

[36] Scottish Funding Council, Report of the Access and Inclusion Committee, 2016, http://www.sfc.ac.uk/web/FILES/CMP_AccessandInclusionCommittee24May2016_24052016/AIC16_13_Annex_G_Chapter_5_Articulation_and_progression.pdf, accessed 26th March, 2018.

improve their labour own market prospects, and their School's League Table performance. Unsurprisingly the percentage reporting no formal qualifications fell from 8-10% to 5%, reflecting the trend towards credentialism in both the education sector and the workplace.

Table 3.6 summarises the mean wage rates for wave 7 (2015) for each qualification level, by gender. This shows both an expected 'pecking order' by qualification, with degrees obtaining the highest rate, and the unqualified achieving the lowest, though note the relatively large standard errors around each mean (all other waves exhibited the same pecking order). Ostensibly, the table indicates a significant gender pay-rate gap in favour of males at each qualification level, though the large standard errors would not necessarily lead to the conclusion that the differences observed were statistically significant.

Table 3.6
Mean Wage Rates (including overtime) by Highest Qualification and Gender: (2015)
(Standard Errors in brackets)

	Highest Qualification					
	Degree	**FE**	**School Higher**	**School Intermediate**	**Other**	**None**
Female	£14.97 (£7.86)	£11.84 (£5.18)	£9.98 (£5.48)	£9.84 (£6.81)	£10.83 (£4.70)	£8.26 (£1.78)
Male	£18.04 (£10.51)	£14.19 (£6.88)	£12.55 (£7.11)	£11.76 (£6.2)	£13.40 (£7.87)	£10.59 (£4.58)

Source: Author, using Understanding Society Data

Recalling that this Study is concerning itself with returns at the quantiles, it is worthwhile noting, as **Figure 3.16** illustrates, that all qualification levels were associated with all centiles of the wage rate distribution.

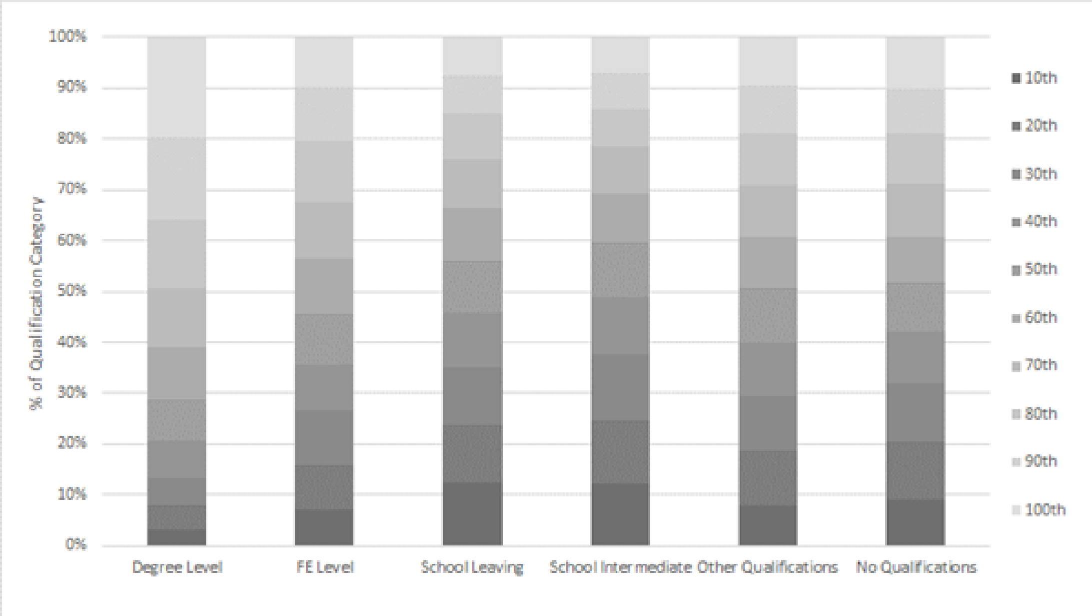


Figure 3.16 Wage-rate-centile composition of each qualification level: Understanding Society Wave 4 (2012)

Even those educated to Degree level are to be found in the bottom Decile, and it was also possible to find people with no (formal) qualifications in the upper echelons of the wage rate distribution.

Looking at the other independent variables included in this Study, the existence of a Public sector and Trades-Union pay premia is usually confirmed in studies on earnings (recall **Figure 3.14**) resulting, perhaps, from the greater prevalence of formal pay scales and associated matching of job profiles to these, in combination with a relatively greater Trades Unions' presence. As **Table 3.7** shows, there is tentative evidence of such a premium in the US survey data.

Both the Public Sector- and Trades Union premia appear to be more significant for females than for males, with a suggestion that both have helped to counteract the negative pay consequences of the 2008 Recession, and its aftermath. Another effect is to narrow the gender pay-rate gap, though this remains in favour of males throughout.

Table 3.7
Public v Private Sector and Trade Union v Non-Trade Union members' Mean Wage Rates by Gender Understanding Society 2009 and 2015

	Public		**Private**		**Public:Private Ratio**		**Female:Male Ratio**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Public**	**Private**
Wave 1 (2009)	£12.35 (£5.69)	£14.17 (£6.72)	£11.28 (£7.45)	£13.86 (£8.47)	1.09	1.02	0.87	0.81
Wave 7 (2015)	£13.28 (£6.26)	£15.11 (£7.14)	£11.54 (£7.32)	£14.67 (£9.33)	1.15	1.03	0.88	0.79
	Trade Union		Not Trade Union		TU:Not TU Ratio		Female:Male Ratio	
	Female	Male	Female	Male	Female	Male	Trade Union	Not Trade Union
Wave 1 (2009)	£12.69 (£5.55)	£13.89 (£7.21)	£11.44 (£5.41)	£13.78 (£6.87)	1.11	1.01	0.91	0.83
Wave 7 (2015)	£13.70 (£6.61)	£15.04 (£6.84)	£12.08 (£6.16)	£14.85 (£7.67)	1.13	1.01	0.91	0.81

Source: Author, using Understanding Society Data

While Trades Union membership appear to be a key factor in determining pay rates, the relatively low number of responses to the Survey's question on Trades Union membership led to the decision to include only a Public Sector dummy in the regressions (discussed more in the next section).

Another commonly-used independent variable in income studies is whether or not a worker has more than one job, normally thought of as having a 'main' job, plus one or more 'other' jobs. It might be assumed that the presence of the 'other' job(s) is due to the 'main' employer not offering them sufficient hours at the observed 'main' job rate, resulting in an insufficient total (net) income. This would not be expected to affect the 'main' job's standard wage rate but may have some bearing on their willingness and ability to accept overtime perhaps paid at a higher rate, for example. As **Table 3.8** indicates, there appeared to be some small variation in the mean wage rates for both males and females at both ends of the Study Period, with an apparent reversal in the effect between the two waves.

Table 3.8

Average Wage Rates for Males and Females with and without a second job

(NB: SDs in brackets and second job wage rate not included in calculation)

	Male		**Female**	
	Has Second Job	**No Second Job**	**Has Second Job**	**No Second Job**
Wave 1 (2009)	£13.02 (£8.26)	£12.86 (£7.87)	£10.63 (£6.64)	£10.47 (£6.22)
Wave 7 (2015)	£13.27 (£7.19)	£14.77 (£8.96)	£12.02 (£6.78)	£12.21 (£6.93)

Source: Author, using Understanding Society Data

Of course, the large standard errors associated with the sub-group's means would lead to the conclusion of no significant effect, but a dummy variable indicating the possession of an 'other' job will be included in the regressions.

Next, we consider the relationship between wage rate and age. **Figure 3.17** charts the mean wage rates by age, by gender for waves 1 (2009) and 7 (2015), with approximating parabolic trend lines fitted. In general, average wages rates for both female and male employees rise as they age through their twenties and thirties, as they are promoted, and climb pay scales. This process continues until they get into their late forties, then the rate falls, perhaps indicating that they are less inclined to continue working antisocial hours or are looking to take on less-stressful/responsible work.

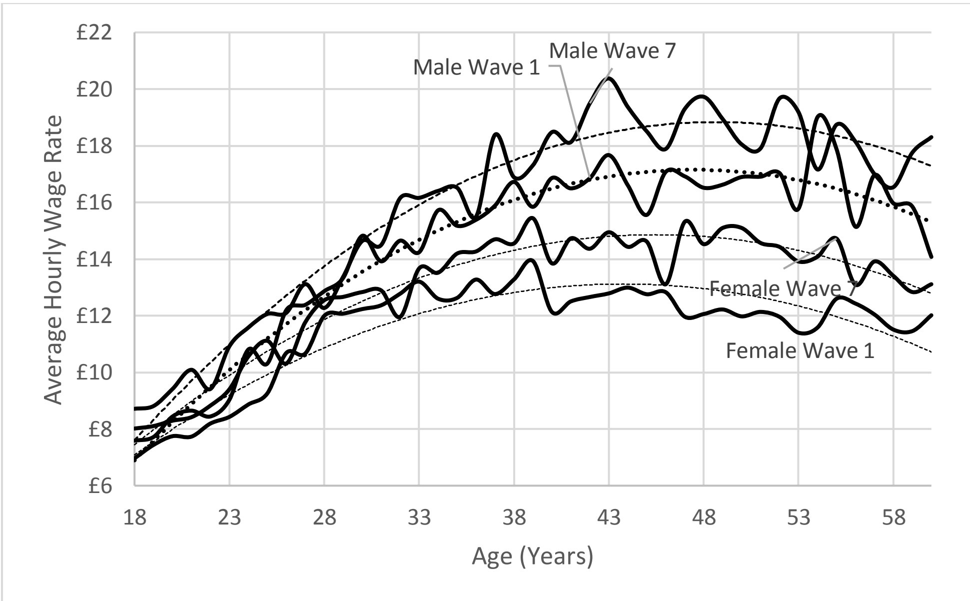


Figure 3.17 Wage Rates by Age by Gender: Understanding Society Wave 1 (2009) and Wave 7 (2015)

Figure 3.18 also shows how the gender gap appears to widen with age (linear trend line fitted for Wave 7). This reflects, perhaps, women's different career trajectories to men, with time out of the labour market to have children, and more periods of part-time work as their children reach school age.

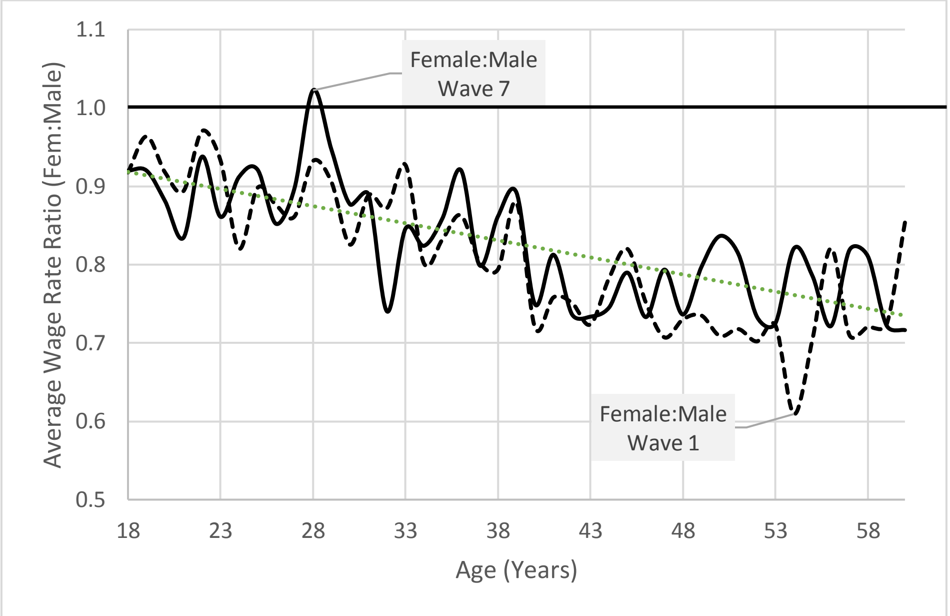


Figure 3.18 Gender Pay Gap and Age: Understanding Society Wave 1 (2009) and Wave 7 (2015)

Although age is clearly a plausible factor in determining wage rates, employers reward Labour Market ('LM') Experience. This is closely allied to age, of course, but there are reasons why people of a certain age may not have the same amount of experience as others of the same age (maternity leave, post-compulsory education, periods of imprisonment, for example). Employers reward an individual's experience in the labour market, perhaps indicating higher positions within the organisation, with more responsibility, and perhaps, more subtly, reflecting the tendency for the individual to have not been 'scarred' by periods of detachment from the same labour market. In order that the wage-rate regressions can identify and experience-related premium, an experience variable was estimated from: (i) the individual's age at the time of the survey minus, (ii) the time they are deemed to have left full-time education. This second part required, in many cases, assumptions to be made about when this actually happened. The US survey includes two variables, *scend* and *feend* that allows individuals to report the age they left school or post-compulsory education (i.e. college of university), respectively. However, of the ~231,000 observations in the seven-wave data file, only ~116,000 reported their school leaving age, and 63,904 the age they completed their post-school education. [37] Accordingly, in order to boost the numbers with implied LM experience, the highest qualifications' dummies were utilised to infer ages at which full-time education ended, viz, Degree (21); Further Education (19); School Leaving (18), School Intermediate and 'Other' (17). Otherwise, it was assumed that individuals left education at 16 and available for work. Given the non-linearity of the relationship between wage rate and age depicted in **Figure 3.17**, the square of LM experience will also be included to allow for any non-linear relationship between it and wage rate to be captured: typically, there is a downturn in the wage rate as workers move into their forties (~20^+ years in the LM).

In common with all labour markets throughout the developed World, the UK has within it, a number of highly differentiated, regional labour markets. These are characterised by different industrial, occupational and demographic profiles, with a consequence for wage rates. **Table 3.9** summarises the median weekly earnings in each of the major UK regions in April 2015.

<table>
<tr><th colspan="4">Table 3.9
Median full-time gross weekly earnings by UK region, April 2015</th></tr>
<tr><th>Region</th><th>£ per week</th><th>Region</th><th>£ per week</th></tr>
<tr><td>1. London</td><td>659.90</td><td>7. North West</td><td>488.80</td></tr>
<tr><td>2. South East</td><td>552.10</td><td>8. North East</td><td>488.70</td></tr>
<tr><td>UK Average</td><td>527.70</td><td>9. Yorkshire and the Humber</td><td>486.40</td></tr>
<tr><td>3. Scotland</td><td>527.00</td><td>10. Northern Ireland</td><td>484.90</td></tr>
<tr><td>4. East</td><td>517.50</td><td>11. East Midlands</td><td>479.10</td></tr>
<tr><td>5. West Midlands</td><td>493.10</td><td rowspan="2">12. Wales</td><td rowspan="2">473.40</td></tr>
<tr><td>6. South West</td><td>492.80</td></tr>
<tr><td colspan="4">Source: ONS (ASHE), 2016</td></tr>
</table>

As might be expected, the highest-paying region is the UK's Capital, followed by the surrounding South East of England area. Unusually, perhaps, there is no smooth

[37] Of course, many of these reported both, so it is not just a matter of adding the two number together to get the number of individuals that reported either.

radiating pattern beyond that, with the northernmost region, Scotland coming third (and at the mean), outperforming regions like the West Midlands, in which Birmingham lies. The lowest-paying Region, Wales, lagged London by over £180 per week on average, or almost 40%. **Table 3.10** confirms that these inter-regional differences and pecking order are reflected in the US survey wage rates for waves 1 and 7, for example. There also seems to be some difference in the gender pay rate gap between the regions, with the low-paying Wales and Northern Ireland regions being relatively egalitarian. compared with England and Scotland.

Table 3.10
Average Regional Wage Rates by Gender: 2009 and 2015
(Standard Errors in brackets)

	Region	**London & the South East**	**Scotland**	**Northern Ireland**	**Rest of England**	**Wales**
Wave 1 (2009)	Female	£11.85 (£7.20)	£10.32 (£5.03)	£11.01 (£12.07)	£9.90 (£5.10)	£9.68 (£5.17)
	Males	£14.43 (£9.75)	£12.53 (£6.80)	£11.58 (£7.37)	£12.39 (£7.07)	£10.67 (£4.54)
	Female:Male Ratio	0.82	0.82	0.95	0.80	0.91
Wave 7 (2015)	Female	£13.78 (£7.86)	£12.33 (£7.18)	£12.00 (£8.81)	£11.55 (£6.21)	£11.19 (£4.61)
	Males	£16.69 (£10.75)	£15.26 (£9.33)	£12.69 (£5.96)	£14.00 (£8.00)	£12.97 (£7.04)
	Female:Male Ratio	0.83	0.81	0.95	0.83	0.86

Source: Author, using Understanding Society Data

Table 3.11 summarises the occupational wage rates from the US data, for Waves One and Seven. The relative average wage rate positions of each occupation are broadly those indicated in **Table 3.1**, though some of the middle- and lower-paying occupations swap positions with each other. Although occupation does appear to be a significant factor in determining wage rates, it will not be included in the regression-based models, due to the inherent endogeneity between it and the education outcome: this will be discussed more in the econometric methodology section.

Table 3.11 **Occupational Wage Rates: US Waves 1 and 7** (Standard Errors in brackets)					
Occupation	**Wave 1 (2009)**	**Wave 7 (2015)**	**Occupation**	**Wave 1 (2009)**	**Wave 7 (2015)**
1. Professional	£16.05 (£7.86)	£17.43 (£8.68)	6. Admin	£9.64 (£4.85)	£11.08 (£5.31)
2. Management	£15.64 (£9.72)	£18.20 (£10.74)	7. Personal	£8.30 (£6.16)	£9.36 (£3.64)
3. Associate Professional	£12.91 (£6.52)	£14.56 (£7.29)	8. Elementary	£7.77 (£3.66)	£8.98 (£3.69)
4. Skilled Trades	£10.90 (£5.59)	£12.36 (£5.69)	9. Sales	£7.47 (£2.72)	£8.85 (£5.87)
5. Process, Plant & Machine Operatives	£9.74 (£3.87)	£11.38 (£5.23)			
Source: Author, using Understanding Society Data					

Table 3.12 overleaf, summarises the mean industry-sector wage rates from the US Data. The relative positions of the sectors accord broadly with those indicated in **Figure 3.14**. Sectors, such as Finance and Insurance, and Information and Communications, pay on average, much more than the poorly-paid accommodation/food and wholesale/retail sectors, while the Public Sector is positioned somewhere in the middle. As with occupation, the sector in which an individual works appears to be a determinant of pay (rate) and appropriate dummies will be included in the regressions.

Public Sector / Trades' Union Membership. The empirical literature usually discerns a significant wage premium associated with union membership, dependent on setting: public-private sectors; size, 'old'/'new' organisations, country of ultimate ownership, level of union membership in a particular workplace, sector or occupation, etc (for example, Bryson, 2002). Respondents in employment were asked to indicate whether or not (i) there was a recognised Union, and (ii) if so, whether or not they were a member of it (variable *tuin1*). A complication is that this question was only asked in Waves two, four and six and not in one, three, five or seven. Accordingly, Excel was used to 'plug' the gaps in the data, assuming that the

Table 3.12
Sectoral Wage Rates: US Waves 1 and 7
(Standard Errors in brackets)

Sector	Wave 1 (2009)	Wave 7 (2015)	Sector	Wave 1 (2009)	Wave 7 (2015)
1. Mining & Quarrying	£17.60 (£12.70)	£21.29 (£11.24)	11. Water, Sewerage & Waste	£11.58 (£8.47)	£14.48 (£9.38)
2. Information & Communications	£15.86 (£9.00)	£19.13 (£10.91)	12. Real Estate	£11.24 (£5.27)	£13.46 (£7.98)
3. Finance & Insurance	£15.51 (£11.35)	£18.28 (£12.65)	13. Health & Social Work	£11.18 (£6.88)	£12.62 (£6.39)
4. Scientific & Technological	£15.17 (£9.23)	£16.69 (£11.01)	14. Admini & Support	£9.65 (£5.13)	£11.46 (£7.15)
5. Power generation & Supply	£12.85 (£6.28)	£16.41 (£7.74)	15. Other Services	£9.48 (£4.60)	£12.24 (£7.61)
6. Public Administration & Defence	£12.83 (£5.78)	£14.37 (£6.37)	16. Arts & Entertainment	£9.42 (£5.03)	£10.83 (£6.41)
7. Construction	£12.76 (£6.85)	£14.29 (£7.28)	17. Wholesale & Retail	£8.54 (£4.45)	£10.35 (£6.54)
8. Education	£12.38 (£7.10)	£13.28 (£6.52)	18. Agriculture, Forestry & Fishing	£8.52 (£2.97)	£12.98 (£6.91)
9. Manufacturing	£11.96 (£6.95)	£14.27 (£7.84)	19. Accomm & Food Services	£7.53 (£5.59)	£8.74 (£3.67)
10. Transport & Storage	£11.66 (£7.03)	£12.97 (£6.85)			

Source: Author, using Understanding Society Data

first observation applied to any gaps before then, as well as any gaps after that observation until the next one. In the event, however, only 44,386 responses were generated for the variable (24,222 member; 20,614 non-member). Given that 167,175 of the responses indicated people that were in paid work, this implies a TU membership of 14.4%, rather less than the 21.7% – 24.5% reported by UK

Government Statistics for the Study Period. [38] This under-reporting may be due to the practise of only asking the question once every two waves and/or a sensitivity in answering questions of this type by some respondents and may introduce an element of bias into the responses, as well as reducing, considerably, the number of observations included in the regressions.

Other household income. One other variable was considered for this Study, namely the effect on the individual's wage rate of the other income coming into the household to which they belong: more specifically, the percentage of their household's total contributed by the individual. The contention here is that someone living in a high-earning household may be (counterfactually) more content to accept a lower wage rate (presumably with less-demanding/responsible work) than someone who appears to be supporting themselves and others. The US Survey includes questions on household, as opposed to individual factors including total household income (see Levy and Jenkins, 2012 for detailed discussion on inclusion in the survey and computation of the gross and net income totals). [39] The Gross Household monthly incomes were retrieved and fed into the Panel using each individual's *household* identifier as the key matching field. The individual's implied contribution to this total was computed as follows:

$$Individual's\ Monthly\ Income = w * (Normal\ Weekly\ Hours + Normal\ Weekly\ OT\ Hours) * 4.333$$

where *w* is the computed wage rate, using **[3.4]** and 4.333 the conversion factor converting a weekly wage to a monthly one.

The percentage contribution of the individual was easily computed from by dividing their total by the household's. In the event however, the household survey files

[38] Trade Union Statistics, 2017, UK Government Department of Business, Energy & Industrial Strategy, https://www.gov.uk/government/statistics/trade-union-statistics-2017, accessed 13th August 2018.

[39] While this pertains to the Survey's successor, the British Household Panel Survey, the same variables and rubric were retained in US.

only returned 86,649 observations of household income which, when combined with the numbers able to provide a wage rate and implied monthly income, fell to only 29,083 (12,958 males and 16,125 females). The Pearson Correlation Coefficients between wage rate and percentage household income were statistically significant ($p<0.05$ for both genders, but otherwise weak-positive (r_{male} = 0.336 ; r_{female} = 0.298). To include the percentage income variable in the regression would therefore result in a large loss of observations (as per the Trade Union variable), so it was (regretfully) decided not to include this variable in the regressions.

The next section will outline the econometric methodology applied to the returns' estimation process, focussing on the functional forms and the sample selection issues.

3.6. Econometric Issues and Models

3.6.1. Introduction

The wage rate premia associated with the different levels of qualifications are estimated by means of a Mincer model of the wage rate in the style of **[3.1].** In keeping with the overarching theme of this Thesis, separate cross-sectional and Panel models of the mean and quantile wage rates are estimated and compared. The dependent variable is the log of the wage rate, and the independent variables are as described in the previous subsection. It is clear that the UK labour market experience of females continues to differ significantly from that of males. One way to deal with this would be to include a gender dummy, interacted in some fashion, with the other independent variables. The judgement here is that rather than do this, separate models for males and females should be estimated, given the clear differences in the wage rates and the potential for very different effects on the wage rate from the presence of dependent children in the household, for instance.

Two sources of potential bias must be considered prior to any estimation involving this kind of Panel Data, namely

(i) *Attrition Bias* and

(ii) *Sample Selection Bias*.

3.6.2. Attrition Bias

Looking first at *Attrition Bias*, it is apparent from the analysis of participation in **3.5.1** that there was considerable attrition from the Panel. In a report authored after the second wave of data had been collected (ISER, 2012), concern was expressed about the high level of attrition. It appeared that attrition was status-dependent, with ethnic minorities and young and elderly age groups exhibiting a greater propensity to leave the Survey (some regional differences were also noted). While the Statisticians assure the Survey's users that they make strenuous efforts to compensate, via like-for-like replenishment and computation of significance weights, it was decided to compute and utilise a set of attrition weights, following the methodology of Becketti, et.al (1988). This involves limiting the Panel Data to each individual's *first* appearance in the survey, and to regress the same dependent variable against the same independent variables, but with the addition of an *attrition dummy* and full interactions between it, and the independent variables. In order to asses this in this Panel, an *attrition* dummy was computed. If a respondent remained in the Panel in every wave following their first appearance (which could be after wave 1, of course), then they were given a dummy value '0', otherwise they were deemed to have been an attritioner and given a score of '1', even if they subsequently returned to the Panel. An F-test was performed to jointly test whether or not the Attrition Dummy and its interactions were jointly significant. If they were, then it should be assumed that attrition was not random and that weights for each individual should be computed and, where permitted, used in the estimation of the parameters. [40]

In the event, the regressions for both males and females indicated significant attrition (Males: F= 1.45 (p=0.0197); Females: F = 3.58 (p=0.0000)) and that it was therefore necessary to compute and use attrition weights. The procedure does this

[40] This has been named the 'BGLW' test, after the test's creators.

by estimating an *'unrestricted'* observation probability (p_u) from a probit of the attrition dummy against the independent variables plus Move- and Occupation-Risk dummies, accompanied by a *'restricted'* observation probability (p_r) from a similar probit but without the Move- and Occupation-Risk dummies.

The Move-risk dummy should indicate the (subjective) likelihood that the individual will not return the following year to participate in the Survey, as sight of them may be lost as they move and fail to keep on board with the Survey in the following waves. It was decided to construct the move-risk dummy using a five-point scale, based on three questions asked in the survey:

(i) whether or not they have ever moved in the (recent) past,

(ii) whether or not they would like to move in the next 12 months and,

(iii) whether it was likely that they would move in the next 12 months.

Affirmative answers to (i) scored one point, while an affirmative answer to (ii) or (iii) attracted two points apiece: the judgement being that moving in the recent past indicates some ability to do it again, but that an expressed wish to do so, or an awareness that this is likely, are stronger indications that they will move and may be lost to the Survey.

Respondents were asked whether or not they intended changing their job within the next 12 months (i.e. before they next scheduled interview). Where they indicated that they intended doing so, and that this may have involved a physical change of location, it was taken as a potential risk indicator of future non-participation in the Survey. Accordingly, the *occupation risk* binary dummy was included in the probit regressions on the attrition variable. Each individual's attrition weight is calculated as $\left(\frac{p_u}{p_r}\right)$ and used in the regressions that permit it. The average weights for males and females were quite similar (1.037 v 1.039), although the males exhibited a greater range: 0.747 versus 0.643 for the females. While the F-tests indicated significant attrition for males and females, those models that allow weights were also estimated without them, to assess the actual effect on the parameter estimates of failing to correct for it, and to allow a more direct comparison against models that STATA does not (currently) permit to be estimated with such weights.

3.6.3 Sample Selection Bias

It is generally assumed that the decision to work is not random, as individuals weigh up the work opportunities available to them, focussing, *inter alia*, on the wage rate associated with these offers. Where the best of these fails to meet the minimum wage rate acceptable to them (their *reservation wage rate*) they remain out of the labour market until they either lower the reservation rate, or a sufficiently enhanced rate is offered to them. This failure to include inactive workers has the potential to significantly bias the parameter estimates. The long-established method to correct for this is due to Heckman (1979) via a two-step process. In the first step, a probit regression of the Work-decision Dummy against a set of variables thought to influence the work decision is performed (as per **[1.7]**). From this is computed the *Inverse Mills Ratio,* λ which is then used in the wage regression as an additional independent variable. In theory, the missing participants are proxied, and the resulting parameters rendered un/less biased. The selection of the variables for the Probit is, of course subjective (a source of criticism of the process), but this Study assumed the following version of **[1.7]** was appropriate:

$$p(w_i) > 0 = \beta_0 + \sum_{j=1}^{5} \beta_j Qual_i$$
$$+ \beta_6 Age_i + \beta_7 Age_i^2 + \sum_{k=8}^{13} \beta_k Region_i$$
$$+ \quad \beta_{14} Partnered\ Status_i + \beta_{15} Housing\ Tenure_i$$
$$+ \beta_{16} Childcare + \beta_{17} Elderly\ care_i + \sum_{p=18}^{24} \beta_p Wave_i + v_i \quad [\mathbf{3.5}]$$

The variables included were selected as being those thought likely to influence an individual's decision to work, but not directly influenced by the wage rate on offer. That is to say, they were status variables, some out of the individual's control (such as age), others a consequence of other non-work-related decisions, such as region and partnered status. While some of these variables will also appear in the (main) wage-rate regressions, there are sufficient numbers of variables excluded from it (for example, *Housing Tenure*) to permit identification.

Separate versions of **[3.5]** were estimated for males and females, allowing different effects of, for example, the responsibility for children to inform the work participation decision. It may be that the effect on males is to *increase* the likelihood of participation, while for females, it may *reduce* the likelihood.

The inverse of the Mills Ratio (λ) is estimated for each individual and used to modify the wage rate regressions. The significance of the Mills Ratios will be reported in each regression involving the correction.

3.6.4. The Econometric Models

Regressions with, and without these corrections are compared in **3.7.2** to assess their significance in relation to this data. The models that will be estimated in this study are grouped as follows:

1. **Pooled models: Means and Quantiles**

$$
\begin{aligned}
ln(w_i|\lambda) = \beta_0 + \sum_{j=1}^{5} \beta_j Qual_i \\
&+ \beta_6 Experience_i + \beta_7 Experience_i^2 + \beta_8 Partnered\ Status_i \\
&+ \beta_9 Public_i + \beta_{10} Has\ 2nd\ Job_i + \sum_{l=11}^{30} \beta_l Sector_i \\
&+ \sum_{m=31}^{36} \beta_m Region_i + \sum_{p=37}^{43} \beta_p Wave_i + v_i
\end{aligned}
$$

$$ln(w_i|\lambda) = \beta_0 + \sum_{j=1}^{5} \beta_j(\tau)Qual_i$$
$$+ \beta_6(\tau)Experience_i + \beta_7(\tau)Experience_i^2$$
$$+ \beta_8(\tau)Partnered\ Status_i + \beta_9(\tau)Public_i$$
$$+ \beta_{10}(\tau)Has\ 2nd\ Job_i + \sum_{l=11}^{30} \beta_l(\tau)Sector_i + \sum_{m=31}^{36} \beta_m(\tau)Region_i$$
$$+ \sum_{p=37}^{43} \beta_p(\tau)Wave_i + v_i$$

2. Panels-based models: Means and Quantiles

$$ln(w_{it}|\lambda) = (\alpha_i + \beta_0) + \sum_{j=1}^{5} \beta_j Qual_{it}$$
$$+ \beta_6 Experience_{it} + \beta_7 Experience_{it}^2 + \beta_8 Partnered\ Status_{it}$$
$$+ \beta_9 Public_{it} + \beta_{10} Has\ 2nd\ Job_{it} + \sum_{l=11}^{30} \beta_l Sector_{it}$$
$$+ \sum_{m=31}^{36} \beta_m Region_{it} + \sum_{p=37}^{43} \beta_p Wave_{it} + v_{it}$$

As a consequence of specifying these variables, the number of useable observations fell from the 121,290 that had a reported, or imputed wage rate to 95,364 with a full set of observations, 43,066 of whom were male and 52,298, female. Models will be estimated using Stata 15: at the time of writing, this supports a standard quantile regression routine (*qreg*) using pooled data and a user-written routine that embodies the Powell Fixed-effects estimator (*qrpd*) based on Powell (2014). In addition, Canay's Fixed-effects estimator will be employed, though requiring some user-written Stata code to mimic his method.

The pooled quantile results will be compared and contrasted with both (Heckman)-corrected and uncorrected regressions of the mean wage rate in order to set the scene for the consideration of, and comparison with, the Panels' results. As the established Panel estimator, Powell's results will then be compared with a means-based Panel result, then against the pooled quantile results. Finally, Powell's and Canay's results will be compared and contrasted.

3.7. Results

3.7.1 Introduction

The results obtained from the various models outlined in the previous section are now considered. Separate models have estimated for males and females and are summarised in turn, and the results compared. Only the results for the qualifications' parameters will be shown, mostly in graphical, rather than tabular form. This makes it easier to compare and contrast the values obtained from the alternative models. Key test statistics will be reported at the appropriate points of the results' discussions.

Before looking at the results, it is worthwhile considering what the coefficients actually mean, in this context, in particular the quantile regressions' parameters. While the focus is on the four levels of qualification there are, in fact six levels assumed in the US Survey. Those not included are (i) the general 'Other Qualifications' category, a 'hotch-potch' collection of job-based qualifications and (ii) the ever-dwindling group of individuals who have no formal qualifications, and are the category excluded in this Study. Thus, all parameters reported are benchmarked against that category and indicate the effect on the wage rate from holding that particular level of qualification, as opposed to no qualification.

It is therefore relatively straightforward to understand what a particular means-based parameter estimates actually means. Any parameter associated with a particular level of qualification that is significantly greater than zero implies the holder's wage rate is increased by a certain amount, compared with an otherwise identical person with no formal qualifications. This may be used to estimate some life-long total value of the degree (the $\Sigma\Delta B$ in **[3.1]**), making assumptions about the numbers of hours per week both individuals would work at their respective rates, and for how long before retirement. The interpretation of conditional quantiles results is a little more nuanced and requires some exposition.

Figure 3.19 is derived from **Figure 3.10**. This depicts a situation where people have one of two qualifications' status: (i) No qualification, (ii) University Degree. The wage-rate distributions for both groups are indicated (NB: they are not necessarily the same shape, as the Figure implies) and the 10th and 90th centiles of each distribution marked with the black dots. Recalling the fundamental assumption made for CQR models that an individual locates at the *same quantile* of their new group, the quantiles of no-qualification and graduates' groups are connected by the upward sloping straight lines [41] Any means-based estimator would focus on the means of either distribution (μ_N and μ_D) and derive β_D (the parameter associated with the Degree Dummy) as the difference between them. The quantiles' estimator, on the other hand, would focus on the differences between the local means at the connected quantiles ($\mu_D^{10} - \mu_N^{10}$ and $\mu_D^{90} - \mu_N^{90}$ for example) to estimate β_D^{10} and β_D^{90}. As the figure shows, the vertical differences between them are not necessarily the same, nor are they necessarily the same as at the mean difference. The extent to which this is true, in practice, clearly depends upon the relative shape and locations of the governing distributions on the *w* axis for each group.

[41] Any other intermediate levels, such as FE, would have its own distribution, and its centiles would connect, in some fashion, with these.

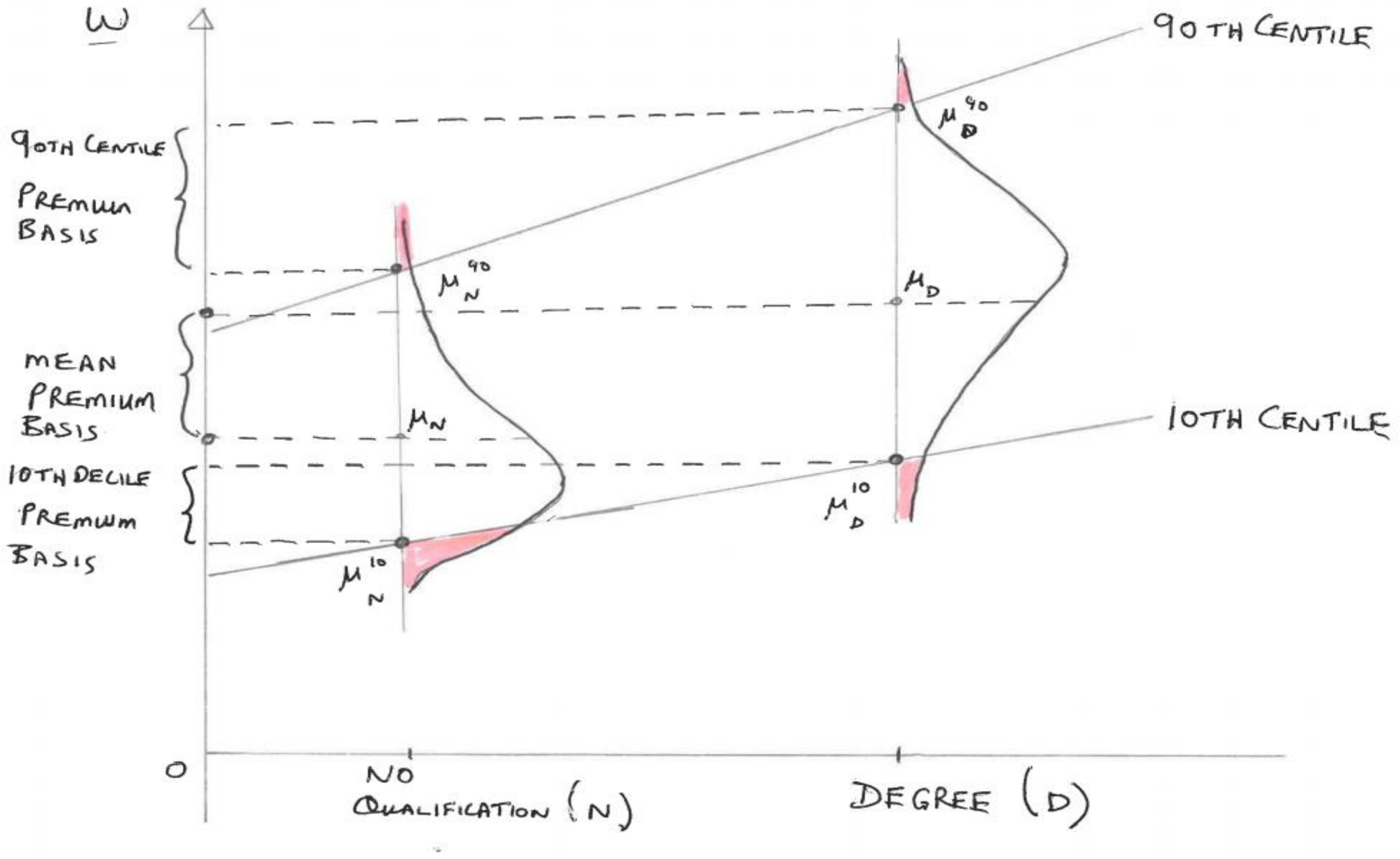


Figure 3.19 Means-based and Quantile Returns

As there are many regression tables generated in this Study (due, in particular to the quantile's regression producing one table per centile requested), the full tables are included in an Appendix at the end of this Chapter. This allows the main body of the text to focus on the key variables of interest (the sign and significance of the qualifications' dummies), that are presented either in tables or graphically.

3.7.2. Assessing the significance of the Attrition and Sample Selection Corrections

In the approach to the consideration of the results, the significance of the attrition- and sample-selection corrections, outlined in **3.6.2** and **3.6.3** are considered. As already noted, the BGLW tests on both Male- and Female models indicate significant attrition bias, and it is normal for earnings' studies to encounter significant sample selection bias.

By way of illustration, **Table 3.13** summarises the coefficients associated with Degrees for males and females at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} centiles with a means-based estimate included for reference. Pooled models with (i) no corrections and (ii) Sample selection only, and Panel models also with (iii) Attrition-bias and (iv) attrition- and sample section bias corrections are reported

Regarding the Pooled models, the effect of the sample selection corrections is to significantly increase the males' parameters (and therefore the wage rate premia), while the females' are left unchanged (note the significant downward effect on the OLS results).

Regarding the Panel models, the general effect of the attrition correction alone, is to lower the premia for both males and females. The sample selection corrections alone, increase the males' premia, but generally leave the females' unchanged. Making both corrections simultaneously nets out to increased premia for the males but reduced premia for the females.

Table 3.13
Effects of Attrition and Sample Selection Corrections on Degree Premia: Pooled & Panel Regressions
(Excluded category: No qualifications) (*** p<0.001)

	Pooled Models (qreg)											
	Males						**Females**					
	10th	**25th**	**50th**	**OLS**	**75th**	**90th**	**10th**	**25th**	**50th**	**OLS**	**75th**	**90th**
No Corrections	0.261*** (0.008)	0.401*** (0.010)	0.535*** (0.012)	0.517*** (0.012)	0.627*** (0.012)	0.661*** (0.023)	0.183*** (0.005)	0.326*** (0.006)	0.495*** (0.006)	0.480*** (0.010)	0.636*** (0.010)	0.664*** (0.022)
Sample Selection	0.278*** (0.010)	0.433*** (0.012)	0.574*** (0.013)	0.432*** (0.013)	0.649*** (0.013)	0.698*** (0.021)	0.180*** (0.005)	0.324*** (0.006)	0.497*** 0.006	0.409*** (0.011)	0.640*** (0.009)	0.661*** (0.018)
	Panel Models (qregpd)											
	Males						**Females**					
	10th	**25th**	**50th**	**Prais-Winsten**	**75th**	**90th**	**10th**	**25th**	**50th**	**Prais-Winsten**	**75th**	**90th**
No Corrections	0.264*** (0.001)	0.402*** (0.001)	0.534*** (0.001)	0.511*** (0.012)	0.624*** (0.001)	0.658*** (0.002)	0.184*** (0.000)	0.323*** (0.000)	0.495*** (0.000)	0.478*** (0.010)	0.637*** (0.001)	0.663*** (0.002)
Attrition	0.246*** (0.001)	0.384*** (0.001)	0.510*** (0.001)	0.495*** (0.013)	0.604*** (0.001)	0.668*** (0.003)	0.166*** (0.001)	0.328*** (0.002)	0.492*** (0.001)	0.469*** (0.011)	0.620*** (0.003)	0.638*** (0.002)
Sample Selection	0.280*** (0.002)	0.435*** (0.002)	0.580*** (0.001)	0.535*** (0.015)	0.647*** (0.003)	0.702*** (0.002)	0.179*** (0.001)	0.323*** (0.002)	0.501 (0.002)	0.478*** (0.010)	0.637*** (0.001)	0.662 (0.002)
Attrition & Sample Selection	0.257*** (0.002)	0.410*** (0.001)	0.553*** (0.007)	0.512*** (0.017)	0.627*** (0.002)	0.671*** (0.002)	0.167*** (0.001)	0.319*** (0.000)	0.492*** (0.001)	0.469*** (0.011)	0.629*** (0.001)	0.655*** (0.001)

It appears, therefore, that the effects of the corrections vary according to gender and quantile, with the males particular affected by sample selection bias, and everyone by the attrition from the Panel, with almost polar opposite net effects between the genders. Notwithstanding the differences between the genders, it is clear that both these corrections are important in this Study.

3.7.3. Pooled Regressions: Quantiles v OLS

The results for the Pooled means-based and quantile regressions are considered in this section. An OLS regression using the attrition weights was performed first, to allow Breusch-Pagan tests for heteroskedasticity to be performed. The null of homoskedasticity was emphatically rejected for both males and females' models, and so, models with robust standard errors were estimated. Rather than present the results as a forest of tables, they are instead presented graphically through the remainder of the Results' section. Where the comparison is between a quantile- and a means-based model, the latter's parameter estimates are displayed as horizontal line, with the 95% confidence band plotted at dotted lines at either side. The central quantile estimates are plotted in black, with the 95% confidences limits plotted in grey.

<u>Males</u>: The sample selection-corrected OLS model has a Wald $X^2 = 16{,}786.72^{***}$ and normally-distributed residuals and was able to comfortably reject the null that qualifications' premia were equal to each other, and collectively equal to zero. Tests of equality of the coefficients between the genders were also able to comfortably reject the null of equality, in favour of higher premia for males. The quantile regressions with the sample selection correction [42](using the same selection equation as the OLS/Heckman) indicated pseudo-R^2 of 25 – 30%. The results for Degrees, FE, School-Leaving and School-Intermediate qualifications are plotted together in **Figure 3.20**.

[42] λ, the Inverse of the Mills' Ratio is λ = minus-0.15 (p = 0.0000) for the Heckman Model, while it was significant to at least p = 0.05 at the 25th, 30th, 40tth and 50th centiles. Full results of the regressions are in **Tables A.3.1** and **3.2** in the Appendix at the end of this chapter.

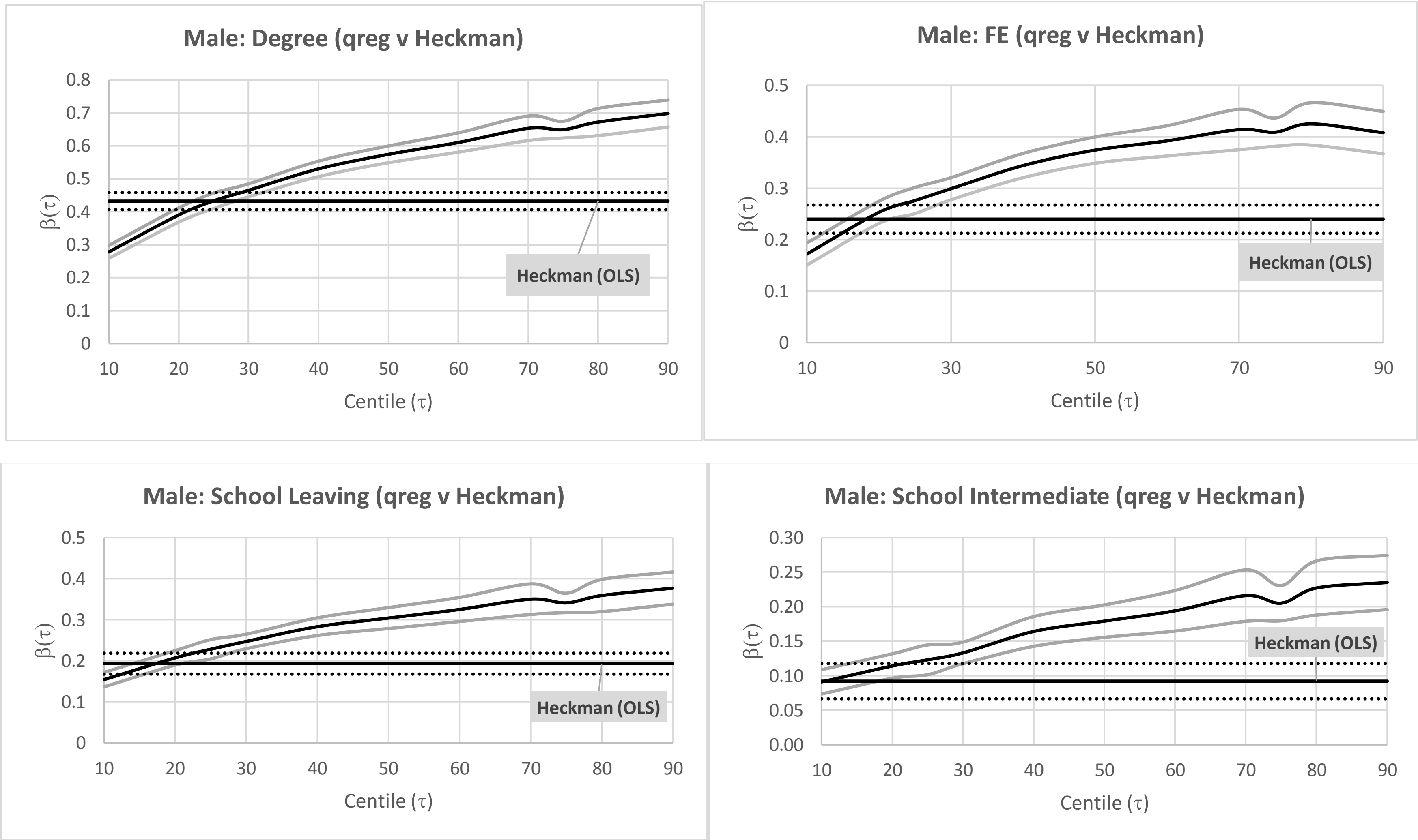


Figure 3.20 Pooled Regression Results: Males

Looking first at the Degree results, both the means-based and quantile models agree with each other that the qualification confers male holders with positive premia right across the wage-rate distribution. The returns rise steadily through the quantiles, with returns significantly below the means'-derived central estimate below the 20th centile, and above it around the 33rd centile. F-tests of equality between the τs of the adjacent key quantiles were only able to reject the Null of equality between the $\beta(10)$ and $\beta(25)$ centiles (p = 0.028) and the $\beta(25)$ and $\beta(50)$ (p = 0.0431).

FE qualifications also confer significant premia to their holders across the full distribution, also rising in the quantile (tests between adjacent centiles were significant). The premia are, however, below those for degrees (as might be expected). The quantiles premia were significantly less below the means'-based estimate until the 15th centile, then significantly above it, after the 28th centile.

School-Leaving qualifications are similar to the FE qualifications: significant wage rate premia across the distribution rising in the quantiles, with concurrence between the quantiles- and means regressions between the 12th and 27th centiles.

School-Intermediate qualifications are also able to confirm a positive wage rate premium though these are lower than all other qualifications. The quantiles' results are never below the means-based ones and are significantly above them after the 30th centile.

Females: The Sample-selection corrected means model has Wald X^2 = 19735.05*** and has normally-distributed residuals. Their results are displayed in **Figure 3.21**. The story for female graduates is similar to the males, insofar that there are always positive returns to all qualifications across their wage rate distribution while rising in the quantile.[43]

[43] λ = minus-0.10 (p = 0.000) for the Heckman Model, while it was significant to at least p = 0.05 at the 10th, 20th, 25th, 30th, 70th, 75th, 80th and 90th centiles. Full results of the regressions are in **Tables A.3.3** and **3.4** in the Appendix at the end of this chapter.

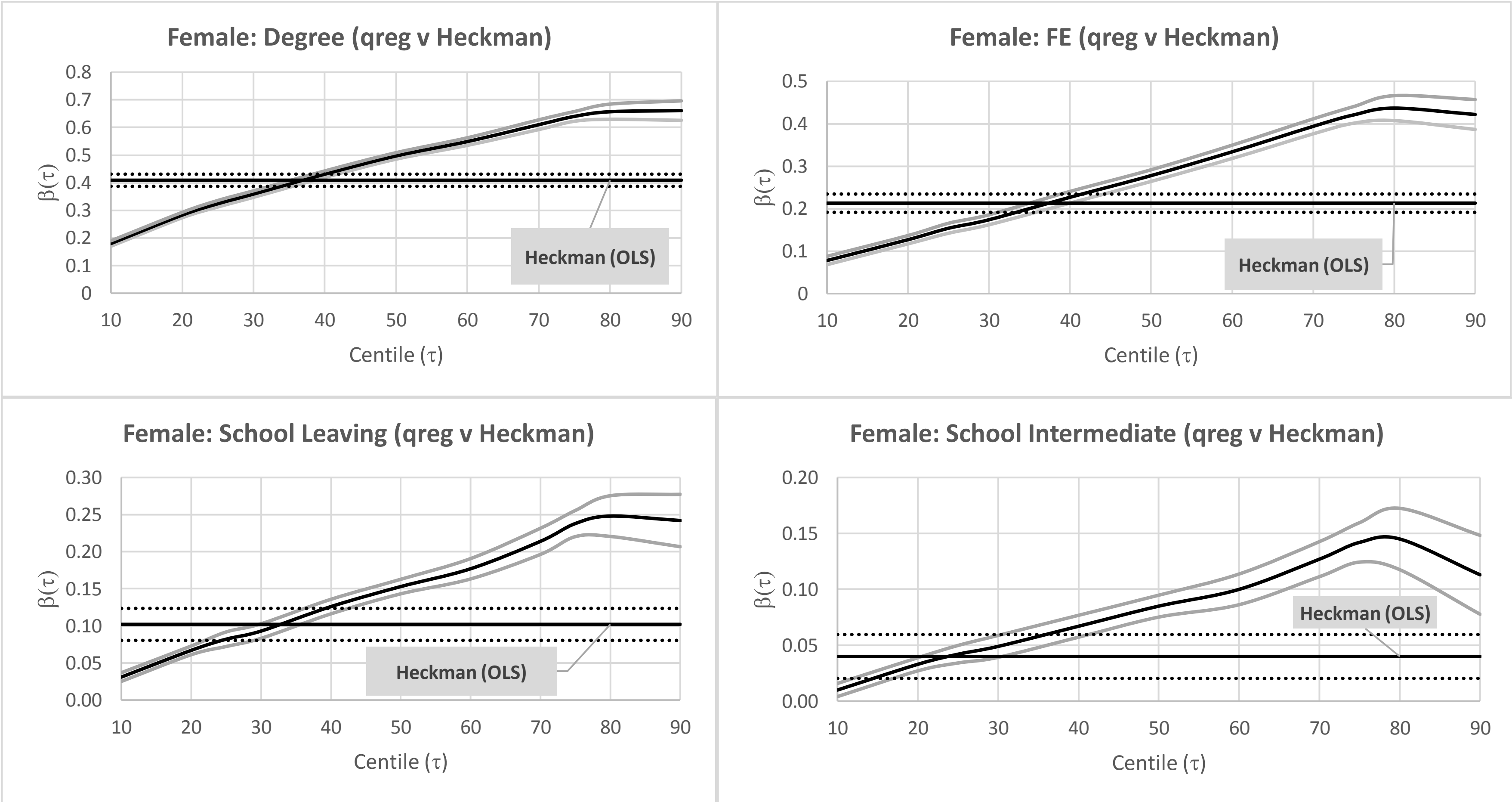


Figure 3.21 Pooled Regression Results: Females

Degrees The returns for the females are effectively the same for males with only a slight difference in the span of quantiles in which the means-based and quantiles' models estimate the same returns. FE returns are also broadly the same as the males, though the quantiles' trajectory is steeper, given the lower premia at the bottom end of the wage rate distribution. There is a very narrow zone of concordance, between the 30th and 40th centiles.

The returns to School Leaving qualifications are lower for females than for the males, though they also confer positive premia throughout the distribution. Similarly, the returns to School Intermediate qualifications lie below those for the males and are only just able to confirm any premium at all at the bottom end of the distribution.

Whilst the purpose of this section is largely to set the scene, insofar as it does not involve any of the Panels'-based models, it is, in itself, a useful exercise as it indicates that:

- All qualifications confer significant wage rate premia on all holders, compare with those with no formal qualifications.
- Male- and female graduates and diplomates have the same premia across the quantiles, though females qualified at the school levels are paid less than their male counterparts.
- Means-based estimates can mask significant differences in premia at different points on a wage rate distribution.

The next section considers the Panel-based estimates and whether, or not, including the unobserved component makes any significant differences to the qualifications' premia.

3.7.4 Panel-based Results: Powell's FE Quantiles versus Means-based (Prais-Winsten)

This section examines the returns to qualifications using Powell's Fixed Effects estimator discussed in earlier. Instead of comparing this directly against OLS, the benchmark will be against a means-based Panel estimator, in particular, the Prais-Winsten model, corrected for attrition, sample selection and serial correlation (and reported for Degrees, in **Table 3.13**). As before, the results for the four qualifications' categories for males are considered first, then those for females.

Males: The Wooldridge test for autocorrelation in the male error terms comfortably rejected the null on no autocorrelation (F= 45.384***), confirming that a model assuming (first-order) serial correlation is appropriate. The results are displayed in **Figure 3.22**. [44] As with the Pooled results, it is apparent that all qualifications confer positive pay-rate premia on males across the distribution, that rise, generally, in the quantile. The pecking order is not quite the same as seen with the Pooled specification, with degrees achieving higher returns than both FE- and School-leaving qualifications, which yield similar returns to each other: with the Pooled model, FE conferred greater premia than School Leaving qualifications. Intermediate School qualifications again, yield the lowest returns. Looking at Degrees, the quantiles' model yields returns significantly below the means'-based result up to the 32nd centile, then significantly greater returns above the median. This contrasts with the lower- and narrower zone of concordance found with the Pooled results for degrees (see top left graph in **Figure 3.20**). With the FE qualification, the quantiles lie below the means-based estimate up to the 30th centile, then above it after the median. As with the degrees, the concordance zone is wider and higher up the distribution than with the Pooled results. The trajectory is also fairly flat after the median, though it was noticeable that it actually *fell* after the upper quartile in the Pooled results.

[44] λ = 0.002 (p=0.061) for the Prais-Winsten model but significant in all the quantiles' models, except at the 50th quantile. Full results of the regressions are in **Tables A.3.5** and **3.6** in the Appendix at the end of this chapter.

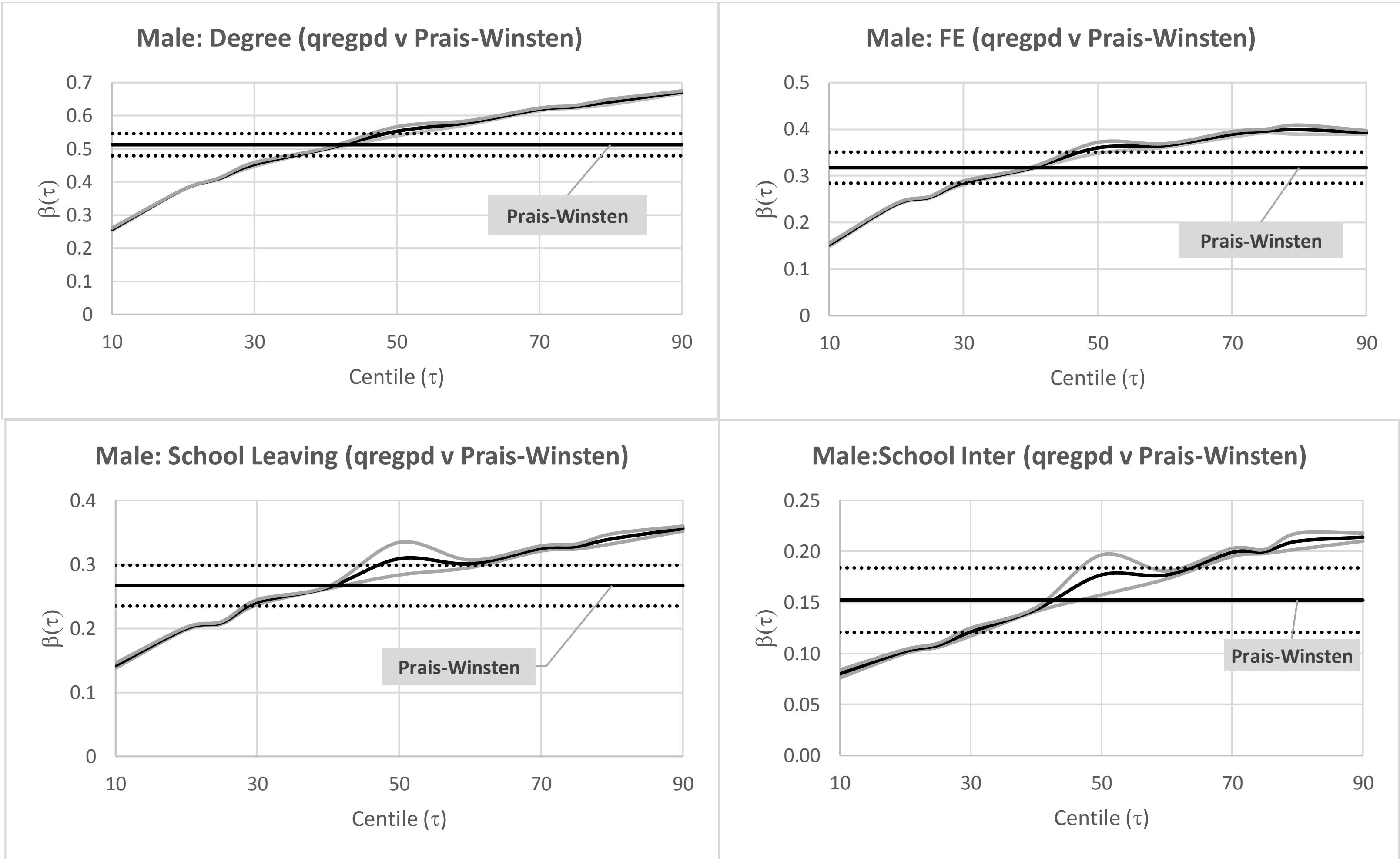


Figure 3.22 Panel Regression Results: Males

The returns to School Leaving qualifications were greater than those estimated by the Pooled models, hence their apparently equal performance with the FE qualification. The concordance zone lies between the 30th and 60th centiles, again in contrast to the pooled results where it ay between the 15th and 30th (approximately).

Similarly, the returns to the lower-level, School Intermediate qualifications are above those estimated by the Pooled models and are also significantly less than those for the School-leaving and FE qualifications. The concordance zone is also broader and higher (30th – 65th) than in the Pooled case.

Females: The Wooldridge test on the females' data also comfortably rejects the null of no serial correlation (F = 22.992***) so the Prais-Winsten means-based estimator is again used to compare against Powell's panel-based estimator. The results are plotted in **Figure 3.23**. [45] There are noticeable 'kinks' in the quantiles' trajectories with relatively wide standard errors around the upper quartile, but otherwise, the premia rise in the quantiles for all qualifications. It is just possible to conclude that, for females, all qualifications confer a wage-rate premium, though the premia for School Intermediate qualification was very close to zero at the bottom end of the distribution. There is the expected pecking order of returns (i.e. Degrees > FE > School Leaving > School Intermediate), whereas for the males, FE and School Leaving were equal, as noted above.

For Degrees, the means-based Panel estimate is significantly higher than the Pooled means-based estimate, but the quantiles' estimates are generally the same across the range. They are also not significantly different to the returns enjoyed by Male graduates. The concordance zone was further to the right than with the Pooled model (~40th Centile to the median), but had the same span, but narrower than for the males.

[45] λ = minus-0.306 (p=0.432) for the Prais-Winsten model but significant in all the quantiles' models, except at the 80th quantile. Full results of the regressions are in **Tables A.3.7** and **3.8** in the Appendix at the end of this chapter.

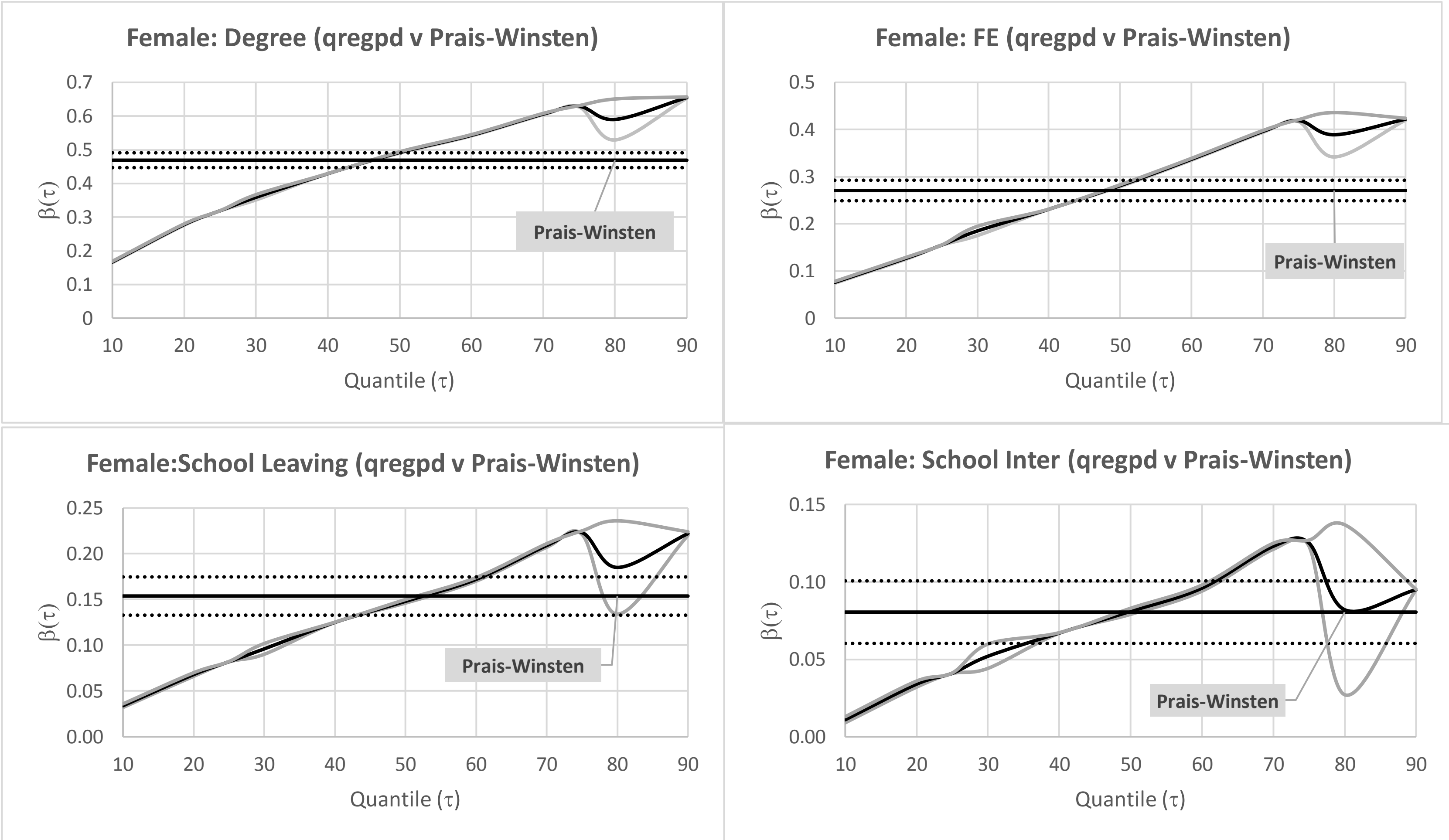


Figure 3.23 Panel Regression Results: Females

For FE, the mean returns were significantly above those estimated by the Pooled model, although significantly lower than for graduates, but above females with School Leaving qualifications. However, the quantiles are not significantly different from those estimated from the pooled data. The concordance zone is slightly higher up the distribution (40th centile – median) than in the Pooled case but has a similar width. It is, however much narrower than for the males.

For those with School leaving qualifications, the mean returns are above those for the Pooled Data, although the quantiles' results are not significantly different. The concordance zone is also further up the wage rate distribution (40th – 60th centile, approximately) and narrower than the males'.

The School Intermediate qualification really struggles to confer any premia at all on the females, particularly below the median. The trajectory across the quantiles is upwards, but struggles to escape the means-based estimate, in contrast to the pooled result which indicate a steady rise until the upper quartile. As a result, the zone of concordance is very wide (35th centile – upper quartile), much further up the distribution than with the pooled estimates, though broadly the same as for their male counterparts.

The Panel results also confirm the Pooled models headline findings that:

- Qualifications significantly increase wage rates for males and females, but there are some differences between the genders, particularly below degree level.
- Means-based results mask significant differences in premia across the wage rate distribution

In addition, there are some differences evident between the Pooled and Panel results:

- The zone of concordance (i.e. where the means-based and quantiles'-based results do not differ, tend to be further up the wage rate distribution.

- The means-based Panel results can estimate higher returns than those from the pooled data

The next section compares these same quantile returns against the standard Cross-section quantile estimates.

3.7.5 Quantiles' estimators: Powell's Fixed Effects Panel (qregpd) versus Cross-sectional returns (qreg)

The comparison brings together results already seen in the sections above but allow us to compare directly the quantiles'-based results for Powell's FE Panel estimator against the Pooled quantiles' estimator. The results are displayed in **Figures 3.24** for the males and **3.25** for the females.

In general, it can be seen that there are no significant differences in the quantiles' premia inferred by both models, for both males and females and for all levels of qualification. The standard errors estimated by the Pooled (qreg) estimator are greater than those estimated by the Panel estimator (qregpd) and are almost always able to encompass the Panel estimators within the 95% confidence interval (i.e. the grey-shaded zones), albeit below the central estimates. The downward trajectory (and large standard errors) experienced by the females at the upper ends of the ware rate distributions is, again, noticeable.

The rather unexciting conclusion, therefore, has to be that the unobserved effects do not exert any significant changes to the estimated premia across the quantiles, and that what really matters is the identification of these premia at each quantile, rather than just at the mean. [46]

[46] . Full results of the regressions are in **Tables A.3.2**, **3.4**, **3.6** and **3.8** in the Appendix at the end of this chapter.

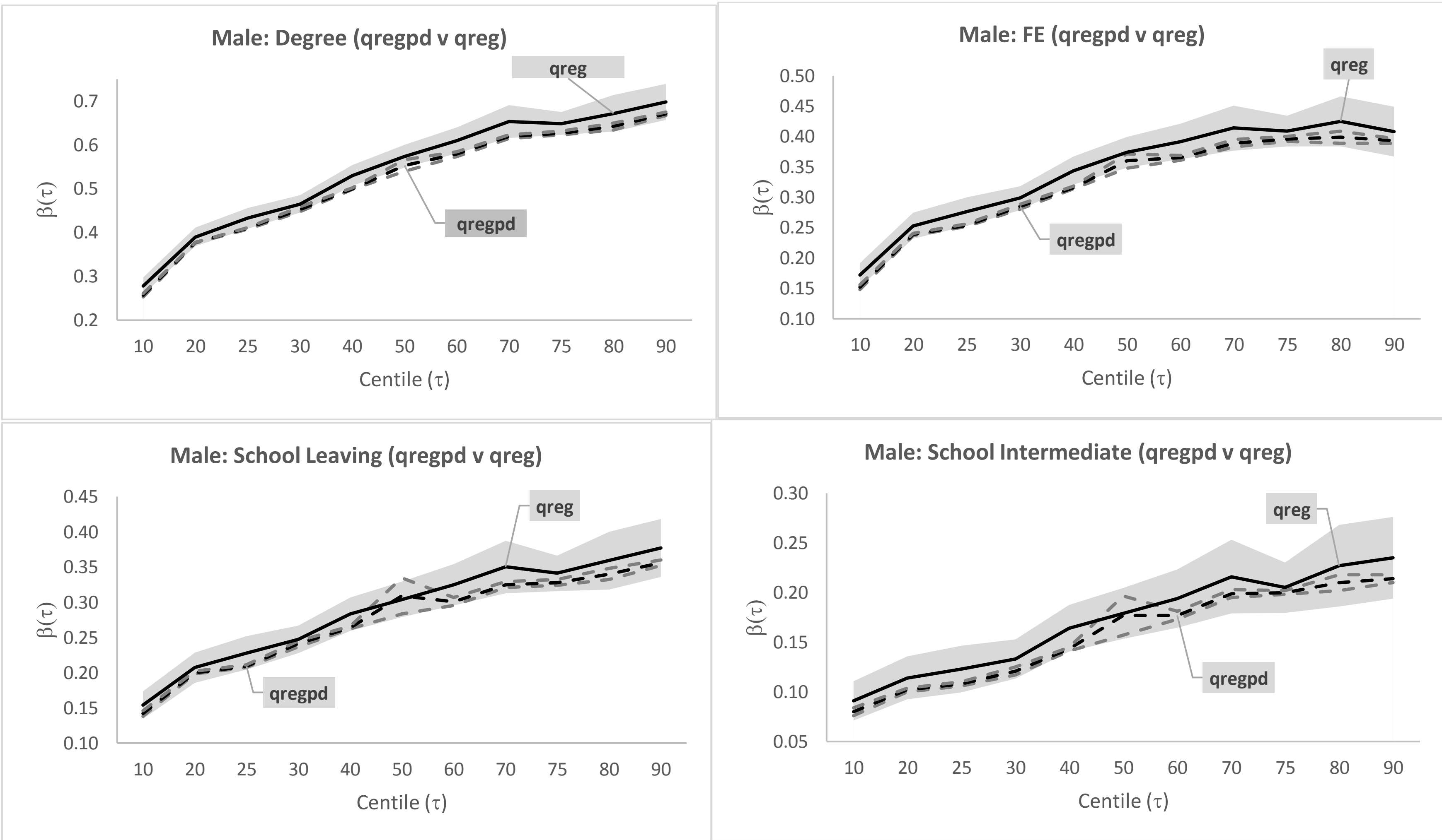


Figure 3.24 Panel and Pooled Quantiles' Results: Males

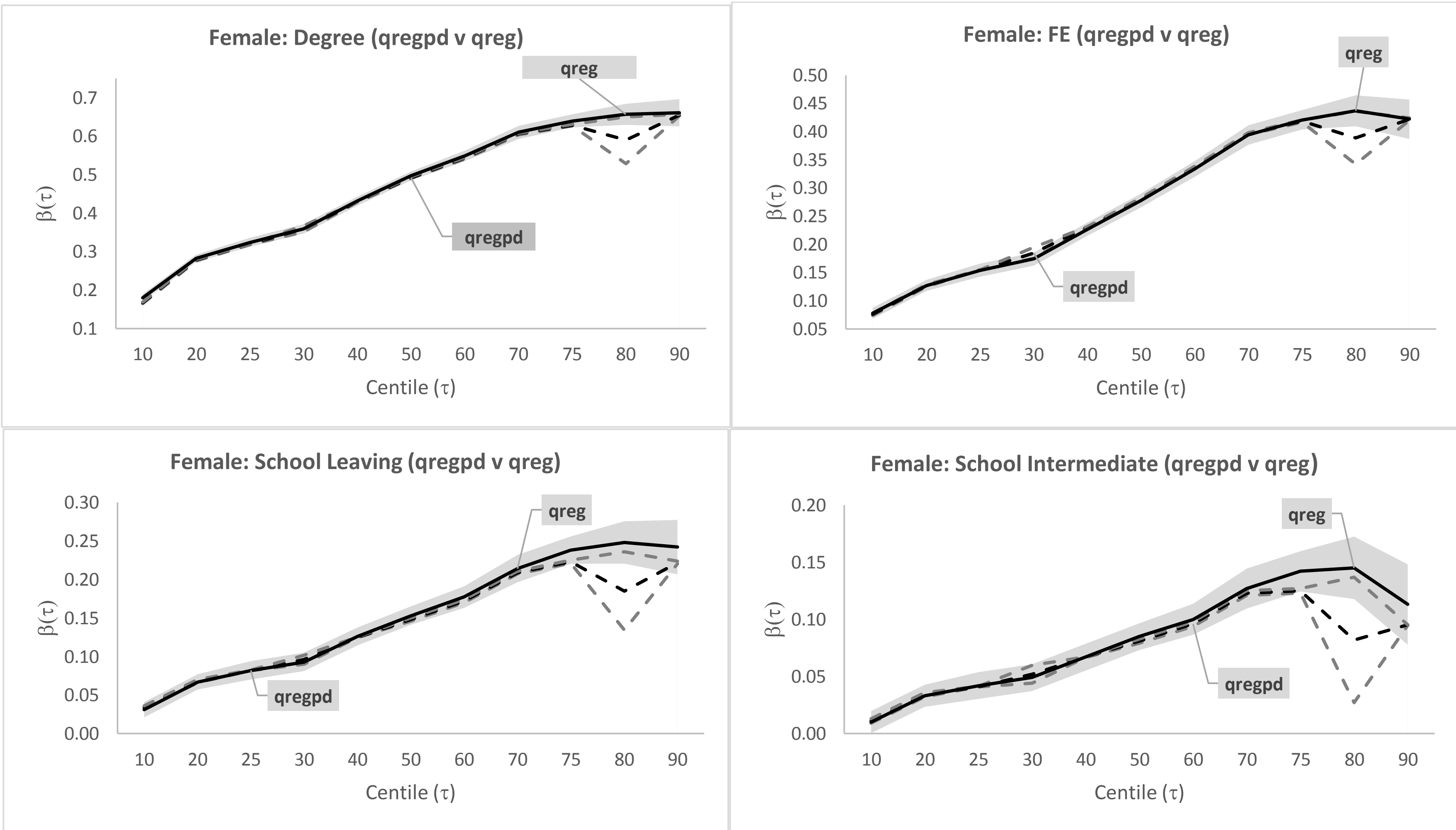


Figure 3.25 Panel and Pooled Quantiles' Results: Females

3.7.6. Non-additive v Additive Fixed Effects Panel Models (Powell v Canay)

The essential difference between Canay's additive and Powell's non-additive Fixed Effects was discussed in **3.3.2**. To reiterate the differences, Powell asserts that separating out the estimated fixed effects from the disturbance term means that any resulting regressions relate to a transformed dependent variable, $\hat{y}_{it} = (y_{it} - \hat{\hat{\alpha}}_i)$ and not the variable of interest, y_{it}. If the fixed effects are significantly correlated with the y_{it}, then the effect will be to alter the shape of the distribution of $\hat{y}_{it}$, even to the extent that the rank order of the individuals within it will alter significantly. As a consequence, the conclusions from these regressions may differ significantly from regressions on y_{it}.

In order to assess this, Canay's two-step procedure was coded in Stata and the fixed effects for the females and males estimated separately. These are plotted as histograms at the top of **Figure 3.26**. These appear broadly similar, centered around zero with a 'normal-esque' distribution with a moderate positive skew. The males' effects range between minus-0.5 to plus-one, while the females' is narrower, ranging from minus-0,5 to plus-0.75. Fixed effects are positively correlated with the wage rate with r_{male} = 0.767 (p=0.000) and r_{female} = 0.757 (p=0.000) and, therefore the tendency to compress both wage rate distributions, and presumably, alter the internal location of the individual within their wage-rate distribution.

The middle- and bottom pairs of histograms indicate the effect of the transformation of the log-wage rate variable, $ln(w_{it})$to $ln(w_{it} - \alpha_i)$. The effect of the transformation is to make each $ln(w_{it} - \alpha_i)$ distribution more 'normal' in appearance and, as expected, to reduce their ranges. What is less clear, is the extent to which individuals leapfrog each other as a result of the transformations. Spearman's coefficients of *rank* correlation (ρ) are ρ_{male} = 0.771 (p=0.000) and ρ_{female} = 0.810 (p=0.000), indicating a general preservation of relative position, but with some internal reordering. There is, therefore, some expectation that the Powell- and Canay estimates may be somewhat different for both males and females.

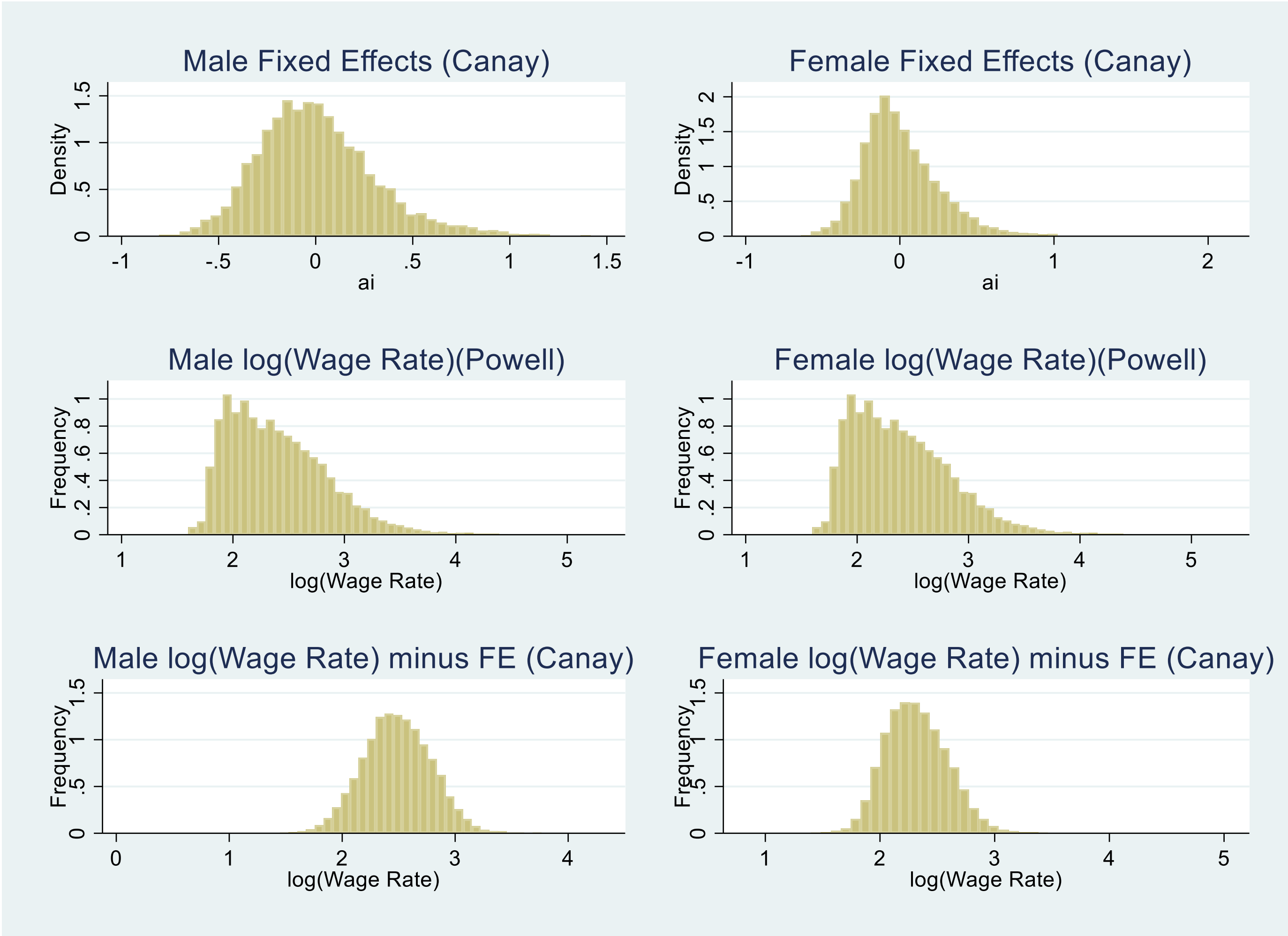


Figure 3.26 Estimated Fixed Effects, ln(wage rate) and [ln(wage rate)-α_i] distributions

This does appear to be the case in this Study, as **Figures 3.27** (Males) and **3.28** (Females) indicate. [47] While both estimators generally confirm the existence of qualifications' premia for males, the trajectories of the Canay-derived results between the 10^{th} and 20^{th} centiles are generally much flatter than Powell's, with crossover occurring around the median (except for School Intermediates). The Story for females is slightly different, in the sense that there is no steep trajectory in the Canay estimates between the 10^{th} and 20^{th} centiles. The post-School qualifications (Degree and FE) are broadly the same as for the males, with a flatter trajectory for the Canay results, crossing from above- to below the Powell estimates around the median. For females with School-leaving and Intermediate qualifications, however, the crossover is never completed. For the Leaving qualification, Canay's returns are above than Powell's until around the 60^{th} centile, then equal thereafter (i.e. they are never lower than Powell's at any centile). For the Intermediate qualification, Canay also exceeds Powell's until the 70^{th} centile, then are equal thereafter. These differences arse from the strange downturn in the females; premia at the 80^{th} centile that affect both Canay and Powell, but particularly, Powell. As these may be data-dependent (the males' results are not similarly afflicted) then it is not possible to draw any special conclusion from this.

In summary, the Canay estimates are less able to find consistently increasing premia across the quantiles for males, and the relatively large wide confidence limits make it difficult to conclude that even $\beta(90) > \beta(10)$. In fact, these trajectories track quite closely the means-based results (**Figure 3.22** and **3.23**) and would deny the importance of estimating quantile- as opposed to means-based estimates of returns.

[47] The Inverse-Mills parameters were statistically significant in the Canay models at the 20th, 25^{th}, 30^{th}, 40^{th}, 50^{th} and 70^{th} Centiles (Males) and the 10^{th}, 20^{th}, 30^{th} and 90^{th} centiles (Females). Full results of the regressions are in **Tables A.3.9** and **3.10** in the Appendix at the end of this chapter.

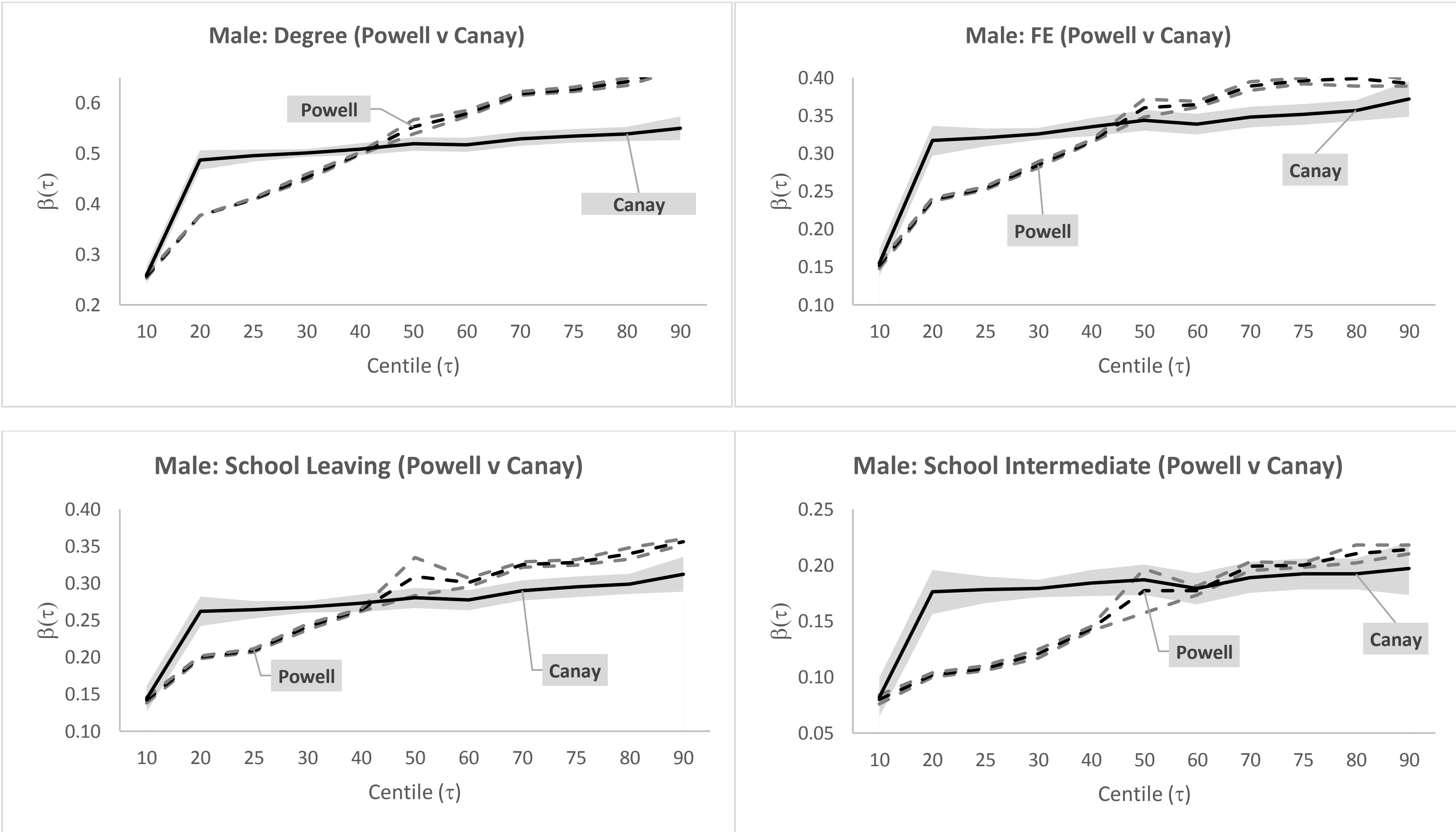


Figure 3.27 Canay v Powell Panel Quantile Estimators (Males)

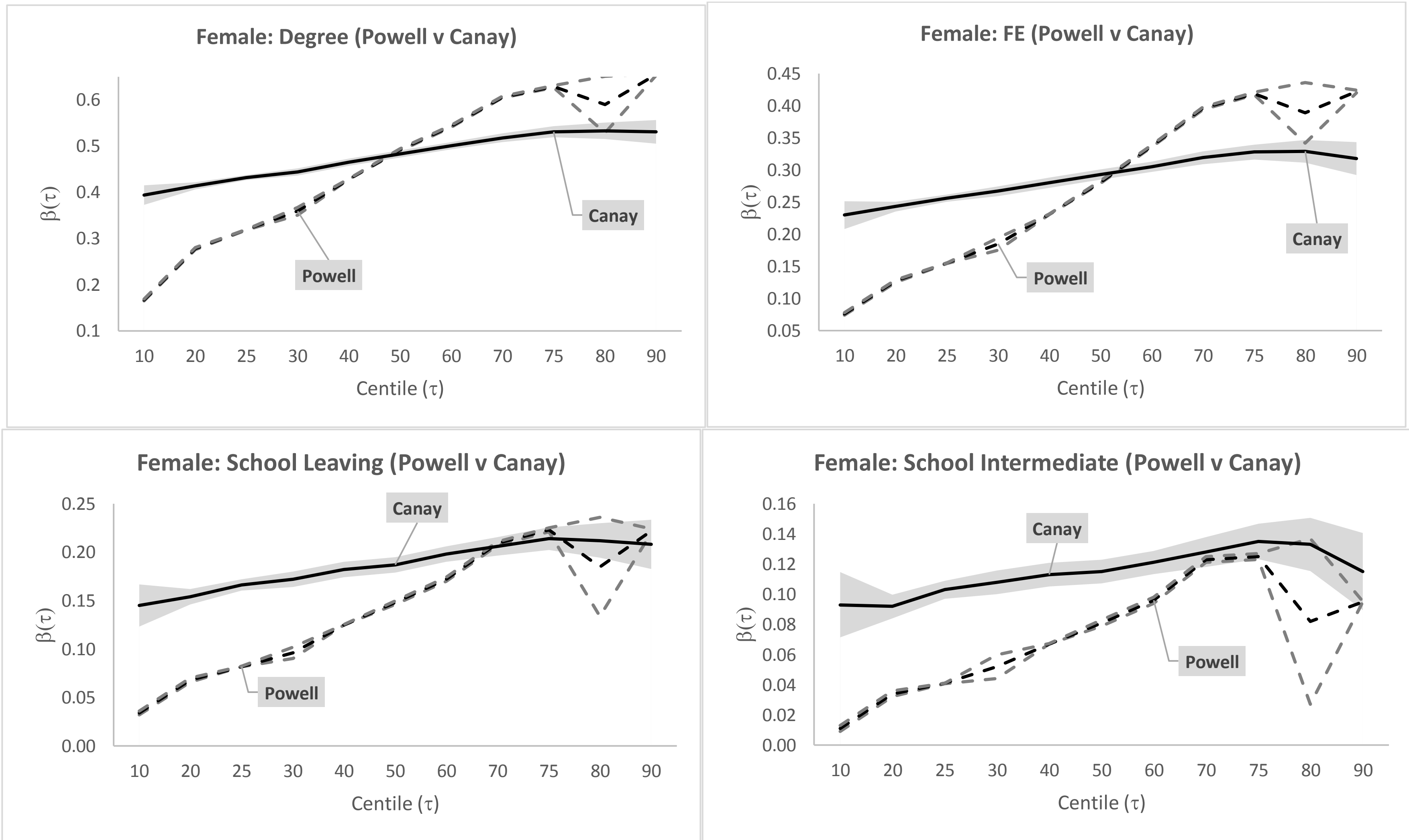


Figure 3.28 Canay v Powell Panel Quantile Estimators (Females)

3.7.7. Relative Premia

The findings above indicate positive premia accruing to qualifications for both males and females, with only just in the case of School Intermediate qualifications for females. It is evident that the premia differ according to the quantile the individual is located at on their wage rate distribution, although it cannot be claimed that the differences between *adjacent* quantiles are always significant. It is the case, however, that the differences between $\beta(90)$ and $\beta(10)$ are always significant, in favour of the higher quantile i.e. premia rise in the quantile. The question arises: how significant to the individual is the premia they receive from their qualification, given where they are on the overall wage-rate distribution? For example, suppose individual A has an FE diploma, and is paid £6 an hour, £2 of which can be assigned to that qualification. Individual B also has the same qualification, but is paid £12 an hour, £3 of which can be assigned to the qualification: a higher absolute premium. However, 33% of A's wage rate can be assigned to the Diploma, but only 25% of B's: the *relative* value of the diploma to A is greater than it is to B. In effect, it pushes A further towards the top end of their centile's distribution than it does B (recall **Figure 3.19**).

The answer to that question is dependent upon many assumptions that have to be made about the individual. Given the variables included in the regressions earlier in this section, these would depend upon:

- The gender of the individual;
- Their labour market experience (in years);
- The sector of the labour market they work in;
- What region of the UK they live in;
- Their domestic partnership arrangement.

As these are deemed to contribute something towards their wage rate. By way of illustration, an illustrative pair of individuals, one male, one female, both working in the low-paying Accommodation & Food Services in the Private Sector, co-habiting with a partner and living in Scotland, but with no second job have their wave 7 (2015) pay rates estimated using the Panel-based attrition and sample selection-corrected results reported in **Table 3.13**. The wage rates for both individuals were computed on the basis that they had no qualifications, then the premia for school intermediate, school leaving, FE diploma and Degree added in turn, each at the 10th, 25th, 50th, 75th and 90th centiles. The percentage increases in each estimated wage rate were computed, in relation to the base case (no formal qualification). The results are summarised in **Table 3.14**.

Table 3.14
Percentage boost to wage rate attributable to qualifications.
Male & Female, 20 years LM experience, working in Scottish-based, Private-sector Accommodation & Food Services company: Wave 7 (2015)

	Centile				
	10th	25th	50th	75th	90th
Degrees					
Males	29.6%	51.3%	71.8%	86.5%	95.8%
Females	17.9%	37.6%	63.6%	87.0%	92.1%
Difference	11.6%	13.7%	8.2%	-0.6%	3.7%
FE					
Males	16.8%	29.6%	42.0%	47.6%	48.0%
Females	7.8%	16.9%	32.4%	51.7%	52.2%
Difference	9.0%	12.7%	9.6%	-4.2%	-4.2%
School Leaving					
Males	15.5%	24.1%	33.4%	38.4%	42.9%
Females	3.3%	8.5%	16.1%	25.1%	24.6%
Difference	12.2%	15.6%	17.3%	13.3%	18.3%
School Intermediate					
Males	8.5%	11.9%	17.5%	21.8%	23.9%
Females	0.9%	4.3%	8.4%	13.7%	10.0%
Difference	7.6%	7.6%	9.0%	8.1%	13.9%

Unsurprisingly, the largest pay rate boost accrues to graduates, with the relative boost rising sharply in the quantile. Recalling that the fundamental assumption of the CQR model is that individuals transit to the same quantile of their new (education) group, the above-median individuals in the unqualified group are therefore transiting to the above-median parts of the graduates' pay-rate distribution: should such people actually exist, the increase in their pay rate will be relatively massive. [48] Female graduates located below the median do not enjoy the same relative boost to their payrates as do their male counterparts, but those above the median close, even reverse, the gap.

There is a similar story for FE diplomates, whose relative pay boost increases with the quantile they are located in (before and after the acquisition of this qualification). Below-median males benefit from a relatively greater pay boost than the females, but the direction reverses at, and above, the upper-quartile (bearing in mind that females are always starting from a lower pay rate at each centile).

School Leaving qualifications are also effective for males in raising their pay rates, in a fashion similar to FE qualifications. For females, however, their performance, though boosting their pay rate, pales in comparison to the males, and the gender gap is not closed in the way it is for post-school qualifications.

The performance of the School Intermediate qualification mirrors that of the higher level but, as might be expected, at a lower level across the distribution. The exceptionally poor performance of this qualification for females, in particular at the bottom end of the distribution is noticeable, with the gender gap not being closed at all. Of course, these specific results pertain to a closely defined pair of individuals, but their general nature and magnitude are robust to changes in the assumptions made. The final section in this Study will review the findings and consider the implications for Policymakers.

[48] In reality, it is unlikely that many such individuals will exist, as the vast majority of individuals transit from unqualified to school intermediate, then school leaving, then to graduate status, with others transiting via FE, perhaps having missed out on one, or both the school stages.

3.8. Final Discussion and Further Work

3.8.1. Discussion

The UK Economy has transited through periods of rapid industrialisation and growth, during which time workers moved off the land into the rapidly-growing towns and cities. The decades following the Second World War can be characterised into periods of rapid technological development and automation, which would have otherwise had the immediate effect of displacing large numbers of workers from the production process, had it not been for the high level of growth, fuelled principally by the boom in consumerism afforded by real increase in workers' wages. Even with the start of the de-industrialisation process of the UK during the late-1960s into the 1970s, with the concomitant breakdown of the apprenticeship system, the labour market, school leavers were still able to find stable employment without necessarily having gained any formal qualifications while at school. However, continued technological development, the relocation of manufacturing jobs outside the UK and the increased participation of female workers in the labour market during the 1980s and 1990s, the need for *credentialism*, i.e. the demonstration of skills via formal qualifications, became ever-more apparent. The policy responses of successive UK governments during that time was to incentivise training in both 'old' and 'new' jobs: 'old' referring to job-creation and apprenticeship subsidies: 'new' to bearing the costs of training in Information Technology, and an overt upskilling of the workforce via an expansion in the post-School Education Sector. As the Public Purse was unwilling to bear the full additional costs of this rapidly-expanding sector, the students became burdened directly with the costs of their tuition and subsistence. This was 'sold' on the basis that their (marginal) lifetime earnings would inevitably exceed the (marginal) cost of acquiring their credentials. It also chimed with the zeitgeist that the person benefitting most from the credential (viz, its holder) should be the one that pays for it. The fact that wider Society would also benefit from process was largely ignored, although the full marginal cost of the tuition was not placed directly on the student, as tuition fees were capped, student loans subsidised and not repayable until a certain income threshold had been achieved, for example.

This transferring of the overt costs of acquiring the credential has helped to sustain academic interest in the labour market value of qualifications, as evidenced in the literature review in this Chapter. It is as if Society is constantly seeking reassurance that a particular qualification is worth acquiring, with either the fear, or the expectation that policymakers may divert resources away from those that appear not to offer 'value for money', as perhaps defined in **[3.1]**, towards those that do. The situation at the time of writing, is that potential students can obtain estimates of the employment probability and/or annual earnings associated with specific programmes, allowing them to make an 'informed' choice as if they were buying car.

The traditional approach to estimating the (marginal) earnings associated with a particular qualification is to estimate a Mincer-style earnings'-rate equation (**[3.2]**) with respondents' highest reported qualifications coded as dummies and regressed, alongside a selection of other independent variables, against an hourly wage rate. The sign and significance of the qualification dummies are taken to be the extent to which any individual's wage rate would be altered were they to acquire that higher level of credential. These 'premia' can then be used in formulae like **[3.1]** to 'guesstimate' the labour market value of the qualification, though this would really require additional estimates of employment probabilities and earnings' trajectories over particular working lives. The majority of the UK-related research easily confirms the existence of significant, positive wage-rate premia accruing to qualifications and that, therefore, all qualifications are worth acquiring. The same research also confirms an expected pecking order, with degrees conferring significantly greater premia than FE Diplomas and School-level qualifications, while FE diplomas may or may not confer significantly greater premia than School-leaving qualifications, but both will perform better than lower-School qualifications. The premium to degrees is currently maintained, despite the evidence of ever-growing graduate *underemployment*, where recent graduates are unable to find 'graduate-level' employment and remuneration for some years after graduation. The effect may be that increased pressure is placed on non-graduates, who displace each other, in turn all the way down the pecking order.

Most of this research focusses on the *average* premia accruing to qualifications, that is the premia accruing at the centre of the wage-rate distribution. This is to imply that whatever relationships appear to pertain between the wage rate and the independent variables (including the qualifications' dummies) apply across the entire wage rate distribution. This is a strong assumption, and may, in this scenario, disguise significantly different relationships between qualifications and wage rates in terms of statistical significance, sign and relative importance (i.e. percentage of the wage rate) at different points of the entire wage rate distribution.

This Study estimated qualifications' premia at various quantiles of the UK's wage rate distribution, using a seven-wave Panel derived from the Understanding Society (US) survey, using a *Conditional Quantiles' Regression* approach. This, critically, assumes that an individual transits from one status to another, but maintains the same relative position in the new status' dependent variable distribution. In the context of this Study, this is to assume that an individual with, say, School Leaving qualifications earning the median wage rate enjoyed by all those with that level of qualification, would earn the median graduate wage rate if they were to obtain a degree. As males and females continue to experience rather different outcomes in the UK labour market, separate models were estimated for both genders. Given the high level of attrition from the US survey, participation was tested and judged to be non-random, necessitating the computation and use of attrition weights in an effort to counteract the effect. These were computed using a number of variables thought likely to influence continued participation, including whether or not the respondent had plans to move before the next scheduled survey. A Heckman two-step procedure was also applied to counter the effects of sample selection bias concerning the work/no-work decision: given the assumed differences in this process that exist between males and females, separate probits were estimated and applied. The probits included a number of standard variables assume to affect the work decision, including age, housing tenure and caring responsibilities. In general, the attrition- and sample selection corrections altered the parameter estimates significantly, so appear justified in this case.

The Mincer models included a number of standard independent variables other than the qualifications' dummies, including labour market experience, Region, partnership status and industrial sector. Occupation was specifically excluded as this is highly correlated with qualifications.

A number of models were estimated, both treating the data as a pooled cross section and as a Panel, using the same attrition- and sample selection corrections. Means-based models were estimated to provide a contrast with the quantiles' results. It was clear from the pooled- and panel-based models that the premia were not the same across the entire wage rate distribution for all qualifications, for both genders, with the associated parameters rising in the quantiles. All qualifications yielded significant, positive premia to all holders, with degrees, unsurprisingly, outperforming all others, for both males and females. School-level qualifications perform relatively better for males than females, with the latter's Intermediate returns hovering close to insignificance at the lower end of the wage rate distribution.

There is no evidence that the estimating the quantile premia with a Fixed-effects Panel specification yielded any significantly different results from the same data, pooled. However, the type of panel estimator assumed does appear to matter, with the Canay estimator tending to 'flatten' the trajectories when compared with the Powell estimator incorporated in Stata (especially for males).

While it is clear that all qualifications offer their holders a significant wage rate premium, it is interesting to note the differences in *relative* premia, varying according to the centile, the qualification and gender of the individual. Of course, receiving such percentage boosts to their wage rates would require the critical assumption that people are projected onto the same centile of their new group's distribution to be realised in practice.

While this is probably unrealistic, the Study has revealed that the assumption that qualifications' premia do not vary across the pay rate distribution is too strong, and that quantile regressions should be employed in studies on returns. There are also observable differences between males and females in the UK Labour market, at least during the Study Period.

3.8.2. Future Work

This Study has made a rare foray into the estimation of returns to qualifications at different points of the wage rate distribution. Further work to this Study would be to relax the assumption inherent in the CQR model, that people transit to the same relative point on their new wage-rate distribution. In order to better model lifetime earnings, more would have to be learned about the actual inter-quartile movements people acquiring qualifications actually make, and from there, what progress they make along their group's distribution in the subsequent periods. In other words, their *lifetime trajectories* with the aim of isolating the long-run impact of the qualification(s) they acquired along the way.

This Study did not consider the impact of the qualifications on the *probability* of employment in the sense that they may have revised their work decision, while employers revise their employment decision. Follow-up work would also consider this key aspect, with the focus moving to *expected* lifetime earnings. It might also be possible to drill down into discipline, even institution/discipline specific returns, as seems to be the current flavour of the month. It is clear, however, that such an enhanced Study would find the Understanding Survey data inadequate for this purpose, on the bases of the high level of attrition from the data, the dropping in and out of certain questions between waves and, critically, the inadequate number of participants. Despite the policy of lifelong learning and 50% graduate targets, relatively few people actually add to their personal qualifications' portfolio year to year. Drilling down further than the broad qualifications' categories utilised in this Study would require much more data than the Survey contains. One solution to this would be the harnessing of various sources of ultra-large and detailed *administrative data* that are rapidly emerging as rich sources for researchers. With millions of individuals observed, rather than the tens of thousands available via this Survey, such detail may be possible.

References used in Chapter Three

Angrist, J.D. and Pischke, J.S., (2009) 'Mostly harmless Econometrics: an empiricist's companion', Princeton University Press, Princeton, NJ, p269

Becker, G (1965) 'Theory of the allocation of time', Economic Journal, 75(299) 493-517

Becketti, S., Gould, W., Lillard, L. and Welch, F. (1988), 'The Panel Study of Income Dynamics after fourteen years: an Evaluation', J. Labor Economics, 6, 472-92

Bolton, P, (2012) 'Education: historical statistics', House of Commons Library, SN/SG/4252, November

Blundell, R., Dearden, L., Goodman, A. and Reed, H. (1997), 'Higher Education, Employment and Earnings in Britain', London: Institute for Fiscal Studies.

Bryson, A., (2002) 'The Union Membership Wage Premium: An Analysis Using Propensity Score Matching', CEPR, London, ISBN 0 7530 1551 X

Budria, S. and Swedberg, P., (2014) 'Host Language proficiency and earnings in Spain: an Instrumental Variable Quantile Regression approach', Working Paper

Canay, I.A., (2011) 'A simple approach to quantile regression for Panel Data', Econometrics Journal, 14, 368-386

References used in Chapter Three (continued)

Carnoy, M., Loyalka, P., Androushchak, G., and Proudnikova, A. (2012) 'The economic returns to higher education in the BRIC countries and their implications for higher education expansion', Stanford University, Graduate School of Education.

Chevalier, Arnaud, (2011), 'Subject choice and earnings of UK graduates', Economics of Education Review, Elsevier, vol. 30(6), pages 1187-1201.

Davies, P., Mangan, J., Hughes, A. and Slack, K. (2013), 'Labour sector motivation and students' choice of degree subject', British Educational Research Journal.32, 2.pp.361-382

Dean, A.J.H., (1976), 'Unemployment among School Leavers: an analysis of the Problem', National Institute of Economic and Social Research Review, 11, 1

Dearden, L., McIntosh, S., Myck, M. and Vignoles, A. (2002) 'The Returns to Academic, Vocational and Basic Skills in Britain', Bulletin of Economic Research, 54, 3, 249-274.

Denman, J and McDonald, P, (1996), 'Unemployment Statistics from 1881 to the present day', Labour Market Trends, January, Labour Market Statistics Group, Central Statistical Office

Department for Education and Skills (DfES), (2007), 'Gender and education: the evidence on pupils in England', ISBN: 978-1-84478-931-3

Department for Education and Skills (DfES), (2016) 'Participation Rates in Higher Education (2006-2015)', SFR45/2016, 15th September

Firpo, S. Fortin, N. and Lemieux, T., (2009), 'Unconditional Quantile Regressions', Econometrica, 77, 953-973

References used in Chapter Three (continued)

Fournier, J-M. and Koske, I. (2012), 'Less Income Inequality and more Growth – are they compatible?', Part 7. The drivers of Labour Earnings Inequality – an Analysis based on Conditional and Unconditional Quantile Regressions, OECD Economics Department Working Papers, 930, OECD Publishing

Freier, R., Schumann, M., and Siedler, T. (2015). 'The earnings returns to graduating with honors—Evidence from law graduates', Labour Economics, 34, 39-50.

Gasteen, A., Houston, J. (2007).'Employability and Earnings Returns to Qualifications in Scotland,' Regional Studies, Taylor and Francis Journals, vol. 41(4), 443-452

Green, F. and Henseke, G (2017) 'Graduates and 'graduate jobs' in Europe: a picture of growth and diversification', UCL Working paper no. 25

Hillman, N. and Robinson, N., (2016) 'Boys to Men: The underachievement of young men in higher education – and how to start tackling it', Higher Education Policy Institute, Report 84, May, ISBN: 978-1-908240-14-9

HMRC 2016, 'Personal Income Statistics', March, HMSO

Kelly, E., O'Connell, P. J., and Smyth, E. (2010) 'The economic returns to field of study and competencies among higher education graduates in Ireland', Economics of Education Review,29, 650–657

Koenker, R. and Bassett, G. (1978), 'Regression Quantiles', Econometrica, 46, 33-50

References used in Chapter Three (continued)

Koenker, R, (2004), 'Quantile Regression for Longitudinal Data', J.Multivariate Analysis, 91, 74-89

Lalley, C, Houston, J. and Gasteen, A. (2018), 'Gender disparities in horizontal mismatch penalties: an examination of professional degrees in the UK (2007-2015)', Studies in Higher Education, forthcoming.

Levy, H. and Jenkins, S.P. (2012), 'Documentation for derived current and Annual Net Household Income Variables, BHPS waves 1-18', UK Data Archive Study Number 3909, Institute for Economic and Social Research, University of Essex

Lindley, J. and McIntosh, S. (2015). 'Growth in within graduate wage inequality: The role of subjects, cognitive skill dispersion and occupational concentration', Labour Economics, 37, 101-111

Mincer, J. (1958), 'Investment in Human Capital and the personal income distribution', J.Political Economy, 66, 281-302

Mosteller, F. and Tukey, J.W. (1977), 'Data Analysis and Regression', New York, Addison-Wesley

O'Leary, N.C., and Sloane, P.J. (2005). 'The return to a university education in Great Britain', National Institute Economic Review 193.1, 75-89.

Powell, D. (2014), 'Documentation for Quantile Regression for Panel Data (qrpd), Rand Corporation

Robinson, P. (1997). 'The Myth of Parity of Esteem: Earnings and Qualifications', London

Scottish Funding Council, (2015), 'Higher Education Students and Qualifiers at Scottish Institutions', 2013-14, SFC Statistical Publication, SFC/ST/04/2015, 25 March

References used in Chapter Three (continued)

Spence, M. (1973). 'Job sector signalling.' The Quarterly Journal of Economics: 355-374.

Strauss, H., De la Maisonneuve, C. (2009) 'The wage premium on tertiary education: New estimates for 21 OECD countries.' OECD Journal: Economic Studies 2009: 183-210.

Walker, I., Zhu, Y. (2011). 'Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales.' Economics of Education Review 30.6: 1177-1186.

World Bank, (2013), World Development Indicators 2013, Washington DC, October

Appendix to Chapter Three: detailed Regression Tables

Table A.3.1

Heckman OLS Regression : Sample Selection corrected : Males (Figure 3.20)

```
Heckman selection model                          Number of obs     =     60,811
(regression model with sample selection)               Selected    =     33,423
                                                       Nonselected =     27,388

                                                 Wald chi2(38)     =   16786.72
Log likelihood = -65967.79                       Prob > chi2       =     0.0000
```

ln_wage_rate_all	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ln_wage_rate_all						
qual_degree	.4223865	.0125482	33.66	0.000	.3977924	.4469806
qual_other_post_school	.2313696	.0129501	17.87	0.000	.2059879	.2567512
qual_school_leaving	.1895514	.0121712	15.57	0.000	.1656962	.2134065
qual_school_intermed	.082403	.0121018	6.81	0.000	.0586839	.1061221
qual_other	.0269557	.013133	2.05	0.040	.0012154	.052696
lm_experience	.0255707	.0007588	33.70	0.000	.0240836	.0270578
lm_experience2	-.0003759	.0000172	-21.86	0.000	-.0004096	-.0003422
partnered	.0645601	.0045155	14.30	0.000	.0557098	.0734104
ind_agric_for_fish	-.0007106	.0300563	-0.02	0.981	-.0596198	.0581987
ind_min_quarry	.3751028	.0299883	12.51	0.000	.3163268	.4338788
ind_manufacting	.1187592	.0149056	7.97	0.000	.0895448	.1479737
ind_power_gen_supply	.2243611	.0241553	9.29	0.000	.1770176	.2717046
ind_water_sewer_waste	.0244214	.022893	1.07	0.286	-.0204479	.0692908
ind_construct	.1841235	.0158638	11.61	0.000	.153031	.2152161
ind_whole_retail	-.0732459	.0150964	-4.85	0.000	-.1028343	-.0436575
ind_trans_storage	.0828146	.0155506	5.33	0.000	.052336	.1132933
ind_accom_foodser	-.1953838	.0171886	-11.37	0.000	-.2290729	-.1616947
ind_info_comm	.2765095	.0162256	17.04	0.000	.2447079	.3083112

Table A.3.1 (continued)

Heckman OLS Regression : Sample Selection corrected : Males (Figure 3.20)

ind_fin_ins	.3464685	.0167495	20.69	0.000	.3136401	.3792968
ind_realest	.0363783	.0236862	1.54	0.125	-.0100458	.0828024
ind_prof_sci_tech	.2419165	.0165374	14.63	0.000	.2095039	.2743291
ind_admin_supp	-.0581425	.0164605	-3.53	0.000	-.0904045	-.0258805
ind_pub_admin_defence	.1280227	.0156652	8.17	0.000	.0973194	.158726
ind_education	.0200303	.0159788	1.25	0.210	-.0112876	.0513482
ind_health_socwork	-.0063515	.015734	-0.40	0.686	-.0371895	.0244866
ind_arts_ent	-.0615095	.0199927	-3.08	0.002	-.1006944	-.0223246
ind_other_serv	0	(omitted)				
scotland	-.1170527	.008426	-13.89	0.000	-.1335673	-.100538
wales	-.1501494	.0101665	-14.77	0.000	-.1700754	-.1302235
nireland	-.1639229	.0095959	-17.08	0.000	-.1827305	-.1451153
restengland	-.1067583	.0047176	-22.63	0.000	-.1160046	-.097512
j2has	-.0393536	.0078763	-5.00	0.000	-.0547909	-.0239162
public	.0329533	.0065371	5.04	0.000	.0201407	.0457658
wave1	-.1139715	.0070016	-16.28	0.000	-.1276944	-.1002486
wave2	-.1096868	.0080224	-13.67	0.000	-.1254104	-.0939633
wave3	-.0813105	.0074474	-10.92	0.000	-.0959071	-.0667138
wave4	-.0632397	.0075378	-8.39	0.000	-.0780134	-.048466
wave5	-.0500091	.0076799	-6.51	0.000	-.0650615	-.0349567
wave6	-.0204028	.0078578	-2.60	0.009	-.0358037	-.0050018
_cons	2.069904	.0229777	90.08	0.000	2.024868	2.114939
select						
age	.0958136	.0030552	31.36	0.000	.0898255	.1018016
age2	-.0012708	.0000374	-33.94	0.000	-.0013442	-.0011974
scotland	.2640153	.0208809	12.64	0.000	.2230894	.3049412
wales	.0971721	.0242145	4.01	0.000	.0497125	.1446317
nireland	.1132031	.0231135	4.90	0.000	.0679014	.1585047
restengland	.1010563	.0110214	9.17	0.000	.0794548	.1226578
qual_degree	.7933465	.0218748	36.27	0.000	.7504727	.8362204
qual_other_post_school	.7121584	.0244174	29.17	0.000	.6643011	.7600157

Table A.3.1 (continued)

Heckman OLS Regression : Sample Selection corrected : Males (Figure 3.20)

qual_school_leaving	.6353468	.0222689	28.53	0.000	.5917006	.678993
qual_school_intermed	.5883824	.022319	26.36	0.000	.5446381	.6321268
qual_other	.434163	.0250934	17.30	0.000	.3849808	.4833452
partnered	.1012617	.0109992	9.21	0.000	.0797036	.1228197
own_house	.3100286	.0107364	28.88	0.000	.2889857	.3310715
kids	-.0865681	.0164658	-5.26	0.000	-.1188405	-.0542958
number_pensioners	-.2706954	.0156703	-17.27	0.000	-.3014086	-.2399821
wave1	.0366965	.0165508	2.22	0.027	.0042575	.0691356
wave2	.0853023	.019255	4.43	0.000	.0475632	.1230415
wave3	.0316903	.0177289	1.79	0.074	-.0030578	.0664383
wave4	.0498248	.0180404	2.76	0.006	.0144663	.0851833
wave5	.0549172	.0184449	2.98	0.003	.0187658	.0910685
wave6	.0351898	.0188529	1.87	0.062	-.0017613	.0721408
_cons	-2.494704	.0625049	-39.91	0.000	-2.617212	-2.372197
/athrho	-.3821539	.020795	-18.38	0.000	-.4229114	-.3413964
/lnsigma	-.9216848	.0056265	-163.81	0.000	-.9327126	-.910657
rho	-.3645766	.018031			-.3993803	-.3287235
sigma	.3978482	.0022385			.3934849	.4022599
lambda	-.1450461	.0078234			-.1603797	-.1297125

LR test of indep. eqns. (rho = 0): chi2(1) = 215.52 Prob > chi2 = 0.0000

Table A3.2 **Male Quantiles with sample selection correction (Figures 3.20 and 3.24)**												
	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
Degree	0.278***	(0.010)	0.390***	(0.011)	0.433***	(0.012)	0.465***	(0.010)	0.530***	(0.012)	0.574***	(0.013)
FE	0.172***	(0.011)	0.253***	(0.011)	0.276***	(0.013)	0.299***	(0.011)	0.344***	(0.012)	0.374***	(0.013)
School Leaving	0.154***	(0.009)	0.207***	(0.009)	0.228***	(0.012)	0.247***	(0.009)	0.283***	(0.011)	0.304***	(0.013)
School intermed	0.091***	(0.009)	0.114***	(0.009)	0.123***	(0.011)	0.133***	(0.008)	0.164***	(0.011)	0.179***	(0.012)
Other quals	0.058***	(0.009)	0.071***	(0.011)	0.079***	(0.012)	0.087***	(0.010)	0.113***	(0.013)	0.119***	(0.014)
LM_experience	0.020***	(0.001)	0.026***	(0.001)	0.028***	(0.001)	0.030***	(0.001)	0.033***	(0.001)	0.034***	(0.001)
lm_experience^2	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.038***	(0.004)	0.056***	(0.005)	0.064***	(0.005)	0.067***	(0.005)	0.075***	(0.005)	0.081***	(0.005)
agric_for_fish	0.100***	(0.018)	0.069	(0.036)	0.083**	(0.027)	0.081**	(0.030)	0.061*	(0.028)	0.055	(0.031)
min_quarry	0.310***	(0.041)	0.351***	(0.023)	0.356***	(0.042)	0.370***	(0.037)	0.352***	(0.035)	0.370***	(0.060)
manufacting	0.159***	(0.011)	0.177***	(0.017)	0.184***	(0.017)	0.199***	(0.011)	0.197***	(0.023)	0.191***	(0.018)
power_gen_supply	0.282***	(0.039)	0.268***	(0.029)	0.279***	(0.029)	0.291***	(0.028)	0.294***	(0.034)	0.272***	(0.028)
water_sewer_waste	0.120***	(0.019)	0.084***	(0.021)	0.086**	(0.029)	0.100***	(0.021)	0.077**	(0.028)	0.042*	(0.020)
construct	0.237***	(0.012)	0.264***	(0.019)	0.266***	(0.019)	0.280***	(0.012)	0.268***	(0.024)	0.251***	(0.018)
whole_retail	0.025**	(0.008)	-0.010	(0.016)	-0.017	(0.017)	-0.010	(0.011)	-0.017	(0.023)	-0.024	(0.018)
trans_storage	0.134***	(0.011)	0.131***	(0.017)	0.131***	(0.018)	0.136***	(0.011)	0.128***	(0.024)	0.109***	(0.018)
accom_foodser	-0.039***	(0.008)	-0.090***	(0.016)	-0.104***	(0.017)	-0.108***	(0.011)	-0.140***	(0.023)	-0.182***	(0.019)
info_comm	0.293***	(0.017)	0.332***	(0.018)	0.326***	(0.019)	0.344***	(0.017)	0.344***	(0.025)	0.338***	(0.020)
fin_ins	0.249***	(0.016)	0.279***	(0.022)	0.299***	(0.022)	0.327***	(0.016)	0.349***	(0.027)	0.356***	(0.023)
realest	0.105***	(0.010)	0.112**	(0.035)	0.106***	(0.031)	0.110***	(0.021)	0.121***	(0.034)	0.084**	(0.029)
prof_sci_tech	0.255***	(0.017)	0.267***	(0.023)	0.280***	(0.021)	0.298***	(0.015)	0.294***	(0.025)	0.279***	(0.019)
admin_supp	0.039***	(0.010)	-0.002	(0.016)	-0.010	(0.018)	-0.006	(0.012)	-0.023	(0.024)	-0.039	(0.020)

Table A3.2 (continued)
Male Quantiles with sample selection correction (Figures 3.20 and 3.24)

	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
pub_admin_defence	0.184***	(0.013)	0.197***	(0.018)	0.201***	(0.019)	0.222***	(0.013)	0.216***	(0.023)	0.192***	(0.018)
education	0.112***	(0.013)	0.097***	(0.018)	0.089***	(0.019)	0.103***	(0.013)	0.089***	(0.024)	0.055**	(0.018)
health_socwork	0.036***	(0.010)	0.010	(0.018)	0.014	(0.018)	0.026	(0.014)	0.036	(0.024)	0.037	(0.019)
arts_ent	0.025*	(0.011)	-0.000	(0.017)	-0.011	(0.023)	0.003	(0.019)	-0.003	(0.028)	-0.022	(0.020)
Scotland	-0.033***	(0.008)	-0.040***	(0.008)	-0.044***	(0.009)	-0.047***	(0.009)	-0.053***	(0.011)	-0.065***	(0.009)
Wales	-0.068***	(0.006)	-0.076***	(0.007)	-0.081***	(0.009)	-0.085***	(0.009)	-0.110***	(0.011)	-0.125***	(0.012)
NIreland	-0.046***	(0.007)	-0.092***	(0.009)	-0.103***	(0.008)	-0.119***	(0.009)	-0.144***	(0.008)	-0.166***	(0.010)
RestEngland	-0.041***	(0.005)	-0.053***	(0.005)	-0.059***	(0.005)	-0.062***	(0.005)	-0.072***	(0.006)	-0.084***	(0.006)
j2has	-0.018**	(0.006)	-0.029***	(0.005)	-0.039***	(0.008)	-0.040***	(0.009)	-0.042***	(0.007)	-0.045***	(0.009)
public	0.073***	(0.006)	0.074***	(0.007)	0.071***	(0.007)	0.068***	(0.007)	0.060***	(0.008)	0.052***	(0.007)
invmills_male	**-0.002**	**(0.002)**	**0.002**	**(0.002)**	**0.003*** **	**(0.001)**	**0.003***	**(0.001)**	**0.003***	**(0.002)**	**0.003*** **	**(0.000)**
wave1	-0.134***	(0.005)	-0.126***	(0.007)	-0.116***	(0.008)	-0.114***	(0.008)	-0.112***	(0.008)	-0.099***	(0.008)
wave2	-0.111***	(0.006)	-0.100***	(0.009)	-0.095***	(0.008)	-0.093***	(0.009)	-0.090***	(0.009)	-0.078***	(0.010)
wave3	-0.076***	(0.005)	-0.070***	(0.007)	-0.069***	(0.006)	-0.073***	(0.008)	-0.072***	(0.008)	-0.065***	(0.009)
wave4	-0.058***	(0.005)	-0.058***	(0.006)	-0.052***	(0.008)	-0.051***	(0.008)	-0.054***	(0.008)	-0.048***	(0.008)
wave5	-0.044***	(0.005)	-0.039***	(0.006)	-0.040***	(0.007)	-0.044***	(0.008)	-0.048***	(0.008)	-0.036***	(0.009)
wave6	-0.017***	(0.005)	-0.018*	(0.007)	-0.016*	(0.007)	-0.017*	(0.008)	-0.025**	(0.008)	-0.016	(0.009)
Constant	1.582***	(0.015)	1.575***	(0.020)	1.579***	(0.021)	1.576***	(0.016)	1.606***	(0.027)	1.656***	(0.023)

Table A3.2 (continued) Male Quantiles with sample selection correction (Figures 3.20 and 3.24)										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	s	b	s	b	s
Degree	0.610***	(0.015)	0.653***	(0.019)	0.649***	(0.013)	0.672***	(0.021)	0.698***	(0.021)
FE	0.392***	(0.015)	0.414***	(0.020)	0.409***	(0.014)	0.425***	(0.021)	0.408***	(0.021)
School Leaving	0.325***	(0.015)	0.350***	(0.019)	0.341***	(0.012)	0.359***	(0.020)	0.377***	(0.020)
School intermed	0.194***	(0.015)	0.216***	(0.019)	0.205***	(0.013)	0.227***	(0.020)	0.235***	(0.020)
Other quals	0.127***	(0.016)	0.139***	(0.020)	0.130***	(0.014)	0.134***	(0.022)	0.133***	(0.023)
LM_experience	0.036***	(0.001)	0.036***	(0.001)	0.036***	(0.001)	0.036***	(0.001)	0.039***	(0.002)
lm_experience^2	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.084***	(0.005)	0.095***	(0.006)	0.100***	(0.006)	0.104***	(0.007)	0.105***	(0.008)
agric_for_fish	0.044	(0.030)	-0.019	(0.045)	-0.019	(0.039)	-0.018	(0.049)	-0.064**	(0.022)
min_quarry	0.411***	(0.042)	0.386***	(0.054)	0.399***	(0.041)	0.410***	(0.055)	0.451***	(0.109)
manufacting	0.192***	(0.020)	0.130***	(0.026)	0.123***	(0.015)	0.119***	(0.027)	0.055**	(0.019)
power_gen_supply	0.285***	(0.026)	0.225***	(0.038)	0.234***	(0.026)	0.229***	(0.040)	0.176***	(0.023)
water_sewer_waste	0.037	(0.031)	-0.010	(0.034)	-0.002	(0.037)	-0.005	(0.031)	0.004	(0.068)
construct	0.247***	(0.021)	0.180***	(0.027)	0.169***	(0.016)	0.161***	(0.028)	0.085***	(0.022)
whole_retail	-0.019	(0.021)	-0.059*	(0.026)	-0.059***	(0.016)	-0.050	(0.027)	-0.099***	(0.020)
trans_storage	0.110***	(0.021)	0.073**	(0.027)	0.073***	(0.016)	0.085**	(0.029)	0.107***	(0.023)
accom_foodser	-0.187***	(0.023)	-0.235***	(0.029)	-0.242***	(0.020)	-0.237***	(0.029)	-0.273***	(0.031)
info_comm	0.336***	(0.022)	0.302***	(0.028)	0.299***	(0.019)	0.312***	(0.029)	0.236***	(0.018)
fin_ins	0.405***	(0.026)	0.403***	(0.032)	0.434***	(0.027)	0.480***	(0.031)	0.496***	(0.029)
realest	0.117**	(0.039)	0.079*	(0.036)	0.063**	(0.021)	0.027	(0.029)	-0.052	(0.050)
prof_sci_tech	0.287***	(0.023)	0.241***	(0.030)	0.256***	(0.021)	0.269***	(0.030)	0.271***	(0.032)
admin_supp	-0.008	(0.024)	-0.051	(0.029)	-0.033	(0.020)	-0.020	(0.031)	-0.043	(0.029)

Table A3.2 (continued) Male Quantiles with sample selection correction (Figures 3.20 and 3.24)										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	s	b	s	b	s
pub_admin_defence	0.189***	(0.021)	0.119***	(0.026)	0.118***	(0.016)	0.111***	(0.027)	0.046*	(0.019)
education	0.059**	(0.021)	-0.014	(0.027)	-0.016	(0.016)	-0.008	(0.029)	-0.039	(0.024)
health_socwork	0.053*	(0.022)	0.016	(0.027)	0.023	(0.017)	0.017	(0.028)	0.027	(0.022)
arts_ent	-0.030	(0.023)	-0.089**	(0.034)	-0.088***	(0.020)	-0.086**	(0.031)	-0.133***	(0.036)
Scotland	-0.071***	(0.009)	-0.097***	(0.010)	-0.104***	(0.011)	-0.114***	(0.012)	-0.128***	(0.015)
Wales	-0.136***	(0.010)	-0.159***	(0.012)	-0.167***	(0.012)	-0.181***	(0.015)	-0.215***	(0.014)
NIreland	-0.175***	(0.012)	-0.190***	(0.013)	-0.197***	(0.013)	-0.210***	(0.014)	-0.262***	(0.014)
RestEngland	-0.094***	(0.006)	-0.108***	(0.007)	-0.107***	(0.007)	-0.115***	(0.007)	-0.138***	(0.010)
j2has	-0.047***	(0.011)	-0.036**	(0.011)	-0.030*	(0.012)	-0.018	(0.014)	0.006	(0.018)
public	0.034***	(0.008)	0.025**	(0.009)	0.010	(0.008)	0.006	(0.010)	-0.034**	(0.012)
invmills_male	**0.003**	**(0.002)**	**0.002**	**(0.002)**	**0.002**	**(0.002)**	**0.001**	**(0.001)**	**0.000**	**(0.006)**
wave1	-0.106***	(0.008)	-0.103***	(0.010)	-0.106***	(0.010)	-0.105***	(0.010)	-0.106***	(0.012)
wave2	-0.083***	(0.009)	-0.088***	(0.011)	-0.090***	(0.011)	-0.083***	(0.012)	-0.092***	(0.013)
wave3	-0.067***	(0.009)	-0.071***	(0.011)	-0.072***	(0.010)	-0.074***	(0.010)	-0.085***	(0.015)
wave4	-0.054***	(0.009)	-0.051***	(0.011)	-0.048***	(0.010)	-0.045***	(0.011)	-0.059***	(0.013)
wave5	-0.041***	(0.009)	-0.040***	(0.011)	-0.037***	(0.011)	-0.041***	(0.011)	-0.046**	(0.014)
wave6	-0.017	(0.009)	-0.018	(0.010)	-0.021	(0.011)	-0.017	(0.011)	-0.019	(0.013)
Constant	1.705***	(0.027)	1.823***	(0.035)	1.887***	(0.023)	1.925***	(0.035)	2.134***	(0.034)

Table A3.3

Heckman OLS Regression : Sample Selection corrected : Females (Figure 3.21)

```
Heckman selection model                         Number of obs     =     78,196
(regression model with sample selection)              Selected    =     41,771
                                                      Nonselected =     36,425

                                                Wald chi2(38)     =   19735.05
Log likelihood = -74526.96                      Prob > chi2       =     0.0000

-----------------------------------------------------------------------------------------
      ln_wage_rate_all |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------------+------------------------------------------------------------------
ln_wage_rate_all      |
           qual_degree |   .3986613   .0103281    38.60   0.000     .3784186    .4189039
qual_other_post_school |   .2093014   .0103579    20.21   0.000     .1890003    .2296026
   qual_school_leaving |   .0900357    .010003     9.00   0.000     .0704302    .1096413
  qual_school_intermed |   .0307915   .0096886     3.18   0.001     .0118022    .0497808
            qual_other |   .0005296   .0108978     0.05   0.961    -.0208296    .0218888
         lm_experience |   .0221127   .0005849    37.81   0.000     .0209664    .0232591
        lm_experience2 |  -.0003969    .000013   -30.54   0.000    -.0004224   -.0003715
             partnered |    .038198    .003404    11.22   0.000     .0315263    .0448696
     ind_agric_for_fish |   .0608465   .0382752     1.59   0.112    -.0141716    .1358646
         ind_min_quarry |   .2169891   .0396999     5.47   0.000     .1391787    .2947995
        ind_manufacting |   .1591762   .0127634    12.47   0.000     .1341603     .184192
    ind_power_gen_supply |   .1893291   .0250573     7.56   0.000     .1402178    .2384405
   ind_water_sewer_waste |   .1635641   .0307891     5.31   0.000     .1032186    .2239097
          ind_construct |   .1927791   .0171916    11.21   0.000     .1590841    .2264741
       ind_whole_retail |  -.0249524   .0114231    -2.18   0.029    -.0473413   -.0025636
      ind_trans_storage |    .131985   .0147513     8.95   0.000      .103073     .160897
      ind_accom_foodser |  -.0916782   .0130286    -7.04   0.000    -.1172138   -.0661426
          ind_info_comm |    .325352   .0155484    20.93   0.000     .2948777    .3558263
            ind_fin_ins |   .2478723   .0133243    18.60   0.000     .2217571    .2739875
            ind_realest |   .1266785   .0178536     7.10   0.000     .0916861     .161671
      ind_prof_sci_tech |   .2068717   .0126632    16.34   0.000     .1820522    .2316911
```

Table A3.3 (continued)

Heckman OLS Regression : Sample Selection corrected : Females (Figure 3.21)

ind_admin_supp	.0349036	.013189	2.65	0.008	.0090536	.0607536
ind_pub_admin_defence	.0974572	.0120386	8.10	0.000	.073862	.1210524
ind_education	-.0249774	.011434	-2.18	0.029	-.0473877	-.0025671
ind_health_socwork	.0394248	.0110093	3.58	0.000	.017847	.0610026
ind_arts_ent	-.0128598	.0154511	-0.83	0.405	-.0431434	.0174238
ind_other_serv	0	(omitted)				
scotland	-.0964109	.0065461	-14.73	0.000	-.109241	-.0835809
wales	-.108933	.0077271	-14.10	0.000	-.1240779	-.093788
nireland	-.0882578	.0075136	-11.75	0.000	-.102984	-.0735315
restengland	-.1139341	.0038328	-29.73	0.000	-.1214462	-.106422
j2has	-.0277989	.0059447	-4.68	0.000	-.0394504	-.0161474
public	.1161572	.0044161	26.30	0.000	.1075018	.1248126
wave1	-.1175899	.0057888	-20.31	0.000	-.1289358	-.1062441
wave2	-.1056915	.0065523	-16.13	0.000	-.1185338	-.0928493
wave3	-.0756432	.0061051	-12.39	0.000	-.0876091	-.0636774
wave4	-.0578894	.0061702	-9.38	0.000	-.0699828	-.045796
wave5	-.0392277	.0062909	-6.24	0.000	-.0515577	-.0268977
wave6	-.0263884	.0064752	-4.08	0.000	-.0390795	-.0136973
_cons	1.985235	.0180381	110.06	0.000	1.949881	2.020589
select						
age	.1260566	.0030558	41.25	0.000	.1200673	.1320458
age2	-.0015806	.0000381	-41.43	0.000	-.0016554	-.0015059
scotland	.1828943	.0178852	10.23	0.000	.1478398	.2179487
wales	.1704759	.0211052	8.08	0.000	.1291105	.2118413
nireland	.2620273	.0209079	12.53	0.000	.2210485	.3030061
restengland	.1423647	.0099646	14.29	0.000	.1228344	.161895
qual_degree	.965078	.0188218	51.27	0.000	.928188	1.001968
qual_other_post_school	.9278636	.020055	46.27	0.000	.8885564	.9671708
qual_school_leaving	.8604549	.0192418	44.72	0.000	.8227416	.8981682
qual_school_intermed	.697168	.0187707	37.14	0.000	.660378	.733958
qual_other	.4880165	.0224878	21.70	0.000	.4439411	.5320918
partnered	-.072613	.0097215	-7.47	0.000	-.0916669	-.0535591

Table A3.3 (continued)

Heckman OLS Regression : Sample Selection corrected : Females (Figure 3.21)

own_house	.4330391	.0097877	44.24	0.000	.4138556	.4522225
kids	-.2305096	.0046156	-49.94	0.000	-.239556	-.2214631
number_pensioners	-.249851	.0141307	-17.68	0.000	-.2775466	-.2221554
wave1	.0081216	.0154624	0.53	0.599	-.0221843	.0384274
wave2	.0740638	.0177707	4.17	0.000	.039234	.1088937
wave3	.02332	.0164511	1.42	0.156	-.0089237	.0555636
wave4	.0499067	.0167474	2.98	0.003	.0170825	.0827309
wave5	.0484714	.0171117	2.83	0.005	.0149332	.0820096
wave6	.0091856	.0175299	0.52	0.600	-.0251723	.0435436
_cons	-3.232649	.0600815	-53.80	0.000	-3.350406	-3.114891
/athrho	-.2904697	.0164523	-17.66	0.000	-.3227157	-.2582238
/lnsigma	-1.063265	.0042445	-250.50	0.000	-1.071584	-1.054945
rho	-.2825671	.0151387			-.3119604	-.2526335
sigma	.3453266	.0014657			.3424657	.3482114
lambda	**-.0975779**	**.0055024**			**-.1083624**	**-.0867935**

LR test of indep. eqns. (rho = 0): chi2(1) = 233.50 Prob > chi2 = 0.0000

Table A3.4 Female Quantiles with sample selection correction (used in Figures 3.21 and 3.25)												
	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
qual_degree	0.180***	(0.005)	0.283***	(0.005)	0.324***	(0.006)	0.359***	(0.006)	0.431***	(0.006)	0.497***	(0.006)
qual_other_post_school	0.078***	(0.005)	0.127***	(0.005)	0.154***	(0.006)	0.174***	(0.006)	0.227***	(0.007)	0.278***	(0.007)
qual_school_leaving	0.031***	(0.003)	0.067***	(0.003)	0.082***	(0.005)	0.093***	(0.005)	0.126***	(0.005)	0.153***	(0.005)
qual_school_intermed	0.010**	(0.003)	0.033***	(0.003)	0.042***	(0.004)	0.049***	(0.005)	0.067***	(0.005)	0.085***	(0.005)
qual_other	-0.003	(0.004)	0.010*	(0.004)	0.017***	(0.005)	0.023***	(0.006)	0.033***	(0.006)	0.048***	(0.006)
LM_experience	0.013***	(0.000)	0.015***	(0.000)	0.015***	(0.000)	0.016***	(0.000)	0.018***	(0.000)	0.021***	(0.000)
lm_experience^2	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)
Partnered	0.015***	(0.003)	0.020***	(0.003)	0.022***	(0.004)	0.025***	(0.004)	0.025***	(0.004)	0.028***	(0.004)
ind_agric_for_fish	0.011	(0.037)	0.007	(0.042)	0.048*	(0.022)	0.064**	(0.023)	0.043	(0.031)	0.054	(0.050)
ind_min_quarry	0.062	(0.066)	0.065	(0.080)	0.146*	(0.061)	0.155***	(0.023)	0.129*	(0.063)	0.183***	(0.032)
ind_manufacting	0.022*	(0.009)	0.054***	(0.011)	0.081***	(0.009)	0.105***	(0.012)	0.140***	(0.014)	0.159***	(0.016)
ind_power_gen_supply	0.102***	(0.020)	0.144***	(0.032)	0.175***	(0.021)	0.177***	(0.028)	0.241***	(0.034)	0.230***	(0.024)
ind_water_sewer_waste	0.115***	(0.025)	0.123***	(0.010)	0.135***	(0.038)	0.140***	(0.016)	0.204***	(0.045)	0.222***	(0.053)
ind_construct	0.106***	(0.018)	0.151***	(0.011)	0.162***	(0.016)	0.194***	(0.017)	0.194***	(0.016)	0.199***	(0.016)
ind_whole_retail	-0.034***	(0.008)	-0.035***	(0.008)	-0.026***	(0.006)	-0.017*	(0.008)	-0.008	(0.012)	-0.009	(0.014)
ind_trans_storage	0.001	(0.012)	0.052***	(0.013)	0.076***	(0.011)	0.100***	(0.019)	0.142***	(0.019)	0.152***	(0.018)
ind_accom_foodser	-0.083***	(0.009)	-0.084***	(0.009)	-0.079***	(0.006)	-0.069***	(0.009)	-0.066***	(0.013)	-0.064***	(0.015)
ind_info_comm	0.136***	(0.018)	0.191***	(0.016)	0.226***	(0.013)	0.248***	(0.020)	0.284***	(0.019)	0.323***	(0.024)
ind_fin_ins	0.091***	(0.011)	0.116***	(0.012)	0.139***	(0.012)	0.171***	(0.013)	0.223***	(0.018)	0.240***	(0.017)
ind_realest	0.043*	(0.018)	0.073***	(0.016)	0.084***	(0.017)	0.094***	(0.017)	0.113***	(0.016)	0.119***	(0.028)
ind_prof_sci_tech	0.121***	(0.011)	0.154***	(0.010)	0.168***	(0.008)	0.181***	(0.010)	0.205***	(0.013)	0.213***	(0.015)
ind_admin_supp	-0.029***	(0.008)	-0.021*	(0.009)	-0.010	(0.009)	-0.001	(0.010)	0.016	(0.014)	0.022	(0.016)

Table A3.4 (continued) Female Quantiles with attrition & sample selection corrections (used in Figures 3.21 and 3.25)												
	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	
ind_pub_admin_defence	0.047***	(0.011)	0.069***	(0.011)	0.084***	(0.009)	0.099***	(0.011)	0.113***	(0.014)	0.117***	(0.016)
ind_education	-0.055***	(0.008)	-0.048***	(0.009)	-0.039***	(0.007)	-0.028**	(0.008)	-0.026*	(0.012)	-0.025	(0.014)
ind_health_socwork	-0.005	(0.009)	0.008	(0.009)	0.022**	(0.007)	0.035***	(0.009)	0.047***	(0.013)	0.045**	(0.014)
ind_arts_ent	-0.026**	(0.008)	-0.021	(0.011)	-0.002	(0.009)	-0.001	(0.007)	-0.006	(0.012)	-0.018	(0.016)
Scotland	-0.024***	(0.003)	-0.035***	(0.006)	-0.039***	(0.006)	-0.041***	(0.006)	-0.056***	(0.005)	-0.076***	(0.006)
Wales	-0.049***	(0.004)	-0.061***	(0.005)	-0.064***	(0.007)	-0.073***	(0.006)	-0.083***	(0.008)	-0.092***	(0.008)
NIreland	-0.034***	(0.006)	-0.036***	(0.006)	-0.039***	(0.006)	-0.048***	(0.007)	-0.057***	(0.007)	-0.064***	(0.009)
RestEngland	-0.042***	(0.002)	-0.057***	(0.003)	-0.061***	(0.004)	-0.071***	(0.004)	-0.084***	(0.004)	-0.094***	(0.004)
j2has	-0.025***	(0.002)	-0.038***	(0.003)	-0.041***	(0.005)	-0.040***	(0.005)	-0.041***	(0.005)	-0.037***	(0.006)
public	0.088***	(0.003)	0.114***	(0.004)	0.123***	(0.004)	0.129***	(0.004)	0.140***	(0.004)	0.146***	(0.005)
invmills_female	**-1.805*****	**(0.231)**	**-1.296*****	**(0.306)**	**-0.829****	**(0.284)**	**-0.758****	**(0.294)**	**-0.196**	**(0.353)**	**0.217**	**(0.356)**
wave1	-0.148***	(0.002)	-0.142***	(0.005)	-0.138***	(0.004)	-0.135***	(0.005)	-0.128***	(0.005)	-0.120***	(0.006)
wave2	-0.131***	(0.004)	-0.121***	(0.005)	-0.115***	(0.005)	-0.113***	(0.005)	-0.106***	(0.005)	-0.102***	(0.007)
wave3	-0.099***	(0.003)	-0.090***	(0.004)	-0.091***	(0.004)	-0.085***	(0.005)	-0.078***	(0.005)	-0.074***	(0.006)
wave4	-0.071***	(0.003)	-0.065***	(0.004)	-0.065***	(0.004)	-0.062***	(0.005)	-0.059***	(0.005)	-0.057***	(0.006)
wave5	-0.051***	(0.003)	-0.046***	(0.004)	-0.047***	(0.004)	-0.043***	(0.005)	-0.039***	(0.005)	-0.037***	(0.006)
wave6	-0.021***	(0.002)	-0.022***	(0.005)	-0.025***	(0.004)	-0.023***	(0.005)	-0.026***	(0.005)	-0.026***	(0.006)
Constant	1.825***	(0.012)	1.811***	(0.014)	1.800***	(0.012)	1.799***	(0.015)	1.786***	(0.018)	1.788***	(0.020)

Table A3.4 (continued) Female Quantiles with attrition & sample selection corrections (used in Figures 3.21 and 3.25)										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	s	b	s	b	s
qual_degree	0.549***	(0.007)	0.610***	(0.009)	0.640***	(0.009)	0.657***	(0.014)	0.661***	(0.018)
qual_other_post_school	0.334***	(0.008)	0.394***	(0.009)	0.421***	(0.010)	0.437***	(0.015)	0.422***	(0.018)
qual_school_leaving	0.177***	(0.007)	0.214***	(0.009)	0.238***	(0.009)	0.248***	(0.014)	0.242***	(0.018)
qual_school_intermed	0.100***	(0.007)	0.127***	(0.008)	0.142***	(0.009)	0.145***	(0.014)	0.113***	(0.018)
qual_other	0.049***	(0.008)	0.071***	(0.009)	0.080***	(0.011)	0.096***	(0.016)	0.065**	(0.022)
LM_experience	0.023***	(0.001)	0.026***	(0.001)	0.028***	(0.001)	0.029***	(0.001)	0.032***	(0.001)
$lm\_experience^2$	-0.000***	(0.000)	-0.000***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.034***	(0.005)	0.036***	(0.005)	0.040***	(0.006)	0.040***	(0.006)	0.041***	(0.008)
ind_agric_for_fish	0.046	(0.030)	0.098	(0.056)	0.069	(0.053)	0.079	(0.044)	0.084***	(0.017)
ind_min_quarry	0.272***	(0.081)	0.316***	(0.085)	0.327***	(0.052)	0.363***	(0.068)	0.679***	(0.091)
ind_manufacting	0.183***	(0.015)	0.208***	(0.020)	0.221***	(0.021)	0.226***	(0.021)	0.292***	(0.022)
ind_power_gen_supply	0.213***	(0.029)	0.241***	(0.027)	0.242***	(0.037)	0.255***	(0.032)	0.293***	(0.035)
ind_water_sewer_waste	0.257***	(0.040)	0.233***	(0.014)	0.215***	(0.048)	0.221***	(0.036)	0.135**	(0.050)
ind_construct	0.207***	(0.018)	0.223***	(0.029)	0.230***	(0.024)	0.218***	(0.027)	0.245***	(0.034)
ind_whole_retail	-0.012	(0.013)	-0.005	(0.016)	-0.002	(0.018)	0.008	(0.019)	0.051**	(0.019)
ind_trans_storage	0.169***	(0.019)	0.194***	(0.023)	0.201***	(0.027)	0.217***	(0.025)	0.283***	(0.028)
ind_accom_foodser	-0.062***	(0.014)	-0.062***	(0.017)	-0.068***	(0.019)	-0.078***	(0.019)	-0.052*	(0.022)
ind_info_comm	0.358***	(0.021)	0.407***	(0.024)	0.412***	(0.025)	0.418***	(0.021)	0.440***	(0.037)
ind_fin_ins	0.276***	(0.017)	0.326***	(0.023)	0.343***	(0.021)	0.361***	(0.026)	0.483***	(0.028)
ind_realest	0.159***	(0.021)	0.176***	(0.028)	0.186***	(0.022)	0.178***	(0.028)	0.217***	(0.045)
ind_prof_sci_tech	0.235***	(0.014)	0.243***	(0.016)	0.249***	(0.018)	0.254***	(0.019)	0.287***	(0.020)
ind_admin_supp	0.039*	(0.016)	0.054**	(0.018)	0.065**	(0.021)	0.072**	(0.022)	0.155***	(0.032)

Table A3.4 (continued) Female Quantiles with attrition & sample selection corrections (used in Figures 3.21 and 3.25)										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	b	s	b	s	b
ind_pub_admin_defence	0.123***	(0.015)	0.144***	(0.019)	0.145***	(0.020)	0.149***	(0.021)	0.200***	(0.021)
ind_education	-0.023	(0.012)	-0.015	(0.015)	-0.010	(0.017)	-0.004	(0.017)	0.042*	(0.017)
ind_health_socwork	0.051***	(0.013)	0.060***	(0.017)	0.060***	(0.018)	0.057**	(0.019)	0.098***	(0.018)
ind_arts_ent	-0.011	(0.014)	-0.004	(0.021)	-0.001	(0.021)	0.003	(0.020)	0.051	(0.028)
Scotland	-0.088***	(0.007)	-0.099***	(0.008)	-0.108***	(0.008)	-0.122***	(0.010)	-0.157***	(0.013)
Wales	-0.097***	(0.009)	-0.108***	(0.008)	-0.115***	(0.011)	-0.113***	(0.012)	-0.146***	(0.014)
NIreland	-0.070***	(0.009)	-0.078***	(0.009)	-0.086***	(0.010)	-0.101***	(0.010)	-0.129***	(0.012)
RestEngland	-0.102***	(0.005)	-0.111***	(0.005)	-0.115***	(0.006)	-0.124***	(0.006)	-0.140***	(0.008)
j2has	-0.033***	(0.008)	-0.022**	(0.008)	-0.017	(0.009)	-0.009	(0.009)	0.025	(0.014)
public	0.142***	(0.005)	0.131***	(0.005)	0.125***	(0.006)	0.117***	(0.007)	0.081***	(0.009)
invmills_female	**0.421**	**(0.382)**	**0.943***	**(0.459)**	**1.060****	**(0.410)**	**1.214***	**(0.503)**	**2.897*****	**(0.645)**
wave1	-0.121***	(0.006)	-0.116***	(0.007)	-0.116***	(0.007)	-0.115***	(0.009)	-0.104***	(0.012)
wave2	-0.100***	(0.008)	-0.098***	(0.007)	-0.100***	(0.008)	-0.092***	(0.010)	-0.087***	(0.013)
wave3	-0.074***	(0.007)	-0.074***	(0.007)	-0.069***	(0.008)	-0.066***	(0.009)	-0.068***	(0.012)
wave4	-0.058***	(0.007)	-0.059***	(0.007)	-0.063***	(0.008)	-0.064***	(0.009)	-0.051***	(0.012)
wave5	-0.044***	(0.007)	-0.044***	(0.008)	-0.040***	(0.008)	-0.038***	(0.009)	-0.027*	(0.012)
wave6	-0.031***	(0.007)	-0.037***	(0.008)	-0.032***	(0.008)	-0.031**	(0.010)	-0.028*	(0.013)
Constant	1.807***	(0.020)	1.810***	(0.024)	1.818***	(0.025)	1.849***	(0.029)	1.945***	(0.033)

Table A.3.5

Linear Prais-Winsten Estimator With heteroscedasticity and serial correlation corrections (Figure 3.22)

```
Prais-Winsten regression, heteroskedastic panels corrected standard errors

Group variable:   pidp                            Number of obs     =     32,167
Time variable:    wave                            Number of groups  =      9,991
Panels:           heteroskedastic (unbalanced)    Obs per group:
Autocorrelation:  common AR(1)                                  min =          1
                                                                avg =  3.2195976
                                                                max =          7
Estimated covariances      =      9991            R-squared         =     0.9416
Estimated autocorrelations =         1            Wald chi2(39)     =   10437.40
Estimated coefficients     =        40            Prob > chi2       =     0.0000

------------------------------------------------------------------------------------------
                       |           Het-corrected
       ln_wage_rate_all |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------------+------------------------------------------------------------------
           qual_degree |   .5124573   .0166636    30.75   0.000     .4797972    .5451174
qual_other_post_school |   .3175015   .0170614    18.61   0.000     .2840617    .3509412
   qual_school_leaving |   .2668545   .0163071    16.36   0.000     .2348931    .2988159
  qual_school_intermed |   .1523225   .0160862     9.47   0.000      .120794    .1838509
            qual_other |   .0789861   .0173969     4.54   0.000     .0448888    .1130834
         lm_experience |   .0322351   .0011385    28.31   0.000     .0300037    .0344666
        lm_experience2 |  -.0005352   .0000246   -21.76   0.000    -.0005833    -.000487
             partnered |   .0761845   .0057937    13.15   0.000      .064829    .0875399
     ind_agric_for_fish |  -.3856973   .0515268    -7.49   0.000    -.4866879   -.2847067
         ind_min_quarry |          0  (omitted)
       ind_manufacting |   -.249404   .0392891    -6.35   0.000    -.3264093   -.1723988
  ind_power_gen_supply |  -.1478077   .0458014    -3.23   0.001    -.2375768   -.0580385
 ind_water_sewer_waste |  -.3358704   .0469051    -7.16   0.000    -.4278027    -.243938
         ind_construct |  -.1871461   .0398232    -4.70   0.000    -.2651981    -.109094
      ind_whole_retail |  -.4402195   .0396042   -11.12   0.000    -.5178423   -.3625967
      ind_trans_storage |   -.293926   .0400022    -7.35   0.000    -.3723289   -.2155232
```

Table A3.5 (continued)

Linear Prais-Winsten Estimator With heteroscedasticity and serial correlation corrections (Figure 3.22)

ind_accom_foodser	-.5855846	.0405544	-14.44	0.000	-.6650698	-.5060994
ind_info_comm	-.1060577	.0406831	-2.61	0.009	-.1857951	-.0263204
ind_fin_ins	-.0265758	.0415381	-0.64	0.522	-.1079891	.0548374
ind_realest	-.3258764	.0452248	-7.21	0.000	-.4145154	-.2372374
ind_prof_sci_tech	-.130243	.0409367	-3.18	0.001	-.2104774	-.0500086
ind_admin_supp	-.4238933	.0403494	-10.51	0.000	-.5029766	-.34481
ind_pub_admin_defence	-.2376165	.0402021	-5.91	0.000	-.3164111	-.1588219
ind_education	-.3486783	.0403867	-8.63	0.000	-.4278347	-.2695218
ind_health_socwork	-.3821719	.0401226	-9.53	0.000	-.4608108	-.3035331
ind_arts_ent	-.4336639	.0428986	-10.11	0.000	-.5177435	-.3495842
ind_other_serv	-.404023	.0463035	-8.73	0.000	-.4947761	-.3132698
scotland	-.0821251	.010172	-8.07	0.000	-.1020619	-.0621883
wales	-.1416501	.0124448	-11.38	0.000	-.1660415	-.1172588
nireland	-.1655096	.0141346	-11.71	0.000	-.1932128	-.1378063
restengland	-.0912953	.0065453	-13.95	0.000	-.1041238	-.0784667
j2has	-.0269653	.0092312	-2.92	0.003	-.0450581	-.0088724
public	.0331178	.0075817	4.37	0.000	.0182579	.0479778
invmills_male	**.0018557**	**.0009894**	**1.88**	**0.061**	**-.0000835**	**.0037949**
wave1	-.1073214	.0086279	-12.44	0.000	-.1242318	-.090411
wave2	-.0944333	.0090315	-10.46	0.000	-.1121347	-.0767319
wave3	-.0744815	.0086374	-8.62	0.000	-.0914106	-.0575524
wave4	-.0588472	.0084718	-6.95	0.000	-.0754517	-.0422427
wave5	-.0453689	.0081182	-5.59	0.000	-.0612804	-.0294575
wave6	-.0174122	.0070256	-2.48	0.013	-.0311822	-.0036423
_cons	2.180462	.0447444	48.73	0.000	2.092765	2.26816
rho	.4120736					

Table A.3.6 Male Quantiles Panel with attrition correction and sample selection corrections (Figures 3.22 and 3.24)												
	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
qual_degree	0.257***	(0.002)	0.377***	(0.000)	0.410***	(0.001)	0.453***	(0.003)	0.500***	(0.001)	0.553***	(0.007)
qual_other_post_school	0.152***	(0.002)	0.239***	(0.001)	0.254***	(0.001)	0.285***	(0.002)	0.316***	(0.001)	0.360***	(0.006)
qual_school_leaving	0.142***	(0.002)	0.200***	(0.001)	0.209***	(0.001)	0.241***	(0.002)	0.264***	(0.001)	0.309***	(0.013)
qual_school_intermed	0.080***	(0.002)	0.102***	(0.001)	0.108***	(0.001)	0.121***	(0.002)	0.143***	(0.001)	0.177***	(0.010)
qual_other	0.046***	(0.003)	0.060***	(0.001)	0.060***	(0.001)	0.081***	(0.003)	0.092***	(0.001)	0.117***	(0.011)
LM_experience	0.018***	(0.000)	0.025***	(0.000)	0.028***	(0.000)	0.030***	(0.000)	0.032***	(0.000)	0.033***	(0.000)
$lm\_experience^2$	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.035***	(0.001)	0.054***	(0.000)	0.057***	(0.000)	0.053***	(0.003)	0.068***	(0.000)	0.079***	(0.003)
ind_agric_for_fish	0.108***	(0.005)	0.076***	(0.004)	0.084***	(0.002)	0.093***	(0.005)	0.056***	(0.002)	-0.040	(0.040)
ind_min_quarry	0.340***	(0.002)	0.410***	(0.004)	0.397***	(0.004)	0.404***	(0.006)	0.368***	(0.004)	0.368***	(0.039)
ind_manufacting	0.178***	(0.001)	0.202***	(0.001)	0.206***	(0.002)	0.221***	(0.002)	0.210***	(0.002)	0.170***	(0.016)
ind_power_gen_supply	0.283***	(0.001)	0.287***	(0.001)	0.298***	(0.003)	0.313***	(0.002)	0.313***	(0.002)	0.205***	(0.051)
ind_water_sewer_waste	0.123***	(0.004)	0.102***	(0.002)	0.110***	(0.002)	0.083***	(0.010)	0.083***	(0.002)	0.013	(0.016)
ind_construct	0.255***	(0.002)	0.293***	(0.001)	0.289***	(0.003)	0.306***	(0.002)	0.285***	(0.001)	0.234***	(0.014)
ind_whole_retail	0.030***	(0.001)	0.011***	(0.001)	-0.006*	(0.002)	0.001	(0.003)	-0.018***	(0.002)	-0.065**	(0.020)
ind_trans_storage	0.148***	(0.002)	0.158***	(0.001)	0.151***	(0.002)	0.146***	(0.007)	0.141***	(0.003)	0.078***	(0.020)
ind_accom_foodser	-0.032***	(0.002)	-0.069***	(0.001)	-0.090***	(0.002)	-0.087***	(0.002)	-0.137***	(0.001)	-0.189***	(0.004)
ind_info_comm	0.294***	(0.002)	0.334***	(0.002)	0.333***	(0.002)	0.337***	(0.010)	0.348***	(0.002)	0.300***	(0.022)
ind_fin_ins	0.263***	(0.001)	0.309***	(0.002)	0.316***	(0.002)	0.339***	(0.002)	0.356***	(0.003)	0.350***	(0.006)
ind_realest	0.127***	(0.004)	0.137***	(0.001)	0.137***	(0.004)	0.146***	(0.002)	0.130***	(0.002)	0.031	(0.028)
ind_prof_sci_tech	0.272***	(0.003)	0.295***	(0.002)	0.305***	(0.001)	0.315***	(0.001)	0.301***	(0.002)	0.246***	(0.017)
ind_admin_supp	0.044***	(0.001)	0.016***	(0.001)	-0.005**	(0.002)	0.019**	(0.007)	-0.023***	(0.002)	-0.077***	(0.017)

Table A.3.6 (continued) **Male Quantiles Panel with attrition correction and sample selection corrections (Figures 3.22 and 3.24)**												
	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
ind_pub_admin_defence	0.205***	(0.001)	0.233***	(0.001)	0.231***	(0.001)	0.262***	(0.004)	0.233***	(0.002)	0.185***	(0.011)
ind_education	0.123***	(0.004)	0.104***	(0.001)	0.091***	(0.002)	0.123***	(0.006)	0.092***	(0.001)	0.039***	(0.011)
ind_health_socwork	0.053***	(0.001)	0.033***	(0.002)	0.025***	(0.002)	0.046***	(0.001)	0.045***	(0.002)	0.033***	(0.002)
ind_arts_ent	0.041***	(0.001)	0.010***	(0.002)	-0.002	(0.002)	0.044***	(0.010)	0.007***	(0.002)	-0.049*	(0.019)
Scotland	-0.032***	(0.003)	-0.050***	(0.001)	-0.042***	(0.001)	-0.052***	(0.001)	-0.051***	(0.000)	-0.059***	(0.005)
Wales	-0.061***	(0.002)	-0.080***	(0.001)	-0.081***	(0.002)	-0.080***	(0.005)	-0.107***	(0.001)	-0.135***	(0.001)
NIreland	-0.048***	(0.001)	-0.082***	(0.001)	-0.095***	(0.002)	-0.119***	(0.002)	-0.153***	(0.000)	-0.173***	(0.003)
RestEngland	-0.040***	(0.001)	-0.056***	(0.001)	-0.060***	(0.002)	-0.063***	(0.001)	-0.076***	(0.000)	-0.090***	(0.001)
j2has	-0.014***	(0.002)	-0.034***	(0.001)	-0.053***	(0.002)	-0.063***	(0.004)	-0.049***	(0.000)	-0.048***	(0.001)
public	0.073***	(0.001)	0.067***	(0.001)	0.071***	(0.001)	0.059***	(0.005)	0.061***	(0.001)	0.046***	(0.004)
wave1	-0.137***	(0.001)	-0.136***	(0.001)	-0.122***	(0.002)	-0.117***	(0.005)	-0.113***	(0.001)	-0.107***	(0.004)
wave2	-0.117***	(0.001)	-0.109***	(0.001)	-0.097***	(0.002)	-0.104***	(0.001)	-0.093***	(0.001)	-0.084***	(0.001)
wave3	-0.079***	(0.001)	-0.083***	(0.002)	-0.075***	(0.001)	-0.080***	(0.003)	-0.074***	(0.000)	-0.064***	(0.003)
wave4	-0.058***	(0.002)	-0.058***	(0.001)	-0.055***	(0.001)	-0.050***	(0.006)	-0.056***	(0.001)	-0.056***	(0.001)
wave5	-0.043***	(0.002)	-0.043***	(0.000)	-0.044***	(0.002)	-0.052***	(0.002)	-0.049***	(0.001)	-0.038***	(0.001)
wave6	-0.012***	(0.001)	-0.014***	(0.001)	-0.009***	(0.002)	-0.020***	(0.003)	-0.020***	(0.001)	-0.015***	(0.001)
invmills_male	**-0.002*** **	**(0.000)**	**0.004*** **	**(0.000)**	**0.004*** **	**(0.000)**	**0.004*** **	**(0.000)**	**0.003*** **	**(0.000)**	**0.000**	**(0.001)**

Table A.3.6 (continued) Male Quantiles Panel with attrition correction and sample selection corrections (Figures 3.22 and 3.24)										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	s	b	s	b	s
qual_degree	0.579***	(0.003)	0.619***	(0.002)	0.627***	(0.002)	0.642***	(0.004)	0.671***	(0.002)
qual_other_post_school	0.365***	(0.002)	0.389***	(0.003)	0.396***	(0.002)	0.399***	(0.005)	0.393***	(0.002)
qual_school_leaving	0.301***	(0.003)	0.325***	(0.002)	0.328***	(0.002)	0.340***	(0.004)	0.356***	(0.002)
qual_school_intermed	0.177***	(0.002)	0.199***	(0.002)	0.200***	(0.001)	0.210***	(0.004)	0.214***	(0.002)
qual_other	0.104***	(0.003)	0.106***	(0.002)	0.110***	(0.002)	0.106***	(0.005)	0.107***	(0.003)
LM_experience	0.035***	(0.000)	0.035***	(0.000)	0.034***	(0.000)	0.035***	(0.000)	0.038***	(0.000)
$lm\_experience^2$	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.080***	(0.000)	0.086***	(0.000)	0.093***	(0.000)	0.095***	(0.001)	0.101***	(0.001)
ind_agric_for_fish	0.012***	(0.003)	-0.046***	(0.003)	-0.000	(0.002)	-0.038***	(0.004)	-0.071***	(0.006)
ind_min_quarry	0.445***	(0.005)	0.413***	(0.002)	0.430***	(0.002)	0.465***	(0.005)	0.467***	(0.004)
ind_manufacting	0.190***	(0.003)	0.129***	(0.001)	0.136***	(0.001)	0.128***	(0.005)	0.067***	(0.003)
ind_power_gen_supply	0.306***	(0.007)	0.244***	(0.001)	0.237***	(0.002)	0.245***	(0.004)	0.183***	(0.005)
ind_water_sewer_waste	0.035***	(0.006)	-0.000	(0.001)	0.011***	(0.002)	0.002	(0.002)	0.036***	(0.003)
ind_construct	0.253***	(0.002)	0.173***	(0.001)	0.176***	(0.001)	0.172***	(0.004)	0.087***	(0.004)
ind_whole_retail	-0.034***	(0.002)	-0.082***	(0.001)	-0.064***	(0.002)	-0.060***	(0.005)	-0.096***	(0.003)
ind_trans_storage	0.107***	(0.003)	0.070***	(0.001)	0.078***	(0.001)	0.083***	(0.005)	0.122***	(0.004)
ind_accom_foodser	-0.198***	(0.002)	-0.246***	(0.001)	-0.240***	(0.001)	-0.240***	(0.004)	-0.315***	(0.003)
ind_info_comm	0.328***	(0.002)	0.298***	(0.001)	0.302***	(0.001)	0.312***	(0.005)	0.238***	(0.003)
ind_fin_ins	0.405***	(0.003)	0.404***	(0.001)	0.454***	(0.002)	0.489***	(0.005)	0.504***	(0.003)
ind_realest	0.098***	(0.004)	0.059***	(0.001)	0.071***	(0.001)	0.031***	(0.004)	-0.012*	(0.005)
ind_prof_sci_tech	0.285***	(0.002)	0.244***	(0.002)	0.267***	(0.002)	0.287***	(0.004)	0.281***	(0.003)
ind_admin_supp	-0.029***	(0.005)	-0.070***	(0.001)	-0.050***	(0.003)	-0.042***	(0.004)	-0.050***	(0.002)

Table A.3.6 (continued) Male Quantiles Panel with attrition correction and sample selection corrections (Figures 3.22 and 3.24)										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	b	s	b	s	b
ind_pub_admin_defence	0.192***	(0.002)	0.117***	(0.001)	0.127***	(0.002)	0.123***	(0.005)	0.062***	(0.003)
ind_education	0.058***	(0.003)	-0.014***	(0.001)	-0.007***	(0.001)	0.008	(0.006)	-0.021***	(0.003)
ind_health_socwork	0.051***	(0.003)	0.012***	(0.001)	0.027***	(0.002)	0.019***	(0.005)	0.040***	(0.003)
ind_arts_ent	-0.024***	(0.005)	-0.089***	(0.001)	-0.069***	(0.001)	-0.071***	(0.003)	-0.132***	(0.005)
Scotland	-0.064***	(0.001)	-0.090***	(0.001)	-0.103***	(0.001)	-0.113***	(0.000)	-0.135***	(0.001)
Wales	-0.132***	(0.002)	-0.155***	(0.001)	-0.160***	(0.001)	-0.177***	(0.001)	-0.210***	(0.002)
NIreland	-0.184***	(0.001)	-0.196***	(0.001)	-0.203***	(0.001)	-0.211***	(0.002)	-0.266***	(0.001)
RestEngland	-0.093***	(0.000)	-0.106***	(0.000)	-0.109***	(0.000)	-0.112***	(0.001)	-0.134***	(0.001)
j2has	-0.057***	(0.001)	-0.048***	(0.000)	-0.030***	(0.001)	-0.026***	(0.001)	0.007***	(0.001)
public	0.030***	(0.001)	0.026***	(0.001)	0.015***	(0.000)	0.004***	(0.001)	-0.041***	(0.001)
wave1	-0.114***	(0.001)	-0.104***	(0.000)	-0.109***	(0.001)	-0.100***	(0.000)	-0.107***	(0.002)
wave2	-0.093***	(0.001)	-0.091***	(0.001)	-0.099***	(0.001)	-0.081***	(0.001)	-0.094***	(0.001)
wave3	-0.075***	(0.001)	-0.079***	(0.001)	-0.077***	(0.001)	-0.069***	(0.001)	-0.085***	(0.001)
wave4	-0.058***	(0.002)	-0.050***	(0.000)	-0.052***	(0.001)	-0.034***	(0.001)	-0.058***	(0.001)
wave5	-0.046***	(0.001)	-0.041***	(0.000)	-0.037***	(0.001)	-0.032***	(0.001)	-0.051***	(0.002)
wave6	-0.017***	(0.000)	-0.017***	(0.001)	-0.022***	(0.000)	-0.010***	(0.001)	-0.021***	(0.002)
invmills_male	**0.002*** **	**(0.000)**	**0.002*** **	**(0.000)**	**0.002*** **	**(0.000)**	**0.002*** **	**(0.000)**	**0.000*** **	**(0.000)**

Table A3.7

Linear Prais-Winsten Estimator With heteroscedasticity and serial correlation corrections Females Figure 3.23

```
Prais-Winsten regression, heteroskedastic panels corrected standard errors

Group variable:   pidp                          Number of obs     =     41,771
Time variable:    wave                          Number of groups  =     13,154
Panels:           heteroskedastic (unbalanced)  Obs per group:
Autocorrelation:  common AR(1)                                min =          1
                                                              avg =   3.175536
                                                              max =          7

Estimated covariances      =     13154          R-squared         =     0.9446
Estimated autocorrelations =         1          Wald chi2(39)     =   16537.40
Estimated coefficients     =        40          Prob > chi2       =     0.0000

------------------------------------------------------------------------------------------
                       |            Het-corrected
      ln_wage_rate_all |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------------+------------------------------------------------------------------
           qual_degree |   .4689612   .0109948    42.65   0.000     .4474118    .4905106
qual_other_post_school |   .2708063   .0111036    24.39   0.000     .2490437    .2925689
   qual_school_leaving |   .1537362   .0106575    14.43   0.000      .132848    .1746244
  qual_school_intermed |   .0804663   .0102995     7.81   0.000     .0602796    .1006529
            qual_other |   .0351354   .0116595     3.01   0.003     .0122833    .0579875
         lm_experience |   .0236227   .0006618    35.69   0.000     .0223256    .0249198
        lm_experience2 |  -.0004237   .0000148   -28.65   0.000    -.0004526   -.0003947
             partnered |   .0355305   .0048781     7.28   0.000     .0259696    .0450915
     ind_agric_for_fish |  -.1646627   .0742636    -2.22   0.027    -.3102166   -.0191088
        ind_min_quarry |          0  (omitted)
       ind_manufacting |  -.0766291   .0591573    -1.30   0.195    -.1925752     .039317
  ind_power_gen_supply |  -.0388112   .0626663    -0.62   0.536    -.1616348    .0840124
 ind_water_sewer_waste |  -.0612276   .0670268    -0.91   0.361    -.1925977    .0701425
         ind_construct |  -.0342136   .0608813    -0.56   0.574    -.1535387    .0851115
      ind_whole_retail |  -.2548355   .0586864    -4.34   0.000    -.3698587   -.1398123
     ind_trans_storage |  -.0935566   .0600332    -1.56   0.119    -.2112194    .0241063
```

Table A3.7 (continued)

Linear Prais-Winsten Estimator With heteroscedasticity and serial correlation corrections Females Figure 3.23

ind_accom_foodser	-.3202139	.058996	-5.43	0.000	-.4358439	-.204584
ind_info_comm	.0867198	.0612761	1.42	0.157	-.0333791	.2068188
ind_fin_ins	.0167815	.0596319	0.28	0.778	-.1000949	.1336578
ind_realest	-.0959254	.0609181	-1.57	0.115	-.2153226	.0234719
ind_prof_sci_tech	-.0158521	.0596496	-0.27	0.790	-.1327631	.1010589
ind_admin_supp	-.1937555	.0593613	-3.26	0.001	-.3101016	-.0774095
ind_pub_admin_defence	-.1225897	.0589676	-2.08	0.038	-.238164	-.0070154
ind_education	-.2419684	.0589327	-4.11	0.000	-.3574744	-.1264623
ind_health_socwork	-.1878125	.0586617	-3.20	0.001	-.3027874	-.0728376
ind_arts_ent	-.2408557	.059808	-4.03	0.000	-.3580772	-.1236341
ind_other_serv	-.2256841	.0602784	-3.74	0.000	-.3438275	-.1075406
scotland	-.07923	.0076206	-10.40	0.000	-.094166	-.064294
wales	-.0993163	.0097535	-10.18	0.000	-.1184329	-.0801998
nireland	-.0706676	.0100311	-7.04	0.000	-.0903283	-.0510069
restengland	-.1022012	.0050508	-20.23	0.000	-.1121005	-.0923018
j2has	-.0171919	.0067338	-2.55	0.011	-.0303899	-.003994
public	.1024863	.0049604	20.66	0.000	.092764	.1122086
invmills_female	**-.3058633**	**.3896331**	**-0.79**	**0.432**	**-1.06953**	**.4578035**
wave1	-.1143949	.0058403	-19.59	0.000	-.1258417	-.1029482
wave2	-.0994419	.0063585	-15.64	0.000	-.1119043	-.0869794
wave3	-.0750714	.0059982	-12.52	0.000	-.0868277	-.0633151
wave4	-.0579075	.0060348	-9.60	0.000	-.0697355	-.0460794
wave5	-.0416436	.0059535	-6.99	0.000	-.0533122	-.0299751
wave6	-.0287866	.00538	-5.35	0.000	-.0393312	-.0182421
_cons	2.060285	.0605025	34.05	0.000	1.941702	2.178867
rho	.3753144					

Table A3.8 Female Quantiles Panel with attrition and sample selection corrections Figure 3.23 and 3.25												
	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
qual_degree	0.167***	(0.001)	0.278***	(0.001)	0.319***	(0.000)	0.359***	(0.004)	0.429***	(0.000)	0.492***	(0.001)
qual_other_post_school	0.076***	(0.001)	0.127***	(0.001)	0.155***	(0.000)	0.185***	(0.005)	0.231***	(0.000)	0.281***	(0.001)
qual_school_leaving	0.034***	(0.001)	0.068***	(0.001)	0.082***	(0.000)	0.096***	(0.003)	0.125***	(0.000)	0.148***	(0.001)
qual_school_intermed	0.011***	(0.001)	0.034***	(0.001)	0.041***	(0.000)	0.052***	(0.004)	0.067***	(0.000)	0.081***	(0.001)
qual_other	-0.003***	(0.001)	0.014***	(0.002)	0.017***	(0.001)	0.021***	(0.004)	0.030***	(0.000)	0.041***	(0.001)
LM_experience	0.011***	(0.000)	0.013***	(0.000)	0.014***	(0.000)	0.016***	(0.000)	0.018***	(0.000)	0.021***	(0.000)
lm_experience^2	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)	-0.000***	(0.000)
Partnered	0.016***	(0.000)	0.021***	(0.000)	0.022***	(0.000)	0.024***	(0.001)	0.023***	(0.000)	0.026***	(0.001)
ind_agric_for_fish	0.003	(0.002)	-0.009***	(0.002)	0.040***	(0.002)	-0.012	(0.015)	0.033***	(0.002)	0.048***	(0.003)
ind_min_quarry	0.008	(0.004)	0.069***	(0.001)	0.149***	(0.001)	0.152***	(0.006)	0.124***	(0.003)	0.179***	(0.003)
ind_manufacting	0.031***	(0.001)	0.054***	(0.001)	0.089***	(0.001)	0.096***	(0.001)	0.148***	(0.001)	0.166***	(0.001)
ind_power_gen_supply	0.122***	(0.002)	0.142***	(0.001)	0.174***	(0.001)	0.209***	(0.012)	0.247***	(0.001)	0.238***	(0.001)
ind_water_sewer_waste	0.112***	(0.000)	0.125***	(0.002)	0.124***	(0.001)	0.166***	(0.016)	0.206***	(0.002)	0.213***	(0.001)
ind_construct	0.133***	(0.000)	0.154***	(0.001)	0.179***	(0.001)	0.196***	(0.005)	0.207***	(0.001)	0.203***	(0.001)
ind_whole_retail	-0.027***	(0.001)	-0.036***	(0.001)	-0.025***	(0.001)	-0.029***	(0.004)	-0.010***	(0.001)	-0.014***	(0.001)
ind_trans_storage	0.013***	(0.000)	0.051***	(0.001)	0.080***	(0.001)	0.087***	(0.003)	0.151***	(0.001)	0.153***	(0.001)
ind_accom_foodser	-0.066***	(0.001)	-0.081***	(0.001)	-0.077***	(0.001)	-0.074***	(0.007)	-0.065***	(0.001)	-0.069***	(0.001)
ind_info_comm	0.145***	(0.000)	0.198***	(0.001)	0.227***	(0.001)	0.240***	(0.002)	0.299***	(0.001)	0.354***	(0.001)
ind_fin_ins	0.090***	(0.000)	0.111***	(0.001)	0.135***	(0.001)	0.157***	(0.001)	0.220***	(0.001)	0.242***	(0.001)

Table A3.8 (continued)
Female Quantiles Panel with attrition and sample selection corrections Figure 3.23 and 3.25

	Centile 10%		Centile 20%		Centile 25%		Centile 30%		Centile 40%		Centile 50%	
	b	s	b	s	b	s	b	s	b	s	b	s
ind_realest	0.059***	(0.001)	0.082***	(0.000)	0.098***	(0.001)	0.105***	(0.001)	0.123***	(0.001)	0.136***	(0.001)
ind_prof_sci_tech	0.121***	(0.000)	0.150***	(0.001)	0.168***	(0.001)	0.158***	(0.003)	0.204***	(0.001)	0.216***	(0.001)
ind_admin_supp	-0.024***	(0.000)	-0.024***	(0.001)	-0.011***	(0.001)	-0.012**	(0.004)	0.015***	(0.001)	0.018***	(0.001)
ind_pub_admin_defence	0.052***	(0.001)	0.066***	(0.001)	0.084***	(0.001)	0.091***	(0.005)	0.110***	(0.001)	0.113***	(0.001)
ind_education	-0.054***	(0.000)	-0.051***	(0.000)	-0.041***	(0.001)	-0.042***	(0.001)	-0.030***	(0.001)	-0.031***	(0.001)
ind_health_socwork	-0.001	(0.001)	0.005***	(0.001)	0.021***	(0.001)	0.017***	(0.002)	0.044***	(0.001)	0.042***	(0.001)
ind_arts_ent	-0.012***	(0.001)	-0.012***	(0.000)	-0.003*	(0.001)	-0.006*	(0.003)	-0.003**	(0.001)	-0.007***	(0.001)
Scotland	-0.024***	(0.000)	-0.035***	(0.000)	-0.039***	(0.000)	-0.044***	(0.001)	-0.053***	(0.001)	-0.080***	(0.000)
Wales	-0.047***	(0.000)	-0.065***	(0.000)	-0.067***	(0.001)	-0.062***	(0.004)	-0.081***	(0.000)	-0.102***	(0.000)
NIreland	-0.029***	(0.000)	-0.032***	(0.001)	-0.040***	(0.000)	-0.044***	(0.001)	-0.052***	(0.000)	-0.066***	(0.000)
RestEngland	-0.041***	(0.000)	-0.058***	(0.000)	-0.065***	(0.000)	-0.066***	(0.003)	-0.085***	(0.000)	-0.099***	(0.000)
j2has	-0.025***	(0.001)	-0.040***	(0.000)	-0.046***	(0.000)	-0.040***	(0.000)	-0.045***	(0.000)	-0.041***	(0.000)
public	0.091***	(0.000)	0.113***	(0.000)	0.124***	(0.000)	0.127***	(0.001)	0.144***	(0.000)	0.147***	(0.000)
wave1	-0.154***	(0.000)	-0.140***	(0.000)	-0.136***	(0.000)	-0.129***	(0.002)	-0.122***	(0.000)	-0.109***	(0.000)
wave2	-0.134***	(0.000)	-0.118***	(0.000)	-0.111***	(0.000)	-0.102***	(0.002)	-0.100***	(0.000)	-0.095***	(0.001)
wave3	-0.101***	(0.000)	-0.089***	(0.000)	-0.086***	(0.000)	-0.081***	(0.000)	-0.075***	(0.000)	-0.067***	(0.001)
wave4	-0.070***	(0.000)	-0.064***	(0.000)	-0.060***	(0.000)	-0.050***	(0.002)	-0.052***	(0.000)	-0.049***	(0.001)
wave5	-0.047***	(0.000)	-0.044***	(0.000)	-0.044***	(0.000)	-0.036***	(0.001)	-0.038***	(0.000)	-0.032***	(0.000)
wave6	-0.018***	(0.000)	-0.018***	(0.000)	-0.021***	(0.000)	-0.014***	(0.002)	-0.023***	(0.000)	-0.022***	(0.001)
invmills_female	**-1.749*****	**(0.016)**	**-1.414*****	**(0.009)**	**-0.968*****	**(0.019)**	**-0.875*****	**(0.078)**	**-0.342*****	**(0.017)**	**-0.045***	**(0.019)**

Table A3.8 (continued)
Female Quantiles Panel with attrition and sample selection corrections Figure 3.23 and 3.25

	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	s	b	s	b	s
qual_degree	0.543***	(0.001)	0.606***	(0.001)	0.629***	(0.001)	0.590***	(0.031)	0.655***	(0.001)
qual_other_post_school	0.337***	(0.001)	0.396***	(0.001)	0.419***	(0.001)	0.389***	(0.024)	0.422***	(0.001)
qual_school_leaving	0.172***	(0.001)	0.209***	(0.001)	0.223***	(0.001)	0.185***	(0.026)	0.222***	(0.001)
qual_school_intermed	0.096***	(0.001)	0.123***	(0.001)	0.125***	(0.001)	0.082**	(0.028)	0.095***	(0.000)
qual_other	0.042***	(0.002)	0.062***	(0.002)	0.063***	(0.001)	0.047***	(0.013)	0.044***	(0.002)
LM_experience	0.023***	(0.000)	0.026***	(0.000)	0.029***	(0.000)	0.031***	(0.001)	0.033***	(0.000)
lm_experience2	-0.000***	(0.000)	-0.000***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.030***	(0.000)	0.034***	(0.000)	0.036***	(0.001)	0.041***	(0.002)	0.038***	(0.001)
ind_agric_for_fish	0.045***	(0.003)	0.105***	(0.002)	0.062***	(0.004)	-0.178	(0.134)	0.055***	(0.004)
ind_min_quarry	0.233***	(0.007)	0.267***	(0.009)	0.338***	(0.006)	0.051	(0.148)	0.571***	(0.003)
ind_manufacting	0.178***	(0.002)	0.184***	(0.002)	0.172***	(0.005)	0.097	(0.059)	0.247***	(0.002)
ind_power_gen_supply	0.240***	(0.003)	0.222***	(0.002)	0.201***	(0.005)	0.066	(0.091)	0.241***	(0.004)
ind_water_sewer_waste	0.238***	(0.001)	0.240***	(0.002)	0.165***	(0.004)	0.134***	(0.030)	0.098***	(0.003)
ind_construct	0.203***	(0.002)	0.213***	(0.002)	0.198***	(0.004)	0.039	(0.093)	0.209***	(0.002)
ind_whole_retail	-0.025***	(0.001)	-0.034***	(0.001)	-0.046***	(0.003)	-0.129*	(0.057)	-0.007***	(0.001)
ind_trans_storage	0.160***	(0.001)	0.169***	(0.002)	0.157***	(0.003)	0.071	(0.054)	0.221***	(0.002)
ind_accom_foodser	-0.077***	(0.002)	-0.098***	(0.002)	-0.115***	(0.003)	-0.211***	(0.056)	-0.104***	(0.002)
ind_info_comm	0.382***	(0.001)	0.395***	(0.001)	0.381***	(0.004)	0.286***	(0.058)	0.414***	(0.001)
ind_fin_ins	0.275***	(0.001)	0.301***	(0.002)	0.293***	(0.005)	0.189*	(0.076)	0.432***	(0.001)
ind_realest	0.167***	(0.001)	0.160***	(0.002)	0.135***	(0.007)	0.081	(0.052)	0.192***	(0.003)

Table A3.8 (continued) Female Quantiles Panel with attrition and sample selection corrections Figure 3.23 and 3.25										
	Centile 60%		Centile 70%		Centile 75%		Centile 80%		Centile 90%	
	b	s	b	s	b	s	b	s	b	s
ind_prof_sci_tech	0.233***	(0.001)	0.231***	(0.001)	0.209***	(0.004)	0.067	(0.084)	0.244***	(0.001)
ind_admin_supp	0.034***	(0.001)	0.037***	(0.001)	0.019***	(0.004)	-0.054	(0.058)	0.126***	(0.002)
ind_pub_admin_defence	0.118***	(0.001)	0.117***	(0.001)	0.102***	(0.002)	-0.005	(0.069)	0.158***	(0.002)
ind_education	-0.033***	(0.001)	-0.040***	(0.001)	-0.056***	(0.003)	-0.128*	(0.051)	-0.003*	(0.001)
ind_health_socwork	0.041***	(0.001)	0.030***	(0.001)	0.009**	(0.003)	-0.079	(0.057)	0.052***	(0.001)
ind_arts_ent	-0.016***	(0.001)	-0.027***	(0.001)	-0.047***	(0.003)	-0.046***	(0.011)	-0.022***	(0.002)
Scotland	-0.088***	(0.000)	-0.105***	(0.001)	-0.112***	(0.001)	-0.115***	(0.009)	-0.164***	(0.001)
Wales	-0.101***	(0.001)	-0.114***	(0.001)	-0.118***	(0.001)	-0.185***	(0.036)	-0.149***	(0.001)
NIreland	-0.066***	(0.001)	-0.077***	(0.001)	-0.087***	(0.002)	-0.134***	(0.020)	-0.133***	(0.001)
RestEngland	-0.103***	(0.000)	-0.115***	(0.000)	-0.116***	(0.001)	-0.153***	(0.016)	-0.140***	(0.000)
j2has	-0.038***	(0.000)	-0.029***	(0.000)	-0.016***	(0.002)	-0.030**	(0.010)	0.019***	(0.002)
public	0.141***	(0.000)	0.134***	(0.000)	0.127***	(0.001)	0.103***	(0.008)	0.079***	(0.000)
wave1	-0.109***	(0.001)	-0.109***	(0.001)	-0.100***	(0.002)	-0.084***	(0.008)	-0.096***	(0.000)
wave2	-0.094***	(0.001)	-0.093***	(0.000)	-0.081***	(0.002)	-0.043*	(0.019)	-0.074***	(0.001)
wave3	-0.069***	(0.001)	-0.065***	(0.000)	-0.063***	(0.000)	-0.054***	(0.000)	-0.062***	(0.001)
wave4	-0.052***	(0.001)	-0.052***	(0.001)	-0.050***	(0.001)	-0.035**	(0.011)	-0.052***	(0.000)
wave5	-0.040***	(0.001)	-0.046***	(0.000)	-0.039***	(0.001)	-0.004	(0.015)	-0.025***	(0.001)
wave6	-0.030***	(0.001)	-0.034***	(0.001)	-0.029***	(0.001)	-0.014	(0.009)	-0.028***	(0.001)
invmills_female	**0.339*** **	**(0.015)**	**0.301*** **	**(0.027)**	**0.605*** **	**(0.037)**	**-0.106**	**(0.635)**	**2.309*** **	**(0.038)**

Table A.3.9
Male Canay with Attrition and Sample Selection Corrections (Figure 3.27)

	Centile 10		Centile 20		Centile 25		Centile 30		Centile 40		Centile 50	
	b	s	b	s	b	s	b	s	b	s	b	s
qual_degree	0.259***	(0.009)	0.487***	(0.010)	0.496***	(0.006)	0.501***	(0.004)	0.509***	(0.006)	0.519***	(0.007)
qual_other_post_school	0.155***	(0.009)	0.317***	(0.009)	0.321***	(0.006)	0.326***	(0.005)	0.335***	(0.006)	0.344***	(0.008)
qual_school_leaving	0.144***	(0.007)	0.262***	(0.009)	0.264***	(0.005)	0.268***	(0.004)	0.273***	(0.006)	0.280***	(0.007)
qual_school_intermed	0.082***	(0.007)	0.176***	(0.009)	0.178***	(0.005)	0.179***	(0.004)	0.184***	(0.006)	0.187***	(0.007)
qual_other	0.051***	(0.007)	0.121***	(0.009)	0.119***	(0.006)	0.120***	(0.004)	0.121***	(0.006)	0.124***	(0.007)
LM_experience	0.019***	(0.001)	0.032***	(0.001)	0.033***	(0.000)	0.033***	(0.000)	0.033***	(0.000)	0.034***	(0.000)
lm_experience2	-0.000***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.036***	(0.005)	0.074***	(0.003)	0.076***	(0.002)	0.077***	(0.002)	0.079***	(0.002)	0.080***	(0.002)
ind_agric_for_fish	0.107***	(0.010)	-0.028	(0.019)	-0.032**	(0.010)	-0.038**	(0.015)	-0.037*	(0.016)	-0.037**	(0.014)
ind_min_quarry	0.340***	(0.093)	0.233***	(0.023)	0.242***	(0.022)	0.250***	(0.016)	0.267***	(0.025)	0.294***	(0.028)
ind_manufacting	0.183***	(0.009)	0.060***	(0.013)	0.064***	(0.009)	0.063***	(0.009)	0.059***	(0.012)	0.059***	(0.010)
ind_power_gen_supply	0.292***	(0.035)	0.109***	(0.020)	0.120***	(0.015)	0.126***	(0.017)	0.139***	(0.017)	0.142***	(0.014)
ind_water_sewer_waste	0.138***	(0.015)	0.003	(0.014)	0.008	(0.012)	0.010	(0.015)	0.008	(0.014)	0.004	(0.014)
ind_construct	0.262***	(0.011)	0.130***	(0.014)	0.134***	(0.009)	0.135***	(0.010)	0.133***	(0.012)	0.132***	(0.011)
ind_whole_retail	0.040***	(0.007)	-0.064***	(0.013)	-0.062***	(0.009)	-0.068***	(0.009)	-0.080***	(0.012)	-0.083***	(0.010)
ind_trans_storage	0.158***	(0.009)	0.021	(0.013)	0.026**	(0.009)	0.029**	(0.010)	0.027*	(0.012)	0.029**	(0.011)
ind_accom_foodser	-0.023**	(0.008)	-0.160***	(0.014)	-0.164***	(0.010)	-0.169***	(0.010)	-0.190***	(0.012)	-0.200***	(0.011)
ind_info_comm	0.296***	(0.017)	0.171***	(0.014)	0.173***	(0.010)	0.176***	(0.010)	0.178***	(0.013)	0.178***	(0.011)
ind_fin_ins	0.269***	(0.016)	0.210***	(0.015)	0.228***	(0.010)	0.228***	(0.011)	0.239***	(0.013)	0.241***	(0.011)
ind_realest	0.123***	(0.011)	0.018	(0.019)	0.019	(0.011)	0.015	(0.011)	0.008	(0.014)	0.005	(0.014)
ind_prof_sci_tech	0.280***	(0.019)	0.130***	(0.015)	0.135***	(0.010)	0.138***	(0.011)	0.137***	(0.013)	0.135***	(0.011)
ind_admin_supp	0.052***	(0.008)	-0.043**	(0.013)	-0.040***	(0.009)	-0.043***	(0.010)	-0.053***	(0.013)	-0.052***	(0.011)

Table A.3.9 (continued)
Male Canay with Attrition and Sample Selection Corrections (Figure 3.27)

	Centile 10		Centile 20		Centile 25		Centile 30		Centile 40		Centile 50	
	b	s	b	s	b	s	b	s	b	s	b	s
ind_pub_admin_defence	0.210***	(0.013)	0.095***	(0.013)	0.100***	(0.009)	0.103***	(0.010)	0.101***	(0.012)	0.101***	(0.010)
ind_education	0.131***	(0.012)	0.000	(0.013)	-0.001	(0.009)	0.001	(0.010)	0.000	(0.012)	-0.003	(0.011)
ind_health_socwork	0.056***	(0.009)	-0.047***	(0.013)	-0.043***	(0.009)	-0.043***	(0.009)	-0.045***	(0.012)	-0.045***	(0.010)
ind_arts_ent	0.045***	(0.008)	-0.088***	(0.017)	-0.079***	(0.013)	-0.074***	(0.012)	-0.081***	(0.014)	-0.081***	(0.012)
Scotland	-0.036**	(0.011)	-0.058***	(0.005)	-0.057***	(0.005)	-0.057***	(0.004)	-0.060***	(0.004)	-0.068***	(0.004)
Wales	-0.065***	(0.008)	-0.117***	(0.005)	-0.120***	(0.004)	-0.122***	(0.004)	-0.126***	(0.004)	-0.135***	(0.005)
NIreland	-0.045***	(0.007)	-0.131***	(0.006)	-0.130***	(0.005)	-0.133***	(0.005)	-0.138***	(0.004)	-0.146***	(0.006)
RestEngland	-0.041***	(0.005)	-0.077***	(0.003)	-0.079***	(0.002)	-0.078***	(0.002)	-0.080***	(0.002)	-0.085***	(0.003)
j2has	-0.017***	(0.004)	-0.019***	(0.005)	-0.021***	(0.004)	-0.020***	(0.004)	-0.016***	(0.004)	-0.016***	(0.004)
public	0.073***	(0.007)	0.040***	(0.004)	0.040***	(0.003)	0.037***	(0.003)	0.034***	(0.003)	0.031***	(0.003)
invmills_male	**-0.001**	**(0.002)**	**0.003****	**(0.001)**	**0.003****	**(0.001)**	**0.003****	**(0.001)**	**0.003***	**(0.001)**	**0.003****	**(0.001)**
wave1	-0.138***	(0.009)	-0.121***	(0.004)	-0.118***	(0.005)	-0.120***	(0.005)	-0.120***	(0.004)	-0.121***	(0.005)
wave2	-0.112***	(0.010)	-0.097***	(0.005)	-0.094***	(0.005)	-0.094***	(0.005)	-0.098***	(0.004)	-0.097***	(0.005)
wave3	-0.078***	(0.008)	-0.075***	(0.005)	-0.072***	(0.005)	-0.075***	(0.005)	-0.080***	(0.004)	-0.081***	(0.005)
wave4	-0.055***	(0.008)	-0.056***	(0.005)	-0.053***	(0.005)	-0.055***	(0.005)	-0.057***	(0.004)	-0.062***	(0.005)
wave5	-0.040***	(0.009)	-0.038***	(0.005)	-0.037***	(0.005)	-0.038***	(0.005)	-0.042***	(0.004)	-0.047***	(0.005)
wave6	-0.010	(0.008)	-0.016***	(0.005)	-0.015**	(0.005)	-0.018***	(0.005)	-0.018***	(0.004)	-0.020***	(0.005)
Constant	1.591***	(0.016)	1.715***	(0.018)	1.722***	(0.012)	1.736***	(0.012)	1.769***	(0.015)	1.793***	(0.015)

Table A.3.9 (continued) Male Canay with Attrition and Sample Selection Corrections (Figure 3.27)										
	Centile 60		Centile 70		Centile 75		Centile 80		Centile 90	
	b	s	b	s	b	s	b	s	b	s
qual_degree	0.517***	(0.007)	0.529***	(0.007)	0.535***	(0.007)	0.539***	(0.007)	0.550***	(0.012)
qual_other_post_school	0.339***	(0.007)	0.348***	(0.008)	0.352***	(0.007)	0.357***	(0.007)	0.372***	(0.013)
qual_school_leaving	0.277***	(0.007)	0.290***	(0.007)	0.295***	(0.007)	0.299***	(0.007)	0.312***	(0.012)
qual_school_intermed	0.179***	(0.007)	0.189***	(0.007)	0.192***	(0.007)	0.192***	(0.006)	0.197***	(0.012)
qual_other	0.116***	(0.007)	0.124***	(0.007)	0.127***	(0.007)	0.128***	(0.007)	0.143***	(0.014)
LM_experience	0.035***	(0.000)	0.035***	(0.000)	0.035***	(0.001)	0.035***	(0.001)	0.034***	(0.001)
lm_experience2	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)	-0.001***	(0.000)
Partnered	0.081***	(0.002)	0.081***	(0.002)	0.083***	(0.003)	0.084***	(0.003)	0.082***	(0.005)
ind_agric_for_fish	-0.031	(0.021)	-0.018	(0.016)	-0.013	(0.015)	-0.023	(0.020)	-0.013	(0.018)
ind_min_quarry	0.319***	(0.018)	0.346***	(0.033)	0.371***	(0.031)	0.385***	(0.029)	0.446***	(0.033)
ind_manufacting	0.059***	(0.009)	0.064***	(0.009)	0.069***	(0.012)	0.064***	(0.013)	0.068***	(0.017)
ind_power_gen_supply	0.148***	(0.012)	0.146***	(0.016)	0.158***	(0.016)	0.155***	(0.017)	0.178***	(0.038)
ind_water_sewer_waste	0.017	(0.017)	0.015	(0.010)	0.013	(0.016)	0.011	(0.016)	0.006	(0.018)
ind_construct	0.137***	(0.009)	0.135***	(0.010)	0.143***	(0.013)	0.141***	(0.013)	0.143***	(0.018)
ind_whole_retail	-0.083***	(0.009)	-0.079***	(0.009)	-0.077***	(0.012)	-0.078***	(0.014)	-0.065***	(0.017)
ind_trans_storage	0.031***	(0.009)	0.038***	(0.009)	0.043***	(0.012)	0.046***	(0.014)	0.055***	(0.017)
ind_accom_foodser	-0.201***	(0.010)	-0.204***	(0.010)	-0.202***	(0.013)	-0.212***	(0.014)	-0.214***	(0.024)
ind_info_comm	0.181***	(0.010)	0.187***	(0.010)	0.190***	(0.013)	0.187***	(0.014)	0.193***	(0.017)
ind_fin_ins	0.243***	(0.010)	0.252***	(0.010)	0.261***	(0.014)	0.259***	(0.014)	0.274***	(0.019)
ind_realest	0.017	(0.013)	0.019	(0.014)	0.026	(0.017)	0.017	(0.019)	0.032	(0.031)

Table A.3.9 (continued) Male Canay with Attrition and Sample Selection Corrections (Figure 3.27)										
	Centile 60		Centile 70		Centile 75		Centile 80		Centile 90	
	b	s	b	s	b	s	b	s	b	s
ind_prof_sci_tech	0.140***	(0.010)	0.144***	(0.010)	0.152***	(0.013)	0.146***	(0.014)	0.139***	(0.018)
ind_admin_supp	-0.052***	(0.010)	-0.047***	(0.009)	-0.044***	(0.013)	-0.047***	(0.014)	-0.038*	(0.019)
ind_pub_admin_defence	0.100***	(0.009)	0.102***	(0.009)	0.105***	(0.012)	0.098***	(0.013)	0.098***	(0.017)
ind_education	0.001	(0.009)	0.010	(0.009)	0.014	(0.012)	0.011	(0.013)	0.020	(0.017)
ind_health_socwork	-0.036***	(0.009)	-0.026**	(0.009)	-0.021	(0.012)	-0.024	(0.013)	-0.019	(0.016)
ind_arts_ent	-0.074***	(0.012)	-0.073***	(0.010)	-0.077***	(0.012)	-0.080***	(0.015)	-0.069*	(0.028)
Scotland	-0.074***	(0.004)	-0.084***	(0.004)	-0.087***	(0.005)	-0.090***	(0.005)	-0.116***	(0.008)
Wales	-0.138***	(0.004)	-0.144***	(0.005)	-0.146***	(0.006)	-0.151***	(0.006)	-0.168***	(0.011)
NIreland	-0.153***	(0.005)	-0.161***	(0.006)	-0.159***	(0.007)	-0.162***	(0.006)	-0.180***	(0.008)
RestEngland	-0.087***	(0.003)	-0.092***	(0.003)	-0.094***	(0.003)	-0.098***	(0.003)	-0.114***	(0.005)
j2has	-0.018***	(0.004)	-0.014**	(0.004)	-0.012*	(0.005)	-0.013**	(0.005)	-0.019*	(0.008)
public	0.029***	(0.003)	0.024***	(0.003)	0.022***	(0.004)	0.018***	(0.004)	0.010	(0.006)
invmills_male	**0.001**	**(0.002)**	**0.002****	**(0.001)**	**0.001**	**(0.001)**	**0.001**	**(0.002)**	**0.002**	**(0.004)**
wave1	-0.120***	(0.004)	-0.119***	(0.004)	-0.117***	(0.006)	-0.125***	(0.005)	-0.127***	(0.009)
wave2	-0.098***	(0.005)	-0.097***	(0.004)	-0.097***	(0.006)	-0.104***	(0.006)	-0.111***	(0.008)
wave3	-0.083***	(0.004)	-0.084***	(0.004)	-0.083***	(0.006)	-0.090***	(0.006)	-0.092***	(0.009)
wave4	-0.064***	(0.004)	-0.067***	(0.004)	-0.066***	(0.006)	-0.075***	(0.005)	-0.082***	(0.009)
wave5	-0.047***	(0.004)	-0.049***	(0.004)	-0.047***	(0.006)	-0.054***	(0.005)	-0.065***	(0.009)
wave6	-0.018***	(0.004)	-0.017***	(0.004)	-0.018**	(0.006)	-0.020***	(0.006)	-0.023*	(0.009)
Constant	1.830***	(0.013)	1.850***	(0.013)	1.864***	(0.016)	1.907***	(0.017)	1.997***	(0.026)

Table A.3.10 Female Canay with Attrition and Sample Selection Corrections (Figure 3.28)												
	Centile 10		Centile 20		Centile 25		Centile 30		Centile 40		Centile 50	
	b	s	b	s	b	s	b	s	b	s	b	s
qual_degree	0.394***	0.011	0.414***	0.004	0.432***	0.003	0.444***	0.004	0.465***	0.004	0.483***	0.004
qual_other_post_school	0.230***	0.011	0.243***	0.004	0.256***	0.004	0.267***	0.004	0.280***	0.004	0.293***	0.004
qual_school_leaving	0.145***	0.011	0.154***	0.004	0.166***	0.003	0.172***	0.003	0.182***	0.003	0.187***	0.004
qual_school_intermed	0.093***	0.01	0.092***	0.003	0.103***	0.003	0.108***	0.003	0.113***	0.003	0.115***	0.004
qual_other	0.064***	0.011	0.059***	0.004	0.067***	0.003	0.069***	0.004	0.071***	0.003	0.074***	0.004
LM_experience	0.022***	0.001	0.021***	0.000	0.022***	0.000	0.022***	0.000	0.023***	0.000	0.023***	0.000
lm_experience2	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000
Partnered	0.042***	0.003	0.041***	0.002	0.041***	0.002	0.042***	0.002	0.041***	0.002	0.040***	0.002
ind_agric_for_fish	-0.056	0.112	-0.007	0.013	0.009	0.007	-0.004	0.029	0.024	0.046	0.050**	0.019
ind_min_quarry	0.146***	0.034	0.167**	0.062	0.164***	0.039	0.167***	0.024	0.203***	0.016	0.212***	0.026
ind_manufacting	0.062***	0.015	0.086***	0.009	0.085***	0.007	0.087***	0.007	0.103***	0.007	0.105***	0.007
ind_power_gen_supply	0.139***	0.016	0.156***	0.027	0.162***	0.014	0.167***	0.013	0.187***	0.023	0.196***	0.010
ind_water_sewer_waste	0.145***	0.012	0.136***	0.01	0.148***	0.021	0.144***	0.012	0.143***	0.007	0.137***	0.012
ind_construct	0.096***	0.015	0.113***	0.012	0.118***	0.009	0.119***	0.008	0.127***	0.008	0.134***	0.009
ind_whole_retail	-0.028*	0.014	-0.026**	0.008	-0.030***	0.007	-0.032***	0.006	-0.026***	0.006	-0.027***	0.006
ind_trans_storage	0.061***	0.016	0.080***	0.012	0.087***	0.009	0.092***	0.008	0.106***	0.007	0.112***	0.008
ind_accom_foodser	-0.085***	0.015	-0.079***	0.009	-0.082***	0.007	-0.085***	0.006	-0.081***	0.006	-0.083***	0.007
ind_info_comm	0.194***	0.015	0.218***	0.012	0.222***	0.011	0.228***	0.009	0.244***	0.008	0.259***	0.009
ind_fin_ins	0.154***	0.015	0.173***	0.01	0.177***	0.008	0.182***	0.007	0.200***	0.007	0.211***	0.007
ind_realest	0.073***	0.017	0.087***	0.013	0.086***	0.012	0.092***	0.008	0.110***	0.010	0.114***	0.008
ind_prof_sci_tech	0.127***	0.014	0.137***	0.008	0.136***	0.007	0.142***	0.006	0.154***	0.006	0.161***	0.006

Table A.3.10 (continued)
Female Canay with Attrition and Sample Selection Corrections (Figure 3.28)

	Centile 10		Centile 20		Centile 25		Centile 30		Centile 40		Centile 50	
	b	s	b	s	b	s	b	s	b	s	b	s
ind_admin_supp	-0.008	0.016	0.014	0.01	0.013	0.007	0.011	0.006	0.020**	0.006	0.025***	0.007
ind_pub_admin_defence	0.092***	0.015	0.102***	0.009	0.100***	0.008	0.100***	0.007	0.113***	0.007	0.116***	0.007
ind_education	-0.017	0.013	-0.004	0.008	-0.007	0.007	-0.006	0.006	0.002	0.006	0.007	0.006
ind_health_socwork	0.016	0.014	0.029***	0.009	0.026***	0.007	0.026***	0.006	0.035***	0.006	0.039***	0.006
ind_arts_ent	-0.017	0.014	-0.015	0.009	-0.012	0.008	-0.017*	0.007	-0.005	0.008	-0.006	0.007
Scotland	-0.052***	0.007	-0.055***	0.004	-0.053***	0.003	-0.053***	0.004	-0.059***	0.003	-0.064***	0.003
Wales	-0.092***	0.006	-0.096***	0.004	-0.098***	0.004	-0.101***	0.004	-0.102***	0.004	-0.107***	0.003
NIreland	-0.064***	0.007	-0.059***	0.005	-0.057***	0.004	-0.056***	0.004	-0.062***	0.004	-0.067***	0.004
RestEngland	-0.080***	0.004	-0.087***	0.002	-0.087***	0.002	-0.088***	0.002	-0.094***	0.002	-0.099***	0.002
j2has	-0.033***	0.005	-0.021***	0.004	-0.022***	0.003	-0.021***	0.003	-0.020***	0.003	-0.020***	0.003
public	0.048***	0.004	0.054***	0.003	0.058***	0.002	0.061***	0.002	0.065***	0.002	0.066***	0.002
invmills_female	**-0.800****	**0.273**	**-0.564****	**0.178**	**-0.520****	**0.172**	**-0.457****	**0.152**	**-0.278**	**0.163**	**-0.15**	**0.163**
wave1	-0.123***	0.006	-0.125***	0.004	-0.127***	0.003	-0.126***	0.003	-0.125***	0.003	-0.123***	0.003
wave2	-0.096***	0.007	-0.100***	0.004	-0.103***	0.003	-0.103***	0.003	-0.100***	0.003	-0.102***	0.003
wave3	-0.074***	0.006	-0.076***	0.004	-0.078***	0.003	-0.078***	0.003	-0.077***	0.003	-0.080***	0.003
wave4	-0.054***	0.006	-0.055***	0.004	-0.057***	0.003	-0.057***	0.003	-0.058***	0.003	-0.061***	0.003
wave5	-0.040***	0.006	-0.041***	0.004	-0.042***	0.003	-0.043***	0.003	-0.042***	0.003	-0.043***	0.003
wave6	-0.029***	0.006	-0.024***	0.004	-0.023***	0.003	-0.024***	0.003	-0.023***	0.003	-0.026***	0.003
Constant	1.733***	0.021	1.774***	0.011	1.780***	0.01	1.784***	0.009	1.788***	0.009	1.803***	0.01

Table A.3.10 (continued)
Female Canay with Attrition and Sample Selection Corrections (Figure 3.28)

	Centile 60		Centile 70		Centile 75		Centile 80		Centile 90	
	b	s	b	s	b	s	b	s	b	s
qual_degree	0.501***	0.004	0.518***	0.005	0.531***	0.006	0.533***	0.009	0.531***	0.013
qual_other_post_school	0.305***	0.004	0.319***	0.005	0.328***	0.006	0.329***	0.009	0.318***	0.013
qual_school_leaving	0.198***	0.004	0.206***	0.005	0.214***	0.006	0.212***	0.009	0.208***	0.013
qual_school_intermed	0.121***	0.004	0.128***	0.004	0.135***	0.006	0.133***	0.009	0.115***	0.012
qual_other	0.076***	0.004	0.084***	0.005	0.093***	0.007	0.092***	0.009	0.076***	0.014
LM_experience	0.023***	0.000	0.024***	0.000	0.024***	0.000	0.024***	0.000	0.024***	0.001
lm_experience2	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000	-0.000***	0.000
Partnered	0.040***	0.002	0.041***	0.002	0.043***	0.003	0.042***	0.003	0.042***	0.004
ind_agric_for_fish	0.060*	0.03	0.074	0.039	0.064***	0.013	0.086*	0.035	0.126	0.083
ind_min_quarry	0.233***	0.021	0.253***	0.038	0.248***	0.025	0.256***	0.013	0.256***	0.016
ind_manufacting	0.113***	0.008	0.115***	0.01	0.109***	0.01	0.123***	0.012	0.115***	0.016
ind_power_gen_supply	0.200***	0.015	0.204***	0.012	0.199***	0.016	0.214***	0.017	0.184***	0.023
ind_water_sewer_waste	0.146***	0.015	0.135***	0.017	0.129***	0.023	0.135***	0.032	0.095***	0.012
ind_construct	0.144***	0.01	0.149***	0.011	0.152***	0.015	0.166***	0.014	0.166***	0.018
ind_whole_retail	-0.026***	0.007	-0.027**	0.009	-0.034***	0.009	-0.022*	0.011	-0.030*	0.015
ind_trans_storage	0.126***	0.009	0.123***	0.011	0.119***	0.011	0.136***	0.014	0.134***	0.019
ind_accom_foodser	-0.081***	0.008	-0.081***	0.01	-0.086***	0.01	-0.076***	0.012	-0.067***	0.018
ind_info_comm	0.266***	0.008	0.268***	0.011	0.267***	0.012	0.279***	0.012	0.288***	0.022
ind_fin_ins	0.215***	0.009	0.223***	0.01	0.220***	0.011	0.234***	0.012	0.231***	0.017
ind_realest	0.112***	0.009	0.128***	0.015	0.124***	0.012	0.132***	0.015	0.123***	0.021

Table A.3.10 (continued) Female Canay with Attrition and Sample Selection Corrections (Figure 3.28)										
	Centile 60		Centile 70		Centile 75		Centile 80		Centile 90	
	b	s	b	s	b	s	b	s	b	s
ind_prof_sci_tech	0.168***	0.007	0.172***	0.009	0.165***	0.009	0.172***	0.011	0.163***	0.015
ind_admin_supp	0.035***	0.008	0.038***	0.01	0.039**	0.012	0.055***	0.012	0.078***	0.019
ind_pub_admin_defence	0.122***	0.008	0.121***	0.01	0.118***	0.01	0.130***	0.012	0.121***	0.016
ind_education	0.014	0.007	0.018*	0.009	0.018*	0.009	0.033**	0.01	0.040**	0.014
ind_health_socwork	0.046***	0.007	0.044***	0.009	0.042***	0.01	0.054***	0.011	0.048**	0.015
ind_arts_ent	-0.005	0.009	-0.006	0.011	-0.01	0.01	0	0.015	0.011	0.017
Scotland	-0.070***	0.003	-0.074***	0.004	-0.078***	0.004	-0.087***	0.004	-0.092***	0.007
Wales	-0.111***	0.004	-0.113***	0.005	-0.111***	0.005	-0.117***	0.006	-0.115***	0.007
NIreland	-0.071***	0.004	-0.074***	0.005	-0.071***	0.005	-0.074***	0.006	-0.072***	0.01
RestEngland	-0.102***	0.002	-0.105***	0.002	-0.106***	0.003	-0.108***	0.003	-0.108***	0.004
j2has	-0.018***	0.004	-0.012***	0.004	-0.010*	0.004	-0.008	0.005	0.005	0.008
public	0.066***	0.002	0.065***	0.003	0.061***	0.003	0.060***	0.003	0.046***	0.005
invmills_female	**-0.063**	**0.165**	**-0.005**	**0.186**	**0.001**	**0.238**	**0.341**	**0.266**	**0.946***	**0.37**
wave1	-0.123***	0.003	-0.122***	0.004	-0.121***	0.004	-0.121***	0.004	-0.121***	0.007
wave2	-0.103***	0.004	-0.103***	0.004	-0.104***	0.005	-0.101***	0.005	-0.105***	0.007
wave3	-0.082***	0.003	-0.084***	0.004	-0.086***	0.004	-0.086***	0.004	-0.089***	0.007
wave4	-0.064***	0.003	-0.069***	0.004	-0.068***	0.004	-0.071***	0.005	-0.080***	0.007
wave5	-0.047***	0.003	-0.051***	0.004	-0.052***	0.004	-0.054***	0.005	-0.056***	0.007
wave6	-0.031***	0.003	-0.031***	0.004	-0.032***	0.004	-0.035***	0.005	-0.032***	0.008
Constant	1.818***	0.01	1.845***	0.013	1.861***	0.014	1.879***	0.017	1.965***	0.024

4. Assessing airline efficiency: a Stochastic Frontier Approach using Panel Data

4.1. Introduction & Aims of the Study

The international airline industry has grown to become one of the World's largest, with a combined revenue of $274 billion in 1995 to $754 billion in 2017 and directly employs in excess of two million people and many more indirectly in ground services, traffic control, aircraft design and manufacture and supply services. [49] The industry has maintained a high rate of long-term growth, though there have been setbacks resulting from Oil Crises, Terrorist Action and Economic Recessions. It has transformed itself from a highly regulated and uncompetitive one into a fully commercialised trans-national industry. The extent and the rate of change in the industry's competitive conditions make it an interesting one to study from both economic and econometric standpoints.

The aims of this Study are to:

- Assemble a Panel of airlines, containing performance data relating to their objective, namely the provision of air services, using aircraft and employees output
- To assess the efficiency of the airlines, relating their actual output, to a hypothesised maximum potential, given the inputs used and standard industry practise.
- To assess what effect on the results there are from using the data as a Panel, as opposed to a Cross-section, allowing for unobserved effects to be incorporated.

The chapter will set the scene by outlining the history of the Commercial Airline Industry since its inception one hundred years ago, focussing on the desire to protect it from full commercial pressure in order that it could both grow, and be used as a symbol of National Power and Prestige in the World, through to the

[49] https://www.statista.com/statistics/278372/revenue-of-commercial-airlines-worldwide/, accessed 11th April, 2018

current day, where regulation focusses instead on non-economic matters, in particular safety and environmental impact. After that, the methodology to be employed and a summary of relevant technical and empirical literature will be discussed, followed by a consideration of the results. The chapter concludes with some discussion and suggestions for future work in this area.

4.2 The competitive and technical conditions in the airline industry

4.2.1. An outline history of competition and regulation in the International Airline Industry

When one talks about the regulation of Air Transport, they are referring to two inter-related, but from an Economics' standpoint, very different concepts. On the one hand, one can talk about the regulation of *aviation*, insofar as it refers to:

- The design and airworthiness of aircraft;
- The fitness and competence of flight, cabin and ground crew;
- The ability of the support infrastructure (air traffic control, airport facilities and security, meteorological services, etc) to allow services to operate at the 'required' level of safety.

Doganis (1991) described such factors as constituting non-economic regulation, insofar as they do not *directly* impinge on the level of fares, service provision (routes and frequencies). Their aim is to ensure that the levels of safety that the Travelling Public requires from air transport, are met. It is certainly true to say that in most of the World, the industry has to meet much higher safety requirements than the Road and Rail transport industries. For instance, in most parts of the World it is acceptable to allow train passengers to stand on journeys, even to sit on the roof, in a few countries. It is also still acceptable for unqualified people to service and maintain motor cars, and for private car drivers to drive for as long as they want without a rest break. There are, undoubtedly *indirect* economic consequences, insofar as any regulation is costly, and would raise unit costs, and affect fares and output, but this Study will be concerned more with the direct effects of Economic Regulation.

It was only 16 years after Orville and Wright had made their first powered flight that the then economically-developed States came together in 1919 to sign what became known as the *Paris Convention*. This was the first such Treaty aimed at regulating the nascent air transport industry and established the principle of the State's sovereignty over its own airspace. Previously, only the land and adjacent sea were regarded as such. This allowed any signatory State to prevent airlines domiciled in another State to overfly its airspace. Thus, in order that an air service could operate between two countries, it was necessary to establish a *bi-lateral agreement* (a 'bi-lateral') between the governments. The process was relatively straightforward where the proposed service did not involve overflying a third country (as between Britain and France, for example). However, where part of the journey involved overflying other countries (as between Britain and Switzerland, for example), then it would be necessary for the origin and destination countries to also have separate bi-laterals with the countries being overflown. The result was a piecemeal and otherwise incomplete system of bi-laterals, occasionally disrupted by conflicts and disputes, which could make it impossible for otherwise uninvolved countries to operate their air service, at least on the shortest route possible. [50]

It wasn't until the 1944 *Chicago Convention* that there was a concerted effort to co-ordinate the development of the emerging, global airline industry. The United States, in particular, wanted to see a completely free 'open skies' system, with no restriction on fares and service levels and complete freedom to overfly any country that happened to lie on the direct route between the origin and the destination, without having to agree specific bi-laterals to do so.[51] Their European Allies, however, preferred that the State should exercise complete control over their fares,

[50] For example, the former Federal Republic of Germany never signed a full bi-lateral with its then neighbour, the German Democratic Republic ('GDR). As a result, services between the UK and GDR had to enter and leave the latter's airspace at its border with Denmark, who had signed bi-laterals with both countries.

[51] Nevertheless, the United States did not wish to extend the same freedoms within its own Domestic system, where the regulatory authority, the Civil Aeronautics Board (CAB) maintained strict control of prices, schedules and what airlines could operate routes. The CAB's power was such that no new 'trunk' operator was authorised to service a domestic route between 1938 and the late 1970s, and it even had to give its permission for an operator to withdraw from a route.

service levels and access rights: thus, the system of piecemeal bi-laterals continued. That said, the bi-laterals gradually became more standardised and started to include specific reference to the five *Freedoms of the Air*, established by the United Nations' International Civil Aviation Authority (ICAO). These are: [52]

1. The right or privilege, in respect of scheduled international air services, granted by one State to another State or States to fly across its territory without landing
2. The right or privilege, in respect of scheduled international air services, granted by one State to another State or States to land in its territory for non-traffic purposes (also known as a Second Freedom Right).
3. The right or privilege, in respect of scheduled international air services, granted by one State to another State to put down, in the territory of the first State, traffic coming from the home State of the carrier (also known as a Third Freedom Right).
4. The right or privilege, in respect of scheduled international air services, granted by one State to another State to take on, in the territory of the first State, traffic destined for the home State of the carrier (also known as a Fourth Freedom Right).
5. the right or privilege, in respect of scheduled international air services, granted by one State to another State to put down and to take on, in the territory of the first State, traffic coming from or destined to a third State (also known as a Fifth Freedom Right).

The bi-laterals that came after Chicago set fares that took into account the cost of providing the service while allowing operators to make a 'reasonable' profit. They also assumed that the service would be divided as equally as possible between the two countries irrespective of their relative size, with the precise service to be operated detailed in the bilateral, devised through what are known as 'inter-airline pooling agreements'. Crucially, the actual fares and service-levels had to be ratified

52 Source: Manual on the Regulation of International Air Transport (Doc 9626, Part 4), ICAO, https://www.icao.int/Pages/freedomsAir.aspx, last accessed 22nd April, 2018

by the host governments who designated which of their 'approved' airlines would actually be permitted to operate its part of the bilateral. Only those 'ICAO-freedoms' necessary to allow the said service to be operated were ever granted by the bi-laterals, with only a very few instances where an airline from X en-route to Y was permitted to land in the intermediate Z, to pick up passengers and/or cargo. Almost unheard of was the right for an airline from X being able to operate a service entirely between Y and Z only, not involving the territory of X in any way.

The near-obsession with international routes can be explained by the small size of the domestic industry until the late 1960s. As this sector grew, the attitude of governments was to ensure that only domestic operators provided those services. This was, in part guaranteed, outside the US, by State Ownership, nationalising, where necessary, established airlines (such as BEA in 1946), or establishing a State Monopoly provider from scratch.

Governments' ability to allocate routes and set fares following World War Two, coincided with the nationalisation of many of the World's airlines, seen as a major plank of the post-War process of reconstruction and reassertion of national identity. Only in the United States was there more than one privately-owned 'Flag Carrier' endowed with the necessary aircraft, staff and ground support capable of operating international services. This allowed these Flag Carriers to charge high fares and otherwise not 'sweat' its assets. Even in the United States, where there has never been a State-owned Airline, its Government designated both Pan-American ('Pan-Am') and Trans World Airways ('TWA') as its only bi-lateral operators. That said, a glimpse of the then future was offered via the special arrangements that existed between the United States and Britain, signing the first *Bermuda Agreement* in 1946. This, and subsequent 'Bermudas' were more liberal than standard bi-laterals, with more 'Fifth Freedom' rights and a more *laissez-faire* attitude to the service levels offered by each airline. Fares would be set by the independent International Air Transport Association (IATA). 'Bermudas' gradually became the industry-default, which also increased the co-ordination and geographical scope of air services whilst increasing the power and influence of IATA. Having achieved this position of power, IATA sought to regulate fares and service levels via a sequence of secret annual ‘conferences’, covering both domestic and international services. Although facilitating the development of an industry Cartel, the ‘conference’ also acted as a useful co-ordination of the world-wide system, and lead to the steady spread of services, taking advantage of the then-rapid improvement in aircraft technology and increased living standards throughout most of the developed World.

Even after the passing of the immediate post-War Two period, by which time most economies had recovered and started to grow again, some Economists continued to argue that the airline industry, in particular, would always require to be regulated (e.g. Wheatcroft (1964); Richmond, (1974). Their logic was that:

1. The service was inherently *homogeneous*, with low barriers to entry and few scale economies. Accordingly, there would be continuous entry and exit from the industry, with resulting price wars and disruption to services. Given tight profit margins, there would be constant downward pressure on costs, with adverse consequences for safety and reliability.
2. Air Transport played a role as a *public utility*, generating many positive externalities, to the extent that governments had cultivated their ‘Flag Carrier’. This patronage has been applied to both domestic- as well as the international services, though many governments had been happy to allow others to pick up some peripheral business (mostly low-value cargo, charters). In addition, many scheduled services were, in effect, a *public service*. The demand conditions on the ‘thin’ routes were such that operating costs could not be covered by revenue and that the resulting losses could only be plugged by State subventions and cross-subsidisation from the more profitable routes. The Scottish Islands’ routes would be one example of this, even to this day.

3. State control of the Flag Carrier meant a captive demand for its domestically-designed and manufactured aircraft (latterly, themselves nationalised until the 1980s). It was even the case, that the Flag Carrier, as lead customer had a strong, and some would say, negative influence on the design of new aircraft, being too tailored to its particular operating requirements, to the detriment of lucrative export sales.

However, from the late 1960s onwards, arguments against regulation of the Air Transport system began to gain ground. Economists like Straszheim (1969) and Eads (1975), argued that:

1. Regulated fares and service frequencies were, from the consumers' standpoint, welfare reducing.
2. The lack of competition was strangling both technical innovation (a Public Good) and the growth of the industry, with adverse implications for both Producer- and Consumer Surplus.

They acknowledged that competition would result in a flood of enthusiastic new entrants but were confident that the incumbents had the necessary financial reserves and self-belief to survive the initial impact. They would also reform their operations in an effort to achieve the necessary efficiency gains needed to compete. The resulting reductions in unit costs would be reflected in a general reduction in real fares and, as a result, consumers would benefit from those and more convenient schedules. This would, in turn, provide an incentive to the designers and manufacturers of aircraft to improve the performance and (cost) efficiency of their designs.

However, the main reason for the subsequent deregulation of the industry during the 1980s was undoubtedly the OPEC Oil Crisis that started in 1973. Accordingly, all airlines' financial performance declined quite markedly in a short space of time. Governments, anxious to protect their Flag Carriers began to loosen controls that afforded the airlines more room to exploit commercial opportunities (such as the

Holiday Charter market). They also began to pursue more vigorously, their own national interests at the IATA conferences. They also permitted off-line discounts to be offered on scheduled fares, and even transatlantic charter flights, albeit with lengthy minimum booking periods. In an effort to fill spare capacity on scheduled flights, operators began to sell blocks of seats to travel agents and 'bucket shops', who then sold them onto the Travelling Public at discounted prices. Resistance to this trend came in the shape of extending bi-laterals to regulate non-scheduled services for the first time. This can now be seen as a recognition that the distinction between scheduled and non-scheduled services had become so blurred, and the latter so important, that they warranted a similar degree of regulation. While the initial reaction to the Oil Price Crisis was to *increase* the level of regulation, this proved to be unsustainable. Attention switched to considering how the *efficiency* of operations could be permanently improved, allowing the airlines to absorb as best they could, permanently higher fuel costs.

Table 4.1 charts the improvement in the fuel efficiency of the workhorse Boeing 747, since it first entered service in 1970, through to the last production version, the 8(00). It is easy to see the improved conversion of fuel into payload transmission, particularly in the ultimate 800-series, emanating from a combination of improved aerodynamics and engine technology. In addition, progressive improvements in engine *reliability* have allowed long-haul flights over water to be serviced by twin-engined aircraft, such as the Boeing 787. Development in computer hardware and software and the Internet led to the increased use of computerised flight management systems; satellite navigation and fly-by-wire technology, displacing the Flight Engineer and Navigator from the cockpit, whilst actually reducing the workload placed on the remaining Pilots. Common-style flight controls allow them to fly more than one type of aircraft in the fleet, permitting more flexibility in rostering flight crew and upping their utilisation while saving on training costs. Computerisation of ticketing and general administrative systems have streamlined passenger and freight handling operations, as well as maintenance procedures.

While there are still national 'Flag Carriers', the trend over the last 30 years has been towards international mega-groupings not identified with any one country in particular. For instance, the International Consolidated Airlines Group ('IAG') was founded in January 2011 to hold the assets of both British Airways and Iberia, both formerly the State-owned airlines of the UK and Spain, respectively.

Table 4.1

Development of the Boeing 747 (1970 to the present)

Series #	Entered service	MTOW 000 kgs	Fuel burn (kgs/hour)	Cruise Speed(kts)	Typical Range (kms)	# Passengers
100	1970	333	12,880	484	13,020	550
200	1975	378	12,990	484	13,690	550
300	1980	378	12,110	490	13,590	600
400	1988	397	11,330	490	13,490	660
8(00)	2008	448	7,280	504	14,320	605
Source: Boeing Aircraft Corporation data						

Since then, it has been joined by the Irish State-owned airline, Aer Lingus, Catalunyan-based airline, Vueling and some other smaller airlines and had a combined turnover of €23billion in 2017 with 63,000 employees, world-wide. [53] IAG's members maintain separate corporate identities and local management, administrative and technical headquarters and even appear to compete against each other on key routes, such as London to Madrid. This, however, masks the practice of *codesharing* where timetables are co-ordinated and flagged with each one's flight numbers to create the illusion that you are flying with 'X', when the service is actually operated by 'Y'. The constituent airlines have a deliberate policy

[53] https://en.wikipedia.org/wiki/International_Airlines_Group, accessed 11th April, 2018

of harmonising their fleets of aircraft, with a view to reducing purchase costs (having with power to negotiate discounts for aircraft manufacturers when making mega fleet orders) and operating costs, via scale economies from maintaining fewer distinct aircraft types and facilitating intra-and inter-fleet exchanges of cockpit- and cabin crew. To compound its internationalisation even further, IAG is currently 20% owned by Qatar Airways.

Other airlines have more explicitly merged their companies into single entities, such as the 'merger' of United and Continental Airlines which was completed in 2012. Usually, the smaller partner (in this case, Continental) is subsumed into that of the larger. Others have preferred to maintain their independence, and have grown, in some senses, organically from past profits with the assistance of venture capital (such as easyJet and Ryanair). While they maintain their Corporate HQs in their countries of origin, they derive most of their revenue outside that territory and can otherwise be androgynous, the sense that their name and corporate colours give no clues as to their country of origin (such as Iceland's WoW airline). For all that they are independent, they have also entered into codeshare agreements with other 'like' airlines, particularly those operating in other territories, with one inter-connecting hub.

One thing all airlines have had in common has been the move towards becoming what are now called euphemistically, *low-cost operators*, characterised by ultra-low fares, a stripped-down service with extras (such as hold baggage and in-flight meals and entertainment) charged for, fast turnaround times, leaner fleets and workforces. The drivers of this change are, in part, technological, pertaining to both aircraft and administrative systems' performance, but are also macro-institutional. It has to be expected that these improvements in technology, pursuit of scale economies and increased freedom to pursue business will have made airlines more technically efficient, but that some will have been more successful than others in pursuing this goal. The incumbent airlines have generally responded to this influx with a combination of a similar stripping-down of service quality and staffing levels, while increasing the aircraft utilisation, but also with loyalty schemes (such as 'Airmiles')

In summary, the post-Second World history of the international airline industry can be split into three broad 'eras':

1. Immediate post-War era to the early 1980s: a period of maximal regulation and State control/influence, akin to Infant Industry Protection (and restoration). The advent of 'Reaganomics' (USA) and 'Thatcherism' (UK), for example, re-established the pre-War consensus of minimal State control of even key strategic industries and increased exposure of the constituent operators to market forces.
2. Early 1980s to mid-2000s: a rapid move towards greater competition, less economic regulation, lower fares and service standards, with tighter profit margins and increased utilisation of aircraft and employees. The effect was to usher in the 'low-cost' carriers and increase competition between airports for traffic. This era of confidence and expansion came to an end with the 2008 Global 'Crash'.
3. Mid-2000s to the present: the emergence of the 'mega-groupings' alliances and merged airlines, in part, as a response by established 'Flag Carriers' like British Airways, to the low-cost operators like Ryanair and EasyJet, but also to reduce operating costs and finance the ever-increasing real cost of aircraft, amongst other factors. Add to this the rapid growth and internationalisation of large airlines in the Middle East (Emirates and Qatar Airlines, for example) and in China. As **Figure 4.1** indicates, the real price of aviation fuel has risen steadily since 1990, punctuated with several spikes after 2005. In an article for the Blogsite '24/7', Ausick (2017) estimated that an original Boeing 747-100 would have cost $149million in then-current prices, compared to a list price of $379mliion for the ultimate production model (the 800 series), a 150% increase in real terms.

4.2.2. Measuring Industry Competitiveness

Industrial Economics has a well-developed body of theory pertaining to the 'Theory of the Firm' in an effort to understand past, and to predict the future structure and performance of industries, the extent of competition within the industry, prices and output/profit shares. In the sense that competition can be equated to the distribution of activity within an industry, popular metrics have been derived from

Figure 4.1. Real Spot Price of Aviation Fuel (US Gulf Coast Kerosene) [54]

output measures (profit, revenue, physical output), such as the m-firm *Concentration Ratio (CR)*, where:

$$CR_m = \sum_{i=1}^{m} \frac{q_i}{Q}$$

Where q_i is the output of firm i, one of the top m of the n firms in the industry, and Q is total industry output (as defined) in that period ($m < n$).

[54] Real Price data: US Energy Information Administration (https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx? n=PET&s=EER_EPJK_PF4_RGC_DPG&f=M,). Fuel Price Indexes: Federal Reserve Economic Data, Economic Research Division (https://fred.stlouisfed.org) Both accessed 23rd July, 2018

In effect, this computes the market share of the top *m* firms in the industry and can be used to compare the same industry in the same location at different periods of time, or the same industry in different locations at the same period of time, or a different industry in the same location at the same period in time, for example. Interpretation is very straightforward: an increase in CR_m implies 'power' in the industry moving to the top players, in effect becoming less competitive and (perhaps) more oligopolistic in nature. The main objection to the CR is that ignores the ($n - m$) firms not included in the calculation and may not capture the full flavour of changed competitive conditions between two periods.

A development from the CR is the Herfindahl-Hirschmann Index (HHI), *H*, that exaggerates the skew to larger, dominant firms in an industry by squaring the market share, i.e.

$$H = \sum_{i=1}^{n} \left(\frac{q_i}{Q}\right)^2 \qquad [\mathbf{4.1}]$$

All *n* firms in the industry are included in the calculation of *H*, overcoming some of the shortcomings of the CR. While there are still shortcomings with the HHI, it remains widely in use by both academics and industry regulators, who use it to estimate the impact on the HHI for a particular industry if, for example, a proposed merger between two existing firms were allowed to go ahead.

Given the value of *n*, it is possible to set limits to H. At one extreme, Monopoly, H would equal:

$$H = 1.\left(\frac{100}{1}\right)^2 = 10{,}000$$

i.e. 100% of the output is in the hands of the monopolist. In all other cases, there is an HHI that would pertain were all output to be shared equally between all *n* firms, as:

$$H = n\left(\frac{100}{n}\right)^2 \qquad [\mathbf{4.2}]$$

$H \to 0 \; as \; n \to \infty$, in other words, an industry can be assessed both in terms of how 'power' is distributed amongst the *n* firms, and at what end of the Monopoly-to-Perfect Competition spectrum it is/was located at a particular moment in time.

Conventionally, an **H** less than 2,000 is regarded as an indication of low concentration (i.e. a high level of competition); 2,000 to 4,000, medium concentration and 4,000 to the maximum 10,000, high concentration (and little or no competition).

Several papers have been written estimating Concentration metrics for airlines, many focussing on specific key routes, rather than whole sectors of the Industry (e.g. Lijesen, et al (2002). One recent (conference) paper by Yasar and Kiraci (2017), estimated airline HHIs for the major regions of the World, although using unspecified output measure of 'market share', covering the period 2006 to 2015. They found that most regions in most years had HHIs less than 1,000, implying a high degree of competition, but that these were increasing steadily. There were considerable differences between the regions, with Australasia and North America having HHIs over 1,000 in 2015, while the Asia-Pacific region had an HHI of 235. While inter-regional differences are not surprising, given the hugely different geographies and demographics, the closeness of the Global Industry to the highly competitive end of the spectrum is surprising, but this may be a result of the assumptions made about the minor players in each Region's industry (as will have to be made in this Study, immediately below).

Airlines, of course, have several outputs: financial outputs such as total revenue and profits, rates of return are often analysed. Physical output such as total passengers carried, total passenger-kilometres flown; tonnes of cargo carried and cargo tonne-kilometres flown, routes flown, airports served, are also used. Given that this Study focusses on *physical* rather than financial output, it is appropriate to consider the trends in measures such as passengers and cargo carried. **Table 4.2** computes the CR_3s and HHIs for the World Airline Industry for 2011 to 2016, where output is taken to be the number of passengers carried on all services in each year. The 10 largest airlines, in terms of passengers carried in 2016 have their squared market shares calculated, and it was assumed that all other airlines shared their residual output, equally. [55]

Table 4.2
Concentration Measures for the World Airline Industry: Passengers Carried 2011 - 2016

Rank (2016)	Airline	2011	2012	2013	2014	2015	2016
1	American Airlines	9.03	8.30	7.69	7.05	17.20	27.39
2	Delta Air Lines	15.78	15.18	14.85	15.31	15.44	23.41
3	Southwest Airlines	14.92	14.03	13.58	15.23	16.75	15.98
4	United Airlines	3.11	9.56	8.30	7.47	7.29	14.22
5	Ryanair	7.13	7.07	6.76	6.82	8.24	9.95
6	China Southern Airlines	7.91	8.29	8.55	9.26	9.57	5.00
7	China Eastern Airlines	6.86	7.06	4.01	4.00	4.52	4.54
8	EasyJet	3.62	3.54	3.48	3.55	3.95	3.71
9	Turkish Airlines	1.30	1.69	2.38	2.64	2.90	2.74
10	Lufthansa	4.84	4.62	4.09	3.27	2.99	2.70
-	All others (average)	0.91	0.88	0.91	0.91	0.85	0.80
HHI (H)		**5558.05**	**5387.63**	**5527.59**	**5524.49**	**5218.03**	**4928.12**
CR_3		**0.108**	**0.105**	**0.103**	**0.105**	**0.122**	**0.141**

Source: Total Outputs: ICAO 2016 Annual Report https://www.icao.int/annual-report-2016/Documents/ARC_2016_Air%20Transport%20Statistics.pdf, accessed 23rd July 2018. HHIs computed by the Author.

The CR_3 measure indicates that the top three airline groupings accounted for a larger share of the passengers' total, particularly after 2014. American Airlines absorbed the large airline, US Airways during 2013 which accounts for its rapid increase in size, to restore its previous position as the largest airline in the World (by virtually all output measures). In so doing, it leapfrogged Delta, whose remarkable upturn is 2016 appears to be organic, rather than achieved by merger (it had previously absorbed Northwest Orient, another mid-sized airline in 2009.

[55] An heroic assumption, but as it is made consistently across all years, should still allow the direction of travel to be ascertained.

It is possible to estimate the minimum possible HHI value, by taking IATA's and ICAO's 2016 reconciled lists of registered airlines which indicates that there were approximately 6,000 registered airlines operating during this period. [56] Using **[4.2]**, the lower limit for *H* is therefore 1.67: the industry, in terms of passengers carried, at least, is clearly neither perfectly competitive or monopolistic and would appear to be located at the oligopolistic part of the spectrum, though some modest movement back towards the 'middle ground' may have taken place in 2016. [57]

Table 4.3 repeats the same exercise, but output is now tonne-kilometres of cargo carried. Unsurprisingly, dedicated cargo airlines and freight divisions of other airlines dominate this sector of the industry.

Table 4.3
Concentration Measures for the World Airline Industry: Cargo Carried 2011 - 2016

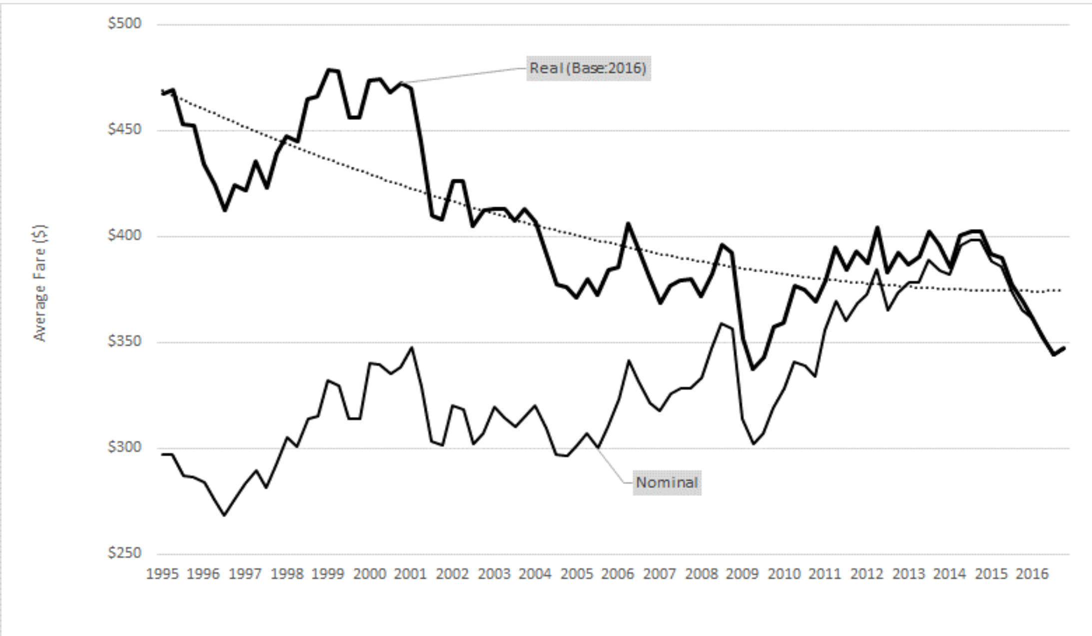
Rank	Airline	2011	2012	2013	2014	2015	2016
1	FedEx Express	72.346	75.454	75.034	67.601	64.093	58.803
2	Emirates Sky Cargo	18.832	25.254	31.560	33.278	37.949	35.861
3	UPS Airlines	31.792	31.550	32.319	31.502	29.989	30.222
4	Cathay Pacific Cargo	23.628	20.681	19.594	23.593	25.345	23.568
5	Qatar Airways Cargo	1.139	3.562	7.132	9.473	15.066	20.253
6	Korean Air Cargo	22.933	19.287	16.955	17.193	15.466	13.998
7	Lufthansa Cargo	16.770	14.971	15.031	13.107	12.183	12.987
8	Cargolux	7.231	9.525	7.876	8.718	10.221	11.268
9	Singapore Airlines Cargo	14.428	13.031	11.234	9.543	9.501	9.590
10	Air China Cargo	5.767	7.270	8.104	7.968	8.395	8.831
-	All others (average)	0.001	0.001	0.001	0.001	0.001	0.001
HHI		**214.87**	**220.58**	**224.84**	**221.98**	**228.21**	**225.88**
CR_3		**0.185**	**0.193**	**0.200**	**0.196**	**0.196**	**0.192**

[56] https://en.wikipedia.org/wiki/List_of_airline_codes (accessed 23 July 2018).

[57] An HHI of 4,000 or more being regarded as the point at which an industry is not competitive. This may be illusory, however, as it critically depends upon the assumption made about how the residual market is shared between the assumed 6,000 non-top 10 firms. The reduction in their average share (squared) resulting from the increased share taken by the top 10 aids to create the impression of a greater spread of activity across the entire industry.

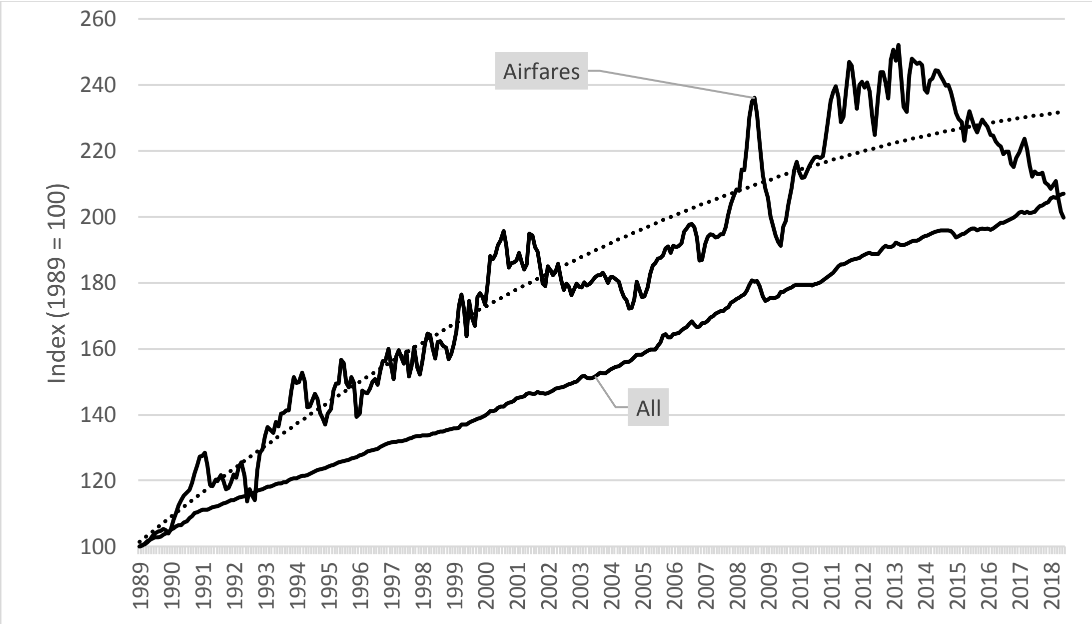
The CR_3 measure indicates both a slight increase in concentration in the top three firms, and a greater degree of concentration than that witnessed above in the passenger carrying sector. As virtually all airlines in the World carry freight (to utilise otherwise empty pace in passenger aircraft), the HHI assumes that there are, in effect, 6,000 firms in the sector, each with a tiny share of the cargo market. The effect is to estimate much lower HHIs than in the passenger sector, implying a greater spread of activity around this sub-sector, though again, this may be largely illusory. There is perhaps some evidence of an increase in concentration towards the bigger players, though no clear, steady trend is apparent.

It appears from this analysis, that the airline industry, particularly the passenger-carrying sub-sector is oligopolistic in nature, as airlines have merged or formed alliances and new, low-cost carriers have grown into major players in a short space of time. One outcome that might have been anticipated as a result, is that real and nominal airfares would have risen. **Figure 4.2** plots the average airfares paid by consumers in the US from 1995 to 2016, while **Figure 4.3** charts the prices indices for airfares and all items in the US from 1989 to 2016.



Source: Bureau of Transportation Statistics, BTS Air Fares, Origin and Destination Survey

Figure 4.2. Real and Nominal Average Airfares in the US: 1995 – 2016



Source: Federal Reserve Economic Data Economic Research Division (https://fred.stlouisfed.org: Accessed 25 July 2018)

Figure 4.3. US Consumer Prices Indices for Airfares and all Items: 1989 – 2018

The figures reveal a picture of considerable fluctuations in nominal airfares through the period, with above-average increases until 2017, clearly affected by global events, most recently, the 2008 global recession. Real fares also fluctuated, though the long-term trend appears to be downwards, thanks, in particular to the deflation in nominal fares in 2015 – 2016, a period when the industry worldwide, and in the US, appeared to be less competitive. [58]

It is difficult, therefore, to apply the simple 'textbook' equivalence of reduced competition therefore higher (real) prices to this industry. Indeed, it seems that the two may actually have occurred simultaneously. This does not necessarily confound the Economic Theory of the Firm, *per se*, as the apparent reduction in competition may have contributed a counterfactual rise in airfares, but that this was more than

[58] The factors determining the prices of air tickets are many and the subject of many papers. One paper by Sengupta and Wiggins (2014) estimated that airfares purchased 'on-line' in the US in 2004 were, on average 11% lower than those purchased 'off-line'. Of course, this was at a time when airlines and consumers were in transition between the pre-Internet days when all tickets were purchased 'off-line' and the situation at the time of writing where virtually all purchases are made 'on-line' and should no longer be a pricing factor. The value of the paper, however, is that it highlighted the many (other) factors in pricing that still pertain, including load-factors, global economic conditions and events, as well as cost-factors, particularly the cost of fuel (**Figure 4.1**).

offset by downward pressures from other factors, such as 'territory-grabs' via price wars. It is also the case that the early advantage that the 'low-cost' entrants had, have largely dissipated, as the incumbents responded to their entry (Tsoukalas, et al., 2008). The industry appears to be in a situation where unit costs have been lowered via scale economies, but that the fewer, larger operators are not (yet) willing to reform the territorial monopolies and cartels they once enjoyed.

Bearing all that in mind, it might be hypothesized that a reduction in competition might reduce the pressure on airlines to be input-efficient in that period. This will be considered in the Results' section of this Study, but to stress (in advance) that the multiplicity of 'other factors' will not allow there to be a definite link made.

4.3. Quantifying airline efficiency: Methodology & Literature

4.3.1. Introduction

The aim of this section is to outline the methods that will be used to assess the efficiency of a sample of airlines, using a Panel assembled by the Author. In particular, the nature and purpose of the *Production Function*, as used by Economists will be described, and from it, the development of the idea of the *Frontier Production Function*, an output ceiling beyond which producer are unable to go, at present, given technology and operating *modus operandi.* On the lead up to that, an outline history of literature on efficiency and studies on Frontier Production Functions will be outlined.

4.3.2. Previous Studies on Airline Efficiency

Academic interest in airline efficiency dates from the mid-1980's when the industry was being transformed as outlined in section **4.2**. The ever-increasing commercialisation of the industry was seen as an opportunity by academics to develop and quantify suitable measures of efficiency. Airlines, however, had been content to calculate and report simple measures of 'efficiency' in their Annual Reports, using output measures like Total Kilometres Performed (TKP) [59], namely:

$$TKP_i = \sum_{n=1}^{k} D_n f_n \overline{P}_n \qquad [\mathbf{4.3}]$$

- where D_n is the distance in kilometres between route *n*'s origin and destination airports

  F_n the annual frequency of route *n*

  $\overline{P}_n$ The average payload in kilogrammes, carried on route *n*

Ratios like TKP/A per employee are still used today as indicators of efficiency, being relatively simple to compute and understand. They have, however, been criticised for being *too* simplistic and therefore, misleading. On the one hand, a measure like TKP per employee tends to favour airlines whose schedules are dominated by long-haul flights (such as Qantas of Australia or Emirates of Dubai). It is relatively easy for them to generate a high TKP using a few large-payload aircraft, like the Airbus 380, or Boeing 747-800 (recall **Table 4.1**). Accordingly, they require fewer employees per TKA than airlines like British Airways, who perform a mixed portfolio of short- and medium- and long-haul flights, also regional airlines, like Scotland's Loganair who operate a high proportion of ultra-short-haul inter-Island flights. The same might be said of cargo operators, who can fill aircraft with weight-dense freight, with no need for cabin crew and passenger-ticketing systems and staff.

Even comparing such measures for an airline through time can be misleading, as many carriers change their business mix in the pursuit of greater profitability and business security. For example, many Flag Carriers prefer to franchise out much of their short-haul domestic work to locally-based carriers. As these shorter routes require more labour input per kilometre, their shedding results in a remarkable upward shift in the Flag Carriers' performance statistics: again, a false impression of greatly improved efficiency is therefore gained. The basic problem is that relating

[59] An alternative measure, Tonne Kilometres *Available* (TKA), relates the potential payload (passengers and freight) that could have been carried, rather than that actually carried.

output only to Labour, excludes on average two-thirds of all inputs to the process. The remainder consists mainly of physical Capital and Fuel: Capital having steadily displaced Labour over the decades. It is inevitable that productivity per employee will have risen, though it does not necessarily imply that *total* efficiency has necessarily risen as well.

Considering the unit cost per TKA/TKP is an improvement, as it *de facto* includes the cost of all inputs. [60] That said, this measure can only be used inter-temporally for a given airline and, at a pinch, for airlines operating in the same country, as input prices vary widely across the World. For example, the average remuneration of a Pilot flying for the US-based United Airlines in 1988 was $116,000, compared with only $29,400 for Air India. There are even wide differences in prices within the same country. For example, in 1989, British Airways paid its pilots an average of £74,000, whilst British Midland paid only £31,000. Fuel prices in 1989 were on average 78% higher in Africa than in North America; 60% higher in South America and (paradoxically) 26% higher in the Middle East.

There have been a few papers that have attempted to increase the sophistication of productivity/efficiency measures in the airline industry. For example, Forsyth, et.al. (1986a; 1986b) attempted to produce a level playing field on which, they claimed, a more equitable comparison of efficiency could be made. Their first step was to convert actual TKA to an adjusted figure. The approach they took was to *reduce* the output of airlines that flew an above-average stage length and/or utilised larger than average aircraft (and *vice versa* for below average lengths / aircraft), all with reference to a hypothetical 'average' airline. Their next step was to standardise input costs by adjusting prices using an index to estimate the costs it would have paid, had it been the 'standard airline'. Their final step was to calculate the adjusted cost per adjusted kilometre. This process changed the 'efficiency' rankings, increasing the imputed per-kilometre cost for long-haul operators, whilst reducing for short-haul. Interestingly, they demonstrated a *decline* in total factor

[60] With the proviso that the depreciation charge in the Accounts usually understates the true amount of capital used in providing their Service.

productivity for almost all airlines over the period 1979-1984, contrary to the impression given by most airlines during that period. Whilst no specific comment on this work was found, it can be criticised on various grounds. Firstly, the adjustments made to output from flying non-average stage lengths and aircraft seem rather subjective. The indexes used to standardise input prices were derived from national aggregates, and not from the prices actually paid by each airline. The results were based on a small sample of big airlines and it is not clear how sensitive the actual results and rankings were to an increase in the sample size. In any case, the airlines have not paid any attention to at least the spirit of this work and have continued to report the traditional measures described earlier.

4.3.3 Production Functions as a measure of producer efficiency

Economists have developed a number of different 'families' of Production Functions in their efforts to model specific production scenarios. Their motivations vary, but are usually done to understand better the competitive situation in specific industries, or to understand the extent of technological development in a particular industry or macro-economy, or to derive factor-demand curves, in particular, for Labour. The adopted method is to assume first that a particular functional form of the relationship between output and input(s) is appropriate then to use historic data to estimate the function's parameters. Given that most production processes are highly complex, perhaps involving thousands of inter-acting chemical and engineering processes that use many different forms of labour and equipment, it is unreasonable to expect the empirical Economist to model the Process perfectly. In any case, there is likely to be a diminishing return to the model's ability to describe the process, adding more and more factors and mining the data to estimate the added parameters. That said, the economist should at least understand the broad underlying process by which the productive unit generated its output, in particular:

- The significant, distinct input factors
- The nature of their interaction with each other and output

- Their ability to substitute for each other, should the circumstances so dictate.

Airlines' output (carrying people and cargo through the air) cannot happen without both aircraft ('capital') and employees ('labour'). Labour, in this industry, would comprise of pilots, cabin attendants (for passenger services), ground and maintenance crew and layers of middle- and senior administrators and management to co-ordinate operations and devise and implement strategy. The absence of any of these factors would render any airline unable to fly beyond the very short term: perhaps as little as a week. This would tend to rule of the simplest class of Production Function, the *Simple Additive Function*, which in its simplest form might look like **[4.4]**

$$q = A + \beta_1 L + \beta_2 K \qquad [\mathbf{4.4}]$$

- where *q* is the level of total output in a fixed time period

  L is the amount of labour applied in the process in that period

  K is the amount of capital applied in the process in that period.

Output may be measured as a physical quantity (as in this Study), which can lead to problems where more than one product is emanating from the chosen system ('apples' and 'bananas'). The normal remedy is to convert all output to a money equivalent, though this is not without problems. For instance, using cost, or price can create illusions of more or less output arising from inflation. The preferred approach here, is to convert all output to a standard physical measure in the style of **[4.3]**.

Labour: as mentioned above, there can be more than one distinct type of Labour involved, which could be expressed as L_1, L_2, and so on in an expanded version of **[4.4]**. Whatever categories of labour are included, it is important to be sure what a 'unit' of labour actually is and to express the amount applied in each period consistently. For example, a unit may be equated to a standard employee, working normally for 35 hours a week for 48 weeks. Employing one person on a half-time contract (17.5 hours) for 24 weeks would therefore equate to ¼ Unit. Employing 10 people for 40 hours a week each, for 50 weeks would therefore equate to (10 * 40/35 * 50/48) = 11.9 units. Adding in the half-timer would imply a labour input of 12.15 units. Similar issues regarding 'addibility' arise, with money sometimes used to combine different types and quality of labour together.

Capital: care and consistency has also to be applied in considering what capital is used in the process. Again, the conversion of actual amounts into 'units' needs to be explicit, as well as the inclusion/exclusion of the elements that make up Capital. Some Studies roll together financial- and physical capital, measured in monetary terms, while others, like this one, include only physical capital (in fact, only one major sub-group of physical capital, the aircraft).

The *A* term in **[4.4]** appears to quantify the output that would be achieved even if none of the included factors were entered into the process, though its true role is to act as a proxy for factors not explicitly modelled, that affect output, in some way. The importance of this term cannot be known in advance and will only emerge after estimation.

[4.4] implies that output is possible from any factor, either in isolation, or in combination with any others: even if all modelled factors are absent (assuming $A>0$). It also implies that each factor's ability to generate a unit of output is not affected by the presence of any other factor: i.e. they are *perfectly substitutable* for each other. Neither do they depend upon the current scale of production. From the observations made already in this Chapter about the airline industry, it is clear that additive functional form is not appropriate for the Study's purposes. What is required, is a functional form where the absence of any single factor input means $q=0$, even if the others are abundant. The most well-known and, empirically, widely-used functional form is the so-called *Cobb-Douglas* function (Cobb and Douglas, 1928), though it can be traced back to Wicksteed (1894). The basic functional form for the two-factor case is:

$$q = AL^{\beta_1}K^{\beta_2} \qquad \textbf{[4.5]}$$

This satisfies the requirement that *all* factors must be present in order for there to be any output. *A* is now a scaling factor that aligns the different units of measurement employed. In order to allow econometric estimation of *A*, β_1 and β_2, it is necessary to log-transform **[4.5]** and estimate **[4.6]**, instead:

$$\ln(q) = \ln(A) + \beta_1 \ln(L) + \beta_2 \ln(K) \qquad \textbf{[4.6]}$$

Interest often focusses on the sum of the exponents ($\beta_1 + \beta_2$) in an effort to discover whether the process exhibits constant (=1), increasing (>1), or diminishing (<1) returns to scale. This will be considered in the Results' section, **4.5**.

While **[4.5 / 4.6]** are more realistic representations of the realities of the production process in the airline industry, it does not allow for any *synergies* between the factors to affect output. For example, adding a log-unit of labour to a process that is currently 'labour-rich' and 'capital poor' would increase output by β_1 log-units. However, so would adding the same log-unit of labour to a process that is, instead, 'labour-poor' and 'capital-rich'. Reality would tend to suggest that labour added to an already 'labour-rich' production process would struggle to find the spare capital they need to allow them to increase output, compared with the identical labour added to a 'capital-rich' environment. The latter would easily be able to find otherwise underutilised capital and add more to output: not because they are superior to the former, but because they have the under-utilised spare capital available. Accordingly, researchers also employ the *Translog* function in their empirical work. The two-factor function might look like **[4.7]**:

$$q = AL^{\beta_1} K^{\beta_2} (L^2)^{\beta_3} (K^2)^{\beta_4} (L.K)^{\beta_5} \qquad [\mathbf{4.7}]$$

- which is estimated, conventionally in the style of **[4.6]** as:

$$\ln(q) = \ln(A) + \beta_1 \ln(L) + \beta_2 \ln(K) + \beta_3 \ln(L^2) + \beta_4 \ln(K^2) + \beta_5 \ln(K.L) \quad \textbf{[4.8]}$$

As well as allowing for a non-linear effect on *q* from changes in either K or L, a joint effect of the coming together (a 'synergy') of both factors is now modellable. It is straightforward to imagine that in this industry, where an idle Aircraft would be infinitely better utilised where more of the relatively scarce factor (Pilots) was provided, for example.

It is clearly one thing to estimate a production function to measure how a producer has converted the available inputs into output. It is another matter to determine whether the producer could have obtained a higher level of output from the same inputs or, alternatively, used fewer inputs to obtain the same level of output. In other words, could its operations have been more *efficient*? Engineer and Scientists understand the concepts of *improved* or *relative* efficiency and also have an idea of what *maximum theoretical* efficiency would look like in the particular context (e.g. aerodynamic perfection leading to frictionless flight). The producer, however, will be concerned with the efficiency of *their* production process, relative to that of its competitors, as a higher level of efficiency transmits itself to lower unit costs that, in turn, may give it a competitive advantage. Economists have, therefore, to measure efficiency *empirically* with reference to *observed*, rather than *theoretical* best practise. Somewhere amongst the various practises, there will be one that was observed that was relatively the best at that time.

It became obvious to some that the Production Function had a role to play in determining the limit, or *frontier* beyond which no producer could currently go. Schmidt (1985) provided a revised definition of the production function as:

> *"... a function giving the maximum possible quantity of some output, given quantities of a set of inputs"*

This introduced the concept of the *Frontier* production function, beyond which by definition, the producer could not currently locate. [61] Farrell (1957), showed that efficiency could be decomposed into *technical* and *allocative* efficiency. Technical inefficiency arises from using more of at least one input in producing a given level of output than might otherwise have been possible, given current best practise. Formally, if *f(x)* is the efficient transformation of *x* a vector of inputs, into *y* a vector

[61] The frontier *cost* function embodies the minimum cost of producing a given level of output, given current factor prices, but will not be considered further in this Study.

of one, or more outputs, then if firm *i* at time *t* has a production plan (y_{it}^0, x_{it}^0) then it is being technically efficient iff $y_{it}^0 = f\left(x_{it}^0\right)$, otherwise it is technically inefficient. [62] Allocative inefficiency arises from not using inputs in the correct proportions as determined by the prevailing factor input prices, which would otherwise permit cost to be minimised given the level of output. Formally, if w_a and w_b are the unit prices of inputs *a* and *b* to a process, then the plan (y_{it}^0, x_{it}^0) is allocatively efficient iff,

$$f\left(\frac{x_{ait}^0}{x_{bit}^0}\right) = \frac{w_a}{w_b}$$

Farrell's model was *non-parametric* in nature, as *f(x)* was not explicitly modelled in relation to inputs and outputs. A *parametric* frontier model, on the other hand, assumed a specific functional form for *f(x),* the most popular being Cobb-Douglas, though others such as the Leontieff and Translog have also been applied, for example, Førsund and Hjalmarsson (1979). Aigner and Chu (1968) developed a model where output was bounded by a *deterministic* industry frontier production function. Assuming a Cobb-Douglas form was appropriate, the model they estimated was:

$$\boldsymbol{ln(q_i) - \left(\beta_0 + \sum_{j=1}^{k} \beta_j ln(x_j)\right) = u_i} \qquad [\mathbf{4.9}]$$

Crucially, u_i was constrained to be non-negative (a one-sided error term), thereby ensuring that all producers would be located on or below the frontier function placed in position by the estimated parameters.

While incorporating a specific functional form allowed Farrell's constant-returns assumption to be relaxed, its introduction was seen by some as forcing the data to conform to the modeller's pre-conceived notion of the relationship between the inputs and output. While subsequent studies used a variety of difference functional forms, purists continued to object to the imposition of any form at all. In addition, the continued use of mathematical programming methods, as distinct from

[62] It cannot, of course exceed *f(x)* as this is the modelled (production) frontier.

econometric ones, retained the problems of the non-parametric approach. Specifically, the technique insisted that at least one of the producers in the data had to be located on the Frontier, even where it was patently inefficient. It also placed an upper limit of the number of firms that could be technically efficient, determined by the number of parameters included in the model. Being non-econometric in nature, the algorithms employed could estimate only single-point parameters and made no assumptions regarding the sampling properties of the data or the distribution of the error term. These models are now known as *Deterministic Frontier Models*.

The Deterministic model was developed into a *statistical* one by Afriat (1972). His model assumed that u_i is iid, with mean μ and that the x_i and u_i are uncorrelated. Again, assuming, a Cobb-Douglas technology is appropriate, **[4.9]** became **[4.10]**:

$$\boldsymbol{ln(q_i) = (\beta_0 - \mu) + \left(\sum_{j=1}^{k} \beta_j ln(x_j)\right) = (u_i - \mu)} \qquad [\mathbf{4.10}]$$

OLS would yield consistent estimates of both (β_0 - μ) and the β_js. β_0 can be consistently estimated by shifting upwards until all residuals are non-positive. There followed a number of papers that examined the effect of assuming a variety of one-sided error distributions (e.g. the half-normal) and using MLE as opposed to OLS, for example Richmond (1974, op.cit) and Schmidt (1976). In spite of some misgivings regarding the consistency and asymptotic qualities of the ML estimates, partially resolved by Greene (1980), this became the preferred method of estimation in most empirical work in the 1970s and early 1980s.

The deterministic model assumed that it was valid to group all firms in the industry as if they all shared a common frontier. However, Aigner, et.al. (1977) argued that a random error term should also be included to allow the total error to be decomposed into:

1. A non-positive, u_i being the output lost over which producer *i has control* and could be reduced by improving its internal processes and,

2. An unsigned v_i which reflects factors *exogenous* to producer *i* over which it *has no control* causing short term movements in the Frontier (for example, an event favourable to a crop farmer like good weather), that may allow the producer to locate *beyond* the deterministic frontier from time to time. This term can also contain errors arising from measurement error of q_i as well as omitted variables' bias.

Given that most data suffer from both measurement error and missing variables, it is important that such an error term be included in the modelled frontier. Their reasoning was that there was a Utopia, in which there was no inefficiency, or 'good or bad luck'. Depending on what governing form was appropriate, then output would come precisely from a function like **[4.5]** (Cobb-Douglas) or **[4.7]** (Translog), for example. However, given that this Utopia does not (cannot) exist, and that producer *i* would be forever frustrated in its efforts to achieve technical perfection, output would be reduced by a factor ϵ_i , $(0 \leq \epsilon_i < 1)$. In essence, the true production function could be related to **[4.5]**, like this:

$$q_i = A\varepsilon_i L_i^{\boldsymbol{\beta_1}} K_i^{\boldsymbol{\beta_2}}$$

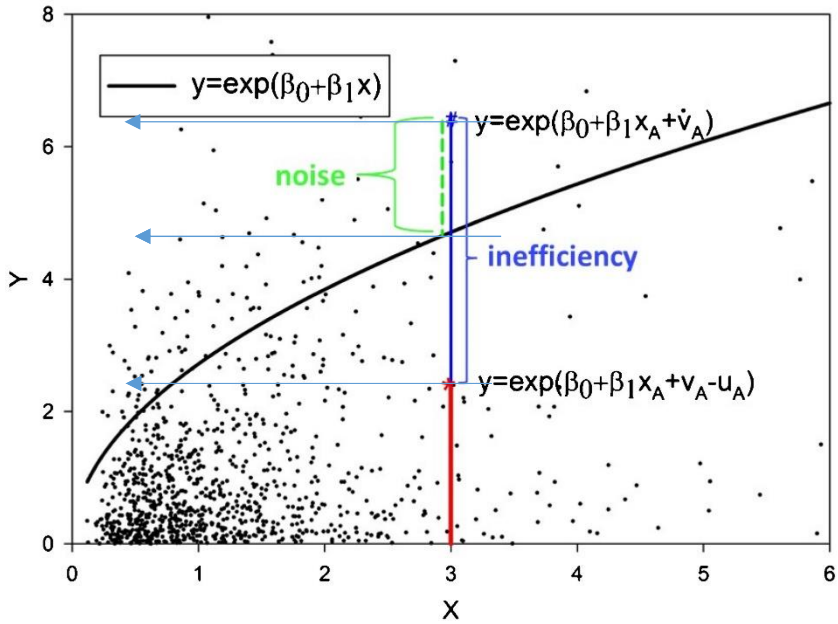
ϵ_i measures the producer's own technical inefficiency and is of primary interest to it (and the researcher). However, the uncontrollable, random shocks v_i, $(-\infty < v_i < \infty)$ that aid or hinder all producers cannot be ignored. If this v_i is also assumed to be multiplicative and non-linear, it can be included in the production function as

$$q_i = A\varepsilon_i e^{v_i} L_i^{\boldsymbol{\beta_1}} K_i^{\boldsymbol{\beta_2}}$$

Taking logs and rebadging $-ln(\varepsilon_i)$ as u_i gives **[4.11]**:

$$\boldsymbol{ln(q_i) = \beta_0 + \left(\sum_{j=1}^{k} \beta_j ln(x_j)\right) + (v_i - u_i)} \qquad [\mathbf{4.11}]$$

This is the *Stochastic Frontier Model* and was first employed by Aigner, et.al (1977, op cit) and Meeusen and v.d. Broeck (1977). **Figure 4.4** illustrates the relationship between the Frontier, u_i and v_i. Several producers are observed lying above (beyond) the Frontier, net beneficiaries of short-term noise working to their advantage, though masking any technical inefficiency they will probably have experienced. Some lie exactly on the Frontier, others well below. In this case, we could imagine two producers each applying an input (X) of three units. One producer obtained output just over two units in the period, while the other achieves just over six. *Prima facie*, the two-unit producer appears to be much less efficient than the six-unit one (as denoted by the inefficiency bracket labelled ''inefficiency' in the Figure). However, the six-unit producer happens to be the short-term beneficiary of some 'good luck' and has actually been able to progress beyond the 'industry-best' Frontier, as measured by the 'Noise' bracket in the Figure.



Source: Mutz, et. al, 2017

Figure 4.4. The Stochastic Production Frontier and Inefficiency and Idiosyncratic elements

The two-unit producer's inefficiency is really only then the gap between the ~2 and the efficient Frontier, which indicates that an output of ~4.5 would be standard to aim for, given the level of input.

Later papers, for example Stevenson (1980) and Lee (1983), experimented with different distributions for u_i. They found that the estimates obtained could depend upon the error distribution assumed, thus making it imperative that the one assumed is tested rigorously against alternative assumptions. As with the deterministic model, MLE is the preferred method of estimation, though Olson, et.al. (1980) argued that OLS could perform as well in most circumstances. A large number of papers appeared, most notably Førsund and Hjalmarsson (1979, op.cit), who estimated stochastic frontiers for the Swedish Milk Processing Industry between 1964 and 1973, to estimate the extent of technical progress in that industry. The same data was used by v.d. Broeck, et.al. (1980) to compare MLE estimates of stochastic and deterministic frontiers. More recent studies include Stevens (2005) who applied the methodology to assess the effectiveness of Universities in England and Wales, and Manonmani (2013) who estimated the efficiencies of producers in the Indian Textiles' Industry.

There is a small literature that has attempted to estimate airline efficiency by estimating production frontiers based on Panel Data. Schmidt and Sickles (1984), estimated a stochastic production function for the US Domestic Airline Industry based on quarterly data for 12 such airlines from 1970 to 1978. Whilst the results were based on a small sample, their methodology is illuminating and is of considerable help in formulating the methodology for this work. They assumed four inputs in the production function, viz,

- *Labour* - direct and indirect costs for all types of labour (Flight & Cabin Crew, Maintenance, Administration and Management) were aggregated.
- *Capital* - they followed the approach of Hoch (1962), in using capital depletion rather than stock, as the input measure. In particular, they calculated a depreciation figure based on an adjusted aircraft and equipment stock added to payments made for aircraft and equipment leased, landing fees and rental payments to airport authorities.
- *Energy* - consumption in (US) gallons converted into BTUs.

- *Materials* - e.g.-passenger services, administration and marketing.

As well as these physical inputs, they also recognised that airlines fly different types of schedule and included load factor % and average stage length as variables in the production function, along with seasonal dummies. Output was Capacity Ton Miles (CTMs),[63] they assumed a Cobb-Douglas technology, assuming distributional forms and correlated errors. [64] They were all broadly in agreement with each other in the sense that they produced similar *relative* efficiencies, though they recognised that the Panel was too small to claim that the actual efficiencies were accurate. In addition, given that the industry was highly regulated at that time, they should probably have estimated *cost*, as opposed to *production* frontiers. Nevertheless, as the authors themselves said, this could have been done with a change in the sign of the error term.

4.4. The Data & Econometric strategy

4.4.1 The Data

While there are plenty of data generated and summarised by airlines for their internal control and management purposes and various industry-related bodies, like the UK's Civil Aviation Authority and the United Nations' International Civil Aviation Organisation (ICAO), public access to it is rather limited. The airlines, obviously, have no incentive to make its data available to anyone as this would be of immense value to its competitors. The industry-level bodies, and related media, seek to make money from the data they collect and aggregate, charging exorbitant prices that only those in the industry can afford. Free, single-source data on the numbers of employees, aircraft and output at the level of the airline are therefore difficult to come by, and after some investigation, this Study was reliant on two main sources:

1. AirFleets website (www.airfleets.net). This is a free-to-use aviation 'enthusiasts' website, where details on the numbers of each main type of employee (flightdeck, cabin, ground, etc) and the number of each aircraft type operated

[63] This is effectively the same as TKAs defined earlier.
[64] In spite of the considerable differences in size of the airlines, there was no significant heteroscedasticity in the models.

each year dating back to 1990, gleaned from airlines' published data are contained. They sell various publications using this data, but the data themselves are downloadable, albeit not via a 'single click'. This phase of the data collection took several weeks' work, as each data item had to be cut and pasted from the site into an Excel file. In the event, the Study was limited to US Airlines only, despite virtually all airlines in the World being recorded on the website, as this was the only country that reported comprehensive output data (see next point).

2. FAA's Output Data: as this Study prefers to model physical, rather than financial inputs and output, it was necessary to discover a data-source that recorded, for each airline, their tkp in each year. The US' Federal Aviation Authority (FAA) publishes semi-aggregated annual data for all US-registered airlines, in the form of downloadable Excel sheets. The rows in each sheet itemise each route operated by the said airline, the number of times it flew that route in the year, the average payload carried each time (in lbs) and the distance flown (in nautical miles), how much was passengers, how much was cargo and how much was mail. It was relatively straightforward to apply **[4.3]** to each route, converting nautical miles to kilometres and lbs to kilogrammes and to estimate what percentage of each airline's tkp involved the carriage of passengers, and what percentage was therefore freight and mail ('cargo').

The next stage was then to marry the employee/capital data for each airline to the output data. This produced a dataset with 276 airlines; 113 of whom had the three key observations (q, K and L) for at least one year, between 1990 and 2014 (the other 159 were mostly missing either K and L, or q, though one or two had neither). Together, they comprised 1,233 observations, implying an average of ~10.6 observations out of a possible 25. 13 airlines featured in all 25 years of the Panel; 20 airlines appeared only once or twice. **Figure 4.5** charts the movement of the mean and extreme values for the output (billions of tkp), number of aircraft used by each airline; number of employees per airline and percentage of tkp that was generated by passenger, as opposed to freight business.

The top-left chart plots the maximum tkp generated by an airline in each year, with the minimum at the bottom. Although the minima were never actually zero, they were close to only one-million tkp, being generated by small, local airlines, flying low-frequency/payload routes over short distances. The mega-carriers, such as American and Delta, operating extensive long- and medium-haul networks with a variety of large- and medium sized aircraft, generated tkps in the billions.

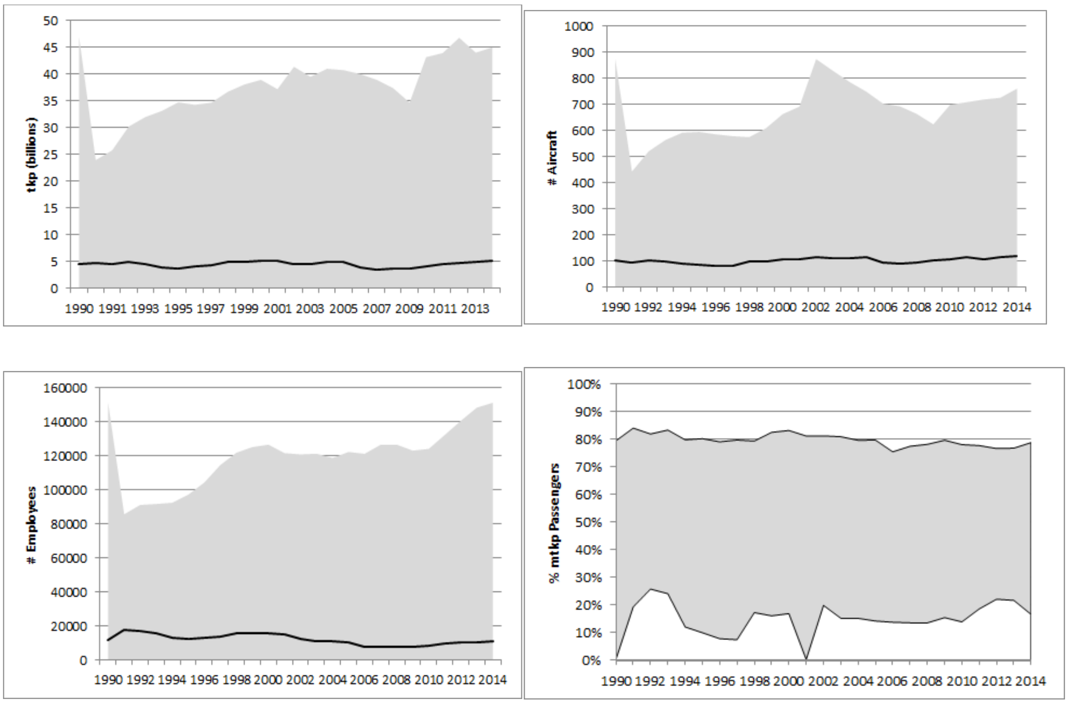


Figure 4.5 Central and extreme values of the dependent and independent variables 1990 - 2014

The solid line plots the annual mean tkp for all the airlines included in *each year's* cross-section. Thus, we see for the first time, the enormous range in airline size, and also the skew of the distribution towards the smaller end of the distribution. This indicates the domination of the industry by a relatively small number of 'mega-carriers'. There is no particular evidence of a trend in the mean, and nothing should be read into the volatile nature of maxima, given the nature of participation in this unbalanced Panel.

The top-right and bottom-left graphs plot the maximum and mean (solid line) numbers of aircraft and employees in each year. As with output, the difference between the 'mega-carriers' with fleets of over 500 aircraft and 80,000 employees, down to the small airlines with only one aircraft and 15-30 employees, can be seen. As with tkp, the means are much closer to the minima than to the maxima, and in the case of aircraft, no particular trend over the 25 years. It may just be possible to discern a downward trend in the mean number of employees, indicating a reduction in staffing, in the face of technological developments both in the aircraft, and in the ground-based administrative processes.

The bottom-right graph plots the mean (top of the shaded area) and minimum percentage of tkp that was due to carrying passengers (the maximum observed each year was 100%). This indicates a fairly wide range of business-mix in the US Industry, with most mixing passenger and cargo services together (often on the same flights). Only specialist carriers, like Federal Express concentrated completely on non-passenger business, though there are clearly some who focus on passenger-only work. The potential importance of business-mix as an indicator of durability, itself related to attrition from the Panel will be discussed later in section **4.4.2**.

Looking briefly at a couple of simple measures of 'efficiency', **Figure 4.6** plots the mean tkp (in millions) per aircraft and employees. Ostensibly, it appears that the output per aircraft fell over the period, but that as indicated in **Figure 4.5** above, the average output per employee rose by about 18%. The aircraft result is only partly surprising: on the one hand, one of the consequences of deregulation and subsequent rise of the low-cost carrier is that aircraft utilisation is much higher than before: the effect should be to raise tkp per aircraft. Ranged against that is the finite limit of aircraft size. Recalling that the Boeing 747 entered service in the early 1970 (**Table 4.1**) it was therefore the case that large aircraft, inherently more able to generate tkp, were already in service by 1990. [65] If anything, the trend since then, has been towards medium- and small aircraft more suited to thinner routes,

[65] The larger Airbus A380 entered service in 2007 but, as at the time of writing (April 2018), had not been ordered by any US Airlines.

or ones on which a high frequency of service is felt to be important. The effect would be to make it more difficult, counterfactually, to generate tkp. In addition, the aircraft used by these airlines are no faster today than the previous generation: the development of speed belongs to the 1960s when propeller-driven aircraft were replaced by jets.

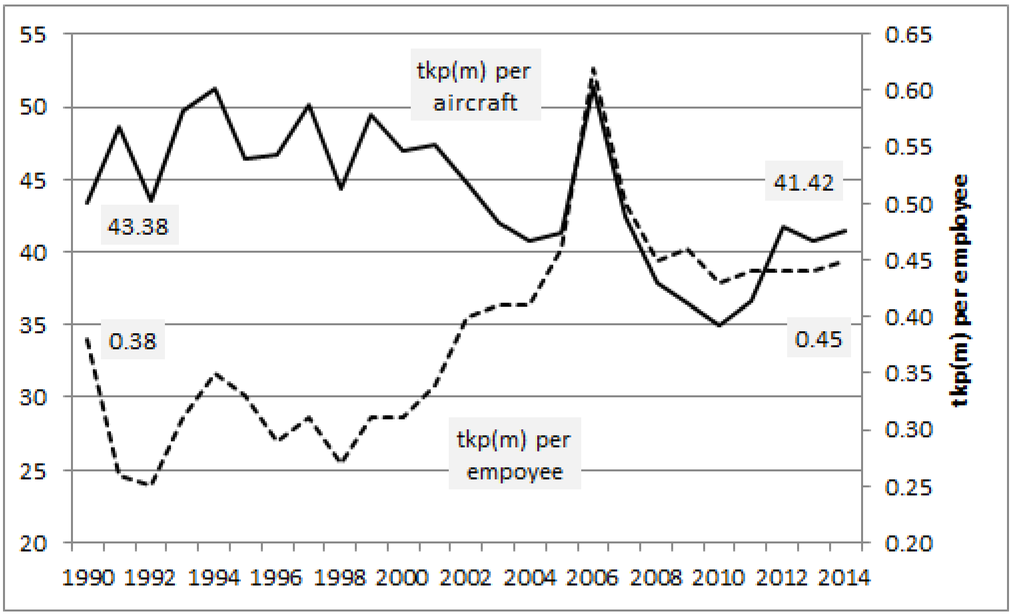


Figure 4.6. Output per aircraft and per employee: US Airlines 1990 – 2014

What cannot be seen is the extent of jet-engine and aerodynamic improvements over the period that have seen rates of fuel burn reduce, both in response to rise in fuel costs and the pressure to reduce harmful emissions into the high atmosphere. If this study was about *cost* or *profit* functions, then should be evident. What is evident, however, is the apparent substitution of labour for capital, allowing employees to generate 18% more tkp each in 2014 than their predecessors could in 1990. Automation in the flight deck has eliminated the ‘third man’, while improved design and reliability has reduced the demand for maintenance technicians. The introduction of internet-based/paperless ticketing has also allowed airlines to slim down airport-based administrators and head-office staff have been replaced by machines as well.

It should be noted, however, that there is considerable variation in these per-capita 'efficiency' measures. As **Figure 4.7** indicates, Polar Air Cargo appeared to be able to generate 50 – 300+ mtkp per aircraft, while Horizon could barely manage 10. The same Polar Air could generate between two and 20 mtkp per employee, while FEDEX couldn't even manage one.

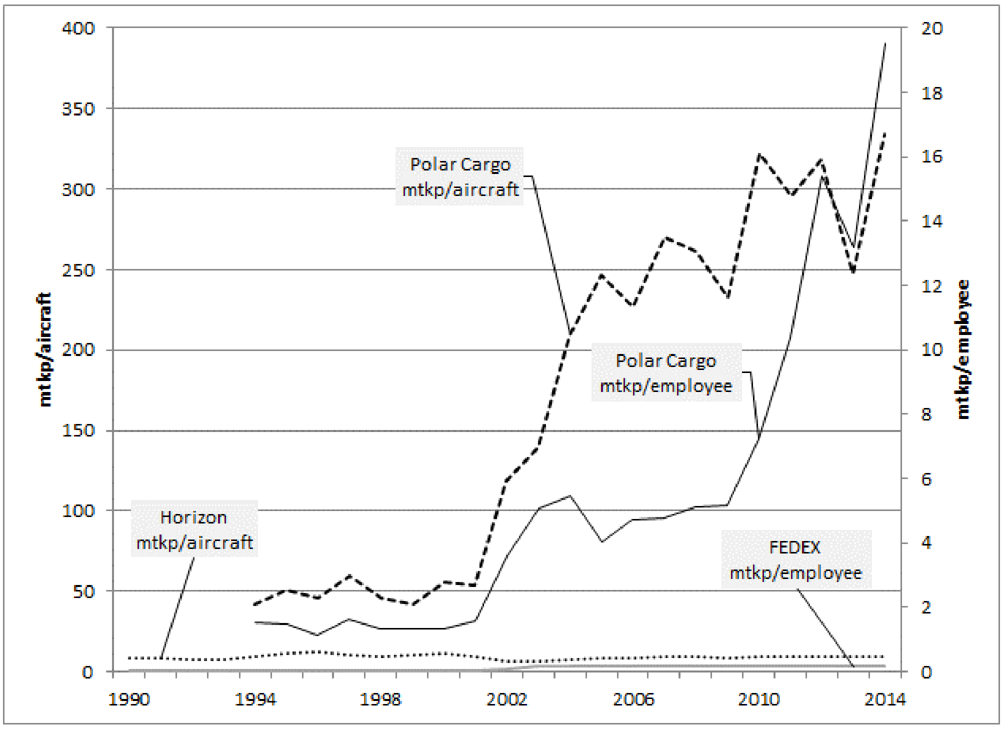


Figure 4.7 Highest and Lowest Average Productivities: Aircraft and Employees

The disparities may lie in the data, as it is hard to imagine two freight-based airlines in close competition with each other having such a large productivity gap. However, it certainly appears that the data will be able to sort the airlines in the Panel into one, or more at or close to the Stochastic Frontier, while others will be located rather far from it. Whether or not, this is the case will become apparent in section **4.5**.

4.4.2 The Econometric Strategy

The Panel will be used to estimate stochastic production frontiers for the US Airlines. In keeping with the common theme of this Thesis, the data will be used as both a cross-section and as a Panel and applied to both Cobb-Douglas and Translog specifications. Stata embodies both **[4.11]** for pooled data, and its Panel-data equivalent, **[4.12**]

$$\boldsymbol{ln(q_{it}) = \beta_0 + \left(\sum_{j=1}^{k} \beta_j ln(x_{jit})\right) + (v_{it} - u_{it})} \qquad [\mathbf{4.12}]$$

-in its *frontier* and *xtfrontier* routines. These will be used in conjunction with the assembled data to estimate and compare the efficiencies from both the pooled and panel estimators.

In the cross-sectional case, Stata allows either half-normal or exponential distributions for *u* and these variants will be estimated for comparative purposes. Stata also permits the assumption of heteroskedasticity in the error terms. Given the huge variation in the size of the airlines in the dataset, it is reasonable to assume that this will be an issue, at least with the non-panel specifications.

While the 'ideal' model would subdivide labour and the aircraft into its main sub-groupings (pilots, cabin crew, maintenance: large- medium- short-haul, etc) the Study only has combined totals for each. Arguably, also, other physical capital (maintenance equipment, computer systems) should be included as well, but the only likely source of this would be the airlines' financial statements. It is unlikely that these will be expressed as anything other than their cost-minus-depreciation amounts and, thus, of no value to this Study. As noted earlier, the annual percentages of each airline's output that emanated from passenger services were estimated. The chosen output variable, tonne kilometres performed is estimated from the FAA's published data, based on statutory returns made to it by the airlines, using **[4.3]**. What is made available to users are subtotals of the total payload in Imperial pounds ('Troy') carried on each of each airline's routes, broken down into passengers and cargo (freight + mail). Pounds-Troy were converted to metric tonnes (dividing by 2204.6) and miles to kilometres (multiplying by 1.60934) to yield tonne-kilometres performed on each route by each airline: this is ICAO's preferred measure of output and is used by most analysts. [66] These were then summed for each airline to give (i) the total tonne kilometres performed in a calendar year, (ii) how much of this was passengers, and (iii) how much was for freight and mail. A variable, 'Passenger Mix (%)' was then calculated by dividing (ii) by (i).

It may have been tempting to include a dummy variable for 'Cargo', as opposed to 'Passenger' airlines in the regressions, particularly as it is relatively easy for freight to generate tkp (space saved on seating, ability to more densely-packed freight and mail compared to the space passengers require). However, the exercise that generated the airlines' tkps for this Study revealed that almost all non-freight airlines in the Panel generated some of their tkp from freight and that even airlines that were overtly cargo-carriers (such as Polar Air Cargo and United Parcel Service) derived some tkp from charter-passenger services. Thus, the generation of a 'cargo' dummy would require an assumption to be made as to the point at which an airline switches between being 'Passenger' and 'Cargo' (50%?). Also, any inherent (dis)advantage from being a cargo operator can be captured better with the continuous 'passenger-mix' variable than a discrete dummy variable could.

Versions of the models with that included as a regressor, alongside Capital and Labour are also estimated, the logic being that it may be fundamentally easier, or more difficult to generate tkp from freight, as opposed to passenger services. One clue to this might be from **Figure 4.7** above, where the cargo airlines Polar Air and Federal Express featured as 'extremes' in the data. Whether or not this has any bearing on the models, in particular the coefficients associated with capital and labour will be apparent from the results.

Finally, on variables, year dummies are used throughout: as the narrative above has demonstrated, the airline industry is constantly evolving, with new generations of more fuel-efficient, reliable aircraft coming into service almost continually, low-cost

[66] The alternative measure *tonne kilometres available ('tka')* can be derived on a similar basis to '*tkp*' but assumes that all aircraft are always filled to their maximum payload. This is a measure of *potential*, rather than *actual* output. The judgement here is that *tkp* is a better output measure from which to estimate efficiency, as it implicitly models the ability of the airline's management to match the demand for its services to its ability to provide appropriate capacity to meet it.

operators entering and growing rapidly while long-established incumbents either go bankrupt or as swallowed up by other incumbents. One consequence of the rise of the low-cost carriers has been the decoupling of many elements of an air journey that were once bundled together into the ticket cost. Meals and drinks were the default on all except the shortest of flights: a generous baggage allowance was granted to all passengers, with only those emigrating with their possessions having to pay for any extra weight. Long-haul (six to 12 hours) and ultra-long-haul flights (> 12 hours) were made easier with 'free' toiletries, flight slippers and in-flight entertainment. These were standard even in Economy ('Coach' or 'Tourist') class, with First Class having more, and better-quality versions and improved seating and ambience. Many of these have now become chargeable extras, while airlines have sought to 'segment the cabin', with different levels of service and, of course, price: 'Economy, 'Economy-plus', 'Business Class'; 'Presidential Class'. Even the on-ground experience has been differentiated, with 'Executive Lounges' for the highest-fare passengers, priority security checks, boarding and disembarkation available, at a price. These developments, and a multiplicity of other factors (such as the intensity of Security and the way ticketing is handled), many unobserved or unobservable, have undoubtedly made the quality of the air traveller's experience rather different with the passing of time. For those that can afford it, the experience is probably no worse than it was for their first-class predecessors in the 1960s and 1970s. For those who can't, the experience is, on the whole, rather less pleasant than it once was (as the Author can attest, having flown regularly since 1975). The output measure 'tkp' simply measures the movement of mass (passengers and cargo) through the air from origin to destination and does not objectively measure the quality of the experience nor, therefore the utility being derived. It must be the case, however, that a closer alignment of facilities and ability to pay should make it easier for an airline in, say 2014, to generate tkp that an airline (even itself) in 1990. The most obvious example would be the space freed up in aircraft Holds as passengers save money by (a) carrying less and, (b) taking what they are bringing with them, into the Cabin. The space freed up can then be filled with cargo and generate more revenue (and of more relevance in this Study), payload. The onus is therefore on the year dummies to absorb these and all the other environmental and technological changes that have taken place (in the US) over the Study Period.

Sample Selection and Attrition Bias. In the quantile wage returns Study (Chapter Three) sample selection bias referred to the non-random decision taken by individuals to work for money. Non-random absence from the wage earners' dataset created the risk of the results from the regressions being biased, necessitating a Heckman-style correction. Selection bias in this Study, is rather different, and refers to the possibility that any skewing of the selected sample emanates from one or more distinct sub-groups not being included in either or both the FAA- and Airfleet's databases. The FAA data is effectively a census, containing the 500+ airlines licensed by it, to operate in the US. Of those, 240 were included in the Airfleet's website, of which 113 had at least one full set of observations. As described in the previous section, the assembled Panel spanned the full range of size, from mega carriers, like Delta and American with 19 – 45 billion tkp per annum each, through to the medium-sized carriers such as Allegiant with ~100 million tkp, down to the single-aircraft operators like Primaris with 1.8 million ktp. In the absence of knowing what the populations of airlines over the Study Period actually looked like, it can never be claimed conclusively that the samples assembled into the Panel are truly representative, but the range of sizes and geographical distribution does indicate a degree of representativeness.

Turning now to *attrition*. As outlined earlier in section **4.2**, there has never been a State-run airline in the US. The closest that there has even been to such a thing were the favoured Flag Carriers, such Pan-Am and TWA, and more recently, American, United, and Delta Airways. However, with the collapse of Pan-Am (1991) and TWA (2001), it is clear that no US airline was ever 'too big to fail' and that the State will not go to the wire to keep them flying, in the same way that a European

Government might have been. [67] Thus, the risk of attrition via bankruptcy is ever-present in the United States. To this can be added the business-normal practices of merger and takeover (such as between United and Continental in 2012) as well as the setting up of brand new airlines. There is also the risk of operating licence revocation by the FAA for operating misdemeanours. Thus, there are various non-random sources of attrition from the Panel: of the 500+ airlines in the entre Panel, 212 were not ever-present over the period 1990-2014. Of those, 54 (28%) were declared bankrupt within the period, followed by another 18 after 2014; eight merged with each other; 107 (55%) were started after 1990 (of which 47 went bankrupt on, or before 2014), two had their operating licence revoked by the FAA for 'operational misdemeanours' and 23 (12%) left for no known reason (probably also bankruptcy). By any measure, this is a significant level of creation and attrition, and implies that the Panel is significantly *unbalanced*. As in the Quantile Returns' Study in the previous chapter, attrition weights will be computed using the same method (after Becketti, et.al, 1988) and the Panels' models estimated with- and without them.

Models' testing will be restricted to the overall significance statistics (R^2; Wald, log-likelihood), the significance of individual parameters and whether or not there appears to be evidence of any technical inefficiency amongst the airlines. This is a test of the null hypothesis that $\sigma_u^2 = 0$, using a modified likelihood ratio test (after Gutierrez, et al (2001)) which Stata only reports for the cross-sectional models. At the risk of drawing the wrong conclusions from the panel models, in this respect, the rejection, or otherwise of this Null will be taken to apply to them as well. Finally, Stata can test the hypothesis that airline 'size' and the variance of the random, idiosyncratic term are correlated and therefore, that the correction for

[67] The French Government injected 20 billion French Francs (~£2bllion) into Air France in 1994, having been given permission by the EU Commission to do so. However, this was ruled to be in breach of the EU's own competition rules, though the airline was not required to repay the money. However, the clear message that applies to this day, is that EU states have to allow their Flag Carrier to fail should circumstances dictate.

heteroskedasticity was necessary. [68] This test is derived from an auxiliary regression on a variable that is assumed encapsulates the size of the unit (in this case, the airline). Accordingly, the dependent variable (mktp) and a constant term were used.

A related topic of interest is whether there was evidence of increasing or diminishing returns to scale from the Production Functions. In the case of the Cobb-Douglas function, this is joint test of the Null hypothesis that the sum of the coefficients is equal to one (constant returns). The situation with the Translog functions is rather more nuanced than this, in the sense that the return is dependent upon the scale of the airline itself. It is possible, therefore, that two airlines of significantly different scale could be experiencing completely different returns to that scale. In order to assess and compare the returns, an Excel Spreadsheet was devised to compute the percentage changes in tkp for the airlines observed in 2014, in response to a simulated 1% increase in aircraft and employees, or both these plus passenger percentage. The results of that exercise will be reported in the results' tables in the next section.

4.5. Results

4.5.1. Introduction

This section considers the results from the various models. **4.5.2** considers the results with no attrition correction for the Cobb-Douglas models, comparing the Panel against the various cross-sectional models (OLS; Frontiers assuming half-normal and exponential distribution) and the models assuming heteroskedasticity, focussing on nature of the returns to scale, and the presence and extent of significant technical (in)efficiency among the airlines include in the model. Section **4.5.3** does the same for the Translog models noting any differences in the returns to

[68] There appears to be no equivalent correction available in Stata's *xtfrontier* routine at the time of writing. From a perusal of the on-line Stata Users' Group's conversations about this, the only remedy available is to estimate an LSDV model, cranking up the general constant β_0 until all the errors, bar one are negative. This is not done here, so the Panel results reported in section **5.7** will be 'moot' in this respect.

scale estimates and efficiencies between the equivalent Cobb-Douglas models. Section **4.5.4** considers the effect of the attrition correction on the Panel-based Cobb-Douglas and Translog models, particularly on the returns to scale. Section **4.5.5** considers whether each airline's ranking is affected by Production Function specification, input variables included, assumptions made about the error terms and whether or not a Panel- or Pooled estimator is used to estimate the location of the Efficient Frontier.

4.5.2. Cobb-Douglas Models: Panel v Cross sections

Table 4.4 summarises the results for the Cobb-Douglas specifications, for the Panel, OLS and Cross-sectional Frontier models. All the models are significant overall, on the Panel-based Passenger-Mix variable is insignificant. Aircraft and Labour parameters are positive to at least 0.001 significance. Apart from in the Panel model, the Passenger-Mix variable is negative impact, implying that shifting service away from cargo to passenger-carrying counterfactually reduces output. This accords with the *apriori* expectation that it is relatively easier to generate tkp with weight-dense cargo, rather than with passengers, who require cabin crew and other services. The insignificance of the passenger-mix variable in the Panel specification may result from the absorption of the particular effect of being a 'cargo' operator', predominantly, into the fixed effects of the relevant airlines (e.g. Polar Air Cargo and FEDEX).

There is evidence of significant technical inefficiency amongst the airlines in the sample, implying that a Frontier does exist. As might be expected, there appears to be heteroskedasticity present in the error terms, and the parameters associated with the inputs alter to some extent. This implies that the returns to scale without the Passenger-Mix variable in the heteroskedastic model are *increasing*, while those not assuming heteroskedasticity, are *decreasing*. The Panel-specifications conclude that returns are *constant*.

Table 4.4

US Airlines: Cobb Douglas Stochastic Frontier Production Function (no attrition correction)

Dependent Variable: ln(Millions of tonne-kilometres performed per annum)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Panel		OLS		Half Normal		Exponential		Heteroskedastic	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP	KL	KLP
ln(Aircraft)	0.422***	0.427***	0.253***	0.283***	0.256***	0.241***	0.265***	0.190***	0.491***	0.257***
ln(Employees)	0.598***	0.590***	0.743***	0.785***	0.719***	0.763***	0.707***	0.838***	0.597***	0.747***
ln(Passenger Mix)	-	0.099	-	-0.764***	-	-1.070***	-	-1.171***	-	-1.035***
Y_1990	-0.867***	-0.859***	-0.771**	-0.750***	-0.798***	-0.742***	-0.804***	-0.787***	-0.357**	-0.770***
Y_1991	-0.775***	-0.771***	-0.675**	-0.662**	-0.706**	-0.712***	-0.723**	-0.740***	-0.340**	-0.775***
Y_1992	-0.570***	-0.567***	-0.442	-0.425*	-0.480*	-0.588***	-0.495*	-0.618***	-0.241**	-0.707***
Y_1993	-0.616***	-0.612***	-0.508*	-0.478*	-0.537*	-0.643***	-0.553*	-0.662***	-0.232**	-0.808***
Y_1994	-0.598***	-0.590***	-0.435*	-0.466*	-0.461*	-0.585***	-0.472*	-0.586***	-0.191*	-0.778***
Y_1995	-0.583***	-0.573***	-0.389	-0.424*	-0.430*	-0.540***	-0.450*	-0.558***	-0.161*	-0.753***
Y_1996	-0.524***	-0.514***	-0.232	-0.271	-0.281	-0.497**	-0.295	-0.512***	-0.097	-0.622***
Y_1997	-0.602***	-0.591***	-0.346	-0.397*	-0.359	-0.505**	-0.349	-0.494**	-0.078	-0.631***
Y_1998	-0.564***	-0.555***	-0.319	-0.374	-0.341	-0.518**	-0.345	-0.518***	-0.103	-0.699***
Y_1999	-0.471***	-0.466***	-0.213	-0.218	-0.252	-0.411*	-0.271	-0.446**	-0.114	-0.626***
Y_2000	-0.413***	-0.409***	-0.256	-0.246	-0.294	-0.384*	-0.313	-0.395**	-0.110*	-0.628***
Y_2001	-0.293*	-0.282*	-0.078	-0.158	-0.120	-0.342*	-0.145	-0.312*	-0.140*	-0.295**
Y_2002	-0.240*	-0.240*	-0.093	-0.079	-0.135	-0.258	-0.164	-0.281	-0.196***	-0.428***
Y_2003	-0.038	-0.035	0.113	0.119	0.075	-0.052	0.044	-0.108	-0.136*	-0.237
Y_2004	0.022	0.029	0.161	0.131	0.136	-0.026	0.112	-0.093	-0.046	-0.263
Y_2005	0.093	0.100	0.205	0.175	0.170	-0.075	0.135	-0.118	0.010	-0.186
Y_2006	-0.011	-0.006	0.094	0.041	0.070	-0.101	0.040	-0.128	0.030	-0.248*
Y_2007	-0.107	-0.101	-0.064	-0.089	-0.067	-0.216	-0.071	-0.220	0.017	-0.304**

Table 4.4 (continued)

US Airlines: Cobb Douglas Stochastic Frontier Production Function (no attrition correction)

Dependent Variable: ln(Millions of tonne-kilometres performed per annum)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Panel		OLS		Half Normal		Exponential		Heteroskedastic	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP	KL	KLP
Y_2008	-0.129	-0.124	-0.071	-0.091	-0.108	-0.279*	-0.136	-0.283*	0.009	-0.356***
Y_2009	-0.114	-0.113	-0.107	-0.108	-0.127	-0.303*	-0.128	-0.265*	-0.042	-0.377***
Y_2010	0.027	0.029	-0.005	-0.028	-0.028	-0.255	-0.040	-0.219	0.021	-0.387**
Y_2011	0.021	0.025	0.022	-0.017	-0.005	-0.232	-0.022	-0.205	0.004	-0.349**
Y_2012	-0.021	-0.017	-0.026	-0.058	-0.038	-0.150	-0.048	-0.153	0.013	-0.299*
Y_2013	-0.071	-0.068	-0.054	-0.090	-0.070	-0.188	-0.082	-0.179	0.002	-0.319**
Constant	3.972*	4.227*	0.265	-0.395*	1.171***	1.026***	1.046***	0.371*	0.692***	1.088***
Observations (Airlines)	1207 (113)									
R^2 / F-statistic	-		0.7709 / 152.67***	0.8093 / 185.35***	-					
Wald (X^2)	1659.978***	1657.671***	-		3096.415***	5033.448***	3549.472***	6665.056***	15019.1***	5610.51***
Log-likelihood	-1140.983	-1140.098	-1699.463	-1588.525	-1696.226	-1501.317	-1691.107	-1473.574	-1551.845	-1451.932
Returns to Scale	Constant			Increasing	Constant	Decreasing	Constant	Decreasing	Increasing	Decreasing
Significant Inefficiency? $\left(\overline{X}^2\right)$ LR	-				6.47**	1.7e02***	16.71***	2.3e02***	-	
Heteroskedasticity? (Z)	-								-24.32***	-5.19***

Stata allows the technical efficiencies for each producer to be estimated using Battese and Coelli's method (1995), where technical efficiency for producer *i* (TE_i) is defined as:

$$TE_i = e^{-u_i} \qquad (0 \leq TE_i \leq 1)$$

where 0 =>complete inefficiency and 1 => 'perfect' efficiency, i.e. on the Frontier.

Figure 4.8 (above) plots these for the Cobb-Douglas models. The OLS plots are, of course, the *total* residuals and only indicators of efficiency if it proved not possible to reject the Null Hypothesis of no differences between the airlines in this respect. The Half-Normal and Exponential models indicate that smaller airlines span the full range of efficiency-to-inefficiency, but that this settles to a factor of 0.6 – 0.7 around one-billion tkp. Even though the errors from the models are most certainly heteroskedastic, the inclusion of the Passenger-mix variable yields a similar result to those assuming homoskedasticity, though with a greater dispersion across the tkp range.

The picture, however, is rather different with the Panel models. The range of efficiencies is rather tighter, peaking at 0.5. While they do agree with the Pooled models that smaller airlines span the full range of (relative) efficiency-to-inefficiency, the larger operators are, perforce highly *inefficient* (round about 0.1). This contrasts with the Pooled models that indicate that the larger airlines coalesce around an average efficiency of 0.6 – 0.7 and certainly not down at 0.1.

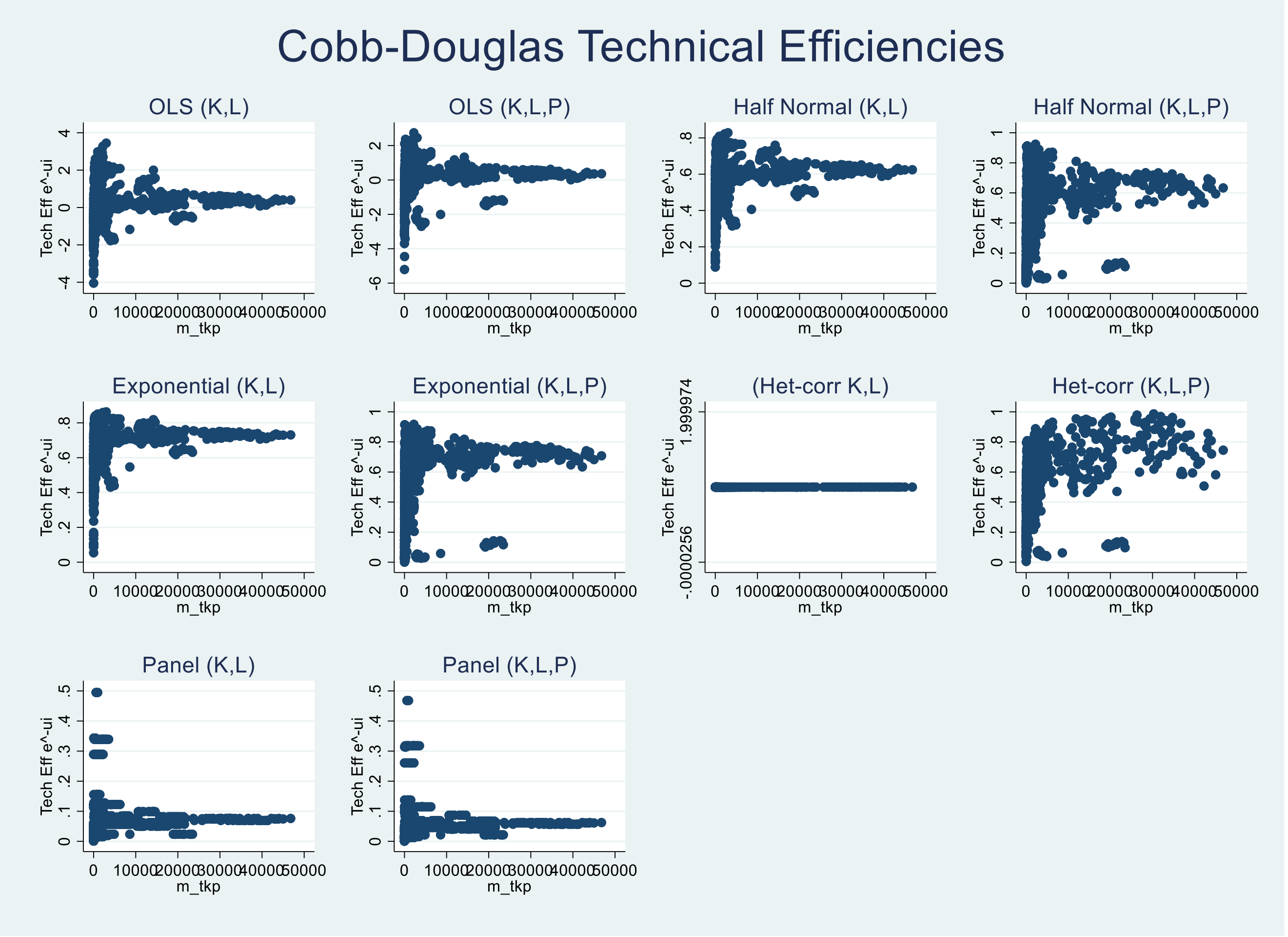


Figure 4.8 Technical Efficiencies: Cobb-Douglas models (No attrition correction)

4.5.3. Translog Models: Panel versus Cross sections

Table 4.5 summarises the results for the Translog specifications, for the Panel, OLS and Cross-sectional frontier models, assuming half-normal and heteroskedastic errors, respectively. As with the Cobb-Douglas models, they are all generally significant, and able to reject that Null that there is no significant technical inefficiency, and also that there is significant heteroskedasticity in the error terms. The significance of most of the quadratic- and interacted factor terms indicate both non-linearities in the relationship between the inputs and output, as well as symbiosis between the factors. Thus, this may be an inherently more realistic modelling of airlines' operations: as pointed out earlier: it is simply not possible to operate an airline without aircraft *and* flight/ground crew.

All the models are able to reject the Null Hypothesis of constant returns to scale, with the Panel specifications indicating *increasing* returns to scale, while the cross-sectional models indicate that they are *decreasing*. [69] This contrasts with the more mixed picture from the Cobb-Douglas models and indicate that the fixed effects, in this case, contribute to alter completely the nature of the returns to scale in the US Industry.

[69] Recalling that the returns with Translog models had to be evaluated at the Airlines' outputs, it proved to be the case that *all* airlines in the sample exhibited either increasing, or decreasing returns, depending on the model applied.

Table 4.5 US Airlines: Translog Stochastic Frontier Production Function (no attrition correction) Dependent Variable: ln(Millions of tonne-kilometres performed per annum) $^{*} p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$										
	Panel		OLS		Half Normal		Exponential		Heteroskedastic	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP	KL	KLP
ln(Aircraft)	0.468**	0.460**	-0.611***	-0.236	-0.527**	0.060	-0.466**	0.018	-0.711***	-0.187
ln(Employees)	0.950***	0.975***	1.147***	1.498***	1.007***	1.104***	0.903***	1.141***	1.687***	1.986***
ln(Aircraft)^2	0.021	0.040*	-0.106***	-0.034	-0.098***	-0.001	-0.092***	-0.003	-0.114***	-0.004
ln(Employees)^2	-0.022	-0.005	-0.073***	-0.040*	-0.061***	-0.009	-0.052**	-0.010	-0.122***	-0.055***
ln(Aircraft)*ln(Employees)	-0.020	-0.052	0.205***	0.067*	0.186***	0.006	0.174***	0.009	0.244***	0.026
ln(Passenger Mix)		-3.470***		-4.757***		-4.551***		-4.671***		-5.950***
ln(Passenger Mix)^2		-0.287***		-0.372***		-0.376***		-0.378***		-0.430***
ln(Aircraft)*ln Passenger Mix))		-0.281***		-0.328***		-0.325***		-0.346***		-0.692***
ln(Employees)* ln(Passenger Mix)		0.513***		0.623***		0.580***		0.605***		0.891***
Y_1990	-0.857***	-0.810***	-0.795***	-0.876***	-0.836***	-0.839***	-0.846***	-0.855***	-0.280*	-0.526***
Y_1991	-0.771***	-0.745***	-0.671**	-0.803***	-0.724**	-0.767***	-0.750***	-0.774***	-0.272**	-0.530***
Y_1992	-0.566***	-0.532***	-0.441	-0.583**	-0.496*	-0.612***	-0.519*	-0.639***	-0.191*	-0.456***
Y_1993	-0.610***	-0.558***	-0.514*	-0.593**	-0.556*	-0.621***	-0.577**	-0.646***	-0.200**	-0.388***
Y_1994	-0.588***	-0.509***	-0.452*	-0.526**	-0.489*	-0.571***	-0.502*	-0.584***	-0.174*	-0.325***
Y_1995	-0.575***	-0.495***	-0.381	-0.460**	-0.426*	-0.537***	-0.450*	-0.561***	-0.142*	-0.313***

Table 4.5 (continued)

US Airlines: Translog Stochastic Frontier Production Function (no attrition correction)

Dependent Variable: ln(Millions of tonne-kilometres performed per annum)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Panel		OLS		Half Normal		Exponential		Heteroskedastic	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP	KL	KLP
Y_1996	-0.513***	-0.460***	-0.256	-0.381*	-0.309	-0.465**	-0.325	-0.493***	-0.083	-0.249***
Y_1997	-0.590***	-0.528***	-0.361	-0.494**	-0.383	-0.488**	-0.374	-0.480**	-0.061	-0.213***
Y_1998	-0.555***	-0.524***	-0.315	-0.504**	-0.344	-0.523***	-0.348	-0.518***	-0.076	-0.304***
Y_1999	-0.457***	-0.419***	-0.224	-0.358*	-0.271	-0.422**	-0.293	-0.454**	-0.089	-0.305***
Y_2000	-0.397**	-0.311**	-0.277	-0.336	-0.324	-0.377*	-0.342	-0.412**	-0.087	-0.250***
Y_2001	-0.281*	-0.151	-0.110	-0.134	-0.160	-0.217	-0.188	-0.255	-0.125*	-0.257***
Y_2002	-0.239*	-0.174	-0.123	-0.151	-0.171	-0.197	-0.205	-0.222	-0.201***	-0.252***
Y_2003	-0.044	-0.003	0.090	0.037	0.042	-0.043	0.003	-0.077	-0.140**	-0.131**
Y_2004	0.021	0.047	0.085	0.033	0.053	-0.016	0.036	-0.069	-0.053	-0.065
Y_2005	0.084	0.114	0.159	0.078	0.118	-0.056	0.086	-0.088	0.004	-0.036
Y_2006	-0.019	-0.005	0.054	-0.053	0.020	-0.116	-0.009	-0.138	0.032	0.011
Y_2007	-0.112	-0.116	-0.114	-0.176	-0.124	-0.222	-0.121	-0.228	0.024	0.013
Y_2008	-0.136	-0.143	-0.099	-0.154	-0.139	-0.247	-0.166	-0.246	0.022	0.024
Y_2009	-0.126	-0.119	-0.122	-0.146	-0.153	-0.248	-0.156	-0.232	-0.025	0.009
Y_2010	0.015	0.025	-0.022	-0.082	-0.053	-0.187	-0.064	-0.174	0.024	0.013
Y_2011	0.010	0.002	0.005	-0.065	-0.029	-0.185	-0.045	-0.172	0.009	-0.028
Y_2012	-0.020	-0.038	-0.068	-0.119	-0.087	-0.168	-0.094	-0.158	0.014	-0.082***
Y_2013	-0.078	-0.093	-0.080	-0.145	-0.103	-0.211	-0.116	-0.197	0.002	0.003
Constant	2.947	0.267	0.286	-3.038***	1.477**	-0.770	1.632**	-1.107**	-1.403**	-5.171***

Table 4.5 (continued)
US Airlines: Translog Stochastic Frontier Production Function (no attrition correction)
Dependent Variable: ln(Millions of tonne-kilometres performed per annum)
*** $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$**

	Panel		OLS		Half Normal		Exponential		Heteroskedastic	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP	KL	KLP
Observations (Airlines)	1207 (113)									
R^2 / F-statistic	-		0.7782/ 142.43***	0.8517/ 204.17***	-					
Wald (X^2)	1667.0***	1727.4***	-		2888.8***	5165.3***	3362.0***	6300.8***	15481.5***	20924.3***
Log-likelihood	-1137.864	-1093.057	-1679.707	-1436.787	-1677.251	-1396.958	-1672.445	-1378.532	-1519.011	-1141.153
Returns to Scale	Increasing		Decreasing							
Significant Inefficiency? $\left(\overline{X}^2\right)$ LR	-				4.91*	79.66***	14.52***	1.2e02***	-	
Heteroskedasticity? (Z)	-								-23.74***	-11.00***

Figure 4.9 plots the Technical efficiencies implied by the models summarised in **Table 4.5**. These mirror closely the patterns seen with the Cobb-Douglas models (**Figure 4.8**), with most of the cross-sectional models indicating that smaller airlines span the full (in)efficiency range, with larger operators being more homogenous in this respect, but all operating off the frontier. Only the heteroskedastic-corrected models appear to conclude that only the larger operators (> 2.5 billion tkp) can be located near the frontier, though not all are.

The Panel models indicate that only a small airline has any hope of being at the frontier, with all (Aircraft + Labour) models concluding that those operators with > 1 billion tkp are certain to be relatively inefficient.

The next section will consider the impact of the attrition correction on the Panel-based models.

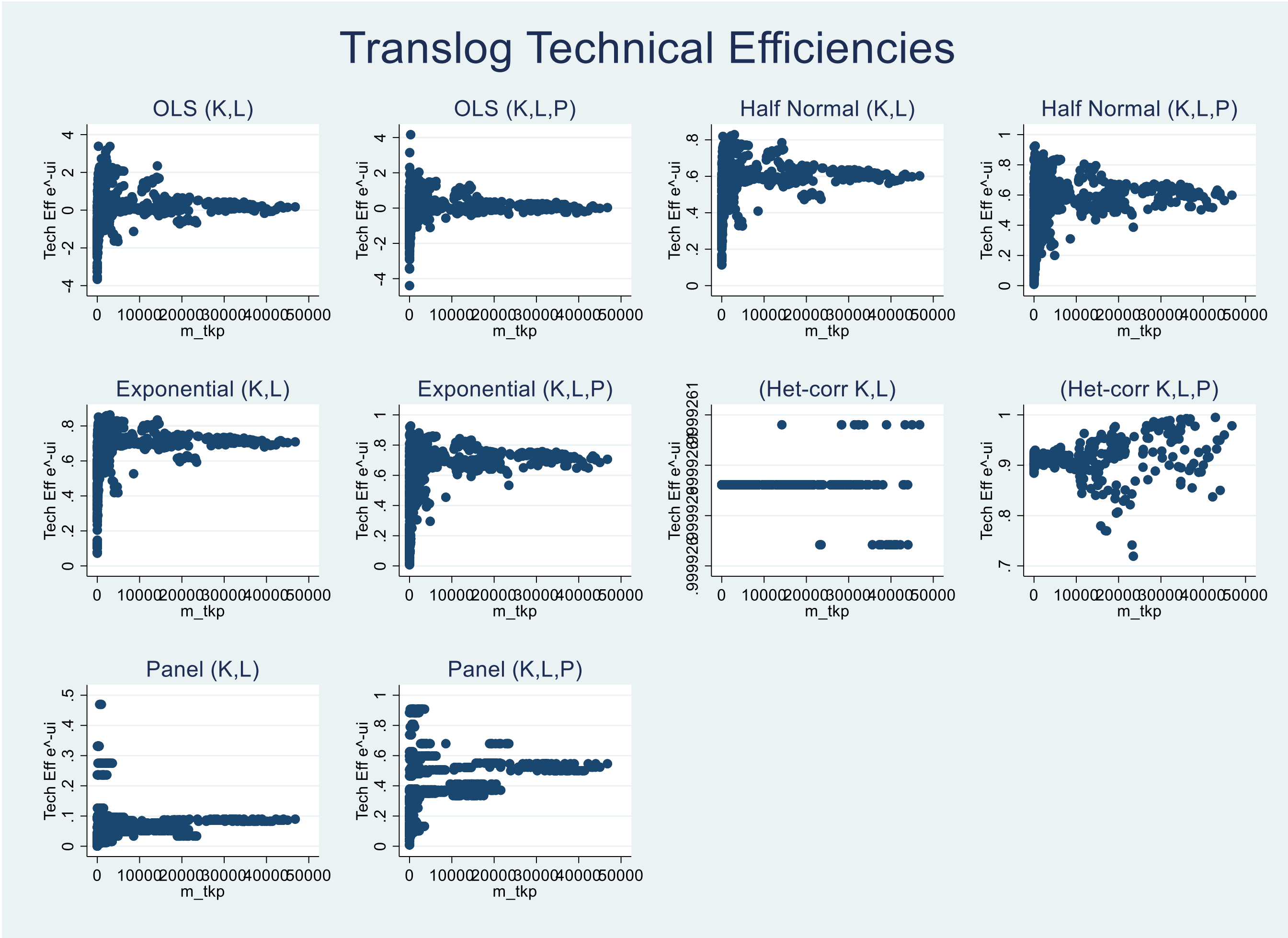


Figure 4.9 Technical Efficiencies: Translog models (No attrition correction)

4.5.4. Panel models, with and without Attrition Correction

As discussed in the section **4.4.2**, it appears that the Panel is significantly unbalanced, via a process of entry and exit (thorough bankruptcy). Following the method of Becketti, et al (1988, op.cit.) a subset of data containing each airlines' *first observation* in the panel was created and an attrition risk score computed. It was assumed that younger airlines generally were more at risk than older ones from going bankrupt (although as the 63-year old Braniff International Airline demonstrated when it ceased operations in 1991, maturity was no guarantee of further longevity). It was also assumed that there was an optimal mix of cargo- and passenger services, where over-reliance on one over the other would reduce an airline's robustness to adverse trading conditions in one of the two distinct markets. Thus, an airline like FEDEX that generates all its tkp from Freight and Braniff that generated all its business from Passenger services, were much more at risk from attrition (by bankruptcy) than an airline like Transcontinental, who generated the sector-average 60% of its tkp from Passenger services in 1999. [70] A quadratic function, with a minimum turning point at 60% passenger mix was assumed and applied to a declining age-factor, to give **[4.13]**.

$$Attriton\_Risk\_Score_{it} = \left(\frac{1000}{Age_{it}}\right)\left(\pi_1\left(P_{it} - \overline{P}\right)^2 - \pi_2\left(P_{it} - \overline{P}\right) \right) \qquad [\mathbf{4.13}]$$

The dependent variable ln(tkp) was then regressed against the independent variables used in the main model, along with the attrition risk score variable, and interactions between it and the other independent variables. An F-test was then performed to jointly test whether or not the attrition variable and the interactions were jointly significantly different to zero. As this tested significant (F = 7.58, p = 0.000) the next stage was to compute the attrition weight, using the same approach in Chapter Three on the wage returns to qualification. This resulted in a set of

[70] In fact, all airlines in the Panel that derived all its tkp from Passenger services went bankrupt during, or after the Study Period.

weights ranging from 0.02 to 5.35, which were included in the Panel regressions, with both capital and labour and capital, labour and passenger-mix variables for both the Cobb-Douglas and Translog specifications. The results, with and without the attrition corrections are in **Table 4.6**.

The parameters for the Cobb-Douglas models change but neither in sign nor significance, and all models indicate *constant* returns to scale. **Figure 4.10** indicates that the attrition correction increases the maximum observed efficiency to 0.5 from 0.8. That said, the general shapes of the distribution are broadly the same, with the larger airlines locating at the inefficient end of the spectrum.

There is also some change in the Translog parameters with one or two of the interacted factor terms moving in/out of statistical significance, though all models find *increasing* returns to scale. **Figure 4.11** confirms similar patterns of efficiencies amongst, though the attrition correction widens the distribution at the top end for the (Aircraft + Labour) version the smaller airlines, which still consigns of the larger airlines to relative inefficiency. However, the effect of the attrition correction with the (Aircraft + Labour + Passenger-mix) version indicates that the larger airlines (>1.5 billion tkp) will locate close to the frontier, though (as before) only small airlines will actually locate very close to it.

Table 4.6

US Airlines: Stochastic Frontier Production Function (Panel specifications)

Dependent Variable: ln(Millions of tonne-kilometres performed per annum) * ***p*** **< 0.05,** ** ***p*** **< 0.01,** *** ***p*** **< 0.001**

	Cobb-Douglas				Translog			
	With Attrition		No Attrition		With Attrition		No Attrition	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP
ln(Aircraft)	0.391***	0.392***	0.422***	0.427***	0.526***	0.537***	0.468**	0.460**
ln(Employees)	0.645***	0.644***	0.598***	0.590***	0.854***	0.807***	0.950***	0.975***
ln(Passenger Mix)	-	0.013	-	0.099	-	-3.023***	-	-3.470***
ln(Aircraft)^2	-	-	-	-	0.016	0.033	0.021	0.040*
ln(Employees)^2	-	-	-	-	-0.011	0.005	-0.022	-0.005
ln(Aircraft)*ln(Employees)	-	-	-	-	-0.028	-0.059*	-0.020	-0.052
ln(Passenger Mix)^2	-	-	-	-	-	-0.236***	-	-0.287***
ln(Aircraft)*ln Passenger Mix))	-	-	-	-	-	-0.278***	-	-0.281***
ln(Employees)* ln(Passenger Mix)	-	-	-	-	-	0.451***	-	0.513***
Y_1990	-0.707***	-0.706***	-0.867***	-0.859***	-0.702***	-0.688***	-0.857***	-0.810***
Y_1991	-0.550***	-0.549***	-0.775***	-0.771***	-0.552***	-0.547***	-0.771***	-0.745***
Y_1992	-0.433***	-0.432***	-0.570***	-0.567***	-0.436***	-0.422***	-0.566***	-0.532***
Y_1993	-0.625***	-0.624***	-0.616***	-0.612***	-0.626***	-0.597***	-0.610***	-0.558***
Y_1994	-0.498***	-0.497***	-0.598***	-0.590***	-0.494***	-0.457***	-0.588***	-0.509***
Y_1995	-0.464***	-0.463***	-0.583***	-0.573***	-0.457***	-0.415***	-0.575***	-0.495***
Y_1996	-0.422***	-0.420***	-0.524***	-0.514***	-0.414***	-0.426***	-0.513***	-0.460***

Table 4.6 (continued)

US Airlines: Stochastic Frontier Production Function (Panel specifications)

Dependent Variable: ln(Millions of tonne-kilometres performed per annum) * ***p* < 0.05,** ** ***p* < 0.01,** *** ***p* < 0.001**

	Cobb-Douglas				Translog			
	With Attrition		No Attrition		With Attrition		No Attrition	
	KL	KLP	KL	KLP	KL	KLP	KL	KLP
Y_1997	-0.353***	-0.352***	-0.602***	-0.591***	-0.342***	-0.323***	-0.590***	-0.528***
Y_1998	-0.331***	-0.330***	-0.564***	-0.555***	-0.319**	-0.300**	-0.555***	-0.524***
Y_1999	-0.256**	-0.255**	-0.471***	-0.466***	-0.242*	-0.217*	-0.457***	-0.419***
Y_2000	-0.331***	-0.329***	-0.413***	-0.409***	-0.316**	-0.264**	-0.397**	-0.311**
Y_2001	-0.129	-0.126	-0.293*	-0.282*	-0.118	-0.024	-0.281*	-0.151
Y_2002	-0.110	-0.110	-0.240*	-0.240*	-0.097	-0.026	-0.239*	-0.174
Y_2003	0.079	0.080	-0.038	-0.035	0.098	0.120	-0.044	-0.003
Y_2004	0.089	0.090	0.022	0.029	0.093	0.111	0.021	0.047
Y_2005	0.140	0.141	0.093	0.100	0.139	0.144	0.084	0.114
Y_2006	0.183	0.184	-0.011	-0.006	0.176	0.156	-0.019	-0.005
Y_2007	0.179	0.180	-0.107	-0.101	0.173	0.122	-0.112	-0.116
Y_2008	0.117	0.118	-0.129	-0.124	0.109	0.072	-0.136	-0.143
Y_2009	0.060	0.060	-0.114	-0.113	0.054	0.049	-0.126	-0.119
Y_2010	0.142	0.142	0.027	0.029	0.136	0.110	0.015	0.025
Y_2011	0.149	0.150	0.021	0.025	0.145	0.104	0.010	0.002
Y_2012	0.058	0.059	-0.021	-0.017	0.060	0.004	-0.020	-0.038
Y_2013	0.014	0.015	-0.071	-0.068	0.011	0.001	-0.078	-0.093
Constant	2.849***	2.873***	3.972*	4.227*	2.157***	0.626	2.947	0.267
Observations (Airlines)	790 (63)		1027 (113)					
Wald (X^2)	2819.496***	2819.212***	1659.978***	1657.671***	2850.287***	3500.177***	1666.959***	1727.382***
Log-likelihood	-914.995	-914.964	-1140.983	-1140.098	-911.131	-855.108	-1137.864	-1093.057
Returns to Scale	Constant				Increasing			

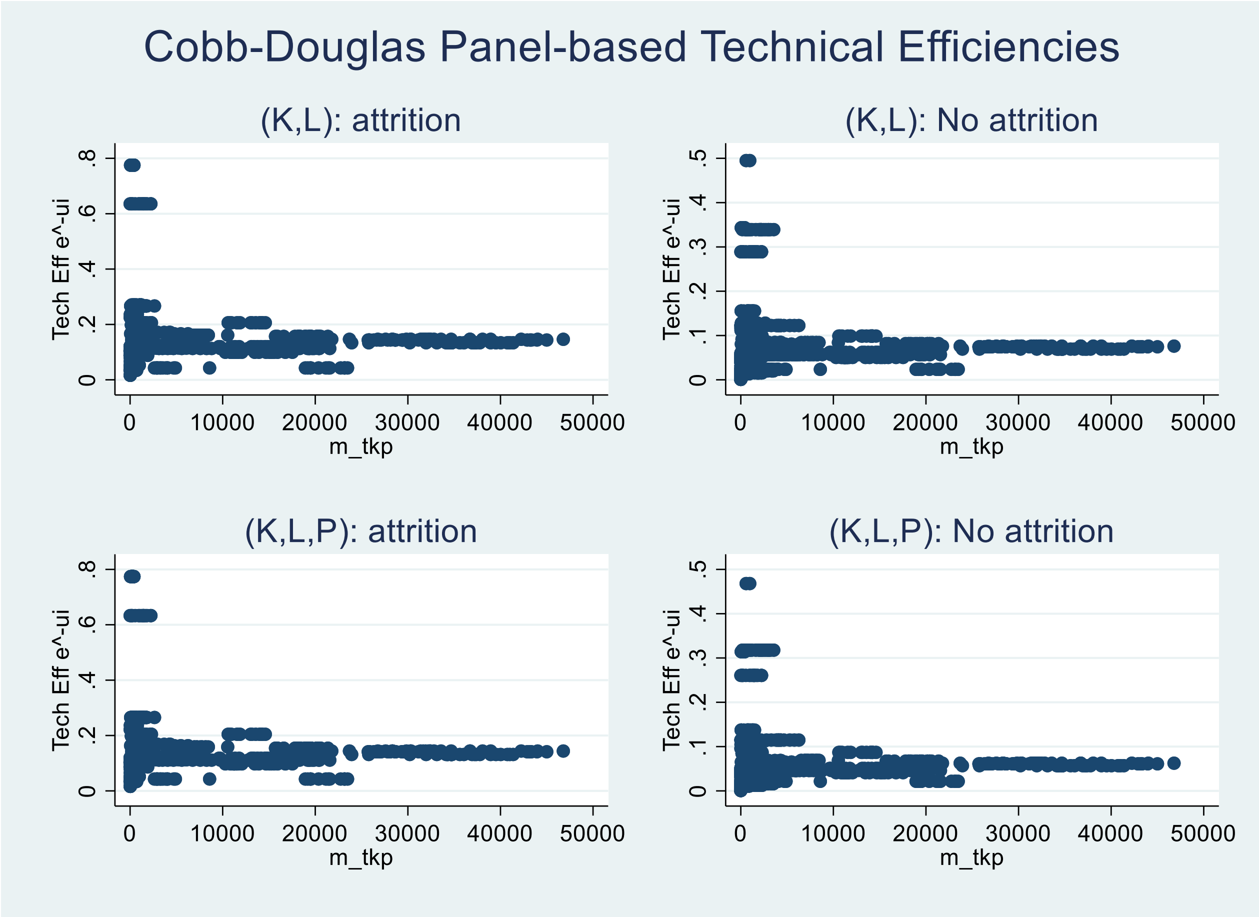


Figure 4.10 Technical Efficiencies: Cobb-Douglas Panel models (with and without attrition correction)

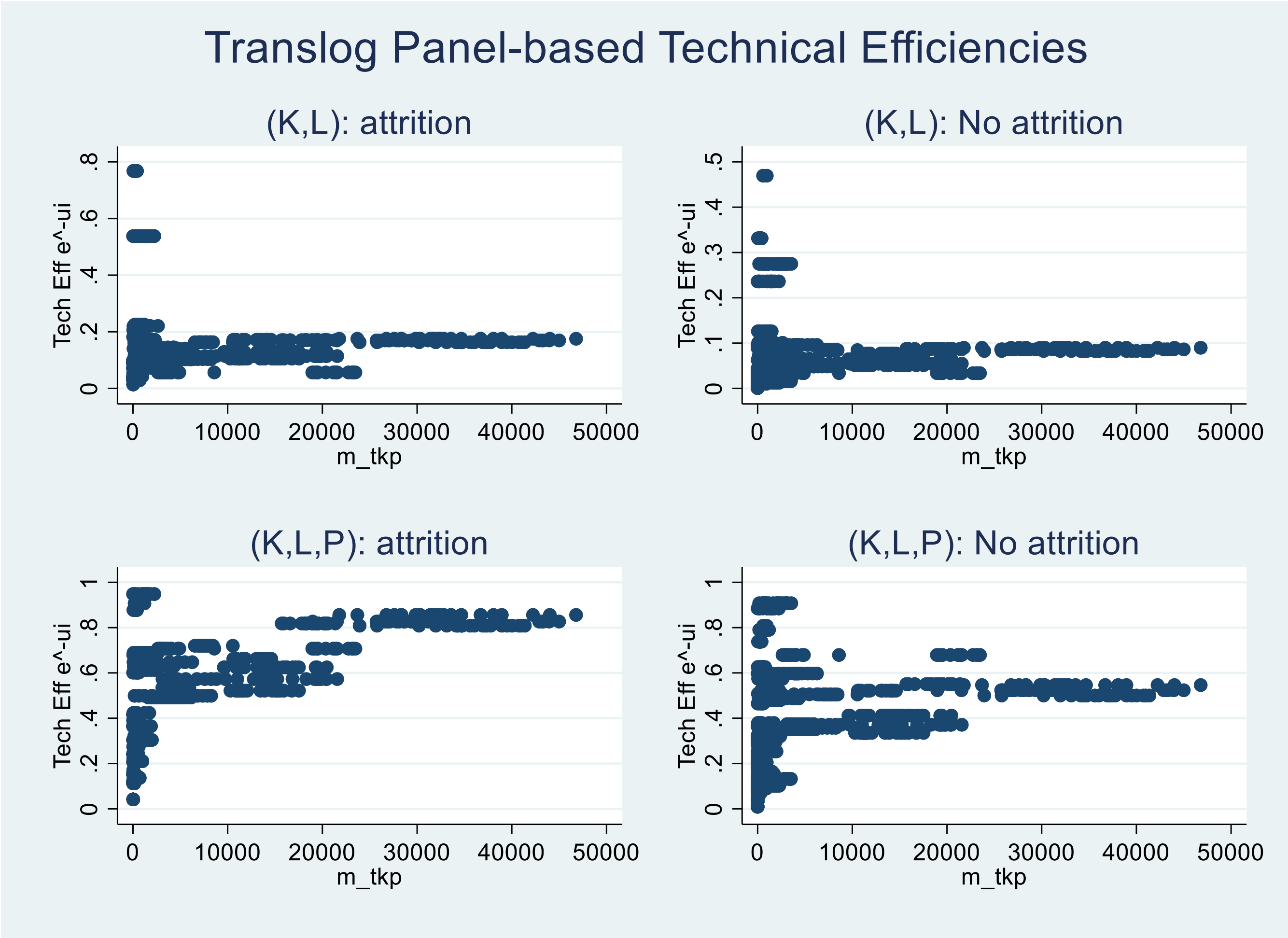


Figure 4.11 Technical Efficiencies: Translog Panel models (with and without attrition correction)

4.5.5 Are there any patterns in the technical efficiencies?

Of course, the apparent similarity in the results plotted in **Figures 4.8** to **4.11** may be illusory, masking a propensity for the different models to move the airlines up and down the ranking, and reducing the robustness of the results. To assess this for the Pooled models, the airlines' all-time [71] technical efficiencies were ranked within each of the models, and Pearson's coefficients of rank correlation (ρ) between each pair of ranks computed. All ρ were statistically significant to at least 0.01, with the minimum ρ = 0.631, the maximum ρ = 0.999 and a mean ρ of 0.841. Only 19 of the 120 model-pairs had a ρ <0.742 (the weaker correlations were between the OLS and Frontier-based models).

When the data are treated as a Panel, each airline is given a *single* efficiency score for the entire period. As with the Pooled model, these were ranked and pair-wise Pearson's ρ computed. All ρ were also significant to at least 0.01, with the minimum ρ =0.668, maximum ρ = 1.000 and mean ρ = 0.826. The weaker correlations were between the Cobb-Douglas and Translog specifications, rather than within a particular specification. These results suggest that the various estimators are, at least, placing the airlines in the same *relative* order in relation to the respective Frontiers. This gives some assurance that the efficiency rankings are, largely independent of the model applied to the data to estimate them.

It might also be supposed that the efficiency of airlines *through time* should improve, assuming the capital (aircraft) is more productive than the labour it is displacing, and that it is more efficient than the capital it is replacing. At the same time, reduced industry regulation, the rise of Alliances, merged airlines and low-cost carriers can be expected to have forced the airlines to make efforts to moving themselves closer to the Frontier. Given the approach taken in this Study, it was only possible to assess the extent of progress through time by assessing the ranks airlines achieved via the Pooled Models at the various points through the Study

[71] i.e. if an airline appeared in each of the 25 years of the Panel, it would have 25 rank scores somewhere in the 1200+ ranks.

Period, 1990 – 2014. By way of illustration, the ranks endowed on the airlines by the (Capital+Labour) model assuming heteroskedasticity were collated for the airlines that were ever-present in the Panel. **Figure 4.12** plots the overall highest-ranked airline (Delta), with the lowest (Federal Express: 'FEDEX') and the annual mean rank between 1990 and 2014. While FEDEX appeared to have worsened its consistently poor rank, the whole-Panel annual means fell through time, as did Delta's after an upward blip in 2003. This may be taken as an indication of time-related technical progress, in the sense that the higher ranks tended to be found towards the end of the Study Period, rather than at the beginning.

The inclusion of the year dummies in **Tables 4.4**, **4.5** and **4.6** also give some insight as to the extent of technical progress over the Study Period (1990 – 2014). There is a fairly clear pattern of a reducing time-penalty as we move from 1990 towards the eb of the Study Period, as the time dummies become less negative, before becoming statistically insignificant. The time at which this happens varies according to the specification assumed. For the Panel-based Cobb-Douglas models with no attrition correction, this was around 2002, while the standard OLS models had this occurring earlier in 1994/97. The Panel-based Translog models had his occurring around 200/02, while the OLS indicate 1994/99. The effect of correcting for attrition on the Panel models is to conclude that this time-based improvement stopped in 2000 rather than 2002.

It is difficult to point to any specific events around the 2000-2002 period that would indicate that a particular phase came to an and (or that one started). As discussed in **4.2.1**, the low-cost carriers were already established, and the incumbents were still losing ground to them: it was not until later in that decade (particularly after the 2008 Global Recession) that they embarked on a programme of mergers and alliance-formation. It may have been the case that prior to those more public responses, that all airlines were 'sweating their assets', introducing new work practices, maintenance schedules, while introducing more modern aircraft, requiring fewer flight crew and time on the ground between flights for example.

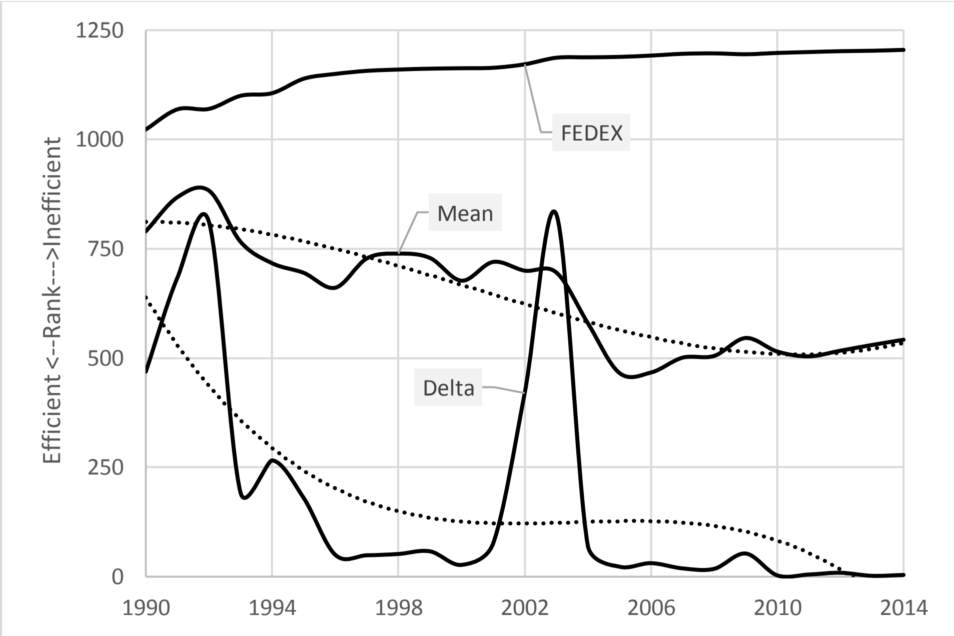


Figure 4.12 All-time Ranks for the top- and bottom ranked airlines

4.6. Discussion & Further work

This chapter has documented the creation of a small, low-cost panel of US Airlines, using publicly-available data, observed over a 25-year period (1990-2014) and the estimation of their technical efficiencies and nature of returns to scale in the industry. A number of different stochastic production frontiers were estimated, based on assumptions about the nature of the error distributions, and comparing Panel-based and Pooled specifications.

It is evident that there are significant differences in technical efficiency between the airlines, and some evidence of an improvement through time, as the competitive conditions in the Industry tightened, as the long-term consequence of the deregulation of the global industry, following the high level of protection it enjoyed in the decades following the Second World War. The nature of the returns to scale in the Industry are specification-dependent, and while the relative rankings in efficiency are robust to the model employed, the Panel models indicate that the larger airlines are extremely inefficient, whereas the Pooled models believe them to be ‘reasonably’ efficient.

While this Study is not the first to employ a Panel to estimate efficiency in the US Airline Industry, it appears to be the first to closely examine the extent of the differences between Panel- and Pooled estimators, and what impact they have on returns and relative efficiency rankings. The Study is limited in geographical scope insofar it is limited to the single largest airline industry in the World, and one that has always been relatively *laissez-faire* at least with respect to its large domestic sector. This, and the location of the Study Period to 1990-2014, a period when many of the technical- and commercialisation initiatives were already well established, mean that the Study does not really reflect the pre- and post-deregulation effects on efficiency. As such, they are limited to the advances that took place post-1990. A follow-up Study would therefore extend itself backward in time, and also increase its geographical scope to include as much of the Developed World, as possible.

The models were, perforce parsimonious in nature: that is to say, only single measures of all Labour and Capital were included, and a simple business-mix variable included in some specifications. Given the heterogeneous nature of airlines' employees, it would be desirable to split these totals out into the different main subgroups, for example Pilots and Cabin Crew; Ground-based maintenance; Administrative and Management. Given the limits on cross-substitutability of these factors (Administrators can't fly aircraft, for example), it would not necessarily expand the number of cross-product variables in the Translog specifications, and refine the returns to each factor, and overall returns in the Industry. The same would be true for the Aircraft: while there is a natural limit applied to how old a fleet can be, an airline that operates newer, more fuel-efficient aircraft than another that uses older-generation ones, close to their retirement age, would be expected to be counterfactually more cost-efficient. An airline that can better tailor its fleet-size mix to the demand conditions on its routes should see its load factors increasing, with the prospect of being able to improve its cost efficiency as well. Given that a refined version of this Study would continue to be about *physical* efficiency, then it is not obvious that it would make any difference if the total were split down into size, or age, or speed categories. However, if the dependent variable were to be total kilometres *available*, as opposed to *performed*, or perhaps the percentage of that available that was actually performed, then this may become a key improvement.

Finally, the related issues of *sample-selection* and *attrition* bias would need to be explored in more depth. This Study has in effect assumed that the sample included in the Panel was representative of the population, taking comfort from the large spread of airline sizes in the Panel. The sample selection risk may not emanate so much from the exclusion of, say, small airlines from the FAA's output database, but more from their exclusion from the Airfleets' capital & labour database. This, bias may have arisen from non-random reporting exclusion. In any case, the inordinate amount of time that had to be devoted to assembling the database, makes it expedient to explore alternatives, such as the ICAO data on Fleets and Workforce. If a follow-up Study were to broaden out into other countries than the US, similar quality output data would have to be available: the lack of this data is what limited this Study to the US. Both financial resource and contacts within the airline industry would have to be established to facilitate this.

More comprehensive data would lessen the attrition through non-availability problem, but not, of course, attrition through the creative-destructive process. The correction attempted here made some bold assumptions about the relationship between airline age, business mix and likelihood of attrition were made. Also, plausible π values were assumed for **[4.13]** to minimise the mix-related attrition risk at 60% (the apparent Industry Average in the Study Period). A follow-up Study would explore more deeply, the causes of attrition through merger and, in particular, bankruptcy and consider, for the first time, *entry* into the Panel as a result of being established as a brand-new venture.

References Used in Chapter Four

Afriat, S, (1972), 'Efficiency Estimation of Production Functions', International Economic Review, 13   568-598

Aigner, D.J. & Chu, S.F., (1968), 'On estimating the industry production function', American Economic Review, 58   226-239

Aigner, D.J., Lovell,C.A. & Schmidt, P., (1977), 'Formulation and estimation of stochastic production function models', J.of Econometrics, 6 1 21-37

Ausick, P., 'How much does a Boeing 747 cost?', Blogpiece for 24/7 Wall Street, 12 January, 2017 (https://247wallst.com/aerospace-defense/2017/01/12/how-much-does-a-boeing-747-cost-2017-may-see-last-flight-on-us-carriers/), accessed 23rd July, 2018

Battese, G E & Coelli, T J, (1995), 'A model for Technical Inefficiency effects in a Stochastic Frontier Production Function for panel data', Empirical Economics, 20, 325-332

Becketti, S., Gould, W., Lillard, L. and Welch, F. (1988), 'The Panel Study of Income Dynamics after fourteen years: an Evaluation', J. Labor Economics, 6, 472-92

Cobb, C. W.; Douglas, P. H. (1928). 'A Theory of Production', American Economic Review. 18 (Supplement): 139–165.

Doganis, R, (1991), 'Flying off-course: the Economics of International Airlines', Routledge

Eads, G.C., *(1975)*, 'Competition in the Domestic Trunk Airline Industry: Too Much or Too Little?' in A.Phillips, ed., Competition and Regulation, Brookings Institution

Farrell, M.J., (1957), 'The measurement of productive efficiency', J.of the Royal Statistical Society, 120   253-266

References Used in Chapter Four (continued)

Førsund, F.R. & Hjalmarsson, L., (1979), 'Frontier Production functions and technical progress: a study of general milk processing in Swedish Dairy Plants', Econometrica, 47, 4 883-900

Forsyth,P.J., Hill,R.D. & Trengrove,C.D., (1986a), 'Measuring airline efficiency', Fiscal Studies, 7 1 61-81

Forsyth,P.J., Hill,R.D. & Trengrove,C.D., (1986b), 'Measuring the productivity of international airlines', IFS Working paper, 85

Greene, W.H., (1980), 'Maximum likelihood estimation of econometric frontier functions', J.of Econometrics, 13 27-56

Hoch, I. (1962), 'Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data', Econometrica, 30, 1

Lee, L F, (1983), 'A test for distributional assumptions for the Stochastic Frontier Functions', J.of Econometrics, 22 245-267

Lijesen, M.G., Nijkamp, P. and Rietveld, P (2002), 'Measuring competition in civil aviation', J.of Air Transport Management, 8 pp189-197

Meussen,W. & van der Broeck,J., (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error', Int.Eco.Review, 18 2 435-444

Manonmani, M., (2013), 'A Stochastic Frontier Production Function approach to (the) Indian Textile Industry', Indian J. of Industrial Relations, 48, 4, 703-710

Mutz, R., Lutz, B. and Hans-Dieter, D. (2017), 'Are there any Frontiers of Research Performance? Efficiency measurement of funded research projects with the Bayesian Stochastic Frontier analysis for Count Data', J.of Informetrics, 11, 3, 613-628

References Used in Chapter Four (continued)

Olson, J.A., Schmidt, P. & Waldman, D.M., (1980), 'A Monte-Carlo study of estimators of Stochastic Frontier production Functions', J.of Econometrics,13 67-82

Richmond, J., (1974), 'Estimating the efficiency of production', Int.Eco.Review, 15 515-521

Schmidt,P., (1985), 'Frontier production functions', Econometrics Reviews, 4 289-328

Schmidt,P., (1976), 'On the statistical estimation of parametric frontier production functions : a rejoinder', Review of Economics & Statistics, 58

Schmidt,P. & Sickles,R.C., (1984), 'Production frontiers and Panel Data', J.of Bus.& Eco.Stats., 2 367-374

Sengupta, A. and Wiggins, S.N. (2014), 'Airline Pricing, Price Dispersion, and Ticket Characteristics on and off the Internet', American Economic Journal: Economic Policy, 6(1): 272–307

Stevens, P.A., (2005), 'A Stochastic Frontier Analysis of English and Welsh Universities', Education Economics, 13,4, 355-374

Stevenson, R., (1980), 'Likelihood functions for generalised stochastic frontier estimation', J.of Econometrics, 13 57-66

Straszheim, R., (1969), 'The International Airline Industry', The Brookings Institution, Washington, DC

References Used in Chapter Four (continued)

Tsoukalas, G., Belobaba, P. and Swelbar, W. (2008), 'Cost convergence in the US airline industry: an analysis of unit costs 1995–2006', J. Air Transport Management, 14, 4, 179-187.

van den Broeck, J, Forsund, F.R., Hjmarllson, L. and Meeusen, W, (1980), 'On the estimation of deterministic and stochastic production functions', J.of Econometrics, 13 26-70

Wheatcroft, S., (1964), 'Air Transport Policy', Michael Joseph Publishers, London

Wicksteed, P, (1894), 'An Essay on the Co-ordination of the Laws of Distribution', MacMillan & Co, London

Yasar, M. and Kiraci, K, (2017) 'Market Share, the number of competitors and concentration: an empirical application on the airline industry', Andalou International Conference in Economics, Eskisehir, Turkey, May.

5. Final Comments, conclusions and observations

The author has engaged in policy-based, applied Economics' research during his career (some of it referenced in Chapter One), and took the opportunity when planning and writing this Thesis to undertake three diverse studies in Economics (and loosely, Labour Economics, insofar as the three may have implications for employment within their associated sectors). The Studies concerned themselves with testing the strength of the evidence in favour certain assertions that might be made, concerning certain causes and effects, using appropriate historic data and econometric models. Of particular interest in these Studies (and the overarching Theme of the Thesis), was the robustness of the findings to using the assembled data as either a standard Cross-section ('Pooled Data') or as a Panel. The essential difference being the identification (or otherwise) of the individual (or 'unit') generating the data when estimating the influence of the independent variables on the dependent variable. Regression-based approaches to modelling aim to identify the statistical significance and magnitude of the effect of each independent variable. However, as any honest modeller should always be willing to admit, the process is always susceptible to a combination of (i) inadequate theory; (ii) inappropriate specification; (iii) the omission of variables, and (iv) significant breaches of the assumptions required by the econometric models. Of course, tackling issue (i) is the primary aim of all the Social Sciences, including Economics, and issues (ii) and (iv) can be tested for and, where possible, corrected. Issue (iii), can also be tested for and, at least, partially addressed by the sourcing and inclusion of additional variables. Nevertheless, it may be that the applied researcher is unable to identify the omitted variables, or that they are unavailable to them or that they are simply incapable of being observed and objectively quantified. The ability to identify a composite 'unobserved' component ('unobserved heterogeneity') is the main (asserted) advantage of a Panel specification. Whether or not these individual effects are significant is a matter of empirical investigation, and not one that can be determined in advance.

Datasets are either deliberately intended to form panels (such as the Understanding Society Survey used in Chapter Three) or can be assembled from otherwise disconnected datasets (such as the Gender Based Violence data used in Chapter Two), or unidentified units can be aggregated into cohorts and a *pseudo-panel* formed. Either way, the applied researcher should always be looking to utilise a panel, even if only to discover that the unobserved component does not exert any significant effect on the parameters.

The three Studies forming this Thesis do share some aspects in common: they are concerned with socio-economic issues of concern, including their potential implications for future employment and aim to model the dependent variable which is likely caused by a large number of independent variables, many/most of whom are unobserved, unobservable, or simply unavailable to this author. That said, they examine very different scenarios, requiring varying degrees of effort to render the available and appropriate data into an amenable Panel. Had the three Studies been so closely related that the *same data* could have been used, then this would have limited the evidence that a Panel is essential, or otherwise. Here, the efficacy of each one can be assessed on its own merits. The topics selected were also of direct interest to the author offering him a variety that would sustain interest and motivation in undertaking the research for each Study.

The first Study tested the *Representative Bureaucracy Hypothesis* in the context of the gendered policing Police Service in England and Wales between 2002 and 2013, and whether or not this appeared to have a significant effect on the number of arrests made for crimes of gender-based violence (GBV). The author had already contributed to research into gendered police *leadership*, that adopted a regression-based approach for the first time anywhere in the World. This Study broadened this out to consider the impact of the increased participation of female officers at *all levels* of the Service in England and Wales, using a self-assembled panel of published Home Office annual data, aggregated at Force level. The overt connection to (labour) economics was the potential employment implications for males and females, should it be discovered that employing proportionately more (or fewer) females had any significant effect on arrests made for GBV. However, the main contribution really pertains to Civic Society, and the extent to which it might be better served through a more closely gender-aligned Police Service. It is within an improved Civic Society, of course, that improvements to economic wellbeing may be achieved, so there is, arguably, an indirect linkage into general Economics. As there is no established quantitative theory into the causes of GBV, but a collection of suppositions regarding socio-economic (such as poverty and population density), legal (such as sanctions for GBV crimes, detection probabilities), the modelling approach was to regress measures of GBV against a small collection of candidate causal variables, namely the numbers of females and males in the Forces and general population and household densities, with the Police Forces as the individual effect using a variety of pooled- and panel models. The analysis, discussed in **2.7.1** failed to find any definitive evidence that altering the gender balance of the Forces unequivocally affected either the absolute numbers or rates of GBV arrests, with the sign and significance of the female police parameters depending upon the econometric model employed. Neither was it clear that a Panel-, as opposed to a Pooled specification made any substantive differences to the conclusions drawn. While this inability to draw any definitive conclusions may be due to data and/or econometric issues, there are plausible reasons why, in the period of the Study, there appeared to be no significant effect of increased female police numbers, and the conclusions to that Study made several recommendations regarding future work, much of which would require significant resourcing and access to highly confidential data.

The second Study looked for evidence of significantly different *wage-rate premia* accruing to qualifications between *different centiles* of the wage-rate distribution. In this case, there is a well-developed theory linking personal productivity to remuneration, with qualifications regarded as productivity- and therefore, remuneration-enhancing. In addition, there is a well-established econometric methodology regarding the decomposition of the wage rate into its components, including the levels of *highest* qualification possessed. There is also a lengthy literature both on methods and empirical application, particularly on the wage(rate) returns to qualifications in a number of territories and epochs. What there seems to be less of, are investigations into the wage rate premia accruing to these qualifications away from the mean of the wage-rate distribution. This Study focussed on the UK, using the large-scale Understanding Society Panel (US) spanning the period 2009-2015, and estimated a number of conditional quantiles' estimates of returns to degrees, diplomas and school qualifications. Again, the dependent variable (log of the hourly wage rate) was regressed against a collection of plausible, 'traditional' independent variables, extracted and rendered from the survey using a standard 'Mincer' wage equation. While the independent variables included covered many aspects likely to affect peoples' work decision and outcomes, there is always the risk that variables *excluded* from the analysis might have exerted some significant effect on the wage rate. Any of these might be unobservable, again either for resource reasons or because they were unobservable/unquantifiable (such as 'talent' or 'motivation'). Consequently, deriving estimates from a panel may result in different outcomes to those from a pooled model. Quantile returns using the data, in turn, as a pooled- then panel dataset were contrasted against means-based estimates to reveal the importance of estimating the returns across the wage rate distribution, also separately for males and females given their continued contrasting labour market experiences. The pooled- and panel results were also contrasted to each other, to reveal few, if any significant differences in returns' estimates for both males and females at all four levels of qualification assumed. Nevertheless, the Study confirmed that all qualifications significantly increase their holders' wage rate, but that the relative increases (i.e. % of the wage rate) varied according to gender, highest qualification and centile occupied on the wage rate distribution. The particular contribution of this Study was to provide evidence of the absolute- and relative returns to qualifications in the UK at different points of the wage rate distribution, as well one of the few studies, anywhere that directly assesses the competing Powell- and Canay Fixed Effects Quantiles' estimators. There was also the conclusion that there may be little mileage in using surveys like US given their reliance of the political appetite necessary to have the funding required to collect and render the data, maintained.

The third, and final Study estimated the *technical efficiencies of a panel of US Airlines* and, as a by-product, the nature of returns to scale in that industry. The data were assembled from a variety of different sources by the author, to span the period 1990-2014. This was a period that followed the rapid liberalisation of the industry, followed by the emergence of the so-called 'low cost operators'. As with the Returns to Qualifications' Study, there is well-established theory relating to how producers convert inputs into output and, more specifically, well-established econometric approach to ascertaining technical (in)efficiency from estimated production functions. Given the resource- and access limitations faced by the author, the only variables observed were (i) the dependent variable, being an aggregate measure of physical output (tonnes of mass moved through the air), (ii) the numbers of aircraft, and (ii) employees, with business-mix and time dummies added to the specifications. The specifications were inadequate in the sense that there are far more variables that might be expected to influence both the output and factor-mix decisions, such as relative factor costs and the level of fares. Again, the Panel specifications might be expected to plug the omitted variables' gap, to some extent. In this case, significant differences in the nature of returns to scale and technical efficiency between the Panel- and Pooled estimators were found, with the Panel methods viewing larger airlines as being extremely efficient when compared with the smaller operators, while the Pooled methods perceived them to be reasonably efficient. This Study updated the small, neglected literature estimating airline efficiency using this approach, and the appears to be the first to compare the results of Pooled and Panel-based estimators using the same data.

Each Study concluded with various *cris de coeur* for more and/or different data, some of which would require privileged access (particularly the GBV Study), all of which would require resources. The Studies have been limited to accessing and assembling publicly-available data and have, in turn, been limited in their ability to draw wider conclusions. With the GBV Study, access to officer-level ('notebook') and station-level data would allow the effect of the female officers on the workplace 'culture' to inform the subsequent on-street arrest decisions. Such data are highly confidential (and voluminous) and would require both considerable resources and a trust relationship to have been established. The returns to qualifications' study could be opened out to consider both *unconditional* quantiles and working-life trajectories in order to better assess the impact of qualifications on *lifetime earnings*. The airlines study could clearly be opened out to more territories than just the US and disaggregate the labour input into key sub-groupings, and included capital, other than aircraft in the production process. Assuming the data were available, the Study could be extended backwards in time, to embrace the pre-deregulation period and assess the impact of that process on general efficiency. Rather than just being limited to *physical* inputs and output, could consider the estimation of *cost frontiers*.

The overall conclusion is that Panels can, but will not necessarily, significantly affect the results obtained, and therefore any policy-related conclusions drawn. The fact that this evidence is mixed, should not dissuade the applied researcher from looking first for an already-constructed Panel, or to assemble one from different sources, then deriving pooled- and panel-based estimates from it. This may be facilitated in the future given the ever-increasing availability of comprehensive, ultra-large *administrative datasets*.